SU SPRENDIMAIS

VILIJANDAS BAGDONAVIČIUS
JULIUS JONAS KRUOPIS
RŪTA LEVULIENĖ

MATEMATINĖS STATISTIKOS UŽDAVINYNAS

VILNIAUS UNIVERSITETO LEIDYKLA

VILIJANDAS BAGDONAVIČIUS

JULIUS JONAS KRUOPIS

RŪTA LEVULIENĖ

MATEMATINĖS STATISTIKOS UŽDAVINYNAS

SU SPRENDIMAIS

VILNIAUS UNIVERSITETO LEIDYKLA

2019

Apsvarstė ir rekomendavo išleisti Vilniaus universiteto Matematikos ir informatikos fakulteto taryba (2019 m. gegužės 31 d., protokolas Nr. (1.5) 110000-TPN-24)

Recenzentai:
prof. habil. dr. Algimantas Bikelis (Vytauto Didžiojo universitetas)
prof. dr. Kęstutis Dučinskas (Klaipėdos universitetas)

Leidinio bibliografinė informacija pateikiama Lietuvos nacionalinės Martyno Mažvydo bibliotekos Nacionalinės bibliografijos duomenų banke (NBDB)

ISBN 978-609-07-0247-5

Turinys

Pratarmė

Matematinės statistikos metodų įsisavinimas yra neįsivaizduojamas, jeigu nebandoma juos pritaikyti sprendžiant įvairaus tipo uždavinius.

Kiekvienoje matematinės statistikos knygoje, monografijoje ar vadovėlyje pateikiama daug įvairių pratimų. Priklausomai nuo knygos paskirties, didesnis dėmesys skiriamas statistinių metodų taikymui realiems duomenims analizuoti arba teoriniams uždaviniams. Klasikinio tipo vadovėliuose įprastai pateikiami įvairaus sudėtingumo ir tipo uždaviniai, didelis dėmesys skiriamas teoriniams pratimams, kuriuos sprendžiant įsigilinama, kaip konstruojami metodai, sukuriamos jų taikymo rekomendacijos. Tik supratus, kaip gaunami matematinės statistikos rezultatai, galima kvalifikuotai interpretuoti konkrečių statistinių duomenų analizės rezultatus ir, susidūrus su nestandartine situacija, sukurti naują arba pakoreguoti esamą sprendimo metodą.

Daugelyje matematinės statistikos knygų pateikiami teorinio pobūdžio uždaviniai papildo ir išplečia dėstomą medžiagą. Neretai juose pateikiami naujausi matematinės statistikos rezultatai, kuriuos galima rasti mokslinėse publikacijose. Tokie uždaviniai daugumai studentų yra sunkiai įveikiami. Iškyla poreikis matematinės statistikos uždavinyno, kuriame būtų pateikti teorinių uždavinių sprendimai arba gana detalūs jų sprendimo nurodymai. Be to, tokiame uždavinyne turėtų būti ir praktinių uždavinių, skirtų realių statistinių duomenų analizei ir gautų analizės rezultatų interpretavimui. Iki šiol neturime tokio uždavinyno lietuvių kalba. Šis uždavinynas ir skirtas šiai spragai užpildyti. Iš užsienio kalbomis išleistų tokio pobūdžio uždavinynų paminėsime [6], [9], [10].

Kad suprastų pateikiamų uždavinių sprendimus, skaitytojas turėtų būti išklausęs universitetinės apimties tikimybių teorijos kursą ir įsisavinęs atitinkamų matematinės statistikos skyrių medžiagą, pavyzdžiui, iš vadovėlio [2], [3], [4]. Uždavinius pateikiame iš minėto vadovėlio, papildydami juos uždaviniais iš kitų matematinės statistikos knygų ir uždavinynų. Sprendžiant pratimus pateikiamos nuorodos į taikomas formules, teoremas ar skyrelius iš minėto vadovėlio. Sąvokas ar apibrėžimus galima pasitikslinti minėtame vadovėlyje pasinaudojus dalykine rodykle.

Uždavinynas skirtas universitetų statistikos ir matematikos studijų krypčių studentams, studijuojantiems matematinės statistikos kursą, taip pat ir kitų mokslinės ar praktinės sričių specialistams, taikantiems matematinės statistikos metodus.

Autoriai

Trumpiniai ir žymenys

A. d. – atsitiktinis dydis;
n. a. d. – nepriklausomi atsitiktiniai dydžiai;
a. v. – atsitiktinis vektorius;
n. a. v. – nepriklausomi atsitiktiniai vektoriai;
TG – tolygiai galingiausias (kriterijus);
TGN – tolygiai galingiausias nepaslinktasis (kriterijus);
DT – didžiausiojo tikėtinumo (funkcija, metodas, įvertinys);
ASE – asimptotinis santykinis efektyvumas (įvertinių, kriterijų);
NMD įvertinys – nepaslinktasis minimalios dispersijos įvertinys
$X,\ Y,\ Z, \dots$ – atsitiktiniai dydžiai;
$\mathbf{X},\ \boldsymbol{Y},\ \boldsymbol{Z}, \dots$ – atsitiktiniai vektoriai;
$\mathbf{X}^T$ – transponuotas vektorius, t. y. vektorius – eilutė;
α_k – pradinis k-osios eilės momentas;
μ_k – centrinis k-osios eilės momentas;
γ_1 – asimetrijos koeficientas;
γ_2 – eksceso koeficientas;
$x(P)$ – P-asis kvantilis;
x_P – P-oji kritinė reikšmė;
pv – P reikšmė;
$\boldsymbol{\Sigma} = [\sigma_{ij}]_{k\times k}$ – kovariacijų matrica;
$\boldsymbol{\rho} = [\rho_{ij}]_{k\times k}$ – koreliacijos koeficientų matrica;
$\mathbf{P}\{A\}$ – įvykio A tikimybė;
$\mathbf{P}\{A|B\}$ – įvykio A sąlyginė tikimybė;
$\mathbf{P}_\theta\{A\}$, $\mathbf{P}\{A|\theta\}$ – tikimybė, priklausanti nuo parametro θ;
$F_\theta(x)$, $F(x;\ \theta)$, $F(x|\theta)$ – pasiskirstymo funkcija, priklausanti nuo parametro θ (analogiškai tankio funkcijai);
$\mathbf{E}X$ – a. d. X vidurkis;
$\boldsymbol{V}X$ – a. d. X dispersija;
$\mathbf{E}_\theta(X)$, $\mathbf{E}(X|\theta)$, $\boldsymbol{V}_\theta(X)$, $\boldsymbol{V}(X|\theta)$ – a. d. X vidurkis ar dispersija, priklausantys nuo parametro θ;
$\mathbf{E}(\mathbf{X})$ – a. v. $\mathbf{X}$ vidurkių vektorius;
$\boldsymbol{V}(\mathbf{X})$ – a. v. $\mathbf{X}$ kovariacijų matrica;
$\mathbf{Cov}\,(X,\ Y)$ – a. d. X ir Y kovariacija;
$\mathbf{Cov}\,(\mathbf{X},\ \boldsymbol{Y})$ – a. v. $\mathbf{X}$ ir $\boldsymbol{Y}$ kovariacijų matrica;
$B(n,\ p)$ – binominis skirstinys su parametrais n ir p;
$B^-(n,\ p)$ – neigiamasis binominis skirstinys su parametrais n ir p;
$\mathcal{P}(\lambda)$ – Puasono skirstinys su parametru λ;
$H(N,\ M,\ n)$ – hipergeometrinis skirstinys su parametrais N, M ir n;
$N(0,\ 1)$ – standartinis normalusis skirstinys;
$N(\mu,\ \sigma^2)$ – normalusis skirstinys su parametrais μ ir σ^2;
$LN(\mu,\ \sigma)$ – lognormalusis skirstinys su parametrais μ ir σ;
$K(\mu,\ \sigma)$ – Koši skirstinys su parametrais μ ir σ;

$\mathcal{E}(\lambda)$ – eksponentinis skirstinys su parametru λ;
$\mathcal{E}(\alpha,\ \lambda)$ – paslinktasis eksponentinis skirstinys su parametrais α ir λ;
$G(\lambda,\ \eta)$ – gama skirstinys su parametrais λ ir η;
$W(\theta,\ \nu)$ – Veibulo skirstinys su parametrais θ ir ν;
$Pa(\alpha,\ \theta)$ – Pareto skirstinys su parametrais α ir θ;
$Be(\gamma,\ \eta)$ – beta skirstinys su parametrais γ ir η;
$U(\alpha,\ \beta)$ – tolygusis skirstinys intervale $(\alpha,\ \beta)$;
$\chi^2(n)$ – chi kvadrato skirstinys su n laisvės laipsnių;
$\chi^2(n;\ \delta)$ – necentrinis chi kvadrato skirstinys su n laisvės laipsnių ir necentriškumo parametru δ;
$S(n)$ – Stjudento skirstinys su n laisvės laipsnių;
$S(n;\ \delta)$ – necentrinis Stjudento skirstinys su n laisvės laipsnių ir necentriškumo parametru δ;
$F(m,\ n)$ – Fišerio skirstinys su m ir n laisvės laipsnių;
$F(m,\ n;\ \delta)$ – necentrinis Fišerio skirstinys su m ir n laisvės laipsnių ir necentriškumo parametru δ;
z_α – standartinio normaliojo skirstinio α kritinė reikšmė;
$t_\alpha(n)$ – Stjudento skirstinio su n laisvės laipsnių α kritinė reikšmė;
$\chi^2_\alpha(n)$ – chi kvadrato skirstinio su n laisvės laipsnių α kritinė reikšmė;
$F_\alpha(\mathrm{m, n})$ – Fišerio skirstinio su m ir n laisvės laipsnių α kritinė reikšmė;
$\mathcal{P}_k(n,\ \boldsymbol{\pi})$ – k-matis polinominis skirstinys su parametrais n ir $\boldsymbol{\pi} = (\pi_1, ..., \pi_k)^T$, $\pi_1 + ... + \pi_k = 1$;
$H_k(N,\ \boldsymbol{M},\ n)$ – k-matis hipergeometrinis skirstinys su parametrais N, $\boldsymbol{M} = (M_1, ..., M_k)^T$ ir n;
$N_k(\boldsymbol{\mu},\ \boldsymbol{\Sigma})$ – k-matis normalusis skirstinys su vidurkių vektoriumi $\boldsymbol{\mu}$ ir kovariacijų matrica $\boldsymbol{\Sigma}$;
$X \sim N(\mu,\ \sigma^2)$ – a. d. X, pasiskirstęs pagal normalųjį dėsnį su parametrais μ ir σ^2 (analogiškai kitų skirstinių atveju);
$X_n \xrightarrow{P} X$ – konvergavimas pagal tikimybę $(n \to \infty)$;
$X_n \xrightarrow{b.t.} X$ – konvergavimas su tikimybe 1 arba beveik tikrai $(n \to \infty)$;
$X_n \xrightarrow{kv.v.} X$ – konvergavimas pagal kvadratinį vidurkį $(n \to \infty)$;
$X_n \xrightarrow{d} X$, $F_n(x) \xrightarrow{d} F(x)$ – konvergavimas pagal pasiskirstymą (silpnasis; $n \to \infty$);
$X_n \xrightarrow{d} X \sim N(\mu,\ \sigma^2)$ – a. d. X_n asimptotiškai $(n \to \infty)$ turi normalųjį skirstinį su parametrais μ ir σ^2;
$X_n \overset{d}{=} Y_n$ – a. d. X_n ir Y_n skirstiniai vienodi;
$||\boldsymbol{x}||$ – kai $\boldsymbol{x} = (x_1, ..., x_k)^T$ yra vektorius, reiškia atstumą $(\boldsymbol{x}^T\boldsymbol{x})^{1/2} = (\sum_i x_i^2)^{1/2}$;
$||\boldsymbol{A}||$ – kai $\boldsymbol{A} = [a_{ij}]$ yra matrica, reiškia $(\sum_i \sum_j a_{ij}^2)^{1/2}$;
$\boldsymbol{A} > \boldsymbol{B}$ $(\boldsymbol{A} \geq \boldsymbol{B})$ – kai $\boldsymbol{A}$ ir $\boldsymbol{B}$ yra vienodos dimensijos kvadratinės matricos, reiškia, kad matrica $\boldsymbol{A} - \boldsymbol{B}$ yra teigiamai (neneigiamai) apibrėžta.

I. Parametrinė statistika

I.1. Empirinės charakteristikos

I.1.1. Statistinis modelis

I.1.1. Yra dvi nepriklausomos paprastosios a. d. $X \sim B(1, p)$ imtys, kurių didumai yra n_1 ir n_2. Tegu vieneto pasirodymų skaičiai šiose imtyse yra X_1 ir X_2. Sudarykite jungtinės imties $\mathbf{X} = (X_1,\ X_2)^T$ statistinį modelį. Raskite statistikos

$$T = \frac{X_1}{n_1} - \frac{X_2}{n_2}$$

vidurkį ir dispersiją.

I.1.2. Objektų, turinčių savybę A, dalys k populiacijose atitinkamai yra $\pi_1, ..., \pi_k$.

a) Atsitiktinai parenkama populiacija ir iš jos atsitiktinai imama grąžinant didumo n imtis.

b) Atsitiktinai parenkama populiacija ir iš jos atsitiktinai imamas grąžinant vienas elementas. Procedūra kartojama n kartų.

Sudarykite šių eksperimentų statistinius modelius. Raskite statistikos X/n vidurkį ir dispersiją, kai X yra skaičius imties objektų, turinčių savybę A.

I.1.3. (**I.1.2** pratimo tęsinys). Iš kiekvienos populiacijos atsitiktinai imama grąžinant didumo n imtis. Sudarykite šio eksperimento statistinį modelį. Raskite statistikos X/n vidurkį ir dispersiją, kai X yra jungtinės nk didumo imties objektų, turinčių savybę A, skaičius.

I.1.4. Iš populiacijos, kurioje požymį A turinčių objektų dalis lygi π, atsitiktinai imama grąžinant didumo n_1 imtis. Iš tos imties atsitiktinai imama grąžinant didumo n_2 imtis. Sudarykite šio eksperimento statistinį modelį.

I.1.5. (**I.1.4** pratimo tęsinys). Tarę, kad eksperimento metu užregistruojami tik skaičiai X_1 ir X_2 pirmosios ir antrosios imties objektų, turinčių savybę A, sudarykite imties $(X_1,\ X_2)^T$ statistinį modelį. Raskite statistikos $X_1/n_1 - X_2/n_2$ vidurkį ir dispersiją.

I.1.6. Išspręskite **I.1.5** pratimą, kai imtys imamos atsitiktinai ir negrąžinant (populiaciją sudaro N elementų, $n_2 < n_1 < N$).

I.1.7. Atrankinei kontrolei pateko N dydžio gaminių partija, kurioje yra nežinomas skaičius M defektinių gaminių. Atsitiktinai imama negrąžinant n gaminių ir nustatomas defektinių skaičius X iš jų. Sudarykite imties X statistinį modelį.

I.1.8. (**I.1.7** pratimo tęsinys). Tegu partija pripažįstama gera, kai $X \leq c$. Kaip priklauso partijos priėmimo tikimybė nuo šios partijos defektinių gaminių skaičiaus M?

I.1.9. Aibę sudaro N objektų, kurie apibūdinami tam tikru požymiu y, t. y. aibė $O_N = \{y_1, ..., y_N\}$ sudaryta iš požymio y reikšmių. Atsitiktinai imame negrąžinant $n \leq N$ objektų. Tegu a. v. $\mathbf{X} = (X_1, ..., X_n)^T$ koordinatė X_i lygi i-ojo parinkto objekto požymio y reikšmei. a) Įrodykite, kad a. v. $\mathbf{X}$ koordinatės yra priklausomi a. d. ir imtis nėra paprastoji; sudarykite imties $\mathbf{X}$ statistinį modelį. b) Raskite a. d. $Y = X_1 + ... + X_n$ vidurkį ir dispersiją. c) Raskite a. d. Y tikimybinį skirstinį, kai y_i, $i = 1, ..., N$, įgyja tik dvi reikšmes: 0 arba 1.

I.1.10. (**I.1.9** pratimo tęsinys). Išspręskite **I.1.9** pratimą tuo atveju, kai objektai imami atsitiktinai grąžinant.

I.1.11. Tegu $\mathbf{X} = (X_1, ..., X_m)^T$ ir $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ yra dvi paprastosios imtys, gautos stebint nepriklausomus a. d. $X \sim \mathcal{P}(\lambda_1)$ ir $Y \sim \mathcal{P}(\lambda_2)$. Sudarykite jungtinės didumo $m + n$ imties statistinį modelį. Kaip pasikeistų šis modelis, jeigu būtų žinoma, kad $\lambda_1 = \lambda_2$?

I.1.12. Tegu $\mathbf{X} = (X_1, ..., X_m)^T$ ir $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ yra dvi paprastosios imtys, gautos stebint nepriklausomus a. d. $X \sim N(\mu_1,\ \sigma_1^2)$ ir $Y \sim N(\mu_2,\ \sigma_2^2)$. Sudarykite jungtinės didumo $m + n$ imties statistinį modelį, jeigu žinoma, kad: a) $\mu_1 = \mu_2$; b) $\sigma_1 = \sigma_2$; c) jokios informacijos apie parametrus nėra.

I.1.13. Tegu $(X_i,\ Y_i)^T, i = 1, ..., n$, yra paprastoji atsitiktinė a. v. $(X, Y)^T$, turinčio dvimatį normalųjį skirstinį, imtis. Sudarykite imties statistinį modelį. Kaip pasikeistų statistinis modelis, jeigu būtų žinoma, kad marginalieji atsitiktinių dydžių X ir Y skirstiniai vienodi?

I.1.14. Imties $\mathbf{X} = (X_1, ..., X_n)^T$ elementai $X_i \sim N(\mu,\ \sigma^2)$, $i = 1, ..., n$. Koreliacijos koeficientas $\rho(X_i,\ X_j) = \rho$, kai $|i - j| = 1$, ir $\rho(X_i,\ X_j) = 0$, kai $|i - j| > 1$. Sudarykite imties $\mathbf{X}$ statistinį modelį.

I.1.15. Tegu $\mathbf{X}_i = (X_{i1}, ..., X_{in_i})^T$, $i = 1, ..., m$, yra paprastosios atsitiktinės imtys, gautos stebint absoliučiai tolydžius n. a. d. $X_1, ..., X_m$. Sudarykite jungtinės imties $\mathbf{X} = (\mathbf{X}_1^T, ..., \mathbf{X}_m^T)^T$ statistinį modelį, kai žinoma, kad: a) a. d. $X_1, ..., X_m$ skirstiniai gali skirtis tik poslinkio parametru; b) gali skirtis tik mastelio parametru; c) gali skirtis poslinkio ir mastelio parametrais; d) jokios papildomos informacijos nėra.

I.1.16. Imties $\mathbf{X} = (X_1, ..., X_n)^T$ nariai yra nepriklausomi absoliučiai tolydūs atsitiktiniai dydžiai. Sudarykite imties $\mathbf{X}$ statistinį modelį. Kaip pasikeistų statistinis modelis, jeigu imtis būtų paprastoji?

I.1.17. Paprastoji atsitiktinė imtis $\mathbf{X}^{(i)} = (X_{1i}, ..., X_{ki})^T$, $i = 1, ..., n$, gauta stebint a. v. $\mathbf{X} = (X_1, ..., X_k)^T$. Sudarykite imties statistinį modelį, jeigu žinoma, kad a. d. $X_1, ..., X_k$ marginalieji skirstiniai yra vienodi.

I.1.18. Taškas tolygiai juda tiese greičiu v. Laiko momentais $t_1, ..., t_n$ užregistruojamas taško nueitas kelias $X_1, ..., X_n$; laiko momentu $t = 0$ taškas buvo taške x_0. Tarkime, kad matavimo rezultatai tarpusavyje nepriklausomi, matavimo paklaidos neturi sisteminės dedamosios, o atsitiktinė dedamoji pasiskirsčiusi pagal normalųjį dėsnį $N(0,\ \sigma^2)$. Sudarykite imties $\mathbf{X} = (X_1, ..., X_n)^T$ statistinį modelį.

I.1.19. (**I.1.18** pratimo tęsinys). Sudarykite statistinį modelį tuo atveju, kai žinoma, kad $x_0 = 0$.

I.1.20. Sveriant tą patį kūną n kartų, gauti svėrimo rezultatai $X_1, ..., X_n$. Tegu jie tarpusavyje nepriklausomi. Pirmųjų k svėrimų sisteminė paklaida 0, o likusiųjų – μ. Atsitiktinės visų svėrimų paklaidų dedamosios turi normaliuosius skirstinius $N(0, \sigma^2)$. Sudarykite imties $\mathbf{X} = (X_1, ..., X_n)^T$ statistinį modelį.

I.1.2. Empirinė pasiskirstymo funkcija ir tankis

I.1.21. Tegu $\hat{F}_n(x)$ yra empirinė pasiskirstymo funkcija, gauta iš paprastosios didumo $n = 100$ imties a. d. $X \sim N(1,\, 4)$. Raskite aproksimaciją tikimybės, kad $\hat{F}_n(0)$ absoliutiniu didumu skirsis nuo teorinės pasiskirstymo funkcijos taške $x = 0$ ne daugiau kaip 0,1; 0,05.

I.1.22. Paprastoji didumo n imtis gauta stebint absoliučiai tolydųjį a. d. su pasiskirstymo funkcija $F(x)$. Raskite aproksimaciją tikimybės, kad empirinės pasiskirstymo funkcijos $\hat{F}_n(x)$ maksimalus nuokrypis nuo teorinės pasiskirstymo funkcijos neviršys ε. Raskite apytikslę šios tikimybės reikšmę, kai $n = 50; 100$ ir $\varepsilon = 0,1; 0,05$.

I.1.23. Paprastoji imtis gauta stebint a. d. $X \sim \mathcal{P}(3)$. Koks apytiksliai turi būti imties didumas n, kad, vertinant tikimybę $\mathbf{P}\{X = 2\}$ santykiniu dažniu, absoliutinė paklaida neviršytų 0,05 su tikimybe, ne mažesne už $0,95$?

I.1.24. Lentelėje pateikiamos $n = 100$ atsitiktinai atrinktų gaminių tam tikro parametro pamatuotos reikšmės (didumo n paprastosios imties realizacija).

24	41	30	37	25	32	28	35	28	51
36	26	43	25	27	39	21	45	39	25
29	43	66	25	24	56	29	31	41	41
36	57	36	48	25	36	48	24	48	22
40	7	31	24	32	53	33	46	22	33
25	37	34	32	41	36	19	32	25	19
19	37	20	21	48	44	35	19	44	34
29	48	38	43	48	35	42	37	35	36
58	45	34	40	37	21	41	11	41	27
50	24	37	39	33	45	39	43	21	34

a) Sugrupuokite stebinius ilgio $h = 10$ intervalais pradėdami nuo 0.
b) Nubraižykite empirinės pasiskirstymo funkcijos realizacijos grafiką ir histogramą.

I.1.25. (**I.1.24** pratimo tęsinys). Palyginkite empirinės pasiskirstymo funkcijos realizacijos grafiką ir histogramą su normaliojo skirstinio pasiskirstymo funkcija ir tankiu (vietoje nežinomų parametrų imkime jų empirinius analogus).

I.1.3. Pozicinės statistikos

I.1.26. Paprastoji didumo $n = 50$ imtis gauta stebint a. d., kurio tankis $f(x|\mu) = 2e^{-2(x-\mu)}, x > \mu$. Raskite tikimybę, kad $X_{(1)} - \mu \leq 0,05$.

I.1.27. Variacinė eilutė $X_{(1)}, ..., X_{(n)}$ gauta stebint paslinktą eksponentinį a. d. $X \sim \mathcal{E}(1/\sigma, \mu)$, kurio tankio funkcija

$$f(x|\mu,\, \sigma) = \frac{1}{\sigma}\exp\{-\frac{x-\mu}{\sigma}\},\ x > \mu.$$

a) Raskite statistikos

$$T = \frac{n(X_{(1)} - \mu)}{W}, \quad (W = \frac{1}{n-1}\sum_{i=2}^{n}(X_{(i)} - X_{(1)}))$$

tikimybinį tankį. b) Įrodykite, kad $X_{(1)}$ skirstinys yra $\mathcal{E}(n/\sigma, \mu)$; a. d. $2(\sum_{i=1}^{n} X_i - nX_{(1)})/\sigma$ turi χ^2 skirstinį $\chi^2(2n-2)$. c) Tarkime, kad paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $\mathcal{E}(\lambda)$. Įrodykite, kad a. d. $Y_1 = nX_{(1)}, Y_2 = (n-1)(X_{(2)} - X_{(1)}), ..., Y_n = X_{(n)} - X_{(n-1)}$ yra nepriklausomi ir turi tą patį eksponentinį skirstinį $\mathcal{E}(\lambda)$.

I.1.28. Variacinė eilutė $X_{(1)}, ..., X_{(n)}$ gauta stebint tolydųjį a. d. X, kurio pasiskirstymo funkcija yra $F(x)$. Įrodykite, kad a. d.

$$Y_i = \left(\frac{F(X_{(i)})}{F(X_{(i+1)})}\right)^i, i = 1, 2, ..., n-1,$$

yra nepriklausomi ir vienodai pasiskirstę pagal $U(0,\ 1)$.

I.1.29. Tegu $\mathbf{X} = (X_1, ..., X_m)^T$ ir $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ yra dvi paprastosios nepriklausomos atsitiktinės imtys, gautos stebint tolygųjį a. d. $U(0,\ 1)$. Raskite tų imčių maksimalių reikšmių santykio $X_{(m)}/Y_{(n)}$ tikimybinį skirstinį.

I.1.30. Imtis, kurios didumas $n = 2k+1$, gauta stebint a. d. $X \sim U(0,\ 1)$. Įrodykite, kad empirinės medianos dispersija lygi $1/(4(2k+3))$.

I.1.31. Variacinė eilutė $X_{(1)}, ..., X_{(n)}$ gauta stebint tolydųjį a. d., kurio pasiskirstymo funkcija yra $F(x)$. Raskite a. v. $(F(X_{(k_1)}), F(X_{(k_2)}))^T$ kovariacijų matricą.

I.1.32. Variacinė eilutė $X_{(1)}, ..., X_{(n)}$ gauta stebint tolydųjį a. d., kurio pasiskirstymo funkcija yra $F(x)$. Įrodykite, kad imties pločio $W_n = X_{(n)} - X_{(1)}$ vidurkis yra

$$\mathbf{E}W_n = \int_{-\infty}^{\infty}(1 - F^n(x) - (1 - F(x))^n)dx,$$

jeigu pasiskirstymo funkcija $F(x)$ tenkina sąlygą $x[1 - F^n(x) - (1 - F(x)^n)] \to 0$, kai $x \to \pm\infty$.

I.1.33. Raskite kraštinių pozicinių statistikų pirmuosius momentus, kai stebimo a. d. skirstinys yra eksponentinis.

I.1.34. Yra k nepriklausomų paprastųjų didumo n imčių, gautų stebint a. d. $X_i \sim U(0,\ 1)$. Tegu i-osios imties maksimali reikšmė yra $X_{(n)}^{(i)}$ ir $V = X_{(n)}^{(1)} ... X_{(n)}^{(k)}$. Raskite a. d. V tikimybių pasiskirstymo dėsnį.

I.1.35. Tegu $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra a. d. $X \sim G(\lambda,\ \eta)$ paprastoji atsitiktinė imtis. Įrodykite, kad $\sum_{i=1}^{n} X_i$ ir $\sum_{i=1}^{n}[\ln X_i - \ln X_{(1)}]$ yra nepriklausomi.

I.1.36. Sakykime, $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra a. d. $X \sim U(a,\ b)$, $-\infty < a < b < \infty$, paprastoji atsitiktinė imtis. Įrodykite, kad $Y_i = (X_{(i)} - X_{(1)})/(X_{(n)} - X_{(1)})$, $i = 2, \ldots, n-1$, yra nepriklausomi nuo $X_{(1)}$ ir $X_{(n)}$.

I.1.37. Tegu $X_1, ...X_{k+1}$ yra n. a. d., turintys gama skirstinius $X_i \sim G(\lambda, \eta_i), i = 1, ..., k+1$. Įrodykite, kad a. v. $(Z_1, ..., Z_k)^T$, kai $Z_i = X_i/(X_1 + ... + X_{k+1})$, $i = 1, 2, ..., k$, turi k-matį Dirichlė skirstinį $D(\eta_1, ..., \eta_k; \eta_{k+1})$. Jeigu $k = 1$, tai $Z_1 \sim Be(\eta_1, \eta_2)$.

I.1.38. Tegu $X_1, ..., X_n$ yra paprastoji imtis a. d. $X \sim U(0, \theta), \theta > 0$. Įrodykite, kad $X_{(n)}$ tankio funkcija yra nx^{n-1}/θ^n, kai $0 < x < \theta$.

I.1.39. Tegu $X_1, ..., X_n$ yra paprastoji imtis a. d. $X \sim U(\mu - \theta/2, \mu + \theta/2), \theta > 0$. Įrodykite, kad imties pločio $X_{(n)} - X_{(1)}$ tankio funkcija

$$(n(n-1)/\theta)(x/\theta)^{n-2}(1 - x/\theta), \quad 0 < x < \theta.$$

I.1.40. (**I.1.39** pratimo tęsinys). Įrodykite, kad a. d. $Z = ((X_{(1)} + X_{(n)})/2 - \mu)/(X_{(n)} - X_{(1)})$ tankio funkcija $f(x) = (n-1)/(1 + 2|x|)^n$, $-\infty < x < \infty$.

I.1.41. Tegu $X_1, X_2, ..., X_n$ yra paprastoji imtis a. d. $X \sim U(0, 1)$. Įrodykite, kad a. d. $Y = (X_1 \cdot ... \cdot X_n)^{1/n}$ tankio funkcija yra

$$f(y) = \frac{n^n y^{n-1}}{(n-1)!}(-\ln y)^{n-1}, \quad 0 < y < 1.$$

I.1.42. Tegu $X_{(k)}$ yra k-oji pozicinė statistika iš didumo m imties, gautos stebint a. d. X su absoliučiai tolydžia pasiskirstymo funkcija $F(x)$. Tarkime, kad gauta kita nepriklausoma didumo n imtis su ta pačia pasiskirstymo funkcija $F(x)$. Pažymėkime Z skaičių tokių antrosios imties elementų, kurie neviršija $X_{(k)}$. Įrodykite, kad a. d. Z skirstinys nusakytas tikimybėmis

$$\mathbf{P}\{Z = z\} = \frac{C_m^k C_n^z}{C_{m+n}^{k+z}} \frac{k}{k+z}, \quad z = 0, 1, ..., n.$$

I.1.43. (**I.1.42** pratimo tęsinys). Įrodykite, kad a. d. Z r-asis faktorialinis momentas

$$\mathbf{E}(Z^{[r]}) = \frac{m!n!(k+r-1)!}{(k-1)!(m+r)!(n-r)!}, \quad 1 \le r \le n.$$

I.1.44. Tegu $Y_1, ..., Y_n$ yra paprastoji imtis a. d. $Y \sim U(0, 1)$. Įrodykite, kad jei $k, n \to \infty$, $k/n \to p$, $0 < p < 1$, tai

$$\sqrt{n}(Y_{(k)} - p) \xrightarrow{d} Y \sim N(0, p(1-p)).$$

I.1.45. Tegu $X_1, ..., X_n$ yra paprastoji imtis a. d. X su tolydžia pasiskirstymo funkcija $F(x)$. Įrodykite, kad jei $F(x)$ tolydžiai diferencijuojama taško $x(p)$, $0 < p < 1$, $0 < p < 1$ aplinkoje ir tankis $f(x(p)) > 0$, $k/n \to p$, $n \to \infty$, tai

$$\sqrt{n}(\hat{x}(p) - x(p)) \xrightarrow{d} Y \sim N(0, p(1-p)/f^2(x(p))).$$

I.1.46. Tegu $X_1, ..., X_n$ yra paprastoji imtis a. d. X su tolydžia pasiskirstymo funkcija $F(x)$. Įrodykite, kad jei $F(x)$ tolydžiai diferencijuojama taškų $x(p_1), x(p_2)$, $0 < p_1 < p_2 < 1$ aplinkose ir tankio reikšmės $f(x(p_1)) > 0, f(x(p_2)) > 0$, tai $n \to \infty$

$$\sqrt{n}(\hat{x}(p_2) - \hat{x}(p_1) - (x(p_2) - x(p_1))) \xrightarrow{d} Y \sim N(0, \sigma^2(p_1, p_2)),$$

$$\sigma^2(p_1, p_2) = \frac{p_1(1-p_1)}{f^2(x(p_1))} - 2\frac{p_1(1-p_2)}{f(x(p_1))f(x(p_2))} + \frac{p_2(1-p_2)}{f^2(x(p_2))}.$$

I.1.47. (**I.1.46** pratimo tęsinys). Jeigu $p_1 = 1/4$, o $p_2 = 3/4$, tai $z = x(3/4) - x(1/4)$ vadinamas interkvartiliniu pločiu. Raskite statistikos $\hat{z} = \hat{x}(3/4) - \hat{x}(1/4)$ asimptotinį skirstinį.

I.1.48. Tarkime, kad paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $X \sim K(\mu, \sigma)$. Šio skirstinio vidurkis neegzistuoja, o parametras μ yra mediana $x(1/2) = \mu$. Kvartiliai $x(1/4) = \mu - \sigma$, $x(3/4) = \mu + \sigma$, ir $(x(3/4) - x(1/4))/2 = \sigma$. Raskite statistikų $\hat{\mu} = \hat{x}(1/2)$ ir $\hat{\sigma} = (\hat{x}(3/4) - \hat{x}(1/4))/2$ asimptotinius skirstinius.

I.1.49. (**I.1.48** pratimo tęsinys). Tegu $0 < p < 1/2$. a) parinkite konstantą $c(p)$ taip, kad $(x(1-p) - x(p))/c(p) = \sigma$; b) raskite tokią parametro p reikšmę, kad statistikos $\hat{\sigma} = (\hat{x}(1-p) - \hat{x}(p))/c(p)$ asimptotinio skirstinio dispersija būtų minimali.

I.1.50. Tarkime, kad paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $X \sim N(\mu, \sigma^2)$. Šio skirstinio mediana sutampa su vidurkiu $x(1/2) = \mu$. Kvartiliai $x(1/4) = \mu - z_{1/4}\sigma$, $x(3/4) = \mu + z_{1/4}\sigma$, $z_{1/4}$ yra standartinio normaliojo skirstinio kritinė reikšmė $z_{1/4} = \Phi^{-1}(3/4)$, ir $(x(3/4) - x(1/4))/(2z_{1/4}) = \sigma$. Raskite statistikų $\hat{\mu} = \hat{x}(1/2)$ ir $\hat{\sigma} = (\hat{x}(3/4) - \hat{x}(1/4))/(2z_{1/4})$ asimptotinius skirstinius.

I.1.51. (**I.1.50** pratimo tęsinys). Tegu $0 < p < 1/2$. a) parinkite konstantą $c(p)$ taip, kad $(x(1-p) - x(p))/c(p) = \sigma$; b) raskite tokią parametro p reikšmę, kad statistikos $\hat{\sigma} = (\hat{x}(1-p) - \hat{x}(p))/c(p)$ asimptotinio skirstinio dispersija būtų minimali.

I.1.52. Tarkime, kad paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. X, turintį Laplaso skirstinį, kurio tankis

$$f(x|\mu, \sigma) = \frac{1}{2\sigma} e^{-|x-\mu|/\sigma}, \quad \sigma > 0, \quad x, \mu \in \mathbf{R}.$$

Šio skirstinio mediana sutampa su vidurkiu $x(1/2) = \mu$. Kvartiliai $x(1/4) = \mu - \sigma \ln 2$, $x(3/4) = \mu + \sigma \ln 2$, ir $(x(3/4) - x(1/4))/(2\ln 2) = \sigma$. Raskite statistikų $\hat{\mu} = \hat{x}(1/2)$ ir $\hat{\sigma} = (\hat{x}(3/4) - \hat{x}(1/4))/(2\ln 2)$ asimptotinius skirstinius.

I.1.53. (**I.1.52** pratimo tęsinys). Tegu $0 < p < 1/2$. a) parinkite konstantą $c(p)$ taip, kad $(x(1-p) - x(p))/c(p) = \sigma$; b) raskite tokią parametro p reikšmę, kad statistikos $\hat{\sigma} = (\hat{x}(1-p) - \hat{x}(p))/c(p)$ asimptotinio skirstinio dispersija būtų minimali.

I.1.54. Empirinė mediana $\hat{x}(1/2)$ gauta iš paprastosios didumo n imties a. d. X su tolydžia pasiskirstymo funkcija $F(x)$. Įrodykite, kad asimptotiškai $(n \to \infty)$

$$2\sqrt{n}(F(\hat{x}(1/2)) - 1/2) \xrightarrow{d} Z \sim N(0,\, 1).$$

I.1.55. Įrodykite, kad eksponentinio skirstinio $\mathcal{E}(\lambda)$ empirinė mediana $\hat{x}(1/2)$, gauta iš didumo $2n+1$ paprastosios imties, turi tokį asimptotinį $(n \to \infty)$ skirstinį

$$\lambda\sqrt{2n+1}(\hat{x}(1/2) - \frac{\ln 2}{\lambda}) \xrightarrow{d} Z \sim N(0,\, 1).$$

I.1.56. Variacinė eilutė $X_{(1)} \leq ... \leq X_{(n)}$ gauta stebint tolydųjį a. d. su pasiskirstymo funkcija $F(x)$. Tegu $Y_{(i)} = F(X_{(i)}), i = 1, ..., n$ ir $Z_i = Y_{(i)} - Y_{(i-1)}, i = 1, ..., n+1, Y_{(0)} = 0, Y_{(n+1)} = 1$. Įrodykite, kad fiksuoto skaičiaus k skirtingų a. d. Z_i suma S_k asimptotiškai $(n \to \infty)$ turi gama skirstinį $G(1, k)$.

I.1.57. (**I.1.56** pratimo tęsinys). Raskite asimptotinius skirstinius statistikų $X_{(k)}$ ir $X_{(n-k+1)}$, kai stebėtas a. d. $X \sim U(a, b)$.

I.1.58. (**I.1.56** pratimo tęsinys). Raskite asimptotinius skirstinius statistikų $X_{(k)}$ ir $X_{(n-k+1)}$, kai stebėtas a. d. $X \sim \mathcal{E}(\lambda)$.

I.1.4. Empiriniai momentai ir jų funkcijos

I.1.59. Pagal didumo $n_1, n_2, ..., n_k$ paprastąsias nepriklausomas imtis gauti empiriniai vidurkiai $\bar{X}_i$ ir nepaslinktos empirinės dispersijos s_i^2, $i = 1, 2, ..., k$. Raskite jungtinės didumo $n = n_1 + ... + n_k$ imties empirinį vidurkį $\bar{X}$ ir nepaslinktąją empirinę dispersiją s^2.

I.1.60. Ekspertų grupė vertino kino kadrų, nufilmuotų dviejų tipų kino juostomis, kokybę. Gauti šie rezultatai:

I tipo kino juosta				II tipo kino juosta			
i	n_i	$\bar{X}_i$	s_i^2	i	n_i	$\bar{X}_i$	s_i^2
1	20	25	6	1	10	21	6
2	10	23	5	2	10	18	25
3	10	21	4	3	10	17	5
4	10	18	4	4	9	17	5
5	10	22	9				

Čia n_i yra i-osios imties didumas, $\bar{X}_i$ – empirinis vidurkis, s_i^2 – nepaslinktasis dispersijos įvertis. Raskite I ir II tipo kino juostų jungtinių imčių empirinius vidurkius ir nepaslinktąsias empirines dispersijas.

I.1.61. Pagal didumo n imtį gauta empirinis vidurkis $\bar{X}$ ir nepaslinktoji empirinė dispersija s^2. Kaip perskaičiuoti šias empirines charakteristikas, jeigu imtis papildyta dar vienu nepriklausomu stebėjimu?

I.1.62. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis ir $\bar{X} = (X_1 + ... + X_n)/n$ – aritmetinis vidurkis. Įrodykite, kad a) jeigu $X_i \sim N(\mu, \sigma^2)$, tai $\bar{X} \sim N(\mu, \sigma^2/n)$; b) jeigu $X_i \sim K(\mu, \sigma)$, tai $\bar{X} \sim K(\mu, \sigma)$; c) jeigu $X_i \sim G(\lambda, \eta)$, tai $n\bar{X} \sim G(\lambda, n\eta)$; d) jeigu $X_i \sim B(k, p)$, tai $n\bar{X} \sim B(nk, p)$; d) jeigu $X_i \sim \mathcal{P}(\lambda)$, tai $n\bar{X} \sim \mathcal{P}(n\lambda)$.

I.1.63. Tegu $\bar{X}$, s_1^2 ir $\bar{Y}$, s_2^2 yra empiriniai vidurkiai ir nepaslinktosios empirinės dispersijos dviejų nepriklausomų didumo m ir n imčių, gautų stebint a. d. $X \sim N(\mu_1, \sigma^2)$ ir $Y \sim N(\mu_2, \sigma^2)$. Įrodykite, kad

$$t = (\bar{X} - \bar{Y} - (\mu_1 - \mu_2)) / \sqrt{\frac{(m-1)s_1^2 + (n-1)s_2^2}{m+n-2} \frac{m+n}{mn}} \sim S(m+n-2).$$

I.1.64. Yra k nepriklausomų a. d. $X \sim N(\mu,\ \sigma^2)$ imčių, kurių didumai – $n_1, ..., n_k$. Tegu $\bar{X}_i$ ir s_i^2 yra i-osios imties empirinis vidurkis ir nepaslinktoji empirinė dispersija. Įrodykite, kad funkcijos

$$U = \sum_{i=1}^{k} (n_i - 1) \frac{s_i^2}{\sigma_i^2}, \qquad V = \sum_{i=1}^{k} n_i \frac{(\bar{X}_i - \bar{X})^2}{\sigma_i^2},$$

$$W = n \frac{(\bar{X} - \mu)^2}{\sigma^2}, \quad (\bar{X} = \frac{\sum_i n_i \bar{X}_i}{n},\ n = \sum_i n_i)$$

yra n. a. d., turintys χ^2 skirstinius.

I.1.65. (**I.1.64** pratimo tęsinys). Įrodykite, kad

$$\frac{W(n-k)}{U} \sim F(1, n-k), \quad \frac{V(n-k)}{U(k-1)} \sim F(k-1, n-k).$$

I.1.66. Tarkime, kad $(X_{1i}, X_{2i})^T, i = 1, 2, ..., n$ yra paprastoji imtis, gauta stebint dvimatį normalųjį a. v. $(X_1, X_2)^T \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\mu} = (\mu_1, \mu_2)^T, \boldsymbol{\Sigma} = [\sigma_{ij}]_{2\times 2}$, $\sigma_{11} = \sigma_1^2, \sigma_{22} = \sigma_2^2, \sigma_{12} = \rho\sigma_1\sigma_2$. Tegu $\bar{X}_1$ ir $\bar{X}_2$ yra empiriniai vidurkiai, o s_1^2 ir s_2^2 – nepaslinktosios empirinės dispersijos ir r – empirinis koreliacijos koeficientas. Įrodykite, kad

$$\sqrt{n}(\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2))/\sqrt{s_1^2 + s_2^2 + rs_1s_2} \sim S(n-1).$$

I.1.67. Paprastosios imtys $\mathbf{X} = (X_1, ..., X_n)^T$ ir $\boldsymbol{Y} = (Y_1, ..., Y_m)^T$ gautos stebint n. a. d. X ir Y su vienodomis dispersijomis σ^2. Raskite statistikos $\bar{Z} - \bar{X}$ vidurkį ir dispersiją, kai $\bar{Z} = (m\bar{Y} + n\bar{X})/(m+n)$.

I.1.68. Tegu $X \sim N(\mu, \sigma^2)$ ir $F(x) = \mathbf{P}\{X \le x|\mu, \sigma\}$. Įrodykite, kad a. d. X ir $F(X)$ koreliacijos koeficientas yra $\sqrt{3/\pi}$.

I.1.69. Tegu $Y_1, ..., Y_n$ yra n. a. d. ir $Y_i \sim N(\alpha + \beta(x_i - \bar{x}), \sigma^2)$, $i = 1, 2, ..., n$; čia $\alpha, \beta, x_1, ..., x_n, \bar{x} = \sum x_i/n$ – konstantos. Tegu

$$SS_E = \min_{\alpha,\beta} SS(\alpha, \beta) = \min_{\alpha,\beta} \sum_{i=1}^{n} (Y_i - \alpha - \beta(x_i - \bar{x}))^2 = \sum_{i=1}^{n} (Y_i - \hat{\alpha} - \hat{\beta}(x_i - \bar{x}))^2.$$

Įrodykite, kad a) $\hat{\alpha}$ ir $\hat{\beta}$ yra nepriklausomi ir turi normaliuosius skirstinius su vidurkiais α ir β ir dispersijomis σ^2/n ir $\sigma^2/\sum_i (x_i - \bar{x})^2$; b) SS_E nepriklauso nuo $\hat{\alpha}$ ir $\hat{\beta}$, o $SS_E/\sigma^2 \sim \chi^2(n-2)$.

I.1.70. Tegu $X_1, ..., X_n$ yra paprastoji imtis, gauta stebint a. d. $X \sim N(\mu, \sigma^2)$. Pažymėkime

$$Y_j = X_{j+1} - \frac{1}{1+\sqrt{n}} X_1 - \frac{n}{n+\sqrt{n}} \bar{X}, \quad j = 1, 2, ..., n-1.$$

Įrodykite, kad a. d. $Y_1, ..., Y_{n-1}$ yra nepriklausomi ir $Y_i \sim N(0, \sigma^2)$.

I.1.71. Tegu $X_1, ..., X_n$ yra paprastoji imtis, gauta stebint a. d. $X \sim N(\mu, \sigma^2)$, ir tegu d empirinis aritmetinis nuokrypis nuo μ

$$d = \frac{1}{n} \sum_{i=1}^{n} |X_i - \mu|.$$

Įrodykite, kad

$$\mathbf{E}d = \sqrt{2/\pi}\sigma, \quad \boldsymbol{V}d = (\sigma^2/n)(1 - 2/\pi).$$

I.1.72. Paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. X, kurio ketvirtas momentas baigtinis $\mathbf{E}(X^4) < \infty$. Raskite empirinio vidurkio $\bar{X}$ vidurkį, dispersiją, asimetrijos ir eksceso koeficientus.

I.1.73. Paprastoji didumo n imtis gauta stebint a. d. X su baigtine dispersija $\mathbf{E}X = \mu$, $\boldsymbol{V}X = \sigma^2 < \infty$. Įrodykite, kad $n \to \infty$

$$\sqrt{n}\frac{\bar{X} - \mu}{\sigma} \xrightarrow{d} Z \sim N(0, 1), \quad \sqrt{n}\frac{\bar{X} - \mu}{\sqrt{m_2}} \xrightarrow{d} Z \sim N(0, 1).$$

I.1.74. Paprastoji didumo n imtis gauta stebint a. d. $X \sim \mathcal{P}(\lambda)$. Įrodykite, kad $n \to \infty$

$$S = \sqrt{n}\frac{\bar{X} - \lambda}{\sqrt{\lambda}} \xrightarrow{d} Z \sim N(0, 1).$$

I.1.75. (**I.1.74** pratimo tęsinys). Įrodykite, kad $n \to \infty$

$$Z = \sqrt{n}(2\sqrt{\bar{X}} - 2\sqrt{\lambda}) \xrightarrow{d} Z \sim N(0,\, 1).$$

I.1.76. (**I.1.74** pratimo tęsinys). Įrodykite, kad $n \to \infty$

$$U = 3\sqrt{n}((\bar{X} + 1/(2n))^{2/3} - \lambda^{2/3})/(2\lambda^{1/6}) \xrightarrow{d} Z \sim N(0,\, 1).$$

I.1.77. Paprastoji didumo n imtis gauta stebint Bernulio a. d. $X \sim B(1, p)$. Įrodykite, kad $n \to \infty$

$$\sqrt{n}\frac{\bar{X} - p}{\sqrt{p(1-p)}} \xrightarrow{d} Z \sim N(0,\, 1), \quad \sqrt{n}\frac{\bar{X} - p}{\sqrt{\bar{X}(1-\bar{X})}} \xrightarrow{d} Z \sim N(0,\, 1).$$

I.1.78. (**I.1.77** pratimo tęsinys). Įrodykite, kad $n \to \infty$

$$\sqrt{n}(\arcsin(2\bar{X} - 1) - \arcsin(2p - 1)) \xrightarrow{d} Z \sim N(0,\, 1).$$

I.1.79. Paprastoji didumo n imtis gauta stebint a. d. $X \sim U(0, \theta), \theta > 0$. Įrodykite, kad

$$\mathbf{E}_\theta(2\bar{X}) = \theta, \quad \boldsymbol{V}_\theta(2\bar{X}) = \frac{\theta^2}{3n} = O(\frac{1}{n}), \quad \sqrt{3n}\frac{2\bar{X} - \theta}{\theta} \xrightarrow{d} Z \sim N(0,\, 1).$$

I.1.80. (**I.1.79** pratimo tęsinys). Įrodykite, kad

$$\mathbf{E}_\theta(\frac{n+1}{n}X_{(n)}) = \theta, \quad \boldsymbol{V}_\theta(\frac{n+1}{n}X_{(n)}) = \frac{\theta^2}{n(n+2)} = O(\frac{1}{n^2}).$$

I.1.81. (**I.1.79** pratimo tęsinys). Įrodykite, kad

$$\sqrt{3n}(\ln(2\bar{X}) - \ln\theta) \xrightarrow{d} Z \sim N(0,\, 1).$$

I.1.82. Paprastoji didumo n imtis gauta stebint geometrinį a. d., kurio skirstinys nusakytas tikimybėmis

$$\mathbf{P}_p\{X = k\} = pq^{k-1}, \quad q = 1 - p, \quad k = 1, 2, ...$$

Įrodykite, kad $n \to \infty$

$$\sqrt{n}(p\bar{X} - 1)/\sqrt{q} \xrightarrow{d} Z \sim N(0,\, 1).$$

I.1.83. (**I.1.82** pratimo tęsinys). Įrodykite, kad $n \to \infty$

$$\sqrt{n}\{\ln[\bar{X}(1 + \sqrt{1 - 1/\bar{X}}) - 1/2] - \ln[(1 + \sqrt{q})/p - 1/2]\} \xrightarrow{d} Z \sim N(0,\, 1).$$

I.1.84. Paprastoji didumo n imtis gauta stebint a. d. X, kurio antras momentas baigtinis $\boldsymbol{V}X = \sigma^2 < \infty$. Patikrinkite, kad empirinės dispersijos vidurkis $\mathbf{E}_\sigma m_2 \neq \sigma^2$. Nepaslinktoji empirinė dispersija $s^2 = nm_2/(n-1)$ ir $\mathbf{E}_\sigma s^2 = \sigma^2$.

I.1.85. (**I.1.84** pratimo tęsinys). Tarkime, kad yra baigtinis ir ketvirtas momentas $\mu_4 = \mathbf{E}(X - \mathbf{E}X)^4 < \infty$. Raskite empirinės dispersijos m_2 dispersiją $\boldsymbol{V}m_2$.

I.1.86. Paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. X, turintį baigtinį momentą $\alpha_{2k} = \mathbf{E}X^{2k}$. Įrodykite, kad $n \to \infty$

$$a_k \xrightarrow{b.t.} \alpha_k, \quad \sqrt{n}(a_k - \alpha_k) \xrightarrow{d} Y \sim N(0, \alpha_{2k} - \alpha_k^2).$$

I.1.87. Paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. X, turintį baigtinį momentą $\alpha_{2k} = \mathbf{E}X^{2k}$. Įrodykite, kad $n \to \infty$

$$\sqrt{n}(\boldsymbol{a} - \boldsymbol{\alpha}) \xrightarrow{d} \boldsymbol{Y} \sim N_k(\mathbf{0}, \boldsymbol{\Sigma}),$$

čia $\boldsymbol{a} = (a_1, a_2, ..., a_k)^T, a_j = \sum_i X_i^j/n, j = 1, ..., k$; $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_k)^T$; $\boldsymbol{\Sigma} = [\sigma_{ij}]_{k\times k}$, $\sigma_{ij} = \alpha_{i+j} - \alpha_i\alpha_j, i, j = 1, ..., k$.

I.1.88. Įrodykite, kad $n \to \infty$ a) $m_2 \xrightarrow{P} \sigma^2$, $s^2 \xrightarrow{P} \sigma^2$.

$$\text{b)}\ \frac{m_2 - \sigma^2}{\sqrt{\boldsymbol{V}m_2}} \xrightarrow{d} Z \sim N(0,\,1), \quad \sqrt{n}\frac{m_2 - \sigma^2}{\sqrt{\mu_4 - \sigma^4}} \xrightarrow{d} Z \sim N(0,\,1),$$

$$\sqrt{n}\frac{s^2 - \sigma^2}{\sqrt{\mu_4 - \sigma^4}} \xrightarrow{d} Z \sim N(0,\,1).$$

c) normaliojo skirstinio atveju

$$nm_2/\sigma^2 \sim \chi^2(n-1), \quad (n-1)s^2/\sigma^2 \sim \chi^2(n-1).$$

I.1.89. Pagal dvi didumo n_1 ir n_2 nepriklausomas paprastąsias imtis, gautas stebint a. d. X ir Y su vidurkiais $\mu_1 = \mathbf{E}X$ ir $\mu_2 = \mathbf{E}Y$, turinčiais baigtines dispersijas, gauti empiriniai vidurkiai ir empirinės dispersijos $\bar{X}, s_1^2$ ir $\bar{Y}, s_2^2$. Įrodykite, kad $n_1, n_2 \to \infty$

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} \xrightarrow{d} Z \sim N(0,\,1).$$

I.1.90. Didumo n paprastoji imtis gauta stebint a. d. X su baigtiniu trečiuoju momentu $\mathbf{E}|X|^3 < \infty$. Raskite a. d. $\bar{X}$ ir s^2 kovariaciją.

I.1.91. Paprastoji didumo n imtis gauta stebint a. d. X su baigtiniu ketvirtu momentu $\mu_4 = \mathbf{E}(X - \mathbf{E}X)^4) < \infty$, $\sigma^2 = \boldsymbol{V}X$. Patikrinkite, kad $\mathbf{E}m_3 \neq \mu_3$, $\mathbf{E}m_4 \neq \mu_4$. Kaip reikia pataisyti šias empirines charakteristikas, kad gautųjų statistikų vidurkiai sutaptų su μ_3 ir μ_4.

I.1.92. Paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. X, turintį baigtinį momentą $\alpha_{2k} = \mathbf{E}X^{2k}$. Įrodykite, kad $n \to \infty$

$$m_k \xrightarrow{b.t.} \mu_k, \quad \sqrt{n}(m_k - \mu_k) \xrightarrow{d} Y \sim N(0, B_k^2),$$

čia $B_k^2 = \mu_{2k} - \mu_k^2 - 2k\mu_{k-1}\mu_{k+1} + k^2\mu_{k-1}^2\sigma^2$, o kai stebėtas normalusis a. d., tai $B_k^2 = \sigma^{2k}[(2k-1)!! - ((k-1)!!)^2]$, kai k lyginis, ir $B_k^2 = \sigma^{2k}[(2k-1)!! - 2k(k-2)!!(k-1)!! + k^2(k-2)!!]$, kai k nelyginis.

I.1.93. (**I.1.92** pratimo tęsinys). Tarkime, egzistuoja momentas $\mathbf{E}(X^{2k+2l}) < \infty$. Įrodykite, kad $n \to \infty$

$$\sqrt{n}(m_k - \mu_k, m_l - \mu_l) \xrightarrow{d} (Y, Z) \sim N_2(\mathbf{0}, \boldsymbol{\Sigma}),$$

čia $\mathbf{\Sigma} = [\sigma_{ij}]_{2\times 2}, \sigma_{11} = B_k^2, \sigma_{22} = B_l^2, \sigma_{12} = B_{kl}$, $B_{kl} = \mu_{k+l} - k\mu_{l+1}\mu_{k-1} - l\mu_{l-1}\mu_{k+1} - \mu_k\mu_l + kl\mu k - 1\mu_{l-1}\sigma^2$.

I.1.94. Tegu paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $X \sim \mathcal{P}(\lambda), \lambda > 0$. Nagrinėkime statistiką $H(\bar{X}) = 1/\bar{X}^3$. Įrodykite, kad $n \to \infty$

$$\sqrt{n}(1/\bar{X}^3 - 1/\lambda^3) \xrightarrow{d} Z \sim N(0, 9/\lambda^7).$$

Pažymėsime, kad funkcijos $H(\bar{X})$ dispersija neegzistuoja, nes vardiklis su teigiama tikimybe įgyja reikšmę 0: $\mathbf{P}\{\bar{X} = 0\} = e^{-n\lambda} > 0$.

I.1.95. Tarkime, $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji didumo n imtis. Įrodykite, kad

a) jeigu egzistuoja $\mu_4 = \mathbf{E}(X_i - \mathbf{E}X_i)^4$, tai

$$\mathbf{E}s = \sigma + O(1/n), \quad \boldsymbol{V}s = \frac{\mu_4 - \sigma^4}{4n\sigma^2} + O(1/n^{3/2}),$$

o kai skirstinys normalusis

$$\mathbf{E}s = \sigma M_{n-1}, \quad \boldsymbol{V}s = \sigma^2(1 - M_{n-1}^2), \quad M_n = \sqrt{\frac{2}{n}}\frac{\Gamma((n+1)/2)}{\Gamma(n/2)};$$

b) jeigu egzistuoja $\mu_6 = \mathbf{E}(X_i - \mathbf{E}X_i)^6$, tai

$$\mathbf{E}g_1 = \gamma_1 + O(1/n), \quad \boldsymbol{V}g_1 = \frac{4\mu_6\sigma^4 - 12\mu_5\mu_3\sigma^2 + 9\mu_4\mu_3^2 + 35\mu_3^2\sigma^4}{4n\sigma^{10}}$$

$$+\frac{9\sigma^4 - 6\mu_4}{n\sigma^4} + O(1/n^{3/2}),$$

o kai skirstinys normalusis

$$\mathbf{E}g_1 = 0, \quad \boldsymbol{V}g_1 = \frac{6(n-1)}{(n+1)(n+3)};$$

c) jeigu egzistuoja $\mu_8 = \mathbf{E}(X_i - \mathbf{E}X_i)^8$, tai

$$\mathbf{E}g_2 = \gamma_2 + O(1/n), \quad \boldsymbol{V}g_2 = \frac{\mu_8\sigma^4 - 4\mu_6\mu_4\sigma^2 + 4\mu_4^3 + 16\mu_4\mu_3^2\sigma^2}{n\sigma^{12}}$$

$$+\frac{16\mu_3^2\sigma^2 - \mu_4^2 - 8\mu_5\mu_3}{n\sigma^8} + O(1/n^{3/2}),$$

o kai skirstinys normalusis

$$\mathbf{E}g_2 = -\frac{6}{n+1}, \quad \boldsymbol{V}g_2 = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}.$$

Visais atvejais statistikų asimptotiniai skirstiniai yra normalieji su pateiktomis asimptotinėmis dispersijomis.

I.1.96. Tegu $(X_i, Y_i)^T, i = 1, 2, ..., n$, yra didumo n paprastoji imtis, gauta stebint a. v. $(X, Y)^T$. Įrodykite, kad

a) jeigu egzistuoja $\mu_{22} = \mathbf{E}[(X - \mathbf{E}X)^2(Y - \mathbf{E}Y)^2]$, tai

$$\mathbf{E}m_{11} = \frac{n-1}{n}\mu_{11}, \quad \boldsymbol{V}m_{11} = \frac{\mu_{22} - \mu_{11}^2}{n} + O(1/n^2);$$

b) jeigu egzistuoja $\mu_{66} = \mathbf{E}[(X - \mathbf{E}X)^6(Y - \mathbf{E}Y)^6]$, tai

$$\mathbf{Cov}\,(m_{20}, m_{11}) = \frac{\mu_{31} - \mu_{20}\mu_{11}}{n} + O(1/n^2), \quad \mathbf{Cov}\,(m_{20}, m_{02}) = \frac{\mu_{22} - \mu_{20}\mu_{02}}{n} + O(1/n^2);$$

o kai skirstinys normalusis

$$\mathbf{Cov}\,(m_{20}, m_{11}) = \frac{2\rho\sigma_1^3\sigma_2}{n} + O(1/n^2), \quad \mathbf{Cov}\,(m_{20}, m_{02}) = \frac{2\rho^2\sigma_1^2\sigma_2^2}{n} + O(1/n^2).$$

I.1.97. (**1.96** pratimo tęsinys). Įrodykite, kad jei egzistuoja

$$\mu_{66} = \mathbf{E}[(X - \mathbf{E}X)^6(Y - \mathbf{E}Y)^6],$$

tai $\mathbf{E}r = \rho + O(1/n)$,

$$\boldsymbol{V}r = \frac{\rho^2}{4n}\left(\frac{\mu_{40}}{\mu_{20}^2} + \frac{\mu_{04}}{\mu_{02}^2} + \frac{2\mu_{22}}{\mu_{20}\mu_{02}} + \frac{4\mu_{22}}{\mu_{11}^2} - \frac{4\mu_{31}}{\mu_{11}\mu_{20}} - \frac{4\mu_{13}}{\mu_{11}\mu_{02}}\right) + O(1/n^{3/2}) =$$

$$b_r^2/n + O(1/n^{3/2}), \quad \sqrt{n}(r - \rho) \xrightarrow{d} Y \sim N(0, b_r^2),$$

o kai skirstinys normalusis, tai $b_r^2 = (1 - \rho^2)^2$.

I.1.98. Tegu r yra empirinis koreliacijos koeficientas, gautas iš didumo n paprastosios dvimačio normaliojo vektoriaus imties. Įrodykite, kad $n \to \infty$ (Fišerio transformacija)

$$\sqrt{n}\left(\frac{1}{2}\ln\frac{1+r}{1-r} - \frac{1}{2}\ln\frac{1+\rho}{1-\rho}\right) \xrightarrow{d} Z \sim N(0,\,1).$$

I.1.5. Sprendimai, nurodymai, atsakymai

I.1.1 skyrelis

I.1.1. A. d. X_1 ir X_2 yra nepriklausomi binominiai a. d. Imties $(X_1,\ X_2)^T$ skirstinys yra diskretusis, nusakomas tikimybėmis $\mathbf{P}_p\{X_1 = m_1,\ X_2 = m_2\} = C_{n_1}^{m_1} C_{n_2}^{m_2} p^{m_1+m_2}(1-p)^{(n_1+n_2-m_1-m_2)}$, $m_1 = 0, 1, ..., n_1,\ m_2 = 0, 1, ..., n_2,\ 0 < p < 1$; $\mathbf{E}_pT = \mathbf{E}_p(X_1/n_1) - \mathbf{E}_p(X_2/n_2) = p - p = 0,\ \boldsymbol{V}_pT = \boldsymbol{V}_pX_1/n_1^2 + \boldsymbol{V}_pX_2/n_2^2 = p(1-p)(n_1+n_2)/(n_1n_2)$.

I.1.2. a) Tegu X_i yra a. d., kuris įgyja reikšmę 1, jeigu i-uoju ėmimu paimtas objektas, turintis savybę A, ir reikšmę 0 – priešingu atveju. Tada imties $\mathbf{X} = (X_1, ..., X_n)^T$ tikimybinis modelis nusakomas tikimybėmis

$$\mathbf{P}\{X_i = j_i,\ i = 1, ..., n\} = \frac{1}{k}\sum_{r=1}^{k}(\pi_r)^{\sum_i j_i} \times (1 - \pi_r)^{n - \sum_i j_i}, \quad j_i = 0,\ 1,$$

$$0 < \pi_j < 1, \quad j = 1, ..., k.$$

Tarkime, kad B_j yra atsitiktinis įvykis, kuris reiškia, kad parinkta j-oji populiacija, $\mathbf{P}\{B_j\} = 1/k, j = 1, ..., k$. Jeigu įvyko B_j, tai sąlyginis X skirstinys yra binominis $(X|B_j) \sim B(n, \pi_j)$; $\mathbf{E}(X|B_j) = n\pi_j$, $\mathbf{E}(X^2|B_j) = n\pi_j(1-\pi_j) + n^2\pi_j^2$. Tada $\mathbf{E}(X/n) = \frac{1}{k}\sum_{r=1}^k \pi_r = \bar{\pi}$, $\boldsymbol{V}(X/n) = \mathbf{E}(X/n)^2 - (\mathbf{E}(X/n))^2 = \frac{1}{kn}\sum_{r=1}^k \pi_r(1-\pi_r) + \frac{1}{k}\sum_{r=1}^k \pi_r^2 - (\frac{1}{k}\sum_{r=1}^k \pi_r)^2$.

b) Eksperimentą galima traktuoti kaip Bernulio eksperimentus, kuriuose įvykis A pasirodo su tikimybe $\bar{\pi}$. Imties $\mathbf{X} = (X_1, ..., X_n)^T$ tikimybinis modelis nusakomas tikimybėmis $\mathbf{P}\{X_1 = j_1, ..., X_n = j_n\} = (\bar{\pi})^{\sum_i j_i}(1-\bar{\pi})^{n-\sum_i j_i}$, $j_i = 0, 1$. A.d. X turi binominį skirstinį $X \sim B(n, \bar{\pi})$; $\mathbf{E}(X/n) = \bar{\pi}$, $\boldsymbol{V}(X/n) = \bar{\pi}(1-\bar{\pi})/n$.

I.1.3. Jungtinės imties $\mathbf{X} = (\mathbf{X}_1^T, ..., \mathbf{X}_n^T)^T$ tikimybinis modelis nusakomas tikimybėmis $\mathbf{P}\{\mathbf{X}_1 = \boldsymbol{x}_1, ..., \mathbf{X}_n = \boldsymbol{x}_n\} = \mathbf{P}\{\mathbf{X}_1 = \boldsymbol{x}_1\}...\mathbf{P}\{\mathbf{X}_n = \boldsymbol{x}_n\}$; $\mathbf{P}\{\mathbf{X}_i = \boldsymbol{x}_i\} = \mathbf{P}\{X_{i1} = j_1, ..., X_{in} = j_n\} = \pi_i^{\sum_i j_i}(1-\pi_i)^{n-\sum_i j_i}$, $j_1, ..., j_n = 0,\ 1$. Atsitiktinis dydis X yra suma $X = X_1 + ... + X_k$, kai $X_1, ..., X_k$ yra nepriklausomi a.d. ir $X_j \sim B(n, \pi_j)$; $\mathbf{E}(X/n) = \sum_{r=1}^k \pi_r$, $\boldsymbol{V}(X/n) = \frac{1}{n}\sum_{r=1}^k \pi_r(1-\pi_r)$.

I.1.4. Jungtinės imties $\mathbf{X} = (X_{11}, ..., X_{1n_1},\ X_{21}, ..., X_{2n_2})^T$ tikimybinis skirstinys nusakomas tikimybėmis $\mathbf{P}_\pi\{X_{1i} = j_i, X_{2l} = k_l;\ i = 1, ..., n_1,\ l = 1, ..., n_2\} = \pi^{\sum_i j_i}(1-\pi)^{n_1-\sum_i j_i}(\sum_i j_i/n_1)^{\sum_l k_l}$, $(1-\sum_i j_i/n_1)^{n_2-\sum_l k_l}$, $j_i, k_l = 0,\ 1,\ 0 < \pi < 1$.

I.1.5. A.d. $X_1 \sim B(n_1, \pi)$, o a.d. X_2 sąlyginis skirstinys, kai $X_1 = m_1$ fiksuotas, yra binominis $B(n_2, m_1/n_1)$. Taigi a.v. $(X_1,\ X_2)^T$ skirstinys nusakomas tikimybėmis

$$\mathbf{P}_\pi\{X_1 = m_1,\ X_2 = m_2\} = C_{n_1}^{m_1}\pi^{m_1}(1-\pi)^{n_1-m_1}C_{n_2}^{m_2}(m_1/n_1)^{m_2}(1-m_1/n_1)^{n_2-m_2},$$

$$m_1 = 0, ..., n_1,\ m_2 = 0, ..., n_2,\ 0 < \pi < 1.$$

Randame

$$\mathbf{E}_\pi X_1 = n_1\pi, \quad \mathbf{E}_\pi X_2 = \mathbf{E}_\pi\{\mathbf{E}(X_2|X_1)\} = \mathbf{E}_\pi(n_2X_1/n_1) = n_2\pi, \quad \mathbf{E}_\pi(\frac{X_1}{n_1} - \frac{X_2}{n_2}) = 0;$$

$$\boldsymbol{V}_\pi(X_1/n_1 - X_2/n_2) = \mathbf{E}_\pi(X_1/n_1)^2 - 2\mathbf{E}_\pi(X_1X_2/(n_1n_2) + \mathbf{E}_\pi(X_2/n_2)^2;$$

$$\mathbf{E}_\pi X_1^2 = \boldsymbol{V}_\pi X_1 + (\mathbf{E}_\pi X_1)^2 = n_1\pi(1-\pi) + n_1^2\pi^2;$$

$$\mathbf{E}_\pi(X_1X_2) = \mathbf{E}_\pi\{X_1\mathbf{E}(X_2|X_1)\} = \mathbf{E}_\pi(n_2X_1^2/n_1) = n_2\pi(1-\pi) + n_1n_2\pi^2;$$

$$\mathbf{E}_\pi X_2^2 = \mathbf{E}_\pi(n_2\frac{X_1}{n_1}(1 - \frac{X_1}{n_1}) + n_2^2(\frac{X_1}{n_1})^2) = n_2\pi + \frac{n_2(n_2-1)}{n_1}\pi(1-\pi) + n_2(n_2-1)\pi^2.$$

Įrašę į dispersijos išraišką ir sutraukę panašiuosius narius, gausime

$$\boldsymbol{V}_\pi(X_1/n_1 - X_2/n_2) = \frac{(n_1-1)\pi(1-\pi)}{n_1n_2}.$$

I.1.6. Populiacijos objektų, turinčių savybę A, skaičius yra $M = N\pi$ ir $\pi = M/N$. Tada a.d. X_1 skirstinys yra hipergeometrinis $X_1 \sim H(N, M, n_1)$, o a.d. X_2 sąlyginis skirstinys, esant sąlygai, kad a.d. X_1 įgijo reikšmę m, taip pat hipergeometrinis $(X_2|X_1 = m) \sim H(n_1, m, n_2)$.

Randame

$$\mathbf{E}_\pi X_1 = n_1\pi, \quad \mathbf{E}_\pi X_2 = \mathbf{E}_\pi\{\mathbf{E}(X_2|X_1)\} = \mathbf{E}_\pi(n_2X_1/n_1) = n_2\pi, \quad \mathbf{E}_\pi(\frac{X_1}{n_1} - \frac{X_2}{n_2}) = 0;$$

$$\boldsymbol{V}_\pi(X_1/n_1 - X_2/n_2) = \mathbf{E}_\pi(X_1/n_1)^2 - 2\mathbf{E}_\pi(X_1X_2/(n_1n_2) + \mathbf{E}_\pi(X_2/n_2)^2;$$

$$\mathbf{E}_\pi X_1^2 = \boldsymbol{V}_\pi X_1 + (\mathbf{E}_\pi X_1)^2 = n_1\pi(1-\pi)\frac{N-n_1}{N-1} + n_1^2\pi^2;$$

$$\mathbf{E}_\pi(X_1X_2) = \mathbf{E}_\pi\{X_1\mathbf{E}(X_2|X_1)\} = \mathbf{E}_\pi(n_2X_1^2/n_1) = n_2\pi(1-\pi)\frac{N-n_1}{N-1} + n_1n_2\pi^2;$$

$$\mathbf{E}_\pi X_2^2 = \mathbf{E}_\pi(n_2\frac{X_1}{n_1}(1-\frac{X_1}{n_1}) + n_2^2(\frac{X_1}{n_1})^2) = n_2\pi + \frac{n_2(n_2-1)}{n_1}\frac{N-n_1}{N-1}\pi(1-\pi) + n_2(n_2-1)\pi^2.$$

Įrašę į dispersijos išraišką ir sutraukę panašiuosius narius, gausime

$$\boldsymbol{V}_\pi(X_1/n_1 - X_2/n_2) = \frac{(n_1-1)N}{n_1n_2(N-1)}\frac{M}{N}(1-\frac{M}{N}), \quad \pi = \frac{M}{N}.$$

I.1.7. A. d. X skirstinys yra hipergeometrinis $X \sim H(N, M, n)$.

I.1.8. Partijos priėmimo tikimybė

$$\mathbf{P}\{X \le c|M\} = \sum_{m=0}^{c} h(m|N,\ M,\ n) = H(c|N,\ M,\ n),$$

čia $h(m|N,\ M,\ n) = C_M^m C_{N-M}^{n-m}/C_N^n$.

I.1.9. a) Atsitiktinio vektoriaus $\mathbf{X} = (X_1, ..., X_n)^T$ koordinatės priklausomos: koordinatė X_i negali įgyti reikšmės y_j, jei kuri nors kita koordinatė jau įgijo šią reikšmę. A. v. $\mathbf{X}$ galimos reikšmės yra visi galimi skirtingi rinkiniai $(y_{i_1}, y_{i_2}, ..., y_{i_n})^T$ iš aibės $\{y_1, ..., y_N\}$ (visi indeksai $i_1, i_2, ..., i_n$ skirtingi); jų įgijimo tikimybės visos vienodos ir lygios $1/C_N^n$.

b)

$$\mathbf{E}Y = \sum_{i=1}^{n}\mathbf{E}X_i = \frac{n}{N}\sum_{i=1}^{N} y_i = n\bar{y};$$

$$\mathbf{E}Y^2 = \mathbf{E}(\sum_{i=1}^{n}X_i^2 + \sum_{i\neq j}X_iX_j) = \frac{n}{N}\sum_{j=1}^{N}y_j^2 + \frac{n(n-1)}{N(N-1)}\sum_{i\neq j}y_iy_j =$$

$$\frac{n}{N}\sum_{j=1}^{N}y_j^2 + \frac{n(n-1)}{N(N-1)}[(\sum_{i=1}^{N}y_i)^2 - \sum_{i=1}^{N}y_i^2].$$

Gauname

$$\boldsymbol{V}Y = \mathbf{E}Y^2 - (\mathbf{E}Y)^2 = \frac{n(N-n)}{N(N-1)}\sum_{i=1}^{N}y_i^2 - \frac{n(N-n)}{N^2(N-1)}(\sum_{i=1}^{N}y_i)^2.$$

Jeigu pažymėsime $s_y^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^2$ aibės reikšmių dispersiją, tai $\boldsymbol{V}Y = ns_y^2(N-n)/(N-1)$.

c) Tarkime, vienetų skaičius aibėje O_N lygus M. Tada $Y \sim H(N,\ M,\ n)$.

I.1.10. a) Atsitiktinio vektoriaus $\mathbf{X} = (X_1, ..., X_n)^T$ koordinatės yra nepriklausomos; a. v. $\mathbf{X}$ reikšmės yra visi galimi rinkiniai $(y_{i_1}, y_{i_2}, ..., y_{i_n})^T$, kurių kiekviena koordinatė nepriklausomai nuo kitų gali įgyti reikšmes $y_1, ..., y_N$ su tikimybėmis $1/N$, t. y. kiekvieno rinkinio tikimybė yra $1/N^n$; b) $\mathbf{E}Y = n\bar{y}$, $\boldsymbol{V}Y = ns_y^2$; c) Jeigu vienetų skaičius aibėje yra M, tai $Y \sim B(n,\ p)$, $p = M/N$.

I.1.11. Jungtinės imties $(X^T,\ Y^T)^T$ tikimybinis skirstinys nusakomas tikimybėmis $\mathbf{P}\{X_1 = k_1, ..., X_m = k_m,\ Y_1 = l_1, ..., Y_n = l_n\} = \lambda_1^{k_1+...+k_m}\ e^{-m\lambda_1}\lambda_2^{l_1+...+l_n}e^{-n\lambda_2}/ (k_1!...k_m!l_1!...l_n!)$, $0 < \lambda_1,\ \lambda_2 < \infty,\ k_j,\ l_i = 0, 1, \cdots$. Jeigu $\lambda_1 = \lambda_2 = \lambda$, tai ši tikimybė yra $\lambda^{k_1+...+k_m+l_1+...+l_n}\ e^{-(m+n)\lambda}/(k_1!...k_m!l_1!...l_n!)$.

I.1.12. c) Jungtinės imties $(X_1, ..., X_m,\ Y_1, ..., Y_n)^T$ tikimybinio tankio funkcija yra $\prod_{i=1}^m \varphi(x_i|\mu_1,\ \sigma_1)\prod_{j=1}^n \varphi(y_j|\mu_2,\ \sigma_2)$, $-\infty < \mu_1,\ \mu_2 < \infty$, $0 < \sigma_1,\ \sigma_2 < \infty$; čia $\varphi(z|\mu,\ \sigma)$ – normaliojo skirstinio $N(\mu,\ \sigma^2)$ tankio funkcija. Atveju a) reikia įrašyti $\mu_1 = \mu_2 = \mu$, o atveju b) $\sigma_1 = \sigma_2 = \sigma$.

I.1.13. Imties tankio funkcija yra $\prod_{i=1}^n \varphi(x_i,\ y_i|\boldsymbol{\mu},\ \boldsymbol{\Sigma})$, $\boldsymbol{\mu} = (\mu_1,\ \mu_2)^T, -\infty < \mu_1,\ \mu_2 < \infty$, $\boldsymbol{\Sigma} = [\sigma_{ij}]_{2\times 2}$, $0 < \sigma_{11},\ \sigma_{22} < \infty$, $\sigma_{12} = \sigma_{21} = \rho\sqrt{\sigma_{11}\sigma_{22}}, -1 < \rho < 1$; o $\varphi(x,\ y|\boldsymbol{\mu},\ \boldsymbol{\Sigma})$ – dvimačio normaliojo skirstinio $N_2(\boldsymbol{\mu},\ \boldsymbol{\Sigma})$ tankio funkcija. Jeigu a. d. X ir Y marginalieji skirstiniai vienodi, tai reikia įrašyti $\mu_1 = \mu_2 = \mu$, $\sigma_{11} = \sigma_{22} = \sigma^2$.

I.1.14. Atsitiktinio vektoriaus $\mathbf{X}$ skirstinys yra n-matis normalusis $N_n(\boldsymbol{\mu},\ \boldsymbol{\Sigma})$, kurio vidurkių vektorius $\boldsymbol{\mu} = (\mu, ..., \mu)^T$, $-\infty < \mu < \infty$; kovariacinės matricos $\boldsymbol{\Sigma} = [\sigma_{ij}]_{n\times n}$ elementai ant pagrindinės įstrižainės yra σ^2, $0 < \sigma < \infty$; elementai ant įstrižainių, gretimų pagrindinei, yra $\rho\sigma^2$, $-1 < \rho < 1$, o visi kiti elementai lygūs 0.

I.1.15. Jungtinės imties $\mathbf{X}$ pasiskirstymo funkcija F priklauso absoliučiai tolydžių $(n_1 + ... + n_m)$-mačių pasiskirstymo funkcijų aibei $\mathcal{F}$: a) $\mathcal{F}$ yra aibė pasiskirstymo funkcijų $\prod_{j=1}^m \prod_{i=1}^{n_j} F(x_{ij} - \mu_j)$, $-\infty < \mu_1, ..., \mu_m < \infty$; b) $\mathcal{F}$ yra aibė pasiskirstymo funkcijų $\prod_{j=1}^m \prod_{i=1}^{n_j} F(x_{ij}/\sigma_j)$, $0 < \sigma_1, ..., \sigma_m < \infty$; c) $\mathcal{F}$ yra aibė pasiskirstymo funkcijų $\prod_{j=1}^m \prod_{i=1}^{n_j} F((x_{ij} - \mu_j)/\sigma_j)$, $-\infty < \mu_1, ..., \mu_m < \infty$, $0 < \sigma_1, ..., \sigma_m < \infty$; d) $\mathcal{F}$ yra aibė pasiskirstymo funkcijų $\prod_{j=1}^m \prod_{i=1}^{n_j} F_j(x_{ij})$; čia F, F_j – vienmatės absoliučiai tolydžios pasiskirstymo funkcijos.

I.1.16. Imties $\mathbf{X}$ pasiskirstymo funkcija F priklauso absoliučiai tolydžių n-mačių pasiskirstymo funkcijų aibei $\mathcal{F}$ pavidalo $\prod_{i=1}^n F_i(x_i)$; čia $F_1, ..., F_n$ – vienmatės absoliučiai tolydžios pasiskirstymo funkcijos. Jeigu imtis paprastoji, tai pasiskirstymo funkcijos $F_1, ..., F_n$ vienodos.

I.1.17. Imties nario $\mathbf{X}^{(i)}$ pasiskirstymo funkcija F priklauso absoliučiai tolydžių n-mačių pasiskirstymo funkcijų aibei $\mathcal{F} = \{F(x_1, ..., x_n) : F(x, \infty, ..., \infty)$
$= F(\infty, x, ..., \infty) = ... = F(\infty, \infty, ..., x), -\infty < x < \infty\}$.

I.1.18. Atsitiktinio vektoriaus $\mathbf{X} = (X_1, ..., X_n)^T$ tikimybinio tankio funkcija yra $(\sqrt{2\pi}\sigma)^{-n}\exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_{t_i} - x_0 - vt_i)^2\}$, $-\infty < x_0, x_{t_1}, ..., x_{t_n} < \infty, v > 0$.

I.1.19. Į **I.1.18** pratimo tankio formulę reikia įrašyti $x_0 = 0$.

I.1.20. A. v. $\mathbf{X} = (X_1, ..., X_n)^T$ tikimybinio tankio funkcija yra

$$(\sqrt{2\pi}\sigma)^{-n}\exp\{-\frac{1}{2\sigma^2}\sum_{i-1}^{k} x_i^2\}\exp\{-\frac{1}{2\sigma^2}\sum_{i=k+1}^{n}(x_i - \mu)^2\},$$

kai $-\infty < x_1, ..., x_n < \infty$, $-\infty < \mu < \infty$, $\sigma > 0$.

I.1.2 skyrelis

I.1.21. Pažymėkime $p = \mathbf{P}\{X \le 0\} = \Phi(-0,5)$. Remiantis [2], 2.1.3 teorema

$$\sqrt{n}\frac{\hat{F}_n(0) - p}{\sqrt{p(1-p)}} \xrightarrow{d} Z \sim N(0,\, 1).$$

Randame

$$\mathbf{P}_p\{|\hat{F}_n(0) - p| \le \varepsilon\} = \mathbf{P}_p\{-\varepsilon\sqrt{n/(p(1-p))} \le \sqrt{n}\frac{\hat{F}_n(0) - p}{\sqrt{p(1-p)}} \le \varepsilon\sqrt{n/(p(1-p))}\}$$

$$\approx 2\Phi(\varepsilon\sqrt{n/(p(1-p))}) - 1.$$

Kai $n = 100$, $\varepsilon = 0,1; 0,05$, gauname apytiksles tikimybių reikšmes 0,9696; 0,7210.

I.1.22. Pažymėkime $D_n = \sup_x |\hat{F}_n(x) - F(x)|$. Tada remiantis Kolmogorovo teorema ([2], 2.1.4 teorema)

$$\mathbf{P}_F\{D_n \le \varepsilon\} = \mathbf{P}\{\sqrt{n}D_n \le \sqrt{n}\varepsilon\} \approx K(\varepsilon\sqrt{n}).$$

Kai $\varepsilon = 0,1; 0,05$, o $n = 50$, tai $\varepsilon\sqrt{n} = 0,7071; 0,35355$. Gauname apytiksles tikimybių reikšmes $K(0,7071) = 0,3006$; $K(0,35355) = 0,0004$. Kai $\varepsilon = 0,1; 0,05$, o $n = 100$, tai $\varepsilon\sqrt{n} = 1; 0,5$. Gauname apytiksles tikimybių reikšmes $K(1) = 0,7300$; $K(0,5) = 0,0361$.

I.1.23. Tegu $p = \mathbf{P}\{X = 2\} = 3^2e^{-3}/2 \approx 0,224$, o $\hat{p}_n = S_n/n$, kai S_n yra imties elementų, įgijusių reikšmę 2, skaičius. Kadangi $S_n \sim B(n,p)$, tai remiantis CRT

$$\sqrt{n}(\hat{p}_n - p) \xrightarrow{d} Z \sim N(0, p(1-p)).$$

Gauname nelygybę n atžvilgiu

$$\mathbf{P}_p\{|\hat{p}_n - p| \le 0,05\} = \mathbf{P}_p\{-0,05\sqrt{n/(p(1-p))} \le$$

$$\le \sqrt{n}\frac{\hat{p}_n - p}{\sqrt{p(1-p)}} \le 0,05\sqrt{n/(p(1-p))}\}$$

$$\approx 2\Phi(0,05\sqrt{n/(p(1-p))}) - 1 \ge 0,95.$$

Iš čia

$$0,05\sqrt{p(1-p)/n} \ge z_{0,025} \Leftrightarrow n \ge z^2_{0,025}p(1-p)/0,05^2 = 267,095.$$

Taigi imties didumas $n \ge 268$.

I.1.24. a) Sugrupavę duomenis ilgio $h = 10$ intervalais gauname grupuotosios imties realizaciją

X'_i	5	15	25	35	45	55	65
n_i	1	6	26	37	24	5	1

Šioje lentelėje X'_i yra i-ojo intervalo vidurys, o n_i – stebinių, patekusių į i-ąjį intervalą, skaičius.

b) Empirinės pasiskirstymo realizacijos $F_n(x)$ grafikas gaunamas braižant funkciją, kuri kinta didumo n_i/n šuoliukais taškuose X'_i.

Histogramos $f_n(x)$ grafikas gaunamas braižant stačiakampius stulpelius, kurių aukštis i-ame intervale $n_i/(nh) = n_i/(10n)$.

Apie grupavimo intervalų ilgio parinkimą ir branduolinius tankio įvertinius žr. [2], 2.6.2 ir 2.6.3 skyrelius.

I.1.25. Nežinomų parametrų įvertinių realizacijos yra $\hat{\mu} = \bar{X} = 34,75$; $\hat{\sigma} = s = 10,56$.

Vizualus palyginimas su normaliojo skirstinio pasiskirstymo funkcija dažnai atliekamas braižant empirinės pasiskirstymo funkcijos grafike funkciją $\Phi((x - \bar{X})/s)$.

Analogiškai, vizualiai lyginant su normaliojo skirstinio tankiu dažnai histogramos grafike pavaizduojama funkcija $\varphi((x - \bar{X})/s)/s$.

Pateiktieji palyginimai nėra vaizdūs. Jeigu paprastoji imtis $X_1, X_2, ..., X_n$ gauta stebint normalųjį a. d. $X \sim N(\mu, \sigma^2)$, tai a. d. $Y_i = \Phi((X_i - \mu)/\sigma) \sim U(0,1)$ turi tolygųjį intervale $[0;1]$ skirstinį. Todėl taškai $(Y_{(i)}; i/n)$, $i = 1, 2, ..., n$, bus išsidėstę apie vienetinio kvadrato įstrižainę.

A. d. $\hat{Y}_i = \Phi(\hat{Z}_i)$, $\hat{Z}_i = (X_i - \bar{X})/s$ nėra tolygiai pasiskirstę intervale $[0;1]$; be to, jie yra priklausomi. Tačiau esant pakankamai dideliam imties didumui n, taškai $(\hat{Y}_{(i)}; i/n)$ irgi bus išsidėstę apie vienetinio kvadrato įstrižainę.

Šį principą galima panaudoti ir vizualiai lyginant histogramą su teoriniu tankiu. Sudalinkime intervalą $[0;1]$ į, pavyzdžiui, 10 vienodo ilgio intervalų. Sugrupuokime a. d. $\hat{Y}_1, \hat{Y}_2, ..., \hat{Y}_n$ į intervalus $[0;0,1]; (0,1;0,2]; ...; (0,9;1,0]$. Tarkime, stebėjimų, patekusių į šiuos intervalus, skaičiai yra $m_1, m_2, ..., m_{10}$. Pagal šiuos duomenis braižoma histograma, t. y. aukščio $m_i/(0,1n) = m_i/10$ stulpelius i-ame intervale. Šią histogramą vizualiai lyginame su skirstinio $U(0,1)$ tankiu $f(x) \equiv 1$, kai $x \in [0;1]$, ir $f(x) \equiv 0$, kai $x < 0$ arba $x > 1$.

Šis palyginimas turi tam tikrų privalumų lyginant su pradinių duomenų grupavimu į vienodo ilgio intervalus. Tada į kraštinius intervalus patenka mažiau stebinių ir palyginimas šiuose intervaluose yra mažiau tikslus negu centriniuose intervaluose. Be to, lyginimas su tiesine funkcija yra vaizdesnis negu su kreivalinijine funkcija $\varphi((x-\bar{X})/s)/s$.

Apie formalius suderinamumo kriterijus, kai tikrinama prielaida, kad turima paprastosios imties realizacija neprieštarauja prielaidai, kad stebimas a. d., turintis konkretų skirstinį ar priklausantis parametrinei skirstinių šeimai, žr. 3 skyriuje.

I.1.3 skyrelis

I.1.26. Statistikos $X_{(1)}$ tankio funkcija

$$f_1(x) = 2ne^{-2n(x-\mu)}, \quad x > \mu.$$

Randame

$$\mathbf{P}_\mu\{X_{(1)} - \mu \leq 0,05\} = \int_\mu^{\mu+0,05} f_1(x)dx = 1 - e^{-2n0,05} = 1 - e^{-5} = 0,9933.$$

I.1.27. a) Atlikime transformaciją

$$\begin{cases} Z_1 = X_{(1)} - \mu, \\ Z_2 = X_{(2)} - X_{(1)}, \\ \ldots\ldots\ldots\ldots\ldots\ldots, \\ Z_n = X_{(n)} - X_{(n-1)}. \end{cases}$$

Atvirkštinė transformacija

$$\begin{cases} X_{(1)} = Z_1 + \mu, \\ X_{(2)} = Z_2 + Z_1 + \mu, \\ \dots\dots\dots\dots\dots\dots, \\ X_{(n)} = Z_n + \dots + Z_1 + \mu. \end{cases}$$

Pakeitimo jakobianas J lygus 1. Variacinės eilutės $(X_{(1)}, ..., X_{(n)})^T$ tankio funkcija

$$f(x_1, ..., x_n) = \frac{n!}{\sigma^n} e^{-(x_1-\mu)/\sigma} \cdot ... \cdot e^{-(x_n-\mu)/\sigma}, \quad \mu < x_1 < ... < x_n < \infty.$$

Atlikę kintamųjų keitimą, gauname a. v. $(Z_1, ..., Z_n)^T$ tankio funkciją

$$g(z_1, ..., z_n) = (n/\sigma)e^{-(n/\sigma)z_1}((n-1)/\sigma)e^{-((n-1)/\sigma)z_2} \cdot ... \cdot (1/\sigma)e^{-(1/\sigma)z_n},$$

kai $0 < z_1, ..., z_n < \infty$. Matome, kad a. d. $Z_1, ..., Z_n$ nepriklausomi; $Z_i \sim \mathcal{E}((n-i+1)/\sigma), i = 1, ..., n$. Gauname $2nZ_1/\sigma = 2n(X_{(1)} - \mu)/\sigma \sim \chi^2(2)$, o suma

$$\sum_{i=2}^{n}(X_{(i)} - X_{(1)}) = (n-1)Z_2 + (n-2)Z_3 + ... + Z_n \sim G(1/\sigma, n-1),$$

nes a. d. $(n-i+1)Z_i \sim G(1/\sigma, 1)$. Tada $(2/\sigma)\sum_{i=2}^{n}(X_{(i)} - X_{(1)}) \sim \chi^2(2(n-1))$.

Gauname

$$T = \frac{n(X_{(1)} - \mu)}{W} = \frac{(2/\sigma)n(X_{(1)} - \mu)/2}{(2/\sigma)\sum_{i=2}^{n}(X_{(i)} - X_{(1)})/(2(n-1))} \sim F(2, 2(n-1)).$$

b) Remiantis p. a) a. d. $X_{(1)} - \mu \sim \mathcal{E}(n/\sigma)$, taigi $X_{(1)} \sim \mathcal{E}(n/\sigma, \mu)$.

c) Imdami p. a) $\mu = 0$ ir $\sigma = 1/\lambda$, gausime, kad a. d. $Y_1, ..., Y_n$ yra nepriklausomi ir $Y_i \sim \mathcal{E}(\lambda), i = 1, ..., n$.

I.1.28. A. v. $(Y_{(1)}, ..., Y_{(n)})^T = (F(X_{(1)}), ..., F(X_{(n)}))^T$ yra variacinė eilutė, gauta stebint a. d. $Y \sim U(0, 1)$; jo tankis lygus $n!$, kai $0 < y_1 < ... < y_n < 1$. Atlikime transformaciją

$$\begin{cases} Z_1 = Y_{(1)}/Y_{(2)}, \\ Z_2 = (Y_{(2)}/Y_{(3)})^2, \\ \dots\dots\dots\dots\dots\dots, \\ Z_{n-1} = (Y_{(n-1)}/Y_{(n)})^{n-1}, \\ Z_n = Y_{(n)}. \end{cases}$$

Atvirkštinė transformacija

$$\begin{cases} Y_{(1)} = Z_n Z_{n-1}^{1/(n-1)} \cdot ... \cdot Z_2^{1/2} Z_1 \\ Y_{(2)} = Z_n Z_{n-1}^{1/(n-1)} \cdot ... \cdot Z_2^{1/2} \\ \dots\dots\dots\dots\dots\dots, \\ Y_{(n-1)} = Z_n Z_{n-1}^{1/(n-1)}, \\ Y_{(n)} = Z_n. \end{cases}$$

Pakeitimo jakobianas $J = z_n^{n-1}/(n-1)!$. Atlikę kintamųjų keitimą gauname, kad a. v. $(Z_1, ..., Z_{n-1}, Z_n)^T$ tankio funkcija yra nz_n^{n-1}. Suintegravę pagal z_n, gauname a. v.

$(Z_1, ..., Z_{n-1})^T$ tankį, kuris lygus 1, kai $0 < z_1, ..., z_{n-1} < 1$. Taigi a. d. $Z_1, ..., Z_{n-1}$ yra nepriklausomi ir kiekvienas iš jų turi tolygųjį skirstinį $U(0, 1)$.

I.1.29. Atsitiktinio vektoriaus $(X_{(m)}, Y_{(n)})^T$ tankio funkcija $f(x, y) = mnx^{m-1}y^{n-1}$ vienetiniame kvadrate $0 < x, y < 1$. Randame

$$\mathbf{P}\{X_{(m)}/Y_{(n)} \leq z\} = \int_0^1 ny^{n-1} \int_0^{zy} mx^{m-1}dxdy = \frac{n}{m+n}z^m, \quad 0 < z < 1;$$

$$\mathbf{P}\{X_{(m)}/Y_{(n)} \leq z\} = 1 - \int_0^1 mx^{m-1} \int_0^{x/z} ny^{n-1}dydx = 1 - \frac{m}{m+n}\frac{1}{z^n}, \quad 1 < z < \infty.$$

Diferencijuodami randame tankio funkciją

$$g(z) = \frac{mn}{m+n}z^{m-1}, \quad \text{kai} \quad 0 < z < 1,$$

$$g(z) = \frac{mn}{m+n}\frac{1}{z^{n+1}}, \quad \text{kai} \quad 1 < z < \infty.$$

I.1.30. Empirinės medianos X_{k+1} skirstinys (žr. [2], 2.3 skyrelį) yra $Be(k+1,\ k+1)$. Tada (žr. 1 priedo 1.P.2 lentelę)

$$\boldsymbol{V}X_{(k+1)} = \frac{k+1}{2(k+1)}\frac{k+1}{2(k+1)}\frac{1}{2(k+1)+1} = \frac{1}{4(2k+3)}.$$

I.1.31. Tegu $1 \leq k_1 < k_2 \leq n$. A. d. $F(X_{(k_1)})$ ir $F(X_{(k_2)})$ yra pozicinės statistikos $Y_{(k_1)}$ ir $Y_{(k_2)}$, gautos stebint a. d. $Y \sim U(0, 1)$. Žinome, kad $Y_{(k_1)} \sim Be(k_1, n-k_1+1)$, $Y_{(k_2)} \sim Be(k_2, n-k_2+1)$. Taigi $\boldsymbol{V}(Y_{(k_1)}) = p_1(1-p_1)/(n+2)$, $\boldsymbol{V}(Y_{(k_2)}) = p_2(1-p_2)/(n+2)$; čia $p_1 = k_1/(n+1)$, $p_2 = k_2/(n+1)$.

Atsitiktinio vektoriaus $(Y_{(k_1)}, Y_{(k_2)})^T$ tankio funkcija (žr. [2], 2.3 skyrelį)

$$\frac{n!}{(k_1-1)!(k_2-k_1-1)!(n-k_2)!}y_1^{k_1-1}(y_2-y_1)^{k_2-k_1-1}(1-y_2)^{n-k_2}, \quad 0 < y_1 < y_2 < 1.$$

Turime

$$\mathbf{Cov}\,(Y_{(k_1)}, Y_{(k_2)}) = \mathbf{E}(Y_{(k_1)}Y_{(k_2)}) - \mathbf{E}Y_{(k_1)}\mathbf{E}Y_{(k_2)}, \quad \mathbf{E}Y_{(k_1)} = \frac{p_1}{n+2}, \mathbf{E}Y_{(k_2)} = \frac{p_2}{n+2}.$$

Randame

$$\mathbf{E}(Y_{(k_1)}Y_{(k_2)}) = \mathbf{E}(Y_{(k_1)} - Y_{(k_1)}(1-Y_{(k_2)})) = \frac{p_1(k_2+1)}{(n+2)},$$

$$\mathbf{Cov}\,(Y_{(k_1)}, Y_{(k_2)}) = p_1\frac{k_2+1}{n+2} - p_1\frac{k_2}{n+1} = p_1\frac{n-k_2+1}{(n+1)(n+2)} = \frac{p_1(1-p_2)}{n+2}.$$

I.1.32. Remiantis [2], 2.3 skyreliu, a. d.

$$\mathbf{E}(W_n) = n\int_{-\infty}^{\infty} x[(F(x))^n - (1-F(x))^n]f(x)dx =$$

$$-\int_{-\infty}^{\infty} xd[1-(F(x))^n-(1-F(x))^n] = \int_{-\infty}^{\infty}[1-(F(x))^n-(1-F(x))^n]dx.$$

I.1.33. Remiantis [2], 2.3 skyreliu, a. d. $X_{(1)}$ tankio funkcija

$$n(1-F(x))^{n-1}f(x) = n\lambda e^{-n\lambda x}; \quad \mathbf{E}_\lambda(X_{(1)}) = 1/(n\lambda).$$

A. d. $X_{(n)}$ tankio funkcija

$$n(F(x))^{n-1}f(x) = n\lambda(1-e^{-\lambda x})^{n-1}e^{-\lambda x} = n\lambda\sum_{j=0}^{n-1} C_{n-1}^j(-1)^j e^{-\lambda(j+1)x},$$

$$\mathbf{E}_\lambda(X_{(n)}) = n\sum_{j=0}^{n-1} C_{n-1}^j(-1)^j/(\lambda(j+1)^2).$$

I.1.34. Randame $-\ln X_{(n)}^{(j)} \sim \mathcal{E}(n),\ j=1,...,k.$ Tada

$$-\ln V = -\ln X_{(n)}^{(1)} - ... - \ln X_{(n)}^{(k)} \sim G(n,k).$$

I.1.35. Atsitiktinio vektoriaus $(X_{(1)},...,X_{(n)})^T$ tankis yra

$$n!(\lambda^\eta/(\Gamma(\eta))^n)x_1^{\eta-1}...x_n^{\eta-1}\exp\{-\lambda(x_1+...+x_n)\}$$

srityje $0<x_1<x_2<...<x_n<\infty$. Atliekame transformaciją

$$\begin{cases} s = x_1+...+x_n, \\ y_2 = \ln x_2 - \ln x_1, \\ \dots\dots\dots\dots, \\ y_n = \ln x_n - \ln x_1. \end{cases}$$

Atvirkštinė transformacija

$$\begin{cases} x_1 = s/(1+e^{y_2}+...+e^{y_n}) \\ x_2 = se^{y_2}/(1+e^{y_2}+...+e^{y_n}) \\ \dots\dots\dots\dots, \\ x_n = se^{y_n}/(1+e^{y_2}+...+e^{y_n}). \end{cases}$$

Pakeitimo jakobianas

$$J = \left[\frac{\mathcal{D}(s,y_2,...,y_n)}{\mathcal{D}(x_1,x_2,...,x_n)}\right]^{-1} = \frac{x_1x_2...x_n}{x_1+...+x_n} = \frac{s^{n-1}e^{1+y_2+...+y_n}}{(1+e^{y_2}+...+e^{y_n})^n}.$$

Įrašę į tankio formulę matome, kad jis lygus funkcijos, priklausančios tik nuo s, ir funkcijos, priklausančios nuo $y_2,...,y_n$, sandaugai. Taigi a. d. $X_1+...+X_n$ ir a. v. $(\ln X_{(2)} - \ln X_{(1)},...,\ln X_{(n)} - \ln X_{(1)})^T$ yra nepriklausomi.

I.1.36. Atsitiktinio vektoriaus $(X_{(1)},...,X_{(n)})^T$ tankio funkcija yra $f(x_1,...,x_n) = n!/(b-a)^n$ srityje $a<x_1<...<x_n<b$. Atliekame transformaciją

$$\begin{cases} y_1 = x_1 \\ y_2 = (x_2-x_1)/(x_n-x_1) \\ \dots\dots\dots\dots, \\ y_{n-1} = (x_{n-1}-x_1)/(x_n-x_1), \\ y_n = x_n. \end{cases}$$

Atvirkštinė transformacija

$$\begin{cases} x_1 = y_1 \\ x_2 = y_2(y_n - y_1) + y_1 \\ \dots\dots\dots\dots\dots\dots\dots, \\ x_{n-1} = y_{n-1}(y_n - y_1) + y_1, \\ x_n = y_n. \end{cases}$$

Pakeitimo jakobianas $J = (y_n - y_1)^{n-2}$.

Atlikę keitimą gauname a. v. $(Y_1, Y_n, Y_2, ..., Y_{n-1})^T$ tankio funkciją

$$\frac{n(n-1)(y_n - y_1)^{n-2}}{(b-a)^{n-2}}(n-2)!, \quad a < y_1 < y_n < b, \quad 0 < y_2, ..., y_{n-2} < 1.$$

Tankis yra sandauga a. v. $(Y_1, Y_n)^T$ tankio ir a. v. $(Y_2, ..., Y_{n-1})^T$ tankio $(n-2)!$. Tai tankis variacinės eilutės, gautos pagal paprastąją a. d. $Y \sim U(0,\ 1)$ imtį, kurios didumas $n - 2$.

I.1.37. A. v. $(X_1, ..., X_{k+1})^T$ tankio funkcija

$$f(x_1, ..., x_{k+1}) = \frac{\lambda^{\eta_1 + ... + \eta_{k+1}}}{\Gamma(\eta_1)\Gamma(\eta_2)...\Gamma(\eta_{k+1})} x_1^{\eta_1 - 1} \cdot ... \cdot x_{k+1}^{\eta_{k+1}-1} e^{-\lambda(x_1 + ... + x_{k+1})}.$$

Atliekame transformaciją

$$\begin{cases} z_1 = x_1/(x_1 + ... + x_{k+1}) \\ z_2 = x_2/(x_1 + ... + x_{k+1}) \\ \dots\dots\dots\dots\dots\dots\dots, \\ z_k = (x_k/(x_1 + ... + x_{k+1}), \\ s = x_1 + ... + x_{k+1}. \end{cases}$$

Atvirkštinė transformacija

$$\begin{cases} x_1 = sz_1 \\ x_2 = sz_2 \\ \dots\dots\dots\dots\dots\dots\dots, \\ x_k = sz_k, \\ x_{k+1} = s(1 - z_1 - ... - z_k). \end{cases}$$

Pakeitimo jakobianas $J = s^k$. Atlikę keitimą gauname tankį

$$\frac{\lambda^{\eta_1 + ... + \eta_{k+1}}}{\Gamma(\eta_1)\Gamma(\eta_2)...\Gamma(\eta_{k+1})}(sz_1)^{\eta_1 - 1} \cdot ... \cdot (sz_k)^{\eta_k - 1}(s(1 - z_1 - ... - z_k))^{\eta_{k+1}-1} e^{-\lambda s} s^k.$$

Suintegravę pagal s, gauname a. v. $(Z_1, ..., Z_k)^T$ tankį

$$g(z_1, ..., z_k) = \frac{\Gamma(\eta_1 + ... + \eta_{k+1})}{\Gamma(\eta_1)\Gamma(\eta_2)...\Gamma(\eta_{k+1})} z_1^{\eta_1 - 1} \cdot ... \cdot z_k^{\eta_k - 1}(1 - z_1 - ... - z_k)^{\eta_{k+1}-1},$$

kai $z_i > 0, z_1 + ... + z_k < 1$. O tai yra Dirichlė skirstinio $D(\eta_1, ..., \eta_k; \eta_{k+1})$ tankio funkcija. Kai $k = 1$, gauname beta skirstinį $Be(\eta_1, \eta_2)$.

I.1.38. Statistikos $X_{(n)}$ tankio funkcija yra $nF^{n-1}(x)f(x), 0 < x < \theta$; kai F ir f yra a. d. X pasiskirstymo funkcija ir tankis. Kadangi $F(x) = x/\theta$, $f(x) = 1/\theta$, tai gauname, kad $X_{(n)}$ tankis yra nx^{n-1}/θ^n, kai $0 < x < \theta$.

I.1.39. Tegu $Y_i = (X_i - \mu + \theta/2)/\theta$. Tada $Y_1, ..., Y_n$ yra paprastoji imtis a. d $Y \sim U(0,1)$. Imties plotis $X_{(n)} - X_{(1)} = (Y_{(n)} - Y_{(1)})\theta$. Rasime a. d. $Y_{(n)} - Y_{(1)}$ pasiskirstymo funkciją

$$G(z) = 1 - \mathbf{P}\{Y_{(n)} - Y_{(1)} > z\} = 1 - \int_0^{1-z} \int_{x+z}^1 n(n-1)(y-x)^{n-2} dy dx =$$

$$1 - \int_0^{1-z} n(n-1)x^{n-1} \int_{(x+z)/x}^{1/x} (t-1)^{n-2} dt dx = nz^{n-1} - (n-1)z^n.$$

Diferencijuodami gauname tankio funkciją

$$g(z) = G'(z) = n(n-1)z^{n-2}(1-z), \quad 0 < z < 1.$$

Taigi imties plotis $Y_{(n)} - Y_{(1)} \sim Be(n-1, 2)$. A. d. $X_{(n)} - X_{(1)}$ tankio funkcija yra $g(z/\theta)/\theta, 0 < z < \theta$.

I.1.40. Tegu $Y_i = (X_i - \mu + \theta/2)/\theta \sim U(0, 1)$, $i = 1, ..., n$. Tada $Z = ((Y_{(1)} + Y_{(n)})/2 - 1/2)/(Y_{(n)} - Y_{(1)})$, $0 < Y_{(1)} < Y_{(n)} < 1$. Atlikime transformaciją

$$U = Y_{(n)} - Y_{(1)}, \quad Z = ((Y_{(1)} + Y_{(n)}) - 1)/(2(Y_{(n)} - Y_{(1)})).$$

Atvirkštinė transformacija

$$Y_{(1)} = (U(2Z-1)+1)/2, \quad Y_{(n)} = (U(2Z+1)+1)/2.$$

Pakeitimo jakobianas $|J| = U$. Gauname a. v. $(U, Z)^T$ tankio funkciją $g(u, z) = n(n-1)u^{n-1}$. Reikia nustatyti argumentų u, z kitimo sritį. Gauname

$$0 < (u(2z-1)+1)/2 < (u(2z+1)+1)/2 < 1,$$

$$-1 < u(2z-1) < u(2z+1) < 1, \quad (1 - 1/u) < 2z < (1/u - 1).$$

Jeigu $z > 0$, tai $0 < u < 1/(1+2z)$; jeigu $z < 0$, tai $0 < u < 1/(1-2z)$. Taigi visais atvejais $0 < u < 1/(1+2|z|)$, $-\infty < z < \infty$.

A. d. Z tankio funkciją gauname integruodami $g(u, z)$ pagal u

$$f(z) = \int_0^{1/(1+2|z|))} n(n-1)u^{n-1} du = (n-1)(1/(1+2|z|))^n, \;\; -\infty < z < \infty.$$

I.1.41. Nagrinėkime logaritmą

$$-n \ln Y = Z = -\ln X_1 - ... - \ln X_n \sim G(1, n),$$

nes a. d. $-\ln X_i \sim \mathcal{E}(1)$. Pereidami prie a. d $Y = e^{-Z/n}$, gauname jo tankį

$$g(y) = \frac{n^n}{\Gamma(n)} y^{n-1} (-\ln y)^{n-1}, \quad 0 < y < 1.$$

I.1.42. Fiksuokime statistiką $X_{(k)} = y$. Tada antrosios imties elementų išsidėstymą galime interpretuoti kaip Bernulio eksperimentus: įgijimas reikšmės, ne didesnės už y su tikimybe y, ir reikšmės, didesnės už y su tikimybe $1-y$. Taigi a. d. Z sąlyginis skirstinys yra binominis $(Z|X_{(k)} = y) \sim B(n, y)$,

$$\mathbf{P}\{Z = z|X_{(k)} = y\} = C_n^z y^z (1-y)^{n-z}, \quad z = 0, 1, ..., n.$$

Norint gauti besąlyginį Z skirstinį, reikia suvidurkinti pagal a. d. $X_{(k)}$ skirstinį $X_{(k)} \sim Be(k, m-k+1)$.

$$\mathbf{P}\{Z = z\} = C_n^z \frac{m!}{(k-1)!(m-k)!} \int_0^1 y^{k+z-1}(1-y)^{m+n-k-z} dy =$$

$$\frac{C_n^z C_m^k}{C_{m+n}^{k+z}} \frac{k}{k+z}, \quad z = 0, 1, ..., n.$$

I.1.43. Binominio a. d. $X \sim B(n,p)$ r-asis faktorialinis momentas $\mathbf{E}X^{[r]} = \mathbf{E}(X(X-1)...(X-r+1)) = n^{[r]}p^r$, $r \leq n$. Taigi a. d. Z sąlyginio skirstinio r-asis faktorialinis momentas $\mathbf{E}(Z^{[r]}|X_{(k)}) = n^{[r]}X_{(k)}^r$. Suvidurkinę pagal $X_{(k)}$ skirstinį, gausime

$$\mathbf{E}Z^{[r]} = n^{[r]} \frac{m!}{(k-1)!(m-k)!} \int_0^1 y^{k+r-1}(1-y)^{m-k} dy = \frac{m!n!(k+r-1)!}{(k-1)!(m+r)!(n-r)!}.$$

I.1.44. Galima pasiūlyti keletą būdų įrodyti šį faktą.

1. Įrodyti, kad a. d. $\sqrt{n}(Y_{(k)} - p)/\sqrt{np(1-p)}$ tankio funkcija konverguoja į standartinio normaliojo skirstinio tankį $\varphi(x)$. Šis būdas panaudotas [2], 2.4.2 teoremoje.

2. Pasinaudosime sąryšiu $\mathbf{P}\{Y_{(k)} < z\} = \mathbf{P}\{Z_n \geq k\}$; čia $Z_n \sim B(n, z)$. Imdami $z = p + x/\sqrt{n}$, gauname $\mathbf{P}\{Y_{(k)} < z\} = \mathbf{P}\{\sqrt{n}(Y_{(k)} - p) < x\} = \mathbf{P}\{Z_n \geq k\}$, kai $Z_n \sim B(n, p + x/\sqrt{n})$. Taikydami CRT, gauname

$$\mathbf{P}\{Z_n \geq k\} \approx 1 - \Phi\left(\frac{k - np - x\sqrt{n}}{\sqrt{n(p + x/\sqrt{n})(1 - p - x/\sqrt{n})}}\right) =$$

$$1 - \Phi\left(\frac{(k-np)/\sqrt{np(1-p)} - x/\sqrt{p(1-p)}}{\sqrt{1 + x/\sqrt{np}}\sqrt{1 - x/\sqrt{n(1-p)}}}\right) \to \Phi(x/\sqrt{p(1-p)}),$$

nes $(k-np)/\sqrt{n} \to 0$, $x/\sqrt{n} \to 0$.

I.1.45. Kadangi $Y_i = F(X_i) \sim U(0, 1)$, tai a. d. $Y \sim U(0,1)$ eilės p kvantilio $y(p)$ empirinis analogas yra $\hat{y}(p) = Y_{(k)}$, kai $k = [np]$, jei np sveikasis skaičius, ir $k = [np] + 1$ kitu atveju; bet kuriuo atveju $k/n \to 0$, kai $n \to \infty$. Remiantis **I.1.44** pratimu

$$\sqrt{n}(\hat{y}(p) - p) \xrightarrow{d} Z \sim N(0, p(1-p)).$$

Kadangi $\hat{x}(p) = F^{-1}(\hat{y}(p))$, $x(p) = F^{-1}(y(p))$, tai remiantis delta metodu

$$\sqrt{n}(\hat{x}(p) - x(p)) \xrightarrow{d} U \sim N(0, \sigma^2(p)),$$

čia

$$\sigma^2(p) = [(F^{-1})'(p)]^2 p(1-p) = p(1-p)/f^2(x(p)).$$

I.1.46. Tvirtinimas yra išvada iš a. v. $(\hat{x}(p_1), \hat{x}(p_2))^T$ asimptotikos, pateiktos [2], 2.4.3 teoremoje.

I.1.47. Pakanka **I.1.46** pratime įrašyti $p_1 = 1/4, p_2 = 3/4$; jeigu tankis simetriškas, tai $f(x(1/4)) = f(x(3/4))$.

I.1.48. Kadangi $f^2(\mu|\mu,\sigma) = 1/(\pi^2\sigma^2)$, tai a. d. $\hat{x}(1/2)$ asimptotinė dispersija yra $\pi^2\sigma^2/(4n)$; $f^2(\mu+\sigma|\mu,\sigma) = 1/(4\pi^2\sigma^2)$, tai a. d. $\hat{\sigma}$ asimptotinė dispersija yra $\pi^2\sigma^2/(4n)$.

I.1.49. a) Prilyginę pasiskirstymo funkciją tikimybei p

$$F(x) = \frac{1}{\pi}\operatorname{arctg}\frac{x-\mu}{\sigma} + \frac{1}{2} = p$$

gauname p kvantilį

$$x(p) = \mu + \sigma\operatorname{tg}(\pi(p-1/2)) = \mu - \sigma\operatorname{tg}(\pi(1/2-p)),$$

$$x(1-p) - x(p) = 2\sigma\operatorname{tg}(\pi(1/2-p)), \quad c(p) = 2\operatorname{tg}(\pi(1/2-p)).$$

Remiantis **I.1.46** pratimu statistikos $\hat{\sigma}$ asimptotinė dispersija

$$\sigma^2(p) = \frac{1}{c^2(p)}\frac{2p(1-2p)}{f^2(x(p))}\frac{\sigma^2}{n} = \frac{\pi^2 2p(1-2p)}{\sin^2(\pi(1-2p))}\frac{\sigma^2}{n}.$$

Pažymėję $z = \pi(1-2p)$, gausime

$$\sigma^2(p) = \frac{z(\pi-z)}{\sin^2 z}\frac{\sigma^2}{n} = h(z)\frac{\sigma^2}{n}.$$

Kai $z = \pi/2$ (t. y. $p = 1/4$), tai $h'(z)$ lygi nuliui ir keičia ženklą iš neigiamo į teigiamą. Asimptotinė dispersija minimali, kai $p = 1/4$; $\sigma^2(1/4) = \pi^2\sigma^2/(4n)$.

I.1.50. Kadangi $f^2(\mu|\mu,\sigma) = 1/(2\pi\sigma^2)$, tai a. d. $\hat{x}(1/2)$ asimptotinė dispersija yra $\pi\sigma^2/(2n)$; $1/f^2(\mu+z_{1/4}\sigma|\mu,\sigma) = 2\pi e^{z_{1/4}^2}\sigma^2$, tai, remiantis **I.1.44** pratimu, a. d. $\hat{\sigma}$ asimptotinė dispersija yra $\pi e^{z_{1/4}^2}\sigma^2/(8nz_{1/4}^2) \approx 1,3605\sigma^2/n$.

I.1.51. a) Kadangi $x(p) = \mu - \sigma z_p$, $x(1-p) = \mu + \sigma z_p$, tai $c(p) = 2z_p$. b) Remiantis **I.1.46** pratimu statistikos $\hat{\sigma}$ asimptotinė dispersija

$$\boldsymbol{V}(\hat{\sigma}) = \pi p(1-2p)e^{z_p^2}\sigma^2/(nz_p^2).$$

Skaitiniais metodais minimizuodami šią funkciją p atžvilgiu, gauname, kad minimumas pasiekiamas, kai $p = 0,0692$. Tada asimptotinė dispersija yra $0,7666\sigma^2/n$. Vietoje kvartilių imdami kvantilius $x(0,0692)$ ir $x(0,9308)$ asimptotinę dispersiją sumažiname nuo $1,3605\sigma^2/n$ iki $0,7666\sigma^2/n$.

Pažymėsime, kad a. d. $s = \sqrt{s^2}$, kai s^2 yra nepaslinktoji empirinė dispersija, asimptotinė dispersija (žr. **I.1.95** pratimą) yra $0,5\sigma^2/n$.

I.1.52. Kadangi $f^2(\mu|\mu,\sigma) = 1/(4\sigma^2)$, tai a. d. $\hat{x}(1/2)$ asimptotinė dispersija yra σ^2/n; $1/f^2(\mu+\sigma\ln 2|\mu,\sigma) = 16\sigma^2$, tai, remiantis **I.1.46** pratimu, a. d. $\hat{\sigma}$ asimptotinė dispersija yra $\sigma^2/(n\ln^2 2) \approx 2,0814\sigma^2/n$.

I.1.53. a) Randame $x(p) = \mu + \sigma \ln(2p)$, $x(1-p) = \mu - \sigma \ln(2p)$ ir $c(p) = -2\ln(2p)$. b) Remiantis **I.1.46** pratimu statistikos $\hat{\sigma}$ asimptotinė dispersija

$$\sigma^2(p) = \frac{1-2p}{2p\ln^2(2p)} \frac{\sigma^2}{n}.$$

Skaitiniais metodais minimizuodami šią funkciją p atžvilgiu, gauname, kad minimumas pasiekiamas, kai $p \approx 0,1015$. Tada asimptotinė dispersija yra $1,5441\sigma^2/n$. Vietoje kvartilių imdami kvantilius $x(0,1015)$ ir $x(0,8985)$ asimptotinę dispersiją sumažiname nuo $2,0814\sigma^2/n$ iki $1,5441\sigma^2/n$.

I.1.54. Statistika $F(\hat{x})$ yra empirinė mediana, gauta stebint a. d. $Y \sim U(0,\,1)$. Tvirtinimas išplaukia iš **I.1.44** pratimo.

I.1.55. A. d. $X \sim \mathcal{E}(\lambda)$ mediana $x(1/2) = \ln 2/\lambda$. Remiantis **I.1.45** pratimu

$$\lambda\sqrt{2n+1}(\tilde{x} - \ln 2/\lambda) \xrightarrow{d} Z \sim N(0,\,1).$$

I.1.56. A. d. $S_k \sim Be(k, n-k+1)$. Remiantis sąryšiu su gama skirstiniu (žr. 1 priedo 1.P.3 lentelę) $S_k = Y/(Y+Z)$; čia Y ir Z nepriklausomi ir $Y \sim G(1,k)$, $Z \sim G(1, n-k+1)$. Jeigu k fiksuotas, o $n \to \infty$, tai

$$nS_k = Y/((Y+Z)/n) \xrightarrow{d} Y \sim G(1,k),$$

nes $Y + Z \sim G(1, n+1)$ ir remiantis didžiuoju skaičių dėsniu

$$\frac{Y+Z}{n} = \frac{Y+Z}{n+1}\frac{n+1}{n} \xrightarrow{P} 1, \quad n \to \infty.$$

I.1.57. Pagal **I.1.56** pratimą a. d. $U_k = nF(X_{(k)}) \xrightarrow{d} Y \sim G(1,k)$. Kadangi $F(x) = (x-a)/(b-a), a < x < b$, tai a. d. $X_{(k)} = a + (b-a)U_k/n$. Taigi a. d. $X_{(k)} - a$ skirstinys aproksimuojamas gama skirstiniu $G(n/(b-a), k)$.

A. d. $V_k = n(1 - F(X_{(n-k+1)})) \xrightarrow{d} Y \sim G(1,k)$, o $X_{(n-k+1)} = b - (b-a)V_k/n$ yra a. d. V_k tiesinė funkcija.

I.1.58. Eksponentinio skirstinio $\mathcal{E}(\lambda)$ pasiskirstymo funkcija $F(x) = 1 - e^{-\lambda x}$. Iš lygybės $U_k = nF(X_{(k)}) = n(1 - e^{-\lambda X_{(k)}})$ gauname $X_{(k)} = -(1/\lambda)\ln(1 - U_k/n) \approx U_k/(n\lambda)$. Kadangi U_k skirstinys aproksimuojamas skirstiniu $G(1,k)$, tai a. d. $X_{(k)}$ skirstinį galima aproksimuoti skirstiniu $G(n\lambda, k)$.

Iš lygybės $V_k = n(1 - F(X_{(n-k+1)}))$ gauname $X_{(n-k+1)} = -(1/\lambda)(\ln V_k - \ln n)$. A. d. V_k skirstinys aproksimuojamas gama skirstiniu $G(1,k)$. Tada $-\ln V_k$ aproksimuojamas skirstiniu, kurio tankis

$$g(y) = \frac{1}{(k-1)!} e^{-ky} e^{-\lambda e^{-y}}.$$

Matome, kad a. d. $X_{(n-k+1)}$ asimptotinis skirstinys yra dvigubo eksponentinio (ekstremalių reikšmių) tipo.

I.1.4 skyrelis

I.1.59. Gauname $\bar{X} = \frac{1}{n}\sum_{i=1}^{k} n_i\bar{X}_i$, $n = n_1 + ... + n_k$;

$$s^2 = \{\sum_{i=1}^{k}(n_i - 1)s_i^2 + \sum_{i=1}^{k} n_i\bar{X}_i^2 - n\bar{X}^2\}/(n-1).$$

I.1.60. Gauname I tipo kino juostoms $\bar{X}^{(1)} = 22,333, (s^{(1)})^2 = 11,278$; II tipo kino juostoms $\bar{X}^{(2)} = 18,282, (s^{(2)})^2 = 12,368$.

I.1.61. Gauname $\bar{X}_{n+1} = (n\bar{X}_n + X_{n+1})/(n+1)$;

$$s_{n+1}^2 = [(n-1)s_n^2 + n\bar{X}_n^2 - (n+1)\bar{X}_{n+1}^2]/n.$$

I.1.62. Remsimės charakteristinių funkcijų išraiškomis iš 1 priedo 1.P.2-1.P.3 lentelių.

a) a. d. $X \sim N(\mu, \sigma^2)$ charakteristinė funkcija $\varphi_X(t) = e^{i\mu t - t^2\sigma^2/2}$; tada a. d. $\bar{X}$ charakteristinė funkcija $\varphi_{\bar{X}}(t) = e^{i\mu t - t^2\sigma^2/(2n)}$, t. y. $\bar{X} \sim N(\mu, \sigma^2/n)$;

b) a. d. $X \sim K(\mu, \sigma)$ charakteristinė funkcija $\varphi_X(t) = e^{i\mu t - |t|\sigma}$; tada a. d. $\bar{X}$ charakteristinė funkcija $\varphi_{\bar{X}}(t) = e^{i\mu t - |t|\sigma}$, t. y. $\bar{X} \sim K(\mu, \sigma)$;

c) a. d. $X \sim G(\lambda, \eta)$ charakteristinė funkcija $\varphi_X(t) = (\lambda/(\lambda - it))^\eta$; tada a. d. $n\bar{X}$ charakteristinė funkcija $\varphi_{n\bar{X}}(t) = (\lambda/(\lambda - it))^{n\eta}$, t. y. $n\bar{X} \sim G(\lambda, n\eta)$;

d) a. d. $X \sim B(k, p)$ charakteristinė funkcija $\varphi_X(t) = (q + pe^{it})^k$; tada a. d. $n\bar{X}$ charakteristinė funkcija $\varphi_{n\bar{X}}(t) = (q + pe^{it})^{kn}$, t. y. $n\bar{X} \sim B(nk, p)$;

e) a. d. $X \sim \mathcal{P}(\lambda)$ charakteristinė funkcija $\varphi_X(t) = e^{\lambda(e^{it}-1)}$; tada a. d. $n\bar{X}$ charakteristinė funkcija $\varphi_{n\bar{X}}(t) = e^{n\lambda(e^{it}-1)}$, t. y. $n\bar{X} \sim \mathcal{P}(n\lambda)$.

I.1.63. A. d. $\bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \sigma^2((1/m) + (1/n)))$ arba $Z = (\bar{X} - \bar{Y} - (\mu_1 - \mu_2))/(\sigma^2((1/m) + (1/n))) \sim N(0, 1)$. Toliau $s_1^2(m-1)/\sigma^2 \sim \chi^2(m-1)$, $s_2^2(n-1)/\sigma^2 \sim \chi^2(n-1)$, tada $V = [s_1^2(m-1) + s_2^2(n-1)]/\sigma^2 \sim \chi^2(m+n-2)$. Pagal Stjudento skirstinio apibrėžimą

$$t = \frac{Z}{\sqrt{V/(m+n-2)}} \sim S(m+n-2).$$

I.1.64. Tegu i-osios imties elementai yra $X_{i1}, ..., X_{in_i}, i = 1, ..., k$. Nagrinėkime dėstinį

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \mu)^2/\sigma^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i + \bar{X}_i - \bar{X} + \bar{X} - \mu)^2/\sigma^2 =$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2/\sigma^2 + \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2/\sigma^2 + n(\bar{X} - \mu)^2/\sigma^2.$$

Kairėje lygybės pusėje turime kvadratinę formą, kuri turi χ^2 skirstinį su n laisvės laipsnių. Dešinėje pusėje pirmasis dėmuo yra $U = \sum_{i=1}^{k}(n_i - 1)s_i^2/\sigma^2 \sim \chi^2(n-k)$. Kadangi $\bar{X} \sim N(\mu, \sigma^2/n)$, tai paskutinis dėmuo $W \sim \chi^2(1)$. Remiantis Fišerio ir Kočreno teorema (žr. [15], 3b.4 skyrelį) vidurinysis dešinės pusės narys $V \sim \chi^2(k-1)$ ir U, V, W yra nepriklausomi.

I.1.65. Išplaukia iš Fišerio skirstinio apibrėžimo.

I.1.66. Tegu $Z_i = X_{1i} - X_{2i}, \mathbf{E}Z_i = \mu_1 - \mu_2, \boldsymbol{V}Z_i = \sigma^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$. A. d. $Z_1, ..., Z_n$ yra paprastoji imtis, gauta stebint a. d. $Z \sim N(\mu_1 - \mu_2, \sigma^2)$. Remiantis [2], 2.5.1 teorema

$$\sqrt{n}(\bar{Z} - (\mu_1 - \mu_2))/s_Z \sim S(n-1), \quad \bar{Z} = \sum_{i=1}^{n} Z_i/n, \quad s_Z^2 = \sum_{i=1}^{n}(Z_i - \bar{Z})^2/(n-1).$$

Tačiau

$$s_Z^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_{1i} - X_{2i} - (\bar{X}_1 - \bar{X}_2))^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_{1i} - \bar{X}_1)^2 + \frac{1}{n-1}\sum_{i=1}^{n}(X_{2i} - \bar{X}_2)^2$$

$$-2\frac{1}{n-1}\sum_{i=1}^{n}(X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2) = s_1^2 + s_2^2 - 2rs_1s_2.$$

I.1.67. Randame

$$\bar{Z} - \bar{X} = \frac{m}{m+n}(\bar{Y} - \bar{X}), \quad \mathbf{E}(\bar{Z} - \bar{X}) = \frac{m}{m+n}(\mathbf{E}Y_i - \mathbf{E}X_i);$$

$$\boldsymbol{V}(\bar{Z} - \bar{X}) = \frac{m^2}{(m+n)^2}(\frac{\sigma^2}{m} + \frac{\sigma^2}{n}) = \frac{m\sigma^2}{n(m+n)}.$$

I.1.68. Randame $\boldsymbol{V}X = \sigma^2$, $F(X) \sim U(0, 1), \boldsymbol{V}(F(X)) = 1/12$,

$$\mathbf{Cov}\,(X, F(X)) = \mathbf{Cov}\,(X - \mu, F(X)) = \mathbf{E}((X - \mu)F(X)) =$$

$$\int_{-\infty}^{\infty} \frac{x-\mu}{\sigma}\Phi(\frac{x-\mu}{\sigma})\varphi(\frac{x-\mu}{\sigma})dx = \frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty} t\Phi(t)e^{-t^2/2}dt =$$

$$-\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\Phi(t)de^{-t^2/2} = \frac{\sigma}{2\pi}\int_{-\infty}^{\infty} e^{-t^2}dt = \frac{\sigma}{2\sqrt{\pi}}.$$

Gauname $\rho(X, F(X)) = \mathbf{Cov}\,(X, F(X))/\sqrt{\boldsymbol{V}X\boldsymbol{V}(F(X))} = \sqrt{3/\pi}$.

I.1.69. Imdami $SS(\alpha, \beta)$ išvestines pagal α ir β ir prilyginę išvestines nuliui, gauname lygčių sistemą

$$2\sum_i(Y_i - \alpha - \beta(x_i - \bar{x})) = 0,$$

$$2\sum_i(Y_i - \alpha - \beta(x_i - \bar{x}))(x_i - \bar{x}) = 0.$$

Išsprendę gauname $\hat{\alpha} = \bar{Y}$, $\hat{\beta} = \sum_i Y_i(x_i - \bar{x})/\sum_i(x_i - \bar{x})^2$. A. d. $\hat{\alpha}$ ir $\hat{\beta}$ turi normaliuosius skirstinius, nes yra normaliųjų a. d. Y_i tiesiniai dariniai. Randame $\boldsymbol{V}(\hat{\alpha}) = \sigma^2/n$, $\boldsymbol{V}(\hat{\beta}) = \sigma^2/\sum_i(x_i - \bar{x})^2$ ir $\mathbf{Cov}\,(\hat{\alpha}, \hat{\beta}) = 0$. Taigi a. d. $\hat{\alpha}$ ir $\hat{\beta}$ yra nepriklausomi. Nesunku patikrinti, kad $\mathbf{Cov}\,(\hat{\alpha}, Y_i - \hat{\alpha} - \hat{\beta}(x_i - \bar{x})) = 0$, $\mathbf{Cov}\,(\hat{\beta}, Y_i - \hat{\alpha} - \hat{\beta}(x_i - \bar{x})) = 0$. Todėl SS_E nepriklauso nuo $\hat{\alpha}$ ir $\hat{\beta}$.

Nagrinėkime kvadratinę formą

$$\sum_i (Y_i - \alpha - \beta(x_i - \bar{x}))^2/\sigma^2 = \sum_i [(Y_i - \hat{\alpha} - \hat{\beta}(x_i - \bar{x})) + (\hat{\alpha} - \alpha) + (\hat{\beta} - \beta)(x_i - \bar{x})]^2/\sigma^2 =$$

$$SS_E/\sigma^2 + n(\hat{\alpha} - \alpha)^2/\sigma^2 + (\hat{\beta} - \beta)^2 \sum_i (x_i - \bar{x})^2/\sigma^2,$$

nes visos dvigubos sandaugos lygios nuliui. Kairėje lygybės pusėje kvadratinė forma turi χ^2 skirstinį su n laisvės laipsnių. Dešinėje pusėje du paskutiniai nariai turi χ^2 skirstinius su 1 laisvės laipsniu. Remiantis Fišerio ir Kočreno teorema (žr. [15], 3b.4 skyrelį), SS_E/σ^2 turi χ^2 skirstinį su $n-2$ laisvės laipsniais.

I.1.70. A. d. Y_j turi normalųjį skirstinį, nes yra normaliųjų a. d. tiesinis darinys. Tiesiogiai patikriname, kad $\boldsymbol{V} Y_j = \sigma^2$, $\mathbf{Cov}\,(Y_j, Y_{j'}) = 0, j \neq j'$. Taigi a. d. $Y_1, ..., Y_{n-1}$ yra nepriklausomi ir vienodai pasiskirstę pagal $N(0, \sigma^2)$.

I.1.71. $\mathbf{E}d = \sum_{i=1}^n \mathbf{E}|X_i - \mu|/n = \mathbf{E}|X_i - \mu|, \boldsymbol{V}d = \boldsymbol{V}|X_i - \mu|/n$. Randame

$$\mathbf{E}|X_i - \mu| = \frac{2}{\sqrt{2\pi}\sigma} \int_0^\infty (x - \mu) e^{-(x-\mu)^2/(2\sigma^2)} dx = \frac{2\sigma}{\sqrt{2\pi}} \int_0^\infty t e^{-t^2/2} dt = \sqrt{\frac{2}{\pi}}\sigma;$$

$$\boldsymbol{V}d = \boldsymbol{V}|X_i - \mu|/n = [\mathbf{E}(X_i - \mu)^2 - (\mathbf{E}|X_i - \mu|)^2]/n = \sigma^2(1 - 2/\pi)/n.$$

I.1.72. Randame

$$\mathbf{E}\bar{X} = \frac{1}{n} \sum_i \mathbf{E}X_i = \mu; \quad \boldsymbol{V}\bar{X} = \frac{1}{n^2} \sum_i \boldsymbol{V}X_i = \frac{\sigma^2}{n}.$$

Pažymėkime $Y_i = X_i - \mu$, $\mathbf{E}Y_i = 0, \mathbf{E}Y_i^2 = \sigma^2, \mathbf{E}Y_i^3 = \mu_3, \mathbf{E}Y_i^4 = \mu_4$. Tada

$$\mu_3(\bar{X}) = \mathbf{E}(\bar{X} - \mu)^3 = \frac{1}{n^3}\mathbf{E}(\sum_i Y_i)^3 = \frac{1}{n^3}\mathbf{E}(\sum_i Y_i^3) = \frac{\mu_3}{n^2},$$

nes kitose sumose bus dėmenys pavidalo $Y_i^2 Y_j, i \neq j$, arba pavidalo $Y_i Y_j Y_l, i \neq j \neq l$.

Kadangi $Y_1, ..., Y_n$ nepriklausomi ir $\mathbf{E}Y_i = 0$, tai tokių sumų vidurkiai lygūs nuliui.

$$\mu_4(\bar{X}) = \frac{1}{n^4}\mathbf{E}(\sum_i Y_i)^4 = \frac{1}{n^4}\mathbf{E}(\sum_i Y_i^4 + 3\sum_{i \neq j} Y_i^2 Y_j^2) = \frac{\mu_4 + 3(n-1)\sigma^4}{n^3}.$$

Randame

$$\gamma_1(\bar{X}) = \frac{\mu_3(\bar{X})}{\boldsymbol{V}(\bar{X})^{3/2}} = \frac{\mu_3}{\sigma^3\sqrt{n}} = \frac{\gamma_1}{\sqrt{n}};$$

$$\gamma_2(\bar{X}) = \frac{\mu_4(\bar{X})}{\boldsymbol{V}(\bar{X})^2} - 3 = \frac{1}{n}(\frac{\mu_4}{\sigma^4} - 3) = \frac{\gamma_2}{n}.$$

Aritmetinio vidurkio asimetrija $\sqrt{n}$, o ekscesas n kartų mažesni už stebėto a. d. asimetriją ir ekscesą.

I.1.73. Pirmasis tvirtinimas yra CRT atskiras atvejis. Antrasis tvirtinimas išplaukia iš to, kad $\sqrt{m_2} \xrightarrow{P} \sigma$.

I.1.74. Tai atskiras atvejis CRT, nes $\bar{X}$ yra suma vienodai pasiskirsčiusių a. d., $\mathbf{E}_\lambda(\bar{X}) = \lambda$, $\boldsymbol{V}_\lambda(\bar{X}) = \lambda/n$.

I.1.75. Kadangi $\sqrt{n}(\bar{X}-\lambda) \xrightarrow{d} Y \sim N(0,\lambda)$, tai remiantis delta metodu (žr. [2], 2.4 skyrelį)

$$\sqrt{n}(g(\bar{X}) - g(\lambda)) \xrightarrow{d} Yg'(\lambda) \sim N(0, \lambda[g'(\lambda)]^2).$$

Kadangi $g(\lambda) = \sqrt{\lambda}$, $[g'(\lambda)]^2 = 1/(4\lambda)$, tai

$$2\sqrt{n}(\sqrt{\bar{X}} - \sqrt{\lambda}) \xrightarrow{d} Z \sim N(0,\ 1).$$

I.1.76. Kadangi $\sqrt{n}(\bar{X} + 1/(2n) - \lambda) \xrightarrow{d} Y \sim N(0,\lambda)$, tai remiantis delta metodu

$$\sqrt{n}(g(\bar{X}) - g(\lambda)) \xrightarrow{d} Yg'(\lambda) \sim N(0, \lambda[g'(\lambda)]^2).$$

Kadangi $g(\lambda) = \lambda^{2/3}$, $[g'(\lambda)]^2 = 4/(9\lambda^{2/3})$, tai

$$\sqrt{n}\frac{(\bar{X} + 1/(2n))^{2/3} - \lambda^{2/3}}{2\lambda^{1/6}/3} \xrightarrow{d} Z \sim N(0,\ 1).$$

I.1.77. Pirmasis tvirtinimas yra CRT atskiras atvejis. Antrasis tvirtinimas išplaukia iš to, kad $\bar{X} \xrightarrow{P} p$.

I.1.78. Kadangi $\sqrt{n}(\bar{X} - p) \xrightarrow{d} Y \sim N(0, p(1-p))$, tai remiantis delta metodu (žr. [2], 2.4 skyrelį)

$$\sqrt{n}(g(\bar{X}) - g(p)) \xrightarrow{d} Yg'(p) \sim N(0, p(1-p)[g'(p)]^2).$$

Kadangi $g(p) = \arcsin(2p-1)$, $[g'(p)]^2 = 1/(p(1-p)$, tai

$$\sqrt{n}(\arcsin(2\bar{X} - 1) - \arcsin(2p-1)) \xrightarrow{d} Z \sim N(0,\ 1).$$

I.1.79. A. d. $X \sim U(0,\theta)$ vidurkis ir dispersija yra $\mathbf{E}_\theta X = \theta/2$, $\boldsymbol{V}_\theta X = \theta^2/12$. Tada $\mathbf{E}_\theta(2\bar{X}) = \theta$, $\boldsymbol{V}_\theta(2\bar{X}) = \theta^2/(3n)$ ir pritaikome CRT.

I.1.80. A. d. $X_{(n)}$ tankio funkcija $f_n(x) = nx^{n-1}/\theta^n, 0 < x < \theta$. Randame

$$\mathbf{E}_\theta X_{(n)} = \frac{n}{n+1}\theta, \quad \mathbf{E}_\theta(X_{(n)}^2) = \frac{n\theta^2}{n+2}, \quad \boldsymbol{V}_\theta(X_{(n)}) = \frac{n\theta^2}{(n+1)^2(n+2)}.$$

Tada

$$\mathbf{E}_\theta(\frac{n+1}{n}X_{(n)}) = \theta, \quad \boldsymbol{V}_\theta(\frac{n+1}{n}X_{(n)}) = \frac{\theta^2}{n(n+2)} = O(\frac{1}{n^2}).$$

I.1.81. Remiantis **I.1.79** pratimu $\sqrt{n}(2\bar{X} - \theta) \xrightarrow{d} Y \sim N(0, \theta^2/3)$; $g(\theta) = \ln\theta$, $[g'(\theta)]^2 = 1/\theta^2$. Remiantis delta metodu

$$\sqrt{3n}(\ln(2\bar{X}) - \ln\theta) \xrightarrow{d} X \sim N(0,\ 1).$$

I.1.82. Kadangi stebimo a. d. pirmieji momentai yra $\mathbf{E}_p X = 1/p$, $\boldsymbol{V}_p X = q/p^2$, tai tvirtinimas yra CRT atskiras atvejis.

I.1.83. Pagal **I.1.82** pratimą $\sqrt{n}(\bar{X} - 1/p) \xrightarrow{d} Y \sim N(0, q/p^2)$, tai remiantis delta metodu (žr. [2], 2.4 skyrelį)

$$\sqrt{n}(g(\bar{X}) - g(1/p)) \xrightarrow{d} Yg'(1/p) \sim N(0, (q/p^2)[g'(1/p)]^2).$$

Kadangi $g(1/p) = \ln[(1+\sqrt{q})/p - 1/2]$, $[g'(1/p)]^2 = p^2/q$, tai įrašę gauname suformuluotą tvirtinimą.

I.1.84. Centriniai empiriniai momentai nekinta keičiant imties elementus X_i į $Y_i = X_i - \mu$. Tada $\mathbf{E}Y_i = 0, \boldsymbol{V}Y_i = \boldsymbol{V}X_i = \sigma^2$

$$m_2 = \frac{1}{n}\sum_{i=1}^{n}(X_1 - \bar{X})^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_1 - \bar{Y})^2 = \frac{1}{n}\sum_{i=1}^{n}Y_i^2 - \bar{Y}^2.$$

Randame

$$\mathbf{E}m_2 = \frac{1}{n}\sum_{i=1}^{n}\mathbf{E}Y_i^2 - \mathbf{E}\bar{Y}^2 = \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n}\sigma^2 \neq \sigma^2.$$

Nepaslinktoji empirinė dispersija

$$s^2 = \frac{n}{n-1}m_2 = \frac{n}{n-1}\frac{1}{n}\sum_{i=1}^{n}(X_1 - \bar{X})^2, \quad \mathbf{E}s^2 = \sigma^2.$$

I.1.85. Turime $\boldsymbol{V}m_2 = \mathbf{E}(m_2)^2 - (\mathbf{E}m_2)^2$, $\mathbf{E}m_2 = (n-1)\sigma^2/n$.

$$\mathbf{E}m_2^2 = \mathbf{E}(a_2 - \bar{Y}^2)^2 = \mathbf{E}(a_2^2 - 2a_2\bar{Y}^2 + \bar{Y}^4).$$

$$\mathbf{E}a_2^2 = \frac{1}{n^2}\mathbf{E}(\sum_j Y_j^2)^2 = \frac{\mu_4 + (n-1)\sigma^4}{n},$$

$$\mathbf{E}(a_2\bar{Y}^2) = \frac{1}{n^3}\mathbf{E}(\sum_j Y_j^2(\sum_i Y_i^2 + \sum_{i\neq k} Y_iY_k)) = \frac{\mu_4 + (n-1)\sigma^4}{n^2},$$

$$\mathbf{E}\bar{Y}^4 = \frac{1}{n^4}\mathbf{E}(\sum_j Y_j^2 + \sum_{i\neq k} Y_iY_k)^2 = \frac{1}{n^4}\mathbf{E}(\sum_j Y_j^4 + \sum_{j\neq l} Y_j^2Y_l^2$$

$$+2\sum_j Y_j^2 \sum_{i\neq k} Y_iY_k + \sum_{i\neq k} Y_iY_k \sum_{i'\neq k'} Y_{i'}Y_{k'}) = \frac{\mu_4 + 3(n-1)\sigma^4}{n^3}.$$

Gauname

$$\mathbf{E}m_2^2 = \frac{(n-1)^2}{n^3}\mu_4 + \frac{(n-1)(n^2-2n+3)}{n^3}\sigma^4.$$

Atėmę $(\mathbf{E}m_2)^2$ ir išrikiavę narius n laipsniais, gauname

$$\boldsymbol{V}m_2 = \frac{\mu_4 - \sigma^4}{n} - \frac{2\mu_4 - 2\sigma^4}{n^2} + \frac{\mu_4 - 3\sigma^4}{n^3} = \frac{\mu_4 - \sigma^4}{n} + O(\frac{1}{n^2}).$$

I.1.86. Statistika na_k yra suma nepriklausomų vienodai pasiskirsčiusių a. d. $X_1^k, ...,$ X_n^k su vidurkiais $\mathbf{E}X_i^k = \alpha_k$ ir dispersijomis $\boldsymbol{V}X_i^k = \mathbf{E}X_i^{2k} - (\mathbf{E}X_i^k)^2 = \alpha_{2k} - \alpha_k^2$. Pakanka pasinaudoti sustiprintu DSD ir CRT.

I.1.87. Statistika $n\boldsymbol{a}$ yra suma nepriklausomų vienodai pasiskirsčiusių a. v. $(X_i^1, X_i^2,$ $..., X_i^k)^T$, $i = 1, 2, ..., n$, su vidurkių vektoriumi $\boldsymbol{\alpha}$ ir kovariacijų matrica $\boldsymbol{\Sigma} = [\sigma_{ij}]_{k\times k}$, $\sigma_{ij} = \mathbf{Cov}\,(X^i, X^j) = \mathbf{E}X^{i+j} - \mathbf{E}(X^i)\mathbf{E}(X^j) = \alpha_{i+j} - \alpha_i\alpha_j, i, j = 1, ..., k$. Pakanka pasinaudoti daugiamate CRT.

I.1.88. a) $m_2 = a_2 - \bar{X}^2$; kadangi $a_2 \xrightarrow{P} \alpha_2$, $a_1^2 \xrightarrow{P} \alpha_1^2$, tai $m_2 = a_2 - a_1^2 \xrightarrow{P} \alpha_2 - \alpha_1^2 = \sigma^2$. b) Nagrinėjant empirinę dispersiją (ar kitus aukštesnių eilių centrinius empirinius momentus), nemažinant bendrumo galima tarti, kad $\mathbf{E}X = \alpha_1 = 0$, ir pradinius momentus α_j galima pakeisti centriniais $\mu_j, j = 2, 3, ...$

Remiantis **I.1.87** pratimu

$$\sqrt{n}(\bar{X} - \alpha_1, a_2 - \alpha_2)^T \xrightarrow{d} \boldsymbol{Y} \sim N_2(\boldsymbol{\mu}, \Sigma);$$

čia $\boldsymbol{\mu} = (0,\ 0)^T$, $\Sigma = [\sigma_{ij}]_{2\times 2}, \sigma_{11} = \mu_2$, $\sigma_{12} = \mu_3, \sigma_{22} = \mu_4 - \sigma^4$. Statistika m_2 yra momentų $\bar{X}$ ir a_2 funkcija; $m_2 = g(\bar{X}, a_2) = a_2 - \bar{X}^2$; $g(\alpha_1, \alpha_2) = \alpha_2 - \alpha_1^2$; $g'_{\alpha_1}(\alpha_1, \alpha_2) =$ 0, $g'_{\alpha_2}(\alpha_1, \alpha_2) = 1$. Remiantis delta metodu

$$\sqrt{n}(m_2 - \sigma^2) \xrightarrow{d} Y \sim N(0, \mu_4 - \sigma^4),$$

arba

$$\sqrt{n}(m_2 - \sigma^2)/\sqrt{\mu_4 - \sigma^4} \xrightarrow{d} Z \sim N(0,\ 1).$$

Remiantis **I.1.85** pratimu $\boldsymbol{V}m_2 = (\mu_4 - \sigma^4)/n + O(1/n^2)$, tai $(\mu_4 - \sigma^4)/n$ galima pakeisti $\boldsymbol{V}m_2$. Kadangi $s^2 = nm_2/(n-1)$, tai m_2 galime pakeisti s^2.

c) Žr. [2], 2.5.1 teoremą.

I.1.89. Kadangi

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \xrightarrow{d} Z \sim N(0,\ 1),$$

tai reikia tik pasinaudoti tuo, kad $s_1^2 \xrightarrow{P} \sigma_1^2$, $s_2^2 \xrightarrow{P} \sigma_2^2$.

I.1.90. $\mathbf{Cov}\,(\bar{X}, s^2) = \mathbf{Cov}\,(\bar{Y}, s^2)$, kai $Y_i = X_i - \mu, \mathbf{E}Y_i = 0, \mathbf{E}Y_i^k = \mu_k$.

$$\mathbf{Cov}\,(\bar{Y}, s^2) = \mathbf{E}(\bar{Y}s^2) = \frac{1}{n(n-1)}\mathbf{E}[\sum_i Y_i(\sum_j Y_j^2 -$$

$$\frac{1}{n}(\sum_j Y_j^2 + \sum_{j\neq l} Y_j Y_l))] = \frac{\mu_3}{n}.$$

I.1.91. Jeigu pažymėsime $Y_i = X_i - \mu$, tai

$$m_3 = \frac{1}{n}\sum_{i=1}^n (Y_i - \bar{Y})^3 = a_3 - 3a_2\bar{Y} + 2\bar{Y}^3.$$

$$\mathbf{E}m_3 = \mathbf{E}[\frac{1}{n}\sum_{i=1}^n Y_i^3 - \frac{3}{n^2}\sum_{i=1}^n Y_i^2 \sum_{j=1}^n Y_j +$$

$$\frac{2}{n^3}(\sum_{i=1}^{n} Y_i^2 + \sum_{i\neq j} Y_iY_j)\sum_{l=1}^{n} Y_l] = \frac{(n-1)(n-2)}{n^2}\mu_3.$$

Analogiškai gauname

$$\mathbf{E}m_4 = \frac{(n-1)(n^2-3n+3)}{n^3}\mu_4 + \frac{3(n-1)(2n-3)}{n^3}\sigma^4.$$

Jeigu imsime statistikas

$$m_3^* = \frac{n^2}{(n-1)(n-2)}m_3, \quad m_4^* = \left[m_4 - \frac{3(2n-3)}{n^2-2n+3}m_2^2\right]\frac{n(n^2-2n+5)}{(n-1(n-2)(n-3))},$$

tai $\mathbf{E}m_3^* = \mu_3$, $\mathbf{E}m_4^* = \mu_4$.

I.1.92. Kadangi

$$m_k = C_k^0 a_k - C_k^1 a_{k-1}a_1 + ... + (-1)^k a_1^k$$

yra pradinių momentų $a_1, ..., a_k$ tolydi funkcija, tai remiantis daugiamate CRT (**I.1.87** pratimas) statistika m_k asimptotiškai turi normalųjį skirstinį. Pagal delta metodą asimptotinės dispersijos pagrindinė dalis $B_k^2 = \boldsymbol{\psi}^T\boldsymbol{\Sigma}\boldsymbol{\psi}$; vektorius $\boldsymbol{\psi} = (\psi_1, ..., \psi_k)^T$ gaunamas imant m_k dalines išvestines taške $\boldsymbol{\alpha}$ (primename, kad nagrinėjant centrinius momentus galima imti $\alpha_1 = 0, \alpha_j = \mu_j$).

Randame $\psi_1 = -k\alpha_{k-1}, \psi_2 = ... = \psi_{k-1} = 0, \psi_k = 1$.

$$nB_k^2 = \psi_1^2\alpha_2 + 2\psi_1\psi_k\alpha_{k+1} + \psi_k^2(\alpha_{2k} - \alpha_k^2) =$$

$$\mu_{2k} - \mu_k^2 - 2k\mu_{k-1}\mu_{k+1} + k^2\mu_{k-1}^2\sigma^2.$$

Normaliojo skirstinio atveju $\mu_{2r} = \sigma^{2r}(2r-1)!!, \mu_{2r+1} = 0$.

I.1.93. Gauname analogiškai ankstesniam pratimui taikydami delta metodą dvimačiu atveju.

I.1.94. Kadangi funkcija $H(\lambda)$ tolydi, o a. d. $\bar{X}$ galioja CRT, t. y. $\sqrt{n}(\bar{X}-\lambda) \xrightarrow{d} Y \sim N(0,\lambda)$, tai, remdamiesi delta metodu, gauname

$$\sqrt{n}\left(\frac{1}{\bar{X}^3} - \frac{1}{\lambda^3}\right) \xrightarrow{d} Y \sim N(0, B^2),$$

$$B^2 = [H'(\lambda)]^2\boldsymbol{V}X = \frac{9}{\lambda^7}.$$

I.1.95. a) Atlikta transformacija $s = H(s^2) = \sqrt{s^2}$. Funkcija H tolydi, o statistikai s^2 galioja CRT. Remiantis delta metodu funkcijai s irgi galioja CRT, o asimptotinės dispersijos pagrindinė dalis $B^2/n = [H'(\sigma^2)]^2\boldsymbol{V}s^) = (\mu_4 - \sigma^4)/(4n\sigma^2)$.

Kai skirstinys normalusis, s momentus galima rasti remiantis tuo, kad $(n-1)s^2/\sigma^2 \sim \chi^2(n-1)$.

b) Statistika $g_1 = H(m_2, m_3) = m_3/\sqrt{m_2^3}$. Tegu $H_1 = H'_{m_2}(\mu_2, \mu_3) = -3\mu_3/(2\sigma^5)$, $H_2 = H'_{m_3}(\mu_2, \mu_3) = 1/(\sigma^3)$. Tada remiantis delta metodu asimptotinės dispersijos pagrindinė dalis $B_1^2/n = [H_1^2\boldsymbol{V}m_2 + 2H_1H_2\mathbf{Cov}\,(m_2, m_3) + H_2^2\boldsymbol{V}m_3]$. Įrašę dispersijų ir kovariacijos išraiškas (**I.1.90**, **I.1.91** pratimai) ir sutraukę panašiuosius narius, gausime pratime pateiktą išraišką.

c) Statistika $g_2 = H(m_2, m_4) = m_4/m_2^2 - 3$. Tegu $H_1 = H'_{m_2}(\mu_2, \mu_4) = -2\mu_4/\sigma^6$, $H_2 = H'_{m_3}(\mu_2, \mu_4) = 1/(\sigma^4)$. Tada remiantis delta metodu asimptotinės dispersijos pagrindinė dalis $B_2^2/n = [H_1^2 \boldsymbol{V} m_2 + 2H_1H_2 \mathbf{Cov}\,(m_2, m_4) + H_2^2 \boldsymbol{V} m_4]$. Įrašę dispersijų ir kovariacijos išraiškas (**I.1.90**, **I.1.91** pratimai) ir sutraukę panašiuosius narius, gausime pratime pateiktą išraišką.

I.1.96. Kaip ir ankstesniuose pratimuose nemažinant bendrumo galima tarti, kad $\mathbf{E}X_i = \mathbf{E}Y_i = 0$. a) Randame

$$\mathbf{E}m_{11} = \mathbf{E}(a_{11} - a_{10}a_{01}) = \mathbf{E}[\frac{1}{n}\sum_{i=1}^n X_iY_i - \frac{1}{n^2}(\sum_{i=1}^n X_iY_i + \sum_{i\neq j} X_iY_j)] = \frac{n-1}{n}\mu_{11};$$

$$\boldsymbol{V} m_{11} = \mathbf{E}(a_{11}^2 - 2a_{11}a_{10}a_{01} + a_{10}^2a_{01}^2) - ((n-1)\mu_{11}/n)^2,$$

$$\mathbf{E}a_{11}^2 = \frac{1}{n^2}\mathbf{E}(\sum_{i=1}^n X_iY_i)^2 = \frac{1}{n^2}\mathbf{E}(\sum_{i=1}^n X_i^2Y_i^2 + \sum_{i\neq j} X_iY_iX_jY_j) = \frac{\mu_{22} + (n-1)\mu_{11}^2}{n},$$

$$\mathbf{E}(a_{11}a_{10}a_{01}) = \frac{1}{n^3}\mathbf{E}[\sum_{i=1}^n X_iY_i(\sum_{j=1}^n X_jY_j + \sum_{l\neq j} X_lY_j)] = \frac{\mu_{22} + (n-1)\mu_{11}^2}{n^2}.$$

Nesunku įsitikinti, kad $\mathbf{E}(a_{10}^2a_{01}^2) = O(1/n^2)$. Sutraukę panašiuosius narius, gauname

$$\boldsymbol{V}(m_{11}) = \frac{\mu_{22} - \mu_{11}^2}{n} + O(\frac{1}{n^2}).$$

b) Analogiškai gauname

$$\mathbf{Cov}\,(m_{20}, m_{11}) = \frac{\mu_{31} - \mu_{11}\mu_{20}}{n} + O(\frac{1}{n^2}), \quad \mathbf{Cov}\,(m_{20}, m_{11}) = \frac{\mu_{22} - \mu_{02}\mu_{20}}{n} + O(\frac{1}{n^2}).$$

I.1.97. Statistika $r = H(m_{11}, m_{20}, m_{02}) = m_{11}/\sqrt{m_{20}m_{02}}$. Tegu $H_1 = H'_{m_{11}}(\mu_{11}, \mu_{20}, \mu_{02}) = \rho/\mu_{11}$, $H_2 = H'_{m_{20}}(\mu_{11}, \mu_{20}, \mu_{02}) = -\rho/\mu_{20}$,
$H_3 = H'_{m_{02}}(\mu_{11}, \mu_{20}, \mu_{02}) = -\rho/\mu_{02}$. Tada remiantis delta metodu asimptotinės dispersijos pagrindinė dalis

$$B_r^2/n = [H_1^2\boldsymbol{V} m_{11} + H_2^2\boldsymbol{V} m_{20} + H_3^2\boldsymbol{V} m_{02} + 2H_1H_2\mathbf{Cov}\,(m_{11}, m_{20}) +$$
$$+ 2H_1H_3\mathbf{Cov}\,(m_{11}, m_{02}) + 2H_2H_3\mathbf{Cov}\,(m_{20}, m_{02})].$$

Įrašę dispersijų ir kovariacijos išraiškas ir sutraukę panašiuosius narius, gausime pratime pateiktą išraišką.

Normaliojo skirstinio atveju tardami, kad dispersijos vienetinės, naudodami charakteristinę funkciją gausime $\mu_{04} = \mu_{40} = 3, \mu_{31} = \mu_{13} = 3\rho, \mu_{11} = \rho, \mu_{22} = 1 + 2\rho^2$. Tada asimptotinės dispersijos pagrindinė dalis $B_r^2/n = (1-\rho^2)^2/n$.

I.1.98. Remiantis **I.1.97** pratimu $\sqrt{n}(r - \rho) \xrightarrow{d} Y \sim N(0, (1-\rho^2)^2)$. Naudojama Fišerio dispersiją stabilizuojanti transformacija $g(r) = (1/2)\ln((1+r)/(1-r))$. Kadangi $(g'(\rho))^2 = 1/(1-\rho^2)^2$, o dispersijos $\boldsymbol{V} r$ pagrindinė dalis $(1-\rho^2)^2/n$, tai asimptotinio skirstinio dispersija nepriklauso nuo parametro ρ:

$$\sqrt{n}(\frac{1}{2}\ln\frac{1+r}{1-r} - \frac{1}{2}\ln\frac{1+\rho}{1-\rho}) \xrightarrow{d} Z \sim N(0,\ 1).$$

I.2. Parametrų įvertiniai

I.2.1. Įvertiniai ir jų klasifikacija

I.2.1. Kokio didumo turi būti a. d. $X \sim N(\mu,\ 4)$ imtis, kad parametro μ įvertinio $\hat{\mu} = \bar{X}$ absoliučioji paklaida būtų ne didesnė už 0,1 su tikimybe, ne mažesne kaip 0,99?

I.2.2. Kokio didumo turi būti normaliojo a. d. imtis, kad dispersijos σ^2 įvertinio $\hat{\sigma}^2 = s^2$ santykinės paklaidos modulis būtų ne didesnis už 0,1 su tikimybe, ne mažesne kaip 0,99?

I.2.3. Kokio didumo turi būti a. d. $X \sim G(1/\lambda,\ 5)$ imtis, kad parametro λ įvertinio $\hat{\lambda} = \bar{X}/5$ santykinės paklaidos modulis būtų ne didesnis už 0,1 su tikimybe, ne mažesne kaip 0,99?

I.2.4. Jūros gylis matuojamas prietaisu, kurio sisteminė paklaida 0, o atsitiktinė paklaida pasiskirsčiusi pagal normalųjį dėsnį su vidutiniu kvadratiniu nuokrypiu $\sigma = 25\,\text{m}$. Kiek kartų reikia nepriklausomai matuoti jūros gylį, kad matavimo paklaida būtų ne didesnė kaip 15 m su tikimybe, ne mažesne už 0,99?

I.2.5. Parametro θ įvertinio $\hat{\theta}_n = T_n = T_n(X_1, ..., X_n)$ poslinkis yra

$$\mathbf{E}T_n - \theta = \sum_{i=1}^{\infty} \frac{a_i}{n^i} = O(\frac{1}{n}), \quad n \to \infty.$$

Įrodykite, kad statistikos $T'_n = nT_n - (n-1)\bar{T}_{n-1}$ poslinkis yra $O(1/n^2)$; statistikos $T''_n = [n^2 T'_n - (n-1)\bar{T}'_{n-1}]/[n^2 - (n-1)^2]$ poslinkis yra $O(1/n^3)$ ir t. t. Čia $\bar{T}_{n-1}$ yra aritmetinis vidurkis statistikų T_{n-1}, apskaičiuotų pagal visus didumo $n-1$ imties poaibius.

I.2.6. $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. X, kurio skirstinys priklauso Veibulo skirstinių šeimai $\mathcal{P} = \{f(x|\rho),\ 0 < \rho < \infty\}$; čia tankio funkcija yra

$$f(x|\rho) = \alpha \rho^{\alpha} x^{\alpha - 1} \mathrm{e}^{-(\rho x)^{\alpha}}, \quad 0 < x < \infty,$$

o $\alpha > 1$ – žinoma konstanta. Įrodykite, kad funkcijos $\gamma(\rho) = \rho^r$ nepaslinktasis įvertinys, kai $n > r/\alpha$, yra

$$\hat{\gamma} = \left(\sum_{i=1}^{n} X_i^{\alpha} \right)^{-r/\alpha} \frac{(n-1)!}{\Gamma(n - r/\alpha)}.$$

I.2.7. Tegu $\bar{\theta} = \bar{\theta}(\mathbf{X})$ yra nepaslinktasis θ įvertinys. Įrodykite, kad bet kuris nepaslinktasis θ įvertinys $\widetilde{\theta}$ yra tokio pavidalo:

$$\widetilde{\theta} = \bar{\theta} - U(\mathbf{X});$$

čia $U(\mathbf{X})$ tenkina sąlygą $\mathbf{E}_{\theta}(U(\mathbf{X})) \equiv 0$.

I.2.8. Tegu $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. $X \sim U(0,\ \theta)$, $0 < \theta < \infty$. Raskite θ įvertinio $cX_{(n)}$ poslinkį ir dispersiją; čia c – žinoma teigiama konstanta. Raskite c tokį, kad $cX_{(n)}$ būtų nepaslinktasis θ įvertinys.

I.2.9. Tarkime, $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. X, kurio skirstinys priklauso šeimai $\mathcal{P} = \{\mathbf{P}_{\theta},\ \theta \in \{\theta_1, ..., \theta_k\}\}$ su fiksuotu natūraliuoju skaičiumi

k. Tegu $T_n(X)$ yra θ įvertinys, įgyjantis reikšmes iš aibės $\{\theta_1, \dots, \theta_k\}$. Įrodykite, kad $T_n(X)$ yra pagrįstasis tada ir tik tada, kai $\mathbf{P}_\theta\{T_n(X) = \theta\} \to 1$, $n \to \infty$.

I.2.10. Tegu $\mathbf{X} = (X_1, \dots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. $X \sim U(\theta - 1/2,\ \theta + 1/2)$, $\theta \in \mathbf{R}$ yra nežinomas. Įrodykite, kad $\bar{\theta} = (X_{(1)} + X_{(n)})/2$ yra $n^\alpha, \alpha < 1$, pagrįstas θ įvertinys, t. y. $n^\alpha(\bar{\theta} - \theta) \xrightarrow{P} 0$, kai $n \to \infty$.

I.2.11. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji a. d. $X \sim B(1, p)$ imtis. Tarkime, statistika T įgyja reikšmę 0, jei ne mažiau kaip pusė visų X_i yra 0; įgyja reikšmę 1, jeigu daugiau negu pusė visų X_i yra 1; įgyja reikšmę 1/2, jeigu lygiai pusė visų X_i yra 0. Įrodykite, kad statistika T nėra pagrįstasis tikimybės p įvertinys.

I.2.12. Tegu $g_1, g_2, \dots$ yra tokios tolydžiosios, apibrėžtos intervale $(a,\ b) \subset \mathbf{R}$, funkcijos, kad $g_n(x) \to g(x)$ tolygiai pagal x bet kuriame intervale, priklausančiame $(a,\ b)$. Tegu T_n yra pagrįstasis $\theta \in (a,\ b)$ įvertinys. Įrodykite, kad $g_n(T_n)$ yra pagrįstasis parametro $\vartheta = g(\theta)$ įvertinys.

I.2.13. Tegu $\mathbf{X} = (X_1, \dots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. X, kurio vidurkis $\mu \in \mathbf{R}$ ir dispersija $\sigma^2 > 0$ nežinomi. Be to, $g(\mu) = 0$, kai $\mu \neq 0$, ir $g(0) = 1$. Nurodykite pagrįstąjį $\vartheta = g(\mu)$ įvertinį.

I.2.14. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji atsitiktinė imtis a. d. $X \sim N(\mu,\ \sigma^2)$, $-\infty < \mu < \infty$, $\sigma > 0$. Įrodykite, kad aritmetinis vidurkis $\bar{X}$ ir empirinė mediana $\hat{x}(0,5)$ yra vidurkio μ pagrįstieji įvertiniai.

I.2.15. Tarkime, $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji atsitiktinė imtis a. d. $X \sim K(\mu,\ \sigma)$, $-\infty < \mu < \infty$, $\sigma > 0$. Įrodykite, kad aritmetinis vidurkis $\bar{X}$ nėra pagrįstasis, o empirinė mediana $\hat{x}(0,5)$ yra pagrįstasis parametro μ įvertinys.

I.2.16. Tegu X ir Y yra nepaslinktieji atitinkamai parametrų θ ir θ^2 įvertiniai. Raskite a. d. X dispersijos $\boldsymbol{V}_\theta X$ nepaslinktąjį įvertinį.

I.2.17. Tegu X ir Y yra nepriklausomi atsitiktiniai dydžiai, vienodai pasiskirstę pagal $N(\mu,\ \sigma^2)$. Ar $|X - Y|$ ir $|X - Y|^2$ yra nepaslinktieji parametrų σ ir σ^2 įvertiniai? Raskite šių įvertinių poslinkius.

I.2.18. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji atsitiktinė imtis a. d. $X \sim U(-\theta,\ \theta)$, $0 < \theta < \infty$. Iš kokios konstantos reikia padauginti statistiką $T = X_{(n)} - X_{(1)}$, kad gautume nepaslinktąjį parametro θ įvertinį? Kokia gautojo įvertinio dispersija?

I.2.2. Pakankamosios statistikos. NMD įvertiniai

I.2.19. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. $X \sim \mathcal{E}(\lambda,\ \mu)$, $0 < \lambda < \infty$, $-\infty < \mu < \infty$. Įrodykite, kad $\boldsymbol{T} = (X_{(1)}, X_{(2)} + ... + X_{(n)})^T$ yra pakankamoji parametro $(\lambda,\ \mu)^T$ statistika.

I.2.20. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. $X \sim W(\eta,\ \sigma), 0 < \eta < \infty$, $0 < \sigma < \infty$, turinčio Veibulo skirstinį. Įrodykite, kad: a) kai η nežinomas, egzistuoja tik triviali pakankamoji statistika; b) kai η žinomas, $T = \sum_i X_i^\eta$ yra parametro σ pakankamoji statistika.

I.2.21. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. X, kurio skirstinys priklauso šeimai $\mathcal{P} = \{f(x;\ \lambda,\ \eta,\ \mu),\ 0 < \lambda,\ \eta < \infty,\ -\infty < \mu < \infty\}$; čia tankio funkcija

$$f(x;\ \lambda,\ \eta,\ \mu) = \frac{\lambda^\eta}{\Gamma(\eta)} x^{\eta-1} \mathrm{e}^{-\lambda(x-\mu)},\ x \geq \mu.$$

Raskite pakankamąją statistiką: a) parametro η, kai μ ir λ žinomi; b) parametro λ, kai μ ir η žinomi; c) parametrų η ir λ, kai μ žinomas; d) parametro μ, kai λ žinomas, o $\eta = 1$.

I.2.22. Tegu $Y_1, ...Y_n$ yra n. a. d. ir $Y_i \sim N(\alpha + \beta x_i,\ \sigma^2)$; čia $x_1, ..., x_n$ – žinomos konstantos, $-\infty < \alpha, \beta < \infty$, $0 < \sigma < \infty$ – nežinomi parametrai. Įrodykite, kad $\boldsymbol{T} = (\sum_i Y_i,\ \sum_i Y_i x_i,\ \sum_i Y_i^2)^T$ yra pilnoji ir pakankamoji parametro $\boldsymbol{\theta} = (\alpha,\ \beta,\ \sigma)^T$ statistika.

I.2.23. Tegu $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d., kurio skirstinys priklauso šeimai $\mathcal{P} = \{P_{\boldsymbol{\theta}},\ \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$. Raskite parametro $\boldsymbol{\theta}$ pakankamąją statistiką, kai $P_{\boldsymbol{\theta}}$ yra:

a) Puasono skirstinys $\mathcal{P}(\lambda)$, $\lambda \in (0, \infty)$;
b) neigiamas binominis skirstinys $B^-(n,\ p)$ su žinomu n, $p \in (0,\ 1)$;
c) eksponentinis skirstinys $\mathcal{E}(\theta)$, $\theta \in (0, \infty)$;
d) gama skirstinys $G(\lambda, \eta)$, $\boldsymbol{\theta}^T = (\lambda, \eta) \in (0, \infty) \times (0, \infty)$;
e) beta skirstinys $Be(\gamma, \eta)$, $\boldsymbol{\theta}^T = (\gamma, \eta) \in (0,\ \infty) \times (0,\ \infty)$;
f) lognormalusis skirstinys $LN(\mu, \sigma)$, $\boldsymbol{\theta}^T = (\mu, \sigma) \in \mathcal{R} \times (0, \infty)$.

I.2.24. Atsitiktinio dydžio X skirstinys priklauso šeimai $\mathcal{P} = \{f(x;\ \theta),\ 0 < \theta < \infty\}$; čia tankis $f(x;\ \theta) = 2(\theta - x)/\theta^2$, kai $0 < x < \theta$, ir lygus 0, kai x įgyja kitas reikšmes. Raskite šeimos $\mathcal{P}$ netrivialią pakankamąją statistiką pagal a. d. X didumo n paprastąją imtį.

I.2.25. Tegu $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. X, kurio tankio funkcija

$$f(x; \theta) = \theta(1 + x)^{-(\theta+1)}, \quad x \geqslant 0;$$

čia $\theta > 0$ nežinomas parametras.

a) Įrodykite, kad $T = \sum_{i=1}^n \ln(1 + X_i)$ yra pakankamoji θ statistika.
b) Raskite T vidurkį ir dispersiją.

I.2.26. Įrodykite: jei T yra pakankamoji statistika ir $T = h(S)$, čia h yra mačioji funkcija, S – kita statistika, tai S taip pat yra pakankamoji statistika.

I.2.27. Tegu $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. X, kurio skirstinys yra Pareto ir tankis

$$f(x|a,\ \theta) = \theta a^\theta / x^{\theta+1},\ 0 < a,\ \theta < \infty,\ a < x < \infty.$$

Raskite parametro $(a,\ \theta)^T$ pakankamąją statistiką.

I.2.28. Tegu $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. X, kurio skirstinys priklauso visų absoliučiai tolydžiųjų skirstinių šeimai $\mathcal{P}$. Įrodykite, kad variacinė eilutė $(X_{(1)}, ..., X_{(n)})^T$ yra šeimos $\mathcal{P}$ pakankamoji statistika.

I.2.29. Įrodykite, kad $\mathcal{P} = \{P_{\boldsymbol{\theta}} :\ \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ yra eksponentinė šeima. Užrašykite jos kanoninį pavidalą ir natūraliąją parametrų erdvę, kai $P_{\boldsymbol{\theta}}$ yra:

a) Puasono skirstinys $\mathcal{P}(\lambda)$, $\lambda \in \Theta = (0, \infty)$;
b) neigiamas binominis skirstinys $B^-(n,\ p)$ su fiksuotu n ir $p \in \Theta = (0,\ 1)$;
c) eksponentinis skirstinys $E(a, \theta)$ su fiksuotu a ir $\theta \in \Theta = (0, \infty)$;
d) gama skirstinys $G(\lambda, \eta)$, $\boldsymbol{\theta}^T = (\lambda, \eta) \in \boldsymbol{\Theta} = (0, \infty) \times (0, \infty)$;
e) beta skirstinys $Be(\gamma, \eta)$, $\boldsymbol{\theta}^T = (\gamma, \eta) \in \boldsymbol{\Theta} = (0,\ 1) \times (0,\ 1)$;
f) Veibulo skirstinys $W(\alpha, \theta)$ su fiksuotu $\alpha > 0$ ir $\theta \in \Theta = (0, \infty)$.

I.2.30. Įrodykite, kad eksponentinių skirstinių šeima $\mathcal{E}(a, \theta)$ su dviem nežinomais parametrais a ir θ nėra eksponentinė šeima.

I.2.31. Įrodykite, kad neigiamų binominių skirstinių šeima $B^{-}(n,\ p)$ su dviem nežinomais parametrais p ir n nėra eksponentinė šeima.

I.2.32. Įrodykite, kad Koši skirstinių šeima $K(\mu, \sigma)$ su dviem nežinomais parametrais μ ir σ nėra eksponentinė šeima.

I.2.33. Įrodykite, kad Veibulo skirstinių šeima $W(\alpha, \theta)$ su dviem nežinomais parametrais α ir θ nėra eksponentinė šeima.

I.2.34. Įrodykite, kad k-mačių normaliųjų skirstinių šeima $N_k(\boldsymbol{\mu},\ \boldsymbol{\Sigma})$ yra eksponentinė šeima. Užrašykite jos kanoninį pavidalą.

I.2.35. Raskite gama skirstinio $G(\lambda, \eta)$ momentų generuojančiąją funkciją.

I.2.36. Diskrečiojo a. d. X skirstinys nusakomas tikimybėmis

$$\mathbf{P}\{X = k\} = \gamma(k)\frac{\theta^k}{c(\theta)}, \quad k = 0, 1, 2, \dots .$$

Įrodykite, kad šie skirstiniai, kai $\theta > 0$, sudaro eksponentinę šeimą, ir raskite X momentų generuojančiąją funkciją.

I.2.37. Tegu X yra a. d., kurio skirstinys priklauso šeimai $\mathcal{P} = \{P_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ ir $f_{\boldsymbol{\theta}}$ yra $P_{\boldsymbol{\theta}}$ tankio funkcija σ-baigtinio mato ν atžvilgiu, o A yra įvykis, kurio $\mathbf{P}_{\boldsymbol{\theta}}\{A\} > 0$. Nagrinėjama nupjautinių skirstinių šeima, $\mathcal{P}_A = \{f_{\boldsymbol{\theta}} I_{\{A\}}/P_{\boldsymbol{\theta}}\{A\},\ \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$. Įrodykite, kad:

a) jeigu $T(X)$ yra pakankamoji šeimos $\mathcal{P}$ statistika, tai ji pakankamoji ir šeimos $\mathcal{P}_A$ statistika;

b) jeigu $T(X)$ yra pilnoji ir pakankamoji šeimos $\mathcal{P}$ statistika, tai ji pilnoji ir pakankamoji ir šeimos $\mathcal{P}_A$ statistika.

I.2.38. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji atsitiktinė imtis a. d. X, kurio tankis

$$f(x|\theta) = c(\theta)\exp\{-\theta x\},\ 0 < x < a,\ \theta > 0,$$

konstanta a žinoma. Raskite parametro θ pilnąją ir pakankamąją statistiką.

I.2.39. Tegu $\mathbf{X} = (X_1, \dots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. $X \sim U(\theta,\ \theta + 1)$, $0 < \theta < \infty$. Įrodykite, kad pakankamoji statistika $(X_{(1)}, X_{(n)})^T$ nėra pilnoji.

I.2.40. Tegu $\psi(x)$, $x \in \mathbf{R}$, yra tokia teigiama Borelio funkcija, kad bet kuriems a ir b, $-\infty < a < b < \infty$, $\int_a^b \psi(x)dx < \infty$. Tegu $\boldsymbol{\theta} = (a,\ b)^T$. Apibrėžkime tankio funkciją

$$f(x|a,\ b) = \frac{\psi(x)}{\int_a^b \psi(u)du},\ \ a < x < b.$$

Įrodykite, kad $(X_{(1)}, X_{(n)})^T$ yra šeimos $\mathcal{P} = \{f(x|\boldsymbol{\theta}),\ \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ pilnoji ir pakankamoji statistika.

I.2.41. Tegu X yra diskretusis a. d., kurio skirstinys nusakytas tikimybėmis

$$\mathbf{P}\{X = k|\theta\} = \begin{cases} \theta, & k = 0, \\ (1-\theta)^2\theta^{k-1}, & k = 1, 2, \dots, \end{cases}$$

čia $\theta \in (0,\,1)$. Įrodykite, kad X nėra pilnoji, tačiau yra aprėžtai pilnoji.

I.2.42. Tegu $(X_1, ..., X_n)^T)$ yra paprastoji imtis, gauta stebint Bernulio a. d. $X \sim B(1,p)$. Įrodykite, kad $S_n = X_1 + ... + X_n$ yra pakankamoji statistika a) remdamiesi pakankamosios statistikos apibrėžimu; b) naudodami Neimano ir Fišerio faktorizacijos kriterijų.

I.2.43. (**I.2.42** pratimo tęsinys). Įrodykite, kad $S_n = X_1 + ... + X_n$ yra pilnoji statistika a) remdamiesi apibrėžimu; b) remdamiesi tuo, kad Bernulio skirstinys priklauso eksponentinio tipo skirstinių šeimai.

I.2.44. (**I.2.42** pratimo tęsinys). Raskite parametro $\gamma = \gamma(p) = p^i(1-p)^j, 0 \le i, j \le n, 1 \le i + j \le n$, dviem būdais a) vidurkindami nepaslinktąjį įvertinį pakankamosios statistikos atžvilgiu; b) spręsdami funkcinę lygtį $\mathbf{E}_p h(S_n) \equiv \gamma(p)$ funkcijos h atžvilgiu.

I.2.45. (**I.2.42** pratimo tęsinys). Įrodykite, kad a) polinomo $t_k(p) = \sum_{i=1}^k a_i p^i$ NMD įvertinys yra $\hat{t}_k(p) = \sum_{i=1}^k a_i S_n^{[i]}/n^{[i]}$; b) parametro p^k, kai $k > n$, NMD įvertinys neegzistuoja.

I.2.46. (**I.2.42** pratimo tęsinys). Atlikus $n = 50$ Bernulio eksperimentų, kai įvykio A tikimybė yra p, įvykis A įvyko 12 kartų. Raskite parametrų $q + p^3, pq, p^{12}q^{38}, q^{50}, p^{49}$ NMD įvertinių realizacijas (įverčius).

I.2.47. Tegu $(X_1, ..., X_n)^T$ yra paprastoji imtis, gauta stebint Puasono a. d. $X \sim \mathcal{P}(\lambda)$. Įrodykite, kad $S_n = X_1 + ... + X_n$ yra pakankamoji statistika a) remdamiesi pakankamosios statistikos apibrėžimu; b) naudodami Neimano ir Fišerio faktorizacijos kriterijų.

I.2.48. (**I.2.47** pratimo tęsinys). Įrodykite, kad $S_n = X_1 + ... + X_n$ yra pilnoji statistika a) remdamiesi apibrėžimu; b) remdamiesi tuo, kad Puasono skirstinys priklauso eksponentinio tipo skirstinių šeimai.

I.2.49. (**I.2.47** pratimo tęsinys). a) Raskite parametro $\gamma = \gamma(\lambda) = \lambda^k$ NMD įvertinį; b) Įrodykite, kad parametro $1/\lambda$ NMD įvertinys neegzistuoja.

I.2.50. (**I.2.47** pratimo tęsinys). a) Raskite tikimybės $\pi_m(\lambda) = \lambda^m e^{-\lambda}/m!$ NMD įvertinį; b) Raskite generuojančios funkcijos $g(s) = e^{\lambda(s-1)}$ NMD įvertinį.

I.2.51. Tegu $(X_1, ..., X_n)^T)$ yra paprastoji imtis, gauta stebint a. d. $X \sim B^-(1,p)$. Įrodykite, kad $S_n = X_1 + ... + X_n$ yra pilnoji ir pakankamoji statistika.

I.2.52. (**I.2.51** pratimo tęsinys). Raskite imties sąlyginį skirstinį, kai pakankamoji statistika $S_n = t$ yra fiksuota.

I.2.53. (**I.2.51** pratimo tęsinys). Raskite parametro $\gamma = \gamma(p) = p^m q^l, l = 0, 1, ..., 1 \le m \le n$, NMD įvertinį.

I.2.54. (**I.2.51** pratimo tęsinys). Raskite parametrų $q/p, p, p^2, pq$ NMD įvertinius.

I.2.55. (**I.2.51** pratimo tęsinys). Atliekant Bernulio eksperimentus įvykis A pirmą kartą įvyko per 14 bandymą. Raskite parametro $p \ln p$ NMD įvertinio realizaciją; čia p yra įvykio A pasirodymo tikimybė per atskirą bandymą.

I.2.56. Tegu a. v. $\mathbf{X} = (X_1, ..., X_n)^T$ turi k-matį polinominį skirstinį $\mathcal{P}_k(n; (p_1, ..., p_k)), 0 < p_i < 1, p_1 + ... + p_k = 1$. a) Įrodykite, kad a. v. $(X_1, ..., X_{k-1})^T$ yra parametro $\boldsymbol{\theta} = (p_1, ..., p_{k-1})^T$ pilnoji ir pakankamoji statistika. b) Raskite parametrų p_i^m ir $p_i^k p_j^l$ NMD įvertinius.

I.2.57. Tegu $(X_1, ..., X_n)^T)$ yra paprastoji imtis, gauta stebint a. d.

$$X \sim U(0, \theta), \quad \theta > 0.$$

Įrodykite, kad $X_{(n)}$ yra pilnoji ir pakankamoji parametro θ statistika.

I.2.58. (**I.2.57** pratimo tęsinys). Raskite parametro θ NMD įvertinį.

I.2.59. Tegu $(X_1, ..., X_n)^T$ yra paprastoji imtis a. d. $X \sim U(\theta, \theta + 1), \theta > 0$. Įrodykite, kad vienmatė pakankamoji statistika neegzistuoja. Dvimatė statistika $\boldsymbol{T} = (X_{(1)}, X_{(n)})^T$ yra pakankamoji, tačiau nėra pilnoji.

I.2.60. Tegu $(X_1, ..., X_n)^T)$ yra paprastoji imtis, gauta stebint a. d.

$$X \sim U(\theta_1, \theta_2), \quad \theta_1 < \theta_2.$$

Įrodykite, kad a) $(X_{(1)}, X_{(n)})^T$ yra pilnoji ir pakankamoji parametro $(\theta_1, \theta_2)^T$ statistika; b) parametrų $\gamma_1 = (\theta_1 + \theta_2)/2$ ir $\gamma_2 = \theta_2 - \theta_1$ NMD įvertiniai yra $\hat{\gamma}_1 = (X_{(1)} + X_{(n)})/2$, $\hat{\gamma}_2 = (n+1)(X_{(n)} - X_{(1)})/(n-1)$, o jų dispersijos $\boldsymbol{V}_{\theta_1,\theta_2}\hat{\gamma}_1 = \gamma_2^2/(2(n+1)(n+2))$, $\boldsymbol{V}_{\theta_1,\theta_2}\hat{\gamma}_2 = 2\gamma_2^2/((n-1)(n+2))$.

I.2.61. Tegu $(X_1, ..., X_n)^T)$ yra paprastoji imtis, gauta stebint a. d.

$$X \sim U(-\theta, \theta), \quad \theta > 0.$$

Įrodykite, kad a) $T = \max(-X_{(1)}, X_{(n)})$ yra pilnoji ir pakankamoji parametro θ statistika; b) parametro θ NMD įvertinys yra $\hat{\theta} = (n+1)T/n$.

I.2.62. Paprastoji imtis $(X_1, ..., X_n)^T$ gauta stebint a. d.

$$X \sim N(\mu, \sigma^2), \quad -\infty < \mu < \infty, \quad \sigma > 0.$$

Įrodykite, kad a) $\boldsymbol{T} = (\bar{X}, s^2)^T$, $\bar{X} = \sum_{i=1}^n X_i/n$, $s^2 = \sum_{i=1}^n (X_i - \bar{X})^2/(n-1)$, yra pilnoji ir pakankamoji parametro $(\mu, \sigma^2)^T$ statistika; b) $\bar{X}$ ir s^2 yra parametrų μ ir σ^2 NMD įvertiniai.

I.2.63. (**I.2.62** pratimo tęsinys). Įrodykite, kad parametro σ NMD įvertinys yra

$$\hat{\sigma} = s/M_{n-1}, \quad M_\nu = \sqrt{2/\nu}\,\Gamma((\nu+1)/2)/\Gamma(\nu/2).$$

I.2.64. (**I.2.62** pratimo tęsinys). Įrodykite, kad a) jeigu parametras μ žinomas, tai parametro σ^2 pilnoji ir pakankamoji statistika yra $s_0^2 = \sum_{i=1}^n (X_i - \mu)^2/n$, o parametrų σ^2 ir σ NMD įvertiniai yra s_0^2 ir s_0/M_n; b) jeigu parametras σ žinomas, parametro μ pilnoji ir pakankamoji statistika yra $\bar{X}$, kuri yra parametro μ NMD įvertinys.

I.2.65. Tegu paprastoji imtis $(X_1, ..., X_n)^T$ gauta stebint a. d.

$$X \sim \mathcal{E}(\mu, \lambda), \quad -\infty < \mu < \infty, \quad \lambda > 0.$$

Įrodykite, kad a) $(X_{(1)}, \sum_{i=1}^n X_i)^T$ yra pilnoji ir pakankamoji parametro $(\mu, \lambda)^T$ statistika; b) $n\lambda(X_{(1)} - \mu) \sim \mathcal{E}(1)$, $\lambda T_2(\mathbf{X}) = \lambda[\sum_{i=1}^n X_i - nX_{(i)}] \sim G(1, n-1)$, o $X_{(1)}$ ir $T_2(\mathbf{X})$ nepriklausomi; c) parametrų $\gamma_1 = \mu^l$ ir $\gamma_2 = 1/\lambda^m$ NMD įvertiniai yra $\hat{\gamma}_1 = X_{(1)}^l - lT_2(\mathbf{X})X_{(1)}^{l-1}/(n(n-1))$ ir $\hat{\gamma}_2 = T_2^m(\mathbf{X})\Gamma(n-1)/\Gamma(n+m-1)$.

I.2.66. (**I.2.65** pratimo tęsinys). Įrodykite, kad a) jei parametras μ žinomas, tai parametro $1/\lambda$ NMD įvertinys yra $\bar{X} - \mu$; b) jeigu λ žinomas, tai parametro μ NMD įvertinys yra $\hat{\mu} = X_{(1)} - 1/(n\lambda)$.

I.2.67. Tegu paprastoji imtis $(X_1, ..., X_n)^T$ gauta stebint a. d. $X \sim G(\lambda, \eta), \lambda, \eta > 0$. Raskite parametro $(\lambda, \eta)^T$ pilnąją ir pakankamąją statistiką. Įrodykite, kad $\bar{X}$ yra parametro $\gamma = \eta/\lambda$ NMD įvertinys.

I.2.68. (**I.2.67** pratimo tęsinys). Įrodykite, kad jei parametras η žinomas, tai parametro λ pilnoji ir pakankamoji statistika yra $T = \sum_{i=1}^n X_i$. Parametro $\gamma = \lambda^k$ NMD įvertinys yra $\hat{\gamma} = \Gamma(n\eta)/[\Gamma(n\eta - k)T^k], n\eta > k$.

I.2.69. (**I.2.67** pratimo tęsinys). Įrodykite, kad jei parametras λ žinomas, tai parametro $\gamma = \Gamma'(\eta)/\Gamma(\eta)$ NMD įvertinys yra $\ln \lambda + (\sum_i \ln X_i)/n$.

I.2.70. Tegu paprastoji imtis $(X_1, ..., X_n)^T$ gauta stebint a. d. X, kurio tankio funkcija $f((x - \theta_1)/\theta_2)/\theta_2$ priklauso nuo poslinkio ir mastelio parametrų $-\infty < \theta_1 < \infty, \theta_2 > 0$. Tarkime, egzistuoja pilnoji ir pakankamoji parametro $\boldsymbol{\theta} = (\theta_1, \theta_2)^T$ statistika $\boldsymbol{T}(\mathbf{X}) = (T_1(\mathbf{X}), T_2(\mathbf{X}))^T$ ir tokie įvertiniai $\hat{\theta}_1, \hat{\theta}_2$, kad statistikos $U(\mathbf{X}) = (X_1 - \hat{\theta}_1)/\hat{\theta}_2$ skirstinys nepriklauso nuo parametro $\boldsymbol{\theta}$. Įrodykite, kad jei egzistuoja funkcija $h(z)$, kad $\mathbf{E}_{\boldsymbol{\theta}}(h(X_1)) = \tau(\boldsymbol{\theta}) < \infty$ su visais $\boldsymbol{\theta}$, tai parametro $\tau(\boldsymbol{\theta})$ NMD įvertinys yra

$$\hat{\tau}(\boldsymbol{T}(\mathbf{X})) = \int_{-\infty}^{\infty} h(u\hat{\theta}_2 + \hat{\theta}_1)g(u)du,$$

čia $g(u)$ yra statistikos $U(\mathbf{X})$ tankis. Jeigu vienas iš parametrų θ_1 arba θ_2 yra žinomas, tai $U(\mathbf{X})$ ir $\hat{\tau}(\boldsymbol{T})$ išraiškose reikia atitinkamą įvertinį pakeisti žinomu parametru.

I.2.71. Tegu paprastoji imtis $(X_1, ..., X_n)^T$ gauta stebint a. d. $X \sim N(\mu, 1), -\infty < \mu < \infty$. Raskite parametrų $\gamma_1 = \Phi(y - \mu)$ ir $\gamma_2 = \varphi(y - \mu)$ NMD įvertinius.

I.2.72. Tegu paprastoji imtis $(X_1, ..., X_n)^T$ gauta stebint a. d. $X \sim N(0, \sigma^2), \sigma > 0$. Raskite parametrų $\gamma_1 = \Phi(y/\sigma)$ ir $\gamma_2 = \varphi(y/\sigma)/\sigma$ NMD įvertinius.

I.2.73. Tegu paprastoji imtis $(X_1, ..., X_n)^T$ gauta stebint a. d. $X \sim N(\mu, \sigma^2)$, su abiem nežinomais parametrais $-\infty < \mu < \infty, \sigma > 0$. Įrodykite, kad parametro $\gamma = \Phi((y - \mu)/\sigma)$ NMD įvertinys yra

$$\hat{\gamma} = \int_{-\sqrt{1-1/n}}^{(y-\bar{X})/S} g(u)du, \quad kai \quad |y - \bar{X}|/S \le \sqrt{1 - 1/n},$$

čia $\bar{X} = \sum_i X_i/n$, $S^2 = \sum_i (X_i - \bar{X})^2$,

$$g(u) = \frac{\Gamma((n-1)/2)}{\sqrt{\pi}\Gamma((n-2)/2)} \sqrt{\frac{n}{n-1}} (1 - \frac{nu^2}{n-1})^{(n-4)/2}, \quad |u| \le \sqrt{1 - 1/n}.$$

I.2.74. Tegu paprastoji imtis $(X_1, ..., X_n)^T$ gauta stebint a. d. $X \sim \mathcal{E}(\theta, 1)$, $-\infty < \theta < \infty$. Įrodykite, kad parametro $\gamma = \mathbf{P}_\lambda\{X > y\} = e^{-(y-\theta)}$ NMD įvertinys yra

$$\hat{\gamma} = \frac{n-1}{n} e^{-(y - X_{(1)})}, \quad y > X_{(1)}.$$

I.2.75. Tegu paprastoji imtis $(X_1, ..., X_n)^T$ gauta stebint a. d. $X \sim \mathcal{E}(a, \lambda)$, $\lambda > 0$, parametras a žinomas. Raskite parametro $\gamma = P\{X > y\}$ NMD įvertinį.

I.2.76. Tegu paprastoji imtis $(X_1, ..., X_n)^T$ gauta stebint a. d. X, kuris turi Pareto skirstinį su tankio funkcija

$$f(x|\alpha, \sigma) = \sigma\alpha^\sigma / y^{\sigma+1}, \quad y > \alpha, \quad \sigma > 0.$$

Įrodykite, kad parametro $\gamma = \alpha^m, m < n\sigma$, NMD įvertinys, kai σ žinomas, yra

$$\hat{\gamma} = [\ln X_{(1)}]^m (1 - \frac{m}{n\sigma}).$$

I.2.77. Tegu paprastoji imtis $(X_1, ..., X_n)^T$ gauta stebint a. d. X, kurio tankio funkcija

$$f(x|\theta) = e^{-(x-\theta)} \exp\{-e^{-(x-\theta)}\}, \quad -\infty < x, \theta < \infty.$$

Raskite parametro $\gamma = f(z|\theta)$ NMD įvertinį.

I.2.78. Didumo N gaminių partijoje defektinių gaminių skaičius M nežinomas. Atsitiktinai negrąžinant atrenkama n gaminių. Tegu $X_i = 1$, jei i-asis atrinktas gaminys defektinis, ir $X_i = 0$ priešingu atveju. Tarsime, kad $0 < n < N$ yra žinomi, o M nežinomas parametras, įgyjantis reikšmes $0, 1, ..., N$. Įrodykite, kad $T = \sum_{i=1}^n X_i$ yra parametro M pilnoji ir pakankamoji statistika.

I.2.79. Turima didumo $n = 1$ imtis diskrečiojo a. d. X, kurio skirstinys, priklausantis nuo nežinomo parametro $\theta, 0 < \theta < 1$, nusakytas tikimybėmis

$$\mathbf{P}\{X = -1\} = \theta, \quad \mathbf{P}\{X = k\} = (1-\theta)^2 \theta^k, \quad k = 0, 1, 2, ...$$

Įrodykite, kad

a) X yra pakankamoji parametro θ statistika, tačiau nėra pilnoji;

b) $T(X) = \mathbf{1}_{\{0\}}(X)$ yra parametro $\gamma = (1-\theta)^2$ NMD įvertinys;

c) $\hat{\theta} = \mathbf{1}_{\{-1\}}(X) + cX$, kai c bet kokia konstanta yra nepaslinktasis parametro θ įvertinys, o NMD įvertinys neegzistuoja.

I.2.80. A. d. X turi nupjautą Puasono skirstinį: $\mathbf{P}\{X = k\} = e^{-\lambda}\lambda^k/(k!(1 - e^{-\lambda}))$, $k = 1, 2, \ldots$. Įrodykite, kad parametro $\gamma = 1 - e^{-\lambda}$ NMD įvertinys $\hat{\gamma}$ yra toks: $\hat{\gamma} = 0$, kai X nelyginis, ir $\hat{\gamma} = 2$, kai X lyginis.

I.2.81. Tarkime, paprastosios imtys $\mathbf{X} = (X_1, ..., X_n)^T$, $\boldsymbol{Y} = (Y_1, ..., Y_m)^T$ gautos stebint n. a. d. $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$. Įrodykite, kad a) $\boldsymbol{T}(\mathbf{X}, \boldsymbol{Y}) = (\sum_i X_i, \sum_i Y_i, \sum_i X_i^2, \sum_i Y_i^2)^T$ yra pilnoji ir pakankamoji parametro $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1, \sigma_2)^T$ statistika; b) jeigu žinoma, kad $\mu_1 = \mu_2 = \mu$, tai $\boldsymbol{T}(\mathbf{X}, \boldsymbol{Y})$ yra pakankamoji parametro $(\mu, \sigma_1, \sigma_2)^T$ statistika, tačiau ji nėra pilna; c) jeigu žinoma, kad $\sigma_1 = \sigma_2 = \sigma$, tai $\boldsymbol{T}(\mathbf{X}, \boldsymbol{Y})$ yra pakankamoji parametro $(\mu_1, \mu_2, \sigma)^T$ statistika, tačiau ji nėra pilna; statistika $\boldsymbol{T}^*(\mathbf{X}, \boldsymbol{Y}) = (\sum_i X_i, \sum_i Y_i, \sum_i X_i^2 + \sum_i Y_i^2)^T$ yra pilnoji ir pakankamoji parametro $(\mu_1, \mu_2, \sigma)^T$ statistika.

I.2.82. Tegu paprastoji imtis $(X_i, Y_i)^T, i = 1, 2, ..., n$, gauta stebint a. v. $(X, Y)^T \sim N_2(\boldsymbol{\mu}, \Sigma)$, $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$, $\Sigma = [\sigma_{ij}]_{2\times 2}, \sigma_{11} = \sigma_1^2, \sigma_{22} = \sigma_2^2, \sigma_{12} = \rho\sigma_1\sigma_2$. Įrodykite, kad $\boldsymbol{T} = (\sum_i X_i, \sum_i Y_i, \sum_i X_i^2, \sum_i Y_i^2, \sum_i X_i Y_i)^T)$ yra pilnoji ir pakankamoji parametro $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)^T$ statistika.

I.2.83. Tarkime, paprastosios imtys $\mathbf{X} = (X_1, ..., X_n)^T$, $\boldsymbol{Y} = (Y_1, ..., Y_m)^T$ gautos stebint n. a. d. $X \sim N(\mu_1, 1)$, $Y \sim N(\mu_2, 1)$. Raskite parametro $\gamma = \mathbf{P}_{\mu_1, \mu_2}\{Y_1 < X_1\}$ NMD įvertinį.

I.2.84. Tarkime, paprastosios imtys $\mathbf{X} = (X_1, ..., X_n)^T$, $\boldsymbol{Y} = (Y_1, ..., Y_m)^T$ gautos stebint n. a. d. $X \sim G(\lambda_1, 1)$, $Y \sim G(\lambda_2, 1)$. Raskite parametro $\gamma = \mathbf{P}_{\lambda_1, \lambda_2}\{Y_1 < X_1\}$ NMD įvertinį.

I.2.85. Paprastoji imtis $(X_1, ..., X_n)^T$ gauta stebint a. d. $X \sim U(0,\theta), \theta > 0$. Tarkime, funkcija $\gamma(\theta)$ turi tolydžią išvestinę ir $\mathbf{E}(\gamma'(X_{(n)})) < \infty$. Raskite parametro $\gamma = \gamma(\theta)$ NMD įvertinį.

I.2.86. Tegu $Y_1, ..., Y_n$ yra n. a. d. ir $Y_i \sim N(\alpha + \beta x_i, \sigma^2)$; čia $x_1, ..., x_n$ – žinomos konstantos, $-\infty < \alpha, \beta < \infty, \sigma > 0$ – nežinomi parametrai. Įrodykite, kad $\boldsymbol{T} = (\sum_i Y_i, \sum_i Y_i x_i, \sum_i Y_i^2)$ yra pilnoji ir pakankamoji parametro $\boldsymbol{\theta} = (\alpha, \beta, \sigma^2)^T$ statistika.

I.2.87. Tarkime, n kartų atliekami kampų matavimai ir k-ojo matavimo metu kampo didumas yra $k\varphi$, $k = 1, 2, ..., n$. Sisteminės matavimų paklaidos lygios 0, o atsitiktinės paklaidos sumuojamos, t. y. k-ojo matavimo atsitiktinė paklaida yra $Z_1 + ... + Z_k$; čia $Z_1, ..., Z_n$ yra vienodai pasiskirstę nepriklausomi atsitiktiniai dydžiai ir $Z_i \sim N(0,\ \sigma^2)$. Raskite parametrų φ ir σ^2 NMD įvertinius.

I.2.88. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. X, kurio skirstinys priklauso šeimai $\mathcal{P} = \{p(i|\theta), \theta > 0\}$; čia $p(i|\theta) = \mathbf{P}\{X = i|\theta\} = a_i\theta^i/f(\theta)$, $i = c,\ c+1, ...$ Įrodykite: a) $T = X_1 + ... + X_n$ yra pilnoji ir pakankamoji šeimos $\mathcal{P}$ statistika, o jos skirstinys yra nusakomas tokio pavidalo tikimybėmis $\mathbf{P}\{T = k|\theta\} = b_k\theta^k/(f(\theta))^n$, $k = nc,\ nc+1, ...$; b) parametro θ^r NMD įvertinys yra $U_r(T) = (b_{T-r})/b_T$, kai $T \geq nc+r$, ir $U_r(T) = 0$, kai $T < nc+r$; c) įvertinio $U_r(T)$ dispersijos NMD įvertinys yra $(U_r(T))^2 - U_{2r}(T)$.

Taip pasiskirstę daugelis diskrečiųjų a. d., kurių galimų reikšmių skaičius yra begalinis (Puasono, geometrinis ir pan., įskaitant ir nupjautinius iš kairės).

I.2.89. Tegu $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. $X \sim N(\mu,\ \sigma^2)$, $0 < \mu < \infty$, $0 < \sigma < \infty$. Vertinamas parametras $\vartheta = \mu^2$. Apskaičiuokite ϑ įvertinio $\bar{X}^2$ poslinkį ir dispersiją. Raskite ϑ NMD įvertinį ir palyginkite jo dispersiją su įvertinio $\bar{X}^2$ dispersija.

I.2.90. Tegu $\mathbf{X} = (X_1, \ldots, X_m)^T$ ir $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ yra paprastosios atsitiktinės imtys, gautos stebint n. a. d. $X \sim N(\mu_x, \sigma_x^2)$ ir $Y \sim N(\mu_y,\ \sigma_y^2)$.

a) Raskite parametrų $\mu_x - \mu_y$ ir $(\sigma_x/\sigma_y)^r$, $r > 0$ NMD įvertinius, kai $\mu_x \in \mathbf{R}$, $\mu_y \in \mathbf{R}$, $\sigma_x > 0$ ir $\sigma_y > 0$.

b) Raskite σ_x^2 ir $(\mu_x - \mu_y)/\sigma_x$ NMD įvertinius, kai $\mu_x \in \mathbf{R}$, $\mu_y \in \mathbf{R}$, $\sigma_x = \sigma_y > 0$.

c) Raskite μ_x NMD įvertinį, kai $\mu_x = \mu_y \in \mathbf{R}$, $\sigma_x > 0$, $\sigma_y > 0$ ir $\sigma_x^2/\sigma_y^2 = \gamma$ yra žinomas.

d) Įrodykite, kad μ_x NMD įvertinys neegzistuoja, kai $\mu_x = \mu_y \in \mathbf{R}$, $\sigma_x > 0$, $\sigma_y > 0$.

e) Raskite $\mathbf{P}\{X_1 \leqslant Y_1\}$ NMD įvertinį, kai $\mu_x = \mu_y \in \mathbf{R}$, $\sigma_x > 0$, $\sigma_y > 0$.

f) Atlikite e) punkto užduotį, kai $\sigma_x = \sigma_y$.

I.2.91. Tegu $\mathbf{X} = (X_1, \ldots, X_m)^T$ ir $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ yra paprastosios atsitiktinės imtys, gautos stebint n. a. d. $X \sim \mathcal{E}(a_x, \theta_x)$ ir $Y \sim \mathcal{E}(a_y,\ \theta_y)$; čia $\theta_x,\ \theta_y > 0$ ir a_x, $a_y \in \mathbf{R}$.

a) Raskite $a_x - a_y$ ir θ_x/θ_y NMD įvertinius.

b) Raskite θ_x ir $(a_x - a_y)/\theta_x$ NMD įvertinius, kai $\theta_x = \theta_y$.

c) Įrodykite, kad a_x NMD įvertinys neegzistuoja, kai $a_x = a_y$.

I.2.3. Rao ir Kramerio nelygybė. Efektyvieji įvertiniai

I.2.92. Tarkime, imties $\mathbf{X}$ skirstinys priklauso nuo parametro $\boldsymbol{\theta} = (\theta_1, ..., \theta_k)^T$, $k > 1$, ir Fišerio informacinė matrica $\boldsymbol{I} = [I_{ij}]_{k\times k}$ neišsigimusi. Pažymėkime $\boldsymbol{I}^{-1} = [I^{ij}]_{k\times k}$

matricos $\boldsymbol{I}$ atvirkštinę matricą. Įrodykite, kad teisinga nelygybė $I^{ii} \geq 1/I_{ii}$.

I.2.93. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. X, kurio skirstinys priklauso binominių skirstinių šeimai $\mathcal{P} = \{B(1,\ p),\ 0 < p < 1\}$. Raskite funkcijų p, pq, p^2 nepaslinktųjų įvertinių dispersijų ribas remdamiesi Rao ir Kramerio nelygybe.

I.2.94. (**I.2.93** pratimo tęsinys). Raskite tikslesnę funkcijų p^2, pq nepaslinktųjų įvertinių dispersijų ribas remdamiesi patikslinta Rao ir Kramerio nelygybe.

I.2.95. (**I.2.93** pratimo tęsinys). Apskaičiuokite funkcijų p, pq, p^2 NMD įvertinių dispersijas ir palyginkite jas su gautomis **I.2.93** ir **I.2.94** pratimuose ribomis.

I.2.96. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. X, kurio skirstinys priklauso Puasono skirstinių šeimai $\mathcal{P} = \{\mathcal{P}(\lambda),\ 0 < \lambda < \infty\}$. Raskite funkcijų λ, λ^2, $\mathrm{e}^{-\lambda}$ nepaslinktųjų įvertinių dispersijų ribas pagal Rao ir Kramerio nelygybę.

I.2.97. (**I.2.96** pratimo tęsinys). Raskite tikslesnes funkcijų $\lambda^2, e^{-\lambda}$ nepaslinktųjų įvertinių dispersijų ribas remdamiesi patikslinta Rao ir Kramerio nelygybe.

I.2.98. (**I.2.96** pratimo tęsinys). Apskaičiuokite funkcijų λ, λ^2, $\mathrm{e}^{-\lambda}$ NMD įvertinių dispersijas ir palyginkite jas su gautomis **I.2.96** ir **I.2.97** pratimuose ribomis.

I.2.99. (**I.2.96** pratimo tęsinys). Įrodykite, kad parametro $\gamma = e^{-\lambda}$ įvertiniai $\tilde{\gamma} = e^{-\bar{X}}$ ir $\hat{\gamma} = (1 - 1/n)^{n\bar{X}}$ yra a) asimptotiškai efektyvūs; b) asimptotiškai efektyvūs pagal Rao.

I.2.100. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. $X \sim N(\mu,\ \sigma^2)$, $-\infty < \mu < \infty,\ 0 < \sigma < \infty\}$. Tarkime, parametras σ žinomas. a) Raskite parametrų μ ir μ^2 NMD įvertinius ir jų dispersijas. b) Raskite parametrų μ ir μ^2 nepaslinktųjų įvertinių dispersijų ribas Rao ir Kramerio nelygybėje. c) Raskite parametro μ^2 nepaslinktojo įvertinio dispersijos ribą patikslinus Rao ir Kramerio nelygybę.

I.2.101. (**I.2.100** pratimo tęsinys). Tarkime, kad parametras μ žinomas. a) Raskite parametrų σ^2 ir σ NMD įvertinius ir jų dispersijas. b) Raskite parametrų σ^2 ir σ nepaslinktųjų įvertinių dispersijų ribas Rao ir Kramerio nelygybėje.

I.2.102. (**I.2.100** pratimo tęsinys). Užrašykite Rao ir Kramerio nelygybę vektorinės funkcijos $\boldsymbol{\theta} = (\mu,\ \sigma^2)^T$ nepaslinktojo įvertinio kovariacijų matricai.

I.2.103. (**I.2.100** pratimo tęsinys). a) Raskite P-osios kritinės reikšmės $x_P = \sigma z_P + \mu$ NMD įvertinį ir jo dispersiją. b) Raskite parametro x_P nepaslinktojo įvertinio dispersijos ribą Rao ir Kramerio nelygybėje.

I.2.104. Atsitiktinio dydžio X skirstinys priklauso ekstremaliųjų skirstinių šeimai. Tankio funkcija

$$f(x|\theta) = \exp\{-(x - \theta) - \exp\{-(x - \theta)\}\},\ -\infty < \theta < \infty.$$

Pagal didumo n paprastąją imtį raskite parametro θ ir parametro $\alpha = \exp\{-\theta\}$ Fišerio informacijos kiekį.

I.2.105. Tarkime, kad $X_1, \ldots, X_n$ n. a. d. Tegu X_i tankio funkcija

$$f_i(x; \beta) = \frac{1}{\beta t_i} \exp(-x/(\beta t_i)), \quad x \geqslant 0;$$

čia $t_1, \ldots, t_n$ – žinomos konstantos.

a) Įrodykite, kad

$$\hat{\beta} = \frac{1}{n}\sum_{i=1}^{n} X_i/t_i$$

yra nepaslinktasis β įvertinys.

b) Apskaičiuokite nepaslinktojo β įvertinio dispersijos ribą Rao ir Kramerio nelygybėje. Ar įvertinio, nurodyto a) punkte, dispersija pasiekia šią ribą?

I.2.106. Tegu X skirstinys priklauso šeimai $\mathcal{P} = \{P_{\boldsymbol{\theta}},\ \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$. Raskite Fišerio informaciją $\boldsymbol{I}(\boldsymbol{\theta})$ pagal didumo n paprastąją imtį, kai $P_{\boldsymbol{\theta}}$ yra:

a) $N(\mu, \sigma^2)$ skirstinys, $\theta = \mu \in \mathcal{R}$; σ žinomas;
b) $N(\mu, \sigma^2)$ skirstinys, $\theta = \sigma^2 > 0$; μ žinomas;
c) $N(\mu, \sigma^2)$ skirstinys, $\theta = \sigma > 0$; μ žinomas;
d) $N(\sigma, \sigma^2)$ skirstinys, $\theta = \sigma > 0$;
e) $N(\mu, \sigma^2)$ skirstinys, $\boldsymbol{\theta} = (\mu, \sigma^2)^T$;
f) neigiamas binominis skirstinys $B^-(k,\ p)$, $\theta = p \in (0,\ 1)$; k žinomas;
g) gama skirstinys $G(\alpha, \gamma)$, $\boldsymbol{\theta}^T = (\alpha, \gamma) \in (0, \infty) \times (0, \infty)$;
h) beta skirstinys $B(\alpha, \beta)$, $\boldsymbol{\theta}^T = (\alpha, \beta) \in (0,\ \infty) \times (0,\ \infty)$.

I.2.107. (**I.2.106** pratimo tęsinys). Raskite $\boldsymbol{\theta}$ funkciją, kurios informacijos kiekis nepriklauso nuo θ, kai P_θ yra:

a) Puasono skirstinys $\mathcal{P}(\theta)$, $\theta > 0$;
b) binominis skirstinys $B(n,\ p)$, $\theta = p \in (0,\ 1)$; n žinomas;
c) gama skirstinys $G(\theta, \gamma)$, $\theta > 0$; γ žinomas.

I.2.108. (**I.2.106** pratimo tęsinys). Raskite Fišerio informacijos matricą, kai $P_{\boldsymbol{\theta}}$ yra:

a) Koši skirstinys $K(\mu, \sigma)$, $\mu \in \mathbf{R}$, $\sigma > 0$;
b) ekstremalių reikšmių skirstinys, kurio parametrai $\mu \in \mathbf{R}$, $\theta > 0$;
c) logistinis skirstinys $LG(\mu, \sigma)$, kurio parametrai $\mu \in \mathbf{R}$, $\sigma > 0$;
d) $F_r(\frac{x-\mu}{\sigma})$, čia F_r yra Stjudento skirstinio su žinomu laisvės laipsnių skaičiumi r pasiskirstymo funkcija, $\mu \in \mathbf{R}$, $\sigma > 0$.

I.2.109. Tegu $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. $X \sim U(0, \theta)$ su $\theta > 0$.

a) Įrodykite, kad Rao ir Kramerio teoremos sąlygos netenkinamos.

b) Įrodykite, kad parametro θ NMD įvertinio dispersija yra eilės $O(1/n^2)$, $n \to \infty$ (reikia pažymėti, kad reguliariu atveju, kai Rao ir Kramerio nelygybė galioja, dispersijos riba Rao ir Kramerio nelygybėje yra eilės $O(1/n)$).

I.2.110. Tegu X yra a. d., turintis ekstremalių reikšmių skirstinį su parametrais $\mu = 0$ ir $\theta > 0$. Raskite nurodytų parametrų NMD įvertinius ir kiekvienu atveju nustatykite, ar NMD įvertinio dispersija pasiekia Rao ir Kramerio nelygybėje nurodytą ribą: a) $\vartheta = \theta$; b) $\vartheta = \theta^r$, čia $r > 1$ žinomas.

I.2.111. Tegu $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. $X \sim N(\mu,\ \sigma^2)$, $-\infty < \mu < \infty$, σ – žinomas.

a) Raskite $\vartheta = \mathrm{e}^{t\mu}$ NMD įvertinį, kai $t \neq 0$ fiksuotas.

b) Nustatykite, ar a) punkte rasto įvertinio dispersija pasiekia Rao ir Kramerio nelygybėje nurodytą ribą.

c) Įrodykite, kad asimptotiškai ($n \to \infty$) tenkinama Rao ir Kramerio nelygybė.

I.2.112. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. $X \sim G(1/\lambda,\ \eta)$, $0 < \lambda, \eta < \infty$. Tarkime, parametras η žinomas. a) Raskite parametrų λ ir λ^2 NMD įvertinius ir jų dispersijas. b) Raskite parametrų λ ir λ^2 nepaslinktųjų įvertinių dispersijų ribas Rao ir Kramerio nelygybėje. c) Raskite parametro λ^2 nepaslinktųjų įvertinių dispersijos ribą patikslinus Rao ir Kramerio nelygybę.

I.2.113. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. $X \sim G(\lambda,\ \eta)$, $0 < \lambda, \eta < \infty$. Tarkime, parametras η žinomas. a) Raskite parametrų λ ir λ^2 NMD įvertinius ir jų dispersijas. b) Raskite parametrų λ ir λ^2 nepaslinktųjų įvertinių dispersijų ribas Rao ir Kramerio nelygybėje.

I.2.114. Tegu $\mathbf{X} = (X_1, \dots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. $X \sim N(\mu,\ 1)$, $\mu \in \mathbf{R}$. Tegu $\vartheta = \mathbf{P}\{X_1 \le c\}$, čia c – fiksuota konstanta. Nagrinėjami tokie ϑ įvertiniai: $T_{1n} = \hat{F}_n(c)$, čia $\hat{F}_n$ yra empirinė pasiskirstymo funkcija, ir $T_{2n} = \Phi(c - \bar{X})$, čia Φ yra standartinio normaliojo skirstinio pasiskirstymo funkcija. Raskite įvertinio T_{1n} ASE T_{2n} atžvilgiu.

I.2.115. Tegu $\mathbf{X} = (X_1, \dots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. $X \sim N(0,\ \sigma^2)$, čia $\sigma > 0$ nežinomas. Vertinama $\vartheta = \sigma$.

Raskite įvertinio $\hat{\sigma}_1 = \sqrt{\pi/2}\sum_{i=1}^n |X_i|/n$ ASE įvertinio $\hat{\sigma}_2 = (\sum_{i=1}^n X_i^2/n)^{1/2}$ atžvilgiu.

I.2.116. Tegu $\mathbf{X} = (X_1, \dots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. X, kurio $\mathbf{E}X = \mu$, $\boldsymbol{V}X = 1$ ir $\mathbf{E}X^4 < \infty$. Tegu $T_{1n} = n^{-1}\sum_{i=1}^n X_i^2 - 1$ ir $T_{2n} = \bar{X}^2 - n^{-1}$ yra $\vartheta = \mu^2$ įvertiniai.

a) Raskite įvertinio T_{1n} ASE atžvilgiu įvertinio T_{2n}.

b) Įrodykite, kad ASE ≤ 1, jeigu $X_i - \mu$ pasiskirstymo funkcija yra simetrinė 0 atžvilgiu.

c) Raskite skirstinį, kurio ASE > 1.

I.2.117. Tegu $\mathbf{X} = (X_1, \dots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. $X \sim B(1,\ p)$, $0 < p < 1$. Vertinamas parametras p. Tegu a ir b yra teigiamos konstantos. Raskite įvertinio $(a + n\bar{X})/(a + b + n)$ ASE įvertinio $\bar{X}$ atžvilgiu.

I.2.118. Tegu $\mathbf{X} = (X_1, \dots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. $X \sim U(0,\ \theta)$, $0 < \theta < \infty$. Nagrinėjami tokie θ įverčiai: $T_{1n} = (n+1)X_{(n)}/n$ ir $T_{2n} = X_{(n)}$. Raskite poslinkius $b_{T_{jn}}(\theta)$, $j = 1, 2$ ir įvertinio T_{1n} ASE įvertinio T_{2n} atžvilgiu.

I.2.119. Tegu paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $X \sim N(0, \sigma^2)$. Dispersijos įvertiniu imkime $n\bar{X}^2$ ir s^2. Koks įvertinio $n\bar{X}^2$ efektyvumas įvertinio s^2 atžvilgiu.

I.2.120. Paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. X, kurio tankio funkcija

$$f(x|\mu) = \frac{\lambda}{2} e{-}\lambda|x - \mu|, \quad -\infty < x, \mu < \infty,$$

parametras $\lambda > 0$ žinomas. Raskite parametro μ įvertinio $\hat{\mu}_2 = \bar{X}$ ASE empirinės medianos $\hat{\mu}_1 = \hat{x}_{1/2}$ atžvilgiu.

I.2.121. Vertinant atsitiktinio dydžio $X \sim N(\mu,\ \sigma^2)$ parametrą μ, gautos trys paprastosios atsitiktinės nepriklausomos imtys, iš kurių gauti įverčiai: $\bar{X}_1 = 17,24$ (didumo $n_1 = 5$ imtis); $\bar{X}_2 = 16,81$ (didumo $n_2 = 10$ imtis); $\bar{X}_3 = 17,22$ (didumo $n_3 = 100$ imtis). Raskite parametro μ NMD įvertinio realizacijos reikšmę naudodamiesi visais matavimais.

I.2.4. Įvertinių radimo metodai

I.2.122. Momentų metodu raskite parametrų α ir β įvertinius pagal didumo n paprastąją atsitiktinę imtį a. d. X, kurio skirstinys yra $N(0,\ 1)$ su tikimybe β ir $N(\alpha,\ 1)$ su tikimybe $1-\beta$.

I.2.123. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. X, kurio skirstinys priklauso šeimai $\mathcal{P} = U(\theta_1,\ \theta_2),\ -\infty < \theta_1 < \theta_2 < \infty\}$. a) Raskite parametrų $\mu = (\theta_1+\theta_2)/2$ ir $\sigma = \theta_2 - \theta_1$ NMD įvertinius. b) Palyginkite jų dispersijas su momentų metodo įvertinių dispersijomis. c) Raskite parametrų μ ir σ DT įvertinius.

I.2.124. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra a. d. X paprastoji imtis. Raskite parametrų DT įvertinius ir palyginkite juos su tų pačių parametrų NMD įvertiniais, kai a. d. X skirstinys priklauso a) normaliųjų skirstinių šeimai $\mathcal{P} = \{N(\mu,\ \sigma^2),\ -\infty < \mu < \infty,\ \sigma > 0\}$;
b) gama skirstinių šeimai $\mathcal{P} = \{G(\lambda,\ \eta),\ 0 < \lambda < \infty\},\ \eta > 0$ – žinoma konstanta;
c) tolygiųjų skirstinių šeimai $\mathcal{P} = \{U(0,\ \theta),\ 0 < \theta < \infty\}$.

I.2.125. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. X, kurio skirstinys priklauso šeimai $\mathcal{P} = \{p(x|\theta),\ 0 < \theta < \infty\}$; čia

$$p(i|\theta) = \mathbf{P}\{X = i|\theta\} = \frac{a_i\theta^i}{f(\theta)},\ i = 0,\ 1,\ 2, \dots .$$

Įrodykite, kad parametro θ DT įvertinys randamas iš lygties

$$\theta f'(\theta)/f(\theta) = \bar{X},$$

kuri sutampa su lygtimi, gaunama θ įvertinio ieškant momentų metodu.

I.2.126. (**I.2.125** pratimo tęsinys). Užrašykite lygtis, iš kurių randami DT parametrų įvertiniai, kai skirstinys yra Puasono, binominis $B(1,\ p)$, logaritminis, taip pat nupjautinis Puasono (praleista reikšmė 0).

I.2.127. (**I.2.125** pratimo tęsinys). Palyginkite binominio ir Puasono skirstinių parametrų p^2 ir λ^2 DT ir NMD įvertinių kvadratinės rizikos funkcijas.

I.2.128. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. X, kurio skirstinys priklauso tolygiųjų skirstinių šeimai: a) $\mathcal{P}_1 = \{U(\theta,\ 2\theta),\ 0 < \theta < \infty\}$; b) $\mathcal{P}_2 = \{U(\theta - 1/2,\ \theta + 1/2),\ -\infty < \theta < \infty\}$. Įrodykite, kad šeimos $\mathcal{P}_1$ parametro θ DT įvertinys yra $\hat{\theta} = X_{(n)}/2$, o šeimos $\mathcal{P}_2$ parametro θ DT įvertinys nėra vienareikšmis, – jis gali būti bet kuri statistika, įgyjanti reikšmes iš intervalo $(X_{(n)} - 1/2,\ X_{(1)} + 1/2)$.

I.2.129. (**I.2.128** pratimo tęsinys). Palyginkite DT įvertinių dispersijas su įvertinių, gautų momentų metodu, dispersijomis.

I.2.130. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. X, kurio skirstinys priklauso Laplaso skirstinių šeimai $\mathcal{P} = \{f(x|\theta),\ -\infty < \theta < \infty\}$; čia $f(x|\theta) = \exp\{-|x-\theta|\}/2,\ -\infty < x < \infty$. Įrodykite, kad parametro DT įvertinys yra empirinė mediana $\hat{x}_{0,5}$ ir

$$\sqrt{n}(\hat{x}_{0.5} - \theta) \xrightarrow{d} Z \sim N(0,\ 1/4),\ n \to \infty.$$

I.2.131. Tegu $\mathbf{X} = \{(X_{1i}, ..., X_{ki})^T,\ i = 1, ..., n\}$ yra paprastoji imtis a. v. $(X_1, ..., X_k)^T$, kurio skirstinys priklauso polinominių skirstinių šeimai $\mathcal{P} = \{\mathcal{P}_k = (1,\ \boldsymbol{\pi})\}$; čia $\boldsymbol{\pi} = (\pi_1, ..., \pi_k)^T$ yra k-mačiai vektoriai, kurių $0 < \pi_i < 1,\ \pi_1 + \pi_2 + ... + \pi_k = 1$. Raskite parametrų $\pi_1, ..., \pi_k$ DT įvertinius, jų dispersijas ir kovariacijas.

I.2.132. (**I.2.131** pratimo tęsinys). Raskite parametro α DT įvertinį, kai $k = 3$, $\pi_1 = (1+\alpha)/2$, $\pi_2 = \pi_3 = (1-\alpha)/4$, ir jo asimptotinį $(n \to \infty)$ skirstinį.

I.2.133. (**I.2.131** pratimo tęsinys). Dviejų berniukų, dviejų mergaičių ir mišrių dvynukų, kai pirmasis gimė berniukas ir pirmoji gimė mergaitė, tikimybės atitinkamai yra $\pi_1 = p^2$, $\pi_2 = (1-p)^2 = q^2$, $\pi_3 = \alpha(1-p^2-q^2)$ ir $\pi_4 = (1-\alpha)(1-p^2-q^2)$. Raskite parametrų p ir α DT įvertinius.

I.2.134. Realizuojant $n = 8000$ kartų nepriklausomus eksperimentus, kurių metu gali įvykti vienas iš trijų nesutaikomų įvykių A, B ir C su tikimybėmis $1/2 - 2\alpha$, $1/2 + \alpha$ ir α, $0 \le \alpha \le 1/4$, užregistruoti šių įvykių dažniai: 2 014, 5 012 ir 974. Raskite parametro α didžiausiojo tikėtinumo įvertį.

I.2.135. Tegu $(X_i,\ Y_i)^T$, $i = 1,...,n$, yra imtis a. v. $(X,\ Y)^T$, kurio skirstinys priklauso dvimačių normaliųjų skirstinių šeimai $\mathcal{P} = N_2(\boldsymbol{\mu},\ \boldsymbol{\Sigma})$; čia $\boldsymbol{\mu} = (\mu_1,\ \mu_2)^T$, $-\infty < \mu_1,\ \mu_2 < \infty$, $\boldsymbol{\Sigma} = [\sigma_{ij}]$, $\sigma_{11} = \sigma_1^2$, $\sigma_{22} = \sigma_2^2$, $\sigma_{12} = \sigma_{21} = \rho\sigma_1\sigma_2$, $0 < \sigma_1,\ \sigma_2 < \infty$, $|\rho| < 1$. Raskite parametrų DT įvertinius. Apskaičiuokite matricos, atvirkštinės informacinei matricai, elementus.

I.2.136. Asimptotiškai $(n \to \infty)$ palyginkite dispersijas ekstremaliųjų reikšmių skirstinio parametrų įvertinių, gautų pagal didumo n paprastąją imtį DT ir momentų metodais.

I.2.137. Asimptotiškai $(n \to \infty)$ palyginkite dispersijas gama skirstinio parametrų įvertinių, gautų pagal didumo n paprastąją imtį DT ir momentų metodais.

I.2.138. Asimptotiškai $(n \to \infty)$ palyginkite dispersijas neigiamojo binominio skirstinio parametrų įvertinių, gautų pagal didumo n paprastąją imtį DT ir momentų metodais.

I.2.139. Asimptotiškai $(n \to \infty)$ palyginkite dispersijas Koši skirstinio parametrų įvertinių, gautų pagal didumo n paprastąją imtį: a) DT metodu, b) grindžiamų statistikomis $\hat{x}_{0,5}$, $\hat{x}_{0,25}$, $\hat{x}_{0,75}$.

I.2.140. Tegu T_n yra pagrįstasis ir asimptotiškai normalusis parametro θ įvertinys, t. y.

$$\sqrt{n}(T_n - \theta) \xrightarrow{d} X \sim N(0,\ \sigma^2(\theta)),\ n \to \infty.$$

Įrodykite, kad įvertinys

$$T_n' = \begin{cases} \alpha T_n, & \text{kai} \quad |T_n| \le n^{-1/4}, \\ T_n, & \text{kai} \quad T_n > n^{-1/4}, \end{cases}$$

taip pat asimptotiškai $(n \to \infty)$ normalusis ir dispersija lygi $\alpha^2\sigma^2(\theta)$, kai $\theta = 0$, ir $\sigma^2(\theta)$, kai $\theta \ne 0$. Taigi įvertinio T_n' dispersija gali būti mažesnė už T_n dispersiją, kai $\theta = 0$, ir lygi T_n dispersijai, kai $\theta \ne 0$.

I.2.141. Tegu $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. $X \sim N(0,\ \sigma^2)$, $\sigma > 0$.

a) Įrodykite, kad $\mathbf{E}(|X_i|) = \sigma\sqrt{2/\pi}$.

b) Naudodamiesi a) punkte gautu rezultatu, momentų metodu raskite σ įvertinį $\hat{\sigma}_n$. Raskite $\sqrt{n}(\hat{\sigma}_n - \sigma)$ asimptotinį skirstinį.

c) Kitas momentų metodu, kai naudojamas antrasis momentas, gautas σ įvertinys yra

$$\tilde{\sigma}_n = \left(\frac{1}{n} \sum_{i=1}^{n} X_i^2 \right)^{1/2}.$$

Raskite $\sqrt{n}(\tilde{\sigma}_n - \sigma)$ asimptotinį skirstinį ir palyginkite jį su b) punkto rezultatu.

I.2.142. Tegu $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. $X \sim \mathcal{E}(\mu,\ 1)$, $\mu \in \mathbf{R}$.

a) Įrodykite, kad $X_{(1)} = \min(X_1, \ldots, X_n)$ yra pakankamoji μ statistika.

b) Įrodykite, kad $X_{(1)} \xrightarrow{P} \mu$, kai $n \to \infty$.

I.2.143. Tegu $X_1, \ldots, X_n$ yra nepriklausomi vienodai pasiskirstę a. d., kurių tankio funkcija

$$f(x; \theta_1, \theta_2) = \begin{cases} a(\theta_1, \theta_2) h(x), & \text{kai } \theta_1 \leqslant x \leqslant \theta_2, \\ 0, & \text{kitais atvejais;} \end{cases}$$

čia $h(x) > 0$ – žinoma tolydžioji funkcija, apibrėžta realiųjų skaičių tiesėje.

Įrodykite, kad θ_1 ir θ_2 DT įvertiniai yra atitinkamai $X_{(1)}$ ir $X_{(n)}$.

I.2.144. Tegu $(X_1, Y_1)^T, \ldots, (X_n, Y_n)^T$ yra nepriklausomos vienodai pasiskirsčiusios normaliųjų a. d. poros; čia $X_i, Y_i \sim N(\mu_i, \sigma^2)$.

a) Raskite parametrų $\mu_1, \ldots, \mu_n$ ir σ^2 DT įvertinius.

b) Įrodykite, kad parametro σ^2 DT įvertinys nėra pagrįstasis. Ar šis rezultatas prieštarauja teorijai apie DT įvertinių pagrįstumą? Kodėl?

c) Stebimi tik $Z_1, \ldots, Z_n$; čia $Z_i = X_i - Y_i$. Raskite σ^2 DT įvertinį, gautą naudojant $Z_1, \ldots, Z_n$, ir įrodykite, kad jis yra pagrįstasis.

I.2.145. Tegu $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ yra nepriklausomi a. d., turintys eksponentinius skirstinius. Tegu X_i tankio funkcija

$$f_i(x) = \lambda_i \theta \exp(-\lambda_i \theta x), \quad x \geqslant 0,$$

o Y_i tankio funkcija

$$g_i(x) = \lambda_i \exp(-\lambda_i x), \quad x \geqslant 0;$$

čia $\lambda_1, \ldots, \lambda_n$ ir θ yra nežinomi parametrai.

a) Įrodykite, kad parametro θ DT įvertinys tenkina lygtį

$$\frac{n}{\hat{\theta}} - 2 \sum_{i=1}^{n} \frac{R_i}{1 + \hat{\theta} R_i} = 0;$$

čia $R_i = X_i / Y_i$.

b) Įrodykite, kad R_i tankio funkcija yra

$$f_R(x; \theta) = \theta(1 + \theta x)^{-2}, \quad x \geqslant 0,$$

o θ DT įvertinys, gautas naudojant $R_1, \ldots, R_n$, sutampa su pateikiamu a) punkte.

c) Tegu $\hat{\theta}_n$ yra DT įvertinys, nurodytas b) punkte. Raskite $\sqrt{n}(\hat{\theta}_n - \theta)$ ribinį skirstinį.

d) Lentelėje pateikiami (X_i, Y_i), $i = 1, \ldots, n$ duomenys. Apskaičiuokite θ DT įvertį artutiniu metodu. Parinkite tinkamą pradinį artinį ir pagrįskite pasirinkimą.

x_i	y_i	x_i	y_i	x_i	y_i	x_i	y_i
0,7	3,8	20,2	2,8	1,1	2,8	15,2	8,8
11,3	4,6	0,3	1,9	1,9	3,2	0,2	7,6
2,1	2,1	0,9	1,4	0,5	8,5	0,7	1,3
30,7	5,6	0,7	0,4	0,8	14,5	0,4	2,2
4,6	10,3	2,3	0,9	1,2	14,4	2,3	4,0

e) Pateikite DT įverčio, gauto d) punkte, standartinės paklaidos pagrįstąjį įvertį.

I.2.146. Tegu $X_1, \ldots, X_n$ yra nepriklausomi vienodai pasiskirstę neneigiami a. d. ir gedimų intensyvumo funkcija

$$\lambda(x) = \begin{cases} \lambda_1, & \text{kai } x \leqslant x_0, \\ \lambda_2, & \text{kai } x > x_0; \end{cases}$$

čia λ_1 ir λ_2 yra nežinomi parametrai, o x_0 – žinoma konstanta.

X_i tankio funkcija

$$f(x; \lambda_1, \lambda_2) = \begin{cases} \lambda_1 \exp(-\lambda_1 x), & \text{kai } x \leqslant x_0, \\ \lambda_2 \exp(-\lambda_2(x - x_0) - \lambda_1 x_0), & \text{kai } x > x_0. \end{cases}$$

Raskite parametrų λ_1 ir λ_2 DT įvertinius ir jų ribinį bendrą skirstinį.

I.2.147. Tegu $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. X, kurio skirstinys priklauso šeimai $\mathcal{P} = \{P_{\boldsymbol{\theta}},\ \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$. Momentų metodu raskite parametrų įvertinius, kai $P_{\boldsymbol{\theta}}$ yra:

a) gama skirstinys $G(\alpha, \gamma)$, $\boldsymbol{\theta} = (\alpha, \gamma)^T$, $\alpha > 0$, $\gamma > 0$;

b) eksponentinis skirstinys $\mathcal{E}(a, \gamma)$, $\boldsymbol{\theta} = (a, \gamma)^T$, $a \in \mathbf{R}$, $\theta > 0$;

c) beta skirstinys $Be(\alpha, \beta)$, $\boldsymbol{\theta} = (\alpha, \beta)^T$, $\alpha > 0$, $\beta > 0$;

d) lognormalusis skirstinys $LN(\mu, \sigma)$, $\boldsymbol{\theta} = (\mu, \sigma)^T$, $\mu \in \mathbf{R}$, $\sigma > 0$;

e) tolygusis skirstinys $U\left(\theta - \frac{1}{2},\ \theta + \frac{1}{2}\right)$, $\theta \in \mathbf{R}$;

f) neigiamas binominis skirstinys $B^-(n,\ p)$, $\boldsymbol{\theta} = (p, n)^T$, $p \in (0,\ 1)$, $n = 1,\ 2, \ldots$;

g) logaritminis skirstinys, kurio parametras $\theta = p \in (0,\ 1)$;

h) chi kvadrato skirstinys $\chi^2(k)$, $\theta = k$, $k = 1,\ 2, \ldots$.

I.2.148. Tegu $\mathbf{X}$ yra imtis iš skirstinio, kurio tankio funkcija yra $f_{\boldsymbol{\theta}}$, o $T(\mathbf{X})$ – pakankamoji $\boldsymbol{\theta}$ statistika. Įrodykite: jeigu egzistuoja DT įvertinys, tai jis yra T funkcija.

I.2.149. Tegu $X_1, \ldots, X_n$ yra nepriklausomi vienodai pasiskirstę a. d., kurių tankio funkcija yra $f_{\boldsymbol{\theta}}$ σ baigtinio mato ν atžvilgiu. Raskite parametro $\boldsymbol{\theta}$ DT įvertinį tokiais atvejais:

a) $f_\theta(x) = 1/\theta$, kai $x = 1, 2, ..., \theta$, θ yra sveikasis skaičius tarp 1 ir θ_0;

b) $f_\theta(x) = \mathrm{e}^{-(x-\theta)}$, $\theta \le x < \infty$, $\theta > 0$;

c) $f_\theta(x) = \theta(1-x)^{\theta-1}$, $0 < x < 1$, $\theta > 1$;

d) $f_\theta(x) = \frac{\theta}{1-\theta} x^{(2\theta-1)/(1-\theta)}$, $0 < x < \infty$, $\theta \in \left(\frac{1}{2},\ 1\right)$;

e) $f_\theta(x) = 2^{-1}\mathrm{e}^{-|x-\theta|}$, $x \in \mathbf{R}$, $\theta > 0$;

f) $f_\theta(x) = \theta x^{-2}$, $\theta < x < \infty$, $\theta > 0$;

g) $f_\theta(x)$ yra tankio funkcija skirstinio $N(\theta, \theta^2)$, $\theta \in \mathbf{R}$;

h) $f_{\boldsymbol{\theta}}(x)$ yra tankio funkcija eksponentinio skirstinio $\mathcal{E}(\mu,\ \sigma)$, $\boldsymbol{\theta}^T = (\mu, \sigma) \in \mathbf{R} \times (0, \infty)$;

i) $f_{\boldsymbol{\theta}}(x)$ yra tankio funkcija lognormalaus skirstinio $LN(\mu,\ \sigma)$, $\boldsymbol{\theta}^T = (\mu, \sigma) \in \mathbf{R} \times (0, \infty)$;

j) $f_\theta(x) = 1,\ x \in (0,\ 1)$, kai $\theta = 0$, ir $f_\theta(x) = (2\sqrt{x})^{-1},\ x \in (0,\ 1)$, kai $\theta = 1$;
k) $f_\theta(x) = \beta^{-\alpha}\alpha x^{\alpha-1},\ 0 < x < \beta,\ \alpha > 0,\ \beta > 0$;
l) $f_\theta(x) = C_\theta^x p^x (1-p)^{\theta - x},\ x = 0, 1, \ldots, \theta,\ \theta = 1, 2, \ldots$; čia $p \in (0,\ 1)$ yra žinomas.

I.2.150. Tegu $(Y_1, Z_1)^T, \ldots, (Y_n, Z_n)^T$ yra nepriklausomi vienodai pasiskirstę a. v., kurių tankio funkcija

$$f(y, z|\lambda,\ \mu) = \lambda^{-1}\mu^{-1}\mathrm{e}^{-y/\lambda}\mathrm{e}^{-z/\mu},\ 0 < y, z < \infty,$$

čia $\lambda > 0$ ir $\mu > 0$.

a) Raskite $(\lambda, \mu)^T$ DT įvertinį.

b) Stebima tik $X_i = \min(Y_i, Z_i)$ ir $\Delta_i = 1$, kai $X_i = Y_i$, ir $\Delta_i = 0$, kai $X_i = Z_i$. Raskite $(\lambda, \mu)^T$ DT įvertinį.

c) Raskite įvertinių asimptotinius skirstinius.

I.2.151. Tegu $X_1, \ldots, X_n$ yra nepriklausomi vienodai pasiskirstę diskretieji a. d., kurių skirstinys nusakytas tikimybėmis

$$\mathbf{P}\{X_1 = x\} = [x!(1 - \mathrm{e}^{-\theta})]^{-1}\theta^x \mathrm{e}^{-\theta},\ x = 1, 2, ...;$$

čia $\theta > 0$. Įrodykite, kad tikėtinumo lygtis turi vienintelę šaknį, kai $\bar{x} > 1$. Ar ši šaknis yra θ DT įvertinys?

I.2.152. Tegu $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. $X \sim \mathcal{E}(a,\ \theta)$, parametrai a ir θ nežinomi. Raskite parametrų a ir θ DT įvertinių ASE jų NMD įvertinių atžvilgiu.

I.2.153. Tegu $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. X, kurio skirstinys yra Pareto su parametrais a ir θ.

a) Raskite (a, θ) DT įvertinį.

b) Raskite parametro a DT įvertinio ASE NMD įvertinio atžvilgiu.

I.2.154. Tegu $\mathbf{X}_1, \ldots, \mathbf{X}_n$ yra nepriklausomi vienodai pasiskirstę k-mačiai a. v., turintys $N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ skirstinį su nežinomais $\boldsymbol{\mu}$ ir $\boldsymbol{\Sigma}$. Raskite $\boldsymbol{\mu}$ ir $\boldsymbol{\Sigma}$ DT įvertinius ir jų asimptotinius skirstinius.

I.2.155. Tegu $X_1, \ldots, X_n$ ir $Y_1, \ldots, Y_n$ yra nepriklausomi a. d., turintys atitinkamai $N(\mu, \sigma^2)$ ir $N(\mu, \tau^2)$ skirstinius su nežinomu $\boldsymbol{\theta} = (\mu, \sigma^2, \tau^2)^T$. Raskite $\boldsymbol{\theta}$ DT įvertinį ir įrodykite, kad jis asimptotiškai efektyvusis.

I.2.156. Tegu $X \sim B(n, p)$, $0 < p < 1$, p žinomas. Raskite parametro n DT įvertinį.

I.2.157. Tegu $X \sim H(N, M, n)$. Raskite parametro M DT įvertinį, kai N ir n žinomi.

I.2.5. Intervaliniai įvertiniai

I.2.158. Tarkime, paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $X \sim N(\mu, \sigma^2)$, $-\infty < \mu < \infty, \sigma > 0$ žinomas. Raskite parametro μ lygmens Q pasikliovimo intervalą.

I.2.159. (**I.2.158** pratimo tęsinys). Tarkime, $\sigma > 0$ nežinomas parametras, o parametras μ žinomas. Raskite parametrų σ^2 ir σ lygmens Q pasikliovimo intervalus.

I.2.160. (**I.2.158** pratimo tęsinys). Tarkime, abu parametrai μ ir $\sigma > 0$ nežinomi. Raskite parametrų μ ir σ^2 lygmens Q pasikliovimo intervalus.

I.2.161. (**I.2.160** pratimo tęsinys). Raskite parametro σ^2 lygmens Q pasikliovimo intervalą, kurio vidutinis santykinis ilgis $b = \mathbf{E}_\sigma(\overline{\sigma^2} - \underline{\sigma^2})/\sigma^2$ būtų mažiausias. Palyginkite jo vidutinį santykinį ilgį su simetriško intervalo vidutiniu santykiniu ilgiu, kai $Q = 0,95$, ir $n = 10; 20; 50; 100$.

I.2.162. (**I.2.160** pratimo tęsinys). Įrodykite, kad

$$C(\bar{X}, s^2) = \{(\mu, \sigma^2) : \sigma^2 > \frac{n(\bar{X} - \mu)^2}{\chi^2_{\alpha_1}(1)}, \frac{s^2(n-1)}{\chi^2_{\alpha_2}(n-1)} < \sigma^2 < \frac{s^2(n-1)}{\chi^2_{1-\alpha_2}(n-1)}\}$$

yra parametro $(\mu, \sigma^2)^T$ lygmens $Q = (1 - \alpha_1)(1 - 2\alpha_2)$ pasikliovimo sritis.

I.2.163. Pagal normaliojo a. d. didumo $n = 100$ paprastąją imtį gauti tokie pasikliovimo intervalo rėžiai: $\underline{\mu} = 1,25$, $\overline{\mu} = 2,05$. Koks to intervalo pasikliovimo lygmuo, jei $\sigma^2 = 4$?

I.2.164. Kokio didumo turi būti atsitiktinio dydžio $X \sim N(\mu, 1)$ imtis, kad parametro μ pasikliovimo intervalo, kai pasikliovimo lygmuo $Q = 0,95$, ilgis būtų ne didesnis kaip 0,1?

I.2.165. Pagal didumo $n = 25$ paprastąją imtį, gautą stebint a. d. $X \sim N(\mu, \sigma^2)$, gautos parametrų μ ir σ^2 NMD įvertinių realizacijos $\hat{\mu} = \bar{X} = 6,334$, $\hat{\sigma^2} = s^2 = 0,012$. Raskite parametrų μ ir σ lygmens $Q = 0,95$ pasikliovimo intervalų realizacijas.

I.2.166. Tarkime, paprastosios imtys $X_1, ..., X_m$ ir $Y_1, ..., Y_n$ gautos stebint n. a. d. $X \sim N(\mu_1, \sigma_1^2)$ ir $Y \sim N(\mu_2, \sigma_2^2)$, $-\infty < \mu_1, \mu_2 < \infty, \sigma_1, \sigma_2 > 0$. a) Raskite parametro $\beta = \sigma_1^2/\sigma_2^2$ lygmens Q pasikliovimo intervalą. b) Tarkime, žinoma, kad $\sigma_1^2/\sigma_2^2 = k$. Raskite parametro $\sigma^2 = \sigma_2^2$ ir parametro $\theta = \mu_1 - \mu_2$ lygmens Q pasikliovimo intervalus.

I.2.167. (**I.2.166.** pratimo tęsinys). Tarkime, apie dispersijas nieko nežinome, o imtys pakankamai didelės. Raskite asimptotinį parametro $\theta = \mu_1 - \mu_2$ pasikliovimo intervalą.

I.2.168. Tarkime, paprastoji imtis $(X_i, Y_i)^T, i = 1, ..., n$, gauta stebint a. v.

$$(X, Y)^T \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad \boldsymbol{\mu} = (\mu_1, \mu_2)^T, \quad -\infty < \mu_1,\ \mu_2 < \infty,$$

$$\boldsymbol{\Sigma} = [\sigma_{ij}]_{2\times 2}, \quad \sigma_{11} = \sigma_1^2, \quad \sigma_{22} = \sigma_2^2, \quad \sigma_{12} = \sigma_1\sigma_2\rho.$$

Raskite parametro $\theta = \mu_1 - \mu_2$ lygmens Q pasikliovimo intervalą.

I.2.169. (**I.2.168** pratimo tęsinys). Raskite koreliacijos koeficiento ρ lygmens Q pasikliovimo intervalą.

I.2.170. (**I.2.168** pratimo tęsinys). Raskite koreliacijos koeficiento ρ aproksimacinį lygmens Q pasikliovimo intervalą.

I.2.171. Tarkime, paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d., turintį Relėjaus skirstinį su parametru $\sigma > 0$ (žr. 1 priedo 1.P.2 lentelę). Raskite parametro σ lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą.

I.2.172. Tarkime, paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d., turintį Maksvelo skirstinį su parametru $\sigma > 0$ (žr. 1 priedo 1.P.2 lentelę). Raskite parametro σ^2 lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą.

I.2.173. Tarkime, imties $X_1, ..., X_n$ nariai yra n. a. d. ir $X_i \sim G(\lambda, \eta_i)$, $i = 1, ..., n$, $\lambda > 0$. Parametrai $\eta_1, ..., \eta_n$ žinomi. Raskite parametro λ lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą.

I.2.174. Tarkime, paprastoji imtis $X_1, ..., X_n$, gauta stebint a. d. turintį ekstremalių reikšmių skirstinį, kurio tankio funkcija

$$f(x|\mu) = e^{x-\mu} \exp\{-e^{x-\mu}\}, \quad -\infty < \mu < \infty.$$

Raskite parametro μ lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą.

I.2.175. Tarkime, paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $X \sim \mathcal{P}(\lambda)$. Raskite parametro λ lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą.

I.2.176. Tarkime, paprastoji imtis $X_1, ..., X_n$ gauta stebint Bernulio a. d. $X \sim B(1, p), 0 < p < 1$. Raskite parametro p lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą.

I.2.177. Tegu paprastosios imtys $X_1, ..., X_n$ ir $Y_1, ..., Y_n$ gautos stebint n. a. d. $X \sim P(\lambda_1)$ ir $Y \sim P(\lambda_2)$, $0 < \lambda_1, \lambda_2 < \infty$. Raskite parametro $\beta = \lambda_1/\lambda_2$ lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą.

I.2.178. Tarkime, imties $X_1, ..., X_n$ koordinatės yra n. a. d. ir $X_i \sim B^-(m_i, p), i = 1, ..., n, 0 < p < 1$. Parametrai $m_1, ..., m_n$ žinomi. Raskite parametro p lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą.

I.2.179. Turime didumo N gaminių partiją, kurioje yra nežinomas skaičius M defektinių. Atsitiktinai negrąžinant atrenkama n gaminių. Tegu $X_i = 1$, jei i-sis atrinktas gaminys defektinis, ir $X_i = 0$ priešingu atveju. Gauname imtį $X_1, ..., X_n$, kuri nėra paprastoji. Tegu N ir n žinomi, o $M = 0, 1, ..., N$ yra nežinomas parametras. Raskite parametro M lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą.

I.2.180. (**I.2.179** pratimo tęsinys). Tarkime, partijos didumas $N = 300$. Patikrinus $n = 50$ gaminių rasti 6 defektiniai. Raskite parametro M taškinį ir intervalinį pasikliovimo lygmens $Q = 0,9$ įverčius.

I.2.181. Tegu $Y_1, ..., Y_n$ yra n. a. d. ir $Y_i \sim N(\beta_0 + \beta_1(x_i - \bar{x}), \sigma^2)$, $i = 1, ..., n$; čia $x_1, ..., x_n, \bar{x} = \sum_i x_i/n$ – žinomos konstantos, o $-\infty < \beta_0, \beta_1 < \infty, \sigma > 0$ nežinomi parametrai.

a) Raskite parametro σ^2 lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą.

b) Raskite parametro $\theta = a\beta_0 + b\beta_1$ lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą.

I.2.182. (**I.2.181** pratimo tęsinys). Fiksuotuose taškuose $\tau_i, i = 1, ..., n$, gauti nepriklausomi tiesinės funkcijos $\beta + \beta_1\tau$ matavimai $X_1, ..., X_n$. Matavimo paklaidos $\varepsilon_i \sim N(0, \sigma^2)$. Raskite parametro $\gamma = 4\beta_0 + 4\beta_1\bar{\tau}$, kai $\beta_0 = \beta + \beta_1\bar{\tau}$, lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą.

I.2.183. Tarkime, paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $X \sim U(0, \theta), 0 < \theta$. Raskite parametro θ lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą.

I.2.184. Tarkime, paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $X \sim U(-\theta, \theta), 0 < \theta$. Raskite parametro θ lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą.

I.2.185. Tarkime, paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $X \sim U(\theta - 1/2, \theta + 1/2)$. Raskite parametro θ lygmens $Q = 1$ pasikliovimo intervalą ir raskite jo vidutinį ilgį.

I.2.186. Tarkime, paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $X \sim U(\theta_1, \theta_1 + \theta_2), -\infty < \theta_1 < \infty, \theta_2 > 0$. Raskite parametro θ_2 lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą.

I.2.187. Tarkime, paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $X \sim \mathcal{E}(a, 1/\theta)$, $-\infty < a < \infty, 0 < \theta$. a) Raskite parametro θ lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą.

b) Raskite parametro a lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą, kai parametras θ žinomas ir kai parametras θ nežinomas. c) Raskite parametro $(a, \theta)^T$ pasikliovimo sritį.

I.2.188. Tarkime, paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $X \sim N(\mu, \sigma^2)$, $|\mu| < \infty, \sigma > 0$. Raskite nepriklausomo tolimesnio stebėjimo X_{n+1} lygmens $Q = 1 - 2\alpha$ prognozės intervalą.

I.2.189. Tarkime, paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $X \sim K(\mu, 1)$, $-\infty < \mu < \infty$. a) Raskite parametro μ lygmens $Q = 1 - 2\alpha$ aproksimacinį intervalą, kai parametro μ taškiniu įverčiu imame empirinę medianą $\hat{x}_{1/2}$. b) Raskite parametro μ lygmens $Q = 1 - 2\alpha$ aproksimacinį intervalą, kai parametro μ įvertinys gaunamas DT metodu. Kuris iš šių intervalų trumpesnis?

I.2.190. Tegu $X_1, ..., X_n$ yra paprastoji imtis a. d. $X \sim \mathcal{P}(\lambda)$. a) Raskite parametro λ lygmens $Q = 1 - 2\alpha$ aproksimacinį pasikliovimo intervalą remdamiesi tuo, kad a. d. $\sqrt{n}(\bar{X} - \lambda)/\sqrt{\bar{X}}$ skirstinys asimptotiškai nepriklauso nuo nežinomo parametro.

b) Raskite parametro λ lygmens $Q = 1 - 2\alpha$ aproksimacinį pasikliovimo intervalą remdamiesi tuo, kad a. d. $\sqrt{n}(\bar{X} - \lambda)/\sqrt{\lambda}$ skirstinys asimptotiškai nepriklauso nuo nežinomo parametro.

c) Raskite parametro λ lygmens $Q = 1 - 2\alpha$ aproksimacinį pasikliovimo intervalą remdamiesi dispersiją stabilizuojančia transformacija, t. y. tuo, kad a. d. $\sqrt{n}(2\sqrt{\bar{X}} - 2\sqrt{\lambda})$ skirstinys asimptotiškai nepriklauso nuo nežinomo parametro.

d) Tegu pagal didumo $n = 50$ imtį gauta NMD įvertinio realizacija $\hat{\lambda} = 2,4$. Raskite lygmens $Q = 0,95$ p. a), b), c) aproksimacinius intervalus ir palyginkite juos su tiksliu pasikliovimo intervalu iš **I.2.175** pratimo.

I.2.191. Tegu $Y_1, \ldots, Y_n$ yra nepriklausomi vienodai pasiskirstę a. d. su baigtiniais $\mu_y = \mathbf{E}Y_1$, $\sigma_y^2 = \boldsymbol{V}(Y_1)$, $\alpha_3 = \mathbf{E}Y_1^3$ ir $\alpha_4 = \mathbf{E}Y_1^4$. Raskite statistiką, kurios skirstinys asimptotiškai $(n \to \infty)$ nepriklausytų nuo nežinomų parametrų, ir sudarykite parametro $\boldsymbol{\theta} = (\mu_y, \sigma_y^2)^T$ aproksimacinę pasikliovimo sritį.

I.2.192. Tegu $X_1, ..., X_n$ yra paprastoji imtis, gauta stebint a. d. $X \sim LN(\mu, \sigma)$.

a) Įrodykite, kad $\theta = \mathbf{E}X = \exp\{\mu + \sigma^2/2\}$, $\gamma = \boldsymbol{V}X = \exp\{2\mu + \sigma^2\}(\exp\{\sigma^2\} - 1)$.

b) Įrodykite, kad parametrų θ ir γ DT įvertiniai yra

$$\hat{\theta} = \exp\{\bar{Y} + m_2/2\}, \quad \hat{\gamma} = \exp\{2\bar{Y} + m_2\}(\exp\{m_2\} - 1),$$

čia

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i, \quad m_2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2, \quad Y_i = \ln X_i, \quad i = 1, ..., n.$$

c) Įrodykite, kad

$$\mathbf{E}_{\mu,\sigma}\hat{\theta} = \theta \exp\{-(n-1)\sigma^2/(2n)\}(n/(n-\sigma^2))^{(n-1)/2} = \theta + O(1/n^2), \quad n \to \infty;$$

$$\boldsymbol{V}_{\mu,\sigma}\hat{\theta} = \exp\{2\mu + \sigma^2/n\}[\exp\{\sigma^2/n\}(n/(n-2\sigma^2))^{(n-1)/2} - (n/(n-\sigma^2))^{n-1}] = \theta^2\sigma^2/n + O(1/n^2).$$

I.2.193. (**I.2.192** pratimo tęsinys). Tarkime, pagal didumo $n = 100$ imtį gauti įverčiai $\bar{Y} = 1,45$, $m_2 = 4,21$.

a) Apskaičiuokite parametrų θ ir γ DT įverčius.

b) Raskite parametro θ asimptotinį pasikliovimo intervalą $(Q = 0,95)$.

I.2.194. Pasėjus kultūrą Petri lėkštelėje po tam tikro laiko registruojamas bakterijų skaičius $Y = X_1 + X_2 + ... + X_N$; čia N – kolonijų skaičius, X_i – bakterijų skaičius i-oje kolonijoje. Tarkime, kad $X_1, X_2, ...$ yra vienodai pasiskirstę n. a. d., turintys Puasono skirstinį su parametru θ, o N yra nepriklausantis nuo $X_1, X_2, ...$ a. d., turintis Puasono skirstinį su parametru λ. Momentų metodu raskite parametrų λ ir θ įvertinius.

I.2.195. (**I.2.194** pratimo tęsinys). Tarkime, pagal didumo $n = 20$ imtį, gautą stebint a. d. Y, apskaičiuoti nepaslinktieji vidurkio $\mu = \mathbf{E}Y$ ir dispersijos $\sigma^2 = \boldsymbol{V}Y$ įverčiai $\hat{\mu} = \bar{Y} = 50$ ir $\widehat{\sigma^2} = s^2 = \sum_{i=1}^{n}(Y_i - \bar{Y})^2/(n-1) = 360$.

a) Apskaičiuokite parametrų θ ir λ momentų metodo įverčius.

b) Raskite vidurkio $\mu = \lambda\theta$ lygmens $Q = 1 - 2\alpha$ asimptotinį pasikliovimo intervalą naudodami aproksimaciją

$$\sqrt{n}(\hat{\mu} - \mu)/s \xrightarrow{d} Z \sim N(0,\ 1)$$

ir aproksimaciją

$$\sqrt{n}(\hat{\mu} - \mu)/\sqrt{\mu(1 + \hat{\theta})} \xrightarrow{d} Z \sim N(0,\ 1).$$

I.2.196. Paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $X \sim N(\mu, \sigma^2)$. Raskite kritinės reikšmės $x_P = \mu + \sigma z_P$ ir funkcijos $F(x) = \Phi((x-\mu)/\sigma)$ aproksimacinius pasikliovimo intervalus.

I.2.197. (**I.2.196** pratimo tęsinys). Tarkime, pagal didumo $n = 100$ paprastąją imtį, gautą stebint a. d. $X \sim N(\mu, \sigma^2)$, apskaičiuotos empirinės charakteristikos $\bar{X} = 1{,}28; s^2 = 4; g_1 = -0{,}25; g_2 = 0{,}1$. Sudarykite parametrų $x(0{,}9); F(0); \gamma_1; \gamma_2$ lygmens $Q = 0{,}95$ aproksimacinius pasikliovimo intervalus.

I.2.198. Paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $X \sim W(\theta, \nu)$, turintį Veibulo skirstinį, kurio pasiskirstymo funkcija

$$F(x|\theta, \nu) = 1 - e^{-(x/\theta)^{\nu}}, \quad x > 0.$$

Raskite eilės P kvantilio $t(P) = \theta(-\ln(1-P))^{1/\nu}$ aproksimacinį pasikliovimo intervalą.

I.2.199. (**I.2.198** pratimo tęsinys). Tarkime, kad gaminio darbo laikas aprašomas Veibulo skirstiniu su parametrais θ ir ν. Pagal paprastąją imtį, gautą išbandžius 100 gaminių, surastos parametrų DT įvertinių realizacijos $\hat{\theta} = 0{,}198, \hat{\nu} = 2{,}15$. Raskite parametrų θ, ν ir kvantilio $t(0{,}9)$ lygmens $Q = 0{,}95$ aproksimacinius pasikliovimo intervalus.

I.2.6. Įverčių radimo pavyzdžiai

I.2.200. Bandant sportinį lėktuvą, gautos šios jo maksimalaus greičio (m/s) reikšmės: 422,2; 418,7; 425,6; 420,3; 425,8; 423,1; 431,5; 428,2; 438,3; 434,0; 411,3; 417,2; 413,5; 441,3; 423,0. Tarę, kad buvo stebimas normalusis a. d., raskite vidurkio ir vidutinio kvadratinio nuokrypio taškinius ir intervalinius ($Q = 0{,}95$) įverčius.

I.2.201. Lentelėje pateikti skaičiai m_i tokių vienodo ploto (0,25 kv. km) pietinės Londono dalies rajonų, į kuriuos Antrojo pasaulinio karo metu pataikė po i lėktuvų-sviedinių.

i	0	1	2	3	4	5	Σ
m_i	229	211	93	35	7	1	576

Tarę, kad buvo stebimas Puasono a. d., raskite parametro λ taškinį ir intervalinį ($Q = 0,95$) įverčius.

I.2.202. Laikas nuo užsakymo pateikimo iki jo gavimo (pristatymo laikas) yra pasiskirstęs pagal gama skirstinį $G(\lambda,\ \eta)$. Lentelėje pateikiamos atsitiktinai parinktų užsakymų pristatymo laikas (i – eilės numeris, X_i – laikas).

i	X_i	i	X_i	i	X_i	i	X_i
1	10	6	7	11	10	16	7
2	10	7	11	12	6	17	6
3	6	8	12	13	13	18	16
4	11	9	12	14	8	19	9
5	8	10	6	15	12	20	5

a) Raskite parametrų λ ir η taškinius įverčius.

b) Tarę, kad parametro η reikšmė lygi 10, DT metodu raskite parametro λ įvertį. Palyginkite jį su NMD įverčiu. Sudarykite parametro λ pasikliovimo intervalą ($Q = 0,95$).

I.2.203. Kiekvienomis iš 100 vienodų staklių gaminami I ir II rūšies gaminiai. Tikrinant produkcijos kokybę, atsitiktinai paimta po 10 gaminių, pagamintų skirtingomis staklėmis, ir nustatytas II rūšies gaminių skaičius. Bandymo rezultatai pateikiami lentelėje (m_i – skaičius imčių, kuriose rasta po i II rūšies gaminių).

i	0	1	2	3	4	5	Σ
m_i	1	10	27	36	25	1	100

Tarę, kad buvo stebimas binominis a. d., raskite parametro p NMD įvertį ir pasikliovimo intervalą ($Q = 0,95$).

I.2.204. Bandant kiekvieną iš 10 prietaisų, nebuvo rasta nė vieno defektinio prietaiso. Raskite tikimybės, kad prietaisas yra defektinis, pasikliovimo intervalą, kai pasikliovimo lygmenys yra 0,8; 0,9; 0,99 ir defektinių prietaisų skaičiaus skirstinys yra binominis.

I.2.205. Nustatant 200 elektros lempučių degimo laiką T, gauti stebiniai, kurie pateikiami lentelėje.

i	$a_{i-1} - a_i$	S_i	i	$a_{i-1} - a_i$	S_i
1	0–300	53	7	1800–2100	9
2	300–600	41	8	2100–2400	7
3	600–900	31	9	2400–2700	5
4	900–1200	22	10	2700–3000	3
5	1200–1500	16	11	3000–∞	2
6	1500–1800	12			

Šioje lentelėje a_{i-1} ir a_i yra i-ojo intervalo galai, o S_i stebinių, patekusių į i-ąjį intervalą, skaičius.

Tarę, kad a. d. T skirstinys yra eksponentinis, raskite parametro λ taškinį ir asimptotinį intervalinį įverčius ($Q = 0,99$).

I.2.206. Pagal tris didumo $n_1 = 20, n_2 = 50, n_3 = 30$ imtis, gautas stebint n. a. d.

$X \sim N(\mu_1, \sigma^2)$, $Y \sim N(\mu_2, \sigma^2)$, $Z \sim N(\mu_3, \sigma^2)$, apskaičiuotos NMD įvertinių realizacijos $\hat{\mu}_1 = \bar{X} = 2,12, s_x^2 = 2,84$, $\hat{\mu}_2 = \bar{Y} = 1,09, s_y^2 = 3,91$, $\hat{\mu}_3 = \bar{Z} = 3,14, s_z^2 = 2,53$.

a) Raskite parametro σ^2 NMD įvertį naudodami visus duomenis; raskite parametro σ lygmens $Q = 0,99$ pasikliovimo intervalą.

b) Raskite parametro $\theta = \mu_1 + \mu_2 - \mu_3$ NMD įvertį; sudarykite parametro θ lygmens $Q = 0,95$ pasikliovimo intervalą.

I.2.207. Krakmolo kiekis bulvėse nustatomas dviem būdais. Norint palyginti tuos būdus, buvo paimta 16 bulvių ir kiekvienos iš jų krakmolo kiekis nustatytas abiem būdais. Gauti stebiniai (krakmolingumas procentais) surašyti lentelėje (X_i – krakmolingumas tiriant i-ąją bulvę pirmu būdu; Y_i – antru būdu).

i	X_i	Y_i	i	X_i	Y_i
1	21,7	21,5	9	14,0	13,9
2	18,7	18,7	10	17,2	17,0
3	18,3	18,3	11	21,7	21,4
4	17,5	17,4	12	18,6	18,6
5	18,5	18,3	13	17,9	18,0
6	15,6	15,4	14	17,7	17,6
7	17,0	16,7	15	18,3	18,5
8	16,6	16,9	16	15,6	15,5

Priėmę normalumo prielaidą, palyginkite šiuos du krakmolingumo nustatymo metodus. a) Raskite parametro $\theta = \mathbf{E}X - \mathbf{E}Y$ lygmens $Q = 0,95$ pasikliovimo intervalą remdamiesi **I.2.160** pratimu. b) Raskite parametro θ lygmens $Q = 0,95$ pasikliovimo intervalą remdamiesi **I.2.168** pratimu. c) Paaiškinkite, kodėl gaunami tokie skirtingi rezultatai. d) Raskite koreliacijos koeficiento ρ taškinį ir intervalinį ($Q = 0,95$) įverčius.

I.2.208. Per pirmą dieną skaitiklis užregistravo 20 026 puasoninio srauto impulsus, o per antrą – 19 580. Raskite intensyvumų santykio $\theta = \lambda_1/\lambda_2$ pasikliovimo intervalą ($Q = 0,99$).

I.2.209. Dviejose nepriklausomose Bernulio bandymų schemose atlikus $n_1 = n_2 = 5000$ bandymų įvykis A įvyko $S_1 = 2602$ ir $S_2 = 2398$ kartus. Tegu įvykio A pasirodymo tikimybė pirmoje bandymų schemoje yra p_1, o antroje – p_2. Raskite parametro $\theta = p_1 - p_2$ asimptotinį pasikliovimo intervalą ($Q = 0,99$). Ar yra pagrindo teigti, kad įvykio A pasirodymo tikimybės abiejose schemose yra vienodos?

I.2.210. Tarkime, kad daugialypis integralas, kurio tikroji reikšmė yra 0,3, buvo skaičiuojamas Monte Karlo metodu. Kiek apytiksliai reikia atlikti nepriklausomų modeliavimų N, kad gautosios reikšmės absoliuti santykinė paklaida su tikimybe, ne mažesne už 0,99, neviršytų a) 0,2; b) 0,1?

I.2.211. Lentelėje pateikti prapuolimo kampai 209 pašto balandžių, kai atliekant bandymą buvo bandoma paveikti jų „vidinį laikrodį“ (žr. [14]).

Duomenys sugrupuoti į 30° ilgio intervalus. Lentelėje nurodyti kampai φ_i, atitinkantys i-ojo intervalo pradžią, ir patekusių į i-ąjį intervalą dažniai n_i, $i = 1, ...12$. Tardami, kad turimus duomenis galima traktuoti kaip paprastosios imties, gautos stebint atsitiktinį kampą $\varphi \sim M(\mu, \theta)$, realizaciją, raskite parametrų μ, θ taškinius įverčius ir sudarykite aproksimacinius intervalus su pasikliovimo lygmeniu $Q = 0,95$.

Kryptis	Dažnis	Kryptis	Dažnis
$0°-$	26	$180°-$	14
$30°-$	22	$210°-$	11
$60°-$	26	$240°-$	12
$90°-$	30	$270°-$	5
$120°-$	29	$300°-$	5
$150°-$	18	$330°-$	11

I.2.212. Lentelėje pateikti duomenys apie užregistruotus susirgimo leukemija atvejus Anglijoje per 1946-1960 metų laikotarpį, sugrupuoti mėnesiniais intervalais (žr. [14]).

Mėnuo	Susirgo	Mėnuo	Susirgo	Mėnuo	Susirgo
Sausis	39	Gegužė	38	Rugsėjis	37
Vasaris	37	Birželis	59	Spalis	47
Kovas	29	Liepa	50	Lapkritis	34
Balandis	45	Rugpjūtis	54	Gruodis	37

Paverskite duomenis kampų stebėjimais sutapatindami metų intervalą su intervalu $(0,\ 2\pi]$, t. y. sausis atitinka sektorių nuo $0°$ iki $30°$; vasaris – sektorių nuo $30°$ iki $60°$ ir t. t. Tardami, kad buvo stebimas atsitiktinis kampas $\varphi \sim M(\mu, \theta)$ a) raskite taškinius parametrų (μ, θ) įverčius; b) sudarykite pasikliovimo lygmens $Q = 0,95$ aproksimacinius pasikliovimo intervalus.

I.2.7. Sprendimai, nurodymai, atsakymai

I.2.1 skyrelis

I.2.1. A. d. $\sqrt{n}(\bar{X} - \mu)/2 \sim N(0,\ 1)$. Imties didumui n rasti gauname nelygybę

$$\mathbf{P}_\mu\{|\bar{X} - \mu| \le 0,1\} = \mathbf{P}_\mu\{-0,1\sqrt{n}/2 \le \sqrt{n}(\bar{X} - \mu)/2 \le 0,1\sqrt{n}/2\} =$$

$$2\Phi(0,1\sqrt{n}/2) - 1 \ge 0,99.$$

Iš šios nelygybės randame

$$\Phi(0,1\sqrt{n}/2) \ge 0,995 \quad \Leftrightarrow 0,1\sqrt{n}/2 \ge z_{0,005} \quad \Leftrightarrow n \ge 2654.$$

I.2.2. A. d. $s^2(n-1)/\sigma^2 \sim \chi^2(n-1)$. Imties didumui n rasti gauname nelygybę

$$\mathbf{P}_\sigma\{|s^2 - \sigma^2|/\sigma^2 \le 0,1\} = \mathbf{P}_\sigma\{0,9(n-1) \le s^2(n-1)/\sigma^2 \le 1,1(n-1)\} \ge 0,99.$$

Iš šios nelygybės randame

$$\mathbf{P}\{0,9(n-1) \le \chi^2_{n-1} \le 1,1(n-1)\} \ge 0,99 \quad \Leftrightarrow n \ge 1330.$$

I.2.3. A. d. $n\bar{X} \sim G(1/\lambda, 5n)$. Remdamiesi gama skirstinio sąryšiu su χ^2 skirstiniu (žr. 1 priedo 1.P.3 lentelę) gauname $2n\bar{X}/\lambda \sim \chi^2(10n)$. Imties didumui n rasti gauname nelygybę

$$\mathbf{P}\lambda\{|\bar{X}/5 - \lambda|/\lambda \le 0,1\} = \mathbf{P}_\lambda\{0,9 \le \bar{X}/(5\lambda) \le 1,1\} =$$

$$\mathbf{P}_\lambda\{9n \le 2n\bar{X}/\lambda \le 11n\} = \mathbf{P}_\lambda\{9n \le \chi^2_{10n} \le 11n\} \ge 0,99.$$

Iš šios nelygybės randame $n \geq 133$.

I.2.4. A. d. $\sqrt{n}(\bar{X}-\mu)/25 \sim N(0,1)$; čia μ – nežinomas jūros gylis. Matavimų skaičiui n rasti turime nelygybę

$$\mathbf{P}_\mu\{\bar{X}-\mu \leq 15\} = \mathbf{P}_\mu\{\sqrt{n}(\bar{X}-\mu)/25 \leq 3\sqrt{n}/5\} = \Phi(3\sqrt{n}/5) \geq 0,99.$$

Išsprendę gauname

$$\Phi(3\sqrt{n}/5) \geq 0,99 \quad \Leftrightarrow 3\sqrt{n}/5 \geq z_{0,01} \quad \Leftrightarrow n \geq 15.$$

I.2.5. Pažymėkime $T_{1n}, T_{2n}, ..., T_{nn}$ imties narius. Yra n aibės $\{1,2,...,n\}$ poaibių, susidedančių iš $n-1$ elemento. Taigi yra n statistikų $T_{n-1}^{(1)}, ..., T_{n-1}^{(n)}$, gautų iš didumo $n-1$ imčių; $\bar{T}_{n-1} = \frac{1}{n}\sum_{j=1}^n T_{n-1}^{(j)}$. Kadangi

$$\mathbf{E}_\theta T_{n-1}^{(j)} = \theta + \sum_{i=1}^{\infty} a_i/(n-1)^i,$$

tai ir statistika $\bar{T}_{n-1}$ tenkina sąlygą

$$\mathbf{E}_\theta \bar{T}_{n-1} - \theta = \sum_{i=1}^{\infty} a_i/(n-1)^i.$$

Taigi

$$\mathbf{E}_\theta T_n' = n\mathbf{E}_\theta T_n - (n-1)\mathbf{E}_\theta \bar{T}_{n-1} = \theta + \sum_{i=2}^{\infty} a_i(1/n^{i-1} - 1/(n-1)^{i-1} =$$

$$\theta + a_2(1/n - 1/(n-1)) + O(1/n^2) = \theta + O(1/n^2).$$

Analogiškai randame $T_n'', T_n''', ...$ poslinkius.

I.2.6. Tegu $Y = X^\alpha$. Tada $Y \sim \mathcal{E}(\rho^\alpha)$. Suma $\sum_{i=1}^n X_i^\alpha \sim G(\rho^\alpha, n)$. Randame vidurkį

$$\mathbf{E}_\rho \hat{\gamma} = \frac{\rho^{n\alpha}}{\Gamma(n-r/\alpha)} \int_0^\infty y^{n-r/\alpha-1} e^{-\rho^\alpha y} dy = \rho^r.$$

I.2.7. Jeigu $\bar{\theta}$ ir $\tilde{\theta}$ yra nepaslinktieji parametro θ įvertiniai, tai $\mathbf{E}_\theta(\bar{\theta} - \tilde{\theta}) \equiv 0$. Taigi nepaslinktieji įvertiniai gali skirtis tik tokia statistika U, kad $\mathbf{E}_\theta U \equiv 0$.

I.2.8. Statistikos $X_{(n)}$ tankio funkcija yra $f_n(x) = nx^{n-1}/\theta^n, 0 < x < \theta$. Įvertinio $cX_{(n)}$ vidurkis ir dispersija

$$\mathbf{E}_\theta(cX_{(n)}) = \frac{cn}{\theta^n}\int_0^\theta x^n dx = \theta\frac{cn}{n+1}, \quad \mathbf{E}_\theta(cX_{(n)})^2 = \frac{c^2 n}{\theta^n}\int_0^\theta x^{n+1} dx = \theta^2\frac{c^2 n}{n+2},$$

$$\boldsymbol{V}_\theta(cX_{(n)}) = c^2\theta^2 \frac{n}{(n+1)^2(n+2)}.$$

Įvertinys nepaslinktas, kai $c = (n+1)/n$. Tada

$$\mathbf{E}_\theta((n+1)X_{(n)}/n) = \theta, \quad \boldsymbol{V}_\theta((n+1)X_{(n)}/n) = \theta^2/(n(n+2)).$$

I.2.9. Tarkime, $\varepsilon < \min|\theta_i - \theta_j|$. Tada

$$\mathbf{P}_\theta\{|T_n(\mathbf{X}) - \theta| \geq \varepsilon\} = 1 - \mathbf{P}_\theta\{|T_n(\mathbf{X}) - \theta| < \varepsilon\} = 1 - \mathbf{P}_\theta\{T_n(\mathbf{X}) = \theta\}.$$

Taigi

$$\mathbf{P}_\theta\{|T_n(\mathbf{X}) - \theta| \geq \varepsilon\} \to 0 \quad \Leftrightarrow \quad \mathbf{P}_\theta\{T_n(\mathbf{X}) = \theta\} \to 1.$$

Jeigu $\forall\ \varepsilon > 0$ tikimybė $\mathbf{P}_\theta\{|T_n(\mathbf{X}) - \theta| \geq \varepsilon\} \to 0$, tai

$$\mathbf{P}_\theta\{|T_n(\mathbf{X}) - \theta| \geq \varepsilon\} \to 0, \quad \forall\ \varepsilon < \min|\theta_1 - \theta_j| \ \Rightarrow \mathbf{P}_\theta\{T_n(\mathbf{X}) = \theta\} \to 1.$$

Atvirkščiai, jeigu $\mathbf{P}_\theta\{T_n(\mathbf{X}) = \theta\} \to 1$, tai $\mathbf{P}_\theta\{|T_n(\mathbf{X}) - \theta| < \varepsilon\} \to 1\ \ \forall\ \varepsilon > 0$. Iš čia gauname, kad $\forall\ \varepsilon > 0$ tikimybė $\mathbf{P}_\theta\{|T_n(\mathbf{X}) - \theta| \geq \varepsilon\} \to 0$.

I.2.10. Tegu $Y_i = X_i - (\theta - 1/2)$, $i = 1, ..., n$. Tada $Y_1, ..., Y_n$ yra paprastoji imtis, gauta stebint a. d. $Y \sim U(0,1)$; $\mathbf{E}_\theta((X_{(1)} + X_{(n)})/2) = \mathbf{E}((Y_{(1)} + Y_{(n)})/2) + \theta - 1/2$; $\boldsymbol{V}_\theta((X_{(1)} + X_{(n)})/2) = \boldsymbol{V}((Y_{(1)} + Y_{(n)})/2)$. Gauname

$$\mathbf{E}(Y_{(1)} + Y_{(n)}) = n\int_0^1 x(1-x)^{n-1}dx + n\int_0^1 x^n dx = nB(2,n) + n/(n+1) = 1,$$

$$\mathbf{E}_\theta((X_{(1)} + X_{(n)})/2)) = \mathbf{E}((Y_{(1)} + Y_{(n)})/2) + \theta - 1/2 = \theta,$$

A. v. $(Y_{(1)}, Y_{(n)})^T$ tankio funkcija (žr. [2], 2.3 skyrelį) yra $n(n-1)(y-x)^{n-2}, 0 < x < y < 1$. Randame

$$\mathbf{E}(Y_{(1)} + Y_{(n)})^2 = n(n-1)\int_0^1\int_0^y (x+y)^2(y-x)^{n-2}dxdy.$$

Atlikę keitimą $x = yt,\ dx = ydt$, gauname

$$\mathbf{E}(Y_{(1)} + Y_{(n)})^2 = n(n-1)\int_0^1 y^{n+1}dy\int_0^1 (1+t)^2(1-t)^{n-2}dt =$$

$$\frac{n(n-1)}{n+1}[B(1,n-1) + 2B(2,n-1) + B(3,n-1)] = \frac{n^2+3n+4}{(n+1)(n+2)};$$

$$\boldsymbol{V}_\theta((X_{(1)} + X_{(n)})/2) = \boldsymbol{V}((Y_{(1)} + Y_{(n)})/2) = 1/(2(n+1)(n+2)).$$

Remdamiesi Čebyšovo nelygybe randame

$$\mathbf{P}_\theta\{n^\alpha|\hat{\theta} - \theta| \geq \varepsilon\} \leq \frac{n^{2\alpha}}{2(n+1)(n+2)\varepsilon^2} \to 0, \quad n \to \infty.$$

I.2.11. Tarkime, $S_n = X_1 + ... + X_n$. Statistika T įgyja reikšmę 0 su tikimybe $P_n = \mathbf{P}\{S_n \leq n/2\}$ ir reikšmę 1 su tikimybe $Q_n = 1 - P_n$. Statistikos T pirmieji momentai $\mathbf{E}T = Q_n, \boldsymbol{V}T = P_nQ_n$. Jeigu $0 < p < 1/2$, tai

$$Q_n = \mathbf{P}_p\{S_n > n/2\} = \mathbf{P}_p\{\frac{S_n - np}{\sqrt{np(1-p)}} > \frac{\sqrt{n}(1/2-p)}{\sqrt{p(1-p)}}\} \to 0,$$

kai $n \to \infty$. Taigi statistika $T \overset{P}{\nrightarrow} p$ su visais $0 < p < 1$.

I.2.12. Iš funkcijos $g_n(x)$ tolygaus konvergavimo išplaukia, kad bet kuriam $\varepsilon > 0$ egzistuoja toks N, kad

$$|g_n(x) - g(x)| < \varepsilon/2, \quad \forall\ x \in (a, b), \quad n > N.$$

Remdamiesi $g_n(x)$ tolydumu ir tolygiu konvergavimu gauname, kad $g(x)$ yra tolydi funkcija. Taigi duotam $\varepsilon > 0$ egzistuoja toks $\delta > 0$, kad

$$|g(T_n) - g(\theta)| < \varepsilon/2,$$

jei $|T_n - \theta| < \delta$. Gauname

$$|g_n(T_n) - g(\theta)| \leq |g_n(T_n) - g(T_n)| + |g(T_n) - g(\theta)| < \varepsilon,$$

jeigu $n > N$ ir $|T_n - \theta| < \delta$. Taigi, kai $n > N$

$$\mathbf{P}_\theta\{|T_n - \theta| < \delta\} \leq \mathbf{P}_\theta\{|g_n(T_n) - g(\theta)| < \varepsilon\};$$

$$\mathbf{P}_\theta\{|T_n - \theta| < \delta\} \to 1 \quad \Rightarrow \quad \mathbf{P}_\theta\{|g_n(T_n) - g(\theta)| < \varepsilon\} \to 1$$
$$\Rightarrow \quad \mathbf{P}_\theta\{g_n(|T_n) - g(\theta)| > \varepsilon\} \to 0, \quad n \to \infty.$$

I.2.13. Tegu $\hat{\vartheta} = 1$, kai $|\bar{X}| \leq c_n$, ir $\hat{\vartheta} = 0$, kai $|\bar{X}| > c_n$; čia $c_n = 1/n^\alpha$, $0 < \alpha < 1/2$. Gauname

$$\mathbf{P}\{|\hat{\vartheta} - g(\mu)| \geq \varepsilon\} = \mathbf{P}\{|1 - g(\mu)| \geq \varepsilon, |\bar{X}| \leq c_n\} + \mathbf{P}\{|g(\mu)| \geq \varepsilon, |\bar{X}| > c_n\}.$$

Tada

$$\mathbf{P}\{|\hat{\vartheta} - g(\mu)| \geq \varepsilon|\mu = 0, \sigma\} = \mathbf{P}\{|\bar{X} - \mathbf{E}\bar{X}| > c_n|\mu = 0, \sigma\} < \frac{\boldsymbol{V}_\sigma \bar{X}}{1/n^{2\alpha}} = \sigma^2 n^{2\alpha - 1} \to 0;$$

$$\mathbf{P}\{|\hat{\vartheta} - g(\mu)| \geq \varepsilon|\mu \neq 0, \sigma\} = \mathbf{P}_{\mu,\sigma}\{\frac{-c_n - \mu}{\sigma\sqrt{n}} \leq \frac{\bar{X} - \mu}{\sigma\sqrt{n}} \leq \frac{c_n - \mu}{\sigma\sqrt{n}}\} \to \Phi(0) - \Phi(0) = 0,$$

nes pagal CRT $(\bar{X} - \mu)/(\sigma\sqrt{n}) \xrightarrow{d} Z \sim N(0,\ 1)$.

I.2.14. Žr. **I.1.50**, **I.1.62** pratimus.

I.2.15. Žr. **I.1.48**, **I.1.62** pratimus.

I.2.16. A. d. X dispersija $\boldsymbol{V}_\theta X = \mathbf{E}_\theta X^2 - (\mathbf{E}_\theta X)^2 = \mathbf{E}_\theta X^2 - \theta^2$. Imkime statistiką $T = X^2 - Y$. Tada $\mathbf{E}_\theta T = \mathbf{E}_\theta X^2 - \mathbf{E}_\theta Y = \mathbf{E}_\theta X^2 - \theta^2 = \boldsymbol{V}_\theta X$. Taigi T yra nepaslinktasis dispersijos $\boldsymbol{V}_\theta X$ įvertinys.

I.2.17. A. d. $X - Y \sim N(0, 2\sigma^2)$. Randame

$$\mathbf{E}_\sigma|X - Y| = \frac{1}{\sqrt{\pi}\sigma} \int_0^\infty x e^{-x^2/(4\sigma^2)} dx = \frac{2\sigma}{\sqrt{\pi}} \neq \sigma;$$

$$\mathbf{E}_\sigma(X - Y)^2 = 2\sigma^2 \neq \sigma^2.$$

I.2.18. Tegu $Y_i = (X_i + \theta)/(2\theta)$. Tada $Y_1, ..., Y_n$ yra paprastoji imtis a. d. $Y \sim U(0, 1)$; $\mathbf{E}_\theta(X_{(n)} - X_{(1)}) = 2\theta\mathbf{E}(Y_{(n)} - Y_{(1)})$, $\boldsymbol{V}_\theta(X_{(n)} - X_{(1)}) = 4\theta^2\boldsymbol{V}(Y_{(n)} - Y_{(1)})$.

Remiantis **I.1.37** pratimu a. d. $Y_{(n)} - Y_{(1)} \sim Be(n-1, 2)$. Tada (žr. 1 priedo 1.P.2 lentelę)

$$\mathbf{E}(Y_{(n)} - Y_{(1)}) = \frac{n-1}{n+1}, \quad \boldsymbol{V}(Y_{(n)} - Y_{(1)}) = \frac{2(n-1)}{(n+1)^2(n+2)}.$$

Nepaslinktasis parametro θ įvertinys $\hat{\theta} = (X_{(n)} - X_{(1)})(n+1)/(2(n-1))$

$$\mathbf{E}_\theta \hat{\theta} = \theta, \quad \boldsymbol{V}_\theta \hat{\theta} = \frac{2\theta^2}{(n-1)(n+2)}.$$

I.2.2 skyrelis

I.2.19. Tikėtinumo funkcija

$$L(\mu, \lambda) = \lambda^n e^{-\lambda \sum_{i=1}^n (X_i - \mu)}, \quad \mu < X_1, ..., X_n < \infty.$$

Arba

$$L(\mu, \lambda) = \lambda^n e^{\lambda n \mu} e^{-\lambda X_{(1)}} e^{-\lambda(X_{(2)} + ... + X_{(n)})} \mathbf{1}_{(\mu, \infty)}(X_{(1)}).$$

Remiantis Neimano ir Fišerio faktorizacijos kriterijumi $\boldsymbol{T} = (X_{(1)}, X_{(2)} + ... + X_{(n)})^T$ yra parametro $(\mu, \lambda)^T$ pakankamoji statistika.

I.2.20. Tikėtinumo funkcija

$$L(\eta, \sigma) = \eta^n (1/\sigma)^{n\eta} (\prod_i X_i)^{\eta - 1} e^{-(1/\sigma^\eta) \sum_i X_i^\eta}.$$

a) jeigu η nežinomas, tai L priklauso nuo visų imties elementų;
b) jeigu η žinomas, tai tikėtinumo funkcija

$$L(\sigma) = \eta^n (\prod_i X_i)^{\eta-1} (1/\sigma^{n\eta}) e^{-(1/\sigma^\eta) \sum_i X_i^\eta} = W(\mathbf{X}) q(T; \sigma),$$

kai $T = \sum_i X_i^\eta$. Remiantis Neimano ir Fišerio faktorizacijos kriterijumi T yra parametro σ pakankamoji statistika.

I.2.21. Tikėtinumo funkcija

$$L(\lambda, \eta, \mu) = \frac{\lambda^{n\eta}}{[\Gamma(\eta)]^n} (\prod_i X_i)^{\eta-1} e^{-\lambda \sum_i (X_i - \mu)} \mathbf{1}_{(\mu, \infty)}(X_{(1)}).$$

a) Kai μ ir λ žinomi, tikėtinumo funkcija

$$L(\eta) = e^{-\lambda \sum_i (X_i - \mu)} \mathbf{1}_{(\mu, \infty)}(X_{(1)}) \frac{\lambda^{n\eta}}{[\Gamma(\eta)]^n} (\prod_i X_i)^{\eta - 1} = W(\mathbf{X}) q(T; \eta),$$

kai $T = \prod_i X_i$. Remiantis Neimano ir Fišerio faktorizacijos kriterijumi T yra parametro η pakankamoji statistika.

b) Kai μ ir η žinomi, tikėtinumo funkcija

$$L(\lambda) = \frac{1}{[\Gamma(\eta)]^n} (\prod_i X_i)^{\eta-1} \mathbf{1}_{(\mu, \infty)}(X_{(1)}) \lambda^{n\eta} e^{-\lambda \sum_i (X_i - \mu)} = W(\mathbf{X}) q(T; \lambda),$$

kai $T=\sum_i(X_i-\mu)$. Remiantis Neimano ir Fišerio faktorizacijos kriterijumi T yra parametro λ pakankamoji statistika.

c) Kai μ žinomas, tikėtinumo funkcija

$$L(\lambda,\eta)=\mathbf{1}_{(\mu,\infty)}(X_{(1)})\frac{\lambda^{n\eta}}{[\Gamma(\eta)]^n}(\prod X_i)^{\eta-1}e^{-\lambda\sum_i(X_i-\mu)}=W(\mathbf{X})q(\boldsymbol{T};\lambda,\eta),$$

kai $\boldsymbol{T}=(\prod_i X_i,\sum_i(X_i-\mu))^T$. Remiantis Neimano ir Fišerio faktorizacijos kriterijumi $\boldsymbol{T}$ yra parametro $(\lambda,\eta)^T$ pakankamoji statistika.

d) Kai λ žinomas, o $\eta=1$, tikėtinumo funkcija

$$L(\mu)=\lambda^n e^{-\lambda\sum_i X_i}e^{\lambda n\mu}\mathbf{1}_{(\mu,\infty)}(X_{(1)})=W(\mathbf{X})q(T;\mu),$$

kai $T=X_{(1)}$. Remiantis Neimano ir Fišerio faktorizacijos kriterijumi T yra parametro μ pakankamoji statistika.

I.2.22. Tikėtinumo funkcija

$$L(\boldsymbol{\theta})=(2\pi)^{-n/2}\exp\{-\frac{1}{2\sigma^2}\sum_i(Y_i-\alpha-\beta x_i)^2-\frac{n}{2}\ln\sigma^2\}=$$

$$(2\pi)^{-n/2}\exp\{-\frac{1}{2\sigma^2}\sum_i Y_i^2+\frac{\alpha}{\sigma^2}\sum_i Y_i+\frac{\beta}{\sigma^2}\sum_i Y_i x_i-B(\boldsymbol{\theta})\}$$

priklauso triparametrei eksponentinio tipo skirstinių šeimai. Parametrų kitimo sritis turi vidinių taškų. Taigi statistika $\boldsymbol{T}$ yra pilnoji ir pakankamoji parametro $\boldsymbol{\theta}$ statistika.

I.2.23. Remdamiesi Neimano ir Fišerio faktorizacijos kriterijumi gauname pakankamas statistikas: a) b) c) $T=\sum_i X_i$; d) $\boldsymbol{T}=(\prod_i X_i,\sum_i X_i)^T$; e) $\boldsymbol{T}=(\prod_i X_i,\prod(1-X_i))^T$; f) $\boldsymbol{T}=(\sum_i\ln X_i,\sum_i\ln^2 X_i)^T$.

I.2.24. Tikėtinumo funkcija

$$L(\theta)=(2/\theta^2)^n\prod_i(\theta-X_i)\mathbf{1}_{(0,\theta)}(X_{(n)}).$$

Egzistuoja tik triviali pakankamoji statistika $\boldsymbol{T}=(X_1,...,X_n)^T$.

I.2.25. a) Tikėtinumo funkcija

$$L(\theta)=\theta^n/\prod_i(1+X_i)^{\theta+1}=\theta^n e^{-(\theta+1)\sum_i\ln(1+X_i)};$$

remiantis Neimano ir Fišerio faktorizacijos kriterijumi $T=\sum_i\ln(1+X_i)$ yra pakankamoji parametro θ statistika; b) nesunku patikrinti, kad a. d. $\ln(1+X_i)\sim\mathcal{E}(\theta)$; $T=\sum_i\ln(1+X_i)\sim G(\theta,n)$. Tada $\mathbf{E}_\theta T=n/\theta$, $\boldsymbol{V}_\theta T=n/\theta^2$.

I.2.26. Pagal Neimano ir Fišerio faktorizacijos kriterijų $L(\theta)=W(\mathbf{X})q(T;\theta)$. Tada $L(\theta)=W(\mathbf{X})q(h(S);\theta)$. Taigi S irgi pakankamoji statistika.

I.2.27. Tikėtinumo funkcija

$$L(\lambda,\theta)=\theta^n\alpha^{n\theta}/(\prod_i X_i)^{\theta+1}\mathbf{1}_{(\alpha,\infty)}(X_{(1)}).$$

Remiantis Neimano ir Fišerio faktorizacijos kriterijumi $\boldsymbol{T} = (X_{(1)}, \prod_i X_i)^T$ yra pakankamoji parametro $(\alpha, \theta)^T$ statistika.

I.2.28. Variacinės eilutės tikėtinumo funkcija $f_n(X_{(1)}, ..., X_{(n)}) = n! f(X_{(1)}) \cdot ... \cdot f(X_{(n)})$. Remiantis Neimano ir Fišerio faktorizacijos kriterijumi $(X_{(1)}, ..., X_{(n)})^T$ yra šeimos $\mathcal{P}$ pakankamoji statistika.

I.2.29. Eksponentinio tipo skirstinių tankio kanoninė forma yra

$$f(x|\boldsymbol{\eta}) = h(x) \exp\{\boldsymbol{\eta}^T \boldsymbol{T}(x) - B(\boldsymbol{\eta})\}.$$

a) $h(x) = 1/x!,\ \eta = \ln\lambda,\ T(x) = x,\ B(\eta) = e^{\eta}; -\infty < \eta < \infty,\ x = 0, 1, ...;$

b) $h(x) = C_{n+x-1}^{n-1},\ \eta = \ln(1-p),\ T(x) = x,\ B(\eta) = -\ln(1-e^{\eta}),\ -\infty < \eta < 0,\ x = 0, 1, ...;$

c) $h(x) = \boldsymbol{I}_{(a,\ \infty)}(x),\ \eta = -\theta,\ T(x) = x,\ B(\eta) = \eta a - \ln(-\eta),\ -\infty < \eta < 0,\ 0 < x < \infty;$

d) $h(x) = 1/x,\ \boldsymbol{\eta} = (\eta_1, \eta_2)^T = (-\lambda,\ \eta)^T,\ \boldsymbol{T}(x) = (x,\ \ln x)^T,\ B(\boldsymbol{\eta}) = \ln\Gamma(\eta_2) - \eta_2 \ln(-\eta_1),\ -\infty < \eta_1 < 0,\ 0 < \eta_2 < \infty, x > 0;$

e) $h(x) = 1/[x(1-x)],\ \boldsymbol{\eta} = (\eta_1, \eta_2)^T = (\gamma,\ \eta)^T,\ \boldsymbol{T}(x) = (\ln x,\ \ln(1-x))^T,\ B(\boldsymbol{\eta}) = \ln[\Gamma(\eta_1)\Gamma(\eta_2)/\Gamma(\eta_1+\eta_2)],\ 0 < \eta_1,\ \eta_2 < \infty,\ 0 < x < 1;$

f) $h(x) = \alpha x^{\alpha-1}, \eta = -1/\theta^{\alpha},\ T(x) = x^{\alpha},\ B(\eta) = \ln(-1/\eta),\ -\infty < \eta < 0,\ 0 < x < \infty.$

I.2.30. Tankio funkcija $f(x|\alpha, \theta) = \theta e^{-\theta(x-\alpha)}, x > \alpha$ negali būti užrašyta eksponentinio tipo tankiu (žr. [2], 3.3.2 apibrėžimą).

I.2.31. Tankio funkcija $f(x|n, p) = C_{n+x-1}^{x} p^n (1-p)^x, x = 0, 1, ...$ negali būti užrašyta eksponentinio tipo tankiu.

I.2.32. Tankio funkcija $f(x|\mu, \sigma) = (1/(\sigma\pi))(1/(1+(x-\mu)^2/\sigma^2))$ negali būti užrašyta eksponentinio tipo tankiu.

I.2.33. Tankio funkcija $f(x|\alpha, \theta) = (\alpha x^{\alpha-1}/\theta^{\alpha}) e^{-(x/\theta)^{\alpha}}$ negali būti užrašyta eksponentinio tipo tankiu.

I.2.34. Tankio funkciją $f(\boldsymbol{x}|\boldsymbol{\mu},\ \boldsymbol{\Sigma}) = (\sqrt{2\pi})^{-k}/\sqrt{|\boldsymbol{\Sigma}|} \exp\{(-1/2)(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\}$ galima perrašyti $f(\boldsymbol{x}|\boldsymbol{\mu},\ \boldsymbol{\Sigma}) = (\sqrt{2\pi})^{-k} \exp\{(-1/2)\sum_i \sigma^{ii} x_i^2 - \sum_{i>j} \sigma^{ij} x_i x_j + \sum_i x_i \sum_j \mu_j \sigma^{ij} -\ (1/2)\sum_i \sum_j \mu_i \sigma^{ij} \mu_j - (1/2)\ln|\boldsymbol{\Sigma}|\}$; čia $\boldsymbol{\Sigma}^{-1} = [\sigma^{ij}]_{k\times k}$. Gauname $(k(k+1)/2)$-matį eksponentinio tipo skirstinį. Kanoninės formos parametrai yra $\boldsymbol{\eta} = (-\sigma^{ii}/2,\ i = 1, ..., k;\ -\sigma^{ij},\ i > j,\ i,\ j = 1, ..., k;\ \sum_j \mu_j \sigma^{ij}, i = 1, ..., k)^T$, $\boldsymbol{T} = (\sum_i X_i^2, i = 1, ..., k; \sum_{i<j} X_i X_j, i > j, i, j = 1, ..., k;\ \sum_i X_i, i = 1, ..., k)^T$, $B(\boldsymbol{\eta}) = (1/2)\sum_i \sum_j \mu_i \sigma^{ij} \mu_j + (1/2)\ln|\boldsymbol{\Sigma}|$.

I.2.35. $M(t) = (1 - t/\lambda)^{-\eta},\ t < \lambda$.

I.2.36. Tankis skaičiuojančiojo mato atžvilgiu yra

$$f(x|\theta) = \gamma(x) e^{x\ln\theta - \ln c(\theta)}, \quad x = 0, 1, ...$$

priklauso eksponentinio tipo skirstinių šeimai. Kadangi

$$\sum_{x=0}^{\infty} \gamma(x)\theta^x = c(\theta),$$

tai

$$M(t) = \mathbf{E}_\theta(e^{tX}) = \frac{1}{c(\theta)} \sum_{x=0}^{\infty} \gamma(x) e^{tx} \theta^x = \frac{c(\theta e^t)}{c(\theta)}.$$

I.2.37. a) Šeimos $\mathcal{P}$ tikėtinumo funkcija

$$L(\theta) = \prod_{i=1}^{n} f(X_i|\theta) = W(\mathbf{X}) q(T;\theta), \quad \theta \in \Theta.$$

Tada šeimos $\mathcal{P}_A$ tikėtinumo funkcija

$$L(\theta) = \prod_{i=1}^{n} f(X_i|\theta)/[P_\theta(A)]^n = W(\mathbf{X}) q(T;\theta)/[P_\theta(A)]^n, \quad \theta \in \Theta.$$

Taigi T yra ir šeimos $\mathcal{P}_A$ pakankamoji statistika. b) Jeigu iš sąlygos $\mathbf{E}_\theta h(T) \equiv 0, \theta \in \Theta$ išplaukia, kad $h(T) = 0$ b. v. skirstinių $\mathcal{P}$ atžvilgiu, tai tuo labiau iš sąlygos $\mathbf{E}_\theta(h(T)|A) \equiv 0, \theta \in \Theta$, išplaukia, kad $h(T) = 0$ b. v. skirstinių $\mathcal{P}_A$ atžvilgiu.

I.2.38. Pilnoji ir pakankamoji statistika $T = X_1 + ... + X_n$. *Nurodymas.* Pasiremkite **I.2.37** pratimu.

I.2.39. Vidurkis $\mathbf{E}_\theta(X_{(n)} - X_{(1)} - (n-1)/(n+1)) \equiv 0, 0 < \theta < \infty$, nors ši funkcija nėra tapačiai lygi 0.

I.2.40. Tikėtinumo funkcija

$$L(a,b) = \prod_{i=1}^{n} \psi(X_i)/[\int_a^b \psi(x)dx]^n \mathbf{1}_{(a,\infty)(X_{(1)})} \mathbf{1}_{(-\infty,b)(X_{(n)})}.$$

Remiantis Neimano ir Fišerio faktorizacijos kriterijumi $\boldsymbol{T} = (X_{(1)}, X_{(n)})^T$ yra parametro $(a,b)^T$ pakankamoji statistika. Statistikos $\boldsymbol{T}$ tankio funkcija

$$g(x,y) = \frac{n(n-1)}{[\int_a^b \psi(x)dx]^n} \left[\int_x^y \psi(u)du\right]^{n-2} \psi(x)\psi(y), \quad a < x < y < b.$$

Tarkime,

$$\mathbf{E}_{a,b} h(X_{(1)}, X_{(n)}) = \int_a^b \int_a^y h(x,y) g(x,y) dxdy \equiv 0, \quad \forall \ a < b.$$

Diferencijuokime tapatybę

$$\int_a^b \int_a^y h(x,y) \left[\int_x^y \psi(u)du\right]^{n-2} \psi(x)\psi(y) dxdy \equiv 0$$

pagal b ir po to pagal a. Gausime

$$\left[\int_a^b \psi(u)du\right]^{n-2} \psi(a)\psi(b) h(a,b) \equiv 0, \quad \forall \ a < b.$$

Matome, kad funkcija $h(X_{(1)}, X_{(n)})$ b. v. lygi 0. Statistika $\boldsymbol{T} = (X_{(1)}, X_{(n)})^T$ yra pilnoji.

I.2.41. Randame

$$\mathbf{E}_\theta X = (1-\theta)^2 \sum_{k=1}^{\infty} k\theta^{k-1} = (1-\theta)^2/(1-\theta)^2 = 1, \quad 0 < \theta < 1.$$

Imdami $\psi(X) = X - 1$, gauname $\mathbf{E}_theta(\psi(X)) \equiv 0, \ \forall \ \ \theta \in (0,1)$. Tačiau $\psi(X)$ įgyja reikšmes $-1, 0, 1, 2, ...$. Taigi statistika X nėra pilnoji.

Iš sąlygos

$$\mathbf{E}_\theta(\psi(X)) = \theta\psi(0) + \sum_{k=1}^{\infty} \psi(k)(1 - 2\theta + \theta^2)\theta^{k-1} =$$

$$\psi(1) + \sum_{k=1}^{\infty} [\psi(k-1) - 2\psi(k) + \psi(k+1)]\theta^k \equiv 0 \quad \forall \ \ \theta \in (0,\ 1),$$

išplaukia, kad

$$\psi(1) = 0, \quad \psi(k) - 2\psi(k+1) + \psi(k+2) = 0, \quad k = 0, 1, 2, ...$$

Nuosekliai spręsdami gauname $\psi(2) = -\psi(0)$, $\psi(3) = -2\psi(0), ..., \psi(k) = -(k-1)\psi(0), ...$

Jeigu $\psi(0) \neq 0$, tai funkcija ψ neaprėžta. Jeigu $\psi(0) = 0$, tai ψ aprėžta ir $\psi(k) = 0, k = 0, 1, 2, ...$ Statistika X yra aprėžtai pilnoji.

I.2.42. a) Reikia įrodyti, kad imties $(X_1, ..., X_n)^T$ sąlyginis skirstinys, kai $S_n = X_1 + ... + X_n = t$ fiksuotas, nepriklauso nuo nežinomo parametro. Randame

$$\mathbf{P}_p\{X_1 = m_1, ..., X_n = m_n | S_n = t\} = \frac{\mathbf{P}_p\{X_1 = m_1\} \cdot ... \cdot \mathbf{P}_p\{X_n = m_n\}}{\mathbf{P}_p\{S_n = t\}} =$$

$$\frac{p^{m_1}(1-p)^{1-m_1} ... p^{m_n}(1-p)^{1-m_n}}{C_n^t p^t (1-p)^{n-t}} = \frac{1}{C_n^t},$$

kai $m_i = 0; 1, i = 1, ..., n$ ir $m_1 + ... + m_n = t$. Sąlyginis skirstinys nepriklauso nuo nežinomo parametro p. Taigi S_n yra pakankamoji parametro p statistika.

b) Tikėtinumo funkcija

$$L(p) = p^{\sum_i X_i}(1-p)^{n - \sum_i X_i}.$$

Remiantis faktorizacijos kriterijumi $S_n = X_1 + ... + X_n$ yra pakankamoji statistika.

I.2.43.(**I.2.42** pratimo tęsinys). a) Tarkime, kad

$$\mathbf{E}_p h(S_n) = \sum_{m=0}^{n} h(m) C_n^m p^m (1-p)^{n-m} \equiv 0, \quad 0 < p < 1.$$

Arba, pažymėjus $x = p/(1-p)$,

$$\sum_{m=0}^{n} h(m) C_n^m x^m \equiv 0, \quad 0 < x < \infty.$$

Kairėje tapatybės pusėje turime n-ojo laipsnio polinomą x atžvilgiu. Polinomas tapatingai lygus 0, kai koeficientai prie x laipsnių lygūs nuliui. Taigi $h(0) = h(1) = ... = h(n) = 0$. Funkcija $h(S_n)$ lygi nuliui su tikimybe 1.

b) Bernulio skirstinio tankis skaičiuojančiojo mato atžvilgiu priklauso eksponentinio tipo skirstinių šeimai; tankis kanonine forma yra

$$f(x|p) = \exp\{x\theta + \ln(1 + e^{\theta})\}, \quad \theta = \ln(p/(1 - p)).$$

Parametro θ kitimo sritis yra intervalas $(-\infty, \infty)$, taigi jai priklauso vidiniai taškai. Statistika S_n yra pilnoji.

I.2.44. (**I.2.42** pratimo tęsinys). a) Statistika $U(\mathbf{X}) = X_1X_2...X_i(1 - X_{i+1})...(1 - X_{i+j})$ įgyja reikšmę 1 su tikimybe $p^i(1 - p)^j$ ir reikšmę 0 priešingu atveju. Statistika $U(\mathbf{X})$ yra nepaslinktasis γ įvertinys.

$$\mathbf{E}(U(\mathbf{X})|S_n = t) = \mathbf{P}\{U(\mathbf{X}) = 1|S_n = t\}.$$

Pastarąją tikimybę galime rasti naudodami klasikinį tikimybės apibrėžimą. Elementariųjų įvykių skaičius lygus skaičiui skirtingų būdų, kuriais galima t simbolių 1 ir $n - t$ simbolių 0 išdėstyti į n vietas, t. y. C_n^t (žr. **I.2.42** pratimą). Palankių įvykių skaičius lygus skaičiui būdų, kuriais galima $t - i$ likusius simbolius 1 ir $n - t - j$ likusius simbolius 0 išdėstyti į likusias $n - i - j$ vietas, t. y. C_{n-i-j}^{t-i}. Taigi

$$\mathbf{E}[U(\mathbf{X})|S_n = t] = C_{n-i-j}^{t-i}/C_n^t = t^{[i]}(n - t)^{[j]}/n^{[i+j]},$$

čia $m^{[k]} = m(m - 1)...(m - k + 1)$. Parametro γ NMD įvertinys yra

$$\hat{\gamma} = S_n^{[i]}(n - S_n)^{[j]}/n^{[i+j]}.$$

Pavyzdžiui, parametrų p; p^2;p(1-p) NMD įvertiniai yra

$$S_n/n; \quad S_n(S_n - 1)/(n(n - 1)); \quad S_n(n - S_n)/(n(n - 1)).$$

b) Tegu h yra funkcija, tenkinanti tapatybę

$$\mathbf{E}_p h(S_n) = \sum_{m=0}^{n} h(m)C_n^m p^m (1 - p)^{n-m} \equiv p^i(1 - p)^j, \quad 0 < p < 1.$$

Padauginkime dešiniąją tapatybės pusę iš 1:

$$\sum_{m=0}^{n} h(m)C_n^m p^m(1 - p)^{n-m} \equiv p^i(1 - p)^j \sum_{k=0}^{n-i-j} C_{n-i-j}^k p^k(1 - p)^{n-i-j-k} =$$

$$\sum_{k=0}^{n-i-j} C_{n-i-j}^k p^{k+i}(1 - p)^{n-i-k} = \sum_{m=i}^{n-j} C_{n-i-j}^{m-i} p^m(1 - p)^{n-m}.$$

Sulyginę koeficientus prie $p^m(1 - p)^{n-m}$, kai $m = i, ..., n - j$, gauname, kad

$$h(m) = C_{n-i-j}^{m-i}/C_n^m = m^{[i]}(n - m)^{[j]}/n^{[i+j]}.$$

Gauname tą patį NMD įvertinį, kaip ir p. a).

I.2.45. (**I.2.42** pratimo tęsinys). a) *Nurodymas.* Pasiremkime **I.2.44** pratimu.
b) Tarkime, kad parametro p^k NMD įvertinys egzistuoja. Tada egzistuoja tokia funkcija $h(m)$, kad

$$\sum_{m=0}^{n} h(m)C_n^m p^m(1-p)^{n-m} \equiv p^k.$$

Kairėje tapatybės pusėje yra n-ojo laipsnio polinomas p atžvilgiu, o dešinėje – polinomas didesnio negu n laipsnio. Tokia tapatybė negalima.

I.2.46. (**I.2.42** pratimo tęsinys). $38/50 + 12 \cdot 11 \cdot 10/(50 \cdot 49 \cdot 48)$; $12 \cdot 38/(50 \cdot 49)$; $12!38!/50!$;0; 0.

I.2.47. a) Imties $(X_1, ..., X_n)^T$ sąlyginis skirstinys, kai $S_n = X_1 + ... + X_n = t$ fiksuotas, nusakomas tikimybėmis

$$\mathbf{P}_\lambda\{X_1 = m_1, ..., X_n = m_n | S_n = t\} =$$

$$= \frac{\mathbf{P}_\lambda\{X_1 = m_1\} \cdot ... \cdot \mathbf{P}_\lambda\{X_n = m_n\}}{\mathbf{P}_\lambda\{S_n = t\}} = \frac{t!}{m_1!...m_n!}\frac{1}{n^t},$$

kai $m_i \geq 0$ ir $m_1 + ... + m_n = t$. Sąlyginis skirstinys yra polinominis $\mathcal{P}_n(t; (1/n, ..., 1/n))$. Sąlyginis skirstinys nepriklauso nuo nežinomo parametro λ. Taigi S_n yra parametro λ pakankamoji statistika.

b) Tikėtinumo funkcija

$$L(\lambda) = \lambda^{\sum_i X_i} e^{-n\lambda}/(X_1!...X_n!).$$

Remiantis faktorizacijos kriterijumi $S_n = X_1 + ... + X_n$ yra pakankamoji statistika.

I.2.48. (**I.2.47** pratimo tęsinys). a) Tarkime, kad

$$\mathbf{E}_\lambda(h(S_n)) = \sum_{k=0}^{\infty} h(k)(n\lambda)^k e^{-n\lambda}/k! \equiv 0, \quad 0 < \lambda < \infty.$$

Arba

$$\sum_{k=0}^{\infty} h(k)n^k\lambda^k/k! \equiv 0, \quad 0 < \lambda < \infty.$$

Laipsninė eilutė tapatingai lygi 0, kai koeficientai prie λ laipsnių lygūs nuliui. Taigi $h(0) = h(1) = ... = h(k) = ... = 0$. Funkcija $h(S_n)$ lygi nuliui su tikimybe 1.

b) Puasono skirstinio tankis skaičiuojančiojo mato atžvilgiu priklauso eksponentinio tipo skirstinių šeimai; tankis kanonine forma yra

$$f(x|\lambda) = \exp\{x\theta - e^\theta\}/x!, \quad \theta = \ln(p/(1-p)).$$

Parametro $\theta = \ln\lambda$ kitimo sritis yra intervalas $(-\infty, \infty)$, taigi jai priklauso vidiniai taškai. Statistika S_n yra pilnoji.

I.2.49. (**I.2.47** pratimo tęsinys). a) Atžvilgiu h sprendžiame funkcinę lygtį

$$\mathbf{E}h(S_n) = \sum_{k=0}^{\infty} h(k)(n\lambda)^k e^{-n\lambda}/k! \equiv \lambda^m,$$

arba
$$\sum_{k=0}^{\infty} h(k)n^k\lambda^k/k! \equiv \lambda^m e^{n\lambda} = \sum_{j=0}^{\infty} \lambda^{j+m}n^j/j! = \sum_{k=m}^{\infty} \lambda^k n^{k-m}/(k-m)!\ .$$
Sulyginę koeficientus prie vienodų laipsnių λ^k, gauname $h(0) = h(1) = ... = h(m-1) = 0$, $h(k) = k^{[m]}/n^m$, kai $k \geq m$.

Taigi parametro $\gamma_m(\lambda) = \lambda^m$ NMD įvertinys yra $\gamma_m(\lambda) = S_n^{[m]}/n^m = S_n(S_n - 1)...(S_n - m + 1)/n^m$.

b) Tarkime priešingai, kad parametro $1/\lambda$ NMD įvertinys egzistuoja. Tada egzistuoja funkcija h, tenkinanti tapatybę
$$\sum_{k=0}^{\infty} h(k)(n\lambda)^k/k! e^{-n\lambda} \equiv 1/\lambda,$$
arba
$$\sum_{k=0}^{\infty} h(k)n^k\lambda^k/k! \equiv \sum_{j=0}^{\infty} n^j\lambda^{j-1}/j!, \quad \forall\ \lambda > 0.$$
Kairėje pusėje yra laipsninė eilutė atžvilgiu λ, kai laipsnio rodikliai yra $0, 1, 2, ...$, o dešinėje tapatybės pusėje laipsnio rodikliai yra $-1, 0, 1, ...$ Tokia tapatybė negalima. Prielaida buvo neteisinga.

I.2.50. (**I.2.47** pratimo tęsinys). a) Atžvilgiu h sprendžiame funkcinę lygtį
$$\mathbf{E}_\lambda h(S_n) = \sum_{k=0}^{\infty} h(k)(n\lambda)^k e^{-n\lambda}/k! \equiv \lambda^m e^{-\lambda}/m!,$$
arba
$$\sum_{k=0}^{\infty} h(k)n^k\lambda^k/k! \equiv \sum_{j=0}^{\infty} \lambda^{j+m}(n-1)^j/(j!m!) = \sum_{k=m}^{\infty} \lambda^k(n-1)^{k-m}/(j!(k-m)!).$$

Sulyginę koeficientus prie vienodų laipsnių λ^k, gauname $h(0) = h(1) = ... = h(m - 1) = 0$ ir
$$h(m) = C_k^m(1/n)^m(1-1/n)^{k-m}, \quad k \geq m.$$

Taigi, kai $S_n \geq m$, tai puasoninė tikimybė $\pi_m(\lambda)$ vertinama binominio skirstinio $B(S_n,\, 1/n)$ tikimybe $C_{S_n}^m(\frac{1}{n})^m(1-\frac{1}{n})^{S_n-m}$.

b) Atžvilgiu h sprendžiame funkcinę lygtį
$$\mathbf{E}_\lambda h(S_n) = \sum_{k=0}^{\infty} h(k)(n\lambda)^k e^{-n\lambda}/k! \equiv e^{\lambda(s-1)},$$
arba
$$\sum_{k=0}^{\infty} h(k)n^k\lambda^k/k! \equiv \sum_{j=0}^{\infty} \lambda^j(n+s-1)^j/j!, \quad h(k) = (1 - \frac{1}{n} + \frac{s}{n})^k.$$
Generuojančios funkcijos $g(s)$ NMD įvertinys yra binominio skirstinio generuojančioji funkcija
$$\hat{g}(s) = (1 - \frac{1}{n} + \frac{s}{n})^{S_n}.$$

I.2.51. A. d. X skirstinio tankis skaičiuojančiojo mato atžvilgiu

$$f(x|p) = pq^x = \exp\{x\theta + \ln(1 - e^\theta)\}, \quad -\infty < \theta = \ln q < 0$$

priklauso vienparametrei eksponentinių skirstinių šeimai. Pakankamoji statistika $S_n = X_1 + ... + X_n$. Kadangi parametro θ kitimo sritis turi vidinių taškų, tai S_n yra pilnoji statistika.

I.2.52. (**I.2.51** pratimo tęsinys). Imties $(X_1, ..., X_n)^T$ sąlyginis skirstinys, kai $S_n = X_1 + ... + X_n = t$ fiksuotas, nusakomas tikimybėmis

$$\mathbf{P}_p\{X_1 = m_1, ..., X_n = m_n | S_n = t\} = \frac{\mathbf{P}_p\{X_1 = m_1\} \cdot ... \cdot \mathbf{P}_p\{X_n = m_n\}}{\mathbf{P}_p\{S_n = t\}} = \frac{1}{C^t_{n+t-1}},$$

kai $m_i \geq 0$ ir $m_1 + ... + m_n = t$. Sąlyginis skirstinys nepriklauso nuo nežinomo parametro. Statistika S_n yra pakankamoji parametro p statistika.

I.2.53. (**I.2.51** pratimo tęsinys). Tegu h yra funkcija, tenkinanti tapatybę

$$\mathbf{E}_p h(S_n) = \sum_{k=0}^{\infty} h(k) C^k_{n+k-1} p^n q^k \equiv p^m q^l, \quad 0 < p < 1.$$

Padauginkime dešiniąją tapatybės pusę iš 1:

$$\sum_{k=0}^{\infty} h(k) C^k_{n+k-1} p^n q^k \equiv p^m q^l \sum_{j=0}^{\infty} C^j_{n-m+j-1} p^{n-m} q^j =$$

$$\sum_{k=l}^{\infty} C^{k-l}_{n+k-m-l-1} p^n q^k.$$

Sulyginę koeficientus prie $p^n q^k$, kai $k \geq l$, gauname

$$h(k) = C^{k-l}_{n+k-m-l-1} / C^k_{n+k-1}, \quad k \geq l.$$

Parametro $\gamma = p^m q^l$ NMD įvertinys

$$\hat{\gamma} = C^{S_n - l}_{n+S_n-m-l-1} / C^{S_n}_{n+S_n-1}, \quad S_n \geq l.$$

I.2.54. (**I.2.51** pratimo tęsinys). Kadangi $\mathbf{E}_p S_n = nq/p$, tai parametro q/p NMD įvertinys yra $\bar{X} = S_n/n$. Kitų parametrų įvertiniai gaunami iš **I.2.53** pratimo

$$\hat{p} = \frac{n-1}{n + S_n - 1}, \quad \hat{p^2} = \frac{(n-1)(n-2)}{(n + S_n - 1)(n + S_n - 2)}, \quad \hat{pq} = \frac{S_n(n-1)}{(n + S_n - 1)(n + S_n - 2)}.$$

I.2.55. (**I.2.51** pratimo tęsinys). Funkcijos h atžvilgiu spręsdami funkcinę lygtį $\mathbf{E}_p(h(X)) \equiv p \ln p$, gauname

$$\mathbf{E}_p(h(X)) = p \sum_{k=1}^{\infty} h(k) q^{k-1} \equiv p \ln(1 - q) = -p \sum_{k=1}^{\infty} q^k / k.$$

Sulyginę koeficientus prie vienodų q^k laipsnių gausime $h(k) = -1/(k-1), k > 1$. Taigi parametro $\gamma = p \ln p$ NMD įvertinys yra $\hat{\gamma} = -1/(X-1), X > 1$. Įvertinio realizacija yra $-1/13$.

I.2.56. a) Tikėtinumo funkcija

$$L(\boldsymbol{\theta}) = W(\mathbf{X}) \exp\{X_1 \ln p_1 + X_2 \ln p_2 + ... + X_k \ln p_k\}, \quad W(\mathbf{X}) = \frac{n!}{X_1!...X_k!},$$

čia $X_k = n - X_1 - ... - X_{k-1}, p_k = 1 - p_1 - ... - p_{k-1}$; priklauso eksponentinio tipo skirstinių šeimai. Statistika $\boldsymbol{T} = (X_1, ...X_{k-1})^T$ yra pakankamoji. Kadangi parametro $(\ln p_1, ..., \ln p_{k-1})^T$ kitimo sritis turi vidinių taškų, tai statistika $\boldsymbol{T}$ yra pilnoji.

b) A. v. $(X_i, X_j)^T$ generuojančioji funkcija

$$\psi(s_1, s_2) = \mathbf{E}(s_1^{X_i} s_2^{X_j} | p_i, p_j) = \sum_{k_1, k_2} \frac{n!}{k_1! k_2! (n - k_1 - k_2)!} (p_i e^{s_1})^{k_1} ((p_j e^{s_2})^{k_2})$$

$$(1 - p_i - p_j)^{n - k_1 - k_2} = (p_i s_1 + p_j s_2 + 1 - p_i - p_j)^n.$$

Imdami funkcijos ψ m-ąją išvestinę pagal s_1 taške $s_1 = s_2 = 1$, gauname faktorialinį momentą $\mathbf{E}(X_i^{[m]} | p_i, p_j) = n^{[m]} p_i^m$; imdami ψ k-ąją išvestinę pagal s_1 ir l-ąją išvestinę pagal s_2 taške $s_1 = s_2 = 1$, gauname $\mathbf{E}(X_i^{[k]} X_j^{[l]} | p_i, p_j) = n^{[k+l]} p_i^k p_j^l$. Taigi parametrų p_i^m ir $p_i^k p_j^l$ NMD įvertiniai yra $X_i^{[m]}/n^{[m]}$ ir $X_i^{[k]} X_j^{[l]}/n^{[k+l]}$.

I.2.57. Tikėtinumo funkcija

$$L(\theta) = 1/\theta^n, \quad \text{kai} \quad 0 < X_1, ..., X_n < \theta,$$

arba

$$L(\theta) = (1/\theta^n) \mathbf{1}_{(0,\theta)}(X_{(n)}).$$

Remiantis faktorizacijos kriterijumi $T = X_{(n)}$ yra pakankamoji statistika.

Statistikos T tankio funkcija

$$f_n(x|\theta) = n x^{n-1}/\theta^n, \quad 0 < x < \theta.$$

Tarkime, kad funkcijos $h(X_{(n)})$ vidurkis tapatingai lygus 0:

$$\mathbf{E}_\theta h(X_{(n)}) = \frac{n}{\theta^n} \int_0^\theta h(x) x^{n-1} dx \equiv 0, \quad 0 < \theta < \infty.$$

Diferencijuodami tapatybę

$$\int_0^\theta h(x) x^{n-1} dx \equiv 0, \quad 0 < \theta < \infty$$

pagal θ gausime $h(\theta)\theta^{n-1} \equiv 0$, kai $0 < \theta < \infty$. Taigi funkcija $h(X_{(n)})$ lygi 0 su tikimybe 1. Statistika T yra pilnoji.

I.2.58. (**I.2.57** pratimo tęsinys). Randame

$$\mathbf{E}_\theta X_{(n)} = \frac{n}{\theta^n} \int_0^\theta x x^{n-1} dx = \frac{n}{n+1}\theta, \quad \mathbf{E}_\theta \hat{\theta} = \mathbf{E}_\theta((n+1) X_{(n)}/n) = \theta.$$

Kadangi $\hat{\theta}$ yra pilnosios ir pakankamosios statistikos $X_{(n)}$ funkcija ir $\mathbf{E}_\theta \hat{\theta} = \theta$, tai $\hat{\theta}$ yra parametro θ NMD įvertinys.

I.2.59. Tikėtinumo funkcija

$$L(\theta) = \mathbf{1}_{(\theta,\infty)}(X_{(1)})\mathbf{1}_{(-\infty,\theta+1)}(X_{(n)}).$$

Pakankamoji statistika $\boldsymbol{T} = (X_{(1)}, X_{(n)})^T$ yra dvimatė.

Randame $\mathbf{E}_\theta X_{(1)} = \theta + 1/(n+1)$, $\mathbf{E}_\theta X_{(n)} = \theta + 1 - 1/(n+1)$. Tokiu būdu vidurkis $\mathbf{E}_\theta(X_{(n)} - X_{(1)} - (n-1)/(n+1)) \equiv 0$, nors pati funkcija nėra tapatingai lygi 0. Statistika $\boldsymbol{T}$ nėra pilnoji.

I.2.60. a) Tikėtinumo funkcija

$$L(\theta_1, \theta_2) = \frac{1}{(\theta_2 - \theta_1)^n}\mathbf{1}_{(\theta_1,\infty)}(X_{(1)})\mathbf{1}_{(-\infty,\theta_2)}(X_{(n)}.$$

Remiantis faktorizacijos kriterijumi statistika $\boldsymbol{T} = (X_{(1)}, X_{(n)})^T$ yra pakankamoji.

Tarkime, kad funkcijos $h(X_{(1)}, X_{(n)})$ vidurkis tapatingai lygus 0:

$$\mathbf{E}_{\theta_1,\theta_2} h(X_{(1)}, X_{(n)}) = \frac{n(n-1)}{(\theta_2 - \theta_1)^n}\int_{\theta_1}^{\theta_2}\int_{\theta_1}^{y} h(x,y)(y-x)^{n-2}dxdy \equiv 0 \quad \theta_1 < \theta_2.$$

Diferencijuokime tapatybę

$$\int_{\theta_1}^{\theta_2}\int_{\theta_1}^{y} h(x,y)(y-x)^{n-2}dxdy \equiv 0 \quad \theta_1 < \theta_2$$

pagal θ_2, po to pagal θ_1. Gauname $h(\theta_1, \theta_2)(\theta_2 - \theta_1)^{n-2} \equiv 0$ su visais $\theta_1 < \theta_2$, t. y. funkcija $h(X_{(1)}, X_{(n)}) \equiv 0$ b. v. Statistika $\boldsymbol{T}$ yra pilnoji.

b) Kadangi $\hat{\gamma}_1$ ir $\hat{\gamma}_2$ yra pilnosios ir pakankamosios statistikos $\boldsymbol{T}$ funkcijos, tai pakanka patikrinti jų nepaslinktumą. Dėl paprastumo atlikime keitimą $Y_i = (X_i - \theta_1)/(\theta_2 - \theta_1) \sim U(0,\ 1)$. Tada

$$\mathbf{E}_{\theta_1,\theta_2}((X_{(n)} + X_{(1)})/2) = \mathbf{E}((Y_{(n)} + Y_{(1)})/2)(\theta_2 - \theta_1) + \theta_1,$$

$$\boldsymbol{V}_{\theta_1,\theta_2}((X_{(n)} + X_{(1)})/2) = \boldsymbol{V}((Y_{(n)} + Y_{(1)})/2)(\theta_2 - \theta_1)^2;$$

$$\mathbf{E}_{\theta_1,\theta_2}((X_{(n)} - X_{(1)})\frac{n+1}{n-1}) = \mathbf{E}((Y_{(n)} - Y_{(1)})\frac{n+1}{n-1})(\theta_2 - \theta_1),$$

$$\boldsymbol{V}_{\theta_1,\theta_2}((X_{(n)} - X_{(1)})\frac{n+1}{n-1}) = \boldsymbol{V}((Y_{(n)} - Y_{(1)})\frac{n+1}{n-1})(\theta_2 - \theta_1)^2.$$

Randame (žr. **I.2.10** ir **I.2.18** pratimus)

$$\mathbf{E}((Y_{(n)} + Y_{(1)})/2) = 1/2, \quad \mathbf{E}_{\theta_1,\theta_2}((X_{(n)} + X_{(1)})/2) = \gamma_1;$$

$$\boldsymbol{V}((Y_{(n)} + Y_{(1)})/2) = \frac{1}{2(n+1)(n+2)}, \quad \boldsymbol{V}_{\theta_1,\theta_2}((X_{(n)} + X_{(1)})/2) = \frac{\gamma_2^2}{2(n+1)(n+2)};$$

$$\mathbf{E}((Y_{(n)} - Y_{(1)})) = \frac{n-1}{n+1}, \quad \mathbf{E}_{\theta_1,\theta_2}\hat{\gamma}_2 = \theta_2 - \theta_1 = \gamma_2;$$

$$\boldsymbol{V}((Y_{(n)} - Y_{(1)})) = \frac{2(n-1)}{(n+1)^2(n+2)}, \quad \boldsymbol{V}_{\theta_1,\theta_2}\hat{\gamma}_2 = \frac{2\gamma_2^2}{(n+2)(n-1)}.$$

I.2.61. Tikėtinumo funkcija

$$L(\theta) = \frac{1}{(2\theta)^n}\mathbf{1}_{(-\theta,\infty)}(X_{(1)})\mathbf{1}_{(-\infty,\theta)}(X_{(n)}) = \frac{1}{(2\theta)^n}\mathbf{1}_{(-\infty,\theta)}(\max(-X_{(1)}, X_{(n)})).$$

Remiantis faktorizacijos kriterijumi $T = \max(-X_{(1)}, X_{(n)})$ yra pakankamoji statistika. Tegu $Y_i = |X_i| \sim U(0,\theta)$, $T = Y_{(n)}$. Statistikos T pilnumas įrodytas **I.2.57** pratime. Ten pat įrodyta, kad $\hat{\theta} = (n+1)T/n$ yra parametro θ NMD įvertinys.

I.2.62. a) A. d. X tankio funkcija

$$f(x|\mu,\sigma) = (2\pi)^{-1/2}\exp\{-\frac{1}{2\sigma^2}x^2 + \frac{\mu}{\sigma^2}x - \frac{\mu^2}{2\sigma^2} - \frac{1}{2}\ln\sigma^2\}$$

priklauso dviparametrei eksponentinio tipo skirstinių šeimai. Iš tankio pavidalo išplaukia, kad $\boldsymbol{U}(\mathbf{X}) = (\sum_i X_i, \sum_i X_i^2)^T$ yra pakankamoji statistika. Tada statistika $\boldsymbol{T} = \boldsymbol{T}(\boldsymbol{U}(\mathbf{X})) = (\bar{X}, s^2)^T$ irgi pakankama. Parametro $(-1/(2\sigma^2), \mu/\sigma^2)^T$ kitimo sritis yra pusplokštumė, todėl jai priklauso vidiniai taškai. Statistika $\boldsymbol{T}$ ne tik pakankamoji, bet ir pilnoji.

b) Kadangi $\mathbf{E}_{\mu,\sigma}\bar{X} = \mu, \mathbf{E}_{\mu,\sigma}s^2 = \sigma^2$, tai $\bar{X}$ ir s^2 yra parametrų μ ir σ^2 NMD įvertiniai.

I.2.63. (**I.2.62** pratimo tęsinys). Žinome, kad $s^2(n-1)/\sigma^2 \sim \chi^2(n-1)$ (žr. [2], 2.5.1 teoremą). Tada

$$\mathbf{E}_{\mu,\sigma}s = \frac{\sigma}{\sqrt{n-1}}\mathbf{E}\sqrt{\chi^2_{n-1}} = \frac{\sigma}{\sqrt{n-1}}\frac{1}{2^{(n-1)/2}\Gamma((n-1)/2)}\int_0^\infty x^{n/2-1}e^{-x/2}dx =$$

$$\frac{\sigma}{\sqrt{n-1}}\frac{\sqrt{2}\Gamma(n/2)}{\Gamma((n-1)/2)} = \sigma M_{n-1}; \quad \mathbf{E}_{\mu,\sigma}(s/M_{n-1}) = \sigma.$$

I.2.64. (**I.2.62** pratimo tęsinys). a) Tikėtinumo funkcija

$$L(\sigma^2) = (2\pi)^{-n/2}\exp\{-\frac{1}{2\sigma^2}\sum_i(X_i-\mu)^2 - \frac{n}{2}\ln\sigma^2\}.$$

Matome, kad pilnoji ir pakankamoji statistika yra s_0^2. Be to, $s_0^2 n/\sigma^2 \sim \chi^2(n)$; t. y. $\mathbf{E}_\sigma s_0^2 = \sigma^2, \mathbf{E}_\sigma(s_0/M_n) = \sigma$.

b) Tikėtinumo funkcija

$$L(\mu) = (2\pi\sigma)^{-n/2}\exp\{-\frac{1}{2\sigma^2}\sum_i X_i^2\}\exp\{\frac{\mu n\bar{X}}{\sigma^2} - \frac{n\mu^2}{2\sigma^2}\}.$$

Pilnoji ir pakankamoji statistika $\bar{X}$; $\mathbf{E}_\mu\bar{X} = \mu$.

I.2.65. a) Tikėtinumo funkcija

$$L(\mu,\lambda) = \lambda^n e^{n\mu\lambda}e^{-\lambda\sum_i X_i}\mathbf{1}_{(\mu,\infty)}(X_{(1)}).$$

Remiantis faktorizacijos kriterijumi $(X_{(1)}, \sum_i X_i)^T$ yra pakankamoji. Pakankamoji yra ir jai ekvivalenti statistika $(X_{(1)}, T_2(\mathbf{X}))^T$.

b) Kadangi $X_{(1)}$ ir T_2 yra nepriklausomi (žr. **I.1.27** pratimą), tai jos pilnumą galima įrodyti nagrinėjant šias statistikas atskirai.

c) Kadangi $\hat{\gamma}_1$ ir $\hat{\gamma}_2$ yra pilnos ir pakankamos statistikos funkcijos, tai pakanka patikrinti jų nepaslinktumą. Randame

$$\mathbf{E}_{\mu,\lambda}\hat{\gamma}_2 = \frac{1}{\lambda^m \Gamma(n+m-1)} \int_0^\infty x^m x^{n-2} e^{-x} dx = \frac{1}{\lambda^m}.$$

Statistikos $X_{(1)}$ tankio funkcija (žr. **I.1.25** pratimą)

$$f(x|\mu,\lambda) = n\lambda e^{-n\lambda(x-\mu)}, \quad \mu < x < \infty.$$

Skaičiuojame vidurkį

$$\mathbf{E}_{\mu,\lambda} X_{(1)}^l = n\lambda \int_\mu^\infty x^l e^{-n\lambda(x-\mu)} dx = n\lambda \int_0^\infty (t+\mu)^l e^{-n\lambda t} dt.$$

Integruodami dalimis gauname

$$\mathbf{E}_{\mu,\lambda} X_{(1)}^l = \mu^l + l \int_0^\infty (t+\mu)^{l-1} e^{-n\lambda t} dt = \mu^l + l/(n\lambda)\mathbf{E}_{\mu,\lambda} X_{(1)}.$$

Kadangi $\mathbf{E}_{\mu,\lambda} T_2/(n-1) = 1/\lambda$, o $X_{(1)}^{l-1}$ ir T_2 nepriklausomi a. d., tai antrasis dėmuo yra vidurkis a. d. $lX_{(1)}^{l-1} T_2/(n(n-1))$. Atėmę šį a. d. iš $X_{(1)}^l$ gausime statistiką, kurios vidurkis lygus μ^l.

I.2.66. (**I.2.65** pratimo tęsinys). a) Tikėtinumo funkcija

$$L(\lambda) = \lambda^n e^{-\lambda \sum_i (X_i - \mu)},$$

taigi $T = \sum_i (X_i - \mu)$ yra pilnoji ir pakankamoji statistika; $X_i - \mu \sim \mathcal{E}(\lambda)$; $\mathbf{E}_\lambda(\bar{X} - \mu) = 1/\lambda$.

b) Tikėtinumo funkcija

$$L(\mu) = \lambda^n e^{-\lambda \sum_i X_i} e^{n\lambda\mu} \mathbf{1}_{(\mu,\infty)}(X_{(1)}),$$

taigi $X_{(1)}$ yra pilnoji ir pakankamoji statistika. **I.2.64** pratime parodėme (imkime $l = 1$), kad $\mathbf{E}_\mu X_{(1)} = \mu + 1/(n\lambda)$. Taigi $\mathbf{E}_\mu(X_{(1)} - 1/(n\lambda)) = \mu$.

I.2.67. Tikėtinumo funkcija

$$L(\lambda,\eta) = \exp\{-\lambda \sum_i X_i + (\eta - 1)\sum_i \ln X_i + n\eta \ln\lambda - n\ln(\Gamma(\eta))\}.$$

Skirstinys priklauso dviparametrei eksponentinio tipo skirstinių šeimai; pakankamoji statistika $\boldsymbol{T} = (\sum_i X_i, \sum_i \ln X_i)^T$. Kadangi parametro kitimo sričiai priklauso vidiniai taškai, tai statistika $\boldsymbol{T}$ pilnoji. A. d. $\bar{X}$ yra pilnosios ir pakankamosios statistikos $\boldsymbol{T}$ funkcija ir $\mathbf{E}(\bar{X}|\lambda,\eta) = \eta/\lambda$, todėl $\bar{X}$ yra parametro η/λ NMD įvertinys.

I.2.68. (**I.2.67** pratimo tęsinys). Kai η žinomas, tikėtinumo funkcija yra

$$L(\lambda) = \frac{1}{(\Gamma(\eta))^n}(\prod_i X_i)^{\eta-1} e^{-\lambda \sum_i X_i + n\eta \ln \lambda}.$$

Pilnoji ir pakankamoji statistika $T = \sum_i X_i \sim G(\lambda, n\eta)$. Randame (atlikę kintamųjų keitimą $\lambda x = y$)

$$\mathbf{E}_\lambda \hat{\gamma} = \frac{\lambda^{n\eta}}{\Gamma(n\eta - k)} \int_0^\infty x^{-k} x^{n\eta-1} e^{-\lambda x} dx = \lambda^k.$$

I.2.69. (**I.2.67** pratimo tęsinys). Kai λ žinomas, tikėtinumo funkcija yra

$$L(\eta) = e^{-\lambda \sum_i X_i} e^{(\eta-1)\sum_i \ln X_i + n\eta \ln \lambda - n \ln \Gamma(\eta)}.$$

Pilnoji ir pakankamoji statistika $T = \sum_i \ln X_i$. Randame

$$\mathbf{E}_\eta(\ln X_i) = \frac{\lambda^\eta}{\Gamma(\eta)} \int_0^\infty \ln x \, x^{\eta-1} e^{-\lambda x} dx = \frac{1}{\Gamma(\eta)} \int_0^\infty (\ln t - \ln \lambda) t^{\eta-1} e^{-t} dt = \frac{\Gamma'(\eta)}{\Gamma(\eta)} - \ln \lambda.$$

I.2.70. Reikia suvidurkinti nepaslinktąjį įvertinį $h(X_1)$ pakankamosios statistikos $\boldsymbol{T}(\mathbf{X})$ atžvilgiu (žr. [2], 3.3.5 skyrelį). Gauname

$$\hat{\tau}(\boldsymbol{T}) = \mathbf{E}(h(X_1)|\boldsymbol{T}) = \mathbf{E}(h\left(\frac{X_1 - \hat{\theta}_1}{\hat{\theta}_2}\hat{\theta}_2 + \hat{\theta}_1\right)|\boldsymbol{T}) =$$

$$\mathbf{E}(h(U(\mathbf{X})\hat{\theta}_2 + \hat{\theta}_1)|\boldsymbol{T}).$$

Remiantis Basu teorema [9] statistikos $U(\mathbf{X})$ ir $\boldsymbol{T} = \boldsymbol{T}(\mathbf{X})$ yra nepriklausomos. Taigi

$$\hat{\tau}(\boldsymbol{T}) = \int_{-\infty}^{\infty} h(u\hat{\theta}_2 + \hat{\theta}_1) g(u) du.$$

I.2.71. Nepaslinktasis parametro γ_1 įvertinys yra $h(X_1) = \mathbf{1}_{(-\infty, y)}(X_1)$, nes

$$\mathbf{E}_\mu h(X_1) = \mathbf{P}_\mu\{X_1 \leq y\} = \Phi(y - \mu).$$

Parametro μ pilnoji ir pakankamoji statistika yra $\bar{X}$. **I.2.70** pratime imkime

$$U = U(\mathbf{X}) = X_1 - \bar{X}, \ \ \hat{\theta}_1 = \bar{X}, \ \ \hat{\theta}_2 = 1.$$

Gauname

$$\hat{\gamma_1} = \mathbf{E}(h(X_1)|T) = \mathbf{E}(h(U + \bar{X})|T) = \int_{-\infty}^{\infty} h(u + \bar{X}|T) g(u) du =$$

$$\int_{-\infty}^{y - \bar{X}} g(u) du = \Phi((y - \bar{X})/\sqrt{1 - 1/n}),$$

nes $U \sim N(0, 1 - 1/n)$. Diferencijuodami tapatybės

$$\mathbf{E}_\mu \hat{\gamma}_1 = \int_{-\infty}^{\infty} \Phi((y - \bar{X})/\sqrt{1 - 1/n}) \sqrt{n} \varphi(\sqrt{n}(x - \mu)) dx \equiv \Phi(y - \mu)$$

abi puses pagal y įsitikiname, kad parametro γ_2 NMD įvertinys yra

$$\hat{\gamma}_2 = \varphi((y - \bar{X})/\sqrt{1-1/n})/\sqrt{1-1/n}.$$

I.2.72. Nepaslinktasis parametro γ_1 įvertinys yra $h(X_1) = \mathbf{1}_{(-\infty,y)}(X_1)$, nes

$$\mathbf{E}_\sigma h(X_1) = \mathbf{P}_\sigma\{X_1 < y\} = \Phi(y/\sigma).$$

Parametro σ pilnoji ir pakankamoji statistika yra $S^2 = X_1^2 + ... + X_n^2$. **I.2.70** pratime imkime $U = U(\mathbf{X}) = X_1/S$, $\hat{\theta_1} = 0, \hat{\theta_2} = S$. Gauname

$$\hat{\gamma_1} = \mathbf{E}(h(X_1)|T) = \mathbf{E}(h(US)|T) = \int_{-\infty}^{y/S} g(u)du.$$

Reikia rasti statistikos $U(\mathbf{X})$ tankio funkciją $g(u)$. Kadangi $X_i^2/\sigma^2 \sim \chi_1^2$, $X_2^2+...+X_n^2 \sim \chi_{n-1}^2$, tai $U^2(\mathbf{X}) = X_1^2/(X_1^2+...+X_n^2) \sim Be(1/2,(n-1)/2)$. Statistika $U(\mathbf{X})$ simetriška. Kai $u > 0$, pasiskirstymo funkcija

$$G(u) = \mathbf{P}\{U < u\} = \frac{1}{2} + \frac{1}{2}\mathbf{P}\{U^2 < u^2\} =$$

$$\frac{1}{2} + \frac{\Gamma(n/2)}{2\Gamma(1/2)\Gamma((n-1)/2)} \int_0^{u^2} v^{-1/2}(1-v)^{(n-3)/2}dv.$$

Diferencijuodami pagal u gausime tankio funkciją

$$g(u) = G'(u) = \frac{\Gamma(n/2)}{\Gamma(1/2)\Gamma((n-1)/2)}(1-u^2)^{(n-3)/2}, \quad |u| < 1.$$

Gauname NMD įvertinį

$$\hat{\gamma_1} = \int_{-1}^{y/S} g(u)du = \frac{\Gamma(n/2)}{\Gamma(1/2)\Gamma((n-1)/2)} \int_{-1}^{y/S} (1-u^2)^{(n-3)/2}du, \quad |y/S| < 1.$$

Tankio funkcijos įvertinys $\hat{\gamma}_2 = g(y/S)/S$.

I.2.73. Nepaslinktasis parametro γ įvertinys yra $h(X_1) = \mathbf{1}_{(-\infty,y)}(X_1)$, nes

$$\mathbf{E}_{(\mu,\sigma)} h(X_1) = \mathbf{P}_{(\mu,\sigma)}\{X_1 < y\} = \Phi((y-\mu)/\sigma).$$

Parametro $(\mu,\sigma)^T$ pilnoji ir pakankamoji statistika yra $\boldsymbol{T} = (\bar{X}, S^2)^T$, $S^2 = \sum_i (X_i - \bar{X})^2$. **I.2.70** pratime imkime $U = U(\mathbf{X}) = (X_1 - \bar{X})/S$, $\hat{\theta_1} = \bar{X}, \hat{\theta_2} = S$. Statistikos $U(\mathbf{X})$ skirstinys nuo nežinomų parametrų nepriklauso, todėl statistikos $U(\mathbf{X})$ ir $\boldsymbol{T}$ nepriklausomos. Gauname

$$\hat{\gamma_1} = \mathbf{E}(h(X_1)|\boldsymbol{T}) = \mathbf{E}(h(US + \bar{X})|\boldsymbol{T}) = \int_{-\infty}^{(y-\bar{X})/S} g(u)du.$$

Lieka rasti statistikos $U(\mathbf{X}) = (X_1 - \bar{X})/S$ tankio funkciją $g(u)$.

Atlikime ortonormuotą transformaciją $\boldsymbol{Y} = \boldsymbol{C}\mathbf{X}$, parinkę matricos $\boldsymbol{C}$ pirmąsias dvi eilutes tokio pavidalo:

$$(1/\sqrt{n}, 1/\sqrt{n}, ..., 1/\sqrt{n});$$

$$(\sqrt{1-1/n}, -1/\sqrt{n(n-1)}, ..., -1/\sqrt{n(n-1)}).$$

Tada $Y_1 = \sqrt{n}\bar{X}$, $Y_2 = \sqrt{n/(n-1)}(X_1 - \bar{X})$, $S^2 = \sum_i X_i^2 - n\bar{X}^2 = \sum_i Y_i^2 - Y_1^2 = Y_2^2 + Y_3^2 + ... + Y_n^2$. Statistika $U(\mathbf{X}) = \sqrt{(n-1)/n}Y_2/\sqrt{Y_2^2 + ... + Y_n^2}$. Kadangi $U(\mathbf{X})$ skirstinys nepriklauso nuo nežinomų parametrų, tai galime imti $\mu = 0, \sigma = 1$. Tada $Y_i^2 \sim \chi_1^2$ ir $U^2(\mathbf{X})/(1-1/n) \sim Be(1/2, n-2)$. Remiantis $U(\mathbf{X})$ simetrija pasiskirstymo funkcija, kai $u > 0$, yra

$$G(u) = \mathbf{P}\{U(\mathbf{X}) < u\} = \frac{1}{2} + \frac{1}{2}\mathbf{P}\{U^2(\mathbf{X}) < u^2\} = \frac{1}{2} +$$

$$\frac{\Gamma((n-1)/2)}{2\Gamma(1/2)\Gamma((n-2)/2)} \int_0^{u^2 n/(n-1)} v^{-1/2}(1-v)^{(n-4)/2} dv.$$

Diferencijuodami pagal u gauname tankio funkciją

$$g(u) = \frac{\Gamma((n-1)/2)}{\Gamma(1/2)\Gamma((n-2)/2)} \sqrt{n/(n-1)}(1 - nu^2/(n-1))^{(n-4)/2}, \quad |u| < \sqrt{1-1/n}.$$

I.2.74. Kadangi $X_{(1)}$ yra pilnoji ir pakankamoji statistika, tai pakanka patikrinti įvertinio $\hat{\gamma}$ nepaslinktumą. A. d. $X_{(1)}$ tankio funkcija $ne^{-n(x-\theta)}, x > \theta$. Gauname

$$\mathbf{E}_\theta \hat{\gamma} = \frac{n-1}{n} e^{-y} \int_\theta^\infty ne^x e^{-n(x-\theta)} dx = e^{-(y-\theta)}.$$

I.2.75. Tikėtinumo funkcija

$$L(\lambda) = \lambda^n e^{-\lambda \sum_i (X_i - a)} \mathbf{1}_{(a,\infty)}(X_{(1)}).$$

Pilnoji ir pakankamoji statistika $T = \sum_i (X_i - a) \sim G(\lambda, n)$. Nepaslinktasis parametro γ įvertinys

$$h(X_1) = \mathbf{1}_{(y,\infty)}(X_1), \quad \mathbf{E}_\lambda h(X_1) = \mathbf{P}_\lambda\{X_1 > y\} = e^{-\lambda(y-a)}.$$

I.2.70 pratime imkime $\hat{\theta}_1 = a, \hat{\theta}_2 = T$. Kadangi $\lambda(X_i - a) \sim \mathcal{E}(1)$, tai $2\lambda(X_i - a) \sim \chi_2^2$ ir statistika $U(\mathbf{X}) = (X_1 - a)/T \sim Be(1, n-1)$ nepriklauso nuo T. Gauname

$$\hat{\gamma} = \int_{-\infty}^\infty h(uT + a) g(u) du = (n-1) \int_{(y-a)/T}^1 (1-x)^{n-2} dx = (1 - (y-a)/T)^{n-1},$$

kai $0 < (y-a)/T < 1$.

I.2.76. A. d. $Y = \ln X \sim \mathcal{E}(\ln \alpha, \sigma)$. Statistika $Y_{(1)} = \ln X_{(1)}$ yra pilnoji ir pakankamoji parametro $\ln \alpha$ statistika. Pakanka patikrinti įvertinio $\hat{\gamma}$ nepaslinktumą.

$$\mathbf{E}_\alpha \hat{\gamma} = \frac{n\sigma - m}{n\sigma} \mathbf{E}_\alpha (\ln X_{(1)})^m = \frac{n\sigma - m}{n\sigma} \int_{\ln \alpha}^\infty n\sigma x^m e^{-n\sigma(x - \ln \alpha)} dx = \alpha^m.$$

I.2.77. A. d. $Y = e^{-X} \sim \mathcal{E}(e^\theta)$. Pasinaudoję **I.2.75** pratimu gauname

$$\hat{f}(z|\theta) = (n-1)e^{-z}(1 - e^{-z}/T)^{n-2}/T, \quad z > -\ln T,$$

čia $T = \sum_i e^{-X_i}$.

I.2.78. Tikėtinumo funkcija

$$L(M) = \frac{M^{[T]}(N-M)^{[n-T]}}{N^{[n]}} = \frac{C_M^T C_{N-M}^{n-T}}{C_N^n};$$

čia $K^{[m]} = K(K-1)...(K-m+1)$, o $T = \sum_{i=1}^n X_i$. Remiantis faktorizacijos kriterijumi T yra parametro M pakankamoji statistika. Statistika T įgyja reikšmes iš aibės $D = \{t : \max(0, n-(N-M)) \le t \le \min(n, M)\}$.

Remiantis statistikos pilnumo apibrėžimu reikia įrodyti, kad iš sąlygos

$$\mathbf{E}(h(T)|M) = \sum_D h(t)\frac{C_M^t C_{N-M}^{n-t}}{C_N^n} \equiv 0, \quad \forall\ M = 0, 1, ..., N,$$

išplaukia, kad $h(t) = 0$ su visais $t \in D$.

Imkime paeiliui $M = 0, 1, ..., N$. Kai $M = 0$, tai aibei D priklauso vienintelis taškas $t = 0$. Taigi

$$\mathbf{E}(h(T)|0) = h(0)\cdot 1 = 0 \quad \Rightarrow h(0) = 0.$$

Kai $M = 1$ ir $1 \le n \le N-1$, tai aibė D susideda iš dviejų taškų $t = 0$ ir $t = 1$. Gauname

$$\mathbf{E}(h(T)|1) = h(0)\frac{C_{N-1}^n}{C_N^n} + h(1)\frac{C_{N-1}^{n-1}}{C_N^n} = 0 \quad \Rightarrow \quad h(1) = 0.$$

Tęsdami šią procedūrą įsitikinsime, kad $h(t) = 0$ su visais $t \in D$. Taigi pakankamoji statistika T yra pilnoji parametro M statistika.

I.2.79. a) Randame

$$\mathbf{E}_\theta X = -\theta + \theta(1-\theta)^2\sum_{k=1}^{\infty} k\theta^{k-1} = -\theta + \theta(1-\theta)^2\frac{1}{(1-\theta)^2} \equiv 0,$$

nors X nėra tapačiai lygi nuliui. Statistika X nėra pilnoji.

b) Parametro $\gamma = (1-\theta)^2$ nepaslinktasis įvertinys turi tokį pavidalą $\hat{\gamma} = T(X) + cX$, kai c bet kokia konstanta, nes

$$\mathbf{E}_\theta\hat{\gamma} = \mathbf{E}_\theta T(X) + c\mathbf{E}_\theta X = (1-\theta)^2.$$

Kadangi $\mathbf{Cov}\,(T(X), X) = \mathbf{E}_\theta(XT(X)) = 0$, tai dispersijos $\boldsymbol{V}_\theta\hat{\gamma} = \boldsymbol{V}_\theta(T(X)) + c^2\boldsymbol{V}_\theta X$ minimumas gaunamas, kai $c = 0$. Taigi $T(X)$ yra parametro γ NMD įvertinys.

c) Įvertinys $\hat{\theta}$ yra nepaslinktasis

$$\mathbf{E}_\theta\hat{\theta} = \mathbf{P}\{X < 1|\theta\} + c\mathbf{E}_\theta X = \theta, \quad c \in \mathbf{R}.$$

Randame

$$\mathbf{E}_\theta\hat{\theta}^2 = \mathbf{E}_\theta\mathbf{1}_{\{-1\}}(X) + 2c\mathbf{E}_\theta(X\mathbf{1}_{\{-1\}}(X)) + c^2\mathbf{E}_\theta X^2;$$

$$\boldsymbol{V}_\theta\hat{\theta} = \theta(1-\theta) - 2c\theta + c^2\mathbf{E}_\theta X^2.$$

Gauname

$$\mathbf{E}_\theta X^2 = \theta + \theta(1-\theta)^2 \sum_{k=1}^{\infty} k^2\theta^{k-1} = \theta + \theta(1-\theta)^2 \frac{1-\theta}{(1-\theta)^3} = \frac{2\theta}{1-\theta}.$$

Tada dispersija

$$\boldsymbol{V}_\theta \hat{\theta} = \theta(1-\theta) - 2c\theta + 2c^2 \frac{\theta}{1-\theta}$$

įgyja minimalią reikšmę, kai $c = (1-\theta)/2$. Tačiau c įeina į įvertinio išraišką, todėl negali priklausyti nuo nežinomo parametro θ. NMD įvertinys neegzistuoja.

I.2.80. Kadangi X yra pilnoji ir pakankamoji statistika, tai pakanka patikrinti įvertinio $\hat{\gamma}$ nepaslinktumą, t. y. patikrinti tapatybę

$$\frac{2e^{-\lambda}}{1-e^{-\lambda}} \sum_{k=1}^{\infty} \frac{\lambda^{2k}}{(2k)!} \equiv 1 - e^{-\lambda}, \quad 0 < \lambda < \infty.$$

Gauname ekvivalenčią tapatybę

$$\sum_{k=1}^{\infty} \frac{\lambda^{2k}}{(2k)!} = (e^{\lambda} + e^{-\lambda} - 2)/2.$$

Skleisdami eksponentes eilute įsitikiname, kad dešinėje tapatybės pusėje gauname kairėje pusėje parašytą sumą.

I.2.81. a) Tikėtinumo funkcija

$$L(\boldsymbol{\theta}) = \frac{1}{(\sqrt{2\pi}\sigma_1)^n} \frac{1}{(\sqrt{2\pi}\sigma_2)^m} \exp\{-\frac{1}{2\sigma_1^2} \sum_i X_i^2 + \frac{\mu_1}{\sigma_1^2} \sum_i X_i -$$
$$\frac{1}{2\sigma_2^2} \sum_i Y_i^2 + \frac{\mu_2}{\sigma_2^2} \sum_i Y_i - \frac{n\mu_1^2}{2\sigma_1^2} - \frac{m\mu_2^2}{2\sigma_2^2}\}$$

priklauso keturparametrei eksponentinio tipo skirstinių šeimai. Parametro kitimo sritis turi vidinių taškų. Statistika $\boldsymbol{T}$ yra pilnoji ir pakankamoji.

b) $\mathbf{E}_{\boldsymbol{\theta}}(\bar{X} - \bar{Y}) \equiv 0$, tačiau statistika $\bar{X} - \bar{Y} \neq 0$ b. v.

c) $\mathbf{E}_{(\theta)}(s_1^2 - s_2^2) \equiv 0$, tačiau statistika $s_1^2 - s_2^2 \neq 0$ b. v., čia $s_1^2 = \sum_i (X_i - \bar{X})^2/(n-1), s_2^2 = \sum_i (Y_i - \bar{Y})^2/(m-1)$. Imdami statistiką $\boldsymbol{T}^*$, gauname triparametrę eksponentinio tipo skirstinių šeimą, kai parametro kitimo sritis turi vidinių taškų. Statistika $\boldsymbol{T}^*$ yra pilnoji ir pakankamoji.

I.2.82. Tikėtinumo funkcija

$$L(\boldsymbol{\theta}) = \frac{1}{(\sqrt{2\pi(1-\rho^2)}\sigma^1\sigma^2)^n} \exp\{-\frac{1}{2(1-\rho^2)}[\frac{1}{\sigma_1^2} \sum_i X_i^2 - \frac{2\mu_1}{\sigma_1^2} \sum_i X_i +$$
$$\frac{1}{\sigma_2^2} \sum_i Y_i^2 - \frac{2\mu_2}{\sigma_2^2} \sum_i Y_i + \frac{4\rho}{\sigma_1\sigma_2}(\mu_1 \sum_i Y_i + \mu_2 \sum_i X_i) + b(\boldsymbol{\theta})]\}$$

priklauso penkiaparametrei eksponentinio tipo skirstinių šeimai. Parametro kitimo sritis turi vidinių taškų. Statistika $\boldsymbol{T}$ yra pilnoji ir pakankamoji parametro $\boldsymbol{\theta}$ statistika.

I.2.83. Parametro $\boldsymbol{\theta} = (\mu_1, \mu_2)^T$ pilnoji ir pakankamoji statistika yra $\boldsymbol{T} = (\bar{X}, \bar{Y})^T$. Gauname

$$\hat{\gamma} = \mathbf{E}[\mathbf{1}_{(-\infty,0)}(Y_1 - X_1)|\bar{X}, \bar{Y}] = \mathbf{P}\{Y_1 < X_1|\bar{X}, \bar{Y}\} = \mathbf{P}\{Y_1 - \bar{Y} < X_1 - \bar{X} + \bar{X} - \bar{Y}|\bar{X}, \bar{Y}\}.$$

Statistikos $X_1 - \bar{X}, Y_1 - \bar{Y}, \bar{X}, \bar{Y}$ nepriklausomos; $X_1 - \bar{X} \sim N(0, 1 - 1/n), Y_1 - \bar{Y} \sim N(0, 1 - 1/m)$. Gauname

$$\hat{\gamma} = \Phi((\bar{X} - \bar{Y})/\sqrt{2 - 1/n - 1/m}).$$

I.2.84. Parametro $\boldsymbol{\theta} = (\lambda_1, \lambda_2)^T$ pilnoji ir pakankamoji statistika yra

$$\boldsymbol{T} = (\sum_i X_i, \sum_i Y_i)^T = (T_1, T_2)^T.$$

Gauname

$$\hat{\gamma} = \mathbf{P}\{Y_1 < X_1|T_1, T_2\} = \mathbf{P}\{\frac{Y_1}{T_2} < \frac{X_1}{T_1}\frac{T_1}{T_2}|T_1, T_2\}.$$

Statistikos $X_1/T_1, Y_1/T_2, T_1, T_2$ nepriklausomos;

$$X_1/T_1 \sim Be(1, m-1), \quad Y_1/T_2 \sim Be(1, n-1).$$

Pažymėkime $Z = T_1/T_2$. Tarkime, $Z \leq 1$. Tada

$$\hat{\gamma} = (m-1)(n-1)\int_0^1 (1-u)^{n-2}\int_0^{uZ}(1-v)^{m-2}dvdu =$$

$$(n-1)\int_0^1 (1-u)^{n-2}(1-(1-uZ)^{m-1})du =$$

$$1-(n-1)\int_0^1 (1-u)^{n-2}(1-uZ)^{m-1}du.$$

Kai $Z > 1$, gauname

$$\hat{\gamma} = 1-(m-1)(n-1)\int_0^1 (1-v)^{m-2}\int_0^{v/Z}(1-u)^{n-2}dudv =$$

$$1-(m-1)\int_0^1 (1-v)^{m-2}(1-(1-v/Z)^{n-1})dv =$$

$$(m-1)\int_0^1 (1-v)^{m-2}(1-v/Z)^{n-1}dv.$$

I.2.85. Parametro θ pilnoji ir pakankamoji statistika yra $X_{(n)}$, kurios tankis

$$f_n(x|\theta) = nx^{n-1}/\theta^n, \quad 0 < x < \theta.$$

Reikia funkcijos h atžvilgiu išspręsti funkcinę lygtį (žr. [2], 3.3 skyrelį)

$$\mathbf{E}_\theta h(X_{(n)}) \equiv \gamma(\theta).$$

Išbandome funkciją $\gamma(X_{(n)})$

$$\mathbf{E}_\theta\gamma(X_{(n)}) = \frac{n}{\theta^n}\int_0^\theta \gamma(x)x^{n-1}dx = \frac{\gamma(x)x^n}{\theta^n}|_0^\theta -$$

$$\frac{1}{\theta^n}\int_0^\theta \gamma'(x)x^n dx = \gamma(\theta) - \mathbf{E}_\theta(X_{(n)}\gamma'(X_{(n)}))/n.$$

Parametro $\gamma(\theta)$ NMD įvertinys

$$\hat{\gamma} = \gamma(X_{(n)}) + X_{(n)}\gamma'(X_{(n)})/n, \quad \mathbf{E}_\theta\hat{\gamma} = \gamma(\theta).$$

I.2.86. Tikėtinumo funkcija

$$L(\alpha,\beta,\sigma^2) = \frac{1}{(\sqrt{2\pi}\sigma)^n}\exp\{-\frac{1}{2\sigma^2}\sum_i Y_i^2 + \frac{\alpha}{\sigma^2}\sum_i Y_i + \frac{\beta}{\sigma^2}\sum_i Y_i x_i - \frac{1}{2\sigma^2}\sum_i(\alpha+\beta x_i)^2\}$$

priklauso triparametrei eksponentinio tipo skirstinių šeimai. Parametro kitimo sritis turi vidinių taškų. Statistika $\boldsymbol{T}$ yra pilnoji ir pakankamoji.

I.2.87. Tegu gautoji imtis $(X_1,...,X_n)^T$. Pažymėkime $Y_1 = X_1, Y_i = X_i - X_{i-1}, i = 2,...,n$. Tada $Y_1,...,Y_n$ yra paprastoji imtis a. d. $Y \sim N(\varphi,\sigma^2)$. Parametrų φ ir σ^2 NMD įvertiniai yra $\hat{\varphi} = \bar{Y} = X_n/n; \hat{\sigma}^2 = \sum_{i=1}^n(X_i - X_{i-1} - \hat{\varphi})^2/(n-1), X_0 = 0$.

I.2.88. a) Tikėtinumo funkcija

$$L(\theta) = (\prod_{i=1}^n a_{X_i})\exp\{\sum_{i=1}^n X_i\ln\theta - n\ln f(\theta)\}$$

priklauso vienparametrei eksponentinio tipo skirstinių šeimai. Statistika $T = \sum_i X_i$ yra pilnoji ir pakankamoji. Jos skirstinys yrs tokio paties eksponentinio tipo: kai $k = 2c, 2c+1, ...$ ir $n = 2$, gauname

$$\mathbf{P}_\theta\{X_1 + X_2 = k\} = \mathbf{P}_\theta\{X_1 = i, X_2 = k-i\} =$$

$$\sum_{i=c}^{k-c}\frac{a_i\theta^i}{f(\theta)}\frac{a_{k-i}\theta^{k-i}}{f(\theta)} = \frac{\theta^k}{f^2(\theta)}\sum_{i=c}^{k-c}a_i a_{k-i} = \frac{b_k\theta^k}{f^2(\theta)}.$$

Naudodami indukciją gausime, kad tokio tipo skirstinys galioja ir kai $n > 2$.

b) Funkcijos h atžvilgiu sprendžiame funkcinę lygtį

$$\mathbf{E}_\theta h(T) = \sum_{k=nc}^\infty b_k\theta^k/(f(\theta))^n \equiv \theta^r, \quad \theta > 0.$$

Padauginę dešiniąją tapatybės pusę iš vieneto, gauname

$$\mathbf{E}_\theta h(T) = \sum_{k=nc}^\infty b_k\theta^k/(f(\theta))^n \equiv \theta^r\sum_{j=nc}^\infty b_j\theta^j/(f(\theta))^n$$

ir, sulyginę koeficientus prie vienodų laipsnių θ^k, matome, kad parametro θ^r

NMD įvertinys yra $U_r(T)$, kai $U_r(T) = 0$, jei $T < nc+r$, $U_r(T) = b_{T-r}/b_T$, jei $T \geq nc+r$.

c) Dispersija $\boldsymbol{V}_\theta(U_r(T)) = \mathbf{E}_\theta(U_r(T))^2 - (\mathbf{E}_\theta(U_r(T)))^2 = \mathbf{E}_\theta(U_r(T))^2 - \theta^{2r}$. Nepaslinktasis dispersijos įvertinys yra $(U_r(T))^2 - U_{2r}(T)$.

I.2.89. $\mathbf{E}_{\mu,\sigma}(\bar{X})^2 = \mu^2 + \sigma^2/n, \boldsymbol{V}_{\mu,\sigma}(\bar{X})^2 = (2\sigma^2/n)(2\mu^2 + \sigma^2/n)$; NMD įvertinys $\hat{\vartheta} = \bar{X}^2 - s^2/n$, $\boldsymbol{V}_{\mu,\sigma}\hat{\vartheta} = (2\sigma^2/n)(2\mu^2 + \sigma^2/(n(n-1)))$.

I.2.90. Tegu $\boldsymbol{\theta} = (\mu_x, \mu_y, \sigma_x^2, \sigma_y^2)$ ir $\bar{X}$, $\bar{Y}$, s_x^2, s_y^2 – parametrų μ_x, μ_y, σ_x^2, σ_y^2 NMD įvertiniai. Tada

a) $\mathbf{E}_{\boldsymbol{\theta}}(\bar{X} - \bar{Y}) = \mu_x - \mu_y$; $(m-1)s_x^2/\sigma_x^2 \sim \chi^2_{m-1}$, $(n-1)s_y^2/\sigma_y^2 \sim \chi^2_{n-1}$;

$$\frac{\sigma_y^2}{\sigma_x^2}\frac{s_x^2}{s_y^2} \sim Fm-1, n-1, \qquad \left(\frac{\sigma_y}{\sigma_x}\right)^r \mathbf{E}_{\boldsymbol{\theta}} \left(\frac{s_x}{s_y}\right)^r = \mathbf{E}\left(F^{r/2}_{m-1,n-1}\right).$$

Parametrų $\mu_x - \mu_y$ ir $(\sigma_x/\sigma_y)^r$ NMD įvertiniai yra $\bar{X} - \bar{Y}$ ir $(s_x/s_y)^r/(\mathbf{E}(F^{r/2}_{m-1,n-1}))$;

b) $s^2 = [s_x^2(m-1) + s_y^2(n-1)] \sim \sigma_x^2\chi^2_\nu, \nu = m+n-2$; $\mathbf{E}_{\boldsymbol{\theta}}(s^2/\nu) = \sigma_x^2$; statistika $(\bar{X} - \bar{Y})/(s\sqrt{1/m + 1/n})$ turi necentrinį Stjudento skirstinį su ν laisvės laipsnių ir necentriškumo parametru $(\mu_x - \mu_y)/(\sigma_x\sqrt{1/m + 1/n})$;

$$\mathbf{E}_{\boldsymbol{\theta}} \frac{\bar{X} - \bar{Y}}{s} = \frac{\mu_x - \mu_y}{\sigma_x}\sqrt{\nu}\frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)}.$$

Parametrų σ_x^2 ir $(\mu_x - \mu_y)/\sigma_x$ NMD įvertiniai yra s^2 ir

$$[(\bar{X} - \bar{Y})/s]/(\sqrt{\nu}\Gamma((\nu-1)/2)/\Gamma((\nu)/2));$$

c) $[m\bar{X} + \gamma n\bar{Y}]/(m + n\gamma)$;

d) pakankamoji statistika $(\bar{X}, \bar{Y}, s_x^2, s_y^2)^T$ nėra pilnoji: statistikos $\bar{X} - \bar{Y}$ vidurkis tapačiai lygus 0, tačiau statistika $\bar{X} - \bar{Y}$ nėra lygi nuliui b. v.

e), f) $\mathbf{P}_{\boldsymbol{\theta}}\{X_1 \leq Y_1\} = \mathbf{P}_{\boldsymbol{\theta}}\{X_1 - Y_1 \leq 0\} = 1/2$, nes $X_1 - Y_1 \sim N(0, \sigma_x^2 + \sigma_y^2)$; tikimybės NMD įvertinys yra 1/2.

I.2.91. Pažymėkime $S_x = \sum_i (X_{(i)} - X_{(1)})$, $S_y = \sum_i (Y_{(i)} - Y_{(1)})$. Tada:

a) $\theta = a_x - a_y$, $\hat{\theta} = X_{(1)} - Y_{(1)} - S_x/(m(m-1)) + S_y/(n(n-1))$; $\gamma = \theta_x/\theta_y$, $\hat{\gamma} = S_x(n-2)/(S_y(m-1))$;

b) $\hat{\theta}_x = (S_x + S_y)/(m+n-2)$; $\eta = (a_x - a_y)/\theta_x$, $\hat{\eta} = (X_{(1)} - Y_{(1)})(m+n-3)/(S_x + S_y) - (n-m)/(mn)$;

c) pakankamoji statistika $(X_{(1)}, Y_{(1)}, S_x, S_y)^T$ nėra pilnoji.

I.2.3 skyrelis

I.2.92. Pasinaudokite tokiu algebros faktu. Tegu $\boldsymbol{A}$ ir $\boldsymbol{D}$ – kvadratinės matricos ir $|\boldsymbol{D}| \neq 0$. Tada determinantas

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix}$$

yra lygus $|\boldsymbol{D}||\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}|$. Suskaidykime matricą $\boldsymbol{I}$ į blokus

$$\begin{vmatrix} I_{11} & \boldsymbol{B}^T \\ \boldsymbol{B} & \boldsymbol{D} \end{vmatrix}$$

Tada

$$I^{11} = \frac{|\boldsymbol{D}|}{|\boldsymbol{I}|} = \frac{|\boldsymbol{D}|}{|\boldsymbol{D}|(I_{11} - \boldsymbol{B}^T\boldsymbol{D}^{-1}\boldsymbol{B})} = \frac{1}{I_{11} - \boldsymbol{B}^T\boldsymbol{D}^{-1}\boldsymbol{B}} \geq \frac{1}{I^{11}}.$$

I.2.93. $\boldsymbol{V}_p\hat{p} \geq pq/n$; $\boldsymbol{V}_p(\hat{pq}) \geq pq(q-p)^2/n$; $\boldsymbol{V}_p(\hat{p^2}) \geq 4p^3q/n$.

I.2.94. Tikėtinumo funkcija

$$L = L(p) = C_n^S e^{S\ln(p/(1-p)) + n\ln(1-p)}, \quad S = X_1 + ... + X_n \sim B(n,p).$$

Tegu

$$Y_1 = \frac{L'}{L} = \frac{S - np}{pq}, \quad Y_2 = \frac{L''}{L} = \frac{(S-np)^2}{p^2q^2} - \frac{(S-np)(1-2p)}{p^2q^2} - \frac{n}{pq}.$$

Randame a. v. $\boldsymbol{Y} = (Y_1, Y_2)^T$ kovariacijų matricą $\boldsymbol{J} = [J_{ij}]_{2\times2}$.

$$J_{11} = \mathbf{E}_p Y_1^2 = n/(pq), \quad J_{12} = \mu_3/(p^3q^3) - n(1-2p)/(p^2q^2) = 0,$$

nes $\mu_3 = \mathbf{E}_p(S - \mathbf{E}_p S)^3 = n\mathbf{E}_p(X_i - \mathbf{E}_p X_i)^3 = npq(1-2p)$.

$$J_{22} = \frac{\mu_4}{p^4q^4} + \frac{\mu_2(1-2p)^2}{p^4q^4} + \frac{n^2}{p^2q^2} - 2\frac{\mu_3(1-2p)}{p^4q^4} - 2\frac{n\mu_2}{p^3q^3} = \frac{2n(n-1)}{p^2q^2},$$

nes $\mu_4 = \mathbf{E}_p(S - \mathbf{E}_p S)^4 = n\mathbf{E}_p(X_i - \mathbf{E}_p X_i)^4 + 3n(n-1)(\boldsymbol{V}_p X_i)^2 = npq(1-3pq) + 3n(n-1)p^2q^2$.

Atvirkštinės matricos $\boldsymbol{J}^{-1} = [J^{ij}]_{2\times2}$ elementai $J^{11} = pq/n$, $J^{12} = J^{21} = 0$, $J^{22} = p^2q^2/(2n(n-1))$. Funkcijų $\gamma_1(p) = p^2$ ir $\gamma_2(p) = pq$ išvestinių vektoriai $\Gamma_1 = (2p, 2)^T$ ir $\Gamma_2 = (1-2p, -2)^T$. Gauname patikslintas Rao ir Kramerio nelygybes

$$\boldsymbol{V}_p\hat{\gamma}_1 \geq \Gamma_1^T\boldsymbol{J}^{-1}\Gamma_1 = \frac{4p^3q}{n} + \frac{2p^2q^2}{n(n-1)},$$

$$\boldsymbol{V}_p\hat{\gamma}_2 \geq \Gamma_2^T\boldsymbol{J}^{-1}\Gamma_2 = \frac{(1-2p)^2pq}{n} + \frac{2p^2q^2}{n(n-1)}.$$

I.2.95. Parametrų NMD įvertiniai (žr. **I.2.53** pratimą) yra $\hat{p} = S/n$, $\hat{p^2} = S(S-1)/[n(n-1)]$, $\hat{pq} = S(n-S)/[n(n-1)]$. $\boldsymbol{V}_p\hat{p} = pq/n$ (Rao ir Kramerio nelygybė virsta lygybe).

$$\boldsymbol{V}_p(\hat{p^2}) = \mathbf{E}_p(S(S-1))^2/(n(n-1))^2 - p^4.$$

Faktorialiniai momentai $\nu_k = \mathbf{E}_p S^{[k]} = p^k n^{[k]}, k = 1, 2, ...$ Pertvarkome $(S(S-1))^2 = S^{[4]} + 4S^{[3]} + 2S^{[2]}$. Tada

$$\boldsymbol{V}_p(\hat{p^2}) = \frac{\nu_4 + 4\nu_3 + 2\nu_2}{(n(n-1))^2} - p^4 = \frac{4p^3q}{n} + \frac{2p^2q^2}{n(n-1)}.$$

Analogiškai $S^2(n-S)^2 = S^{[2]}(n-S)^{[2]} + S^{[2]}(n-S) + S(n-S)^{[2]} + S(n-S)$, $\nu_{kl} = \mathbf{E}(S^{[k]}(n-S)^{[l]}) = p^k q^l n^{[k+l]}, k, l = 1, 2, ...$

Randame

$$\boldsymbol{V}_p(\hat{pq}) = \frac{\nu_{22}+\nu_{21}+\nu_{12}+\nu_{11}}{(n(n-1))^2} - p^2q^2 = \frac{(1-2p)^2pq}{n} + \frac{2p^2q^2}{n(n-1)}.$$

Abiem atvejais patikslinta Rao ir Kramerio nelygybė virsta lygybe.

I.2.96. Dispersijos ribos Rao ir Kramerio nelygybėje $\boldsymbol{V}_\lambda\hat{\lambda} \geq \lambda/n$; $\boldsymbol{V}_\lambda(\hat{\lambda^2}) \geq 4\lambda^3/n$; $\gamma = \mathrm{e}^{-\lambda}$, $\boldsymbol{V}_\lambda(\hat{\gamma}) \geq \lambda e^{-2\lambda}/n$.

I.2.97. Tikėtinumo funkcija

$$L = L(\lambda) = (1/\prod_i X_i!)e^{S\ln(\lambda)-n\ln(\lambda)}, \quad S = X_1 + ... + X_n \sim \mathcal{P}(\lambda).$$

Tegu

$$Y_1 = \frac{L'}{L} = \frac{S-n\lambda}{\lambda}, \quad Y_2 = \frac{L''}{L} = \frac{(S-n\lambda)^2}{\lambda^2} - \frac{(S-n\lambda)}{\lambda^2} - \frac{n}{\lambda}.$$

Randame a. v. $\boldsymbol{Y} = (Y_1, Y_2)^T$ kovariacijų matricą $\boldsymbol{J} = [J_{ij}]_{2\times 2}$.

$$J_{11} = \mathbf{E}_\lambda Y_1^2 = n/\lambda, \quad J_{12} = \mu_3/\lambda^3 - \mu_2/\lambda^3 = 0,$$

nes $\mu_3 = \mathbf{E}_\lambda(S - \mathbf{E}_\lambda S)^3 = n\lambda$, $\mu_2 = \mathbf{E}_\lambda(S-\mathbf{E}_\lambda S)^2 = n\lambda$.

Skaičiuojant centrinius momentus μ_k patogu juos išreikšti faktorialiniais momentais $\nu_k = \mathbf{E}_\lambda S^{[k]} = n^k\lambda^k, k = 1, 2, ...$;

$$\mu_3 = \nu_3 - 3\nu_2\nu_1 + 3\nu_2 + 2\nu_1^3 - 3\nu_1^2 + \nu_1 = n\lambda.$$

$$J_{22} = \frac{\mu_4}{\lambda 4} + \frac{\mu_2}{\lambda^4} + \frac{n^2}{\lambda^2} - 2\frac{\mu_3}{\lambda^4} - 2\frac{n\mu_2}{\lambda^3} = \frac{2n^2}{\lambda^2},$$

nes $\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_3 + 6\nu_2\nu_1^2 - 12\nu_2\nu_1 + 7\nu_2 - 3\nu_1^4 + 6\nu_1^3 - 4\nu_1^2 + \nu_1 = 3n^2\lambda^2 + n\lambda$.

Atvirkštinės matricos $\boldsymbol{J}^{-1} = [J^{ij}]_{2\times 2}$ elementai $J^{11} = \lambda/n$, $J^{12} = J^{21} = 0$, $J^{22} = \lambda^2/(2n^2)$. Funkcijų $\gamma_1 = \lambda^2$ ir $\gamma_2 = e^{-\lambda}$ išvestinių vektoriai $\Gamma_1 = (2\lambda, 2)^T$ ir $\Gamma_2 = (-e^{-\lambda}, e^{-\lambda})^T$. Gauname patikslintas Rao ir Kramerio nelygybes

$$\boldsymbol{V}_\lambda\hat{\gamma}_1 \geq \Gamma_1^T\boldsymbol{J}^{-1}\Gamma_1 = \frac{4\lambda^3}{n} + \frac{2\lambda^2}{n^2},$$

$$\boldsymbol{V}_\lambda\hat{\gamma}_2 \geq \Gamma_2^T\boldsymbol{J}^{-1}\Gamma_2 = \frac{\lambda e^{-2\lambda}}{n} + \frac{\lambda^2 e^{-2\lambda}}{2n^2}.$$

I.2.98. Parametrų NMD įvertiniai (žr. **I.2.49**, **I.2.50** pratimus) yra $\hat{\lambda} = S/n$, $\hat{\lambda^2} = S(S-1)/n^2$, $\gamma = e^{-\lambda}$, $\hat{\gamma} = (1-1/n)^S$.

$\boldsymbol{V}_\lambda\hat{\lambda} = \lambda/n$ (Rao ir Kramerio nelygybė virsta lygybe).

$$\boldsymbol{V}_\lambda\hat{\lambda^2} = \mathbf{E}(S(S-1))^2/n^4 - \lambda^4.$$

Pertvarkę analogiškai **I.2.95** pratimui, gauname

$$\boldsymbol{V}_\lambda\hat{\lambda^2} = (\nu_4 + 4\nu_3 + 2\nu_2 - n^4\lambda^4)/n^4 = \frac{4\lambda^3}{n} + \frac{2\lambda^2}{n^2}.$$

Dispersija sutampa su patikslintos Rao ir Kramerio nelygybės riba.

$$\mathbf{E}_\lambda \hat{\gamma} = \sum_{k=0}^{\infty}(1-1/n)2k\frac{(n\lambda)^k}{k!}e^{-n\lambda} = e^{-\lambda(1/n-2)},$$

$$\boldsymbol{V}_\lambda \hat{\gamma} = e^{-2\lambda}(e^{\lambda/n}-1) = \frac{e^{-2\lambda}\lambda}{n} + \frac{e^{-2\lambda}\lambda^2}{n^2} + O(\frac{1}{n^3}).$$

Įvertinio $\hat{\gamma}$ dispersija skiriasi nuo klasikinės Rao ir Kramerio nurodytos ribos eilės $O(1/n^2)$ nariu; patikslinus nelygybę panaudojant antrąją L išvestinę, skirtumas yra $O(1/n^3)$. Matyt, nelygybę galima dar patikslinti panaudojant aukštesnių eilių išvestines.

I.2.99. a) Kadangi $\sqrt{n}(\bar{X}-\lambda) \xrightarrow{d} Y \sim N(0,\lambda)$, tai remiantis delta metodu $\sqrt{n}(\tilde{\gamma} - \gamma) \xrightarrow{d} Z \sim N(0, \lambda e^{-2\lambda})$ ir $\lambda[\gamma'(\lambda)]^2/(i(\gamma) = 1$; $\tilde{\gamma} = \hat{\gamma} + O_P(1/n)$.

b) Reikia rasti tokią funkciją $B(\lambda)$, kad

$$\sqrt{n}(e^{-\bar{X}} - e^{-\lambda} - B(\lambda)\dot{\ell}(\lambda)/n) \xrightarrow{P} 0.$$

Kadangi $\dot{\ell}(\lambda)/n = (\bar{X}-\lambda)/\lambda$, tai, imdami $B(\lambda) = -\lambda e^{-\lambda}$, gausime

$$\sqrt{n}(e^{-\bar{X}} - e^{-\lambda} + e^{-\lambda}(\bar{X}-\lambda)) = e^{-\lambda}\sqrt{n}(\bar{X}-\lambda)[(e^{-(\bar{X}-\lambda)}-1)/(\bar{X}-\lambda)+1].$$

Kadangi $\bar{X} \xrightarrow{P} \lambda$, tai laužtiniuose skliaustuose parašytas reiškinys pagal tikimybę artėja į 0. Įvertinys $e^{-\bar{X}}$ asimptotiškai efektyvus pagal Rao.

I.2.100. a) $\hat{\mu} = \bar{X}$, $\widehat{\mu^2} = \bar{X}^2 - \sigma^2/n$; $\boldsymbol{V}_\mu \hat{\mu} = \sigma^2/n$, $\boldsymbol{V}_\mu \widehat{\mu^2} = \mathbf{E}_\mu((\bar{X}-\mu)^2 + 2\mu(\bar{X} - \mu) - \sigma^2/n)^2 = 4\mu^2\sigma^2/n + 2\sigma^4/n^2$.

b) $\boldsymbol{V}_\mu \hat{\mu} \geq \sigma^2/n$, $\boldsymbol{V}_\mu \widehat{\mu^2} \geq 4\mu^2\sigma^2/n$. Įvertinys $\hat{\mu} = \bar{X}$ yra efektyvusis. Įvertinys $\widehat{\mu^2} = \bar{X}^2 - \sigma^2/n$ asimptotiškai efektyvus pagal Rao, nes

$$\sqrt{n}(\bar{X}^2 - \mu^2 - B(\mu)(\bar{X}-\mu)/\sigma^2) = \sqrt{n}(\bar{X}-\mu)(\bar{X}+\mu-B(\mu)/\sigma^2) \xrightarrow{P} 0,$$

kai $B(\mu) = 2\mu\sigma^2$.

c) Tikėtinumo funkcija

$$L = L(\mu) = (\sqrt{2\pi}\sigma)^{-n}\exp\{-\sum_{i=1}^{n} X_i^2/(2\sigma^2)\}\exp\{n\mu\bar{X}/\sigma^2 - n\mu^2/(2\sigma^2)\}.$$

Tegu

$$Y_1 = \frac{L'}{L} = \frac{n}{\sigma^2}(\bar{X}-\mu), \quad Y_2 = \frac{L''}{L} = \frac{n^2}{\sigma^4}(\bar{X}-\mu)^2 - \frac{n}{\sigma^2}.$$

Randame a. v. $\boldsymbol{Y} = (Y_1, Y_2)^T$ kovariacijų matricą $\boldsymbol{J} = [J_{ij}]_{2\times 2}$.

$$J_{11} = \frac{n^2}{\sigma^4}\mathbf{E}_\mu(\bar{X}-\mu)^2 = \frac{n}{\sigma^2}, \quad J_{12} = J_{21} = 0,$$

$$J_{22} = \frac{n^4}{\sigma^8}\mathbf{E}_\mu(\bar{X}-\mu)^4 - 2\frac{n^3}{\sigma^6}\mathbf{E}_\mu(\bar{X}-\mu)^2 + \frac{n^2}{\sigma^4} = \frac{2n^2}{\sigma^4}.$$

Atvirkštinės matricos $\boldsymbol{J}^{-1}$ elementai $J^{11} = \sigma^2/n, J^{12} = J^{21} = 0,\ J^{22} = \sigma^4/(2n^2)$. Funkcijos μ^2 išvestinių vektorius $\Gamma = (2\mu, 2)^T$. Gauname patikslintą Rao ir Kramerio nelygybę

$$\boldsymbol{V}_\mu \hat{\mu^2} \geq \Gamma^T \boldsymbol{J}^{-1}\Gamma = \frac{4\mu^2\sigma^2}{n} + \frac{2\sigma^4}{n^2}.$$

NMD įvertinio $\hat{\mu^2}$ dispersija sutampa su patikslinta dispersijos riba.

I.2.101. a) $\hat{\sigma^2} = s_0^2 = \sum_i (X_i - \mu)^2/n,\ \boldsymbol{V}_\sigma s_0^2 = 2\sigma^4/n;\ \hat{\sigma} = s_0/M_n,\ M_n = \sqrt{2/\pi}\Gamma((n+1)/2)/\Gamma(n/2)$ (žr. **I.2.63** pratimą), $\boldsymbol{V}_\sigma \hat{\sigma} = \sigma^2(1/M_n^2 - 1) \approx \sigma^2(1/(2n) + 1/(4n^2))$.

b) $\boldsymbol{V}_\sigma \hat{\sigma^2} \geq 2\sigma^4/n,\ \boldsymbol{V}_\sigma \hat{\sigma} \geq \sigma^2/(2n)$. Įvertinys s_0^2 efektyvus, o $\hat{\sigma}$ asimptotiškai efektyvus.

I.2.102. $\boldsymbol{V}_{\mu,\sigma}(\hat{\mu}, \hat{\sigma^2})^T \geq \boldsymbol{I}^{-1} = [I^{kl}]_{2\times 2};\ I^{11} = \sigma^2/n,\ I^{12} = I^{21} = 0,\ I^{22} = 2\sigma^4/n.$

I.2.103. a) $\hat{x}_P = \bar{X} + z_P s/M_{n-1},\ \boldsymbol{V}_{\mu,\sigma}\hat{x}_P = \sigma^2/n + z_P^2\sigma^2(1/M_{n-1}^2 - 1)$.
b) $\boldsymbol{V}_{\mu,\sigma}\hat{x}_P \geq \sigma^2/n + z_P^2\sigma^2/(2n)$.

I.2.104.

$$i(\theta) = -\mathbf{E}_\theta(\frac{\partial^2 \ln f(X|\theta)}{\partial\theta^2}) = \mathbf{E}e^{-(X-\theta)} = \int_\theta^{+\infty} e^{-2(x-\theta)} \exp\left(-e^{-(x-\theta)}\right) =$$

$$= |e^{-(x-\theta)} = y;\ dx = dy/y|$$

$$= -\int_1^0 y^2 e^{-y} dy/y = \int_0^1 ye^{-y}dy = -\int_0^1 yd(e^{-y}) = -ye^{-y}|_0^1 + \int_0^1 e^{-y}dy =$$

$$= 1/e - 1/e + 1 = 1; \quad I(\theta) = ni(\theta) = n.$$

Nagrinėkime parametrą $\alpha = \exp\{-\theta\}$. Gauname $\theta = -\ln\alpha$

$$i(\alpha) = -\mathbf{E}\left(\left(\dot{\ell}_\theta\, \dot{\theta}_\alpha\right)'_\alpha\right) = -\mathbf{E}\left(\ddot{\ell}_{\theta\theta}\, \dot{\theta}_\alpha\, \dot{\theta}_\alpha + \dot{\ell}_\theta \ddot{\theta}_{\alpha\alpha}\right) =$$

$$= \left(\dot{\theta}_\alpha\right)^2 \left(-\mathbf{E}\left(\ddot{\ell}_{\theta\theta}\right)\right) + \ddot{\theta}_{\alpha\alpha}\left(-\mathbf{E}\left(\dot{\ell}_\theta\right)\right) = \left(\dot{\theta}_\alpha\right)^2 i(\theta)$$

$$\dot{\theta}_\alpha = -\frac{1}{\alpha}$$

$$I(e^{-\theta}) = n\left(-\frac{1}{\alpha}\right)^2 = n/e^{-2\theta}.$$

I.2.105. a) $\mathbf{E}_\beta\hat{\beta} = \beta$; b) $\boldsymbol{V}_\beta\hat{\beta} = \beta^2/n$; kadangi $I(\beta) = n/\beta^2$, tai įvertinio $\hat{\beta}$ dispersija lygi Rao ir Kramerio nelygybėje nurodytai ribai.

I.2.106. a) n/σ^2; b) $n/(2\sigma^4)$; c) $2n/\sigma^2$; d) $3n/\sigma^2$; e) $\boldsymbol{I}(\boldsymbol{\theta}) = [I_{kl}]_{2\times 2},\ I_{11} = n/\sigma^2,\ I_{22} = n/(2\sigma^4),\ I_{12} = I_{21} = 0$; f) $nk/(qp^2)$; g) $\boldsymbol{I}(\alpha,\gamma) = [I_{ij}]_{2\times 2};\ I_{11} = n\lambda/\alpha^2,\ I_{12} = I_{21} = -n/\alpha,\ I_{22} = n[\ln\Gamma(\lambda)]''$; h) $\boldsymbol{I}(\alpha,\beta) = [I_{ij}]_{2\times 2};\ I_{11} = n[\ln B(\alpha,\beta)]''_{\alpha^2},\ I_{12} = I_{21} = n[\ln B(\alpha,\beta)]''_{\alpha,\beta},\ I_{22} = n[\ln B(\alpha,\beta)]''_{\beta^2}$.

I.2.107. a) $2\sqrt{\lambda}$; b) $\arcsin(2p-1)$; c) $\sqrt{\gamma}\ln\theta$.

I.2.108. a) $\boldsymbol{I}(\mu,\ \sigma) = [I_{kl}]_{2\times 2}$; $I_{11} = I_{22} = 2\sigma^2$; $I_{12} = I_{21} = 0$; b) $\boldsymbol{I}(\mu,\ \theta) = [I_{kl}]_{2\times 2}$; $I_{11} = 1/\theta^2$, $I_{12} = I_{21} = (1+\Gamma'(1))/\theta^2$; $I_{22} = (1+2\Gamma'(1)+\Gamma''(1))/\theta^2$; c) $I_{11} = n/(3\theta^2)$, $I_{12} = I_{21} = -\mu/(3\theta^3)$, $I_{22} = 1/\theta^2 + 2\mu/\theta^3 + \mu^2/(3\theta^4)$; d) $I_{11} = (r+1)\sqrt{r+4}/(\sigma^2(r+3)\sqrt{r})$; $I_{22} = (1+I_{11}/\sigma^2)/\sigma^2$, $I_{12} = I_{21} = I_{11}/\sigma$.

I.2.109. a) Aibė $\{x : f(x|\theta) > 0\}$ priklauso nuo θ. b) Parametro θ NMD įvertinys (žr. **I.2.55** pratimą) yra $\hat{\theta} = (n+1)X_{(n)}/n$, $\boldsymbol{V}\hat{\theta} = \theta^2/((n+1)(n+2))$.

I.2.110. a) $\hat{\vartheta} = -X/\Gamma'(1)$; $\boldsymbol{V}\hat{\vartheta} = \theta^2(\Gamma''(1) - \Gamma'^2(1))/\Gamma'^2(1) \geq I^{-1}(\theta) = \theta^2(2 + 2\Gamma'(1) + \Gamma''(1))$; b) $\hat{\vartheta} = (-1)^r X^r/\Gamma^{(r)}(1)$; $\boldsymbol{V}(\hat{\vartheta}) = \theta^{2r}(\Gamma^{(2r)}(1) - [\Gamma^{(r)}(1)]^2)/\ [\Gamma^{(r)}(1)]^2)$ $> I^{-1}(\theta^r) = \theta^{2r}r^2/(1+2\Gamma'(1)+\Gamma''(1))$.

I.2.111. a) $\hat{\vartheta} = \exp\{t\bar{X} - t^2\sigma^2/2n\}$; b) $\boldsymbol{V}\hat{\vartheta} = \exp\{2t\mu\}(\exp\{t^2\sigma^2/n\} - 1) > I^{-1}(\vartheta) = \exp\{2t\mu\}\sigma^2 t^2/n$. Rao ir Kramerio nelygybėje nurodyta riba nepasiekiama. $\boldsymbol{V}\hat{\vartheta} - I^{-1}(\vartheta) = O(1/n^2)$. c) Įvertinys $\hat{\vartheta}$ asimptotiškai efektyvus.

I.2.112. Tikėtinumo funkcija

$$L(\lambda) = (\prod_i X_i)^{\eta-1}/[\Gamma(\eta)]^n \exp\{-\frac{1}{\lambda}\sum_{i=1}^n X_i - n\eta\ln\lambda\}, \quad \lambda > 0$$

priklauso vienparametrei eksponentinio tipo skirstinių šeimai. Statistika $T = \sum_i X_i \sim G(1/\lambda, n\eta)$ yra pilnoji ir pakankamoji statistika.

a) Parametrų λ ir λ^2 NMD įvertiniai $\hat{\lambda} = T/(n\eta)$, $\hat{\lambda^2} = T^2/(n\eta(n\eta+1))$. A. d. T pradiniai momentai $\alpha_k = \mathbf{E}_\lambda T^k = \lambda^k\Gamma(n\eta+k)/\Gamma(n\eta), k = 1, 2, ...$ Gauname

$$\mathbf{E}_\lambda\hat{\lambda} = \lambda,\ \ \boldsymbol{V}_\lambda\hat{\lambda} = \frac{\lambda^2}{n\eta};\quad \mathbf{E}_\lambda\hat{\lambda^2} = \lambda^2,\ \ \boldsymbol{V}_\lambda\hat{\lambda^2} = \frac{4\lambda^4}{n\eta} + \frac{2\lambda^4}{n\eta(n\eta+1)}.$$

b) Randame Fišerio informaciją

$$I(\lambda) = -\mathbf{E}_\lambda(\ln L(\lambda))'' = \mathbf{E}_\lambda(2T/\lambda^3 - n\eta/\lambda^2) = n\eta/\lambda^2.$$

Parametrų λ ir λ^2 nepaslinktųjų įvertinių dispersijų ribos Rao ir Kramerio nelygybėje yra

$$\boldsymbol{V}_\lambda(\hat{\lambda}) \geq \lambda^2/(n\eta), \quad \boldsymbol{V}_\lambda\hat{\lambda^2} \geq 4\lambda^4/(n\eta).$$

c) Tegu

$$Y_1 = \frac{L'}{L} = \frac{T}{\lambda^2} - \frac{n\eta}{\lambda}, \quad Y_2 = \frac{L''}{L} = \frac{T^2}{\lambda^4} - 2\frac{T(n\eta+1)}{\lambda^3} + \frac{n\eta(n\eta+1)}{\lambda^2}.$$

Randame a. v. $\boldsymbol{Y} = (Y_1, Y_2)^T$ kovariacinę matricą

$$J_{11} = \mathbf{E}_\lambda(\frac{T^2}{\lambda^4} - 2\frac{Tn\eta}{\lambda^3} + \frac{(n\eta)^2}{\lambda^2}) = \frac{\alpha_2}{\lambda^4} - 2\frac{\alpha_1 n\eta}{\lambda^3} + \frac{(n\eta)^2}{\lambda^2} = \frac{n\eta}{\lambda^2};$$

$$J_{12} = \frac{\alpha_3}{\lambda^6} - 2\frac{\alpha_2(n\eta+1)}{\lambda^5} - \frac{\alpha_2 n\eta}{\lambda^5} + 3\frac{\alpha_1 n\eta(n\eta+1)}{\lambda^4} - \frac{(n\eta)^2(n\eta+1)}{\lambda^3} = 0;$$

$$J_{22} = \frac{\alpha_4}{\lambda^8} - 4\frac{\alpha_3(n\eta+1)}{\lambda^7} + 2\frac{\alpha_2(3n\eta+2)(n\eta+1)}{\lambda^6} -$$

$$4\frac{\alpha_1 n\eta(n\eta+1)^2}{\lambda^5}+\frac{(n\eta)^2(n\eta+1)^2}{\lambda^4}=\frac{2n\eta(n\eta+1)}{\lambda^4}.$$

Atvirkštinės matricos $\boldsymbol{J}^{-1}=[J^{ij}]_{2\times 2}$ elementai $J^{11}=\lambda^2/(n\eta)$, $J^{12}=J^{21}=0$, $J^{22}=\lambda^4/(2n\eta(n\eta+1))$. Funkcijos λ^2 išvestinių vektorius $\Gamma=(2\lambda,2)^T$. Gauname patikslintą Rao ir Kramerio nelygybę

$$\boldsymbol{V}_\lambda\hat{\lambda^2}\geq\Gamma^T\boldsymbol{J}^{-1}\Gamma=\frac{4\lambda^4}{n\eta}+\frac{2\lambda^4}{n\eta(n\eta+1)}.$$

Įvertinio $\hat{\lambda^2}$ dispersija lygi patikslintai Rao ir Kramerio nelygybės ribai.

I.2.113. Tikėtinumo funkcija

$$L(\lambda)=(\prod_i X_i)^{\eta-1}/[\Gamma(\eta)]^n\exp\{-\lambda\sum_{i=1}^n X_i+n\eta\ln\lambda\},\quad \lambda>0$$

priklauso vienparametrei eksponentinio tipo skirstinių šeimai. Statistika $T=\sum_i X_i\sim G(\lambda,n\eta)$ yra pilnoji ir pakankamoji statistika.

a) Parametrų λ ir λ^2 NMD įvertiniai $\hat{\lambda}=(n\eta-1)/T$, $\hat{\lambda^2}=(n\eta-1)(n\eta-2)/T^2$. A. d. T pradiniai momentai $\alpha_k=\mathbf{E}_\lambda T^k=\lambda^{-k}\Gamma(n\eta-k)/\Gamma(n\eta), k=1,2,...$ Gauname

$$\mathbf{E}_\lambda\hat{\lambda}=\lambda,\ \ \boldsymbol{V}_\lambda\hat{\lambda}=\frac{\lambda^2}{n\eta-2};\quad \mathbf{E}_\lambda\hat{\lambda^2}=\lambda^2,\ \ \boldsymbol{V}_\lambda\hat{\lambda^2}=2\lambda^4\frac{2n\eta-5}{(n\eta-3)(n\eta-4)},\quad n\eta>4.$$

b) Randame Fišerio informaciją

$$I(\lambda)=-\mathbf{E}_\lambda(\ln L(\lambda))''=n\eta/\lambda^2.$$

Parametrų λ ir λ^2 nepaslinktųjų įvertinių dispersijų ribos Rao ir Kramerio nelygybėje yra

$$\boldsymbol{V}_\lambda(\hat{\lambda})\geq\lambda^2/(n\eta),\quad \boldsymbol{V}_\lambda\hat{\lambda^2}\geq 4\lambda^4/(n\eta).$$

Įvertiniai nėra efektyvūs, tačiau yra asimptotiškai efektyvūs.

I.2.114. $\boldsymbol{V}_\mu(T_{1n})=\Phi(c-\mu)(1-\Phi(c-\mu)/n$; $\boldsymbol{V}_\mu(T_{2n})=[\Phi'(c-\mu)]^2/n+O(1/n\sqrt{n})$; $ASE=[\Phi'(c-\mu)]^2/(\Phi(c-\mu)(1-\Phi(c-\mu))$.

I.2.115. Gauname įvertinių dispersijas

$$\boldsymbol{V}_\sigma\hat{\sigma}_2=\boldsymbol{V}_\sigma(\sum_i X_i^2)^{1/2}=\sigma^2/(2n)+O(1/(n\sqrt{n})),$$

$$\boldsymbol{V}_\sigma\hat{\sigma}_1=\boldsymbol{V}_\sigma(\sqrt{\pi/2}|X_i|)/n=\sigma^2(\pi-2)/(2n).$$

Tada $E_{21}=1/(\pi-2)$.

I.2.116. a) $ASE=4\mu^2/(4\mu^2+\mu_4-1+4\mu_3\mu)$; b) $ASE\leq 1$, kai $\mu_3=0$; c) pvz., a. d. X tankio funkcija $f(x)=kx^2\boldsymbol{I}_{(0,\,a)}(x)$, $a>0$.

I.2.117. $ASE=1$.

I.2.118. $b_{T_{1n}}=0$, $b_{T_{2n}}=-\theta/(n+1)$; $ASE=1$.

I.2.119. ASE $= 2/3$.

I.2.120. Randame

$$\sqrt{n}(\hat\mu_1 - \mu) \xrightarrow{d} Z \sim N(0, 1/\lambda^2),$$

$$\sqrt{n}(\hat\mu_2 - \mu) \xrightarrow{d} Y \sim N(0, 2/\lambda^2),$$

$E_{21} = 1/2$.

I.2.121. $\hat\mu = 17,185$.

I.2.4 skyrelis

I.2.122. Turime lygčių sistemą $\mathbf{E}_{\alpha,\beta}X = (1-\beta)\alpha = \bar X$, $\mathbf{E}_{\alpha,\beta}X^2 = \beta + (1-\beta)(1+\alpha^2) = a_2$. Išsprendę randame $\hat\alpha = (a_2-1)/\bar X$, $\hat\beta = 1 - \bar X^2/(a_2-1)$.

I.2.123. a) Parametrų μ ir σ NMD įvertiniai (žr. **I.2.60** pratimą) yra $\hat\mu = (X_{(1)} + X_{(n)})/2$, $\hat\sigma = (n+1)(X_{(n)} - X_{(1)})/(n-1)$; jų dispersijos $\boldsymbol{V}_{\theta_1,\theta_2}\hat\mu = (\theta_2-\theta_1)^2/(2(n+1)(n+2))$, $\boldsymbol{V}_{\theta_1,\theta_2}\hat\sigma = 2(\theta_2-\theta_1)^2/((n-1)(n+2))$.

b) Momentų metodu gauname įvertinius $\tilde\mu = \bar X, \tilde\sigma = 2\sqrt{3}s$. Dispersijos

$$\boldsymbol{V}_{\theta_1,\theta_2}\tilde\mu = (\theta_2-\theta_1)^2/(12n); \quad \boldsymbol{V}_{\theta_1,\theta_2}\tilde\sigma = 12\frac{\mu_4-\mu_2^2}{4n\mu_2} + O(1/n\sqrt{n}).$$

Tegu $Y = (X-\theta_1)/(\theta_2-\theta_1) \sim U(0,\ 1)$; $Y - 1/2 = (X - \mathbf{E}X)/(\theta_2-\theta_1)$. Tada

$$\mu_4 = (\theta_2-\theta_1)^4\mathbf{E}(Y-1/2)^4 = (\theta_2-\theta_1)^4\int_{-1/2}^{1/2} z^4 dz = (\theta_2-\theta_1)^4/80.$$

Gauname

$$\boldsymbol{V}_{\theta_1,\theta_2}\tilde\sigma = 12\frac{\mu_4-\mu_2^2}{4n\mu_2} + O(1/n\sqrt{n}) = \frac{(\theta_2-\theta_1)^2}{5n} + O(\frac{1}{n\sqrt{n}}).$$

Momentų metodu gautų įvertinių dispersijos yra eilės $O(1/n)$, o NMD įvertinių eilės $O(1/n^2)$.

c) Tikėtinumo funkcija

$$L(\theta_1,\theta_2) = \frac{1}{(\theta_2-\theta_1)^n}\mathbf{1}_{(-\infty,\theta_2)}(X_{(n)})\mathbf{1}_{(\theta_1,\infty)}(X_{(1)})$$

įgyja maksimumą taške $\hat\theta_1 = X_{(1)}, \hat\theta_2 = X_{(n)}$. Remiantis invariantiškumo principu parametrų μ ir σ DT įvertiniai yra $(X_{(1)} + X_{(n)})/2$ ir $X_{(n)} - X_{(1)}$. Antrasis įvertinys yra paslinktas. Poslinkį galima atitaisyti (žr. p. a).

I.2.124. a) DT įvertiniai yra $\bar X$ ir m_2, o NMD įvertiniai yra $\bar X$ ir s^2;
b) DT įvertinys yra $n\eta/S$, $S = X_1 + ... + X_n$, o NMD įvertinys yra $(n\eta-1)/S$;
c) DT įvertinys yra $X_{(n)}$, o NMD įvertinys yra $(n+1)X_{(n)}/n$.

I.2.125. Tikėtinumo funkcija

$$L = L(\theta) = \prod_i (a_{X_i})\exp\{\sum_i X_i\ln\theta - n\ln(f(\theta))\},$$

$$\ell = \ln L = \ln(\prod_i (a_{X_i})) + n\bar{X}\ln\theta - n\ln(f(\theta)),$$

$$\dot{\ell} = 0 \quad \Leftrightarrow \quad \theta f'(\theta)/f(\theta) = \bar{X}.$$

Pažymėję $\eta = \ln\theta$, gauname kanoninio pavidalo eksponentinių skirstinių šeimą

$$L = L(\eta) = \prod_i (a_{X_i}) \exp\{\sum_i X_i\eta - n\ln(f(e^\eta))\}.$$

Tada momentų metodo lygtis

$$\mathbf{E}_\theta\bar{X} = (\ln(f(e^\eta)))' = e^\eta f'(e^\eta)/f(e^\eta) = \theta f'(\theta)/f(\theta) = \bar{X}.$$

I.2.126. $\hat{\lambda} = \bar{X}$; $\hat{p} = \bar{X}$; $-q/(p\ln p) = \bar{X}$; $\hat{\lambda}\exp\{\hat{\lambda}\}/(\exp\{\hat{\lambda}\} - 1) = \bar{X}$.

I.2.127. NMD įvertinius ir jų dispersijas žr. **I.2.95** ir **I.2.98** pratimuose. Remiantis invariantiškumo principu parametrų p^2 ir λ^2 DT įvertiniai yra $\widehat{p^2} = S^2/n^2$, $\widehat{\lambda^2} = S^2/n^2$. Kvadratinės rizikos funkcijos

$$\mathbf{E}_p(\widehat{p^2} - p^2) = \frac{4p^3q}{n} - \frac{p^2q(11p-7)}{n^2} + \frac{pq(1-6pq)}{n^3},$$

$$\mathbf{E}_\lambda(\widehat{\lambda^2} - \lambda^2) = \frac{4\lambda^3}{n} + \frac{5\lambda^2}{n^2} + \frac{\lambda}{n^3}.$$

I.2.128. a) Tikėtinumo funkcija

$$L(\theta) = \frac{1}{\theta^n}, \quad \theta < X_{(1)} \leq ... \leq X_{(n)} < 2\theta,$$

įgyja maksimalią reikšmę $(2/X_{(n)})^n$, kai $\hat{\theta} = X_{(n)}/2$.

b) Tikėtinumo funkcija

$$L(\theta) = 1, \quad \text{kai} \quad \theta - 1/2 < X_{(1)} \leq X_{(n)} < \theta + 1/2,$$

igyja maksimalią reikšmę 1, kai $X_{(n)} - 1/2 < \theta < X_{(1)} + 1/2$, ir reikšmę 0 kitais atvejais.

I.2.129. a) DT įvertinys $\hat{\theta} = X_{(n)}/2$ yra paslinktasis; $\mathbf{E}_\theta(\hat{\theta}) = \theta(2n+1)/(2n+2)$; imdami nepaslinktąjį įvertinį $\tilde{\theta} = X_{(n)}(n+1)/(2n+1)$, gauname $\boldsymbol{V}_\theta\tilde{\theta} = n\theta^2/((n+2)(2n+1)^2$; momentų metodo įvertinys $\hat{\theta} = 2\bar{X}/3$, $\boldsymbol{V}_\theta\hat{\theta} = \theta^2/(27n)$;

b) $\boldsymbol{V}_\theta((X_{(n)} + X_{(1)})/2) = n\theta^2/((n+1)^2(n+2))$; momentų metodo įvertinys $\hat{\theta} = \bar{X}$ ir $\boldsymbol{V}_\theta\hat{\theta} = \theta^2/(12n)$.

I.2.130. Tikėtinumo funkcijos

$$L(\theta) = (1/2^n)\exp\{-\sum_i |X_i - \theta|\}$$

maksimumas gaunamas, kai $\sum_i |X_i - \theta|$ yra minimali. Suma įgyja minimalią reikšmę, kai $\hat{\theta} = X_{k+1}$, jei $n = 2k+1$ nelyginis, ir kai $\theta \in (X_{(k)}, X_{(k+1)})$, jei $n = 2k$ lyginis, t. y. parametro θ DT įvertinys yra empirinė mediana $\hat{x}_{1/2}$. Remiantis [2], 2.4.2 teorema

$$\sqrt{n}(\hat{x}(0,5) - x(0,5)) \xrightarrow{d} Y \sim N(0, 1/(4f^2(x(0,5)))).$$

Kadangi $x(0,5) = 0, f^2(0) = 1$, tai

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} Y \sim N(0, 1/4).$$

I.2.131. Pažymėkime $\boldsymbol{S} = (S_1, ..., S_k)^T = \sum_i (X_{1i}, ..., X_{ki})^T$. Tada tikėtinumo funkcija

$$L(\pi_1, ..., \pi_{k-1}) = \prod_{i=1}^{n} (\pi_1^{X_{1i}} \cdot ... \cdot \pi_k^{X_{ki}}) = \pi_1^{S_1} \cdot ... \cdot \pi_k^{S_k} =$$

$$\exp\{\sum_{i=1}^{k-1} S_i \ln(\pi_i/\pi_k) + n \ln \pi_k\},$$

čia $\pi_k = 1 - \pi_1 - ... - \pi_{k-1}$, priklauso $(k-1)$-parametrei eksponentinio tipo skirstinių šeimai.

Statistika $\boldsymbol{T} = (S_1, ..., S_{k-1})^T$ yra pilnoji ir pakankamoji parametro $(\pi_1, ..., \pi_{k-1})^T$ statistika. Gauname DT lygčių sistemą

$$\frac{\partial \ell}{\partial \pi_i} = \frac{S_i}{\pi_i} - \frac{S_k}{\pi_k} \quad \Leftrightarrow \quad \frac{S_i}{\pi_i} = \frac{S_j}{\pi_j}, \ i, j = 1, ..., k.$$

Sumuodami pagal i lygybę $S_i \pi_j = \pi_i S_j$ ir turėdami omenyje, kad $\sum_i S_i = n, \sum_i \pi_i = 1$, gausime DT įvertinius $\hat{\pi}_j = S_j/n, j = 1, ..., k$;

$$\boldsymbol{V}(\hat{\pi}_j | \pi_j = \pi_j(1 - \pi_j)/n, \quad \mathbf{Cov}\,(\hat{\pi}_i, \hat{\pi}_j | \pi_i, \pi_j) = -\pi_i \pi_j / n, \quad i, j = 1, ..., k.$$

I.2.132. $\hat{\alpha} = (S_1 - S_2 - S_3)/n$; $S_j = \sum_i X_{ji}, j = 1, 2, 3$; $\sqrt{n}(\hat{\alpha} - \alpha) \xrightarrow{d} Y \sim N(0,\ (1 - \alpha^2))$.

I.2.133. $\hat{\alpha} = (S_3/(S_3 + S_4)$; $\hat{p} = (2S_1 + S_3 + S_4)/(2n)$.

I.2.134. $\hat{\alpha} = 0,1235$.

I.2.135. $\hat{\mu}_1 = \bar{X}$, $\hat{\mu}_2 = \bar{Y}$; $\hat{\sigma_{11}} = \sum_i (X_i - \bar{X})^2/n$, $\hat{\sigma_{22}} = \sum_i (Y_i - \bar{Y})^2/n$, $\hat{\sigma_{12}} = \hat{\sigma_{21}} = \sum_i (X_i - \bar{X})(Y_i - \bar{Y})/n$, $\hat{\rho} = \hat{\sigma_{12}}/\sqrt{\hat{\sigma_{11}}\hat{\sigma_{22}}}$. Parametro $\boldsymbol{\theta} = (\mu_1,\ \mu_2,\ \sigma_1^2,\ \sigma_2^2,\ \rho)^T$ informacinės matricos atvirkštinės $\boldsymbol{I}^{-1}(\boldsymbol{\theta}) = [I^{kl}]_{5\times5}$ elementai yra $I^{11} = \sigma_1^2/n$, $I^{22} = \sigma_2^2/n$, $I^{12} = I^{21} = \rho\sigma_1\sigma_2/n$; $I^{33} = 2\sigma_1^4/n$, $I^{44} = 2\sigma_2^4/n$, $I^{34} = I^{43} = 2\rho^2\sigma_1^2\sigma_2^2/n$; $I^{55} = (1 - \rho^2)^2/n$, $I^{35} = I^{53} = \rho(1 - \rho^2)\sigma_1^2/n$, $I^{45} = I^{54} = \rho(1 - \rho^2)\sigma_2^2/n$; $I^{kl} = 0$, kai $k = 1, 2$, o $l = 3, 4, 5$.

I.2.136. A. d. $Z = (X - \mu)/\sigma$ pradiniai momentai yra

$$\alpha_k = \mathbf{E}Z^k = \int_0^\infty (\ln x)^k e^{-x} dx = \Gamma^{(k)}(1), \quad k = 1, 2, ...,$$

čia $\Gamma^{(k)}(1)$ yra gama funkcijos k-oji išvestinė taške 1. Turėdami šiuos momentus, jais galime išreikšti ir centrinius a. d. Z momentus $\mu_k = \mathbf{E}(Z - \mathbf{E}Z)^k$:

$$\mu_2 = \alpha_2 - \alpha_1^2 = \Gamma^{(2)}(1) - [\Gamma^{(1)}(1)]^2.$$

$$\mu_3 = \Gamma^{(3)}(1) - 3\Gamma^{(2)}(1)\Gamma^{(1)}(1) + 2(\Gamma^{(1)}(1))^3,$$

$$\mu_4 = \Gamma^{(4)}(1) - 4\Gamma^{(3)}(1)\Gamma^{(1)}(1) + 6\Gamma^{(2)}(1)(\Gamma^{(1)}(1))^2 - 3(\Gamma^{(1)}(1))^4.$$

A. d. X pirmieji momentai yra

$$\mathbf{E}_{\mu,\sigma}X = \mu - \gamma\sigma, \quad \mathbf{E}_{\mu,\sigma}(X-\mu)^2 = \sigma^2\frac{\pi^2}{6}.$$

Randame parametrų μ ir σ įvertinius momentų metodu:

$$\hat{\mu} = \bar{X} + \gamma\hat{\sigma}, \quad \hat{\sigma} = \frac{\sqrt{6}}{\pi}s.$$

Remdamiesi [2], 2.5.5 skyreliu, gauname, kad šie įvertiniai asimptotiškai nepaslinktieji ir normalieji.

$$\sqrt{n}(\hat{\mu}, \hat{\sigma}) \xrightarrow{d} \boldsymbol{Y} \sim N_2(\mathbf{0},\ \boldsymbol{\Sigma}).$$

Kovariacinės matricos $\boldsymbol{\Sigma} = [\sigma_{ij}]_{2\times 2}$ elementai yra tokie:

$$\sigma_{11} = \sigma^2\left(\mu_2 + \frac{3\gamma^2(\mu_4 - \mu_2^2)}{2\pi^2\mu_2} - \frac{\gamma\sqrt{6}\mu_3}{\pi\sqrt{\mu_2}}\right),$$

$$\sigma_{22} = \sigma^2\frac{3(\mu_4 - \mu_2^2)}{2\pi^2\mu_2},$$

$$\sigma_{12} = \sigma_{21} = \sigma^2\left(-\frac{3\gamma(\mu_4 - \mu_2^2)}{2\pi^2\mu_2} + \frac{\sqrt{6}\mu_3}{2\pi\sqrt{\mu_2}}\right).$$

Ieškant parametro $\boldsymbol{\theta} = (\mu,\ \sigma)^T$ DT įvertinio $\widetilde{\boldsymbol{\theta}} = (\widetilde{\mu},\ \widetilde{\sigma})^T$, reikia μ ir σ atžvilgiu spręsti lygčių sistemą

$$\begin{cases} \frac{1}{n}\sum_i \exp\{\frac{X_i-\mu}{\sigma}\} = 1, \\ \frac{1}{n}\sum_i \frac{X_i-\mu}{\sigma}(1 + \exp\{\frac{X_i-\mu}{\sigma}\}) = 1. \end{cases}$$

DT įvertinių savybes galime gauti pasinaudoję [2], 3.5.3 teorema, kurios sąlygos tenkinamos.

DT įvertinys $\widetilde{\boldsymbol{\theta}} = (\widetilde{\mu},\ \widetilde{\sigma})^T$ yra asimptotiškai efektyvusis ir normalusis:

$$\sqrt{n}(\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} \boldsymbol{Z} \sim N_2(\mathbf{0},\ \boldsymbol{i}^{-1}(\boldsymbol{\theta})).$$

Informacinės matricos $\boldsymbol{i} = [i_{kl}]_{2\times 2}$ elementai yra $i_{11} = 1/\sigma^2$, $i_{12} = (1 + \Gamma'(1))/\sigma^2$, $i_{22} = (1 + 2\Gamma'(1) + \Gamma''(1))/\sigma^2$.

I.2.137. Momentų metodu gauname įvertinius

$$\hat{\lambda} = \frac{\bar{X}}{s^2},\ \hat{\eta} = \frac{\bar{X}^2}{s^2}.$$

Remiantis [2], 2.5.5 teorema, šie įvertiniai asimptotiškai ($n \to \infty$) nepaslinktieji ir normalieji

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} \boldsymbol{Z} \sim N_2(\mathbf{0},\ \boldsymbol{\Sigma}),$$

čia $\boldsymbol{\Sigma} = [\sigma_{ij}]_{2\times 2}$; $\sigma_{11} = \lambda^2(3 + 2\eta)/\eta,\ \sigma_{22} = 2\eta(1+\eta),\ \ \sigma_{12} = \sigma_{21} = 2\lambda(1+\eta)$.

DT metodu taškinius įvertinius randame iš lygčių sistemos

$$\begin{cases} n\widetilde{\eta} - \widetilde{\lambda} T_1 = 0, \\ n \ln \widetilde{\lambda} - n(\ln \Gamma(\widetilde{\eta}))' + \ln T_2 = 0. \end{cases}$$

Šie įvertiniai (remiantis [2], 3.5.3 teorema) asimptotiškai ($n \to \infty$) efektyvieji ir normalieji

$$\sqrt{n}(\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} \boldsymbol{Y} \sim N_2(\mathbf{0},\ \boldsymbol{G});$$

čia $\boldsymbol{G} = [g_{ij}]_{2\times 2}$; $g_{11} = a\lambda^2/(a\eta - 1)$, $g_{22} = \eta/(a\eta - 1)$, $g_{12} = g_{21} = \lambda/(a\eta - 1)$, $a = (\ln \Gamma(\eta))''$.

I.2.138. Momentų metodu gauname gana paprastą lygčių sistemą, iš kurios randame įvertinius

$$\hat{p} = \frac{\bar{X}}{s^2}, \quad \hat{k} = \frac{\bar{X}^2}{(s^2 - \bar{X})},$$

čia $\bar{X}$ ir s^2 – empiriniai vidurkio ir dispersijos analogai.

Remiantis [2], 3.5.1 skyreliu, šie įvertiniai asimptotiškai ($n \to \infty$) nepaslinktieji ir normalieji

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} \boldsymbol{Z} \sim N_2(\mathbf{0},\ \boldsymbol{\Sigma}), \quad \boldsymbol{\theta} = (p,\ k)^T,$$

čia $\boldsymbol{\Sigma} = [\sigma_{ij}]_{2\times 2}$; $\sigma_{11} = [p^2(1-p) + 2(1+k)]/k$, $\sigma_{22} = 2k(1+k)/(1-p)^2$, $\sigma_{12} = \sigma_{21} = 2p(1+k)/(1-p)$.

DT metodu gauname sudėtingesnę lygčių sistemą

$$\begin{cases} \widetilde{k} - \widetilde{p}(\widetilde{k} + \bar{X}) = 0, \\ \ln \widetilde{p} + \bar{Z} = 0; \end{cases}$$

čia

$$\bar{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i, \quad Z_i = \frac{\partial}{\partial \widetilde{k}} \ln \frac{\Gamma(\widetilde{k} + X_i)}{\Gamma(\widetilde{k})}.$$

DT įvertinio $\widetilde{\boldsymbol{\theta}} = (\widetilde{k},\ \widetilde{p})^T$ asimptotinį ($n \to \infty$) skirstinį gauname remdamiesi [2], 3.5.3 teorema

$$\sqrt{n}(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} \boldsymbol{Y} \sim N_2(\mathbf{0},\ \boldsymbol{i}^{-\mathbf{1}}(\boldsymbol{\theta}_\mathbf{0})).$$

Informacinės matricos $\boldsymbol{i} = [i_{kl}]_{2\times 2}$ elementai yra $i_{11} = k/(qp^2)$, $i_{12} = -1/p$, $i_{22} = \boldsymbol{V}_{p,k}(\Gamma'(k+X)/\Gamma(k+X))$.

I.2.139. Net pirmasis šio skirstinio momentas neegzistuoja. Parametras μ yra skirstinio mediana, σ – mastelio parametras.

Parametro μ taškiniu įvertiniu imkime empirinę medianą $\hat{\mu} = \hat{x}_{0,5} = X_{([n/2]+1)}$. Asimptotiškai ($n \to \infty$) šis įvertinys yra normalusis

$$\sqrt{n}(\hat{x}_{0,5} - \mu) \xrightarrow{d} Y \sim N(0,\ \pi^2\sigma^2/4).$$

Parametro σ įvertinys

$$\hat{\sigma} = (\hat{x}_{0,25} - \hat{x}_{0,75})/2,$$

asimptotiškai ($n \to \infty$) yra normalusis:

$$\sqrt{n}(\hat\sigma - \sigma) \xrightarrow{d} Y \sim N(0,\ \pi^2\sigma^2/4).$$

Jeigu parametrus μ ir σ vertintume DT metodu, tai gautieji įvertiniai $\widetilde{\mu}$ ir $\widetilde{\sigma}$ būtų asimptotiškai normalieji ir efektyvieji. Įvertinių $\widetilde{\mu}$ ir $\widetilde{\sigma}$ asimptotinės dispersijos lygios $2\sigma^2/n$. Taigi įvertinių $\hat\mu$ ir $\hat\sigma$ ASE, palyginti su DT įvertiniais $\widetilde{\mu}$ ir $\widetilde{\sigma}$, yra $8/\pi^2 \approx 0,81$. Tai yra, naudojant DT metodą, reikia apie 19 procentų mažiau stebinių tam pačiam tikslumui pasiekti.

I.2.141. b) $\hat\sigma = \sqrt{\pi/2}\sum_i |X_i|/n$, $\sqrt{n}(\hat\sigma - \sigma) \xrightarrow{d} Y \sim N(0,\ (\pi-2)\sigma^2/2)$;
c) $\sqrt{n}(\hat\sigma - \sigma) \xrightarrow{d} Y \sim N(0,\ \sigma^2/2)$.

I.2.142. a) Tikėtinumo funkcija

$$L(\mu) = e^{-\sum_i X_i} e^{n\mu} \mathbf{1}_{(\mu,\infty)}(X_{(1)}).$$

Remiantis faktorizacijos kriterijumi $X_{(1)}$ yra parametro μ pakankamoji statistika.

b) A. d. $X_{(1)}$ tankio funkcija yra $ne^{-n(x-\mu)}, x > \mu$; $\mathbf{E}_\mu X_{(1)} = \mu + 1/n$, $\boldsymbol{V}_\mu X_{(1)} = \mathbf{E}_\mu (X_{(1)} - \mu - 1/n)^2 = 1/n^2$;

$$\mathbf{P}_\mu\{|X_{(1)} - \mu| > \varepsilon\} = \mathbf{P}_\mu\{|X_{(1)} - \mathbf{E}_\mu X_{(1)} + 1/n| > \varepsilon\} =$$

$$\mathbf{P}_\mu\{X_{(1)} - \mathbf{E}_\mu X_{(1)} > \varepsilon - 1/n, X_{(1)} - \mathbf{E}_\mu X_{(1)} < -\varepsilon - 1/n\} \le$$

$$\mathbf{P}_\mu\{|X_{(1)} - \mathbf{E}_\mu X_{(1)}| > \varepsilon - 1/n\} \le \frac{\boldsymbol{V}_\mu X_{(1)}}{(\varepsilon - 1/n)^2} = \frac{1}{n^2(\varepsilon - 1/n)^2} \to 0,$$

kai $n \to \infty$, t. y. $X_{(1)} \xrightarrow{P} \mu$.

I.2.143. Tikėtinumo funkcija

$$L(\theta_1, \theta_2) = [a(\theta_1, \theta_2)]^n \prod_{i=1}^n h(X_i) \mathbf{1}_{(\theta_1,\infty)}(X_{(1)}) \mathbf{1}_{(-\infty,\theta_2)}(X_{(n)})$$

įgyja maksimalią reikšmę, kai $\theta_1 = X_{(1)}, \theta_2 = X_{(n)}$, nes $a(\theta_1, \theta_2) = 1/\int_{\theta_1}^{\theta_2} h(x)dx$. Parametrų θ_1 ir θ_2 DT įvertiniai yra $\hat\theta_1 = X_{(1)}$ ir $\hat\theta_2 = X_{(n)}$.

I.2.144. a)$\hat\mu_j = (X_j + Y_j)/2,\ j = 1, 2, ..., n$, $\widehat{\sigma^2} = \sum_i (X_i - Y_i)^2/(4n)$;

b) $\mathbf{E}_\sigma(\widehat{\sigma^2}) = \sigma^2/2$. Parametro $(\mu_1, ..., \mu_n, \sigma)^T$ dimensija $n+1$ priklauso nuo n, todėl teorema apie ribines DT įvertinių savybes netaikytina;

c) $\widehat{\sigma^2} = \sum_i (X_i - Y_i)^2/(2n)$; $\widehat{\sigma^2} \xrightarrow{P} \sigma^2$, nes $\mathbf{E}_\sigma(\widehat{\sigma^2}) = \sigma^2$, $\boldsymbol{V}_\sigma(\widehat{\sigma^2}) = \sigma^4/(2n)$.

I.2.145. a) Tikėtinumo funkcija

$$L(\lambda_1, ..., \lambda_n, \theta) = \exp\{-\theta \sum_i (\lambda_i X_i) + \sum_i \ln(\theta X_i) - \sum_i (\lambda_i Y_i) + \sum_i \ln \lambda_i\}.$$

DT įvertiniams rasti gauname lygčių sistemą

$$\dot\ell_{\lambda_i}(\hat\lambda_1, ..., \hat\lambda_n, \hat\theta) = -\hat\theta X_i - Y_i + 2/\hat\lambda_i = 0, \quad i = 1, 2, ..., n,$$

$$\dot{\ell}_\theta(\hat\lambda_1,...,\hat\lambda_n,\hat\theta) = -\sum_i(\lambda_i X_i) + n/\hat\theta = 0.$$

Išsprendę iš pirmųjų lygčių $\hat\lambda_i, i=1,...,n$ ir įrašę į paskutiniąją, gauname

$$\dot{\ell}_\theta(\hat\lambda_1,...,\hat\lambda_n,\hat\theta) = -\sum_i \frac{2X_i}{Y_i+\hat\theta X_i} + \frac{n}{\hat\theta} = 0 \quad\Rightarrow\quad \frac{n}{\hat\theta} - 2\sum_i \frac{R_i}{1+\hat\theta R_i} = 0.$$

b) A. d. R_i pasiskirstymo funkcija

$$F(x|\theta) = \mathbf{P}_\theta\{X_i \le xY_i\} = \int_0^\infty\int_{xv}^\infty \lambda^2\theta e^{-\lambda\theta u}e^{-\lambda v}dudv = \frac{1}{1+\theta x}$$

ir tankio funkcija

$$f(x|\theta) = F'(x|\theta) = \frac{\theta}{(1+\theta x)^2}, \quad x>0.$$

Imties $R_1,...,R_n$ tikėtinumo funkcija

$$L(\theta) = \theta^n\prod_i\frac{1}{(1+\theta R_i)^2} = \exp\{-2\sum_i\ln(1+\theta R_i) + n\ln\theta\}.$$

Prilyginę jos logaritmo išvestinę pagal θ nuliui, DT įvertiniui rasti gauname lygtį

$$\dot{\ell}_\theta(\hat\theta) = \frac{n}{\hat\theta} - 2\sum_i\frac{R_i}{1+\hat\theta R_i} = 0.$$

c) Informacijos kiekis

$$I(\theta) = -\mathbf{E}\ddot{\ell}_{\theta^2} = \frac{n}{\theta^2} - 2n\int_0^\infty\frac{\theta x^2dx}{(1+\theta x)^4} = \frac{3\theta^2}{n}.$$

Gauname

$$\sqrt{n}(\hat\theta-\theta) \xrightarrow{d} Y \sim N(0,3\theta^2).$$

d) $\hat\theta = 2,0102$; kadangi $\mathbf{E}_\theta(R_i)$ neegzistuoja, tai pradinį artinį $\hat\theta_0$ galima parinkti, pvz., naudojant empirinę medianą $\hat\theta_0 = 1/\hat x_{0,5}$;

e) $\hat\sigma_{\hat\theta} = \sqrt{3}\hat\theta/\sqrt{n} = 0,7785$.

I.2.146. $\hat\lambda_1 = K/(S_1+(n-K)x_0)$, $\hat\lambda_2 = (n-K)/S_2$; $S_1 = \sum_{i=1}^K X_{(i)}$, $S_2 = \sum_{i=K+1}^n(X_{(i)}-x_0)$; K – gedimų, mažesnių už x_0, skaičius; $\sqrt{n}(\hat\lambda_1-\lambda_1,\hat\lambda_2-\lambda_2) \xrightarrow{d} \boldsymbol{Y} \sim N_2(\mathbf{0},\ \mathbf{\Sigma})$; $\mathbf{\Sigma} = [\sigma_{ij}]_{2\times2}$; $\sigma_{11} = \lambda_1^2/(1-p)$, $\sigma_{12}=\sigma_{21}=0$; $\sigma_{22} = \lambda_2^2/p$, $p = e^{-\lambda_1 x_0}$.

I.2.147. a) $\hat\alpha = \bar X/m_2$, $\hat\gamma = \bar X^2/m_2$;

b) $\hat\theta = 1/\sqrt{m_2}$, $\hat a = \bar X + \sqrt{m_2}$;

c) $\hat\alpha = \bar X(1-\bar X(1-\bar X)/m_2)$, $\hat\beta = (1-\bar X)(1-\bar X(1-\bar X)/m_2)$;

d) $\hat\mu = \bar Y$, $\hat\sigma = \sqrt{m_2}$, $m_2 = (1/n)\sum_i(Y_i-\bar Y)$, $Y_i = \ln X_i$, $i=1,...,n$;

e) $\hat\theta = \bar X$; f) $\hat p = \bar X/m_2$, $\hat n = \bar X^2/(m_2-\bar X)$;

g) $-\hat q/\hat p\ln\hat p = \bar X$; h) $\hat k = \bar X$.

I.2.148. Pagal faktorizacijos kriterijų $\boldsymbol{T}$ yra pakankamoji parametro $\boldsymbol{\theta}$ statistika tada ir tik tada, kai tikėtinumo funkcija $L(\boldsymbol{\theta}) = q(\boldsymbol{T};\boldsymbol{\theta})W(\mathbf{X})$ ir $W(\mathbf{X})$ nuo parametro $\boldsymbol{\theta}$

nepriklauso. Funkcijos L maksimizavimas pagal $\boldsymbol{\theta}$ ekvivalentus funkcijos q maksimizavimui.

I.2.149. a) $\hat{\theta} = X_{(1)}$; b) $\hat{\theta} = X_{(1)}$;
c) $\hat{\theta} = n/\sum_j \ln(1 - X_j)$;
d) $\hat{\theta} = n/(n - \sum_j \ln X_j)$, $1/2 < \hat{\theta} < 1$;
e) $\hat{\theta} = \hat{x}_{1/2}$;
f) $\hat{\theta} = X_{(1)}$;
g) $\hat{\theta} = (\sqrt{\bar{X}^2 + 4a_2} - \bar{X})/2$;
h) $\hat{\mu} = X_{(1)}$, $\hat{\sigma} = n/\sum_i (X_i - X_{(1)})$;
i) $\hat{\mu} = (1/n)\sum_j \ln X_j = \bar{Y}$, $\hat{\sigma} = [\sum_j (\ln X_j - \bar{Y})^2/n]^{1/2}$;
j) $\hat{\theta} = 0$, kai $2^n \prod_j \sqrt{X_j} > 1$, ir $\hat{\theta} = 1$ priešingu atveju;
k) $\hat{\beta} = X_{(n)}$, $\hat{\alpha} = n/(n \ln \hat{\beta} - \sum_j \ln X_j)$;
l) $\hat{\theta} = X_{(n)}$, n>1.

I.2.150. a) $\hat{\lambda} = \bar{Y}$, $\hat{\mu} = \bar{Z}$;
b) Atsitiktinio vektoriaus $(X_i, \Delta_i)^T$ pasiskirstymo funkcija

$$F(x, k) = \int_0^x f_1^k(u, \lambda)[1 - F_1(u, \lambda)]^{1-k} f_2^{1-k}(u, \mu)[1 - F_2(u, \mu)]^k du.$$

Tikėtinumo funkcija

$$L(\lambda, \mu) = \prod_i \{(\frac{1}{\lambda} e^{-X_i/\lambda})^{\Delta_i} (e^{-X_i/\lambda})^{1-\Delta_i} (\frac{1}{\mu} e^{-X_i/\mu})^{1-\Delta_i} (e^{-X_i/\mu})^{\Delta_i}\}.$$

Randame parametrų λ ir μ DT įvertinius

$$\hat{\lambda} = \sum_i X_i / \sum_i \Delta_i, \quad \hat{\mu} = \sum_i X_i/(n - \sum_i \Delta_i).$$

I.2.151. Tikėtinumo funkcija

$$L(\theta) = \frac{1}{(1 - e^{-\theta})^n} \theta^{\sum_i X_i} e^{-n\theta},$$

iš kurios parametro θ DT įvertiniui rasti gauname lygtį

$$-\frac{1}{e^\theta - 1} + \frac{\bar{X}}{\theta} - 1 = 0 \quad \Leftrightarrow \quad \bar{X} = \frac{\theta}{1 - e^{-\theta}}.$$

Nagrinėkime funkciją $g(\theta) = \theta/(1 - e^{-\theta})$;

$$g'(\theta) = \frac{1 - e^{-\theta} - \theta e^{-\theta}}{(1 - e^{-\theta})^2}.$$

Skaitiklis yra didėjanti funkcija, nes $(1 - e^{-\theta} - \theta e^{-\theta})' = \theta e^{-\theta} > 0$. Skaitiklis lygus 0, kai $\theta = 0$. Taigi skaitiklis teigiamas su visais $\theta > 0$. Taške $\theta = 0$ funkcija $g(\theta) = 1$. Funkcija $g(\theta)$ didėja nuo 1 iki ∞, kai θ kinta nuo 0 iki ∞.

Lygtis $\bar{X} = \theta/(1 - e^{-\theta})$ turi vienintelį sprendinį, kai $\bar{X} > 1$.

I.2.152. $ASE = 1$.

I.2.153. $\hat{\alpha} = X_{(1)}$, $\hat{\theta} = n/S$, $S = \sum_j (\ln X_j - \ln X_{(1)})$; NMD įvertiniai $\hat{\alpha} = X_{(1)}(1 - S/n)$, $\hat{\theta} = (n-1)/S$; $ASE = 1$.

I.2.154. $\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}} \sim N_k(\boldsymbol{\mu},\ \boldsymbol{\Sigma}/n)$, $\hat{\boldsymbol{\Sigma}} = \sum_j (\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})^T/n = \boldsymbol{S}/n$; pažymėkime $\boldsymbol{\theta} = (\sigma_{ij},\ i \le j)$ vektorių, sudarytą iš skirtingų kovariacinės matricos $\boldsymbol{\Sigma}$ elementų, o $\hat{\boldsymbol{\theta}}$ DT įvertinių (t. y. atitinkamų matricos $\boldsymbol{S}/n$ elementų) vektorių. Tada $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} \boldsymbol{Z} \sim N_{k(k-1)/2}(\mathbf{0},\ \boldsymbol{\Gamma})$; čia kovariacinės matricos $\boldsymbol{\Gamma} = [\gamma_{ijrs}]_{k(k-1)/2 \times k(k-1)/2}$ elementai yra $\gamma_{iiii} = 2\sigma_{ii}^2$, $\gamma_{iijj} = 2\sigma_{ij}^2$, $\gamma_{iiij} = 2\sigma_{ii}\sigma_{ij}$, $i \ne j$, $\gamma_{ijrs} = (\sigma_{ir}\sigma_{js} + \sigma_{is}\sigma_{jr})$, $i \ne j \ne r \ne s$.

I.2.155. $\hat{\sigma}^2 = \sum_j (X_j - \hat{\mu})^2/n$, $\hat{\tau}^2 = \sum_j (Y_j - \hat{\mu})^2/n$, o įvertinys $\hat{\mu}$ randamas iš lygties $\hat{\mu} = [\bar{X} \sum_j (Y_j - \hat{\mu})^2 + \bar{Y} \sum_j (X_j - \hat{\mu})^2]/[\sum_j (Y_j - \hat{\mu})^2 + \sum_j (X_j - \hat{\mu})^2]$; šie įverčiai asimptotiškai nekoreliuoti ir normalieji: $\sqrt{n}(\hat{\sigma}^2 - \sigma^2) \xrightarrow{d} Z \sim N(0,\ 2\sigma^4)$, $\sqrt{n}(\hat{\tau}^2 - \tau^2) \xrightarrow{d} Y \sim N(0,\ 2\tau^4)$, $\sqrt{n}(\hat{\mu} - \mu) \xrightarrow{d} V \sim N(0,\ \sigma^2\tau^2/(\sigma^2 + \tau^2))$.

I.2.156. Palyginę tikimybes $b(X|n,p) = C_n^X p^X (1-p)^{n-X}$ ir $b(X|n+1,p)$ gausime, kad maksimumas pasiekiamas, kai $\hat{n} = [X/p]$, jei X/p nėra sveikasis skaičius, ir $\hat{n} = X/p - 1$, jei X/p – sveikasis skaičius.

I.2.157. Palyginę tikimybes $h(X|N,M,n) = C_M^X C_{N-M}^{n-X}/C_N^n$ ir $h(X|N,M+1,n)$ gausime, kad maksimumas pasiekiamas, kai $\hat{M} = [X(N+1)/n]$, jei $X(N+1)/n$ nėra sveikasis skaičius, ir $\hat{M} = X(N+1)/n - 1$, jei $X(N+1)/n$ – sveikasis skaičius.

I.2.5 skyrelis

I.2.158. Parametro μ pilnoji ir pakankamoji statistika $\bar{X}$ ir funkcijos $T(\bar{X};\mu) = \sqrt{n}(\bar{X} - \mu)/\sigma \sim N(0,\ 1)$ skirstinys nepriklauso nuo nežinomo parametro μ. Parinkime tokius skaičius $k_1 < k_2$, kad

$$\mathbf{P}_\mu\{k_1 < \sqrt{n}(\bar{X} - \mu)/\sigma < k_2\} = Q = 1 - 2\alpha.$$

Išsprendę nelygybes μ atžvilgiu, gausime parametro μ pasikliovimo intervalą. Intervalo ilgis bus tuo mažesnis, kuo mažesnis $k_2 - k_1$, t. y. taškai k_1, k_2 turėtų būti simetriški 0 atžvilgiu. Gauname $k_1 = -z_\alpha, k_2 = z_\alpha$. Pasikliovimo intervalas

$$(\underline{\mu}; \overline{\mu}) = (\bar{X} - z_\alpha\sigma/\sqrt{n}; \bar{X} + z_\alpha\sigma/\sqrt{n}).$$

I.2.159. Parametro σ^2 pilnoji ir pakankamoji statistika $s_0^2 = \sum_i (X_i - \mu)^2/n$ ir funkcijos $T(s_0^2;\sigma^2) = s_0^2 n/\sigma^2 \sim \chi^2(n)$ skirstinys nepriklauso nuo nežinomo parametro σ^2. Parinkime konstantas $k_1 < k_2$, kad

$$\mathbf{P}_\sigma\{k_1 < s_0^2 n/\sigma^2 < k_2\} = Q = 1 - 2\alpha.$$

Imdami simetriškus rėžius $k_1 = \chi^2_{1-\alpha}(n), k_2 = \chi^2_\alpha(n)$, gauname pasikliovimo intervalą

$$(\underline{\sigma^2}; \overline{\sigma^2}) = \left(\frac{s_0^2 n}{\chi^2_\alpha(n)}, \frac{s_0^2 n}{\chi^2_{1-\alpha}(n)}\right).$$

Kadangi $\sigma = h(\sigma^2) = \sqrt{\sigma^2}$ yra monotoniška funkcija, tai gauname

$$(\underline{\sigma}; \overline{\sigma}) = (\sqrt{\underline{\sigma^2}}; \sqrt{\overline{\sigma^2}}).$$

I.2.160. Parametro $(\mu, \sigma^2)^T$ pilnoji ir pakankamoji statistika $(\bar{X}, s^2)^T$. Remiantis [2], 2.4.1 teorema funkcijų $s^2(n-1)/\sigma^2 \sim \chi^2(n-1)$, $\sqrt{n}(\bar{X}-\mu)/s \sim S(n-1)$ skirstiniai nepriklauso nuo nežinomų parametrų. Analogiškai **I.2.158** ir **I.2.159** pratimams gauname lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalus

$$(\underline{\mu}; \overline{\mu}) = (\bar{X} - t_\alpha(n-1)s/\sqrt{n}; \bar{X} + t_\alpha(n-1)s/\sqrt{n}),$$

$$(\underline{\sigma^2}; \overline{\sigma^2}) = \left(\frac{s^2(n-1)}{\chi^2_\alpha(n-1)}, \frac{s^2(n-1)}{\chi^2_{1-\alpha}(n-1)} \right).$$

I.2.161. Mažiausio vidutinio santykinio ilgio intervalą gauname, kai $\alpha_1 = \alpha_1^*, \alpha_2 = 1 - Q - \alpha_1^*$, o α_1* randamas iš sąlygos

$$\min_{0 \le \alpha_1 \le 1-Q} b(\alpha_1) = b(\alpha_1^*), \quad b(\alpha_1) = (1/\chi^2_{1-\alpha_1}(n-1) - 1/\chi^2_{1-Q-\alpha_1}(n-1)).$$

Simetrišku atveju $\alpha_1 = \alpha_2 = (1-Q)/2$.

Mažiausio vidutinio santykinio ilgio pasikliovimo intervalo ilgio santykis su simetriško intervalo vidutiniu santykiniu ilgiu, kai $Q = 0,95$ ir $n = 10; 20; 50; 100$, yra 0,8378; 0,9156; 0,9652; 0,9826.

I.2.162. *Nurodymas.* Pasiremkite tuo, kad a. d. $\bar{X}$ ir s^2 yra nepriklausomi.

I.2.163. Randame

$$\overline{\mu} - \underline{\mu} = 0,4, \quad z_\alpha = 0,8, \quad z_\alpha = 2, \quad Q = 1 - 2\alpha = 2\Phi(2) - 1 \approx 0,9544.$$

I.2.164. Gauname

$$\overline{\mu} - \underline{\mu} = 2z_{0,025}/\sqrt{n} \le 0,1 \Leftrightarrow n \ge (20z_{0,025})^2 = 1537.$$

I.2.165. Tegu $\alpha = (1-Q)/2 = 0,025$. Gauname

$$(\underline{\mu}; \overline{\mu}) = (\bar{X} - t_\alpha(24)s/5; \bar{X} + t_\alpha(24)s/5) = (6,289; 6,379);$$

$$(\underline{\sigma}; \overline{\sigma}) = (\sqrt{\underline{\sigma^2}}; \sqrt{\overline{\sigma^2}}) = (0,0855; 0,1524).$$

I.2.166. a) Funkcija $T = \beta s_2^2/s_1^2 \sim F(n-1, m-1)$ yra monotoniška pagal β, o jos skirstinys nuo nežinomų parametrų nepriklauso. Remdamiesi [2], 3.6.4 pastaba gauname lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą

$$(\underline{\beta}; \overline{\beta}) = \left(\frac{s_1^2}{s_2^2} F_{1-\alpha}(n-1, m-1), \frac{s_1^2}{s_2^2} F_\alpha(n-1, m-1) \right).$$

b) Parametro σ^2 NMD įvertinys $s^2 = (s_1^2(m-1)/k + s_2^2(n-1))/(m+n-2)$, $s^2(m+n-2)/\sigma^2 \sim \chi^2(m+n-2)$. Analogiškai **I.2.159** pratimui gauname

$$(\underline{\sigma^2}; \overline{\sigma^2}) = \left(\frac{s^2(m+n-2)}{\chi^2_\alpha(m+n-2)}, \frac{s^2(m+n-2)}{\chi^2_{1-\alpha}(m+n-2)} \right).$$

Parametro $\theta = \mu_1 - \mu_2$ NMD įvertinys $\hat\theta = \bar X - \bar Y \sim N(\theta, k\sigma_2^2/m + \sigma_2^2/n)$. Tada

$$\frac{\bar X - \bar Y - \theta}{\sigma_{\hat\theta}} \sim S(m+n-2), \quad \sigma_{\hat\theta}^2 = s^2\frac{kn+m}{mn};$$

$$(\underline{\theta}; \overline{\theta}) = (\bar X - \bar Y - t_\alpha(m+n-2)\sigma_{\hat\theta}, \bar X - \bar Y + t_\alpha(m+n-2)\sigma_{\hat\theta}).$$

I.2.167. Parametro θ NMD įvertinys $\hat\theta = \bar X - \bar Y \sim N(\theta, \sigma_1^2/m + \sigma_2^2/n)$. Kadangi $s_1^2 \xrightarrow{P} \sigma_1^2$ ir $s_2^2 \xrightarrow{P} \sigma_2^2$, tai

$$\frac{\bar X - \bar Y - \theta}{\sqrt{s_1^2/m + s_2^2/n}} \xrightarrow{P} Z \sim N(0,\ 1), \quad m, n \to \infty.$$

Gauname lygmens $Q = 1 - 2\alpha$ asimptotinį pasikliovimo intervalą

$$(\underline{\theta}; \overline{\theta}) = (\bar X - \bar Y - z_\alpha\sqrt{s_1^2/m + s_2^2/n}, \bar X - \bar Y + z_\alpha\sqrt{s_1^2/m + s_2^2/n}).$$

I.2.168. A. v. $(X, Y)^T$ skirstinys priklauso penkių parametrų eksponentinio tipo skirstinių šeimai. Pilnoji ir pakankamoji statistika

$$\boldsymbol{T} = \left(\sum_i X_1, \sum_i Y_i, \sum_i X_i^2, \sum_i Y_i^2, \sum_i X_i Y_i\right)^T.$$

Parametro θ NMD įvertinys $\hat\theta = \bar X - \bar Y \sim N(\theta, \sigma^2)$, $\sigma^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$. Tegu $Z_i = X_i - Y_i$, tada $Z_1, ..., Z_n$ yra paprastoji imtis a. d. $Z \sim N(\theta, \sigma^2)$. Parametro θ lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalas randamas analogiškai **2.160** pratimui

$$(\underline{\theta}; \overline{\theta}) = (\bar X - \bar Y - t_\alpha(n-1)s/\sqrt{n}; \bar X - \bar Y + t_\alpha(n-1)s/\sqrt{n}),$$

čia $s^2 = \sum_i (Z_i - \bar Z)^2/(n-1)$.

I.2.169. Parametro ρ DT įvertinys $\hat\rho$ yra empirinis koreliacijos koeficientas

$$\hat\rho = r = \frac{\sum_i (X_i - \bar X)(Y_i - \bar Y)}{\sqrt{\sum_i (X_i - \bar X)^2 \sum_i (Y_i - \bar Y)^2}}.$$

Statistikos r tankio funkcija $f(r|\rho)$ pateikta [5], 4.1 skyrelyje

$$f(r|\rho) = \frac{2^{n-3}(1-\rho^2)^{(n-1)/2}(1-r^2)^{(n-4)/2}}{\pi(n-3)!}\sum_{j=0}^\infty \frac{(2r\rho)^j}{j!}\Gamma^2(\frac{n-1+j}{2}).$$

Pasiskirstymo funkcija $F(r|\rho) = \int_{-1}^r f(u|\rho)du$ yra monotoniškai didėjanti pagal ρ. Remiantis Bolševo teorema (žr. [2], 3.6.1 teoremą) lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalo rėžiai randami iš lygčių

$$F(r|\underline{\rho}) = 1 - \alpha, \quad F(r|\bar\rho) = \alpha.$$

I.2.170. Statistikos r skirstinio aproksimacija

$$\sqrt{n}\frac{r - \rho}{1 - \rho^2} \xrightarrow{d} Z \sim N(0,\ 1)$$

gali būti netiksli, ypač kai ρ artimas ± 1 ir r skirstinys yra labai asimetriškas.

Rekomenduojama taikyti Fišerio pasiūlytą dispersiją stabilizuojančią transformaciją

$$V = V(r) = \frac{1}{2}\ln\frac{1+r}{1-r}, \quad \sqrt{n-3}(V(r) - V(\rho)) \xrightarrow{d} Z \sim N(0,\, 1).$$

Pritaikę šią aproksimaciją ir remdamiesi tuo, kad $V(\rho)$ monotoniška funkcija, gauname lygmens $Q = 1 - 2\alpha$ aproksimacinį intervalą

$$\underline{\rho} = \frac{e^{2V_1-1}}{e^{2V_1+1}}, \quad \bar{\rho} = \frac{e^{2V_2-1}}{e^{2V_2+1}},$$

$$V_1 = \frac{1}{2}\ln\frac{1+r}{1-r} - z_\alpha\frac{1}{\sqrt{n-3}}, \quad V_2 = \frac{1}{2}\ln\frac{1+r}{1-r} + z_\alpha\frac{1}{\sqrt{n-3}}.$$

I.2.171. Parametro σ^2 NMD įvertinys

$$\hat{\sigma^2} = \sum_{i=1}^{n} X_i^2/(2n), \quad 2n\hat{\sigma^2}/\sigma^2 \sim \chi^2(2n).$$

Analogiškai **I.2.159** pratimui gauname pasikliovimo intervalą $\alpha = (1-Q)/2$

$$(\underline{\sigma}; \overline{\sigma}) = (\sqrt{2n\hat{\sigma^2}/\chi^2_\alpha(2n)}, \sqrt{2n\hat{\sigma^2}/\chi^2_{1-\alpha}(2n)}).$$

I.2.172. Parametro σ^2 NMD įvertinys

$$\hat{\sigma^2} = \sum_{i=1}^{n} X_i^2/(3n), \quad 3n\hat{\sigma^2}/\sigma^2 \sim \chi^2(3n).$$

Analogiškai **I.2.159** pratimui gauname pasikliovimo lygmens $Q = 1 - 2\alpha$ intervalą

$$(\underline{\sigma^2}; \overline{\sigma^2}) = (3n\hat{\sigma^2}/\chi^2_\alpha(3n), 3n\hat{\sigma^2}/\chi^2_{1-\alpha}(3n)).$$

I.2.173. Parametro λ pilnoji ir pakankamoji statistika yra $T = X_1 + ... + X_n \sim G(\lambda, \eta)$, $\eta = \eta_1 + ... + \eta_n$. Funkcija $Y = 2\lambda T \sim \chi^2(2\eta)$ yra monotoniška pagal λ, o jos skirstinys nuo nežinomo parametro nepriklauso. Remdamiesi [2], 3.6.4 pastaba gauname pasikliovimo intervalą

$$\underline{\lambda} = \chi^2_{1-\alpha}(2\eta)/(2T), \quad \bar{\lambda} = \chi^2_\alpha(2\eta)/(2T).$$

I.2.174. A. d. $Y = e^X$ turi eksponentinį skirstinį su parametru $\lambda = e^{-\mu}$. Sudarę parametro λ pasikliovimo intervalą (žr. **I.2.173** pratimą)

$$\underline{\lambda} = \chi^2_{1-\alpha}(2n)/(2T), \quad \overline{\lambda} = \chi^2_\alpha(2n)/(2T), \quad T = Y_1 + ... + Y_n,$$

gauname ir parametro μ pasikliovimo intervalą

$$\underline{\mu} = \ln(1/\overline{\lambda}), \quad \overline{\mu} = \ln(1/\underline{\lambda}).$$

I.2.175. Pilnosios ir pakankamosios statistikos $T = X_1 + ... + X_n \sim \mathcal{P}(n\lambda)$ pasiskirstymo funkcija

$$F(t; \lambda) = \sum_{k=0}^{t} \frac{(n\lambda)^k}{k!} e^{-n\lambda} = \mathbf{P}\{\chi^2_{2t+2} > 2n\lambda\}$$

yra mažėjanti pagal λ, kai $t > 0$. Remiantis Bolševo teorema (žr. [2], 3.6.1 teoremą) pasikliovimo intervalo rėžiai randami iš lygčių

$$F(T;\bar{\lambda}) = \mathbf{P}\{\chi^2_{2T+2} > 2n\bar{\lambda}\} = \alpha,$$

$$F(T;\underline{\lambda}) = \mathbf{P}\{\chi^2_{2T} > 2n\underline{\lambda}\} = 1-\alpha$$

ir $\underline{\lambda} = 0$, kai $T = 0$. Išsprendę gauname

$$\underline{\lambda} = \frac{1}{2n}\chi^2_{1-\alpha}(2T), \quad \overline{\lambda} = \frac{1}{2n}\chi^2_{\alpha}(2T+2).$$

I.2.176. Pilnosios ir pakankamosios statistikos $T = X_1 + ... + X_n \sim B(n,p)$ pasiskirstymo funkcija

$$F(t;p) = \sum_{m=0}^{t} C_n^m p^t (1-p)^{n-t} = I_{1-p}(m-t, t+1)$$

yra mažėjanti pagal p, kai $0 < t < n$, čia $I_x(\gamma,\eta) = \mathbf{P}\{U < x\}$, $U \sim Be(\gamma,\eta)$. Remiantis Bolševo teorema pasikliovimo intervalo rėžiai randami iš lygčių

$$F(T;\overline{p}) = I_{1-\overline{p}}(m-T, T+1) = \alpha,$$

$$F(T-0;\underline{p}) = I_{1-\underline{p}}(m-T+1, T) = 1-\alpha$$

ir $\underline{p} = 0$, kai $T = 0$, $\overline{p} = 1$, kai $T = n$. Pažymėję $\beta_\alpha(\gamma,\eta)$ beta skirstinio su parametrais γ ir η lygmens α kritinę reikšmę, gauname

$$\underline{p} = 1 - \beta_\alpha(m-T+1, T), \quad \text{kai} \quad T > 0,$$

$$\overline{p} = 1 - \beta_{1-\alpha}(m-T, T+1), \quad \text{kai} \quad T < n.$$

I.2.177. Parametro $(\lambda_1, \lambda_2)^T$ pilnoji ir pakankamoji statistika yra $\boldsymbol{T} = (T_1, T_2)^T$, $T_1 = X_1 + ... + X_m \sim \mathcal{P}(m\lambda_1)$, $T_2 = Y_1 + ... + Y_n \sim \mathcal{P}(n\lambda_2)$. A. d. T_1 sąlyginis skirstinys, kai $T_1 + T_2$ fiksuotas, yra binominis

$$(T_1 | T_1 + T_2 = N) \sim B(N, p), \quad p = m\lambda_1/(m\lambda_1 + n\lambda_2).$$

Suradę tikimybės p pasikliovimo intervalą $(\underline{p}, \overline{p})$, gauname ir parametro β pasikliovimo intervalą

$$\underline{\beta} = \frac{n}{m}\frac{\underline{p}}{1-\underline{p}}, \quad \overline{\beta} = \frac{n}{m}\frac{\bar{p}}{1-\bar{p}}.$$

I.2.178. Pilnosios ir pakankamosios statistikos $T = X_1 + ... + X_n \sim B^-(m,p)$, $m = m_1 + ... + m_n$, pasiskirstymo funkcija

$$F(t;p) = \sum_{i=0}^{t} C_{m+i-1}^{m-1} p^m (1-p)^i = I_p(m, t+1)$$

yra didėjanti pagal p, kai $0 < T$. Remiantis Bolševo teorema (žr. [2], 3.6.1 teoremą) pasikliovimo intervalo rėžiai randami iš lygčių

$$F(T-0;\overline{p}) = I_{\overline{p}}(m, T) = \alpha,$$

$$F(T;\underline{p}) = I_{\underline{p}}(m, T+1) = 1-\alpha$$

ir $\overline{p} = 0$, kai $T = 0$. Pažymėję $\beta_\alpha(\gamma, \eta)$ beta skirstinio su parametrais γ ir η lygmens α kritinę reikšmę, gauname

$$\underline{p} = \beta_{1-\alpha}(m, T+1), \quad \overline{p} = \beta_\alpha(m, T).$$

I.2.179. Parametro M pilnoji ir pakankamoji statistika $T = X_1 + ... + X_n \sim H(N, M, n)$. Pasiskirstymo funkcija

$$F(t; M) = \sum_{m=0}^{t} \frac{C_M^m C_{N-M}^{n-m}}{C_N^n} = H(t|N, M, n)$$

yra monotoniškai mažėjanti pagal M. Remiantis Bolševo teorema parametro M apatinis lygmens $(1-\alpha)$ pasikliovimo rėžis $\underline{M}$ yra didžiausias sveikasis skaičius, tenkinantis nelygybę

$$F(T-1, \underline{M}) = H(T-1|N, \underline{M}, n) \geq 1-\alpha,$$

o viršutinis lygmens $(1-\alpha)$ pasikliovimo rėžis $\bar{M}$ yra mažiausias sveikasis skaičius, tenkinantis nelygybę

$$F(T, \bar{M}) = H(T|N, \bar{M}, n) \leq \alpha.$$

I.2.180. Parametro M NMD įvertinys yra $\hat{M} = NT/n$; jo realizacija $\hat{M} = 300 \times 6/50 = 36$. Pasikliovimo rėžius gauname spręsdami **I.2.179** pratime pateiktas nelygybes

$$H(5|300, \underline{M}, 50) \geq 0,95 \quad \Leftrightarrow \quad \underline{M} \geq 13, \quad H(6|300, \overline{M}, 50) \leq 0,05 \quad \Leftrightarrow \quad \overline{M} \leq 58.$$

Pasikliovimo rėžiai $\underline{M} = 13$, $\overline{M} = 58$.

I.2.181. Parametrų $\beta_0, \beta_1, \sigma^2$ NMD įvertiniai (žr. **I.1.69** pratimą) $\hat{\beta}_0 = \bar{Y} \sim N(\beta, \sigma^2/n)$, $\hat{\beta}_1 = \sum_i Y_i(x_i - \bar{x})/\sum_i(x_i-\bar{x})^2 \sim N(\beta_1, \sigma^2/\sum_i(x_i-\bar{x})^2)$;

$$\hat{\sigma^2} = s^2 = SS_E/(n-2), \quad SS_E = \sum_i (Y_i - \hat{\beta}_0 - \hat{\beta}_1(x_i - \bar{x}))^2, \quad s^2(n-2)/\sigma^2 \sim \chi^2(n-2).$$

a) Analogiškai **I.2.160** pratimui gauname lygmens $Q = 1-2\alpha$ pasikliovimo intervalą

$$(\underline{\sigma^2}, \overline{\sigma^2}) = (s^2(n-2)/\chi^2_\alpha(n-2), s^2(n-2)/\chi^2_{1-\alpha}(n-2)).$$

b) Įvertiniai $\hat{\beta}_0, \hat{\beta}_1, s^2$ nepriklausomi (žr. **I.1.69** pratimą). Tada

$$\hat{\theta} = a\hat{\beta}_0 + b\hat{\beta}_1 \sim N(\theta, \sigma^2_\theta), \quad \sigma^2_\theta = \sigma^2(a^2/n + b^2/\sum_i(x_1 - \bar{x})^2)$$

ir $(\hat{\theta} - \theta)/\sigma_{\hat{\theta}} \sim S(n-2)$, $\sigma^2_{\hat{\theta}} = s^2(a^2/n + b^2/\sum_i(x_1-\bar{x})^2)$. Gauname pasikliovimo intervalą

$$(\ underline\theta, \overline{\theta}) = (\hat{\theta} - t_\alpha(n-2)\sigma_{\hat{\theta}}, \hat{\theta} + t_\alpha(n-2)\sigma_{\hat{\theta}}).$$

Imdami $a = 1, b = 0$; $a = 0, b = 1$; $a = 1, b = x - \bar{x}$ gausime parametrų $\beta_0, \beta_1, \beta_0 + \beta_1(x - \bar{x})$ pasikliovimo intervalus.

I.2.182. Pasikliovimo intervalą gauname **I.2.181** pratimo p. b) imdami $a = 4, b = 4\bar{\tau}$.

I.2.183. Parametro θ pilnoji ir pakankamoji statistika yra $T = X_{(n)}$. A. d. $T = X_{(n)}$ pasiskirstymo funkcija yra $F(t|\theta) = (t/\theta)^n, 0 < t < \theta$. Fiksuotam $t > 0$ funkcija $t/\theta^n, \theta > t$ yra mažėjanti θ atžvilgiu. Pagal Bolševo teoremą lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalo rėžiai gaunami iš lygčių

$$(X_{(n)}/\underline{\theta})^n = 1 - \alpha, \quad (X_{(n)}/\overline{\theta})^n = \alpha.$$

Gauname pasikliovimo intervalą

$$(\underline{\theta}; \overline{\theta}) = (X_{(n)}/(1-\alpha)^{1/n}; \overline{\theta} = X_{(n)}/\alpha^{1/n}).$$

I.2.184. Pilnoji ir pakankamoji statistika yra $T = |X|_{(n)}$ ir $|X| \sim U(0, \theta)$. Gauname analogišką intervalą, kaip ir **I.2.183** pratime.

I.2.185. Kadangi $\mathbf{P}_\theta\{X_{(n)} - 1/2 < \theta < X_{(1)} + 1/2\} = 1$, tai intervalas $(\underline{\theta}; \overline{\theta}) = (X_{(n)} - 1/2; X_{(1)} + 1/2)$ uždengia nežinomą parametrą θ su tikimybe 1. Šio intervalo ilgis $X_{(1)} + 1/2 - (X_{(n)} - 1/2) = 1 - (X_{(n)} - X_{(1)}) = Y_{(1)} + 1 - Y_{(n)}$, kai $Y_i \sim U(0,\ 1)$. Taigi intervalo ilgis turi beta skirstinį $Be(2, n-1)$ ir jo vidurkis $2/(n+1) = O(1/n)$.

I.2.186. Pažymėkime $Y_i = (X_i - \theta_1)/\theta_2 \sim U(0,\ 1)$. Tada funkcijos $T = (X_{(n)} - X_{(1)})/\theta = Y_{(n)} - Y_{(1)} \sim Be(n-1, 2)$ tankio funkcija $f(x; n) = n(n-1)x^{n-2}(1-x)$ ir pasiskirstymo funkcija $F(x; n) = nx^{n-1} - (n-1)x^n$. Pažymėkime $T_{1-\alpha}$ ir T_α šios pasiskirstymo funkcijos eilės $1-\alpha$ ir α kritines reikšmes. Jos randamos iš lygčių

$$nT_{1-\alpha}^{n-1} - (n-1)T_{1-\alpha}^n = \alpha, \quad nT_\alpha^{n-1} - (n-1)T_\alpha^n = 1 - \alpha.$$

Gauname lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą

$$(\underline{\theta}_2; \overline{\theta}_2) = ((X_{(n)} - X_{(1)})/T_\alpha; (X_{(n)} - X_{(1)})/T_{1-\alpha}).$$

I.2.187. Pilnoji ir pakankamoji statistika yra $(T_1, T_2)^T = (X_{(1)}, n(\bar{X} - X_{(1)}))^T$ (žr. **I.1.27** pratimą). a) Funkcija $2T_2/\theta \sim \chi^2(2n-2)$. Gauname lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą

$$(\underline{\theta}; \overline{\theta}) = (2T_2/\chi^2_\alpha(2n-2); 2T_2/\chi^2_{1-\alpha}(2n-2)).$$

b) Funkcija $T_1 - a = X_{(1)} - a \sim \mathcal{E}(n/\theta)$. Šios funkcijos pasiskirstymo funkcija $F(x) = 1 - e^{-nx/\theta}$. Randame eilės $1-\alpha$ ir α kritines reikšmes $u_{1-\alpha}$ ir u_α:

$$F(u_{1-\alpha}) = \alpha, \quad F(u_\alpha) = 1 - \alpha \quad \Rightarrow \quad u_{1-\alpha} = -\frac{\theta}{n}\ln(1-\alpha), \;\; u_\alpha = -\frac{\theta}{n}\ln\alpha.$$

Randame parametro a lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą

$$(\underline{a}; \overline{a}) = (X_{(1)} + (\theta/n)\ln\alpha; X_{(1)} + (\theta/n)\ln(1-\alpha)).$$

Kai ir parametras θ nežinomas, tai remsimės tuo, kad T_1 ir T_2 nepriklausomi ir

$$(2n/\theta)(X_{(1)} - a) \sim \chi^2(2), \quad 2T_2/\theta \sim \chi^2(2n-2).$$

Tada funkcija $n(n-2)(X_{(1)} - a)/T_2 \sim F(2, 2n-2)$. Iš čia lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalas yra

$$(\underline{a}; \overline{a}) = (X_{(1)} - (T_2/(n(n-2)))F_{1-\alpha}(2, 2n-2), X_{(1)} - (T_2/(n(n-2)))F_\alpha(2, 2n-2)).$$

c) Tegu

$$C(T_1,T_2) = \{(a,\theta) : X_{(1)} + (\theta/n)\ln\alpha_1 < a < X_{(1)} + (\theta/n)\ln(1-\alpha), \underline{\theta} < \theta < \overline{\theta}\}.$$

Tada $\mathbf{P}_{a,\theta}\{(a,\theta) \in C(T_1,T_2)\} = (1-2\alpha_1)(1-2\alpha)$.

I.2.188. A. d. $X_{n+1} - \bar{X} \sim N(0, \sigma^2(1+1/n))$. Tada $(X_{n+1} - \bar{X})/s\sqrt{1+1/n} \sim S(n-1)$. Iš čia randame

$$(\underline{X}_{n+1}; \overline{X}_{n+1}) = (\bar{X} - t_\alpha(n-1)s\sqrt{1+1/n}; \bar{X} + t_\alpha(n-1)s\sqrt{1+1/n}).$$

I.2.189. a) $\sqrt{n}(\tilde{\mu} - \mu) \xrightarrow{d} Z \sim N(0, \pi^2/4)$. Gauname lygmens $Q = 1 - 2\alpha$ aproksimacinį pasikliovimo intervalą

$$(\underline{\mu}; \overline{\mu}) = (\tilde{\mu} - z_\alpha\pi/(2\sqrt{n}); \tilde{\mu} + z_\alpha\pi/(2\sqrt{n})).$$

b) $\sqrt{n}(\hat{\mu} - \mu) \xrightarrow{d} Y \sim N(0,\ 2)$. Gauname lygmens $Q = 1-2\alpha$ aproksimacinį pasikliovimo intervalą

$$(\underline{\mu}; \overline{\mu}) = (\hat{\mu} - z_\alpha\sqrt{2/n}); \hat{\mu} + z_\alpha\sqrt{2/n})).$$

Intervalo a) ilgis yra $\pi/(2\sqrt{2}) \approx 1,11$ karto didesnis už intervalo b) ilgį.

I.2.190. a) $(\underline{\lambda}; \overline{\lambda}) = (\bar{X} - z_\alpha\sqrt{\bar{X}/n}; \bar{X} + z_\alpha\sqrt{\bar{X}/n})$.

b) Spręsdami antrojo laipsnio nelygybę, gauname

$$(\underline{\lambda}; \overline{\lambda}) = (\bar{X} + z_\alpha^2/(2n) - z_\alpha\sqrt{\bar{X}/n + z_\alpha^2/(4n^2)}; \bar{X} + z_\alpha^2/(2n) - z_\alpha\sqrt{\bar{X}/n + z_\alpha^2/(4n^2)}).$$

c) Randame

$$(\underline{\lambda}; \overline{\lambda}) = ((\sqrt{\bar{X}} - z_\alpha/(2\sqrt{n}))^2; (\sqrt{\bar{X}} + z_\alpha/(2\sqrt{n}))^2).$$

d) Randame a) (1,9706; 2,8294); b) (2,0073; 2,8695); c) (1,9898; 2,8486); remdamiesi **I.2.175** pratimu gauname (1,9898; 2,8698). Intervalas a) mažiau tikslus.

I.2.191. $C(Y_1, ..., Y_n) \approx \{(\mu_y,\ \sigma_y^2) : n[(\bar{Y} - \mu_y)^2/s^2 + (s^2 - \sigma_y^2)^2/(m_4 - s^4)] < \chi_\alpha^2(2)\}$.

I.2.192. *Nurodymas.* a) ir c) pasinaudokite vidurkio ir dispersijos apibrėžimu bei faktu, kad $Y = ln(X)\ N(\mu, \sigma)$. b) Raskite parametrų μ ir σ DT įvertinius $\hat{\mu}$ ir $\hat{\sigma}$. Tada parametrų θ ir γ DT įvertiniai gaunami įrašius $\hat{\mu}$ ir $\hat{\sigma}$ į θ ir γ išraiškas.

I.2.193. a) $\hat{\theta} = 34,988$, $\hat{\gamma} = 81230,195$. b) Naudodami aproksimaciją $\sqrt{n}(\hat{\theta} - \theta)/(\theta\sqrt{m_2}) \xrightarrow{d} Z \sim N(0,\ 1)$, gauname pasikliovimo intervalą $(\underline{\theta}; \overline{\theta}) = (\hat{\theta}/(1 + z_\alpha\sqrt{m_2/n})$; $\hat{\theta}/(1 - z_\alpha\sqrt{m_2/n})$, $\alpha = (1-Q)/2$; pagal turimus duomenis randame realizaciją $(\underline{\theta}; \overline{\theta}) = (24,953;\ 58,524)$.

I.2.194. Kai $N = n$ fiksuotas, tai $(Y|N = n) \sim \mathcal{P}(n\theta)$; generuojančioji funkcija $\psi(s|n) = \mathbf{E}_\theta(s^Y|N = n) = e^{n\theta(s-1)}$. Besąlyginė a. d. Y generuojančioji funkcija

$$\psi(s) = \mathbf{E}_{\theta,\lambda}s^Y = \sum_{k=0}^{\infty} e^{k\theta(s-1)}\frac{\lambda^k}{k!}e^{-\lambda} = \exp\{\lambda(e^{\theta(s-1)} - 1)\}.$$

Randame momentus

$$\mathbf{E}_{\theta,\lambda}Y = \psi'(1) = \lambda\theta, \quad \boldsymbol{V}_{\theta,\lambda}Y = \psi''(1) + \psi'(1) - [\psi'(1)]^2 = \lambda\theta(\theta+1).$$

Prilyginę empiriniam vidurkiui $\bar{Y}$ ir empirinei dispersijai s^2, gauname momentų metodo įvertinius

$$\hat{\theta} = (s^2 - \bar{Y})/\bar{Y}, \hat{\lambda} = \bar{Y}^2/(s^2 - \bar{Y}).$$

I.2.195. a) $\hat{\theta} = 6,2$, $\hat{\lambda} = 8,0645$.

b) Pagal pirmąją aproksimaciją gauname lygmens $Q = 1 - 2\alpha$ aproksimacinį pasikliovimo intervalą intervalą

$$(\underline{\mu}; \overline{\mu}) = (\hat{\mu} - z_\alpha s/\sqrt{n}; \hat{\mu} + z_\alpha s/\sqrt{n}),$$

o pagal antrąją – intervalą $(\underline{\mu}; \overline{\mu})$:

$$\underline{\mu} = (\hat{\mu} + (1 + \hat{\theta})z_\alpha^2)/(2n) - \sqrt{\hat{\mu}(1 + \hat{\theta})z_\alpha^2/n + (1 + \hat{\theta})^2 z_\alpha^4/(4n)},$$

$$\overline{\mu} = (\hat{\mu} + (1 + \bar{\theta})z_\alpha^2)/(2n) + \sqrt{\hat{\mu}(1 + \hat{\theta})z_\alpha^2/n + (1 + \hat{\theta})^2 z_\alpha^4/(4n)}.$$

c) Pagal turimus duomenis gauname intervalus $(41,685;\ 58,315)$, $(42,347;\ 59,036)$.

I.2.196. Parametrų μ ir σ^2 DT įvertiniai yra empirinis vidurkis $\bar{X}$ ir empirinė dispersija m_2. Asimptotinis įvertinio $(\bar{X},\ m_2)$ skirstinys yra dvimatis normalusis su nekoreliuotomis koordinatėmis

$$\sqrt{n}(\bar{X} - \mu) \xrightarrow{d} Y \sim N(0,\ \sigma^2),\ \sqrt{n}(m_2 - \sigma^2) \xrightarrow{d} Z \sim N(0,\ 2\sigma^4).$$

Nagrinėjamų charakteristikų DT įvertiniai yra

$$\hat{x}_P = \bar{X} + z_P\sqrt{m_2},\ \hat{F}(x; \mu, \sigma) = \Phi\left(\frac{x - \bar{X}}{\sqrt{m_2}}\right).$$

Pakeitę dispersijas jų įvertiniais ir pažymėję, kad

$$\frac{\partial x_P}{\partial \mu} = 1,\ \frac{\partial x_P}{\partial \sigma^2} = \frac{z_P}{2\sigma},$$

naudodami delta metodą gauname įvertinio $\hat{x}_P$ asimptotinės dispersijos įvertinį

$$\hat{\sigma}^2_{\hat{x}_P} = \frac{m_2}{n} + \left(\frac{z_P}{2\sqrt{m_2}}\right)^2 \frac{2m_2^2}{n} = \frac{m_2}{n}\left(1 + \frac{z_P^2}{2}\right).$$

Parametro x_P aproksimacinis pasikliovimo intervalas, kai pasikliovimo lygmuo $Q = 1 - 2\alpha$, yra

$$(\hat{x}_P - z_\alpha \hat{\sigma}_{\hat{x}_P},\ \ \hat{x}_P + z_\alpha \hat{\sigma}_{\hat{x}_P}).$$

Analogiškai gauname

$$\frac{\partial F}{\partial \mu} = -\frac{1}{\sigma}\varphi\left(\frac{x - \mu}{\sigma}\right),\ \frac{\partial F}{\partial \sigma^2} = -\frac{x - \mu}{2\sigma^3}\varphi\left(\frac{x - \mu}{\sigma}\right),$$

$$\hat{\sigma}^2_{\hat{F}} = \frac{1}{n}\left(1 + \frac{(x - \bar{X})^2}{2m_2}\right)\varphi^2\left(\frac{x - \bar{X}}{\sqrt{m_2}}\right),$$

$$\underline{F} = \hat{F}(x) - z_\alpha \hat{\sigma}_{\hat{F}}, \ \overline{F} = \hat{F}(x) + z_\alpha \hat{\sigma}_{\hat{F}}.$$

I.2.197. Parametrų $x(0,9)$ ir $F(0)$ taškiniai įverčiai yra

$$\hat{x}(0,9) = \bar{X} + sz(0,9) = 3,8432; \quad \hat{F}(0) = \Phi(-1,28/2) = 0,2611.$$

Asimptotinių dispersijų įverčiai

$$\hat{\sigma}^2_{\hat{x}(0,9)} = 4(1 + z^2(0,9)/2)/n = 0,0728; \quad \hat{\sigma}^2_{\hat{F}(0)} = (1 + 1,28^2/8)\varphi^2(0,64)/n = 0,00127.$$

Gauname asimptotinius pasikliovimo intervalus

$$\underline{x}(0,9) = \hat{x}(0,9) - z_{0,025}\hat{\sigma}_{\hat{x}(0,9)} = 3,3144, \quad \overline{x}(0,9) = \hat{x}(0,9) + z_{0,025}\hat{\sigma}_{\hat{x}(0,9)} = 4,3720;$$

$$\underline{F}(0) = \hat{F}(0) - z_{0,025}\hat{\sigma}_{\hat{F}(0)} = 0,1913, \quad \overline{F}(0) = \hat{F}(0) + z_{0,025}\hat{\sigma}_{\hat{F}(0)} = 0,3309.$$

Statistikų g_1 ir g_2 pirmųjų momentų išraiškos pateiktos **I.1.95** pratime. Naudodami šias išraiškas gauname asimptotinius pasikliovimo intervalus

$$\underline{\gamma}_1 = g_1 - z_{0,025}\sqrt{\hat{\boldsymbol{V}} g_1} = -0,7184, \ \overline{\gamma}_1 = g_1 + z_{0,025}\sqrt{\hat{\boldsymbol{V}} g_1} = 0,2184;$$

$$\underline{\gamma}_2 = g_2 - 6/(n+1) - z_{0,025}\sqrt{\hat{\boldsymbol{V}} g_2} = -0,8507, \ \overline{\gamma}_2 = g_2 - 6/(n+1) + z_{0,025}\sqrt{\hat{\boldsymbol{V}} g_2} = 0,9319.$$

I.2.198. Pažymėkime $u_p = -\ln(1-p)$. Išvestinės

$$\frac{\partial t(p)}{\partial \theta} = u_p^{1/\mu}, \quad \frac{\partial t(p)}{\partial \nu} = -\frac{\theta}{\nu^2} u_p^{1/\nu} \ln u_p,$$

taigi

$$\hat{\sigma}^2_{\hat{t}(p)} = u_p^{2/\hat{\nu}} \left(\hat{I}^{11} - 2\hat{I}^{12} \frac{\hat{\theta}}{\hat{\nu}^2} \ln u_p + \hat{I}^{22} \frac{\hat{\theta}^2}{\hat{\nu}^4} \ln^2 u_p \right);$$

čia $\hat{I}^{ij}$ yra matricos $\hat{\boldsymbol{I}}_n^{-1}$ elementai. Informacinės matricos $\boldsymbol{I}$ pagrįstojo įvertinio $\hat{\boldsymbol{I}}_n = [\hat{I}_{ij}]_{2\times 2}$ elementai

$$\hat{I}_{11} = n\frac{\hat{\nu}^2}{\hat{\theta}^2}, \quad \hat{I}_{22} = n\frac{1 + \Gamma''(2)}{\hat{\nu}^2}, \quad \hat{I}_{12} = \hat{I}_{21} = n\frac{\Gamma'(2)}{\hat{\sigma}}.$$

Intervalas

$$\left(\hat{t}(p) - \hat{\sigma}_{\hat{t}(p)}\, z_{1-\alpha/2}, \quad \hat{t}(p) + \hat{\sigma}_{\hat{t}(p)}\, z_{1-\alpha/2}\right)$$

yra parametro $t(p)$ aproksimacinis pasikliovimo intervalas, kai pasikliovimo lygmuo $Q = 1 - 2\alpha$.

I.2.199. Parametrų θ ir ν įvertinių vidutinius kvadratinius nuokrypius įvertinkime taip:

$$\hat{\sigma}_{\hat{\theta}} = 1/\sqrt{\hat{\boldsymbol{I}}^{11}} = \hat{\theta}/(\hat{\nu}\sqrt{n}) = 0,0092, \quad \hat{\sigma}_{\hat{\nu}} = 1/\sqrt{\hat{\boldsymbol{I}}^{22}} = \hat{\nu}/\sqrt{n(1 + \Gamma''(2))} = 0,1592.$$

Gauname asimptotinius pasikliovimo intervalus

$$\underline{\theta} = \hat{\theta} - z_{0,025}\hat{\sigma}_{\hat{\theta}} = 0,0568, \quad \overline{\theta} = \hat{\theta} + z_{0,025}\hat{\sigma}_{\hat{\theta}} = 0,3392;$$

$$\underline{\nu} = \hat{\nu} - z_{0,025}\hat{\sigma}_{\hat{\nu}} = 1,8380, \quad \overline{\nu} = \hat{\nu} + z_{0,025}\hat{\sigma}_{\hat{\nu}} = 2,4620.$$

Parametro $t(0,9)$ taškinio įvertinio realizacija yra $\hat{t}(0,9) = \hat{\theta}(-\ln 0,1)^{1/\hat{\nu}} = 0,2918$. Vertinant dispersiją reikia rasti Fišerio informacinės matricos atvirkštinę. Gauname $\hat{\boldsymbol{I}}^{11} = 0,0094/n$, $\hat{\boldsymbol{I}}^{22} = 2,8104/n$, $\hat{\boldsymbol{I}}^{12} = -0,0509/n$. Tada $\hat{\sigma}_{\hat{t}(0,9)} = 0,0172$ ir pasikliovimo intervalas

$$\underline{t}(0,9) = \hat{t}(0,9) - z_{0,025}\hat{\sigma}_{\hat{t}(0,9)} = 0,2580, \quad \overline{t}(0,9) = \hat{t}(0,9) + z_{0,025}\hat{\sigma}_{\hat{t}(0,9)} = 0,3427.$$

I.2.6 skyrelis

I.2.200. $\hat{\mu} = \bar{X} = 424,933$, $\hat{\sigma} = s = 8,598$; $(\underline{\mu}, \overline{\mu}) = (420,172;\ 429,695)$, $(\underline{\sigma},;\ \overline{\sigma}) = (6,295;\ 13,560)$.

I.2.201. $\hat{\lambda} = \bar{X} = 0,9288$; $(\underline{\lambda};\ \overline{\lambda}) = (0,850;\ 1,009)$.

I.2.202. a) Momentų metodu gauname įverčius $\hat{\lambda} = 1,073$, $\hat{\eta} = 9,928$;
DT įverčius galime rasti maksimizuodami tikėtinumo funkcijos $L(\lambda, \eta)$ logaritmą

$$\max_{\lambda,\eta} = \max_{\lambda,\eta}[-\lambda T_1 + (\eta - 1)T_2 + n\eta \ln\lambda - n\ln(\Gamma(\eta))] = \ell(\tilde{\lambda}, \tilde{\eta});$$

čia $T_1 = X_1 + ... + X_n, T_2 = \sum_i \ln X - i$.
b) DT įvertinys $\hat{\lambda} = \eta/\bar{X}$ paslinktasis; NMD įvertis $\hat{\lambda} = (n\eta - 1)/(n\bar{X}) = 1,076$; $(\underline{\lambda};\ \overline{\lambda}) = (0,936;\ 1,236)$.

I.2.203. $\hat{p} = 0,277$; $(\underline{p};\ \overline{p}) = (0,2495;\ 0,3059)$.

I.2.204. $(\underline{p};\ \overline{p}) = (0,000;\ 0,206)$, $(0,000;\ 0,259)$, $(0,000;\ 0,411)$.

I.2.205. Parametro λ įvertį galime rasti maksimizuodami tikėtinumo funkcijos $L(\lambda)$ logaritmą

$$\max_{\lambda} \ell(\lambda) = \max_{\lambda} \sum_{i=1}^{11} S_i \ln(\pi_i(\lambda)) = \ell(\hat{\lambda}),$$

čia $\pi_i(\lambda) = e^{-\lambda a_{i-1}} - e^{-\lambda a_i}$, $i = 1, ..., 11$, $a_0 = 0, a_{12} = \infty$.
Asimptotinį pasikliovimo intervalą gauname naudodami aproksimaciją

$$\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{d} Y \sim N(0, 1/i(\hat{\lambda})),$$

$$i(\lambda) = -\mathbf{E}(\ddot{\ell}_{\lambda^2}(\lambda))/n = \sum_i [(\dot{\pi}_i(\lambda))^2/\pi_i(\lambda)].$$

Asimptotinis pasikliovimo intervalas

$$(\underline{\lambda}; \overline{\lambda}) = (\hat{\lambda} - z_{0,025} i(\hat{\lambda})/n; \hat{\lambda} + z_{0,025} i(\hat{\lambda})/n).$$

I.2.206. a) $\hat{\sigma}^2 = s^2 = [s_x^2(n_1 - 1) + s_y^2(n_2 - 1) + s_z^2(n_3 - 1)]/\nu$, $\nu = n_1 + n_2 + n_3 - 3$; $s^2\nu/\sigma^2 \sim \chi^2(\nu)$. Remdamiesi turimais duomenimis, gauname realizacijas $s^2 = 3,288$,

$(\underline{\sigma};\overline{\sigma}) = (\sqrt{s^2\nu/\chi^2_{0,995}(\nu)};\sqrt{s^2\nu/\chi^2_{0,005}(\nu)}) = (1,528;2,217)$. b) $\hat\theta = \bar X + \bar Y - \bar Z \sim N(\theta,\sigma^2(1/n_1+1/n_2+1/n_3))$, $(\hat\theta-\theta)/\sigma_{\hat\theta} \sim S(\nu)$, $\sigma_{\hat\theta} = s\sqrt{1/n_1+1/n_2+1/n_3}$. Remdamiesi turimais duomenimis, gauname realizacijas $\hat\theta = 0,070$, $(\underline{\theta};\overline{\theta}) = (-1,087;1,227)$.

I.2.207. a) $\hat\theta = 0,075$, $s_x^2 = 3,9686$, $s_y^2 = 3,8783$; $(\underline{\theta};\overline{\theta}) = (-1,355;1,505)$. Kadangi nulis yra per vidurį šio intervalo, tai darytina išvada, kad a. d. X ir Y vidurkiai nesiskiria.

b) $s_z^2 = \sum_i(X_i - Y_i - (\bar X - \bar Y))^2/(n-1) = 0,02867$; $(\underline{\theta};\overline{\theta}) = (-0,015;0,165)$. Intervalas trumpesnis, o nulis yra šio intervalo krašte. Todėl išvada dėl a. d. X ir Y vidurkių lygybės kelia abejonių.

c) Pagal prasmę a. d. X ir Y turėtų būti priklausomi (jei bulvė turi daugiau krakmolo, tai abu metodai turėtų duoti didesnes reikšmes ir atvirkščiai). Todėl lentelės duomenis reikėtų traktuoti kaip dvimačio a. v. $(X,Y)^T$ realizacijas.

d) Taškinis koreliacijos koeficiento įvertis $\hat\rho = r = 0,9964$. Remdamiesi **I.2.171** pratimu gauname pasikliovimo intervalą $(\underline{\rho};\overline{\rho}) = (0,9894;0,9988)$. A. d. X ir Y yra stipriai priklausomi.

I.2.208. $(\underline{\theta};\overline{\theta}) = (0,9966;1,0497)$.

I.2.209. $\hat\theta = 0,0408$; $(\underline{\theta};\overline{\theta}) = (0,0151;0,0665)$; yra pagrindo tvirtinti, kad $p_1 > p_2$.

I.2.210. Taikydami normaliąją aproksimaciją gauname: a) $N \geq 316$; b) $N \geq 1263$.

I.2.211. Randame (žr. [2], 4.7.12 skyrelį) $\bar C = -0,0345$, $\bar S = 0,3222$, $\bar R = 0,3240$; $\hat\mu = 96,1605°$, $\hat\theta = 0,6854$. Aproksimaciniai pasikliovimo intervalai: $(\underline{\mu};\ \overline{\mu}) = (79,6766°;\ 112,6444°)$; $(\underline{\theta};\ \overline{\theta}) = (0,4768;\ 0,8940)$.

I.2.212. a) $\hat C = -0,0954$, $\hat S = -0,0316$, $\hat R = 0,1005$; $\hat\mu = 198,34°$, $\hat\theta = 0,2021$. b) $(\underline{\mu};\ \overline{\mu}) = (163,31°;\ 233,37°)$; $(\underline{\theta};\ \overline{\theta}) = (0,0779;\ 0,3263)$.

I.3. Parametrinių hipotezių tikrinimas

I.3.1. Statistiniai kriterijai

I.3.1. Psichologas nori nustatyti, ar pelės atmintis pasikeičia pašalinus tam tikrą smegenų dalį. Iš pradžių jis įpratino 6 peles neklystamai rasti labirinte vienintelį kelią prie maisto. Paskui pašalino smegenų dalį. Galima tarti, kad neturinti atminties pelė teisingą kelią ras su tikimybe 0,2. Psichologas teigia, kad pelė atminties neturi, jei teisingą kelią randa ne daugiau kaip 2 pelės iš 6 pelių. Suformuluokite uždavinį hipotezių tikrinimo terminais. Kokia psichologo taikomo kriterijaus kritinė sritis? Koks šio kriterijaus reikšmingumo lygmuo? Apskaičiuokite kriterijaus galią, kai $p = 0,4,\ 0,6,\ 0,8,\ 1,0$.

I.3.2. Gaminant tam tikrą chemikalą, pageidautina, kad vartojamo vandens tūrio vienete būtų vidutiniškai ne daugiau kaip m_0 bakterijų. Nuo per didelės jų koncentracijos apsisaugoma taip: imama n vienodo tūrio V vandens pavyzdžių, kiekvieno iš tų pavyzdžių vanduo supilamas į kolbą su maitinamąja terpe. Jeigu pavyzdžio vanduo užterštas (t. y. jame yra nors viena bakterija), tai bakterijų kolonija plėsis ir buvęs skaidrus tirpalas susidrums.

Vanduo laikomas pakankamai švariu ir vartojamas gamyboje, jeigu kolbų su susidrumstusiu vandeniu skaičius X ne didesnis už skaičių t. Priešingu atveju vanduo valomas.

Tegu:

a) $m_0 = 1,\ V = 1,\ n = 10,\ t = 7;$

b) $m_0 = 1,\ V = 2,\ n = 8,\ t = 7.$

Suformuluokite uždavinį hipotezės tikrinimo terminais ir apskaičiuokite pirmosios rūšies klaidos tikimybę.

Nurodymas. Vandens tūrį, iš kurio imami pavyzdžiai, laikykite daug kartų didesnį už pavyzdžio tūrį. Tiksliau kalbant, bakterijų skaičiaus pasiskirstymą tūryje V aproksimuokite Puasono skirstiniu.

I.3.3. Sėklų pirkėjas ir pardavėjas susitarė, kad sėklų partijos kainą nustatys po bandymo. Iš partijos buvo paimta 250 sėklų ir pasėta. Tegu X yra iš 250 sėklų išaugusių augalų, turinčių tam tikrų savybių, skaičius. Jeigu $X \leq 25$, tai nusprendžiama, kad sėklų partijos kaina bus pripažinta normalia; priešingu atveju sėklų partijos kaina sumažinama. Kiek procentų sėklų turi turėti minėtas savybes, kad tikimybė, jog partijos kaina bus normali, būtų lygi (apytiksliai) 0,95?

I.3.4. Parduodant elektros lempučių partijas ($N = 10\,000$) atliekama atrankinė kontrolė: atsitiktinai negrąžinant atrenkama n lempučių ir nustatomas defektinių lempučių skaičius X; jeigu $X \leq d$, partija pripažįstama gera, priešingu atveju – defektine. Suformuluokite uždavinį hipotezių tikrinimo terminais. Nurodykite kriterijaus kritinę sritį.

I.3.5.(**I.3.4** pratimo tęsinys). Tarkime, pirkėjui nepriimtina, jei defektuotų lempučių procentas ne mažesnis už 10; jis reikalauja, kad partija būtų atmesta su tikimybe 0,95, jeigu joje defektuotų lempučių yra 10 procentų. Pardavėjas įsitikinęs, kad defektuotų lempučių procentas ne didesnis už 5. Jis nori, kad kiekvieną kartą, kai defektuotų lempučių yra 5 procentai, partija būtų priimama su tikimybe 0,9. Pasiūlykite kriterijų, kuris tenkintų ir pirkėją, ir pardavėją.

I.3.6. Nustatyta, kad gamyklos produkcijoje vidutiniškai yra 5 procentai broko. Per pamainą pagaminama 500 gaminių. Atliekant kontrolę nustatomas per pamainą pagamintų defektinių gaminių skaičius X. Jeigu X įgyja didelę reikšmę, tai daroma išvada, kad gamybos procesas sutriko. Suformuluokite uždavinį hipotezių tikrinimo terminais.

I.3.7. (**I.3.6** pratimo tęsinys). a) Nurodykite ribą, kurią defektinių gaminių skaičius viršija su tikimybe, ne didesne už 0,01, kai defektinio gaminio pagaminimo tikimybė lygi 0,05. Apskaičiuokite tikimybę, kad tą ribą viršys defektinių gaminių skaičius, kai defektinio gaminio pagaminimo tikimybė lygi 0,08; 0,10; 0,12. b) Nurodykite kontrolinę ribą ir minėtas tikimybes, jei tikrinama tik 50 atsitiktinai paimtų gaminių.

I.3.8. Tegu paprastoji imtis $\mathbf{X} = (X_1, ..., X_n)^T$ gauta stebint a. d. $X \sim N(\mu, \sigma^2)$, σ žinomas. a) Tikriname hipotezę $H : \mu = \mu_0$, kai alternatyva yra $\bar{H}_1 : \mu > \mu_0$ arba $\bar{H}_2 : \mu < \mu_0$. Suformuluokite kriterijus remdamiesi pasikliovimo intervalais (žr. [2], 4.5 skyrelį); suformuluokite kriterijus P reikšmių terminais (žr. [2], 4.1.2 skyrelį).

b) Tikriname hipotezę $H : \mu = \mu_0$, kai alternatyva yra $\bar{H}_1 : \mu \neq \mu_0$. Suformuluokite kriterijų remdamiesi pasikliovimo intervalais; suformuluokite kriterijų P reikšmių terminais.

I.3.2. Neimano ir Pirsono lema

I.3.9. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. X, kurio skirstinys priklauso šeimai $\mathcal{P} = \{f(x;\theta), \theta = \theta_0, \theta = \theta_1\}$; čia tankio funkcija $f(x;\theta) = \theta \exp\{-\theta x\}, x > 0$. Raskite galingiausią hipotezės $H : \theta = \theta_0$, kai alternatyva yra $\bar{H} : \theta = \theta_1$, tikrinimo kriterijų (reikšmingumo lygmuo lygus α). Apskaičiuokite rastojo kriterijaus galią.

I.3.10. Tegu X yra atsitiktinis dydis, kurio skirstinys priklauso Koši skirstinių šeimai $\mathcal{P} = \{f(x;\theta), \theta = 0, \theta = 1\}$; čia

$$f(x;\ \theta) = \frac{1}{\pi}\frac{1}{1 + (x - \theta)^2},\ -\infty < x < \infty.$$

Raskite galingiausią kriterijų hipotezei $H : \theta = 0$, kai alternatyva $\bar{H} : \theta = 1$, tikrinti. Imties didumas $n = 1$.

I.3.11. Raskite galingiausią kriterijų hipotezei H, kad a. d. X yra pasiskirstęs pagal standartinį normalųjį skirstinį $N(0,\ 1)$, esant alternatyvai $\bar{H}$, kad a. d. X skirstinys yra: a) Laplaso, kurio tankis

$$\frac{1}{2}\mathrm{e}^{-|\mathrm{x}|},\ -\infty < \mathrm{x} < \infty;$$

b) tolygus intervale $(-\delta,\ \delta)$, tikrinti. Imties didumas lygus 1.

I.3.12. Paprastosios imties realizacija yra $-0,260$; $-0,114$; $-0,325$; $0,196$; $-0,174$. Sudarykite galingiausiąjį kriterijų hipotezei H: stebimo a. d. skirstinys yra normalusis $N(0,\ 0,025)$, esant alternatyvai $\bar{H}$: stebimo a. d. skirstinys yra tolygusis $U(-0,5,\ 0,5)$, tikrinti. Ar atmetama hipotezė H pagal turimą realizaciją, jei reikšmingumo lygmuo $\alpha = 0,1$?

I.3.13. Pagal paprastąją didumo n imtį raskite galingiausiąjį kriterijų hipotezei H: stebimo a. d. skirstinys $N(0,\ 1)$, esant alternatyvai $\bar{H}$: stebimo a. d. skirstinys yra $N(1,\ 1)$, tikrinti. Koks turėtų būti mažiausias imties didumas n, kad abiejų rūšių klaidų tikimybės neviršytų 0,01?

I.3.14. Yra $n = 1$ didumo imtis atsitiktinio dydžio $X \sim \mathcal{P}(\lambda)$, turinčio Puasono skirstinį. Tikrinama hipotezė $H : \lambda = \lambda_0$, kai alternatyva $\bar{H} : \lambda = \lambda_1$. Raskite galingiausiojo kriterijaus galią, kai $\lambda_0 = 0,1, \lambda_1 = 0,2$; $\lambda_0 = 1, \lambda_1 = 2$; $\lambda_0 = 10, \lambda_1 = 20$; $\lambda_0 = 0,1, \lambda_1 = 0,4$, o kriterijaus reikšmingumo lygmuo $\alpha = 0,1$.

I.3.15. Imties $\mathbf{X}$ skirstinys priklauso šeimai $\mathcal{P} = \{f_\theta(\boldsymbol{x}),\ \theta = 1,\ 2\}$. Tikrinama paprastoji hipotezė $H : \theta = 1$, esant paprastajai alternatyvai $\bar{H} : \theta = 2$. Tegu η įgyja reikšmę 1, jeigu priimta hipotezė H, reikšmę 2, jeigu priimta alternatyva $\bar{H}$, ir reikšmę 0, jeigu atsisakoma sprendimo, kuri iš hipotezių teisinga. Pažymėkime

$$\varphi_i(\boldsymbol{x}) = \mathbf{P}\{\eta = i|\mathbf{X} = \boldsymbol{x}\},\ i = 1,\ 2,\ 0;$$

$$0 \le \varphi_i(\boldsymbol{x}) \le 1;\ \varphi_1(\boldsymbol{x}) + \varphi_2(\boldsymbol{x}) + \varphi_0(\boldsymbol{x}) \equiv 1.$$

Pirmosios ir antrosios rūšies klaidų tikimybės yra

$$\alpha_{12} = \mathbf{P}\{\eta = 1|\theta = 2\} = \mathbf{E}(\varphi_1(\mathbf{X})|\theta = 2),\ \alpha_{21} = \mathbf{P}\{\eta = 2|\theta = 1\} = \mathbf{E}(\varphi_2(\mathbf{X})|\theta = 1).$$

Raskite kriterijų (t. y. raskite funkcijas $\varphi_i(\mathbf{X})$, $i = 1,\ 2,\ 0$), tenkinantį sąlygas $\alpha_{12} \le \alpha$, $\alpha_{21} \le \beta$ ir minimizuojantį tikimybes

$$\alpha_{02} = \mathbf{P}\{\eta = 0|\theta = 2\} = \mathbf{E}(\varphi_0(\mathbf{X})|\theta = 2),\ \alpha_{01} = \mathbf{P}\{\eta = 0|\theta = 1\} = \mathbf{E}(\varphi_0(\mathbf{X})|\theta = 1).$$

I.3.16. (**I.3.15** pratimo tęsinys). Tegu apriorinės hipotezių H ir $\bar{H}$ tikimybės yra ω_1 ir ω_2 ($\omega_1 + \omega_2 = 1$) ir β_{21} ir β_{12} – aposteriorinės tikimybės:

$$\beta_{21} = \mathbf{P}\{\theta = 2|\eta = 1\} = \frac{\alpha_{12}\omega_2}{\alpha_{12}\omega_2 + (1 - \alpha_{01} - \alpha_{21})\omega_1},$$

$$\beta_{12} = \mathbf{P}\{\theta = 1|\eta = 2\} = \frac{\alpha_{21}\omega_1}{\alpha_{21}\omega_1 + (1 - \alpha_{02} - \alpha_{12})\omega_2}.$$

Raskite kriterijų, tenkinantį sąlygas $\beta_{21} \le b_2$, $\beta_{12} \le b_1$ ir minimizuojantį tikimybes α_{01}, α_{02}.

I.3.17. Tegu H_0 ir H_1 yra paprastosios hipotezės ir reikšmingumo lygmuo $\alpha \in (0,\, 1)$. Be to, φ_* yra tolygiai galingiausias α lygmens kriterijus hipotezei H_0, kai alternatyva yra H_1, tikrinti, o kriterijaus galia $\beta < 1$, kai H_1 teisinga. Įrodykite, kad $1 - \varphi_*$ yra tolygiai galingiausias $1 - \beta$ lygmens kriterijus hipotezei H_1, kai alternatyva yra H_0, tikrinti.

I.3.18. Tegu X yra vienetinė imtis iš skirstinio, kurio tankio funkcija lygi $f_\theta(x)$. Raskite galingiausią lygmens $\alpha \in (0,\, 1/2)$ kriterijų hipotezei $H_0 : \theta = \theta_0$, kai alternatyva yra $H_1 : \theta = \theta_1$, tikrinti tokiais atvejais:

a) $f_\theta(x) = 2\theta^{-2}(\theta - x)$, $0 < x < \theta$, $\theta_0 < \theta_1$;

b) $f_\theta(x) = 2[\theta x + (1 - \theta)(1 - x)]$, $0 < x < 1$, $0 \leqslant \theta_1 < \theta_0 \leqslant 1$;

c) $f_{\theta_0}(x) = 4xI_{(0,\,1/2)}(x) + 4(1 - x)I_{(1/2,\,1)}(x)$ ir $f_{\theta_1}(x) = I_{(0,\,1)}(x)$.

I.3.19. Tegu $X_1, \ldots, X_n$ nepriklausomi vienodai pasiskirstę a. d., kurių tankio funkcija $f_\theta(x)$. Raskite galingiausią lygmens α kriterijų hipotezei $H_0 : \theta = \theta_0$, kai alternatyva yra $H_1 : \theta = \theta_1$, tikrinti tokiais atvejais:

a) $f_\theta(x) = \mathrm{e}^{-(\mathrm{x}-\theta)}$, $\theta < \mathrm{x} < \infty$, $\theta_0 < \theta_1$;

b) $f_\theta(x) = \theta x^{-2}$, $\theta < x < \infty$, $\theta_0 > \theta_1$.

I.3.20. Imtis $\mathbf{X} = (X_1, ..., X_n)^T$ gauta stebint Bernulio a. d. $B(1, p)$. Sudarykite kriterijų hipotezei $H : p = 0$, kai alternatyva yra $\bar{H} : p = 0,01$, tikrinti. Raskite mažiausiąjį imties didumą n, kad pirmosios ir antrosios rūšies klaidų tikimybės neviršytų 0,01.

I.3.3. Skirstiniai, priklausantys nuo vieno parametro

I.3.21. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. $X \sim N(\mu, \sigma^2)$, $-\infty < \mu < \infty$, o $0 < \sigma$ – žinomas. a) Raskite reikšmingumo lygmens α hipotezės $H : \mu = \mu_0$, kai alternatyvos yra $\bar{H}_1 : \mu > \mu_0$ arba $\bar{H}_2 : \mu < \mu_0$, tikrinimo TG kriterijus ir jų galios funkcijas. b) Raskite TGN kriterijų, kai alternatyva dvipusė $\bar{H}_3 : \mu \neq \mu_0$, ir jo galios funkciją.

I.3.22. (**I.3.21** pratimo tęsinys). Suformuluokite **I.3.21** pratime gautus kriterijus pasikliovimo intervalų terminais.

I.3.23. (**I.3.21** pratimo tęsinys). Suformuluokite **I.3.21** pratime gautus kriterijus P reikšmių terminais.

I.3.24. (**I.3.21** pratimo tęsinys). Remiantis tuo, kad kriterijų galios funkcijos priklauso ne tik nuo skirtumo $\mu - \mu_0$, bet ir nuo imties didumo n, išspręskite tokį eksperimento planavimo uždavinį. Tikrinama hipotezė $H : \mu = \mu_0$, kai alternatyva yra $\bar{H}_1 : \mu > \mu_0$. a) Raskite tokį imties didumą n, kad hipotezė H būtų priimama su tikimybe, ne didesne už α', kai $\mu \geq \mu' > \mu_0$. b) Raskite imties didumą, kai $\alpha = 0,05; \alpha' = 0,01; \sigma^2 = 4; \mu' = \mu_0 + 0,5$.

I.3.25. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. $X \sim N(\mu, \sigma^2)$, $0 < \sigma$, o vidurkis μ žinomas. a) Raskite reikšmingumo lygmens α hipotezės $H : \sigma^2 = \sigma_0^2$, kai alternatyvos yra $\bar{H}_1 : \sigma^2 > \sigma_0^2$ arba $\bar{H}_2 : \sigma^2 < \sigma_0^2$, tikrinimo TG kriterijus ir jų galios funkcijas. b) Raskite TGN kriterijų, kai alternatyva dvipusė $\bar{H}_3 : \sigma^2 \neq \sigma_0^2$. c) Dvipusės alternatyvos atveju suformuluokite paslinktą, tačiau paprasčiau randamą simetrišką kriterijų.

I.3.26. (**I.3.25** pratimo tęsinys). Dvipusės alternatyvos atveju palyginkite TGN ir simetriško kriterijaus galią, kai $n = 10; 20; 50$ ir $\lambda = \sigma^2/\sigma_0^2 = 0,5; 0,75; 1; 1,25; 1,5$, o kriterijaus reikšmingumo lygmuo $\alpha = 0,05$.

I.3.27. (**I.3.25** pratimo tęsinys). Suformuluokite **I.3.25** pratime gautus kriterijus a) pasikliovimo intervalų terminais; b) P reikšmių terminais.

I.3.28. (**I.3.25** pratimo tęsinys). Reikšmingumo lygmens $\alpha = 0,05$ kriterijumi tikrinama hipotezė $H : \sigma^2 = \sigma_0^2$, kai alternatyva yra $\bar{H}_1 : \sigma^2 > \sigma_0^2$. Raskite tokį imties didumą n, kad hipotezė H būtų atmetama su tikimybe, ne mažesne už $0,9$, kai tikroji parametro σ^2 tenkina nelygybę $\sigma^2 \geq \mathbf{1},5\sigma_0^2$.

I.3.29. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. X, turintis Relėjaus skirstinį (žr. 1 priedo 1P1 lentelę) su parametru $\sigma > 0$. Raskite TG arba TGN kriterijus hipotezei $H : \sigma^2 = \sigma_0^2$ tikrinti vienpusių ir dvipusės alternatyvų atvejais.

I.3.30. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. X, turintį Maksvelo skirstinį (žr. 1 priedo 1P1 lentelę) su parametru $\sigma > 0$. Raskite TG arba TGN kriterijus hipotezei $H : \sigma^2 = \sigma_0^2$ tikrinti vienpusių ir dvipusės alternatyvų atvejais.

I.3.31. Tarkime, imties $X_1, ..., X_n$ elementai yra n. a. d., turintys gama skirstinius $X_i \sim G(\lambda, \eta_i)$, $0 < \lambda$, o $\eta_1, ..., \eta_n$ žinomi, $\eta_1 + ... + \eta_n = \eta$. a) Raskite reikšmingumo lygmens α hipotezės $H : \lambda = \lambda_0$, kai alternatyvos yra $\bar{H}_1 : \lambda > \lambda_0$ arba $\bar{H}_2 : \lambda < \lambda_0$, TG kriterijus. b) Raskite TGN kriterijų, kai alternatyva dvipusė $\bar{H}_3 : \lambda \neq \lambda_0$. c) Dvipusės alternatyvos atveju suformuluokite paslinktą, tačiau paprasčiau randamą simetrišką kriterijų.

I.3.32. Tarkime, paprastoji imtis $X_1, ..., X_n$ gauta stebint a. d. $X \sim \mathcal{P}(\lambda)$, $0 < \lambda$. a) Raskite reikšmingumo lygmens α TG kriterijus hipotezei $H : \lambda = \lambda_0$, kai alternatyvos yra $\bar{H}_1 : \lambda > \lambda_0$ arba $\bar{H}_2 : \lambda < \lambda_0$. b) Remdamiesi [2], 4.1.3 pastaba suformuluokite nerandomizuotus kriterijus.

I.3.33. (**I.3.32** pratimo tęsinys). a) Raskite TGN kriterijų dvipusės alternatyvos $\bar{H}_3 : \lambda \neq \lambda_0$ atveju. b) Suformuluokite simetrišką kriterijų atsisakius randomizacijos.

I.3.34. (**I.3.32** pratimo tęsinys). Tarkime, pagal didumo $n = 40$ paprastąją imtį gauta parametro λ NMD įvertinio realizacija $\hat{\lambda} = 2,0$. Reikšmingumo lygmens $\alpha = 0,05$ kriterijumi patikrinkite hipotezę $H : \lambda \geq 2,5$, kai alternatyva yra $\bar{H}_2 : \lambda < 2,5$.

I.3.35. Tarkime, paprastoji imtis $X_1, ..., X_n$ gauta stebint Bernulio a. d. $X \sim B(1, p)$, $0 < p < 1$. a) Raskite reikšmingumo lygmens α TG kriterijus hipotezei $H : p = p_0$, kai alternatyvos yra $\bar{H}_1 : p > p_0$ arba $\bar{H}_2 : p < p_0$, tikrinti ir TGN kriterijų, kai alternatyva dvipusė. b) Suformuluokite nerandomizuotus ir simetriškus kriterijus.

I.3.36. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. X, kurio skirstinys priklauso Bernulio skirstinių šeimai $\mathcal{P} = \{B(1,\ p),\ 0 < p < 1\}$. Reikia patikrinti hipotezę

$H : p \le p_0$, kai alternatyva yra $\bar{H} : p > p_0$. Raskite TG kriterijaus galios $\beta(p)$ reikšmes taškuose $p = 0,3,\ 0,4,\ 0,5,\ 0,6$, kai $n = 6$, $p_0 = 0,25$ ir kriterijaus reikšmingumo lygmuo $\alpha = 0,05;\ 0,1;\ 0,2$. Naudodami nerandomizuotą kriterijų raskite minimalų imties didumą n, kad kriterijaus galia $\beta(p)$ tenkintų nelygybę $\beta(p) \ge 0,9$, kai $p \ge p_1$, ir: a) $p_0 = 0,2,\ p_1 = 0,4$; b) $p_0 = 0,02,\ p_1 = 0,04$, o kriterijaus reikšmingumo lygmuo $\alpha = 0,1$.

I.3.37. Eksperimentais nustatyta, kad gamykla pagamina vidutiniškai 5 procentus defektinės produkcijos. Iš 50 atsitiktinai paimtų gaminių 6 gaminiai buvo defektuoti. Patikrinkite prielaidą, kad defektuotų gaminių procentas padidėjo. Kriterijaus reikšmingumo lygmuo $\alpha = 0,05$.

I.3.38. Tegu $X_1, \dots, X_n$ yra nepriklausomi vienodai pasiskirstę Bernulio a. d. $B(1,p)$. Raskite α lygmens TGN kriterijų hipotezei $H_0 : p = p_0$, kai alternatyva yra $H_1 : p \ne p_0$, tikrinti, kai

a) $n = 10$, $\alpha = 0,1$ ir $p_0 = 0,2$;

b) $n = 10$, $\alpha = 0,05$ ir $p_0 = 0,4$.

I.3.39. Tegu X yra a. d., turintis geometrinį skirstinį. Raskite α lygmens TGN kriterijų hipotezei $H_0 : p = p_0$, kai alternatyva yra $H_1 : p \ne p_0$, tikrinti.

I.3.40. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. X, kurio skirstinys priklauso tolygiųjų skirstinių šeimai $\mathcal{P} = \{U(0,\ \theta),\ 0 < \theta < \infty\}$. Raskite TG kriterijus hipotezei $H : \theta = \theta_0$, esant vienpusėms ir dvipusei alternatyvoms, tikrinti.

I.3.41. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. X, kurio skirstinys priklauso eksponentinių skirstinių šeimai $\mathcal{P} = \{f(x,\ \mu,\ \sigma), -\infty < \mu < \infty,\ 0 < \sigma < \infty\}$; čia tankis

$$f(x;\ \mu,\ \sigma) = \sigma e^{-\sigma(x-\mu)},\ \ \mu < x < \infty.$$

Raskite TG kriterijų hipotezėms

a) $H : \sigma = \sigma_0$, kai μ žinomas;

b) $H : \mu = \mu_0$, kai σ žinomas,

esant vienpusėms alternatyvoms, tikrinti.

I.3.42. Tegu $X_1, \dots, X_n$ yra nepriklausomi vienodai pasiskirstę a. d., kurių tankio funkcija $f_\theta(x)$, $\theta \in \Theta \subset \mathbf{R}$. Raskite tolygiai galingiausią α lygmens kriterijų hipotezei $H_0 : \theta \leqslant \theta_0$, kai alternatyva yra $H_1 : \theta > \theta_0$, tikrinti tokiais atvejais:

a) $f_\theta(x) = \theta^{-1}\mathrm{e}^{-\mathrm{x}/\theta}$, $0 < \mathrm{x} < \infty$, $\theta > 0$;

b) $f_\theta(x) = \theta x^{\theta-1}$, $0 < x < 1$, $\theta > 0$;

c) f_θ yra $N(1,\theta)$ skirstinio tankio funkcija;

d) $f_\theta(x) = \theta^{-c} c x^{c-1} \mathrm{e}^{-(\mathrm{x}/\theta)^{\mathrm{c}}}$, $0 < \mathrm{x} < \infty$, $\theta > 0$; čia $c > 0$ – žinomas.

I.3.43. Tegu $X_1, \dots, X_n$ yra nepriklausomi vienodai pasiskirstę a. d., turintys tolygųjį skirstinį $U(\theta, \theta+1)$, $\theta \in \mathcal{R}$. Tegu $n \geqslant 2$.

a) Raskite bendrą $X_{(1)}$ ir $X_{(n)}$ skirstinį.

b) Tegu tikrinant hipotezę $H : \theta = \theta_0$ taikomas toks kriterijus: hipotezė H priimama, kai $X_{(n)} - 1 < \theta_0 < X_{(1)}$. Koks šio kriterijaus reikšmingumo lygmuo?

c) Raskite p. b) pateikto kriterijaus galią, kai $\theta > \theta_0$.

d) Raskite imties didumą, kad p. b) apibrėžtas kriterijus atmestų hipotezę H su tikimybe, ne mažesne už 0,99, jei tikroji parametro θ reikšmė tenkina nelygybę $\theta \ge \theta_0 + 0,1$.

I.3.44. Tegu $\mathbf{X} = (X_1, \ldots, X_n)^T$ yra paprastoji atsitiktinė imtis a. d. X, turinčio diskretųjį tolygųjį skirstinį, sukoncentruotą taškuose $0,\ 1, ..., \theta$, kai nežinomas $\theta = 1,\ 2, ...$

a) Tegu tikrinama hipotezė $H_0 : \theta \leqslant \theta_0$, kai alternatyva yra $H_1 : \theta > \theta_0$. Įrodykite, kad

$$\varphi_*(\mathbf{X}) = \begin{cases} 1, & X_{(n)} > \theta_0, \\ \alpha, & X_{(n)} \leqslant \theta_0 \end{cases}$$

yra α lygmens TG kriterijus.

b) Tegu tikrinama hipotezė $H_0 : \theta = \theta_0$, kai alternatyva yra $H_1 : \theta \neq \theta_0$. Įrodykite, kad

$$\varphi_*(\mathbf{X}) = \begin{cases} 1, & X_{(n)} > \theta_0 \text{ arba } X_{(n)} \leqslant \theta_0 \alpha^{1/n}, \\ 0, & \text{priešingu atveju} \end{cases}$$

yra α lygmens TG kriterijus.

c) Įrodykite, kad a) ir b) punktai išlieka teisingi, diskretųjį tolygųjį skirstinį pakeitus tolygiuoju skirstiniu $U(0, \theta)$, $\theta > 0$.

I.3.45. Įrodykite, kad šios šeimos turi monotoninį tikėtinumo santykį:

a) Laplaso skirstinių šeima $\{L(\mu, \theta)\}$, kai θ žinomas;

b) paslinktųjų eksponentinių skirstinių šeima $\{\mathcal{E}(\theta, c)\}$, kai c žinomas;

c) logistinių skirstinių šeima $\{LG(\theta, c)\}$, kai c žinomas;

d) tolygiųjų skirstinių šeima $\{U(\theta, \theta + 1)\}$;

e) hipergeometrinių skirstinių šeima $\{H(N, M, n)\}$, kai n ir N žinomi.

I.3.46. Įrodykite, kad šeima $\{f_\theta : \theta \in \mathbf{R}\}$, kai $f_\theta(x) = c(\theta)h(x)$, $a(\theta) < x < b(\theta)$, turi monotoninį tikėtinumo santykį; čia $h(x)$ yra pagal Lebego matą integruojama teigiama funkcija, $a(\theta < b(\theta)$ yra nemažėjančios θ funkcijos.

I.3.47. Tegu X turi vienparametrį eksponentinio tipo skirstinį. Įrodykite, kad TG kriterijus, kai alternatyvos dvipusės, neegzistuoja.

I.3.48. Paprastoji imtis $\mathbf{X} = (X_1, ..., X_n)^T$ gauta stebint a. d. $X \sim N(\mu, k\mu)$, $\mu > 0$, k – žinoma teigiama konstanta. Sudarykite kriterijų hipotezei $H : \mu = \mu_0$, kai alternatyva yra $\bar{H} : \mu > \mu_0$, tikrinti. Tarę, kad n didelis, raskite apytiksles kritinės srities ir galios funkcijos išraiškas.

I.3.49. (**I.3.48** pratimo tęsinys.) Tarę, kad žinomas tik vidurkis $\bar{X}$, raskite TG kriterijų hipotezei $H : \mu = \mu_0$, kai alternatyva yra $\bar{H} : \mu > \mu_0$, tikrinti, grindžiamą statistika $\bar{X}$. Raskite kriterijaus galios funkciją.

I.3.50. Atsitiktinis dydis X įgyja sveikąsias neneigiamas reikšmes ir jo pasiskirstymo funkcija

$$F(x|\beta) = \mathbf{P}_\beta\{X \leq x\} = 1 - \beta^x,\ 0 < \beta < 1,\ x = 0, 1, 2, ...$$

Raskite kriterijų hipotezei $H : \beta = \beta_0$, kai alternatyva yra $\bar{H} : \beta > \beta_0$, tikrinti.

I.3.4. Skirstiniai, priklausantys nuo keleto parametrų

I.3.51. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. $X \sim N(\mu, \sigma^2), -\infty < \mu < \infty$, $0 < \sigma$, kai abu parametrai nežinomi. Raskite reikšmingumo lygmens α hipotezės $H : \mu = \mu_0$, kai alternatyvos yra $\bar{H}_1 : \mu > \mu_0$ arba $\bar{H}_2 : \mu < \mu_0$, tikrinimo TGN kriterijus.

I.3.52. (**I.3.51** pratimo tęsinys). Raskite reikšmingumo lygmens α TGN kriterijų tikrindami hipotezę $H : \mu = \mu_0$, kai alternatyva dvipusė $\bar{H}_3 : \mu \neq \mu_0$.

I.3.53. (**I.3.51** pratimo tęsinys). Raskite reikšmingumo lygmens α TGN kriterijų tikrindami hipotezę $H : \sigma^2 = \sigma_0^2$, kai alternatyvos vienpusės

$$\bar{H}_1 : \sigma^2 > \sigma_0^2 \text{ arba } \bar{H}_2 : \sigma^2 < \sigma_0^2.$$

I.3.54. (**I.3.51** pratimo tęsinys). Raskite reikšmingumo lygmens α TGN kriterijų tikrindami hipotezę $H : \sigma^2 = \sigma_0^2$, kai alternatyva dvipusė $\bar{H}_3 : \sigma^2 \neq \sigma_0^2$.

I.3.55. Remiantis didumo n imtimi, tikrinama hipotezė $H : \mu = 0$ apie normaliojo skirstinio vidurkio reikšmę, esant alternatyvai $\bar{H}_1 : \mu > 0$ arba alternatyvai $\bar{H}_3 : \mu \neq 0$, kai σ nežinomas. Įrodykite, kad kriterijaus galia yra didėjanti μ/σ funkcija, kai alternatyva yra $\bar{H}_1$, ir didėjanti $|\mu|/\sigma$ funkcija, kai alternatyva yra $\bar{H}_3$.

I.3.56. (**I.3.55** pratimo tęsinys). Įrodykite, kad kriterijus, kurio reikšmingumo lygmuo yra α ir kurio galia, esant visoms alternatyvoms $\{(\mu,\ \sigma) : \mu > \mu_1 > 0\}$, yra ne mažesnė už β, $\beta > \alpha$, neegzistuoja.

I.3.57. (**I.3.55** pratimo tęsinys). Kai alternatyva yra $\bar{H}_1$, Stjudento kriterijaus galią palyginkite su atitinkamo kriterijaus, kai σ žinomas, galia, jei $n = 5;\ 10;\ 15$ ir $\mu/\sigma = 0,8;\ 1,0;\ 1,2$ (kriterijų reikšmingumo lygmuo $\alpha = 0,05$).

I.3.58. Tegu $X_1, X_2, ..., X_n$ yra nepriklausomi normalieji atsitiktiniai dydžiai $X_i \sim N(\mu_i, \sigma^2)$, $i = 1, ..., s$ ir $X_i \sim N(0,\ \sigma^2)$, $i = s+1, ..., n$; čia $-\infty < \mu_i < \infty$, $i = 1, ..., s$; $0 < \sigma < \infty$. Raskite TGN kriterijus hipotezei $H : \mu_1 = \mu_1^0$, esant vienpusėms ir dvipusei alternatyvoms, tikrinti.

I.3.59. Tegu $X_i = \beta_0 + \beta_1(t_i - \bar{t}) + \varepsilon_i$; čia t_i yra fiksuotos konstantos (ne visos vienodos), $\bar{t} = \sum_i t_i/n$, ε_i yra nepriklausomi vienodai pasiskirstę a. d., turintys normalųjį skirstinį $N(0, \sigma^2)$, β_0, β_1 ir σ^2 yra nežinomi parametrai. Raskite α lygmens TGN kriterijų hipotezėms tikrinti:

a) $H_0 : \beta_0 \leqslant \theta_0$, kai alternatyva $H_1 : \beta_0 > \theta_0$;
b) $H_0 : \beta_0 = \theta_0$, kai alternatyva $H_1 : \beta_0 \neq \theta_0$;
c) $H_0 : \beta_1 \leqslant \theta_0$, kai alternatyva $H_1 : \beta_1 > \theta_0$;
d) $H_0 : \beta_1 = \theta_0$, kai alternatyva $H_1 : \beta_1 \neq \theta_0$.

I.3.60. Tegu $X_1, \ldots, X_n$ yra nepriklausomi vienodai pasiskirstę a. d., turintys gama skirstinį $G(\lambda, \eta)$ su nežinomais λ ir η.

a) Įrodykite, kad tikrinant hipotezę $H_0 : \gamma \leq \gamma_0$, kai alternatyva yra $H_1 : \gamma > \gamma_0$, TGN kriterijus atmeta H_0, kai $\prod_{i=1}^n X_i > g(\bar{X})$; čia g – tam tikra reali funkcija.

b) Įrodykite, kad tikrinant hipotezę $H_0 : \eta \leqslant \eta_0$, kai alternatyva yra $H_1 : \eta > \eta_0$, TGN kriterijus atmeta H_0, kai $\bar{X} > h\left(\prod_{i=1}^n X_i\right)$; čia h – tam tikra reali funkcija.

I.3.61. Tarkime, $\mathbf{X} = (X_1, ..., X_n)^T$ ir $\boldsymbol{Y} = (Y_1, ..., Y_m)^T$ yra nepriklausomos paprastosios imtys, gautos stebint n. a. d. $\boldsymbol{X} \sim N(\mu_1, \sigma_1^2)$ ir $Y \sim N(\mu_2, \sigma_2^2)$. Raskite reikšmingumo lygmens α TGN kriterijus tikrindami hipotezę $H : \mu_1 - \mu_2 = \beta_0$ su vienpusėmis ir dvipuse alternatyvomis, kai dispersijos yra žinomos $\sigma_1^2 = \sigma_{10}^2, \sigma_2^2 = \sigma_{20}^2$.

I.3.62. (**I.3.61** pratimo tęsinys). Raskite reikšmingumo lygmens α TGN kriterijus tikrindami hipotezę $H : \mu_1 - \mu_2 = \beta_0$ su vienpusėmis ir dvipuse alternatyvomis, kai dispersijos yra lygios $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

I.3.63. (**I.3.61** pratimo tęsinys). Raskite reikšmingumo lygmens α TGN kriterijus tikrindami hipotezę $H : \sigma_1^2 = \sigma_2^2$ su vienpusėmis ir dvipuse alternatyvomis.

I.3.64. Tarkime, $(X_i, Y_i)^T, i = 1, ..., n$, yra paprastoji imtis a. v. $(X, Y)^T \sim N_2(\boldsymbol{\mu}, \Sigma)$, $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$, $\Sigma = [\sigma_{ij}]_{2\times 2}, \sigma_{11} = \sigma_1^2, \sigma_{22} = \sigma_2^2, \sigma_{12} = \rho\sigma_1\sigma_2$. Raskite TGN reikšmingumo lygmens α kriterijus a. d. X ir Y nepriklausomumo hipotezei $H : \rho = 0$ tikrinti vienpusių ir dvipusės alternatyvų atvejais.

I.3.65. (**I.3.64** pratimo tęsinys). a) Raskite reikšmingumo lygmens α tikėtinumų santykio kriterijų hipotezei $H : \rho = \rho_0 \neq 0$ tikrinti, kai alternatyva $\bar{H}_3 : \rho \neq \rho_0$, b) Pasiūlykite paprasčiau randamą simetrišką kriterijų. c) Pateikite kriterijus vienpusių alternatyvų atvejais.

I.3.66. (**I.3.64** pratimo tęsinys). Raskite reikšmingumo lygmens α apytikslius kriterijus hipotezei $H : \rho = \rho_0 \neq 0$ tikrinti vienpusių ir dvipusės alternatyvų atvejais naudodami Fišerio dispersijų stabilizuojančią transformaciją.

I.3.67. (**I.3.64** pratimo tęsinys). Raskite reikšmingumo lygmens α kriterijus hipotezei $H : \mu_1 = \mu_2 = \beta_0$ tikrinti vienpusių ir dvipusės alternatyvų atvejais.

I.3.68. Tarkime, paprastosios imtys $\mathbf{X} = (X_1, ..., X_m)^T$ ir $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ gautos stebint n. a. d. $X \sim \mathcal{P}(\lambda_1)$ ir $Y \sim \mathcal{P}(\lambda_2)$. Raskite reikšmingumo lygmens α TGN kriterijus hipotezei $H : \lambda_1/\lambda_2 = c_0$ tikrinti vienpusių ir dvipusės alternatyvų atvejais.

I.3.69. (**I.3.68** pratimo tęsinys). Per pirmąją ir antrąją valandas į komutatorių buvo kreiptasi atitinkamai 15 ir 13 kartų. Kitą dieną per 5 valandas buvo kreiptasi 45 kartus. Tarus, kad iškvietimų skaičiai pasiskirstę pagal Puasono dėsnį su parametru λ (vidutinis iškvietimų skaičius per valandą), reikia patikrinti, ar iškvietimų intensyvumas nepakito (kriterijaus reikšmingumo lygmuo $\alpha = 0,05$).

I.3.70. Tegu X_1 ir X_2 yra nepriklausomi a. d., turintys Puasono skirstinius $\mathcal{P}(\lambda_1)$ ir $\mathcal{P}(\lambda_2)$.

a) Raskite α lygmens TGN kriterijų hipotezei $H_0 : \lambda_1 \geqslant \lambda_2$, kai alternatyva yra $H_1 : \lambda_1 < \lambda_2$, tikrinti.

b) Apskaičiuokite punkte a) gauto kriterijaus galią, kai $\alpha = 0,1$, $(\lambda_1, \lambda_2) = (0,1,\ 0,2)$; $(1,\ 2)$; $(10,\ 20)$; $(0,1,\ 0,4)$.

I.3.71. Tarkime, paprastosios imtys $\mathbf{X} = (X_1, ..., X_m)^T$ ir $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ gautos stebint n. a. d. $X \sim B(1, p_1)$ ir $Y \sim B(1, p_2))$. Raskite reikšmingumo lygmens α TGN kriterijus hipotezei $H : p_1 = p_2$ tikrinti vienpusių ir dvipusės alternatyvų atvejais.

I.3.72. (**I.3.71** pratimo tęsinys). Vienos brigados 20 darbininkų buvo paskiepyti nuo gripo, per 6 mėnesius iš jų susirgo 6 darbininkai. Tos pačios brigados 5 darbininkai skiepytis atsisakė, 4 iš jų susirgo per tą patį 6 mėnesių laikotarpį. Ar galima daryti išvadą apie teigiamą priešgripinio serumo poveikį?

I.3.73. Tegu $\mathbf{X}_1 = X_{11}, \ldots, X_{1n_1}$ ir $\mathbf{X}_2 = X_{21}, \ldots, X_{2n_2}$ yra dvi nepriklausomos imtys vienodai pasiskirsčiusių n. a. d., turinčių atitinkamai gama skirstinius $\Gamma(\theta_1, \gamma_1)$ ir $\Gamma(\theta_2, \gamma_2)$.

Tegu γ_1 ir γ_2 žinomi. Įrodykite, kad, tikrinant hipotezę $H_0 : \theta_1 \leqslant \theta_2$, kai alternatyva yra $H_1 : \theta_1 > \theta_2$, ir hipotezę $H_0 : \theta_1 = \theta_2$, kai alternatyva $H_1 : \theta_1 \neq \theta_2$, egzistuoja TGN kriterijai, kurių statistikų skirstiniai išreiškiami beta skirstiniais.

I.3.74. Tegu $(X_i, Y_i)^T, i = 1, ..., n$, yra paprastoji atsitiktinė imtis dvimačio normaliojo a. v. $(X,\ Y)^T \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\mu} = (\mu_1,\ \mu_2)^T$, $-\infty < \mu_1, \mu_2 < \infty$, $\boldsymbol{\Sigma} = [\sigma_{ij}]_{2\times 2}, \sigma_{11} = \sigma_1^2,\ \sigma_{22} = \sigma_2^2,\ \sigma_{12} = \sigma_{21} = \rho\sigma_1\sigma_2,\ 0 < \sigma_1,\ \sigma_2 < \infty,\ -1 < \rho < 1$.

Be to, $S_{11} = \sum_i (X_i - \bar{X})^2$, $S_{22} = \sum_i (Y_i - \bar{Y})^2$ ir $S_{12} = \sum_i (X_i - \bar{X})(Y_i - \bar{Y})$.

a) Įrodykite, kad TGN kriterijus hipotezei $H_0 : \sigma_2/\sigma_1 = \Delta_0$, kai alternatyva yra $H_1 : \sigma_2/\sigma_1 \neq \Delta_0$, tikrinti atmeta H_0, kai

$$R = |\Delta_0^2 S_{11} - S_{22}| / \sqrt{(\Delta_0^2 S_{11} + S_{22})^2 - 4\Delta_0^2 S_{12}^2} > c.$$

b) Raskite R iš a) punkto skirstinį, kai $\sigma_2/\sigma_1 = \Delta_0$.

c) Tegu $\sigma_1 = \sigma_2$. Įrodykite, kad TGN kriterijus hipotezei $H_0 : \mu_1 = \mu_2$, kai alternatyva yra $H_1 : \mu_1 \neq \mu_2$, tikrinti atmeta H_0, kai

$$V = |\bar{X}_2 - \bar{X}_1| / \sqrt{S_1^2 + S_2^2 - 2S_{12}} > c.$$

d) Raskite punkto c) a. d. V skirstinį, kai $\mu_1 = \mu_2$.

I.3.75. Tegu $X_1, \ldots, X_n$ yra nepriklausomi vienodai pasiskirstę a. d., turintys paslinktąjį eksponentinį skirstinį $\mathcal{E}(a, \theta)$ su nežinomais a ir θ.

a) Įrodykite, kad tikrinant $H_0 : \theta = 1$, kai alternatyva yra $H_1 : \theta \neq 1$, α lygmens TGN kriterijus atmeta H_0, kai $V < c_1$ arba $V > c_2$; čia $V = 2\sum_{i=1}^n (X_i - X_{(1)}) = 2n(\bar{X} - X_{(1)})$, o c_i apibrėžti taip:

$$\int_{c_1}^{c_2} f(x|2n-2)dx = \int_{c_1}^{c_2} f(x|2n)dx = 1 - \alpha,$$

$f(x|\nu)$ yra χ^2 skirstinio su ν laisvės laipsnių tankio funkcija.

b) Įrodykite, kad, tikrinant hipotezę $H_0 : a = 0$, kai alternatyva yra $H_1 : a \neq 0$, lygmens α TGN kriterijus atmeta H_0, kai $X_{(1)} < 0$ arba $2nX_{(1)}/V > c(n-1)$; čia c randamas iš lygties

$$(n-1)\int_0^c (1+v)^{-n} dv = 1 - \alpha.$$

I.3.5. Hipotezių tikrinimas, kai imtys didelės

I.3.76. Tegu $X_{i1}, ..., X_{in_i}$, $i = 1, ..., k$, $n_i \geq 2$, yra k paprastųjų nepriklausomų imčių eksponentinių a. d. $X_1, ..., X_k$, kurių tikimybiniai tankiai yra

$$\frac{1}{\sigma_i}\exp\{-\frac{x-\theta_i}{\sigma_i}\},\ \theta_i < x < \infty,\ i = 1, ..., k;$$

čia $0 < \sigma_i < \infty$, $-\infty < \theta_i < +\infty$, $i = 1, ..., k$.

Raskite tikėtinumų santykį, kai tikrinama hipotezė: a) $H_1 : \theta_1 = \theta_2 = ... = \theta_k$; b) $H_2 : \sigma_1 = ... = \sigma_k$; c) $H_3 : \theta_1 = ... = \theta_k$, kai visi σ_i yra lygūs.

I.3.77. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. $X \sim N(\mu, \sigma^2)$, $-\infty < \mu < +\infty$, $0 < \sigma < +\infty$. Įrodykite, kad: a) tikėtinumų santykio kriterijus hipotezei $H : \mu = \mu_0$ tikrinti yra ekvivalentus Stjudento kriterijui; b) tikėtinumų santykio kriterijus hipotezei $H : \sigma^2 = \sigma_0^2$ tikrinti yra ekvivalentus χ^2 kriterijui.

I.3.78. Tegu $(X_{1i}, ..., X_{ki})^T$, $i = 1, ..., n$, yra imtis vektoriaus $(X_1, ..., X_k)^T$, kurio skirstinys priklauso polinominių skirstinių šeimai $\{\mathcal{P}_k(1,\ \boldsymbol{\pi})\}$; čia $\boldsymbol{\pi} = (\pi_1, ..., \pi_k)^T$ yra

k-matis vektorius, kurio koordinatės tenkina sąlygas $0 < \pi_i < 1,\ \pi_1 + ... + \pi_k = 1$. Įrodykite, kad tikėtinumų santykis hipotezei $H : \pi_1 = \pi_1^0, ..., \pi_k = \pi_k^0$ tikrinti yra

$$\Lambda = \left(\prod_{i=1}^{k} \left(\frac{\pi_i^0}{\hat{\pi}_i} \right)^{\hat{\pi}_i} \right)^n;$$

čia $\hat{\pi}_i = V_i/n,\ V_i = X_{i1} + ... + X_{in},\ i = 1, ..., k$.

I.3.79. (**I.3.78** pratimo tęsinys). Įrodykite, kad statistikų $-2\ln\Lambda$ ir $\sum_i (V_i - n\pi_i^0)^2/n\pi_i^0$ skirstiniai, kai hipotezė H yra teisinga, silpnai konverguoja į χ^2 skirstinį su $k-1$ laisvės laipsnių.

I.3.80. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ ir $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ yra nepriklausomos paprastosios a. d. $X \sim N(\mu_1, \sigma_1^2)$ ir $Y \sim N(\mu_2, \sigma_2^2)$ imtys. Įrodykite, kad: a) tikėtinumų santykio kriterijus hipotezei $H : \mu_1 = \mu_2$, kai alternatyva yra $\bar{H} : \mu_1 \neq \mu_2$, tikrinti, kai $\sigma_1^2 = \sigma_2^2$, yra ekvivalentus Stjudento kriterijui; b) tikėtinumų santykio kriterijus hipotezei $H : \sigma_1^2 = \sigma_2^2$, kai alternatyva yra $\bar{H} : \sigma_1^2 \neq \sigma_2^2$, tikrinti ekvivalentus Fišerio kriterijui.

I.3.81. Tegu $(X_{i1}, ..., X_{in_i})^T,\ i = 1, ..., k$, yra k paprastųjų nepriklausomų imčių, gautų stebint normaliuosius a. d. $X_i \sim N(\mu_i,\ \sigma_i^2)$. Įrodykite, kad tikėtinumų santykio kriterijus hipotezei $H : \mu_1 = ... = \mu_k$ tikrinti, kai $\sigma_1^2 = ... = \sigma_k^2$, yra ekvivalentus kriterijui, grindžiamam statistika

$$F = \frac{(n-k)\sum_{i=1}^{k} n_i(\bar{X}_{i.} - \bar{X})^2}{k\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_{i.})^2},$$

kurios skirstinys, kai H teisinga, yra Fišerio $F(k,\ n-k),\ n = n_1 + ... + n_k$,

$$\bar{X}_{i.} = \frac{1}{n_i}\sum_{j-1}^{n_i} X_{ij},\ \bar{X} = \frac{1}{n}\sum_{i=1}^{k} n_i \bar{X}_{i.}.$$

I.3.82. Tegu $(X_1, ..., X_{n_1})^T$ ir $(Y_1, ..., Y_{n_2})^T$ yra paprastosios nepriklausomos imtys a. d. $X \sim B(1, p_1),\ 0 < p_1 < 1$, ir $Y \sim B(1, p_2),\ 0 < p_2 < 1$. Raskite tikėtinumų santykį Λ hipotezei $H : p_1 = p_2$ tikrinti ir įrodykite, kad $-2\ln\Lambda \xrightarrow{d} \chi^2(1)$, kai $n_1,\ n_2 \to \infty$, ir H yra teisinga.

I.3.83. Apibendrinkite **I.3.82** pratimą ir jo sprendimą tuo atveju, kai imčių skaičius didesnis už 2.

I.3.84. Tegu $(X_1, ..., X_{n_1})^T$ ir $(Y_1, ..., Y_{n_2})^T$ yra paprastosios nepriklausomos imtys a. d. $X \sim \mathcal{P}(\lambda_1),\ 0 < \lambda_1 < \infty$, ir $Y \sim \mathcal{P}(\lambda_2),\ 0 < \lambda_2 < \infty$. Raskite tikėtinumų santykį Λ hipotezei $H : \lambda_1 = \lambda_2$ tikrinti ir įrodykite, kad $-2\ln\Lambda \xrightarrow{d} \chi^2(1)$, kai $n_1,\ n_2 \to \infty$, ir H yra teisinga.

I.3.85. Apibendrinkite **I.3.84** pratimą tuo atveju, kai imčių skaičius didesnis už 2.

I.3.86. Tegu $(X_1, ..., X_{n_1})^T$ ir $(Y_1, ..., Y_{n_2})^T$ yra paprastosios nepriklausomos imtys a. d. X ir Y, turinčių eksponentinius skirstinius $X \sim \mathcal{E}(1/\theta_1)$ ir $Y \sim \mathcal{E}(1/\theta_2),\ 0 < \theta_1,\ \theta_2 < \infty$. Raskite statistikos $\bar{X}/\bar{Y}$ skirstinį. Įrodykite, kad tikėtinumų santykio hopotezei $H : \theta_1 = \theta_2$ tikrinti statistikos yra $\bar{X}/\bar{Y}$ funkcijos. Raskite kriterijų galią.

I.3.87. Atlikta 500 nepriklausomų stebėjimų ir jie sugrupuoti į intervalus; m_i – stebėjimų, patekusių į atitinkamą intervalą, skaičius.

Intervalas	m_i
$(-\infty,\ -3/2)$	2
$[-3/2,\ -1/2)$	78
$[-1/2,\ 1/2)$	339
$(1/2,\ \infty)$	81

Ar neprieštarauja šie duomenys prielaidai, kad buvo sugrupuota paprastosios atsitiktinio dydžio $X \sim N(0,\ 1/4)$ imties realizacija?

I.3.88. Lentelėje iš 2 000 atsitiktinių skaičių skaitmuo 0 aptinkamas 160 kartų, skaitmuo 3 – 247 kartus, skaitmuo 6 – 191 kartą, o likusieji skaitmenys – 1 402 kartus. Ar neprieštarauja šie duomenys prielaidai, kad skaitmenys 0,1,...,9 pasitaiko su vienodomis tikimybėmis 1/10?

I.3.89. Tarp 2 020 šeimų, turinčių du vaikus, užregistruota 527 šeimos, kuriose abu vaikai berniukai; 476 šeimos, kur abu vaikai mergaitės, o likusiose 1 017 šeimų – vienas berniukas ir viena mergaitė. Patikrinkite prielaidą apie berniuko ir mergaitės gimimo tikimybių lygybę. Patikrinkite prielaidą, kad berniukų skaičius X šeimose, turinčiose du vaikus, yra binominis $X \sim B(2,\ p)$.

I.3.90. Atlikus 200 nepriklausomų bandymų, įvykiai A, B ir C pasirodė atitinkamai 49, 93 ir 58 kartus. Patikrinkite hipotezę, pagal kurią $\mathbf{P}\{A\} = \mathbf{P}\{C\} = p$, $\mathbf{P}\{B\} = 1 - 2p$, $0 < p < 1/2$.

I.3.91. Atlikus 8 000 nepriklausomų bandymų, įvykiai A, B ir C įvyko atitinkamai 2 018, 5 012 ir 970 kartų. Patikrinkite hipotezę, pagal kurią $\mathbf{P}\{A\} = 1/2 - 2p$, $\mathbf{P}\{B\} = 1/2 + p$, $\mathbf{P}\{C\} = p$, $0 < p < 1/4$.

I.3.6. Hipotezių tikrinimo pavyzdžiai

I.3.92. Tikrinama hipotezė $H : \mu = 1$, kai alternatyva yra $\bar{H} : \mu \neq 1$, remiantis a. d. $X \sim N(\mu,\ 4)$ paprastąja imtimi. Kokio didumo turi būti imtis, kad hipotezė H būtų atmetama su tikimybe 0,05, kai ji teisinga, ir priimama, kai tikroji parametro reikšmė tenkina nelygybę $|\mu - 1| \geq 1$ su tikimybe, ne didesne kaip 0,01?

I.3.93. Remiantis $n = 50$ didumo normaliojo skirstinio $N(0, \sigma^2)$ imtimi, tikrinama hipotezė $H : \sigma^2 = \sigma_0^2$, kai alternatyva yra $\bar{H} : \sigma^2 > \sigma_0^2$. Kokia tikimybė, kad ta hipotezė bus atmesta, jei tikroji parametro reikšmė σ^2 tenkina nelygybę $\sigma^2 > 1,5\sigma_0^2$, o kriterijaus reikšmingumo lygmuo $\alpha = 0,05$? Kokio didumo turi būti imtis, kad ta tikimybė būtų ne mažesnė už 0,95?

I.3.94. Tegu hipotezei $H : \sigma^2 = \sigma_0^2$, kai alternatyva yra $\bar{H} : \sigma^2 \neq \sigma_0^2$, tikrinti taikomas TGN kriterijus ir paslinktasis kriterijus. Raskite, kokio didumo turi būti imtys, kad kriterijų galios funkcijos būtų ne mažesnės už 0,9, kai $\sigma^2 \geq 2\sigma_0^2$ ir $\sigma^2 \leq \sigma_0^2/2$, o kriterijų reikšmingumo lygmuo yra 0,05.

I.3.95. Lentelėje pateikti duomenys apie dviejose fermose vienodo amžiaus kiaulių svorio prieaugį per tam tikrą laiką. Pirmoje fermoje pamatuota $n_1 = 16$ kiaulių svorio prieaugis $X_i, i = 1, ..., 16$; antroje fermoje $n_2 = 15$ kiaulių svorio prieaugis $Y_i, i = 1, ..., 15$. Reikia patikrinti hipotezę, kad vidutinis svorio priaugis nesiskiria, kai alternatyva yra, jog pirmoje fermoje vidutinis svorio priaugis yra didesnis.

I ferma				II ferma			
i	X_i	i	X_i	i	Y_i	i	Y_i
1	109,95	9	108,86	1	81,45	9	85,63
2	103,54	10	98,69	2	94,63	10	90,92
3	104,58	11	97,51	3	73,70	11	95,58
4	114,43	12	100,48	4	87,36	12	71,52
5	90,92	13	96,76	5	89,12	13	108,85
6	104,59	14	102,77	6	96,69	14	87,36
7	103,85	15	100,47	7	83,93	15	99,48
8	88,23	16	99,48	8	86,49		

Nurodymas. Iš lentelės matyti, kad a. d. skirstiniai asimetriški. Todėl reikėtų atlikti stebimojo dydžio transformaciją, kad naujo a. d. skirstinys būtų patenkinamai aprašomas normaliuoju skirstiniu, paskui remtis Stjudento kriterijumi. Nesunku įsitikinti, kad nagrinėjamame pavyzdyje stebėjimų logaritmai tiksliau aprašomi normaliuoju skirstiniu. Kitaip sakant, stebimasis a. d. tiksliau aprašomas lognormaliuoju skirstiniu.

I.3.96. Užregistruota 100 metų duomenys apie vidutinę liepos mėnesio temperatūrą. Remiantis šiais duomenimis, gauta $\bar{X} = 16,482$, $s = 1,6145$. Naudojant šio laikotarpio 30 pirmųjų metų duomenis, gauti įverčiai $\bar{X}_1 = 16,893$, $s_1 = 1,5904$, o pagal paskutiniųjų 30 metų duomenis – įverčiai $\bar{X}_2 = 15,963$, $s_2 = 1,6531$. Patikrinkite hipotezes, kad šių dviejų laikotarpių vidutinė temperatūra nesiskiria nuo vidurinio laikotarpio vidutinės temperatūros, tarę, kad vidutinę temperatūrą galima aprašyti normaliuoju skirstiniu.

I.3.97. Pagal dvi nepriklausomas $n_1 = n_2 = 50$ imtis, gautas stebint n. a. d. $X \sim N(\mu_1,\ 1)$ ir $Y \sim N(\mu_2,\ 1)$, gauti įverčiai $\bar{X} = 0,103$ ir $\bar{Y} = 0,368$. Sudarykite TG kriterijų hipotezei $H : \mu_1 = \mu_2$, kai alternatyva yra $\bar{H} : \mu_1 < \mu_2$, tikrinti. Ar ši hipotezė atmetama pagal turimas realizacijas, jeigu kriterijaus reikšmingumo lygmuo $\alpha = 0,05$?

I.3.98. Yra dvi nepriklausomos paprastosios vienodo didumo n imtys, gautos stebint nepriklausomus normaliuosius a. d., ir, remiantis Fišerio kriterijumi, tikrinama hipotezė $H : \sigma_1^2 = \sigma_2^2$, kai alternatyva yra $\bar{H} : \sigma_1^2/\sigma_2^2 > 1$. Raskite tokį imties didumą n, kad kriterijaus galia būtų ne mažesnė už 0,9, jei kriterijaus reikšmingumo lygmuo yra $\alpha = 0,05$ ir $\sigma_1^2/\sigma_2^2 = 1,5;\ 2;\ 3$.

I.3.99. Dviejose laboratorijose buvo matuojamas sieros dyzeliniame kure kiekis pagal identiškus pavyzdžius, kuriuose sieros kiekis buvo 0,870. Atlikus 8 nepriklausomus matavimus, pirmoje laboratorijoje gauti tokie rezultatai: 0,869; 0,874; 0,867; 0,875; 0,870; 0,869; 0,864; 0,872. Kitoje laboratorijoje atlikus 10 matavimų, gauti tokie rezultatai: 0,865; 0,870; 0,866; 0,871; 0,868; 0,870; 0,871; 0,870; 0,869; 0,874. Tarę, kad matavimo paklaidos turi normaliuosius skirstinius, patikrinkite dispersijų lygybės hipotezę. Tarę, kad dispersijos vienodos, patikrinkite laboratorijų paklaidų vidurkių vienodumo hipotezę.

I.3.100. Tikrinama hipotezė, kad impulso atpažinimo paklaidos dispersija nepriklauso nuo jo intensyvumo. Buvo atlikti du nepriklausomi eksperimentai. Impulsas, kurio intensyvumas 10 sąlyginių vienetų, buvo įvertintas taip: 9, 9, 8, 10, 12, 12, 13, 10, 10; impulsas, kurio intensyvumas 20 sąlyginių vienetų, – taip: 15, 16, 17, 23, 22, 20, 21, 24, 27. Ar šie duomenys neprieštarauja iškeltajai hipotezei (tarkite, kad buvo stebimi nepriklausomi normalieji a. d.)?

I.3.101. (**I.2.206** pratimo tęsinys). **I.2.206** pratimo sąlygomis a) patikrinkite trijų dispersijų lygybės hipotezę $H : \sigma_x^2 = \sigma_y^2 = \sigma_z^2$; b) patikrinkite hipotezę $H : \theta = 0$.

I.3.102. Tikrinant keturias didumo $n_1 = 20, n_2 = 38, n_3 = 25, n_4 = 50$ lempučių partijas, gautos jų darbo laiko iki gedimo vidutinės reikšmės $\bar{T}_1 = 154,3$, $\bar{T}_2 = 165,1$, $\bar{T}_3 = 159,0$, $\bar{T}_4 = 175,5$. Tardami, kad i-osios partijos lemputės darbo laikas iki gedimo turi eksponentinį skirstinį $\mathcal{E}(1/\lambda)$, patikrinkite hipotezę $H : \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4$.

I.3.103. Tiriant specialios sėjamosios efektyvumą, 10 sklypelių buvo sėjama paprasta sėjamąja ir 10 sklypelių – specialia sėjamąja, paskui buvo lyginamas derlingumas. Dvidešimt vienodo ploto sklypelių buvo taip sugrupuoti poromis, kad būtų greta vienas kito. Metant monetą buvo pasirenkama, kuriame iš dviejų sklypelių sėti specialia sėjamąja. Rezultatai pateikti lentelėje.

Eil. Nr.	Speciali	Paprasta	Eil. Nr.	Speciali	Paprasta
1	8,0	5,6	6	7,7	6,1
2	8,4	7,4	7	7,7	6,6
3	8,0	7,3	8	5,6	6,0
4	6,4	6,4	9	5,6	5,5
5	8.6	7,5	10	6,2	5,5

Patikrinkite hipotezę, kad abiejų sėjamųjų efektyvumas vienodas: a) taikydami dviejų imčių Stjudento kriterijų; b) taikydami Stjudento kriterijų atitinkamų sklypelių derlingumų skirtumams; c) paaiškinkite, kodėl gaunamos skirtingos išvados.

I.3.104. (**I.3.103** pratimo tęsinys). Įvertinkite koreliacijos koeficientą ir patikrinkite koreliacijos koeficiento lygybės 0 hipotezę.

I.3.105. (**I.2.207** pratimo tęsinys). **I.2.207** pratimo sąlygomis patikrinkite prielaidą, kad a. d. X ir Y vidurkiai nesiskiria.

I.3.106. Lentelėje nurodyta 10 pacientų, vartojusių migdomuosius vaistus A ir B, papildomo miego trukmė X ir Y (valandomis).

i	X_i	Y_i	i	X_i	Y_i
1	$1,9$	$0,7$	6	$4,4$	$3,4$
2	$0,8$	$-1,6$	7	$5,5$	$2,7$
3	$1,1$	$-0,2$	8	$1,6$	$0,8$
4	$0,1$	$-1,2$	9	$4,6$	$0,0$
5	$-0,1$	$-0,1$	10	$3,4$	$2,0$

Patikrinkite hipotezę, kad vaistų poveikis vienodas, tarę, kad buvo stebimas normalusis atsitiktinis vektorius.

I.3.107. (**I.2.202** pratimo tęsinys). **I.2.202** pratimo sąlygomis tarę, kad parametras $\eta = 10$, patikrinkite hipotezę $H : \lambda \leq 1$, kai alternatyva yra $\bar{H} : \lambda > 1$.

I.3.108. (**I.2.205** pratimo tęsinys). **I.2.205** pratimo sąlygomis tarę, kad parametras $\eta = 10$, patikrinkite hipotezę $H : \lambda = \lambda_0 = 0,001$, kai alternatyva yra $\bar{H} : \lambda \neq 0,001$. $(\alpha = 0,01)$.

I.3.109. (**I.2.201** pratimo tęsinys). **I.2.201** pratimo sąlygomis reikšmingumo lygmens $\alpha = 0,05$ kriterijumi patikrinkite hipotezę $H : \lambda = \lambda_0 = 1$, kai alternatyva yra $\bar{H} : \lambda \neq 1$.

I.3.110. Tegu $X_1, X_2, ..., X_7$ yra firmoje užregistruotų klientų skambučių skaičiai per 7 savaitės dienas. Tarę, kad a. d. $X_1, ..., X_7$ yra nepriklausomi ir turi Puasono skirstinius

$X_i \sim \mathcal{P}(\lambda_i)$, a) patikrinkite hipotezę $H : \lambda_1 = ... = \lambda_7$ remdamiesi a. d. $X_1, ..., X_7$ realizacija: 52; 65; 60; 71; 75; 43; 40. b) Matome, kad savaitgalį skambučių skaičius yra mažesnis. Patikrinkite hipotezę $H : \lambda_1 = ... = \lambda_5$, kad skambučių intensyvumas darbo dienomis yra vienodas.

I.3.111. (**I.2.203** pratimo tęsinys). **I.2.203** pratimo sąlygomis patikrinkite hipotezę, kad II rūšies gaminių dalis neviršija 0,25.

I.3.112. Per pirmą valandą skaitiklis užregistravo 150 tam tikrų kosminių dalelių, per tolesnes dvi valandas – 250 dalelių. Patikrinkite hipotezę, kad dalelių srauto intensyvumas nepakito.

I.3.113. Du nepriklausomi a. d., pasiskirstę pagal Puasono dėsnį, įgijo atitinkamai reikšmes 75 ir 200. Patikrinkite hipotezę $H : \lambda_1 = \lambda_2/2$, kai alternatyva $\bar{H} : \lambda_1 < \lambda_2/2$.

I.3.114. Per pirmąją dieną skaitiklis užregistravo 20 026 puasoninio srauto impulsus, o per antrąją dieną – 19 580 impulsų. Ar yra pagrindo teigti, kad impulsų srauto intensyvumas sumažėjo?

I.3.115. Ar galima teigti, kad dviejose nepriklausomose Bernulio bandymų schemose įvykio A tikimybė vienoda, jeigu atlikus $n_1 = n_2 = 5000$ bandymų įvykis A įvyko 2 602 ir 2 398 kartus?

I.3.116. Patikrinus 5 vienodo didumo $n = 200$ gaminių partijas, jose buvo surasta atitinkamai 15; 10; 6; 12; 4 defektiniai gaminiai. Tegu defektinių gaminių skaičius j-oje partijoje turi binominį skirstinį $B(n, p_i), i = 1, ..., 5$. Patikrinkite hipotezę $H : p_1 = ... = p_5$.

I.3.117. (**I.2.211** pratimo tęsinys). **I.2.211** pratimo sąlygomis patikrinkite hipotezę, kad prapuolimo kampas turi tolygųjį pasiskirstymą.

I.3.118. (**I.2.212** pratimo tęsinys). **I.2.212** pratimo sąlygomis patikrinkite hipotezę, kad susirgimai leukemija tolygiai pasiskirstę per metus.

I.3.119. Lentelėje pateikta smėlio grūdelių orientacija plokštumoje (žr. [14]).

Kampas	Kiekis	Kampas	Kiekis	Kampas	Kiekis
$0°-$	244	$60°-$	326	$120°-$	322
$10°-$	262	$70°-$	340	$130°-$	295
$20°-$	246	$80°-$	371	$140°-$	230
$30°-$	290	$90°-$	401	$150°-$	256
$40°-$	284	$100°-$	382	$160°-$	263
$50°-$	314	$110°-$	332	$170°-$	281

Kampai sugrupuoti į ilgio $10°$ intervalus (nurodoma grupavimo intervalo pradžia). Gretimuose stulpeliuose nurodomi smėlio grūdelių, kurių orientacija patenka į atitinkamus intervalus, skaičiai.

Padvigubinę kampus, perveskite duomenis į intervalą $[0° - 360°]$. Patikrinkite kampų skirstinio tolygumo hipotezę.

I.3.7. Sprendimai, nurodymai, atsakymai

I.3.1 skyrelis

I.3.1. Pagal didumo $n = 6$ paprastąją imtį $(X_1, ..., X_6)^T$ gautą stebint a. d. $X \sim B(1, p)$, tikrinama hipotezė $H : p = 0,2$ kai alternatyva yra $\bar{H} : p > 0,2$. Kriterijaus kritinė sritis $K = \{(X_1, ..., X_6) : S = X_1 + ... + X_6 \geq 3\}$. Reikšmingumo lygmuo $\alpha = \mathbf{P}\{S \geq 3 | p = 0,2\} = 0,0989$. Kriterijaus galia: 0,4557; 0,8208; 0,9830; 1,000.

I.3.2. Tegu bakterijų skaičius tūrio vienete turi Puasono skirstinį $\mathcal{P}(\lambda)$. Tada bakterijų skaičius N tūrio V kolboje turi Puasono skirstinį $\mathcal{P}(V\lambda.)$. Tikimybė, kad vanduo kolboje susidrums, $p = \mathbf{P}\{N > 0\} = 1 - e^{-V\lambda}$. Patikrinus n kolbų susidrumstusių skaičius $X \sim B(n, p)$. Remiantis a. d. X tikrinama hipotezė $H : p \leq p_0 = 1 - e^{-Vm_o}$, kai alternatyva yra $\bar{H} : p > p_0$. Kritinė sritis $K = \{X : X > t\}$.

a) $p_0 = 1 - e^{-1}, n = 10$; kritinė sritis $K = \{X : X > 7\}$; pirmosios rūšies klaidos tikimybė

$$\alpha \leq \mathbf{P}\{X > 7 | p_0\} = \sum_{m=8}^{10} C_{10}^m p_0^m (1 - p_0)^{10-m} = 0,2247.$$

b) $p_0 = 1 - e^{-2}, n = 8$; kritinė sritis $K = \{X : X > 7\}$; pirmosios rūšies klaidos tikimybė

$$\alpha \leq \mathbf{P}\{X > 7 | p_0\} = p_0^8 = 0,3126.$$

I.3.3. Pagal didumo $n = 250$ paprastąją imtį, gautą stebint Bernulio a. d. $X \sim B(1, p)$, tikrinama hipotezė $H : p \leq p_0$, kai alternatyva $\bar{H} : p > p_0$. Kriterijaus priėmimo sritis $S = \sum_i X_i \leq 25$. Reikia rasti tokį p_0, kad $\mathbf{P}\{S \leq 25 | p = p_0\} = 0,95$. Naudodami SAS, MS Excel, R arba SPSS programą randame $p_0 \approx 7,39\,\%$.

I.3.4. Tegu defektinių gaminių skaičius partijoje yra M. Tikrinama hipotezė $H : M/N \leq p_0$, kai alternatyva yra $\bar{H} : M/N > p_0$ remiantis a. d. $X \sim H(N, M, n)$. Hipotezės priėmimo sritis $A = \{X : X \leq d\}$ ir kritinė sritis $K = \{X : X > d\}$.

I.3.5. Kadangi partijos didelės, o tikrinama, kaip matysime, nedidelė partijos dalis, tai a. d. X skirstinį galima aproksimuoti binominiu $B(n, p), p = M/N$. Tada atrankinės kontrolės plano charakteristikoms n ir d rasti turime nelygybių sistemą

$$\mathbf{P}\{X \leq d | p = p_0 = 0,05\} = \sum_{m=0}^{d} C_n^m p_0^m (1 - p_0)^{n-m} \geq 0,9;$$

$$\mathbf{P}\{X \leq d | p = p_1 = 0,1\} = \sum_{m=0}^{d} C_n^m p_1^m (1 - p_1)^{n-m} \leq 0,05.$$

Gauname, kad minimalus imties didumas $n = 251$; partija priimama, kai $X \leq 17$.

I.3.6. Turime didumo $n = 500$ paprastąją imtį, gautą stebint Bernulio a. d. $X \sim B(1, p)$. Tada defektinių gaminių skaičius $S = X_1 + ... + X_n \sim B(n, p)$. Tikrinama hipotezė $H : p \leq p_0 = 0,05$, kai alternatyva $\bar{H} : p > p_0$. Taikomo kriterijaus kritinė sritis $K = \{S : S \geq k\}$.

I.3.7. a) $k = 37$; 0,6527; 0,9726; 0,9995; b) $k = 7$; 0,0438; 0,1221; 0,2467.

I.3.8. a) Parametro μ pasikliovimo lygmens $Q = 1 - \alpha$ dešininis vienpusis pasikliovimo intervalas yra $(\underline{\mu}, \infty)$, kai $\underline{\mu} = \bar{X} - z_\alpha \sigma/\sqrt{n}$. Taigi hipotezės H priėmimo sritis $A_1 = \{\bar{X} : \underline{\mu} < \mu_0 < \infty\} \;\Leftrightarrow\; \{\bar{X} : Z = \sqrt{n}(\bar{X} - \mu_0)/\sigma < z_\alpha\}$ (žr. [2], 4.5 skyrelį). Kritinė sritis $K_1 = \{\bar{X} : Z > z_\alpha\}$. P reikšmių terminais kriterijus formuluojamas taip: hipotezė atmetama, kai $pv = \mathbf{P}\{Z > z|\mu = \mu_0\} = 1 - \Phi(z) \le \alpha$; čia z yra statistikos Z realizacija (žr. [2], 4.1.2 skyrelį).

Analogiškai, kai alternatyva $\bar{H}_2 : \mu \le \mu_0$, gauname kritinę sritį $K_2 = \{\bar{X} : \mu_0 > \bar{\mu} = \bar{X} + z_\alpha \sigma/\sqrt{n}\} \;\Leftrightarrow\; \{\bar{X} : Z < -z_\alpha\}$. P reikšmių terminais hipotezė atmetama, kai $pv = \Phi(z) < \alpha$.

b) Parametro μ pasikliovimo lygmens $Q = 1 - \alpha$ pasikliovimo intervalas yra $(\underline{\mu}, \overline{\mu}) = (\bar{X} - z_{\alpha/2}\sigma/\sqrt{n}, \bar{X} + z_{\alpha/2}\sigma/\sqrt{n})$. Jį atitinkanti hipotezės H priėmimo sritis $A_3 = \{\bar{X} : \underline{\mu} < \mu_0 < \bar{\mu}\}$. Kritinė sritis $K_3 = \{\bar{X} : |Z| > -z_{\alpha'2}\}$. P reikšmių terminais hipotezė atmetama, kai

$$pv = 2\min(\Phi(z), 1 - \Phi(z)) = 2(1 - \Phi(|z|)) < \alpha.$$

I.3.2 skyrelis

I.3.9. Remiantis Neimano ir Pirsono lema (žr. [2], 4.2.1 skyrelį) hipotezė H atmetama, kai

$$\frac{L(\theta_1; \mathbf{X})}{L(\theta_0; \mathbf{X})} = \frac{\theta_1^n}{\theta_0^n} e^{-(\theta_1 - \theta_0)S_n} > c \;\Leftrightarrow\; -(\theta_1 - \theta_0)S_n > d;$$

čia $S_n = X_1 + ... + X_n \sim G(\theta_0, n); 2\theta_0 S_n \sim \chi^2(2n)$, kai hipotezė H teisinga; konstanta d randama iš sąlygos

$$\mathbf{P}_{\theta_0}\{-(\theta_1 - \theta_0)S_n > d\} = \alpha.$$

Tarkime, kad $\theta_1 > \theta_0$.

$$\mathbf{P}_{\theta_0}\{-(\theta_1 - \theta_0)S_n > d\} = \alpha \quad\Rightarrow\quad \mathbf{P}_{\theta_0}\{(\theta_1 - \theta_0)S_n < -d\} = \alpha \quad\Rightarrow$$

$$\mathbf{P}_{\theta_0}\{2\theta_0 S_n < -\frac{d 2\theta_0}{\theta_1 - \theta_0}\} = \alpha \quad\Rightarrow\quad -\frac{d 2\theta_0}{\theta_1 - \theta_0} = \chi^2_{1-\alpha}(2n) \quad\Rightarrow$$

$$d = -\chi^2_{1-\alpha}(2n)\frac{\theta_1 - \theta_0}{2\theta_0};$$

$$-(\theta_1 - \theta_0)S_n > d \quad\Rightarrow\quad -(\theta_1 - \theta_0)S_n > -\chi^2_{1-\alpha}(2n)\frac{\theta_1 - \theta_0}{2\theta_0}$$

$$\Rightarrow\quad S_n < \chi^2_{1-\alpha}(2n)/(2\theta_0).$$

Jeigu $\theta_1 > \theta_0$, tai H atmetame, kai $S_n < \chi^2_{1-\alpha}(2n)/(2\theta_0)$.

Kriterijaus galia:

$$\beta(\theta_1) = \mathbf{P}_{\theta_1}\left(S_n < \chi^2_{1-\alpha}(2n)/(2\theta_0)\right) = \mathbf{P}_{\theta_1}\left(2\theta_1 S_n < \frac{\chi^2_{1-\alpha}(2n)2\theta_1}{2\theta_0}\right) =$$

$$= \mathbf{P}\left(\chi^2_{2n} < \frac{\chi^2_{1-\alpha}(2n)\theta_1}{\theta_0}\right)$$

Analogiškai, jeigu $\theta_1 < \theta_0$, tai H atmetame, kai $S_n > \chi^2_\alpha(2n)/(2\theta_0)$; kriterijaus galia $\beta(\theta_1) = \mathbf{P}\{\chi^2_{2n} > (\theta_1/\theta_0)\chi^2_\alpha(2n)\}$.

I.3.10. Tegu kriterijaus reikšmingumo lygmuo α tenkina nelygybę $\alpha < 1/2 - \mathrm{arctg}(1/2)/\pi \approx 0,352$. Remiantis Neimano ir Pirsono lema hipotezė atmetama, kai

$$\frac{f(X;1)}{f(X;0)} = \frac{1+X^2}{1+(X-1)^2} > c \quad \Leftrightarrow \quad (c-1)X^2 - 2cX + 2c - 1 < 0.$$

Taigi hipotezė atmetama, kai $x_1 < X < x_2$;

$$x_1 = [c - \sqrt{c-(c-1)^2}]/(c-1), \quad x_2 = [c + \sqrt{c-(c-1)^2}]/(c-1);$$

konstanta c randama iš sąlygos $(\mathrm{arctg}(x_2) - \mathrm{arctg}(x_1))/\pi = \alpha$.

I.3.11. a) Hipotezė atmetama, kai $|X| > z_{\alpha/2}$, jeigu $\alpha < 0,0455$; hipotezės atmetimo sritis: $|X| < x_1$ arba $|X| > x_2$, $x_1 = 1 - \sqrt{1+c}, x_2 = 1 + \sqrt{1+c}$; konstanta c randama iš sąlygos $\alpha = 2[1 - \Phi(x_2)] + 2\Phi(x_1) - 1$, jeigu $\alpha > 0,0455$.

b) Hipotezė atmetama, kai $\delta - c < |X| < \delta$; konstanta c randama iš sąlygos $2[\Phi(\delta) - \Phi(\delta - c)] = \alpha$; jeigu toks $c < \delta$ neegzistuoja, tai H atmetama, kai $|X| < \delta$.

I.3.12. Imties $\mathbf{X} = (X_1, ..., X_5)^T$ tankis, kai teisinga hipotezė, yra

$$\varphi(\boldsymbol{x}) = (1/\sqrt{2\pi}\sigma)^5 \exp\{-(x_1^2 + ... + x_5^2)/(2\sigma^2)\};$$

esant teisingai alternatyvai skirstinys yra tolygus vienetiniame penkiamačiame kube $-1/2 < x_1, ..., x_5 < 1/2$. Remiantis Neimano ir Pirsono lema, H priimama, kai

$$\max(|X_1|, ..., |X_5|) > 1/2,$$

arba kai a. v. $\mathbf{X}$ patenka į penkiamatę sferą su centru koordinačių pradžioje: $X_1^2 + ... + X_5^2 < r^2$. Tikimybė $\mathbf{P}\{\max(|X_1|, ..., |X_5|) > 1/2\} = 1 - (2\Phi(1/0,316) - 1)^5 = 0,0077$. Iš sąlygos $\mathbf{P}\{X_1^2 + ... + X_5^2 < r^2\} = \mathbf{P}\{\chi^2_5 < r^2/0,025\} = 0,9 - 0,0077 = 0,8923$ randame $r^2 = 0,2258$, t. y. penkiamatė sfera patenka į kubo vidų. Kadangi $X_1^2 + ... + X_5^2 = 0,2549$, tai H atmetama.

I.3.13. Remiantis Neimano ir Pirsono lema H atmetama, kai $\bar{X} > c$. Abiejų klaidų tikimybės vienodos, kai $c = 1/2$. Tada imties didumui rasti turime nelygybę

$$\mathbf{P}\{\bar{X} > 1/2 | \mu = 0\} = \mathbf{P}\{\sqrt{n}\bar{X} > \sqrt{n}/2 | \mu = 0\} = 1 - \Phi(\sqrt{n}/2) \leq 0,01.$$

Išsprendę gauname, kad $n \geq 22$.

I.3.14. TG kriterijus atmeta hipotezę, kai $X > m$, ir atmeta su tikimybe γ, kai $X = m$. Konstantos m ir γ randamos iš sąlygos

$$\mathbf{P}\{X > m | \lambda_0\} + \gamma \mathbf{P}\{X = m | \lambda_0\} = 0,1.$$

Pirmuoju atveju gauname $m = 0, \gamma = 0,00535$. Kriterijaus galia

$$\beta(\lambda_1) = \mathbf{P}\{X > 0 | \lambda_1\} + \gamma \mathbf{P}\{X = 0 | \gamma_1\} = 0,1856.$$

Analogiškai kitais atvejais kriterijaus galia $0,3523; 0,9074; 0,3655$.

I.3.15. Neatsižvelgiant į randomizaciją, kriterijus yra toks: $\varphi_1(\boldsymbol{x}) = 1$, kai $f_1(\boldsymbol{x}) > c_1 f_2(\boldsymbol{x})$; $\varphi_2(\boldsymbol{x}) = 1$, kai $f_2(\boldsymbol{x}) > c_2 f_1(\boldsymbol{x})$; $\varphi_0(\boldsymbol{x}) = 1 - \varphi_1(\boldsymbol{x}) - \varphi_2(\boldsymbol{x})$. Konstanta c_1 randama iš sąlygos $\alpha_{12} = \alpha$, o konstanta c_2 – iš sąlygos $\alpha_{21} = \beta$, jeigu sprendiniai c_1 ir c_2 tenkina sąlygą $c_1 c_2 > 1$. Priešingu atveju sprendinys neegzistuoja (tokiu atveju galima minimizuoti, pvz., sumą $\alpha_{01} + \alpha_{10}$).

I.3.16. Sprendinys yra tokio pat pavidalo kaip ir **I.3.15** pratime. Konstantos c_1, c_2 randamos iš sąlygų $\beta_{12} = b_1$, $\beta_{21} = b_2$, jei tik $c_1 c_2 > 1$ (žr. [4], 6.1.3 skyrelį).

I.3.18. a) H_0 atmetame, kai $X > \theta_0(1-\sqrt{\alpha})$; b) H_0 atmetame, kai $X < [-(1-\theta_0)+\sqrt{(1-\theta_0)^2 + \alpha(2\theta_0 - 1)}]/(\theta_0 - 1/2)$ ir $\theta_0 \neq 1/2$; atmetama, kai $X < \alpha$, jeigu $\theta_0 = 1/2$; c) hipotezė atmetama, kai $X < \sqrt{\alpha}/2$ arba kai $X > 1 - \sqrt{\alpha}/2$.

I.3.19. a) H_0 atmetama, kai $X_{(1)} > \theta_0 - (\ln \alpha)/n$; b) hipotezę atmetame, kai $X_{(1)} > \theta_0 \alpha^{-1/n}$, esant alternatyvai $\theta_1 > \theta_0$; hipotezę atmetame su tikimybe 1, kai $X_{(1)} < \theta_0$, ir atmetame su tikimybe α, kai $X_{(1)} \geq \theta_0$, esant alternatyvai $\theta_1 < \theta_0$.

I.3.20. Remiantis TG kriterijumi H atmetama su tikimybe 1, kai $S = X_1 + ... + X_n \geq 1$, ir atmetama su tikimybe $\alpha = 0,01$, kai $S = 0$. Abiejų rūšių klaidos neviršija 0,01, kai $n \geq 458$.

I.3.3 skyrelis

I.3.21. a) Kadangi tikėtinumo funkcija

$$L(\mu) = W(\mathbf{X}) e^{\bar{X}\eta(\mu) - b(\mu)}, \quad \eta(\mu) = \frac{n\mu}{\sigma^2}, \quad b(\mu) = \frac{n\mu^2}{2\sigma^2}$$

priklauso vienparametrei eksponentinio tipo skirstinių šeimai, $T = \bar{X}$ yra pakankamoji statistika, tai egzistuoja TG kriterijus (žr. [2], 4.3.2 pastabą), kai alternatyva $\bar{H}_1$, kurio kritinė sritis

$$K_1 = \{\mathbf{X} : \bar{X} > c\} \;\Leftrightarrow\; \{\mathbf{X} : Z = \sqrt{n}\frac{\bar{X} - \mu_0}{\sigma_0} > z_\alpha\}.$$

Kriterijaus galia

$$\beta_1(\mu) = \mathbf{P}_\mu\{\sqrt{n}\frac{\bar{X} - \mu_0}{\sigma_0} > z_\alpha\} = \mathbf{P}_\mu\{\sqrt{n}\frac{\bar{X} - \mu}{\sigma_0} > z_\alpha - \lambda\} = \Phi(\lambda - z_\alpha),$$

čia $\lambda = \sqrt{n}(\mu - \mu_0)/\sigma_0$, $\beta_1(\mu) \to 1$, kai $\lambda \to \infty$.

Analogiškai, kai alternatyva yra $\bar{H}_2$, TG kriterijaus kritinė sritis yra

$$K_2 = \{\mathbf{X} : Z = \sqrt{n}\frac{\bar{X} - \mu_0}{\sigma_0} < -z_\alpha\}.$$

Šio kriterijaus galia

$$\beta_2(\mu) = \mathbf{P}_\mu\{\sqrt{n}\frac{\bar{X} - \mu_0}{\sigma_0} < -z_\alpha\} = \Phi(-\lambda - z_\alpha)$$

monotoniškai artėja prie 1, kai $\mu - \mu_0 \to -\infty$.

b) Kai alternatyva $\bar{H}_3 : \mu \neq \mu_0$ yra dvipusė, TG kriterijus neegzistuoja. Remiantis [2], 4.3.2 teorema, TGN kriterijaus kritinė sritis

$$K_3 = \{\mathbf{X} : |Z| = \sqrt{n}\frac{|\bar{X} - \mu_0|}{\sigma_0} > z_{\alpha/2}\},$$

o galios funkcija

$$\beta_3(\mu) = \Phi(-\lambda - z_{\alpha/2}) + \Phi(\lambda - z_{\alpha/2})$$

artėja prie 1, kai $|\mu - \mu_0| \to \infty$ (arba $|\lambda| \to \infty$).

I.3.22. Patekti į kritinę sritį K_1 yra ekvivalentu nelygybei $\mu_0 < \underline{\mu}$, o patekti į K_2 – nelygybei $\mu_0 > \overline{\mu}$; čia $\underline{\mu}$ ir $\overline{\mu}$ yra parametro μ pasikliovimo intervalo, kai pasikliovimo lygmuo $Q = 1 - 2\alpha$, viršutinis ir apatinis rėžiai. Patekti į kritinę sritį K_3 ekvivalentu nelygybėms $\mu_0 < \underline{\mu}$ arba $\mu_0 > \overline{\mu}$, kai intervalo pasikliovimo lygmuo $Q = 1 - \alpha$ (žr. [2], 4.5 skyrelį).

I.3.23. Tarkime, z yra statistikos $Z = \sqrt{n}(\bar{X} - \mu_0)/\sigma$ realizacija. Tada P reikšmių terminais kriterijai, kurių kritinės sritys K_1, K_2, K_3, formuluojami taip: hipotezė H atmetama, kai atitinkamai (žr. [2], 4.1.2 skyrelį)

$$pv = 1 - \Phi(z) < \alpha, \quad pv = \Phi(z) < \alpha, \quad pv = 2(1 - \Phi(|z|)) < \alpha.$$

I.3.24. a) Imties didumui n rasti gauname nelygybę

$$\beta_1(\mu') = \Phi(\sqrt{n}\frac{\mu_1 - \mu_0}{\sigma_0} - z_\alpha) \geq 1 - \alpha'.$$

Iš čia gauname, kad imties didumas n tenkina nelygybę

$$n \geq \frac{\sigma_0^2}{(\mu_1 - \mu_0)^2}(z_{\alpha'} + z_\alpha)^2.$$

Matome, kad imties didumas tiesiog proporcingas dispersijai ir atvirkščiai proporcingas atstumo $\mu_1 - \mu_0$ kvadratui.

b) Įrašę z_α ir $z_{\alpha'}$ iš 2 priedo 2P2 lentelės, gauname

$$n \geq \frac{\sigma^2}{(\mu_1 - \mu_0)^2}(z_\alpha + z_{\alpha'})^2 = \frac{4}{0,5^2}(z_{0,05} + z_{0,01})^2 = 252,33.$$

Taigi imties didumas turi būti ne mažesnis už 253.

I.3.25. a) Imties $\mathbf{X} = (X_1, ..., X_n)^T$ skirstinys priklauso vienparametrei eksponentinio tipo skirstinių šeimai, o tankis

$$f(\boldsymbol{x},\ \theta) = h(\boldsymbol{x}) \exp\{\eta(\theta)T(\boldsymbol{x}) - b(\theta)\};$$

čia

$$\eta(\theta) = -\frac{1}{2\sigma^2}, \quad T = T(\boldsymbol{x}) = \sum_i (x_i - \mu_0)^2, \quad b(\theta) = n \ln \sigma, \quad h(\boldsymbol{x}) = (\sqrt{2\pi})^{-n}.$$

Remiantis [2], 4.3.2 pastaba, egzistuoja TG kriterijai vienpusių alternatyvų $\bar{H}_1, \bar{H}_2$ atvejais. Jų kritinės sritys

$$K_1 = \{\mathbf{X} : \frac{T(\mathbf{X})}{\sigma_0^2} > \chi^2_\alpha(n)\}, \quad K_2 = \{\mathbf{X} : \frac{T(\mathbf{X})}{\sigma_0^2} < \chi^2_{1-\alpha}(n)\}.$$

Šių kriterijų galios funkcijos $\beta_1(\sigma^2), \beta_2(\sigma^2)$ ir išreiškiamos χ^2 skirstinio pasiskirstymo funkcija

$$\beta_1(\sigma^2) = \mathbf{P}_{\sigma^2}\{\frac{T(\mathbf{X})}{\sigma_0^2} > \chi^2_\alpha(n)\} = \mathbf{P}\{\chi^2_n > \frac{\sigma_0^2}{\sigma^2}\chi^2_\alpha(n)\},$$

$$\beta_2(\sigma^2) = \mathbf{P}\{\chi^2_n < \frac{\sigma_0^2}{\sigma^2}\chi^2_{1-\alpha}(n)\}.$$

b) Kai alternatyva $\bar{H}_3 : \sigma^2 \neq \sigma_0^2$ yra dvipusė, TG kriterijus neegzistuoja. Tačiau, naudojantis [2], 4.3.2 teorema, galima rasti TGN kriterijų, kurio kritinė sritis

$$K_3 = \{\mathbf{X} : \frac{T(\mathbf{X})}{\sigma_0^2} < c_1 \quad \text{arba} \quad \frac{T(\mathbf{X})}{\sigma_0^2} > c_2\};$$

čia konstantos c_1 ir c_2 randamos iš lygčių sistemos

$$\mathbf{P}_{\sigma_0^2}\{c_1 < T/\sigma_0^2 < c_2\} = 1 - \alpha,$$

$$\mathbf{E}_{\sigma_0^2}[(T/\sigma_0^2)\mathbf{1}_{(c_1,c_2)}(T/\sigma_0^2)] = (1-\alpha)\mathbf{E}_{\sigma_0^2}(T/\sigma_0^2) = (1-\alpha)n.$$

Kadangi $\mathbf{E}_{\sigma_0^2}[(T/\sigma_0^2)\mathbf{1}_{(c_1,c_2)}(T/\sigma_0^2)] = n\mathbf{P}\{c_1 < \chi^2_{n-2} < c_2\}$, tai konstantoms c_1 ir c_2 rasti gauname lygčių sistemą

$$\begin{cases} \mathbf{P}\{c_1 < \chi^2_n < c_2\} = 1 - \alpha, \\ \mathbf{P}\{c_1 < \chi^2_{n+2} < c_2\} = 1 - \alpha. \end{cases}$$

Kriterijaus galios funkcija

$$\beta_3(\sigma^2) = 1 - \mathbf{P}\{\frac{\sigma_0^2}{\sigma^2}c_1 < \chi^2_n < \frac{\sigma_0^2}{\sigma^2}c_2\}.$$

c) Simetrinio kriterijaus kritinė sritis

$$K_3^* = \{\mathbf{X} : \frac{T(\mathbf{X})}{\sigma_0^2} < \chi^2_{1-\alpha/2}(n), \quad \text{arba} \quad \frac{T(\mathbf{X})}{\sigma_0^2} > \chi^2_{\alpha/2}(n)\}.$$

I.3.26. Sprendžiame **I.3.25** pratime pateiktą lygčių sistemą, kai $n = 10, 20, 50$. TGN kriterijaus galios $\beta_3(\sigma^2)$ ir simetriško kriterijaus galios $\tilde{\beta}_3(\sigma^2)$ išraiškas žr. **I.3.25** pratime.

I.3.27. a) Kritines sritis K_1 ir K_2 atitinka kriterijai: hipotezė atmetama, kai atitinkamai $\sigma_0^2 < \underline{\sigma^2}$ ir $\sigma_0^2 > \overline{\sigma^2}$; čia $\underline{\sigma^2}$ ir $\overline{\sigma^2}$ yra lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalo apatinis ir viršutinis rėžiai. Kritinę sritį K_3^* atitinka kriterijus: hipotezė atmetama, kai teisingos nelygybės $\sigma_0^2 < \underline{\sigma^2}$ arba $\sigma_0^2 > \overline{\sigma^2}$; čia $(\underline{\sigma^2}, \overline{\sigma^2})$ yra lygmens $Q = 1 - \alpha$ pasikliovimo intervalas.

b) Tegu t yra statistikos T/σ_0^2 realizacija, o $F(x|\nu)$ yra a. d. χ^2_ν pasiskirstymo funkcija. Tada kriterijus K_1, K_2, K_3^* P reikšmių terminais atitinka tokios taisyklės: hipotezė atmetama, kai atitinkamai

$$pv = 1 - F(y|n-1) \leq \alpha, \quad pv = F(y|n-1) \leq \alpha,$$

$$pv = 2\min(1 - F(y|n-1), F(y|n-1)) \leq \alpha.$$

I.3.28. Remdamiesi **I.3.25** pratime pateikta galios funkcija, imties didumui rasti turime nelygybę

$$\beta_1(1,5\sigma_0^2) = \mathbf{P}\{\chi_n^2 > \frac{2}{3}\chi^2_{0,05}(n)\} \geq 0,9 \quad \Leftrightarrow \quad n \geq 104.$$

I.3.29. Imties $\mathbf{X}$ tankis

$$f(\mathbf{X},\ \sigma) = (\prod_i X_i)\exp\{\eta(\sigma)T(\mathbf{X}) - B(\sigma)\},$$

čia

$$\eta(\sigma) = -\frac{1}{2\sigma^2}, \quad T(\mathbf{X}) = \sum_i X_i^2, \quad B(\sigma) = 2n\ln\sigma.$$

Kadangi $T(\mathbf{X})/\sigma_0^2 \sim \chi^2(2n)$, kai hipotezė teisinga, tai palyginę su **I.3.25** pratimu darome išvadą, kad kriterijai bus nusakyti kritinėmis sritimis K_1, K_2, K_3, K_3^*, kuriose n reikia pakeisti į $2n$.

I.3.30. Imties $\mathbf{X}$ tankis

$$f(\mathbf{X}, \sigma) = (2/\pi)^{n/2}(\prod_i X_i^2)\exp\{\eta(\sigma)T(\mathbf{X}) - B(\sigma)\},$$

čia

$$\eta(\sigma) = -\frac{1}{2\sigma^2}, \quad T(\mathbf{X}) = \sum_i X_i^2, \quad B(\sigma) = 2n\ln\sigma.$$

Kadangi $T(\mathbf{X})/\sigma_0^2 \sim \chi^2(3n)$, kai $\sigma = \sigma_0$, tai palyginę su **I.3.25** pratimu darome išvadą, kad kriterijai bus nusakyti kritinėmis sritimis K_1, K_2, K_3, K_3^*, kuriose n reikia pakeisti į $3n$.

I.3.31. a) Imties $\mathbf{X}$ tankio funkcija

$$f(\mathbf{X}, \lambda) = \exp\{\eta(\lambda)T - B(\lambda)\}h(\mathbf{X}),$$

čia

$$\eta(\lambda) = -\lambda, \quad T = \sum_i X_i, \quad B(\lambda) = -\eta\ln\lambda,$$

priklauso vienparametrei eksponentinio tipo skirstinių šeimai; statistika $V = 2\lambda_0 T \sim \chi^2(2\eta)$, kai $\lambda = \lambda_0$. Remiantis [2], 4.3.2 pastaba, egzistuoja TG kriterijai, esant vienpusėms alternatyvoms $\bar{H}_1$ ir $\bar{H}_2$. Kritinės sritys yra tokios:

$$K_1 = \{\mathbf{X} : 2\lambda_0 T < \chi^2_{1-\alpha}(2\eta)\}, \quad K_2 = \{\mathbf{X} : 2\lambda_0 T > \chi^2_\alpha(2\eta)\}.$$

Pažymėkime t statistikos T realizaciją ir tegu $F(x|\nu)$ yra χ^2 skirstinio su ν laisvės laipsniais pasiskirstymo funkcija. Tada kriterijus galima suformuluoti P reikšmių terminais: hipotezė atmetama, kai atitinkamai teisingos nelygybės

$$pv = F(2\lambda_0 t|2\eta) \le \alpha, \quad pv = 1 - F(2\lambda_0 t|2\eta) \le \alpha.$$

b) Kai alternatyva $\bar{H}_3 : \lambda \ne \lambda_0$ yra dvipusė, remiantis [2], 4.3.2 teorema egzistuoja TGN kriterijus, kurio kritinė sritis

$$K_3 = \{\mathbf{X} : 2\lambda_0 T < c_1 \text{ arba } 2\lambda_0 T > c_2\}.$$

Analogiškai **I.3.25** pratimui konstantos c_1 ir c_2 randamos iš tokios lygčių sistemos:

$$\begin{cases} \mathbf{P}\{c_1 < \chi^2_{2\eta} < c_2\} = 1 - \alpha, \\ \mathbf{P}\{c_1 < \chi^2_{2\eta+2} < c_2\} = 1 - \alpha. \end{cases}$$

c) Simetriško kriterijaus kritinė sritis

$$K_3^* = \{\mathbf{X} : 2\lambda_0 T < \chi^2_{1-\alpha/2}(2\eta) \quad \text{arba} \quad 2\lambda_0 T > \chi^2_{\alpha/2}(2\eta)\},$$

arba P reikšmių terminais hipotezė H_3 atmetama, kai

$$pv = 2\min(F(2\lambda_0 t|2\eta), 1 - F(2\lambda_0 t|2\eta)) \le \alpha.$$

I.3.32. Imties tankis (skaičiuojančiojo mato atžvilgiu)

$$f(\mathbf{X},\ \lambda) = \exp\{\eta(\lambda)T - B(\lambda)\}h(\mathbf{X}),$$

čia

$$\eta(\lambda) = \ln\lambda, \quad T = \sum_i X_i, \quad B(\lambda) = n\lambda,$$

priklauso vienparametrei eksponentinio tipo skirstinių šeimai.

Remiantis [2], 4.3.2 pastaba, egzistuoja TG kriterijus, kai alternatyva yra $\bar{H}_1$, tikrinti. Kriterijus nusakomas taip:

$$\varphi(T) = \begin{cases} 1, & \text{kai } T > k, \\ \gamma, & \text{kai } T = k, \\ 0, & \text{kai } T < k, \end{cases}$$

čia konstantos k ir γ randamos iš sąlygos

$$\mathbf{E}_{\lambda_0}(\varphi(T)) = \mathbf{P}_{\lambda_0}\{T > k\} + \gamma\mathbf{P}_{\lambda_0}\{T = k\}.$$

Kadangi statistika $T \sim \mathcal{P}(n\lambda_0)$, kai $\lambda = \lambda_0$, tai ši sąlyga reiškia, kad

$$\sum_{m=k+1}^{\infty} \frac{(n\lambda_0)^m}{m!}\mathrm{e}^{-\mathrm{n}\lambda_0} + \gamma\frac{(\mathrm{n}\lambda_0)^{\mathrm{k}}}{\mathrm{k}!}\mathrm{e}^{-\mathrm{n}\lambda_0} = \alpha.$$

Statistikos T skirstinys yra diskretusis, todėl daugumai λ_0 reikšmių TG kriterijus bus randomizuotas, t. y. $\gamma \ne 0;\ 1$. Norint, kad reikšmingumo lygmuo būtų tiksliai lygus α, tenka įtraukti γ dydžio taško k „dalį“ ir taip papildyti reikšmingumo lygmenį iki α.

Kaip minėjome (žr. [2], 4.2.3 pastabą), paprastai nereikalaujama, kad reikšmingumo lygmuo būtų tiksliai lygus α. Todėl dažniau naudojami nerandomizuoti kriterijai, gaunami šiek tiek sumažinus reikšmingumo lygmenį. Tiksliau, parenkamas toks mažiausias sveikasis skaičius k', kad

$$\mathbf{P}_{\lambda_0}(T \geq k') = \sum_{m=k'}^{\infty} \frac{(n\lambda_0)^m}{m!} \mathrm{e}^{-\mathrm{n}\lambda_0} = 1 - \mathbf{P}\{\chi^2_{\mathrm{2k'}} > 2\mathrm{n}\lambda_0\} \leq \alpha.$$

Tada kriterijaus kritinė sritis

$$K_1 = \{\mathbf{X} : T \geq k'\}.$$

Pažymėkime t statistikos T realizaciją ir tegu $F(x|\nu)$ yra χ^2 skirstinio su ν laisvės laipsniais pasiskirstymo funkcija. Tada kriterijus P reikšmių terminais formuluojamas taip: hipotezė atmetama, kai teisinga nelygybė

$$pv = \mathbf{P}_{\lambda_0}\{T \geq t\} = 1 - F(2n\lambda_0|2t) \leq \alpha.$$

Pagaliau kritinę sritį K_1 galima užrašyti parametro λ pasikliovimo intervalo terminais:

$$K_1 = \{\mathbf{X} : \lambda_0 < \underline{\lambda} = \frac{1}{2n}\chi^2_{1-\alpha}(2T)\};$$

čia intervalo pasikliovimo lygmuo $Q = 1 - 2\alpha$.

Analogiškai sudarome TG kriterijų, kai alternatyva yra $\bar{H}_2$. Atsižvelgus į ankstesnes pastabas, kriterijaus kritinę sritį galima užrašyti trimis ekvivalenčiais pavidalais. Tegu k'' – didžiausias sveikasis skaičius, kuriam

$$\mathbf{P}_{\lambda_0}(T \leq k'') = \mathbf{P}\{\chi^2_{2k''+2} > 2n\lambda_0\} \leq \alpha.$$

Tada kriterijaus kritinė sritis

$$K_2 = \{\mathbf{X} : T \leq k''\} \Leftrightarrow pv = F(2n\lambda_0|2t+2) \leq \alpha \Leftrightarrow$$

$$\Leftrightarrow \{\mathbf{X} : \lambda_0 > \overline{\lambda} = \frac{1}{2n}\chi^2_{\alpha}(2T+2)\};$$

čia intervalo pasikliovimo lygmuo $Q = 1 - 2\alpha$.

I.3.33. a) Kai alternatyva dvipusė $\lambda \neq \lambda_0$, tai, remiantis [2], 4.3.2 teorema, egzistuoja TGN kriterijus

$$\varphi(T) = \begin{cases} 1, & \text{kai } T < c_1, \text{ arba } T > c_2, \\ \gamma, & \text{kai } T = c_i, \ i = 1, \ 2, \\ 0, & \text{kai } c_1 < T < c_2. \end{cases}$$

Konstantos c_1, c_2, γ_1, γ_2 randamos iš sąlygų

$$\mathbf{E}_{\lambda_0}(\varphi(T)) = \alpha, \quad \mathbf{E}_{\lambda_0}(T\varphi(T)) = \alpha\mathbf{E}_{\lambda_0}(T).$$

Kadangi

$$\mathbf{E}_{\lambda_0}(T) = n\lambda_0, \quad \mathbf{E}_{\lambda_0}(T\mathbf{1}_{[a,\, b]}(T)) = n\lambda_0\mathbf{P}_{\lambda_0}\{a-1 \leq T \leq b-1\},$$

tai lygtis konstantoms rasti galime perrašyti taip:

$$\sum_{k=c_1+1}^{c_2-1} \frac{(n\lambda_0)^k}{k!} \mathrm{e}^{-\mathrm{n}\lambda_0} + \sum_{\mathrm{i}=1}^{2} (1-\gamma_\mathrm{i}) \frac{(\mathrm{n}\lambda_0)^{\mathrm{c_i}}}{\mathrm{c_i}!} \mathrm{e}^{-\mathrm{n}\lambda_0} = 1-\alpha,$$

$$\sum_{k=c_1}^{c_2-2} \frac{(n\lambda_0)^k}{k!} \mathrm{e}^{-\mathrm{n}\lambda_0} + \sum_{\mathrm{i}=1}^{2} (1-\gamma_\mathrm{i}) \frac{(\mathrm{n}\lambda_0)^{\mathrm{c_i}-1}}{(\mathrm{c_i}-1)!} \mathrm{e}^{-\mathrm{n}\lambda_0} = 1-\alpha.$$

b) Atsisakius randomizacijos ir imant simetriškus kriterijus kritinė sritis gali būti užrašyta trimis ekvivalenčiais pavidalais. Tegu k'' ir k' yra tokie didžiausias ir mažiausias sveikieji skaičiai, kad

$$\mathbf{P}_{\lambda_0}\{T \le k''\} \le \frac{\alpha}{2}, \quad \mathbf{P}_{\lambda_0}\{T \ge k'\} \le \frac{\alpha}{2}.$$

Tada kriterijaus kritinė sritis yra

$$K_3 = \{\mathbf{X} : T \le k'' \text{ arba } T \ge k'\} \Leftrightarrow$$

$$\Leftrightarrow pv = 2\min(1 - F(2n\lambda_0|2t), F(2n\lambda_0|2t+2)) \le \alpha \Leftrightarrow$$

$$\Leftrightarrow \{\mathbf{X} : \lambda_0 < \underline{\lambda} = \frac{1}{2n}\chi^2_{1-\alpha/2}(2T), \text{ arba } \lambda_0 > \overline{\lambda} = \frac{1}{2n}\chi^2_{\alpha/2}(2T+2)\},$$

pastarojo intervalo pasikliovimo lygmuo $Q = 1-\alpha$.

I.3.34. Kai hipotezė H teisinga ir $\lambda = \lambda_0 = 2,5$, tai a. d. $T = \sum_i X_i$ turi Puasono skirstinį su parametru $n\lambda_0 = 100$. Kadangi $\mathbf{P}\{T \le 83|\lambda = \lambda_0\} = 0,0463 < 0,05$, o $\mathbf{P}\{T \le 84|\lambda = \lambda_0\} = 0,0575 > 0,05$, tai reikšmingumo lygmens TG kriterijus yra toks: hipotezė H atmetama, kai $T \le 83$, atmetama su tikimybe $\gamma = 0,327$, kai $T = 84$, ir priimama kitais atvejais. Kadangi šiame pavyzdyje statistikos T realizacija yra $T = \hat{\lambda}n = 80$, tai hipotezė H atmetama.

Dažniau naudojami nerandomizuoti kriterijai, gaunami šiek tiek sumažinus reikšmingumo lygmenį. Šiame pavyzdyje nerandomizuotu kriterijumi hipotezė H būtų atmetama, kai $T \le 83$. Kriterijaus reikšmingumo lygmuo $\alpha' = 0,0463 < 0,05$.

Kriterijų galima suformuluoti P reikšmių terminais: hipotezė atmetama, kai P reikšmė $pv = \mathbf{P}\{T \le 80|\lambda = \lambda_0\} = 0,0226$ yra mažesnė už reikšmingumo lygmenį $\alpha = 0,05$. Hipotezė atmetama.

Pagaliau kriterijų galime suformuluoti pasikliovimo intervalo terminais. Randame parametro λ pasikliovimo lygmens $Q = 1-2\alpha = 0,9$ pasikliovimo intervalą

$$(\underline{\lambda};\ \overline{\lambda}) = \left(\frac{1}{2n}\chi^2_{1-\alpha}(2T);\ \frac{1}{2n}\chi^2_{\alpha}(2T+2)\right) =$$

$$= \left(\frac{1}{80}\chi^2_{0,95}(160);\ \frac{1}{80}\chi^2_{\alpha}(162)\right) = (1,6470;\ 2,4088).$$

Kadangi $\overline{\lambda} < 2,5$, t. y. pasikliovimo intervalas pasislinkęs į kairę nuo hipotetinės reikšmės $\lambda_0 = 2,5$, tai hipotezę atmetame.

I.3.35. a) Imties tankis (skaičiuojančiojo mato atžvilgiu) yra

$$f(\mathbf{X},\ p) = \exp\{\theta T(\mathbf{X}) - B(\theta)\},$$

čia
$$\theta = \ln\frac{p}{1-p}, \quad T = T(\mathbf{X}) = X_1 + \cdots + X_n, \quad B(\theta) = -n\ln(1-p),$$
priklauso vienparametrei eksponentinio tipo skirstinių šeimai. Todėl egzistuoja TG kriterijai hipotezei dėl p reikšmės tikrinti, kai alternatyvos vienpusės, ir TGN kriterijus – kai alternatyva dvipusė.

Kai alternatyva yra $\bar{H}_1$, TG kriterijus yra
$$\varphi(T) = \begin{cases} 1, & \text{kai } T > m, \\ \gamma, & \text{kai } T = m, \\ 0, & \text{kai } T < m; \end{cases}$$
čia konstantos m ir γ randamos iš sąlygos
$$\mathbf{E}_{p_0}(\varphi(T)) = \sum_{k=m+1}^{n} C_n^k p_0^k (1-p_0)^{n-k} + \gamma C_n^m p_0^m (1-p_0)^{n-m} = \alpha.$$

Analogiškai, kai alternatyva yra $\bar{H}_2$, TG kriterijus yra
$$\varphi(T) = \begin{cases} 1, & \text{kai } T < m, \\ \gamma, & \text{kai } T = m, \\ 0, & \text{kai } T > m; \end{cases}$$
čia konstantos m ir γ randamos iš sąlygos
$$\mathbf{E}_{p_0}(\varphi(T)) = \sum_{k=0}^{m-1} C_n^k p_0^k (1-p_0)^{n-k} + \gamma C_n^m p_0^m (1-p_0)^{n-m} = \alpha.$$

Kai alternatyva dvipusė, egzistuoja TGN kriterijus
$$\varphi(T) = \begin{cases} 1, & \text{kai } T < m_1 \text{ arba } T > m_2, \\ \gamma_i, & \text{kai } T = m_i, \ i = 1,\ 2, \\ 0, & \text{kai } m_1 < T < m_2; \end{cases}$$
čia konstantos $m_1,\ m_2,\ \gamma_1,\ \gamma_2$ randamos iš lygčių sistemos ($q_0 = 1 - p_0$)
$$\begin{cases} \sum_{k=m_1+1}^{m_2-1} C_n^k p_0^k q_0^{n-k} + \sum_{i=0}^{2} (1-\gamma_i) C_n^{m_i} p_0^{m_i} q_0^{n-m_i} = 1-\alpha, \\ \sum_{k=m_1+1}^{m_2-1} C_{n-1}^{k-1} p_0^{k-1} q_0^{n-k} + \sum_{i=0}^{2} (1-\gamma_i) C_{n-1}^{m_i-1} p_0^{m_i-1} q_0^{n-m_i} = 1-\alpha. \end{cases}$$

Nereikalaujant, kad reikšmingumo lygmuo būtų tiksliai lygus α, t. y. įtraukiant taškus $T = m_i$ į priėmimo sritį, nerandomizuotus kriterijus galima suformuluoti taip. Tegu m' yra mažiausias sveikasis skaičius, kuriam $\mathbf{P}_{p_0}\{T \ge m'\} \le \alpha$. Nerandomizuotas kriterijus gali būti užrašytas tokiais trimis ekvivalenčiais pavidalais:
$$K_1 = \{\mathbf{X} : T \ge m'\} \Leftrightarrow pv = I_{p_0}(t,\ n-t+1) \le \alpha \Leftrightarrow$$
$$\Leftrightarrow \{\mathbf{X} : p_0 \le \underline{p} = X_{1-\alpha}(T,\ n-T+1)\};$$
čia intervalo pasikliovimo lygmuo $Q = 1 - 2\alpha$, o t yra statistikos T realizacija, $I_x(\gamma,\eta)$ beta skirstinio su parametrais γ ir η, o $X_\alpha(\gamma,\eta)$ šio skirstinio eilės α kritinė reikšmė.

Pažymėję m'' didžiausią sveikąjį skaičių, kuris tenkina nelygybę $\mathbf{P}_{p_0}\{T \leq m''\} \leq \alpha$, gauname nerandomizuotą kriterijų, kuris taip pat gali būti užrašytas trimis ekvivalenčiais pavidalais

$$K_2 = \{\mathbf{X} : T \leq m''\} \Leftrightarrow pv = 1 - I_{p_0}(t+1,\ n-t) \leq \alpha \Leftrightarrow$$
$$\Leftrightarrow \{\mathbf{X} : p_0 > \overline{p} = X_{1-\alpha}(T+1,\ n-T)\};$$

čia intervalo pasikliovimo lygmuo $Q = 1 - 2\alpha$.

Pažymėkime m'' didžiausią sveikąjį skaičių, kuriam $\mathbf{P}_{p_0}\{T \leq m''\} \leq \alpha/2$, o m'– mažiausią sveikąjį skaičių, kuriam $\mathbf{P}_{p_0}\{T \geq m'\} \leq \alpha/2$. Tada gauname nerandomizuotą paslinktąjį, tačiau lengviau randamą kriterijų:

$$K_3 = \{\mathbf{X} : T \leq m'' \text{ arba } T \geq m'\} \Leftrightarrow$$
$$\Leftrightarrow pv = 2\min(I_{p_0}(t,\ n-t+1),\ I_{p_0}(t+1,\ n-t)) \leq \alpha \Leftrightarrow$$
$$\Leftrightarrow \{\mathbf{X} : p_0 < \underline{p} = X_{1-\alpha/2}(T,\ n-T+1) \text{ arba } p_0 > \overline{p} = X_{\alpha/2}(T+1,\ n-T)\};$$

čia intervalo pasikliovimo lygmuo $Q = 1 - \alpha$.

I.3.36. 0,0879; 0,2052; 0,3731; 0,5703, kai $\alpha = 0,05$; 0,1581; 0,3101; 0,4917; 0,6752, kai $\alpha = 0,1$; 0,2891; 0,4877; 0,6804; 0,8350, kai $\alpha = 0,2$. Imties didumas: a) $n \geq 39$; b) $n \geq 471$.

I.3.37. Prielaida atmetama, nes P reikšmė $pv = \mathbf{P}\{x \geq 6 | n, p_0\} = 0,0378 < 0,05$.

I.3.38. a) H_0 atmetame, kai $S = X_1 + ... + X_{10} \geq 5$; atmetame su tikimybe $\gamma_1 = 0,4657$, kai $S = 0$; atmetame su tikimybe $\gamma_2 = 0,1954$, kai $S = 4$; b) H_0 atmetame, kai $S = 0$; 8; 9; 10; atmetame su tikimybe $\gamma_1 = 0,4702$, kai $S = 1$; atmetame su tikimybe $\gamma_2 = 0,2992$, kai $S = 7$.

I.3.39. Hipotezę H_0 atmetame, kai $X < k_1$ arba $X > k_2$; atmetame su tikimybe γ_i, kai $X = k_i$, $i = 1,\ 2$; konstantos $k_1,\ k_2,\ \gamma_1,\ \gamma_2$ randamos iš lygčių sistemos:

$$p_0 \left[\sum_{k=k_1+1}^{k_2-1} q_0^{k-1} + (1-\gamma_1) q_0^{k_1} + (1-\gamma_2) q_0^{k_2} \right] = 1 - \alpha,$$

$$p_0^2 \left[\sum_{k=k_1+1}^{k_2-1} k q_0^{k-1} + (1-\gamma_1) k_1 q_0^{k_1-1} + (1-\gamma_2) k_2 q_0^{k_2-1} \right] = 1 - \alpha.$$

I.3.40. Tikėtinumo funkcija

$$L(\theta) = \mathbf{1}_{(0,\theta)}(X_{(n)})/\theta^n.$$

Tegu $\theta_1 < \theta_2$. Tada santykis $L(\theta_2)/L(\theta_1)$ yra nemažėjanti statistikos $X_{(n)}$ funkcija. Remiantis [2], 4.3.1 teorema TG kriterijus alternatyvos $\bar{H}_1 : \theta > \theta_0$ atveju atmeta hipotezę, kai $X_{(n)} > (1-\alpha)^{1/n}\theta_0$.

Analogiškai, alternatyvos $H_2 : \theta < \theta_0$ atveju TG kriterijus atmeta hipotezę, kai $X_{(n)} < \alpha^{1/n}\theta_0$.

Pastebėkime, kad abiejų alternatyvų atveju TG kriterijai ekvivalentūs kriterijui su kritine sritimi

$$K = \{X_{(n)} : X_{(n)} < \alpha^{1/n}\theta_0 \quad \text{arba} \quad X_{(n)} > \theta_0\}.$$

Iš tikrųjų

$$\mathbf{P}_{\theta_0}\{X_{(n)} < \alpha^{1/n}\theta_0\} + \mathbf{P}_{\theta_0}\{X_{(n)} > \theta_0\} = \alpha.$$

Jeigu $\theta > \theta_0$, tai pirmojo TG kriterijaus galia

$$\mathbf{P}_{\theta}\{X_{(n)} > (1-\alpha)^{1/n}\theta_0\} = 1 - (1-\alpha)(\theta_0/\theta)^n,$$

o kriterijaus su kritine sritimi K galia tokia pat:

$$\mathbf{P}_{\theta}\{X_{(n)} < \alpha^{1/n}\theta_0\} + \mathbf{P}_{\theta}\{X_{(n)} > \theta_0\} = \alpha(\theta_0/\theta)^n + 1 - (\theta_0/\theta)^n = 1 - (1-\alpha)(\theta_0/\theta)^n.$$

Jeigu $\theta < \theta_0$, tai antrojo TG kriterijaus galia

$$\mathbf{P}_{\theta}\{X_{(n)} < \alpha^{1/n}\theta_0\} = \alpha(\theta_0/\theta)^n,$$

kai $\alpha(\theta_0/\theta)^n < 1$, ir lygi 1 priešingu atveju. Kriterijaus su kritine sritimi K galia tokia pat, nes $\mathbf{P}_{\theta}\{X_{(n)} > \theta_0\} = 0$.

Taigi kriterijus su kritine sritimi K yra TG ir dvipusės alternatyvos atveju.

I.3.41. a) Kai alternatyva yra $\bar{H} : \sigma > \sigma_0$, tai H atmetame, kai $T = \sum_i (X_i - \mu) < \chi^2_{1-\alpha}(2n)/(2\sigma_0^2)$; jeigu alternatyva yra $\bar{H} : \sigma < \sigma_0$, tai H atmetame, kai $T > \chi^2_{\alpha}(2n)/(2\sigma_0^2)$; b) Jeigu alternatyva yra $\bar{H} : \mu > \mu_0$, tai kriterijaus funkcija $\varphi(\mathbf{X}) = 1$, kai $X_{(1)} \geq \mu_0 - (\ln\alpha)/(\sigma n)$, ir $\varphi(\mathbf{X}) = 0$ kitais atvejais; jeigu alternatyva yra $\mu < \mu_0$, tai kriterijaus funkcija $\varphi(\mathbf{X}) = 1$, kai $X_{(1)} < \mu_0$, ir $\varphi(\mathbf{X}) = \alpha$, kai $X_{(1)} \geq \mu_0$.

I.3.42. H_0 atmetame, kai: a) $S = X_1 + ... + X_n > \theta_0\chi^2_{\alpha}(2n)/2$; b) $S = -(\ln X_1 + ... + \ln X_n) < \chi^2_{1-\alpha}(2n)/(2\theta_0)$; c) $\sum_i (X_i - 1)^2 > \theta_0\chi^2_{\alpha}(n)$; d) $S = X_1^c + ... + X_n^c > \theta_0^c\chi^2_{\alpha}(2n)/2$.

I.3.43. a) $(X_{(1)}, X_{(n)}) \sim f(x, y) = n(n-1)(y-x)^{n-2}$, $\theta < x < y < \theta + 1$. b) 0. c) $\beta(\theta) = 1 - (1 - (\theta - \theta_0))^n$, kai $\theta_0 \leq \theta \leq \theta_0 + 1$; $\beta(\theta) = 1$, kai $\theta > \theta_0 + 1$. d) $n \geq 44$.

I.3.48. TG kriterijus atmeta H, kai $a_2 = \sum_i X_i^2/n > c$; konstanta c randama iš sąlygos $\mathbf{P}\{a_2 > c|\mu_0\} = \alpha$. Kadangi $\mathbf{E}(a_2|\mu) = \mu^2 + k\mu$ ir $\boldsymbol{V}(a_2|\mu) = 2k\mu^2(k+2\mu)$, tai esant dideliems n konstanta $c \approx \mu_0^2 + k\mu_0 + \mu_0 z_\alpha\sqrt{2k(k+2\mu_0)/n}$; kriterijaus galia $\beta(\mu) \approx 1 - \Phi(\sqrt{n}(c - \mu^2 - k\mu)/(\mu\sqrt{2k(k+2\mu)})$, $\mu > \mu_0$.

I.3.49. Tai atskiras **I.3.48** pratimo atvejis, kai imties didumas lygus 1, o vietoje k imama k/n, t. y. **I.3.48** pratimo atsakymuose reikia įrašyti $\bar{X}^2$ vietoje a_2 ir k/n vietoje k.

I.3.50. Kadangi $\mathbf{P}_\beta(0) = 0$, tai a. d. X įgyja reikšmes 1,2,... Tankio funkcija $f(x|\beta) = \beta^{x-1} - \beta^x, x = 1, 2, ...$ Kai $\beta_1 > \beta_0$, tai santykis

$$\frac{f(x|\beta_1)}{f(x|\beta_0)} = (\frac{\beta_1}{\beta_0})^{X-1}\frac{1-\beta_1}{1-\beta_0}$$

didėja X atžvilgiu. Remiantis [2], 4.3.1 teorema TG kriterijus turi pavidalą: $\varphi(X) = 1$, kai $X > c$; $\varphi(X) = \gamma$, kai $X = c$; $\varphi(X = 0) = 0$, kai $X < c$, ir kriterijaus reikšmingumo lygmuo $\alpha = \mathbf{E}_{\theta_0}\varphi(X)$. Randame, kad c yra mažiausias sveikasis skaičius, tenkinantis nelygybę $p = \beta_0^c \leq \alpha$ ir $\gamma = (\alpha - p)/(\beta_0^{c-1}(1-\beta_0))$.

I.3.4 skyrelis

I.3.51. Pereidami prie a. d. $Y_i = X_i - \mu_0$ nesiaurindami prasmės galime tarti, kad $\mu_0 = 0$. Tikėtinumo funkcija

$$L(\mu, \sigma^2) = (\sqrt{2\pi})^{-n} \exp\{\theta U + \vartheta T\} \exp\{-n\mu^2/(2\sigma^2) - n \ln \sigma\}$$

priklauso dviparametrei eksponentinio tipo skirstinių šeimai su parametru $\theta = n\mu/\sigma^2$, trukdančiuoju parametru $\vartheta = -1/(2\sigma^2)$ ir statistikomis $U = \bar{X}, T = \sum_i X_i^2$.

Kai $\mu = 0$, statistikos

$$V = h(U, T) = \frac{\sqrt{n(n-1)U}}{\sqrt{T - nU^2}} = \frac{\sqrt{n}\bar{X}}{s} \sim S(n-1), \quad s^2 = \frac{1}{n-1} \sum_i (X_i - \bar{X})^2$$

skirstinys nepriklauso nuo ϑ ir monotoninis pagal U. Remiantis [2], 4.4.3 teorema egzistuoja TGN kriterijus hipotezei $H : \theta = 0$, kai alternatyva yra $\bar{H}_1 : \theta > 0$, tikrinti, kurio kritinė sritis

$$K_1 = \{\mathbf{X} : t(\mathbf{X}) = \sqrt{n}\bar{X}/s > t_\alpha(n-1)\}.$$

Kai $\theta > 0$ statistika $t(\mathbf{X})$ turi necentrinį Stjudento skirstinį su $n - 1$ laisvės laipsniu ir necentriškumo parametru $\delta = \sqrt{n}\theta/\sigma$. Todėl kriterijaus galia

$$\beta_1(\mu, \sigma^2) = \mathbf{P}\{t_{\delta;n-1} > t_\alpha(n-1)\}.$$

Analogiškai, kai alternatyva yra $\bar{H}_2 : \theta < 0$, TGN kriterijaus kritinė sritis

$$K_2 = \{\mathbf{X} : t(\mathbf{X}) = \sqrt{n}\bar{X}/s < -t_\alpha(n-1)\}.$$

Grįždami prie hipotezės $H : \mu = \mu_0$ gausime, kad vienpusių alternatyvų atvejais kriterijai nusakomi kritinėmis sritimis K_1, K_2, kuriose reikia imti $t(\mathbf{X}) = \sqrt{n}(\bar{X} - \mu_0)/s$.

Tegu t yra statistikos $t(\mathbf{X})$ realizacija, o $F(x|\nu)$ Stjudento skirstinio su ν laisvės laipsnių pasiskirstymo funkcija. Tada kriterijai su kritinėmis sritimis K_1, K_2 P reikšmių terminais formuluojami taip: hipotezė H atmetama, kai atitinkamai

$$pv = 1 - F(t|n-1) < \alpha, \quad pv = F(t|n-1) < \alpha.$$

I.3.52. Statistika V nėra tiesinė statistikos U atžvilgiu. Imkime statistiką

$$W = \frac{U}{\sqrt{T}} = \frac{\bar{X}}{\sqrt{\sum_i X_i^2}},$$

kurios skirstinys, kai $\theta = 0$, nepriklauso nuo σ^2 ir ji tiesinė U atžvilgiu. Remiantis [2], 4.4.3 teorema egzistuoja TGN kriterijus, kurio priėmimo sritis yra

$$c_1 < W < c_2.$$

Statistikos W skirstinys, kai $\theta = 0$, yra simetriškas 0 atžvilgiu. Todėl kritinė sritis yra $|W| > c$, kai c randamas iš sąlygos $\mathbf{P}_0\{W > c\} = \alpha/2$. Statistikos W ir $t(\mathbf{X})$ susietos lygybe

$$t(\mathbf{X}) = \frac{W\sqrt{n(n-1)}}{\sqrt{1 - nW^2}},$$

iš kurios matome, kad $|t|$ yra didėjanti $|W|$ funkcija. Todėl α lygmens kriterijaus kritinę sritį $|W| > c$ galima perrašyti taip

$$K_3 = \{\mathbf{X} : |t(\mathbf{X})| = \sqrt{n}|\bar{X} - \mu_0|/s > t_{\alpha/2}(n-1)\}.$$

P reikšmių terminais hipotezė atmetama, kai

$$pv = 2(1 - F(|t||n-1)) < \alpha.$$

I.3.53. Perrašykime tikėtinumo funkciją tokiu būdu

$$L(\mu, \sigma^2) = (\sqrt{2\pi})^{-n} \exp\{\theta U + \vartheta T\} \exp\{-n\mu^2/(2\sigma^2) - n \ln \sigma\},$$

kai dominantis parametras $\theta = -1/(2\sigma^2)$, trukdantysis parametras $\vartheta = n\mu/\sigma^2$ ir statistikos $U = \sum_i X_i^2, T = \bar{X}$.

Statistika

$$V = h(U, T) = U - nT^2 = \sum_i (X_i - \bar{X})^2$$

yra tiesinė U funkcija, o jos skirstinys, kai σ fiksuotas, nepriklauso nuo ϑ, taip pat ir nuo T; statistika $V/\sigma_0^2 \sim \chi^2(n-1)$, kai H teisinga.

Remiantis [2], 4.4.3 teorema alternatyvos $\bar{H}_1 : \sigma^2 > \sigma_0^2$ atveju egzistuoja TGN kriterijus, kurio kritinė sritis yra

$$K_1 = \{\mathbf{X} : V/\sigma_0^2 > \chi^2_\alpha(n-1)\}.$$

Jeigu v yra statistikos V/σ_0^2 realizacija, tai P reikšmių terminais hipotezė atmetama, kai

$$pv = \mathbf{P}\{\chi^2_{n-1} > v\} < \alpha.$$

Analogiškai, kai alternatyva $\bar{H}_2 : \sigma^2 < \sigma_0^2$, egzistuoja TGN kriterijus, kurio kritinė sritis yra

$$K_2 = \{\mathbf{X} : V/\sigma_0^2 < \chi^2_{1-\alpha}(n-1)\},$$

arba P reikšmių terminais hipotezė atmetama, kai

$$pv = \mathbf{P}\{\chi^2_{n-1} < v\} < \alpha.$$

Kriterijų galios funkcijos

$$\beta_1(\sigma^2) = \mathbf{P}\{\chi^2_{n-1} > (\sigma_0^2/\sigma^2)\chi^2_\alpha(n-1)\}, \quad \beta_2(\sigma^2) = \mathbf{P}\{\chi^2_{n-1} < (\sigma_0^2/\sigma^2)\chi^2_{1-\alpha}(n-1)\}.$$

I.3.54. Statistika V tiesinė U atžvilgiu. Remiantis [2], 4.4.3 teorema egzistuoja TGN kriterijus, kurio priėmimo sritis

$$A_3 = \{\mathbf{X} : c_1 < V/\sigma_0^2 < c_2\}.$$

Analogiškai **I.3.25** pratimui konstantos c_1 ir c_2 gali būti surastos iš lygčių sistemos

$$\mathbf{P}\{c_1 < \chi^2_{n-1} < c_2\} = 1 - \alpha,$$

$$\mathbf{P}\{c_1 < \chi^2_{n+1} < c_2\} = 1 - \alpha.$$

Imdami paslinktą, tačiau paprasčiau randamą simetrinį kriterijų gausime kritinę sritį

$$K_3^* = \{\mathbf{X} : V/\sigma_0^2 < \chi^2_{1-\alpha/2}(n-1), \text{ arba } V/\sigma_0^2 > \chi^2_{\alpha/2}(n-1)\};$$

P reikšmių terminais hipotezė atmetama, kai

$$pv = 2\min(\mathbf{P}\{\chi^2_{n-1} < v\}, \mathbf{P}\{\chi^2_{n-1} > v\}) < \alpha.$$

Kriterijaus galios funkcija

$$\beta_3(\sigma^2) = 1 - \mathbf{P}\{(\sigma_0^2/\sigma^2)\chi^2_{1-\alpha/2} < \chi^2_{n-1} < (\sigma_0^2/\sigma^2)\chi^2_{\alpha/2}\}.$$

I.3.55. Žr. **I.3.51** pratimą.

I.3.56. Žr. **I.3.51** pratimą.

I.3.57. Jeigu σ žinoma, tai kriterijaus galia lygi 0,5573; 0,7228; 0,8505, kai $n = 5$; 0,8119; 0,9354; 0,9842, kai $n = 10$; 0,9270; 0,9871; 0,9987, kai $n = 15$. Jeigu σ nežinoma, tai kriterijaus galia lygi 0,5869; 0,7280; 0,8393, kai $n = 5$; 0,8077; 0,9275; 0,9795, kai $n = 10$; 0,9216; 0,9844; 0,9961, kai $n = 15$.

I.3.58. Užrašę tikėtinumo funkciją tokiu būdu

$$L(\mu_1, ..., \mu_s, \sigma^2) = \exp\{\theta_1 X_1 + \sum_{j=2}^{n} \theta_j X_j - \frac{1}{2\sigma^2}\sum_{i=1}^{n} X_i^2 - B(\theta_1, ..., \theta_s, \sigma^2)\},$$

čia $\theta_i = \mu_i/\sigma^2, i = 1, ..., s$, matome, kad ji priklauso $(s+1)$-mačių eksponentinių skirstinių šeimai. Dominantis parametras θ_1, o trukdantysis $\boldsymbol{\vartheta} = (\theta_2, ..., \theta_s, -1/(2\sigma^2))^T$; pakankamosios statistikos $U = U(\mathbf{X}) = X_1$, $\boldsymbol{T} = \boldsymbol{T}(\mathbf{X}) = (X_2, ..., X_s, \sum_i X_i^2)^T$. Nemažinant bendrumo galima tarti, kad tikrinama hipotezė $H : \theta_1 = 0$, t. y. $\mu_1^0 = 0$ (priešingu atveju vietoje X_1 imsime $Y_1 = X_i - \mu_1^0$). Statistikos $V(U, \boldsymbol{T}) = (X_1 - \mu_1)/\sqrt{\hat{\sigma}^2} \sim S(n-n)$, čia $\hat{\sigma}^2 = \sum_{i=s+1}^{n} X_i^2/(n-s)$, skirstinys, kai $\theta_1 = 0$, nepriklauso nuo $\boldsymbol{\vartheta}$. Remiantis [2], 4.4.3 teorema, kai alternatyva yra $\bar{H}_1 : \theta_1 > 0$, reikšmingumo α TGN kriterijus atmeta hipotezę, kai $T = (X_1 - \mu_1^0)/\sqrt{\hat{\sigma}^2} > t_\alpha(n-s)$. Alternatyvos $\bar{H}_2 : \theta_1 < 0$ atveju hipotezė atmetama, kai $T < -t_\alpha(n-s)$, o dvipusės alternatyvos $\bar{H}_3 : \theta_1 \neq 0$ atveju – kai $|T| > t_{\alpha/2}(n-s)$.

I.3.59. Pažymėkime $\hat{\beta}_1 = \sum_i X_i(t_i - \bar{t})/\sum_i(t_i - \bar{t})^2$, $\hat{\beta}_0 = \bar{X}$, $\hat{\sigma}^2 = s^2 = \sum_i(X_i - \hat{\beta}_0 - \hat{\beta}_1(t_i - \bar{t}))^2/(n-2)$. Tada: a) H_0 atmetame, kai $\sqrt{n}(\hat{\beta}_0 - \theta_0)/(s) > t_\alpha(n-2)$; b) H_0 atmetame, kai $|\hat{\beta}_0 - \theta_0|/(s) > t_{\alpha/2}(n-2)$; c) H_0 atmetame, kai $(\hat{\beta}_1 - \theta_0)/(sd) > t_\alpha(n-2)$, $d^2 = 1/\sum_i(t_i - \bar{t})^2$; d) H_0 atmetame, kai $|\hat{\beta}_1 - \theta_0|/(sd) > t_{\alpha/2}(n-2)$.

I.3.60. a) Tikėtinumo funkcija

$$L(\lambda, \eta) = \exp\{\theta U + \vartheta T\}\exp\{n\eta\ln\lambda - n\ln\Gamma(\eta)\}/\prod_i X_i$$

priklauso dviparametrei eksponentinio tipo skirstinių šeimai. Dominantis parametras $\theta = \eta$, trukdantysis parametras $\vartheta = -n\lambda$ ir statistikos $U = \sum_i \ln X_i, T = \bar{X}$. Remiantis

[2], 4.4.2 teorema egzistuoja TGN kriterijus, kurį apibrėžiame kaip sąlyginį paviršiuose $T = t$. Kritinė sritis

$$K = \{\mathbf{X} : U > c(T)\} \quad \Leftrightarrow \quad \{\mathbf{X} : \prod_i X_i > g(\bar{X})\}.$$

b) Analogiškai p. a).

I.3.61. Jungtinės imties $(\mathbf{X}^T, \boldsymbol{Y}^T)^T$ tankis

$$f(\mathbf{X}, \boldsymbol{Y}; \boldsymbol{\theta}) = C \exp\{-\frac{1}{2\sigma_1^2} \sum_i X_i^2 + \frac{\mu_1}{\sigma_1^2} \sum_i X_i - \frac{1}{2\sigma_2^2} \sum_i Y_i^2 + \frac{\mu_2}{\sigma_2^2} \sum_i Y_i - b(\boldsymbol{\theta})\}$$

priklauso keturparametrei eksponentinio tipo skirstinių šeimai, parametras $\boldsymbol{\theta} = (\mu_1,\ \mu_2,\ \sigma_1,\ \sigma_2)^T$, $-\infty < \mu_1,\ \mu_2 < \infty$, $0 < \sigma_1,\ \sigma_2 < \infty$, $C = (2\pi)^{-(m+n)/2}$.

Pertvarkykime tankį taip, kad išsiskirtų parametras $\mu_1 - \mu_2 - \beta_0$:

$$f(\mathbf{X}, \boldsymbol{Y}; \theta, \vartheta) = \exp\{\theta U(\mathbf{X}, \boldsymbol{Y}) + \vartheta T(\mathbf{X}, \boldsymbol{Y}) - B(\theta, \vartheta)\},$$

$$\theta = \frac{(\mu_1 - \mu_2 - \beta_0)mn}{m\sigma_{10}^2 + n\sigma_{20}^2}, \quad \vartheta = \frac{n\sigma_{20}^2(\mu_1 - \beta_0) + m\sigma_{10}^2\mu_2}{m\sigma_{10}^2 + n\sigma_{20}^2},$$

$$U = U(\mathbf{X}, \boldsymbol{Y}) = \bar{X} - \bar{Y}, \quad T = T(\mathbf{X}, \boldsymbol{Y}) = \frac{n}{\sigma_{10}^2}\bar{X} + \frac{m}{\sigma_{20}^2}\bar{Y}.$$

Statistikos

$$Z = \frac{(\bar{X} - \bar{Y} - \beta_0)\sqrt{mn}}{\sqrt{m\sigma_{10}^2 + n\sigma_{20}^2}}$$

skirstinys, kai $\mu_1 - \mu_2 = \beta_0$, yra $N(0,\ 1)$ ir nepriklauso nuo ϑ, tai jis nepriklauso ir nuo T. Remiantis [2], 4.4.2 teorema, egzistuoja TGN kriterijai dėl parametro θ (kartu dėl skirtumo $\mu_1 - \mu_2$) reikšmių, kai yra vienpusės ar dvipusė alternatyvos. Kadangi Z yra monotoniškai didėjanti ir tiesinė pagal U, tai alternatyvų $\bar{H}_1 : \mu_1 - \mu_2 > \beta_0$, $\bar{H}_2 : \mu_1 - \mu_2 < \beta_0$, $\bar{H}_3 : \mu_1 - \mu_2 \neq \beta_0$ atvejais kritinės sritys yra

$$K_1 = \{(\mathbf{X}, \boldsymbol{Y}) : Z > z_\alpha\}, \quad K_2 = \{(\mathbf{X}, \boldsymbol{Y}) : Z < -z_\alpha\},$$

$$K_3 = \{(\mathbf{X}, \boldsymbol{Y}) : |Z| > z_{\alpha/2}\}.$$

Kriterijų galios funkcijos išreiškiamos standartinio normaliojo skirstinio pasiskirstymo funkcija $\Phi(x)$.

Pažymėkime z statistikos Z realizaciją. Tada P reikšmių terminais kriterijai K_1, K_2, K_3 formuluojami taip: hipotezė atmetama, kai atitinkamai teisingos nelygybės

$$pv = 1 - \Phi(z) \leq \alpha, \quad pv = \Phi(z) \leq \alpha,$$

$$pv = 2\min(1 - \Phi(z), \Phi(z)) = 2(1 - \Phi(|z|)) \leq \alpha.$$

I.3.62. Pertvarkykime tankį taip, kad išsiskirtų dominantis parametras $\mu_1 - \mu_2 - \beta_0$:

$$f(\mathbf{X}, \boldsymbol{Y}; \theta, \vartheta) = \exp\{\theta U + \vartheta_1 T_1 + \vartheta_2 T_2 - B(\theta, \vartheta_1, \vartheta_2)\},$$

$$U = \bar{X} - \bar{Y}, \quad T_1 = n\bar{X} + m\bar{Y}, \quad T_2 = \sum_i X_i^2 + \sum_i Y_i^2,$$

$$\theta = \frac{(\mu_1 - \mu_2 - \beta_0)mn}{(m+n)\sigma^2}, \quad \vartheta_1 = \frac{n(\mu_1 - \beta_0) + m\mu_2}{(m+n)\sigma^2}, \quad \vartheta_2 = -\frac{1}{2\sigma^2}.$$

Nagrinėkime statistiką

$$V = V(U, T_1, T_2) = \frac{\bar{X} - \bar{Y} - \beta_0}{\sqrt{\sum_i (X_i - \bar{X})^2 + \sum_i (Y_i - \bar{Y})^2}} = \frac{U - \beta_0}{\sqrt{T_2 - \frac{1}{m+n}T_1^2 - \frac{mn}{m+n}U^2}}.$$

Funkcija V yra monotoniškai didėjanti pagal U. Be to, jos skirstinys, kai $\theta = 0$, nepriklauso nuo parametrų μ_1, μ_2, σ (kartu nuo T_1, T_2). Tuo įsitikinti galima pažymėjus, kad V reikšmė nepakinta, kai X_i pakeičiama į $(X_i - \mu_1)/\sigma$ ir Y_i pakeičiama į $(Y_i - \mu_2)/\sigma$.

Remiantis [2], 4.4.2 teorema, egzistuoja TGN kriterijus, kai alternatyva yra $\overline{H}_1 : \theta > 0$. Jo kritinė sritis yra $V \geq c_1$, arba ekvivalenčia forma $T \geq c_1'$,

$$T = V\sqrt{\frac{mn(m+n-2)}{m+n}} = \frac{\bar{X} - \bar{Y} - \beta_0}{\sqrt{s_1^2(n-1) + s_2^2(m-1)}}\sqrt{\frac{mn(m+n-2)}{m+n}};$$

čia s_1^2 ir s_2^2 yra nepaslinktieji dispersijos σ^2 įvertiniai, sudaryti pagal imtį $\mathbf{X}$ ir imtį $\boldsymbol{Y}$. Kadangi statistikos T skirstinys, kai $\mu_1 - \mu_2 = \beta_0$, yra Stjudento su $m + n - 2$ laisvės laipsnių, TGN kriterijaus kritinė sritis yra

$$K_1 = \{(\mathbf{X}, \boldsymbol{Y}) : T > t_\alpha(m+n-2)\}.$$

Analogiškai, kai alternatyva yra $\overline{H}_2 : \mu_1 - \mu_2 < \beta_0$, TGN kriterijaus kritinė sritis yra

$$K_2 = \{(\mathbf{X}, \boldsymbol{Y}) : T < -t_\alpha(m+n-2)\}.$$

Tikrinant hipotezę $H_3 : \theta = 0$ (arba $\mu_1 - \mu_2 = \beta_0$), kai alternatyva $\overline{H}_3 : \theta \neq 0$ yra dvipusė, tiesiogiai pritaikyti [2], 4.4.3 teoremos negalima, nes V nėra tiesinė U funkcija. Imkime funkciją

$$W = \frac{\bar{X} - \bar{Y} - \beta_0}{\sqrt{\sum_i X_i^2 + \sum_i Y_i^2 - \frac{1}{m+n}(\sum_i X_i + \sum_i Y_i)^2}} = \frac{U}{\sqrt{T_2 - \frac{1}{m+n}T_1^2}},$$

kuri tiesiškai priklauso nuo U. Kadangi W ir V susieti lygybe

$$V = \frac{W}{\sqrt{1 - \frac{mn}{m+n}W^2}},$$

tai W skirstinys, kai $\theta = 0$, taip pat nepriklauso nuo T_1, T_2. Be to, W skirstinys, kai $\theta = 0$, yra simetrinis taško $\theta = 0$ atžvilgiu, todėl, remiantis 4.4.2 teorema, TGN kriterijaus kritinė sritis yra $|W| > c$.

Funkcija $|V|$ yra monotoniškai didėjanti pagal $|W|$. Todėl šią kritinę sritį galima perrašyti $|V|$ (kartu ir $|T|$) terminais. Gauname kritinę sritį

$$K_3 = \{(\mathbf{X}, \boldsymbol{Y}) : |T| > t_{\alpha/2}(m+n-2)\}.$$

Pažymėkime t statistikos T realizaciją ir tegu $F(x|\nu)$ yra Stjudento skirstinio su ν laisvės laipsniais pasiskirstymo funkcija. Tada P reikšmių terminais kriterijai K_1, K_2, K_3 formuluojami taip: hipotezė atmetama, kai atitinkamai teisingos nelygybės

$$pv = 1 - F(t|m+n-2) \leq \alpha, \quad pv = F(t|m+n-2) \leq \alpha,$$

$$pv = 2\min(1 - F(t|m+n-2), F(t|m+n-2)) = 2(1 - F(|t||m+n-2)) \le \alpha.$$

I.3.63. Tankį pertvarkykime taip:

$$f(\mathbf{X}, \boldsymbol{Y}; \theta, \vartheta) = \exp\{\theta \mathbf{U} + \vartheta_1 \mathbf{T_1} + \vartheta_2 \mathbf{T_2} + \vartheta_3 \mathbf{T_3} - \mathbf{B}(\theta, \vartheta)\},$$

$$\theta = -\frac{1}{2\sigma_2^2} + \frac{1}{2\sigma_1^2}, \quad \vartheta_1 = -\frac{1}{2\sigma_1^2}, \quad \vartheta_2 = \frac{n\mu_1}{\sigma_1^2}, \quad \vartheta_3 = \frac{n\mu_2}{\sigma_2^2},$$

$$U = \sum_i Y_i^2, \quad T_1 = \sum_i X_i^2 + \sum_i Y_i^2, \quad T_2 = \bar{X}, \quad T_3 = \bar{Y}.$$

Nagrinėkime šių statistikų funkciją

$$F = \frac{(m-1)\sum_i (X_i - \bar{X})^2}{(n-1)\sum_i (Y_i - \bar{Y})^2} = \frac{s_1^2}{s_2^2} = \frac{(m-1)(T_1 - U - nT_2^2)}{(n-1)(U - nT_3^2)},$$

kuri monotoninė pagal U. Jos skirstinys, kai $\theta = 0$ (t. y. $\sigma_1^2 = \sigma_2^2$), yra Fišerio skirstinys su $n-1$ ir $m-1$ laisvės laipsnių. Taigi F skirstinys, kai $\theta = 0$, nepriklauso nuo parametrų $\vartheta_1, \vartheta_2, \vartheta_3$, o kartu nuo statistikų T_1, T_2, T_3. Remiantis [2], 4.4.3 teorema, egzistuoja TGN kriterijai, kai alternatyvos yra $\bar{H}_1 : \sigma_1^2 > \sigma_2^2$, $\bar{H}_2 : \sigma_1^2 < \sigma_2^2$ vienpusės. Jų kritinės sritys yra

$$K_1 = \{(\mathbf{X}, \boldsymbol{Y}) : F > F_\alpha(n-1, m-1)\},$$

$$K_2 = \{(\mathbf{X}, \boldsymbol{Y}) : F < F_{1-\alpha}(n-1, m-1)\};$$

čia $F_\alpha(\nu_1, \nu_2)$ yra Fišerio skirstinio su ν_1 ir ν_2 laisvės laipsnių α kritinė reikšmė.

Pažymėkime f statistikos F realizaciją. Tada P reikšmių terminais kriterijai K_1, K_2 formuluojami taip: hipotezė atmetama, kai atitinkamai teisingos nelygybės

$$pv = 1 - \mathbf{P}\{F_{n-1,m-1} \ge f\} \le \alpha, \quad pv = \mathbf{P}\{F_{n-1,m-1} \le f\} \le \alpha.$$

Kriterijų K_1, K_2 galia išreiškiama Fišerio skirstinio pasiskirstymo funkcija, kai jos argumentas priklauso nuo dispersijų santykio $\lambda = \sigma_1^2/\sigma_2^2$:

$$\beta_1(\lambda) = \mathbf{P}\{F_{n-1,m-1} > \frac{1}{\lambda} F_\alpha(n-1, m-1)\}, \quad \lambda > k,$$

$$\beta_2(\lambda) = \mathbf{P}\{F_{n-1,m-1} < \frac{1}{\lambda} F_{1-\alpha}(n-1, m-1)\}, \quad \lambda < k.$$

Kai alternatyva $\bar{H}_3 : \theta \ne 0$ yra dvipusė, tiesiogiai pritaikyti [2], 4.4.3 teoremos negalima, nes F nėra tiesinė U funkcija. Apibrėžkime statistiką W lygybe

$$W = \frac{(n-1)F}{m-1+(n-1)F} = \frac{T_1 - kU - nT_2^2}{(T_1 - nT_2^2 - knT_3^2)}.$$

Statistika W yra tiesinė pagal U. Kadangi ji išreiškiama per F, tai jos skirstinys, kai $\theta = 0$, taip pat nepriklauso nuo T_1, T_2, T_3. Remiantis [2] 4.4.3 teorema, hipotezė H_3 atmetama, kai

$$W < c_1' \quad \text{arba} \quad W > c_2'.$$

Funkcija W yra monotoniškai didėjanti pagal F, kai $0 < F < \infty$. Todėl kritinę sritį galime perrašyti statistikos F terminais: $F < c_1$, arba $F > c_2$. Konstantos c_1 ir c_2 randamos iš sąlygų

$$\mathbf{E}(\varphi(F)|\sigma_1^2 = \sigma_2^2) = \mathbf{P}\{F_{n-1,m-1} < c_1\} + \mathbf{P}\{F_{n-1,m-1} > c_2\} = \alpha,$$

$$\mathbf{E}(F\varphi(F)|\sigma_1^2 = \sigma_2^2) = \alpha \mathbf{E}(F|\sigma_1^2) = \sigma_2^2.$$

Tada

$$\mathbf{E}(F|\sigma_1^2 = \sigma_2^2) = \mathbf{E}(F_{n-1,m-1}) = \frac{m-1}{m-3},$$

$$\mathbf{E}(F\varphi(F)|\sigma_1^2 = \sigma_2^2) = \frac{m-1}{m-3}[\mathbf{P}\{F_{n+1,m-3} < c_1\} + \mathbf{P}\{F_{n+1,m-3} > c_2\}].$$

Lygčių sistema parametrams c_1 ir c_2 rasti gali būti užrašyta taip:

$$\begin{cases} \mathbf{P}\{c_1 < F_{n-1,m-1} < c_2\} = 1 - \alpha, \\ \mathbf{P}\{c_1 < F_{n+1,m-3} < c_2\} = 1 - \alpha. \end{cases}$$

Praktiškai vietoje K_3 TGN kriterijaus dažniau naudojamas paslinktasis, tačiau paprasčiau randamas simetrinis kriterijus, kurio kritinė sritis

$$\widetilde{K}_3 = \{(\mathbf{X}, \boldsymbol{Y}) : F < F_{1-\alpha/2}(n-1, m-1), \text{ arba } F > F_{\alpha/2}(n-1, m-1)\},$$

P reikšmių terminais kriterijus formuluojamas taip: hipotezė atmetama, kai teisinga nelygybė

$$pv = 2\min(1 - \mathbf{P}\{F_{n-1,m-1} \le f\}, \mathbf{P}\{F_{n-1,m-1} \ge f\}) \le \alpha.$$

I.3.64. Tikėtinumo funkciją $L(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ užrašykime tokiu pavidalu

$$L(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \exp\{\theta U + \vartheta_1 T_1 + \vartheta_2 T_2 + \vartheta_3 T_3 + \vartheta_4 T_4 - B(\theta, \boldsymbol{\vartheta})\},$$

čia $\theta = \rho/(\sigma_1\sigma_2(1-\rho^2))$ yra dominantis parametras

$$\vartheta_1 = \frac{-1}{2\sigma_1^2(1-\rho^2)}, \ \vartheta_2 = \frac{-1}{2\sigma_2^2(1-\rho^2)}, \ \vartheta_3 = \frac{1}{(1-\rho^2)}(\frac{\mu_1}{\sigma_1^2} - \frac{\rho\mu_2}{\sigma_1\sigma_2}),$$

$$\vartheta_4 = \frac{1}{(1-\rho^2)}(\frac{\mu_2}{\sigma_2^2} - \frac{\rho\mu_1}{\sigma_1\sigma_2})$$

trukdantieji parametrai; statistikos

$$U = \sum_i X_i Y_i, \ T_1 = \sum_i X_i^2, \ T_2 = \sum_i Y_i^2, \ T_3 = \sum_i X_i, \ T_4 = \sum_i Y_i.$$

Turime penkiaparametrę eksponentinio tipo skirstinių šeimą. Hipotezė $H : \rho = 0$ ekvivalenti hipotezei $H : \theta = 0$. Nagrinėkime statistiką

$$r = r(U, \boldsymbol{T}) = \frac{s_{12}}{s_1 s_2} = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_i (X_i - \bar{X})^2 \sum_i (Y_i - \bar{Y})^2}} = \frac{U - T_3 T_4/n}{\sqrt{(T_1 - T_3^2/n)(T_2 - t_4^2/n)}}.$$

Statistikos r skirstinys nepriklauso nuo parametrų $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$. Kai hipotezė H : $\rho = 0$ yra teisinga, tai r skirstinys nepriklauso nuo parametro $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2, \vartheta_3, \vartheta)^T$, o kartu pačiu ir nuo statistikos $\boldsymbol{T} = (T_1, T_2, T_3, T_4)^T$. Kadangi r yra tiesinė U funkcija, tai remiantis [2] 4.4.3 teorema, tikrinant hipotezę $H : \rho = 0$, kai alternatyvos yra $\bar{H}_1 : \rho > 0$ arba $\bar{H}_2 : \rho < 0$, egzistuoja TGN kriterijai, kurių atmetimo sritys $r > c_2$, arba $r < c_1$. Dvipusės alternatyvos $\bar{H}_3 : \rho \neq 0$ atveju, remiantis statistikos r skirstinio simetriškumu 0 atžvilgiu, TGN kriterijus atmeta hipotezę, kai $|r| > d$. Jeigu kriterijaus reikšmingumo lygmuo α, tai konstantos c_1, c_2, d randamos iš sąlygų

$$F(c_1|0) = \alpha, \quad 1 - F(c_2|0) = \alpha, \quad F(-d|0) + 1 - F(d|0) = \alpha,$$

čia $F(x|\rho)$ – statistikos r pasiskirstymo funkcija; tankio funkcija $f(x|\rho)$ pateikta **I.2.169** pratime. Kai $\rho = 0$, tankis yra

$$f(x|0) = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n-2}{2})} (1 - r^2)^{\frac{n-4}{2}}.$$

Atlikę a. d. r transformaciją: $t = \sqrt{n-2}r/\sqrt{1-r^2}$, gauname, kad $t \sim S(n-2)$, kai $\rho = 0$. Kadangi t yra monotoniškai didėjanti r funkcija, tai pateiktieji reikšmingumo lygmens α TGN kriterijai įgauna tokį pavidalą

$$t > t_\alpha(n-2), \quad t < -t_\alpha(n-2), \quad |t| > t_{\alpha/2}(n-2).$$

Tegu $S(x|\nu)$ yra Stjudento skirstinio su ν laisvės laipsniais pasiskirstymo funkcija. Tada P reikšmių terminais kriterijai formuluojami taip: hipotezė atmetama, kai atitinkamai teisingos nelygybės

$$pv = 1 - S(t|n-2) \leq \alpha, \quad pv = S(t|n-2) \leq \alpha,$$

$$pv = 2\min(1 - S(t|n-2), S(t|n-2)) = 2(1 - S(|t||n-2)) \leq \alpha;$$

šiose nelygybėse t yra statistikos realizacija.

I.3.65. a) Tikėtinumų santykis

$$\Lambda = \frac{\max_{\boldsymbol{\mu}, \boldsymbol{\Sigma}: \rho = \rho_0} L(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\max_{\boldsymbol{\mu}, \boldsymbol{\Sigma}} L(\boldsymbol{\mu}, \boldsymbol{\Sigma})} = \frac{L(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}})}{L(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})};$$

čia $\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}$ yra parametrų $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ DT įvertiniai, o $\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}$ DT įvertiniai, surasti esant sąlygai $\rho = \rho_0$.

Parametrų $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ DT įvertiniai (žr. [4], 1.2 skyrelį) yra $\hat{\mu}_1 = \bar{X}, \hat{\mu}_2 = \bar{Y}$, $\hat{\boldsymbol{\Sigma}} = \boldsymbol{S}/n; \boldsymbol{S} = [S_{ij}]_{2\times 2}$, $S_{11} = \sum_i (X_i - \bar{X})^2, S_{22} = \sum_i (Y_i - \bar{Y})^2, S_{12} = S_{21} = \sum_i (X_i - \bar{X})(Y_i - \bar{Y})$, o tikėtinumo funkcijos maksimumas

$$L(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}) = (\sqrt{2\pi})^{-n} n^n e^{-n} |\boldsymbol{S}|^{-n/2}.$$

Užrašę tikėtinumo funkciją, kai $\rho = \rho_0$, tokiu pavidalu

$$L(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(\sqrt{2\pi(1-\rho_0^2)}\sigma_1\sigma_2)^n} \exp\left\{ -\frac{1}{2(1 - rho_0^2)} \frac{S_{11}}{\sigma_1^2} - 2\rho_0 \frac{S_{12}}{\sigma_1\sigma_2} + \frac{S_{22}^2}{\sigma_2^2} \right\}$$

$$\exp\left\{ -\frac{n}{2(1-\rho_0^2)} \left[\frac{(\bar{X} - \mu_1)^2}{\sigma_1^2} - 2\rho_0 \frac{(\bar{X} - \mu_1)(\bar{Y} - \mu_2)}{\sigma_1\sigma_2} + \frac{(\bar{Y} - \mu_2)^2}{\sigma_2^2} \right] \right\}$$

matome, kad paskutinis daugiklis įgyja maksimalią reikšmę 1, kai $\tilde{\mu}_1 = \bar{\mu}_1 = \bar{X}, \tilde{\mu}_2 = \bar{\mu}_2 = \bar{Y}$. Lieka rasti daugiklio, priklausančio nuo σ_1 ir σ_2, maksimumą. Logaritmuodami gauname

$$h(\sigma_1, \sigma_2) = -n \ln \sigma_1 - n \ln \sigma_2 - \frac{1}{2(1-\rho_0^2)}[\frac{S_{11}}{\sigma_1^2} - 2\rho_0 \frac{S_{12}}{\sigma_1\sigma_2} + \frac{S_{22^2}}{\sigma_2^2}].$$

Prilyginę funkcijos h išvestines pagal σ_1 ir σ_2 nuliui, gauname lygčių sistemą

$$\frac{S_{11}}{\sigma_1^2} - \rho_0 \frac{S_{12}}{\sigma_1\sigma_2} = n(1-\rho_0^2),$$

$$\frac{S_{22}}{\sigma_2^2} - \rho_0 \frac{S_{12}}{\sigma_1\sigma_2} = n(1-\rho_0^2).$$

Atėmę iš pirmosios lygties antrąją gauname, kad $S_{11}/\sigma_1^2 = S_{22}/\sigma_2^2$. Įrašę į pirmąją lygtį randame

$$\frac{S_{11}}{\sigma_1^2} - \rho_0 r \frac{S_{11}}{\sigma_1^2} \quad \Rightarrow \tilde{\sigma}_1^2 = \frac{S_{11}(1-\rho_0 r)}{n(1-\rho_0^2)} = \hat{\sigma}_1^2 \frac{1-\rho_0 r}{1-\rho_0^2}.$$

Analogiškai iš antrosios lygties gauname

$$\frac{S_{22}}{\sigma_2^2} - \rho_0 r \frac{S_{22}}{\sigma_2^2} \quad \Rightarrow \tilde{\sigma}_2^2 = \frac{S_{22}(1-\rho_0 r)}{n(1-\rho_0^2)} = \hat{\sigma}_2^2 \frac{1-\rho_0 r}{1-\rho_0^2}.$$

Naudodami šiuos įvertinius gauname tikėtinumo funkcijos sąlyginį maksimumą

$$L(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}) = \frac{\sqrt{2\pi}^{-n} n^n e^{-n} (1 - rho_0^2)^{n/2}}{(S_{11}S_{22})^{n/2}(1 - rho_0 r)^n}.$$

Tada

$$\Lambda^{2/n} = \frac{(1-\rho_0^2)(1-r^2)}{(1-\rho_0 r)^2}.$$

Tikėtinumų santykio kriterijus atmeta hipotezę, kai Λ įgyja mažas reikšmes (žr. [2], 4.6.1 skyrelį). Spręsdami nelygybę

$$(1-\rho_0^2)(1-r^2) < c(1-\rho_0 r)^2$$

r atžvilgiu gausime, kad hipotezė atmetama, kai

$$r < d_1(c) = \frac{c\rho_0 - (1-\rho_0^2)\sqrt{1-c}}{1-\rho_0^2(1-c)} \ \text{ arba } \ r > d_2(c) = \frac{c\rho_0 + (1-\rho_0^2)\sqrt{1-c}}{1-\rho_0^2(1-c)}.$$

Parametrą c reikia parinkti taip, kad

$$F(d_1(c)|\rho_0) + 1 - F(d_2(c)|\rho_0) = \alpha;$$

čia $F(x|\rho)$ yra statistikos r pasiskirstymo funkcija.

b) Hipotezę atmetame, kai $r < c_1$ arba $r > c_2$. Konstantos c_1 ir c_2 randamos iš sąlygų

$$F(c_1|\rho_0) = \alpha/2, \quad 1 - F(c_2|\rho_0) = \alpha/2,$$

arba P reikšmių terminais hipotezė atmetama, kai

$$pv = 2\max(F(r|\rho_0), 1 - F(r|\rho_0)) < \alpha,$$

čia r yra empirinio koreliacijos koeficiento realizacija.

c) Vienpusių alternatyvų $\bar{H}_1 : \rho > \rho_0$, $\bar{H}_2 : \rho < \rho_0$ atvejais natūralu naudoti vienpusius kriterijus. Hipotezė atmetama, kai atitinkamai

$$pv = 1 - F(r|\rho_0) < \alpha, \quad pv = F(r|\rho_0) < \alpha, \quad pv = 2\min(F(r|\rho_0), 1 - F(r|\rho_0)) < \alpha.$$

Šiose nelygybėse r yra empirinio koreliacijos koeficiento realizacija.

Įrodyta (žr. [12]), kad pateiktieji kriterijai yra TG tarp kriterijų, kurių statistikos invariantiškos poslinkio ir mastelio transformacijų atžvilgiu.

I.3.66. Praktiškai dažniau naudojami apytikslieji kriterijai, kurie grindžiami statistikos r Fišerio transformacija

$$Z = Z(r) = \frac{1}{2}\ln\frac{1+r}{1-r}$$

ir a. d. Z asimptotiniu ($n \to \infty$) normalumu

$$V_n(\rho_0) = \sqrt{n-3}\left(\frac{1}{2}\ln\frac{1+r}{1-r} - \frac{1}{2}\ln\frac{1+\rho_0}{1-\rho_0}\right) \xrightarrow{d} Y \sim N(0,\ 1),$$

kai tikroji parametro reikšmė yra ρ_0. Apytikslieji kriterijai gaunami lyginant $V_n(\rho_0)$ su standartinio normaliojo skirstinio kritinėmis reikšmėmis. Gauname tokias apytikslių kriterijų kritines sritis

$$V_n(\rho_0) > z_\alpha, \quad V_n(\rho_0) < -z_\alpha, \quad |V_n(\rho_0)| > z_{\alpha/2}.$$

Pažymėkime v statistikos $V_n(\rho_0)$ realizaciją. Tada kriterijus galima suformuluoti asimptotinių P reikšmių pv_a terminais: hipotezė atmetama, kai atitinkamai teisingos nelygybės

$$pv_a = 1 - \Phi(v) \le \alpha, \quad pv_a = \Phi(v) \le \alpha,$$
$$pv_a = 2\min(1 - \Phi(v), \Phi(v)) = 2(1 - \Phi(|v|)) \le \alpha.$$

I.3.67. Pažymėkime $Z_i = X_i - Y_i$, $i = 1, 2, ..., n$. Tada $\boldsymbol{Z} = (Z_1, ..., Z_n)^T$ yra paprastoji imtis a. d. $Z \sim N(\mu,\ \sigma^2)$, $\mu = \mu_1 - \mu_2$, $\sigma^2 = \boldsymbol{V}Z = \sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2$. Vietoje hipotezės H galime tikrinti analogišką hipotezę dėl normaliojo skirstinio vidurkio μ reikšmės naudodamiesi imtimi $(Z_1, ..., Z_n)^T$, kai dispersija σ^2 yra nežinoma (žr. **I.3.51**,**I.3.52** pratimus).

I.3.68. Jungtinės imties tankis (skaičiuojančiojo mato atžvilgiu) priklauso dviparametrei eksponentinio tipo skirstinių šeimai. Jį galime pertvarkyti taip:

$$f(\mathbf{X}, \boldsymbol{Y}|\lambda_1, \lambda_2) = \exp\{\theta U + \vartheta T - B(\theta, \vartheta)\}h(\mathbf{X}, \boldsymbol{Y}),$$

$$\theta = \ln\frac{\lambda_2}{\lambda_1}, \quad \vartheta = \ln\lambda_1, \quad U = \sum_i X_i, \quad T = \sum_i X_i + \sum_i Y_i,$$

$$B(\theta,\ \vartheta) = m\lambda_1 + n\lambda_2.$$

Remiantis [2], 4.4.2 teorema, egzistuoja TGN kriterijai hipotezėms dėl θ (kartu ir dėl santykio λ_1/λ_2) reikšmių tikrinti. Jie sudaromi kaip sąlyginiai paviršiuose $T = t$. Kadangi sąlyginis U skirstinys, kai $T = t$, $\lambda_1/\lambda_2 = c_0$, yra binominis $B(t, mc_0/(mc_0 + n))$, tai kriterijus hipotezei $H : \lambda_1/\lambda_2 = c_0$ tikrinti sudarome kaip ir kriterijus dėl binominio skirstinio tikimybės p reikšmių (žr. **I.3.35** pratimą), kur tikimybės p hipotetine reikšme reikia laikyti $p_0 = mc_0/(mc_0 + n)$. Tikrinant hipotezę reikia tarti, kad Bernulio eksperimentų skaičius t, o teigiamų įvykių skaičius yra U.

I.3.69. Turime dvi n. a. d. X ir Y, pasiskirsčiusių pagal Puasono dėsnį su parametrais λ_1 ir λ_2, imtis, kurių didumai $m = 2$ ir $n = 1$. Reikia patikrinti hipotezę $H_3 : \lambda_1/\lambda_2 = 1/5$, kai alternatyva $\bar{H}_3 : \lambda_1/\lambda_2 \neq 1/5$. Vietoje šios hipotezės tikriname hipotezę $H_3' : p = p_0 = 2/7$, kai alternatyva yra $\bar{H}_3' : p \neq 2/7$, o bandymų skaičius yra a. d. $T = X_1 + X_2 + Y$, jo realizacija $t = 73$; teigiamo įvykio įvykimų skaičiaus $U = X_1 + X_2$ realizacija $u = 28$. Randame tikimybės pasikliovimo intervalą, kurio pasikliovimo lygmuo $Q = 0,95$. Gauname $(\underline{p}, \overline{p}) = (0,272,\ 0,505)$. Kadangi $p_0 = 0,286$ patenka į šį intervalą, tai darome išvadą, kad turimi stebėjimai suformuluotai hipotezei neprieštarauja. Naudodami kriterijų P reikšmių terminais, randame $pv = 2\min(I_{5/7}(45,\ 29), I_{2/7}(28,\ 46)) = 2\min(0,9736,\ 0,0454) = 0,0908 > 0,05$. Hipotezė neatmetama.

I.3.70. a) A. d. X_1 sąlyginis skirstinys $(X_1|X_1 + X_2 = N) \sim B(N, p), p = \lambda_1/(\lambda_1 + \lambda_2)$. Hipotezę H_0 atmetame, kai $X_1 < k$, ir atmetame su tikimybe γ, kai $X_1 = k$; čia k yra mažiausias sveikasis skaičius, kuriam galioja nelygybė $\sum_{m=0}^{k-1} C_N^m/2^N = p \leq \alpha$; $\gamma = (\alpha - p)/(C_N^k/2^N)$; $N = X_1 + X_2$.

b) $\beta = \sum_{m=0}^{k-1} C_N^m 2^{N-m}/3^N + \gamma C_N^k 2^{N-k}/3^N$ pirmaisiais trimis atvejais ir

$$\beta = \sum_{m=0}^{k-1} C_N^m 4^{N-m}/5^N + \gamma C_N^k 4^{N-k}/5^N$$

paskutiniu atveju.

I.3.71. Jungtinės imties tankis (skaičiuojančiojo mato atžvilgiu) priklauso dviparametrei eksponentinio tipo skirstinių šeimai. Jį galime pertvarkyti taip:

$$f(\mathbf{X}, \boldsymbol{Y}, |p_1, p_2) = \exp\{\theta U + \vartheta T - B(\theta,\ \vartheta)\},$$

$$\theta = \ln\left(\frac{p_1(1 - p_2)}{(1 - p_1)p_2}\right), \quad \vartheta = \ln\frac{p_2}{1 - p_2}, \quad U = \sum_i X_i, \quad T = \sum_i X_i + \sum_i Y_i.$$

Remiantis [2], 4.4.1 teorema, egzistuoja TGN kriterijai hipotezėms dėl θ reikšmių tikrinti. Jie sudaromi kaip sąlyginiai paviršiuose $T = t$. Hipotezė $H : \theta = 0$ yra ekvivalenti hipotezei $H' : p_1 = p_2$. Nesunkiai patikriname, kad kai tikimybės vienodos $p_1 = p_2$, a. d. U sąlyginis skirstinys, kai $T = t$, yra hipergeometrinis $H(n + m,\ n,\ t)$.

Vadinasi, hipotezė H, kai alternatyva yra $\bar{H}_1 : p_1 > p_2$, taikant nerandomizuotą kriterijų, yra atmetama, kai $U \geq k'$; čia k' yra mažiausias sveikasis skaičius, tenkinantis nelygybę

$$1 - H(k' - 1|n + m,\ n,\ t) = \sum_{i=k'}^{\min(n,t)} \frac{C_n^i C_m^{t-i}}{C_{n+m}^t} \leq \alpha.$$

Norint įsitikinti, kad U pateko į kritinę sritį, galima apskaičiuoti P reikšmę, t. y. hipotezę atmesti, kai $1 - H(u-1|n+m,\ n,\ t) \leq \alpha$; čia u yra statistikos U realizacija.

Analogiškai, kai alternatyva yra $\bar{H}_2 : p_1 < p_2$, kritinė sritis gali būti užrašyta taip: $pv = H(u|n+m,\ n,\ t) \leq \alpha$.

Kai alternatyva $\bar{H}_3 : p_1 \neq p_2$ yra dvipusė, H atmetama, kai teisinga kuri nors iš pateiktų nelygybių, kuriose α reikia pakeisti į $\alpha/2$.

I.3.72. Tarkime, p_1 ir p_2 yra tikimybės, kad susirgs pirmosios ir antrosios grupių darbininkai. Reikia patikrinti hipotezę $H_1 : p_1 = p_2$, kai alternatyva yra $\bar{H}_1 : p_1 < p_2$. Parinkime reikšmingumo lygmenį $\alpha = 0,05$. Tada $n+m = 25$, $t = 6+4 = 10$, $u = 6$, $n = 20$. Naudojantis **I.3.71** pratimo kriterijumi, hipotezę reikia atmesti, jeigu suma

$$pv = H(u|n+m,\ n,\ t) = \sum_{i=\max(0,t-m)}^{u} \frac{C_n^i C_m^{t-i}}{C_{n+m}^t} \leq 0,05.$$

Tada (kadangi $t - m = 5$, tai suma turi tik du dėmenis)

$$pv = H(6|25,\ 10,\ 20) = \frac{C_{20}^5 C_5^5}{C_{25}^{10}} + \frac{C_{20}^6 C_5^4}{C_{25}^{10}} = 0,064.$$

Taigi iškeltosios hipotezės neatmetame.

I.3.73. Pažymėkime $S_i = \sum_{j=1}^{n_i} X_{ij}$, $i = 1,2$. Tikėtinumo funkcija

$$L = h(\mathbf{X}_1, \mathbf{X}_2) \exp\{-\theta S_1 - \vartheta(S_1 + S_2) - B(\theta, \vartheta)\}, \quad \theta = \theta_1 - \theta_2, \vartheta = \theta_2.$$

Kadangi esant teisingai hipotezei $H : \theta_1 = \theta_2$ (arba $\theta = 0$) statistika $S_1/(S_1+S_2) \sim Be(n_1\gamma_1, n_2\gamma_2)$, tai remdamiesi [2], 4.4.3 teorema TGN kriterijus galime suformuluoti naudodami beta skirstinį.

I.3.74. a) Atlikime transformaciją $U_i = \Delta_0 X_i + Y_i, Z_i = X_i - Y_i/\Delta_0, i = 1,...,n$. Tada hipotezė $H_0 : \sigma_2/\sigma_1 = \Delta_0$ yra ekvivalenti hipotezei, kad koreliacijos koeficientas $\rho = \rho(U_i, Z_i) = 0$. Statistika R yra empirinio koreliacijos koeficiento, apskaičiuoto pagal imtį $(U_i, Z_i)^T, i = 1,...,n$, modulis; b) $(n-2)R^2/(1-R^2) \sim F(1, n-2)$; c) Atlikime transformaciją $U_i = X_i + Y_i, Z_i = X_i - Y_i, i = 1,...,n$. Tada statistika $V = |\bar{Z}|/\sqrt{\sum_i (Z_i - \bar{Z})^2}$; d) $(n-1)V^2 \sim F(1, n-1)$.

I.3.75. a) Remiantis **I.2.13** pratimu $V/\theta \sim \chi^2(2n-2)$. Hipotezė atmetama, kai $V < c_1$ arba $V > c_2$. Lygtys konstantoms c_1, c_2 apskaičiuoti randamos analogiškai **I.3.30** pratimui. Imant simetriškus kriterijus hipotezė atmetama, kai $V < \chi^2_{1-\alpha/2}$ arba $V > \chi^2_{\alpha/2}$.

b) Remiantis **I.2.13** ir **I.2.14** pratimais $X_{(1)}$ ir V nepriklausomi; $X_{(1)} \sim \mathcal{E}(a, \theta/n)$. Jei teisinga hipotezė $H : a = 0$, tai $X_{(1)} \sim \mathcal{E}(\theta/n)$ ir $n(n-2)X_{(1)}/V \sim F(2, 2n-2)$. Hipotezė atmetama simetrišku reikšmingumo lygmens α kriterijumi, kai

$$\frac{n(n-2)X_{(1)}}{V} < F_{1-\alpha/2}(2, 2n-2), \quad \text{arba} \quad \frac{n(n-2)X_{(1)}}{V} > F_{\alpha/2}(2, 2n-2).$$

I.3.5 skyrelis

I.3.76. Pažymėkime $n = n_1 + ... + n_k$,

$$X_{i(1)} = \min(X_{i1}, ..., X_{in_i}), \quad X_{(1)} = \min(X_{1(1)}, ..., X_{k(1)}).$$

Tada

a) tikėtinumų santykis $\Lambda = \prod_{i=1}^{k} (\hat{\sigma}_i/\widetilde{\sigma}_i)^{n_i}$,

$$\hat{\sigma}_i = \sum_{j=1}^{n_i} (X_{ij} - X_{i(1)})/n_i, \quad \widetilde{\sigma}_i = \sum_{j=1}^{n_i} (X_{ij} - X_{(1)})/n_i;$$

b) $\Lambda = \prod_{i=1}^{k} (\hat{\sigma}_i/\hat{\sigma})^{n_i}$, $\hat{\sigma} = \sum_{i=1}^{k} n_i \hat{\sigma}_i/n$; c) $\Lambda = (\hat{\sigma}/\widetilde{\sigma})^n$, $\widetilde{\sigma} = \sum_{i=1}^{k} n_i \widetilde{\sigma}_i/n$.

I.3.77. a) Tikėtinumų santykis

$$\Lambda = \frac{\max_{\mu=\mu_0,\sigma^2} L(\mu, \sigma^2)}{\max_{\mu,\sigma^2} L(\mu, \sigma^2)} = \frac{L(\mu_0, \tilde{\sigma^2})}{L(\hat{\mu}, \hat{\sigma^2})} = \left(\frac{\hat{\sigma^2}}{\tilde{\sigma^2}}\right)^{n/2};$$

čia $\hat{\sigma^2} = m_2 = (n-1)s^2/n = \sum_i (X_i - \bar{X})^2/n$ yra dispersijos σ^2 DT įvertinys, o $\tilde{\sigma^2} = \sum_i (X_i - \mu_0)^2/n$ yra dispersijos σ^2 DT įvertinys, kai $\mu = \mu_0$. Gauname

$$\Lambda^{2/n} = \frac{1}{1 + n(\bar{X} - \mu_0)^2/((n-1)s_2)} = \frac{1}{1 + t^2(\mathbf{X})/(n-1)}.$$

Taigi

$$\{\mathbf{X} : \Lambda < c\} \quad \Leftrightarrow \quad \{\mathbf{X} : |t(\mathbf{X})| > t_{\alpha/2}(n-1)\}.$$

b) Tikėtinumų santykis

$$\Lambda = \frac{\max_{\mu,\sigma^2=\sigma_0^2} L(\mu, \sigma^2)}{\max_{\mu,\sigma^2} L(\mu, \sigma^2)} = (m_2/\sigma_0^2)^{n/2} e^{-nm_2/(2\sigma_0^2)+n/2}.$$

Funkcija $h(x) = x^{n/2}e^{-nx/2}$ didėjanti intervale $(0,\ 1)$ ir mažėjanti intervale $1, \infty$. Gauname

$$\Lambda < c \quad \Leftrightarrow \quad \frac{nm_2}{\sigma_0^2} < c_1 \text{ arba } \frac{nm_2}{\sigma_0^2} > c_2.$$

I.3.78. Tikėtinumų santykis

$$\Lambda = \frac{\max_{\pi_i=\pi_i^0} L(\pi_1, ..., \pi_k)}{\max_{\pi_i} L(\pi_1, ..., \pi_k)} = \frac{L(\pi_1^0, ..., \pi_k^0)}{L(\hat{\pi}_1, ..., \hat{\pi}_k)} = \frac{\prod_{i=1}^{k} (\pi_i^0)^{V_i}}{\prod_{i=1}^{k} (V_i/n)^{V_i}};$$

čia $\hat{\pi}_i = V_i/n$ yra parametro π_i DT įvertinys (žr. **I.2.131** pratimą). Gauname

$$\Lambda = \left(\prod_{i=1}^{k} (\pi_i^0/\hat{\pi}_i)^{\hat{\pi}_i}\right)^n.$$

I.3.79. Remiantis [2], 3.5.4 skyreliu $-2\ln\Lambda \xrightarrow{d} \chi_{k-1}^2$. Skleidžiant $-2\ln\Lambda$ eilute (žr. [4], 2.2.1 teoremą) gaunama, kad $-2\ln\Lambda = \sum_i (V_i - n\pi_i^0)^2/(n\pi_i^0) + o_P(1)$.

I.3.80. a) Tikėtinumų santykis

$$\Lambda = \frac{\max_{\mu_1=\mu_2,\sigma^2} L(\mu_1, \mu_2, \sigma^2)}{\max_{\mu_1,\mu_2,\sigma^2} L(\mu_1, \mu_2, \sigma^2)} = \frac{L(\tilde{\mu}_1, \tilde{\mu}_2, \tilde{\sigma^2})}{L(\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma^2})} = \left(\frac{\hat{\sigma}^2}{\tilde{\sigma}^2}\right)^n,$$

čia $\hat{\mu}_1 = \bar{X}, \hat{\mu}_2 = \bar{Y}$, $\hat{\sigma^2} = [\sum_i (X_i - \bar{X})^2 + \sum_i (Y_i - \bar{Y})^2]/(2n)$ yra parametrų μ_1, μ_2, σ^2 DT įvertiniai, o $\tilde{\mu}_1 = \tilde{\mu}_2 = (\bar{X} + \bar{Y})/2$,

$$\tilde{\sigma}^2 = [\sum_i (X_i - \tilde{\mu}_1)^2 + \sum_i (Y_i - \tilde{\mu}_2)^2]/(2n) = (n-1)(s_1^2 + s_2^2)/(2n) + n(\bar{X} - \bar{Y})^2/4$$

yra parametrų μ_1, μ_2, σ^2 DT įvertiniai, kai $\mu_1 = \mu_2$.

Vienodo didumo imčių atveju Stjudento statistika

$$T = T(\mathbf{X}, \boldsymbol{Y}) = \sqrt{n}(\bar{X} - \bar{Y})/\sqrt{s_1^2 + s_2^2} \sim S(2n-2),$$

kai hipotezė teisinga. Gauname

$$\Lambda^{1/n} = \frac{1}{1 + nT^2(\mathbf{X}, \boldsymbol{Y})/(2(n-1))}.$$

Tikėtinumų santykis Λ yra monotoniškai mažėjantis T^2 atžvilgiu. Taigi

$$\Lambda < c \quad \Leftrightarrow \quad |T(\mathbf{X}, \boldsymbol{Y})| > t_{\alpha/2}(2(n-1)).$$

b) Tikėtinumų santykis

$$\Lambda = \frac{\max_{\mu_1,\mu_2,\sigma_1^2=\sigma_2^2} L(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)}{\max_{\mu_1,\mu_2,\sigma_1^2,\sigma_2^2} L(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)} = \frac{(\hat{\sigma}_1^2)^{n/2}(\hat{\sigma}_2^2)^{n/2}}{((\hat{\sigma}_1^2 + \hat{\sigma}_2^2)/2)^n};$$

čia $\hat{\sigma^2}_1 = \sum_i (X_i - \bar{X})^2/n = (n-1)s_1^2/n$, $\hat{\sigma^2}_2 = \sum_i (Y_i - \bar{Y})^2/n = (n-1)s_2^2/n$ yra parametrų σ_1^2 ir σ_2^2 DT įvertiniai. Fišerio statistika $F = s_1^2/s_2^2 \sim F(n-1, n-1)$, kai hipotezė teisinga. Gauname

$$\Lambda^{1/n} = \frac{2s_1 s_2}{s_1^2 + s_2^2} = 2\frac{\sqrt{F}}{1+F}.$$

Nelygybė $\Lambda < c$ ekvivalenti nelygybei $2\sqrt{F}/(1+F) > d, d > 0$. Funkcija $h(x) = 2x/(1+x^2)$, kai $x > 0$, didėja nuo 0 iki 1, kai x kinta nuo 0 iki 1, ir mažėja nuo 1 iki 0, kai x kinta nuo 1 iki ∞. Taigi

$$\Lambda < c \quad \Leftrightarrow \quad F < c_1 \text{ arba } F > c_2.$$

I.3.81. Tikėtinumų santykis $\Lambda = [1/(1 + kF/(n-k))]^{n/2}$.

I.3.82. Tikėtinumų santykis

$$\Lambda = \hat{p}^{S_1}(1-\hat{p})^{n_1-S_1}\hat{p}^{S_2}(1-\hat{p})^{n_2-S_2}/[\hat{p}_1^{S_1}(1-\hat{p}_1)^{n_1-S_1}\hat{p}_2^{S_2}(1-\hat{p}_2)^{n_2-S_2}];$$

$S_1 = X_1+...+X_{n_1},\ S_2 = Y_1+...+Y_{n_2},\ \hat{p}_1 = S_1/n_1,\ \hat{p}_2 = S_2/n_2,\ \hat{p} = (S_1+S_2)/(n_1+n_2)$. Remiantis **I.3.79** pratimu $-2\ln\Lambda \xrightarrow{d} \chi_1^2$.

I.3.83. $\Lambda = \hat{p}^S(1-\hat{p})^{n-S}/(\prod_{i=1}^k \hat{p}_i^{S_i}(1-\hat{p}_i)^{n_i-S_i})$, $S = S_1+...+S_k$, $n = n_1+...+n_k$, $\hat{p} = S/n$, $\hat{p}_i = S_i/n_i$, $i = 1,2,...,k$; $-2\ln\Lambda \xrightarrow{d} \chi_{k-1}^2$, kai $n_i \to \infty$, $i = 1,2,...,k$, ir hipotezė $H : p_1 = p_2 = ... = p_k$ yra teisinga.

I.3.84. Tikėtinumų santykis $\Lambda = (\hat{\lambda}^{S_1+S_2}/(\hat{\lambda}_1^{S_1}\hat{\lambda}_2^{S_2}))\exp\{n_1\hat{\lambda}_1 + n_2\hat{\lambda}_2 - (n_1+n_2)\hat{\lambda}\}$; $S_1 = X_1+...+X_{n_1},\ S_2 = Y_1+...+Y_{n_2},\ \hat{\lambda}_1 = S_1/n_1,\ \hat{\lambda}_2 = S_2/n_2,\ \hat{\lambda} = (S_1+S_2)/(n_1+n_2)$. Remiantis **I.3.79** pratimu $2\ln\Lambda \xrightarrow{d} \chi_1^2$.

I.3.85. $\Lambda = \hat{\lambda}^S/(\prod_{i=1}^k \hat{\lambda}_i^{S_i})$, $S = S_1+...+S_k$, $n = n_1+...+n_k$, $\hat{\lambda} = S/n$, $\hat{\lambda}_i = S_i/n_i$, $i = 1,2,...,k$; $-2\ln\Lambda \xrightarrow{d} \chi_{k-1}^2$, kai $n_i \to \infty$, $i = 1,2,...,k$, ir hipotezė $H : \lambda_1 = \lambda_2 = ... = \lambda_k$ yra teisinga.

I.3.86. A. d. $\bar{X}/\bar{Y} \sim (\theta_1/\theta_2)F_{2n_1,2n_2}$; čia $F_{m,n}$ – a. d., turintis Fišerio skirstinį su m ir n laisvės laipsnių. Tikėtinumų santykis

$$\Lambda = \frac{\max_{\theta_1=\theta_2} L(\theta_1,\theta_2)}{\max_{\theta_1,\theta_2} L(\theta_1,\theta_2)} = \frac{L(\tilde{\theta},\tilde{\theta})}{L(\hat{\theta_1},\hat{\theta_2})},$$

čia $\hat{\theta_1} = \bar{X}, \hat{\theta}_2 = \bar{Y}$ yra parametrų θ_1, θ_2 DT įvertiniai, $\tilde{\theta} = (n_1\bar{X} + n_2\bar{Y})/(n_1+n_2)$ yra parametro θ įvertinys, kai $\theta_1 = \theta_2 = \theta$. Gauname

$$\Lambda = \frac{(n_1+n_2)^{n_1+n_2}}{n_1^{n_1}n_2^{n_2}} \frac{(n_1\bar{X})^{n_1}(n_2\bar{Y})^{n_2}}{(n_1\bar{X}+n_2\bar{Y})^{n_1+n_2}} = (\frac{n_1+n_2}{n_2})^{n_1+n_2} \frac{F^{n_1}}{1+n_1F/n_2}^{n_1+n_2}$$

$$\Lambda < c \quad \Leftrightarrow \quad F < c_1 \quad \text{arba} \quad F > c_2.$$

Tegu alternatyva yra $\bar{H} : \theta_1 > \theta_2$. Tada hipotezė H atmetama, kai

$$\bar{X}/\bar{Y} > F_\alpha(2n_1, 2n_2),$$

o kriterijaus galia $\beta(\theta_2) = \mathbf{P}\{F_{2n_1,2n_2} > (\theta_2/\theta_1)F_\alpha(2n_1,2n_2)\}$, $\theta_2 < \theta_1$.

I.3.87. *Nurodymas.* Pasinaudokime tuo, kad a. d. $V = \sum_i (m_i - np_i^0)^2/(np_i^0)$ apytiksliai turi χ^2 skirstinį su 3 laisvės laipsniais; čia p_i^0 – tikimybė, kad a. d. $X \sim N(0, 1/4)$ įgijo reikšmę iš i-ojo intervalo: $p_1^0 = 0,00135; p_2^0 = 0,1573; p_3^0 = 0,6827; p_4^0 = 0,15865$. Atsitiktinis dydis V įgijo reikšmę 2,6578. Asimptotinė P-reikšmė $pv_a = \mathbf{P}\{\chi_3^2 > 2,6578\} = 0,4474$. Atmesti hipotezę nėra pagrindo.

I.3.88. Analogiškai kaip ir **I.3.87** pratime gauname, kad a. d. V, kuris apytiksliai turi χ^2 skirstinį su 3 laisvės laipsniais, įgijo reikšmę 19,453. Asimptotinė P-reikšmė $pv_a = \mathbf{P}\{\chi_3^2 > 19,453\} = 0,00022$. Hipotezę atmetame.

I.3.89. Atsitiktinis dydis V, kuris apytiksliai turi χ^2 skirstinį su 2 laisvės laipsniais, įgijo reikšmę 5,3446. Asimptotinė P-reikšmė $pv_a = \mathbf{P}\{\chi_2^2 > 5,3446\} = 0,0691$. Hipotezę atmetame, jei kriterijaus reikšmingumo lygmuo viršija 0,0691. Tikrindami hipotezę $H : X \sim B(2,p)$, apskaičiuojame statistikos $\hat{V} = (m_0 - n\hat{p}^2)^2/(n\hat{p}^2) + (m_1 - n2\hat{p}\hat{q})^2/(n2\hat{p}\hat{q}) + (m_2 - n\hat{q}^2)^2/(n\hat{q}^2)$, kuri apytiksliai turi χ^2 skirstinį su 1 laisvės laipsniu, reikšmę; čia $\hat{p}$ yra tikimybės p DT įvertinys. Gauname $\hat{p} = = (1017 + 1054)/4040 = 0,5126$; $\hat{V} =$

$0,2317$. Asimptotinė P-reikšmė $pv_a = \mathbf{P}\{\chi_1^2 > 0,2317\} = 0,6303$. Atmesti hipotezę nėra pagrindo.

I.3.90. Parametro p DT įvertis $\hat{p} = 0,2675$. Analogiškai kaip **I.3.89** pratime statistikos $\hat{V}$, kuri esant teisingai hipotezei apytiksliai turi χ^2 skirstinį su 1 laisvės laipsniu, reikšmė yra $\hat{V} = 0,757$. Asimptotinė P-reikšmė $pv_a = \mathbf{P}\{\chi_1^2 > 0,757\} = 0,3843$. Todėl atmesti hipotezę nėra pagrindo.

I.3.91. Parametro p DT įvertis $\hat{p} = 0,1232$. Analogiškai kaip **I.3.89** pratime statistikos $\hat{V}$, kuri esant teisingai hipotezei apytiksliai turi χ^2 skirstinį su 1 laisvės laipsniu, reikšmė lygi $0,444$. Todėl atmesti hipotezę nėra pagrindo.

I.3.6 skyrelis

I.3.92. $n \geq 74$.

I.3.93. $\leq 0,6737$; $n \geq 133$.

I.3.94. Taikant TGN kriterijų $n \geq 45$, taikant simetrinį kriterijų $n \geq 47$.

I.3.95. Pereiname prie logaritmų. Kadangi atmesti dispersijų lygybės hipotezę nėra pagrindo, tai taikome kriterijų esant vienodoms dispersijoms. Statistikos T realizacija $t = 4,2745$; P reikšmė $pv = 9,45 \times 10^{-5}$. Hipotezė atmetama.

I.3.96. Pagal turimus duomenis randame vidurinio laikotarpio vidurkio ir kvadratinio nuokrypio įverčius $\bar{X}_0 = 16,563, s_0 = 1,5362$ ir taikome Stjudento dviejų imčių kriterijų. Lygindami pirmą ir vidurinį laikotarpius, gauname statistikos T realizaciją 0,8761, o lyginant vidurinį ir paskutinį laikotarpius statistikos T realizacija yra $-1,5652$; atitinkamos P reikšmės esant dvipusei alternatyvai yra 0,3841 ir 0,1222. Atmesti hipotezę nėra pagrindo.

I.3.97. Neatmetama.

I.3.98. $n \geq 211$; $n \geq 74$; $n \geq 31$.

I.3.99. Dispersijų lygybės hipotezė neatmetama, nes a. d., turintis Fišerio skirstinį $F(7,9)$, įgijo reikšmę 1,9584; sisteminių paklaidų vienodumo hipotezė neatmetama, nes a. d., turintis Stjudento skirstinį $S(16)$, įgijo reikšmę 0,4099.

I.3.100. Tikriname hipotezę $H : \sigma_1^2 = \sigma_2^2$, kai alternatyva yra $\bar{H} : \sigma_1^2 < \sigma_2^2$. Statistika $F = s_1^2/s_2^2$, kuri esant teisingai hipotezei turi Fišerio skirstinį su $\nu_1 = 10$ ir $\nu_2 = 10$ laisvės laipsnių, įgijo reikšmę 0,1783. P reikšmė $pv = \mathbf{P}\{F_{10;10} < 0,1783\} = 0,0059$. Hipotezę atmetame, kai $\alpha \geq 0,0059$.

I.3.101. a) Tikėtinumų santykio statistika $\tilde{R}_{TS}$ įgijo reikšmę 1,8931; asimptotinė P reikšmė $pv_a = \mathbf{P}\{\chi_2^2 > 1,8931\} = 0,3881$; atmesti hipotezę nėra pagrindo. b) Hipotezė neatmetama.

I.3.102. Tikėtinumų santykio statistika įgijo reikšmę 0,3086; atmesti hipotezę nėra pagrindo.

I.3.103. a) Statistika Z įgijo reikšmę 1,882. Tegu $F(x|\nu)$ yra Stjudento skirstinio su ν laisvės laipsnių pasiskirstymo funkcija. Tada P-reikšmė $pv = 2[1 - F(1,882|18)] = 0,0761$. Reikšmingumo lygmens $\alpha = 0,05$ kriterijumi hipotezė neatmetama. b) Statistika T įgijo reikšmę 3,2143. P reikšmė $pv = 2[1 - F(3,2143|8)] = 0,0106$; hipotezę atmetame, kai kriterijaus reikšmingumo lygmuo α viršija 0,0106. c) Pagal prasmę a. d.X ir Y yra teigiamai priklausomi (jei gretimų sklypelių žemė derlingesnė, tai abu a. d. X ir Y turės tendenciją įgyti didesnes reikšmes ir atvirkščiai). Todėl lentelės duomenis reikia traktuoti kaip 10 dvimačio a. v. $(X, Y)^T$ realizacijų.

I.3.104. Koreliacijos koeficiento įvertis $\hat{\rho} = r = 0,7055$. Statistika U įgijo reikšmę 2,8157. Jeigu hipotezės $H : \rho = 0$ alternatyva yra $\bar{H} : \rho > 0$, tai P reikšmė $pv = 1-F(2,8157|8) = 0,0113$; nepriklausomumo hipotezė $H : \rho = 0$ atmetama, kai kriterijaus reikšmingumo lygmuo $\alpha \geq 0,0113$.

I.3.105. Statistika T įgijo reikšmę 1,7718. Jeigu alternatyva yra $\bar{H} : \theta >0$, tai P reikšmė $pv = 1 - F(1,7718|15) = 0,0484$. Hipotezė $H : \theta = 0$ atmetama, kai kriterijaus reikšmingumo lygmuo $\alpha \geq 0,0484$.

I.3.106. Tikriname hipotezę dėl a. d. X ir Y vidurkių lygybės. Statistika T įgijo reikšmę 4,1212. Jeigu alternatyva dvipusė, tai P reikšmė $pv = 2[1 - F(4,1212|9) = 0,0026$. Hipotezė $H : \theta = 0$ atmetama, kai kriterijaus reikšmingumo lygmuo $\alpha \geq 0,0026$.

I.3.107. Statistika $T = X_1 + ... + X_n$ įgijo reikšmę $t = 185$. P reikšmė $pv = \mathbf{P}\{\chi^2_{400} < 2t\} = 0,1435$. Hipotezę atmesti nėra pagrindo.

I.3.108. Hipotezė neatmetama. Pasikliovimo intervalas uždengia hipotetinę reikšmę λ_0.

I.3.109. Hipotezė neatmetama. Pasikliovimo intervalas uždengia hipotetinę reikšmę λ_0.

I.3.110. a) Tikėtinumų santykio statistika R_{TS} įgijo reikšmę 19,3386; asimptotinė P reikšmė $pv_a = \mathbf{P}\{\chi^2_6 > 19,3386\} = 0,0036$. Hipotezė atmetama, jei kriterijaus reikšmingumo lygmuo viršija 0,0036. b) Tikėtinumų santykio statistika R_{TS} įgijo reikšmę 5,1783; asimptotinė P reikšmė $pv_a = \mathbf{P}\{\chi^2_4 > 5,1783\} = 0,2695$. Hipotezė neatmetama.

I.3.111. Antros rūšies gaminių skaičius 277; nerandomizuoto kriterijaus P reikšmė $pv = \mathbf{P}\{S \geq 277\}$; čia $S \sim B(1000; 0,25)$; randame $pv = 0,0275$. Hipotezė atmetama, jei kriterijaus reikšmingumo lygmuo viršija 0,0275.

I.3.112. Tikrinama hipotezė $H : \lambda_1 = \lambda_2$. Jei alternatyva yra $\bar{H} : \lambda_1 > \lambda_2$, tai P reikšmė $pv = \mathbf{P}\{S \geq 150\}$; čia $S \sim B(400; 1/3)$. Randame $pv = 0,0442$. Hipotezė atmetama, jei kriterijaus reikšmingumo lygmuo viršija 0,0442.

I.3.113. P reikšmė $pv = \mathbf{P}\{S \leqslant 75\}$; čia $S \sim B(275; 1/3)$. Randame $pv = 0,0181$. Hipotezė atmetama, jei kriterijaus reikšmingumo lygmuo viršija 0,0181.

I.3.114. Tikrinama hipotezė $H : \lambda_1 = \lambda_2$. Jei alternatyva yra $\bar{H} : \lambda_1 > \lambda_2$, tai P reikšmė $pv = \mathbf{P}\{S \geq 20026\}$; čia $S \sim B(39606; 1/2)$. Randame $pv = 0,0123$. Hipotezė atmetama, jei kriterijaus reikšmingumo lygmuo viršija 0,0123.

I.3.115. Statistika $\sqrt{n}(\hat{p}_1 - \hat{p}_2)/\sqrt{\hat{p}_1\hat{q}_1 + \hat{p}_2\hat{q}_2}$, kuri esant teisingai hipotezei asimptotiškai turi standartinį normalųjį skirstinį, įgijo reikšmę 4,0834; asimptotinė P reikšmė $pv_a = 2(1 - \Phi(4,0834)) = 4,44 \times 10^{-5}$. Hipotezė atmetama.

I.3.116. Tikėtinumų santykio statistika R_{TS} įgijo reikšmę 9,3113; asimptotinė P reikšmė $pv_a = \mathbf{P}\{\chi^2_4 > 9,3113\} = 0,0538$. Hipotezė atmetama, jei kriterijaus reikšmingumo lygmuo viršija 0,0538.

I.3.117. Statistika $2n\bar{R}^2$ (žr. [2], 4.7.12 skyrelį) įgijo reikšmę 43,888; asimptotinė P reikšmė $pv_a = \mathbf{P}\{\chi^2_2 > 43,888\} = 2,9 \cdot 10^{-10}$; statistika $-2\ln\Lambda$ įgijo reikšmę 55,099; asimptotinė P reikšmė $pv_a = \mathbf{P}\{\chi^2_{11} > 55,099\} = 7,4 \cdot 10^{-8}$. Hipotezė atmetama.

I.3.118. Statistika $2n\bar{R}^2$ įgijo reikšmę 10,2248; asimptotinė P reikšmė $pv_a = \mathbf{P}\{\chi^2_2 > 10,2248\} = 0,006$; statistika $-2\ln\Lambda$ įgijo reikšmę 20,0797; asimptotinė P reikšmė $pv_a = \mathbf{P}\{\chi^2_{11} > 20,0797\} = 0,0443$. Reikšmingumo lygmens $\alpha = 0,05$ kriterijumi hipotezė atmetama.

I.3.119. Hipotezė atmetama, nes statistika $2n\bar{R}^2$ įgijo reikšmę 112,8, o statistika $-2\ln\Lambda$ įgijo reikšmę 137,9.

II. Tiesiniai modeliai

II.1. Gauso ir Markovo tiesinis modelis

II.1.1. Mažiausiųjų kvadratų metodas

II.1.1. Turime tiesinį Gauso ir Markovo modelį, t. y. imtis $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ turi tokią struktūrą

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{\beta} + \boldsymbol{e},$$

čia $\boldsymbol{A} = [a_{ij}]_{n\times m}$ eksperimentų plano matrica su žinomais elementais a_{ij}; $\boldsymbol{\beta} = (\beta_1, ..., \beta_m)^T$ nežinomų parametrų vektorius; $\boldsymbol{e} = (e_1, ..., e_n)^T$ paklaidų vektorius, $\mathbf{E}(\boldsymbol{e}) = \mathbf{0}, \boldsymbol{V}(\boldsymbol{e}) = \sigma^2\boldsymbol{I}$. Tegu matricos $\boldsymbol{A}$ rangas lygus m. a) Raskite parametro $\boldsymbol{\beta}$ mažiausiųjų kvadratų (MK) įvertinį $\hat{\boldsymbol{\beta}}$. b) Raskite įvertinio $\hat{\boldsymbol{\beta}}$ pirmuosius du momentus. c) Raskite k-mačio parametro $\boldsymbol{\theta} = \boldsymbol{H}\boldsymbol{\beta}$ įvertinio $\hat{\boldsymbol{\theta}} = \boldsymbol{H}\hat{\boldsymbol{\beta}}$ vidurkį ir kovariacijų matricą (čia $\boldsymbol{H}$ yra dimensijos $k \times m$ rango $k \leq m$ matrica).

II.1.2. (**II.1.1** pratimo tęsinys). a) Įrodykite, kad liekamoji kvadratų suma

$$SS_E = \min_{\boldsymbol{\beta}} SS(\boldsymbol{\beta}) = (\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}})^T(\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}})$$

gali būti surasta tokiu būdu

$$SS_E = \boldsymbol{Y}^T\boldsymbol{Y} - \hat{\boldsymbol{\beta}}^T\boldsymbol{A}^T\boldsymbol{Y}.$$

b) Raskite kvadratų sumos SS_E vidurkį ir parametro σ^2 nepaslinktąjį įvertinį.

II.1.3. Norint ištirti kviečių derlingumo priklausomybę nuo jų veislės, atliekamas toks eksperimentas. Atsitiktinai parinktų n_1 sklypelių apsėjama pirmąja kviečių veisle; n_2 – antrąja ir t. t., ir pagaliau n_m sklypelių – m-ąja veisle. Pažymėkime Y_{ij} i-osios kviečių veislės derlingumą j-ajame sklypelyje. Tarkime, kad derlingumo pokyčiai pereinant nuo vienos kviečių veislės prie kitos nekeičia dispersijos, o gali keisti tik vidurkį. a) Užrašykite imtį tiesinio modelio (žr. **I.1.1** pratimą) pavidalu. b) Raskite modelio parametrų įvertinius.

II.1.4. Norint ištirti vyrų sistolinio kraujo spaudimo Y priklausomybę nuo jų svorio X_1 ir amžiaus X_2, atsitiktinai atrenkama n vyrų ir jiems pamatuojamos a. v. $(Y, X_1, X_2)^T$ reikšmės $(Y_i, x_{1i}, x_{2i})^T$, $i = 1, ..., n$. a) Užrašykite imtį tiesinio modelio (žr. **II.1.1** pratimą) pavidalu. b) Raskite modelio parametrų įvertinius.

II.1.5. Tarkime, imties $\boldsymbol{Y} = (Y_1, Y_2, Y_3, Y_4)^T$ elementai turi tokią struktūrą $Y_1 = \alpha - \beta_2 + e_1$, $Y_2 = \alpha + 2\beta_1 - \beta_2 + e_2$, $Y_3 = \alpha + \beta_2 + e_3$, $Y_4 = \alpha - 2\beta_1 + \beta_2 + e_4$;

čia $e_1, ..., e_4$ yra n. a. d., turintys nulinius vidurkius ir vienodas dispersijas σ^2. Raskite parametro $\boldsymbol{\beta} = (\alpha, \beta_1, \beta_2)^T$ MK įvertinį ir jo realizaciją (įvertį), kai a. v. $\boldsymbol{Y}$ realizacija yra $(-1,9; 2,05; 3,15; 0,1)^T$.

II.1.6. (**II.1.5** pratimo tęsinys). Raskite liekamosios kvadratų sumos SS_E realizaciją.

II.1.7. (**II.1.5** pratimo tęsinys). Raskite įvertinio $\hat{\boldsymbol{\beta}}$ kovariacijų matricą ir dispersijos σ^2 įvertį.

II.1.8. (**II.1.5** pratimo tęsinys). Tarkime, kad domina parametras $\boldsymbol{\theta} = (\theta_1, \theta_2)^T$, $\theta_1 = 2\alpha - \beta_1$, $\theta_2 = \alpha + \beta_1 - \beta_2$. Raskite parametro $\boldsymbol{\theta}$ įvertinį ir jo kovariacijų matricą.

II.1.9. Tegu $Y_1 = \alpha + e_1$, $Y_2 = 2\alpha - \beta + e_2$, $Y_3 = \alpha + 2\beta + e_3$; čia $\{e_i\}$ nepriklausomi a. d., $\mathbf{E}(e_i) = 0$, $\boldsymbol{V}(e_i) = \sigma^2$. Raskite parametrų α ir β mažiausiųjų kvadratų įvertinius ir jų dispersijas.

II.1.10. Turime tiesinį modelį

$$\mathbf{E}(Y_i) = \beta_0 + \beta_1 x_i + \beta_2(3x_i^2 - 2), \quad i = 1, 2, 3;$$

čia $x_1 = -1$, $x_2 = 0$, $x_3 = 1$. Raskite parametrų $\beta_0, \beta_1, \beta_2$ įvertinius. Įrodykite, kad parametrų β_0, β_2 įvertiniai modelyje, kuriame $\beta_1 = 0$, turi tą patį pavidalą.

II.1.11. Parametrams α ir β įvertinti turime stebėjimus: m stebėjimų a. d. Y_1, kurio $\mathbf{E}(Y_1) = \alpha$; m stebėjimų a. d. Y_2, kurio $\mathbf{E}(Y_2) = \alpha + \beta$, ir n stebėjimų a. d. Y_3, kurio $\mathbf{E}(Y_3) = \alpha - 2\beta$. Stebėjimų paklaidos nekoreliuotos ir turi vienodas dispersijas. Įrodykite, kad mažiausiųjų kvadratų įvertiniai $\hat{\alpha}$ ir $\hat{\beta}$ nekoreliuoti, kai $m = 2n$.

II.1.12. Mažiausiųjų kvadratų metodu parenkami pirmojo ir antrojo laipsnio polinomai pagal didumo n imtį $(X_i,\ Y_i)^T$, $i = 1, 2, ..., n$. Tegu ω ir Ω yra šitokios prielaidos:

$$\omega: \ Y_i = \alpha + \beta X_i + e_i,$$

$$\Omega: \ Y_i = \alpha + \beta X_i + \gamma X_i^2 + e_i';$$

čia e_i, e_i' – n. a. d. su nuliniais vidurkiais ir vienodomis dispersijomis σ^2. Sudarykite normaliųjų lygčių sistemas ir raskite parametrų α, β ir α, β, γ mažiausiųjų kvadratų įvertinius.

II.1.13 (**II.1.12** pratimo tęsinys). Raskite parametrų įvertinių dispersijas. Įrodykite: jeigu prielaidoje ω vietoje $\alpha + \beta X_i$ imsime $\delta + \beta(X_i - \bar{X})$, tai $\mathbf{Cov}\,(\hat{\delta},\ \hat{\beta}) = 0$.

II.1.14. Gauti nepriklausomi ir vienodo tikslumo keturkampio ABCD kampų ABD, DBC, ABC, BCD, CDB, BDA, CDA, DAB matavimai (laipsniais): 50,78; 30,25; 78,29; 99,57; 50,42; 40,59; 88,87; 89,96. Raskite kampų β_1 =ABD, β_2 =DBC, β_3 =CDB, β_4 =BDA mažiausiųjų kvadratų įverčius ir matavimo paklaidos dispersijos σ^2 nepaslinktąjį įvertį.

II.1.15. $\hat{\theta}_1, ..., \hat{\theta}_k$ yra vienmačio parametro θ nepaslinktieji įvertiniai ir $\mathbf{Cov}\,(\hat{\theta}_i,\ \hat{\theta}_j) = \sigma_{ij}$. Raskite tiesinę $\hat{\theta}_1, ..., \hat{\theta}_k$ funkciją, kuri būtų nepaslinktasis θ įvertinys ir turėtų minimalią dispersiją. Raskite tos dispersijos reikšmę.

II.1.16. (**II.1.15** pratimo tęsinys). Tegu $\mathbf{Cov}\,(\hat{\theta}_i,\ \hat{\theta}_j) = 0$, $i \neq j$, $\boldsymbol{V}\hat{\theta}_i = \sigma_i^2$, $i = 1, ..., k$. Įrodykite, kad $\boldsymbol{V}(c_1\hat{\theta}_1 + ... + c_k\hat{\theta}_k)$, $c_1 + ... + c_k = 1$, yra minimali, kai $c_i = \sigma_i^{-2}/(\sigma_1^{-2} + ... + \sigma_k^{-2})$, ir raskite tos dispersijos reikšmę.

II.1.17. Įrodykite, kad jeigu $\hat{\theta}_1, ..., \hat{\theta}_k$ yra parametrų $\theta_1, ..., \theta_k$ nepriklausomi NMD įvertiniai, tai $c_1\hat{\theta}_1 + ... + c_k\hat{\theta}_k$ yra parametro $\theta = c_1\theta_1 + ... + c_k\theta_k$ NMD įvertinys.

II.1.18. $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. X, kurio $\mathbf{E}X = \mu$ ir $\boldsymbol{V}X = \sigma^2$. Užrašykite stebėjimus kaip tiesinį modelį. Raskite parametro μ mažiausiųjų kvadratų įvertinį.

II.1.2. Normaliojo skirstinio atvejis

II.1.19. Tarkime, **II.1.1** pratimo sąlygomis papildomai priimama prielaida, kad paklaidų vektorius $\boldsymbol{e} \sim N_n(\mathbf{0}, \sigma^2\boldsymbol{I})$ (arba $\boldsymbol{Y} \sim N_n(\boldsymbol{A\beta}, \sigma^2\boldsymbol{I})$. a) Įrodykite, kad MK metodas parametro $\boldsymbol{\beta}$ įvertiniui rasti yra ekvivalentus DT metodui. b) Įrodykite, kad a. v. $\boldsymbol{T} = ((\boldsymbol{A}^T\boldsymbol{Y})^T, \boldsymbol{Y}^T\boldsymbol{Y})^T$ yra parametro $(\beta_1, ..., \beta_m, \sigma^2)^T$ pilnoji ir pakankamoji statistika. c) Įrodykite, kad $\hat{\beta}_i, \boldsymbol{L}^T\hat{\boldsymbol{\beta}}, s^2$ yra parametrų $\beta_i, \boldsymbol{L}^T\boldsymbol{\beta}, \sigma^2$ NMD įvertiniai.

II.1.20. (**II.1.19** pratimo tęsinys). a) Raskite įvertinių $\hat{\boldsymbol{\beta}}$, $\hat{\theta} = \boldsymbol{L}^T\hat{\boldsymbol{\beta}}, s^2$ tikimybinius skirstinius. b) Raskite parametrų $\beta_i, \theta, \sigma^2$ pasikliovimo intervalus. c) Sudarykite hipotezių $H : \beta_i = \beta_i^0$, $H : \theta = \theta_0$, $H : \sigma^2 = \sigma_0^2$ tikrinimo kriterijus.

II.1.21. (**II.1.19** pratimo tęsinys). a) Raskite k-mačio a. v. $\hat{\boldsymbol{\theta}} = \boldsymbol{H}\hat{\boldsymbol{\beta}}$ (žr. **II.1.1** pratimą, p. c) tikimybinį skirstinį. b) Sudarykite parametro $\boldsymbol{\theta}$ pasikliovimo sritį. c)Raskite reikšmingumo lygmens α kriterijų hipotezei $H : \boldsymbol{H\beta} = \boldsymbol{\theta}_0$ tikrinti. d) Užrašykite p. c) kriterijų naudodami kvadratinės formos $SS(\boldsymbol{\beta})$ sąlyginį minimumą, surastą, kai H teisinga.

II.1.22. (**II.1.19** pratimo tęsinys). a) Raskite parametrų $\beta_1, ..., \beta_m$ pasikliovimo intervalų rinkinį, kad jie dengtų visus parametrus su tikimybe, ne mažesne už Q, remdamiesi **II.1.21** pratime surasta pasikliovimo sritimi. b) Raskite tokį pasikliovimo intervalų rinkinį pasiremdami Bonferonio nelygybe.

II.1.23. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. d. $X \sim N(\mu, \sigma^2)$. Raskite parametro μ MK įvertinį ir liekamąją kvadratų sumą SS_E.

II.1.24. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ ir $\boldsymbol{Y} = (Y_1, ..., Y_m)^T$ yra paprastosios imtys n. a. d. X ir Y, kurių $\mathbf{E}X = \mu_1$, $\mathbf{E}Y = \mu_2$ ir $\boldsymbol{V}X = \boldsymbol{V}Y = \sigma^2$. Užrašykite stebėjimus tiesinio modelio pavidalu. Raskite parametrų μ_1, μ_2 mažiausiųjų kvadratų įvertinius.

II.1.25. (**II.1.24** pratimo tęsinys). Tarkime, kad stebėti normalieji a. d. $X \sim N(\mu_1,\ \sigma^2)$ ir $Y \sim N(\mu_2,\ \sigma^2)$. Raskite liekamąją kvadratų sumą SS_E ir jos skirstinį.

II.1.26. (**II.1.24** pratimo tęsinys). Reikia patikrinti sudėtingąją hipotezę $H : \mu_1 = \mu_2$. Užrašykite šią hipotezę matriciniu pavidalu kaip **II.1.21** pratime. Raskite SS_{EH} ir $SS_{EH} - SS_E$.

II.1.27. (**II.1.12** pratimo tęsinys). Sukurkite kriterijų hipotezei $H : \gamma = 0$ tikrinti, kai a. d. e_i' pasiskirstę pagal normalųjį skirstinį.

II.1.28. (**II.1.5** pratimo tęsinys). Tarkim, kad **II.1.5** pratime paklaidos turi normaliuosius skirstinius. Raskite parametrų $\alpha, \beta_1, \beta_2, \sigma^2$ lygmens $Q = 0,95$ pasikliovimo intervalus.

II.1.29. (**II.1.28** pratimo tęsinys). Raskite parametrų α, β_1, β_2 pasikliovimo intervalų rinkinius, kurie uždengtų visus parametrus su tikimybe, ne mažesne už $Q = 0,95$. a) Remdamiesi **II.1.22** pratimo p. a). b) Remdamiesi **II.1.22** pratimo p. b).

II.1.30. Kiek kartų **II.1.22** parametro p. a) intervalų ilgiai yra didesni už **II.1.22** p. b) intervalų ilgius, kai $n = 20,\ 100$; $m = 2,\ 5,\ 10$, o pasikliovimo lygmuo $Q = 0,95$.

II.1.31. (**II.1.7** pratimo tęsinys). Tarkime, kad paklaidos normaliosios. a) Raskite parametro $\boldsymbol{\theta} = (\theta_1, \theta_2)^T$, $\theta_1 = 2\alpha - \beta_1, \theta_2 = \alpha + \beta_1 - \beta_2$ lygmens Q pasikliovimo sritį. b) Reikšmingumo lygmens $\alpha = 0,05$ kriterijumi patikrinkite hipotezę $H : \boldsymbol{\theta} = \mathbf{0}$.

II.1.32. (**II.1.14** pratimo tęsinys). Tardami, kad kampų matavimų paklaidos turi normaliuosius skirstinius $N(0, \sigma^2)$, patikrinkite hipotezę $H : \beta_1 + \beta_2 = 90, \beta_3 + \beta_4 = 90$, t. y. kad keturkampis yra stačiakampis.

II.1.33. Tarkime, kad tiesinio modelio matrica $\boldsymbol{A}^T\boldsymbol{A} = [a_{ij}]_{(m)\times(m)}$ neišsigimusi ir diagonaliniai elementai a_{ii}, $i = 1, ..., m$ fiksuoti. Įrodykite, kad

a) parametro β_i įvertinio dispersija tenkina nelygybę $\boldsymbol{V}\hat{\beta}_i \geq 1/a_{ii}$;

b) įvertinių $\hat{\beta}_i$ dispersijos minimalios, kai plano matricos A stulpeliai ortogonalūs, t. y. $\boldsymbol{A}^T\boldsymbol{A}$ – diagonalioji matrica.

II.1.34. Turime m objektų, kurių svoriai $\beta_1, ..., \beta_m$ yra nežinomi. Objektų svoris nustatomas sveriant juos lėkštelinėmis svarstyklėmis dviem būdais.

1) Kiekvienas objektas sveriamas r kartų ir jo svorio įvertiniu imamas gautų rezultatų aritmetinis vidurkis.

2) Sveriant keli objektai dedami ant vienos lėkštelės, keli objektai – ant kitos ir pridedamas svarelis y, kad svarstyklės būtų pusiausviros. Tada k-ajam svėrimui aprašyti turime tiesinį modelį:

$$y_k = a_{k1}\beta_1 + a_{k2}\beta_2 + ... + a_{km}\beta_m + e_k,\ k = 1, ..., n;$$

čia plano matrica $A = [a_{kj}]_{n\times m}$, elementas $a_{kj} = +1,\ -1$ arba 0, atsižvelgiant į tai, ar j-asis objektas padėtas ant kairės, dešinės lėkštutės arba apskritai nedalyvauja sveriant. Tarkime, kad matavimo paklaidos e abiem svėrimo būdais yra vienodai pasiskirstę n. a. d. su ta pačia dispersija σ^2.

Įrodykite, kad antruoju būdu didžiausio tikslumo pasiekiama, kai plano matricos A elementai yra arba $+1$, arba -1 ir jos stulpeliai ortogonalūs.

II.1.35. (**II.1.34** pratimo tęsinys). Tarkime, kad reikia įvertinti $m = 4$ objektų svorius $\boldsymbol{V}\hat{\beta}_i = \sigma^2/4$ tikslumu. Tada pirmuoju būdu reikėtų atlikti $mr = 16$ svėrimų. Kiek kartų galima sumažinti svėrimų skaičių antruoju būdu? Raskite tokių minimalaus skaičiaus svėrimų plano matricos A pavidalus.

II.1.36. (**II.1.34** pratimo tęsinys). Sveriant 4 objektus antruoju būdu gauti dviejų nepriklausomų serijų po 4 svėrimus rezultatai

y_k	a_{k1}	a_{k2}	a_{k3}	a_{k4}
20,2	$+1$	$+1$	$+1$	$+1$
8,1	$+1$	-1	$+1$	-1
9,7	$+1$	$+1$	-1	-1
1,9	$+1$	-1	-1	$+1$

y_k	a_{k1}	a_{k2}	a_{k3}	a_{k4}
19,9	$+1$	$+1$	$+1$	$+1$
8,3	$+1$	-1	$+1$	-1
10,2	$+1$	$+1$	-1	-1
1,8	$+1$	-1	-1	$+1$

a) Raskite parametrų $\beta_1, ..., \beta_4$ įverčius pagal vieno ir kito eksperimento rezultatus. Ar galima pagal tas atskiras eksperimentų serijas įvertinti dispersiją σ^2?

b) Sujunkite šiuos abu eksperimentus ir įvertinkite parametrus $\beta_1, ..., \beta_4$, σ^2. Kiek kartų reikėtų padidinti svėrimų skaičių naudojant pirmąjį būdą, kad parametrų $\beta_1, ..., \beta_4$ įvertiniai būtų tokio paties tikslumo?

c) Tarę, kad paklaidų skirstiniai yra normalieji, palyginkite dispersijos σ^2 įvertinių dispersijas pirmuoju ir antruoju būdu, kai parametrų $\beta_1, ..., \beta_4$ įvertinių tikslumas yra vienodas.

II.1.37. Tarkime, $\boldsymbol{Y} = \boldsymbol{A\beta} + \boldsymbol{e}$, $\mathbf{E}(\boldsymbol{e}) = \mathbf{0}, \boldsymbol{V}(\boldsymbol{e}) = \sigma^2\boldsymbol{\Lambda}$; čia $\mathrm{Rang}(\boldsymbol{A}^T\boldsymbol{A}) = m$, o $\boldsymbol{\Lambda} = [\lambda_{ij}]_{n\times n}$ – žinoma teigiamai apibrėžta matrica. Raskite parametro $\boldsymbol{\beta}$ mažiausiųjų kvadratų įvertinį ir jo kovariacinę matricą.

II.1.38. (**II.1.37** pratimo tęsinys). Tegu $Y_i, i = 1, ..., n$, yra nepriklausomi a. d., kurių $\mathbf{E}(Y_i) = \theta$, $\boldsymbol{V}(Y_i) = \sigma^2/\omega_i$; ω_i – žinomi. Raskite parametro θ tiesinį nepaslinktąjį įvertinį su minimalia dispersija. Raskite šios dispersijos išraišką.

II.1.39. (**II.1.37** pratimo tęsinys). Tegu $Y_1, ..., Y_n$ yra nepriklausomi a. d. ir $Y_i \sim N(i\theta, i^2\sigma^2)$, $i = 1, ..., n$. Raskite parametro θ NMD įvertinį ir įrodykite, kad jo dispersija lygi σ^2/n.

II.1.40. (**II.1.37** pratimo tęsinys). Tarkime, kad **II.1.37** pratimo sąlygomis $\boldsymbol{e} \sim N_n(\mathbf{0}, \sigma^2\boldsymbol{\Lambda})$. Raskite parametrų $\boldsymbol{\beta}$ ir σ^2 įvertinius ir jų skirstinius.

II.1.41. (**II.1.37** pratimo tęsinys). Reikia įvertinti skysčio tankį d atliekant nepriklausomus įvairaus tūrio skysčio svėrimus. Tegu Y_i yra gautas tūrio X_i skysčio svoris; $\mathbf{E}(Y_i) = dX_i$, $\boldsymbol{V}(Y_i) = \sigma^2 f(X_i)$, $i = 1, ..., n$. Raskite parametro d mažiausiųjų kvadratų įvertinį, kai a) $f(X_i) = 1$; b) $f(X_i) = X_i$; c) $f(X_i) = X_i^2$.

II.1.42. (**II.1.37** pratimo tęsinys). Tegu $Y_i = \beta_0 + \beta_1 X_i + e_i$, $i = 1, 2, 3$; $\mathbf{E}(\boldsymbol{e}) = \mathbf{0}$, $\boldsymbol{V}(\boldsymbol{e}) = \sigma^2\boldsymbol{\Lambda}$; čia

$$\boldsymbol{\Lambda} = \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix},$$

o ρ yra žinomas. Raskite parametrų β_0 ir β_1 mažiausiųjų kvadratų įvertinius ir jų dispersijas.

II.1.43. Tegu imties elementai $Y_1, ..., Y_n$, aprašomi tiesiniu modeliu su normaliosiomis paklaidomis, turi vienodas dispersijas $\boldsymbol{V}(Y_i) = \sigma^2$ ir vienodas kovariacijas $\mathbf{Cov}\,(Y_i, Y_j) = \rho\sigma^2, i \neq j$. Atliekame ortogonalią tiesinę transformaciją, pervedančią a. v. $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ į vektorių $\boldsymbol{Z} = (Z_1, ..., Z_n)^T$, kai $Z_1 = (Y_1 + ... + Y_n)/n$. Įrodykite, kad vektoriaus $\boldsymbol{Z}_2 = (Z_2, ..., Z_n)^T$ koordinatės nekoreliuotos ir turi vienodas dispersijas $\sigma^2(1 - \rho)$. Įrodykite, kad parametrinių funkcijų nepaslinktieji tiesiniai įvertiniai su minimalia dispersija yra tiesinio modelio $\boldsymbol{Z}_2 = \boldsymbol{U\beta} + \boldsymbol{e}, \boldsymbol{e} \sim N_{n-1}(\mathbf{0}, \sigma^2(1-\rho)\boldsymbol{I})$ mažiausiųjų kvadratų įvertiniai.

II.1.44. Tarkime, kad į **II.1.1** pratimo tiesinį modelį įtraukiame papildomai r kovariančių. Tada gauname išplėstą tiesinį modelį

$$\boldsymbol{Y} = \boldsymbol{A\beta} + \boldsymbol{B\gamma} + \boldsymbol{e} = (\boldsymbol{A}\vdots\boldsymbol{B})\begin{pmatrix} \boldsymbol{\beta} \\ \cdots \\ \boldsymbol{\gamma} \end{pmatrix} = \boldsymbol{W\delta} + \boldsymbol{e}.$$

Tarę, kad $\mathrm{Rang}(\boldsymbol{W}) = m + r$, gauname išplėstinio modelio parametro $\boldsymbol{\delta}$ įvertinį

$$\hat{\boldsymbol{\delta}} = \begin{pmatrix} \boldsymbol{\beta}^* \\ \cdots \\ \boldsymbol{\gamma}^* \end{pmatrix} = (\boldsymbol{W}^T\boldsymbol{W})^{-1}\boldsymbol{W}^T\boldsymbol{Y},$$

kuris yra sistemos, susidedančios iš $m + r$ lygčių, sprendinys. Įrodykite, kad

a) $\boldsymbol{\beta}^* = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T(\boldsymbol{Y} - \boldsymbol{B\gamma}^*)$, $\boldsymbol{\gamma}^* = (\boldsymbol{B}^T\boldsymbol{RB})^{-1}\boldsymbol{B}^T\boldsymbol{RY}$, $\boldsymbol{R} = \boldsymbol{I} - \boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T$, t. y. galima spręsti dvi sistemas, susidedančias iš m ir r lygčių;

b) $SS_E = (\boldsymbol{Y} - \boldsymbol{B}\boldsymbol{\gamma}^*)^T \boldsymbol{R}(\boldsymbol{Y} - \boldsymbol{R}\gamma^*) = \boldsymbol{Y}^T \boldsymbol{R}\boldsymbol{Y} - (\gamma^*)^T \boldsymbol{B}^T \boldsymbol{R}\boldsymbol{Y}$;
c) $\boldsymbol{V}(\boldsymbol{\beta}^*) = \sigma^2[(\boldsymbol{A}^T\boldsymbol{A})^{-1} + \boldsymbol{L}\boldsymbol{M}\boldsymbol{L}^T], \quad \boldsymbol{V}(\gamma^*) = \sigma^2\boldsymbol{M},$

$$\mathbf{Cov}\,(\boldsymbol{\beta}^*, \gamma^*) = -\sigma^2\boldsymbol{L}\boldsymbol{M}; \quad \boldsymbol{L} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{B}, \quad \boldsymbol{M} = (\boldsymbol{B}^T\boldsymbol{R}\boldsymbol{B})^{-1}.$$

II.1.3. Atsakymai, nurodymai, sprendimai

II.1.1 skyrelis

II.1.1. a) Mažiausiųjų kvadratų įvertinys gaunamas minimizuojant pagal $\boldsymbol{\beta}$ kvadratinę formą (žr. [3], 1.2 skyrelį)

$$SS(\boldsymbol{\beta}) = (\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{\beta})^T(\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{\beta}) \to \min_{\boldsymbol{\beta}}.$$

Diferencijuodami $SS(\boldsymbol{\beta})$ pagal $\boldsymbol{\beta}$ ir prilyginę išvestinę nuliui, gauname lygčių sistemą

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{\beta} = \boldsymbol{A}^T\boldsymbol{Y},$$

iš kurios randame parametro $\boldsymbol{\beta}$ MK įvertinį

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y}.$$

b) Įvertinys $\hat{\boldsymbol{\beta}}$ yra tiesinė a. v. $\boldsymbol{Y}$ funkcija. Taigi

$$\mathbf{E}(\hat{\boldsymbol{\beta}}) = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{\beta} = \boldsymbol{\beta},$$

$$\boldsymbol{V}(\hat{\boldsymbol{\beta}}) = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{V}(\boldsymbol{Y})\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A})^{-1} = \sigma^2(\boldsymbol{A}^T\boldsymbol{A})^{-1}.$$

c) Įvertinys $\hat{\boldsymbol{\theta}}$ yra tiesinė a. v. $\boldsymbol{Y}$ funkcija. Gauname

$$\mathbf{E}(\hat{\boldsymbol{\theta}}) = \boldsymbol{H}(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{\beta} = \boldsymbol{H}\boldsymbol{\beta} = \boldsymbol{\theta}, \quad \boldsymbol{V}(\hat{\boldsymbol{\theta}}) = \sigma^2\boldsymbol{H}(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{H}^T.$$

II.1.2. a) Randame

$$SS_E = \boldsymbol{Y}^T\boldsymbol{Y} - \boldsymbol{Y}^T\boldsymbol{A}\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}^T\boldsymbol{A}^T\boldsymbol{Y} + \hat{\boldsymbol{\beta}}^T\boldsymbol{A}^T\boldsymbol{A}\hat{\boldsymbol{\beta}} =$$

$$\boldsymbol{Y}^T\boldsymbol{Y} - \hat{\boldsymbol{\beta}}^T\boldsymbol{A}^T\boldsymbol{Y} + \hat{\boldsymbol{\beta}}^T[\boldsymbol{A}^T\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{A}^T\boldsymbol{Y}] = \boldsymbol{Y}^T\boldsymbol{Y} - \hat{\boldsymbol{\beta}}^T\boldsymbol{A}^T\boldsymbol{Y},$$

nes (**II.1.1** pratimo p. a)) laužtiniuose skliaustuose parašytas reiškinys lygus nuliui.
b) Turime

$$SS(\boldsymbol{\beta}) = (\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}} + \boldsymbol{A}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}))^T(\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}} + \boldsymbol{A}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})) =$$

$$SS_E + (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T\boldsymbol{A}^T\boldsymbol{A}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}),$$

nes

$$(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T\boldsymbol{A}^T(\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}}) = (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T\boldsymbol{A}^T(\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{a}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T)\boldsymbol{Y} = 0.$$

Kadangi $\mathbf{E}(SS(\boldsymbol{\beta})) = \mathbf{E}(e_1^2 + ... + e_n^2) = n\sigma^2$, o (žr. [3], 7.3 skyrelį)

$$\mathbf{E}((\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T\boldsymbol{A}^T\boldsymbol{A}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})) = Tr(\boldsymbol{V}(\hat{\boldsymbol{\beta}})\boldsymbol{A}^T\boldsymbol{A}) = \sigma^2 Tr(\boldsymbol{I}) = m\sigma^2,$$

tai $\mathbf{E}(SS_E) = (n-m)\sigma^2$. Nepaslinktasis dispersijos σ^2 įvertinys

$$s^2 = SS_E/(n-m), \quad \mathbf{E}(s^2) = \sigma^2.$$

II.1.3. a) Sujungę stebėjimus Y_{ij} į vieną bendrą vektorių

$$\boldsymbol{Y} = (Y_{11}, ..., Y_{1n_1}, Y_{21}, ..., Y_{2n_2}, ..., Y_{m1}, ..., Y_{mn_m})^T,$$

pažymėję nežinomų parametrų vektorių $\boldsymbol{\beta} = (\mu_1, ..., \mu_m)^T$ (čia μ_i yra i-osios kviečių veislės vidutinis derlingumas) ir tarę, kad paklaidų vektorius $\boldsymbol{e}$ tenkina sąlygas $\mathbf{E}(\boldsymbol{e}) = \mathbf{0}$, $\boldsymbol{V}(\boldsymbol{e}) = \sigma^2\boldsymbol{I}$, gausime modelio aprašymą matricine forma $\boldsymbol{Y} = \boldsymbol{A\beta} + \boldsymbol{e}$. Matrica $\boldsymbol{A}$ turi $n = n_1 + \cdots + n_m$ eilučių ir m stulpelių. Pirmosios n_1 eilutės turi pavidalą $(1, 0, \ldots, 0)$, paskui n_2 eilučių turi pavidalą $(0, 1, 0, \ldots, 0)$, pagaliau paskutinės n_m eilučių turi pavidalą $(0, 0, \ldots, 0, 1)$.

b) Matrica $\boldsymbol{A}^T\boldsymbol{A}$ diagonali su diagonaliniais elementais $n_i, i = 1, ..., m$; vektoriaus $\boldsymbol{A}^T\boldsymbol{Y}$ elementai yra

$$\sum_{j=1}^{n_i} Y_{ij} = n_i\bar{Y}_{i\cdot}, \quad i = 1, ..., m.$$

Gauname parametrų įvertinius

$$\hat{\mu}_i = \bar{Y}_{i\cdot}, \quad i = 1, ..., m, \quad \hat{\sigma^2} = s^2 = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij} - Y_{i\cdot})^2/(n-m).$$

II.1.4. a) Tarkime, kad stebėjimai, kai skirtingi $i = 1, ..., n$ yra nekoreliuoti a. d.; Y sąlyginio skirstinio, kai $\mathbf{X} = (X_1, X_2)^T = (x_1, x_2)^T$ yra fiksuotas, dispersija lygi σ^2 ir nepriklauso nuo $\boldsymbol{x} = (x_1, x_2)^T$, o Y vidurkis yra tiesinė funkcija: $\mathbf{E}(Y|\mathbf{X} = \boldsymbol{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$. Tardami, kad a. v. $(X_1, X_2)^T$ realizacijos $(x_{1i}, x_{2i}), i = 1, ..., n$ yra fiksuotos, gauname modelį

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + e_i, \quad i = 1, ..., n.$$

Pažymėjus $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ ir $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ nežinomų parametrų vektorių, modelį galima užrašyti matriciniu pavidalu $\boldsymbol{Y} = \boldsymbol{A}^T\boldsymbol{\beta} + \boldsymbol{e}$ su matrica $\boldsymbol{A}$, kuri turi n eilučių ir 3 stulpelius: i-oji eilutė turi pavidalą $(1, x_{1i}, x_{2i})$.

b) Parametro $\boldsymbol{\beta}$ MK įvertinys yra $\hat{\boldsymbol{\beta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y}$, kai matricos $\boldsymbol{A}^T\boldsymbol{A} = [a_{ij}]_{3\times 3}$ elementai yra $a_{11} = n, a_{22} = \sum_i x_{1i}^2, a_{33} = \sum_i x_{2i}^2$, $a_{12} = \sum_i x_{1i}, a_{13} = \sum_i x_{2i}, a_{23} = \sum_i x_{1i}x_{2i}$; vektorius $\boldsymbol{A}^T\boldsymbol{Y} = (\sum_i Y_i, \sum_i Y_i x_{1i}, \sum_i Y_i x_{2i})^T$;

$$\hat{\sigma^2} = s^2 = \sum_i (Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \hat{\beta}_2 x_{2i})^2/(n-3).$$

II.1.5. Turime tiesinį Gauso ir Markovo modelį $\boldsymbol{Y} = \boldsymbol{A\beta} + \boldsymbol{e}$, kuriame

$$\boldsymbol{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \alpha \\ \beta_1 \\ \beta_2 \end{pmatrix}, \quad \boldsymbol{A} = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 2 & -1 \\ 1 & 0 & 1 \\ 1 & -2 & 1 \end{pmatrix}, \quad \boldsymbol{e} = \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{pmatrix}.$$

Gauname

$$(\boldsymbol{A}^T\boldsymbol{A})^{-1} = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \quad \boldsymbol{A}^T\boldsymbol{Y} = \begin{pmatrix} Y_1 + Y_2 + Y_3 + Y_4 \\ 2Y_2 - 2Y_4 \\ -Y_1 - Y_2 + Y_3 + Y_4 \end{pmatrix}$$

ir

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y} = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} Y_1 + Y_2 + Y_3 + Y_4 \\ -Y_1 + Y_2 + Y_3 - Y_4 \\ -2Y_1 + 2Y_3 \end{pmatrix}.$$

Pagal turimą imties realizaciją gauname įvertį

$$\boldsymbol{A}^T\boldsymbol{Y} = \begin{pmatrix} 4,2 \\ 3,9 \\ 3,9 \end{pmatrix}, \quad \hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 4,2 \\ 7,8 \\ 11,7 \end{pmatrix} = \begin{pmatrix} 1,05 \\ 1,95 \\ 2,93 \end{pmatrix}.$$

II.1.6. Gauname

$$SS_E = (Y_1 - \hat{\alpha} + \hat{\beta}_2)^2 + (Y_2 - \hat{\alpha} - 2\hat{\beta}_1 + \hat{\beta}_2)^2 + (Y_3 - \hat{\alpha} - \hat{\beta}_2)^2 + (Y_4 - \hat{\alpha} + 2\hat{\beta}_1 - \hat{\beta}_2)^2 = 0,0025.$$

Liekamąją kvadratų sumą galima rasti ir pagal **II.1.2** pratimo p. a) formulę

$$SS_E = Y_1^2 + Y_2^2 + Y_3^2 + Y_4^2 - 4,2\hat{\alpha} - 3,9\hat{\beta}_1 - 3,9\hat{\beta}_2 = 0,0025.$$

II.1.7. Remiantis **II.1.1** pratimu

$$\boldsymbol{V}(\hat{\boldsymbol{\beta}}) = \sigma^2(\boldsymbol{A}^T\boldsymbol{A})^{-1} = \frac{\sigma^2}{4}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix},$$

iš čia $\boldsymbol{V}\hat{\alpha} = \sigma^2/4, \boldsymbol{V}\hat{\beta}_1 = \sigma^2/4, \boldsymbol{V}\hat{\beta}_2 = \sigma^2/2$, $\mathbf{Cov}\,(\hat{\alpha}, \hat{\beta}_1) = 0$, $\mathbf{Cov}\,(\hat{\alpha}, \hat{\beta}_2) = 0$, $\mathbf{Cov}\,(\hat{\beta}_1, \hat{\beta}_2) = \sigma^2/4$. Dispersijos σ^2 nepaslinktasis įvertis

$$s^2 = SS_E/(n-m) = SS_E = 0,0025.$$

II.1.8. Parametras $\boldsymbol{\theta}$ yra tiesinė $\boldsymbol{\beta}$ funkcija

$$\boldsymbol{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{L}^T\boldsymbol{\beta} \\ \boldsymbol{K}^T\boldsymbol{\beta} \end{pmatrix}, \quad \boldsymbol{L} = (2,\ -1,\ 0)^T, \ \ \boldsymbol{K} = (1,\ 1,\ -1)^T.$$

Remiantis [3], 1.2.3 teorema vienintelis mažiausios dispersijos įvertinys tiesinių nepaslinktųjų įvertinių klasėje yra

$$\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{L}^T\hat{\boldsymbol{\beta}} \\ \boldsymbol{K}^T\hat{\boldsymbol{\beta}} \end{pmatrix} = \begin{pmatrix} 2\hat{\alpha} - \hat{\beta}_1 \\ \hat{\alpha} + \hat{\beta}_1 - \hat{\beta}_2 \end{pmatrix},$$

o jo kovariacinė matrica

$$\boldsymbol{V}(\hat{\boldsymbol{\theta}}) = \sigma^2\begin{pmatrix} \boldsymbol{L}^T(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{L} & \boldsymbol{L}^T(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{K} \\ \boldsymbol{K}^T(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{L} & \boldsymbol{K}^T(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{K} \end{pmatrix} = \frac{\sigma^2}{4}\begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}.$$

Įvertinio $\hat{\boldsymbol{\theta}}$ realizacija yra $(0,15,\ \ 0,075)$.

II.1.9. $\hat{\alpha} = (Y_1 + 2Y_2 + Y_3)/6$, $\hat{\beta} = (2Y_3 - Y_2)/5$, $\boldsymbol{V}\hat{\alpha} = \sigma^2/6, \boldsymbol{V}\hat{\beta} = \sigma^2/5$, $\mathbf{Cov}\,(\hat{\alpha}, \hat{\beta}) = 0$.

II.1.10. $\hat{\beta}_0 = (Y_1 + Y_2 + Y_3)/3$, $\hat{\beta}_1 = (Y_3 - Y_1)/2$, $\hat{\beta}_2 = (Y_1 - 2Y_2 + Y_3)/6$.

II.1.11. Matricos $\boldsymbol{A}^T\boldsymbol{A} = [c_{ij}]_{2\times 2}$ elementas $c_{12} = c_{21} = m - 2n$.

II.1.12. Tarkime, kad $X_i = x_i, i = 1,...,n$, reikšmės yra fiksuotos. Atveju ω parametro $\boldsymbol{\beta} = (\alpha, \beta)^T$ įvertinys $\hat{\boldsymbol{\beta}}$ gaunamas sprendžiant dviejų lygčių sistemą:

$$\boldsymbol{A}^T\boldsymbol{A}\hat{\boldsymbol{\beta}} = \boldsymbol{A}^T\boldsymbol{Y};$$

čia matrica $\boldsymbol{A}^T\boldsymbol{A} = [a_{ij}]_{2\times 2}$, $a_{11} = n, a_{12} = a_{21} = \sum_i x_i, a_{22} = \sum_i x_i^2$, o vektorius $\boldsymbol{A}^T\boldsymbol{Y} = (\sum_i Y_i, \sum_i Y_i x_i)^T$. Atveju Ω parametro $\boldsymbol{\beta} = (\alpha, \beta, \gamma)^T$ įvertinys $\tilde{\boldsymbol{\beta}}$ gaunamas sprendžiant trijų lygčių sistemą: $\boldsymbol{A}^T\boldsymbol{A}\tilde{\boldsymbol{\beta}} = \boldsymbol{A}^T\boldsymbol{Y}$; čia matrica $\boldsymbol{A}^T\boldsymbol{A} = [a_{ij}]_{3\times 3}$, $a_{13} = a_{31} = \sum_i x_i^2, a_{23} = a_{32} = \sum_i x_i^3$, $a_{33} = \sum_i x_i^4$, o vektorius $\boldsymbol{A}^T\boldsymbol{Y} = (\sum_i Y_i, \sum_i Y_i x_i, \sum_i Y_i x_i^2)^T$.

II.1.13. Įvertinių dispersijos yra matricos $\sigma^2(\boldsymbol{A}^T\boldsymbol{A})^{-1}$ diagonaliniai elementai. Jeigu atveju ω vietoje $\alpha + \beta X_i$ imsime $\delta + \beta(X_i - \bar{X})$, tai matrica $\boldsymbol{A}^T\boldsymbol{A}$ diagonali.

II.1.14. Imties $\boldsymbol{Y} = (Y_1, ..., Y_8)^T$ elementai turi tokią struktūrą $Y_1 = \beta_1 + e_1, Y_2 = \beta_2 + e_2$, $Y_3 = \beta_1 + \beta_2 + e_3, 180 - Y_4 = \beta_2 + \beta_3 + e_4$, $Y_5 = \beta_3 + e_5, Y_6 = \beta_4 + e_6$, $Y_7 = \beta_3 + \beta_4 + e_7, 180 - Y_8 = \beta_1 + \beta_4 + e_8$. Pažymėkime $Z_4 = 180 - Y_4, Z_8 = 180 - Y_8$, $Z_i = Y_i, i \neq 4,\ 8$. Tada turime tiesinį modelį $\boldsymbol{Z} = \boldsymbol{A\beta} + \boldsymbol{e}$, kuriame plano matrica

$$\boldsymbol{A}^T = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

Randame

$$(\boldsymbol{A}^T\boldsymbol{A})^{-1} = \frac{1}{15}\begin{pmatrix} 7 & -3 & 2 & -3 \\ -3 & 7 & -3 & 2 \\ 2 & -3 & 7 & -3 \\ -3 & 2 & -3 & 7 \end{pmatrix}, \boldsymbol{A}^T\boldsymbol{Z} = \begin{pmatrix} Z_1 + Z_3 + Z_8 \\ Z_2 + Z_3 + Z_4 \\ Z_4 + Z_5 + Z_7 \\ Z_6 + Z_7 + Z_8 \end{pmatrix} = \begin{pmatrix} 219,21 \\ 188,97 \\ 219,72 \\ 219,60 \end{pmatrix}.$$

Gauname parametrų įverčius $\hat{\boldsymbol{\beta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Z} = (49,88;\ 29,68;\ 50,05;\ 39,89)^T$. Liekamoji kvadratų suma

$$SS_E = \sum_{i=1}^{8} Z_i^2 - \hat{\boldsymbol{\beta}}^T\boldsymbol{A}^T\boldsymbol{Z} = 5,1465.$$

Dispersijos σ^2 nepaslinktasis įvertis $s^2 = SS_E/4 = 1,2866$.

II.1.15. Pažymėkime $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, ..., \hat{\theta}_k)^T$ ir $\boldsymbol{\Sigma}^{-1}$ kovariacinės matricos $\boldsymbol{\Sigma} = \boldsymbol{V}(\hat{\boldsymbol{\theta}})$ atvirkštinę matricą. Tiesinės formos $\hat{\theta} = \boldsymbol{L}^T\hat{\boldsymbol{\theta}}$ dispersija $\boldsymbol{V}\theta = \boldsymbol{L}^T\boldsymbol{\Sigma}\boldsymbol{L}$. Įvertinys $\hat{\theta}$

nepaslinktas, kai $\boldsymbol{L}^T\mathbf{1} = 1$; čia $\mathbf{1} + (1,\ 1, ..., 1)^T$ yra vektorius su vienetinėmis koordinatėmis. Reikia rasti vektorių $\boldsymbol{L}$ tenkinantį sąlygas

$$\begin{cases} \boldsymbol{L}^T\boldsymbol{\Sigma}\boldsymbol{L} \to \min_{\boldsymbol{L}}, \\ \boldsymbol{L}^T\mathbf{1} = 1. \end{cases}$$

Remiantis Lagranžo neapibrėžtinių daugiklių metodu reikia minimizuoti $\boldsymbol{L}^T\boldsymbol{\Sigma}\boldsymbol{L} - 2\lambda(\boldsymbol{L}^T\mathbf{1} - 1)$. Diferencijuodami pagal $\boldsymbol{L}^T$ gauname

$$\boldsymbol{\Sigma}\boldsymbol{L} - \lambda\mathbf{1} = 0 \quad \Leftrightarrow \quad \boldsymbol{L} = \lambda\boldsymbol{\Sigma}^{-1}\mathbf{1}.$$

Iš sąlygos $\boldsymbol{L}^T\mathbf{1} = 1$ randame $\lambda = 1/(\mathbf{1}^T\boldsymbol{\Sigma}^{-1}\mathbf{1})$.

Taigi minimali dispersija $\boldsymbol{V}(\boldsymbol{L}^T\hat{\boldsymbol{\theta}}) = 1/(\mathbf{1}^T\boldsymbol{\Sigma}^{-1}\mathbf{1})$.

II.1.16. Remiantis **II.1.15** pratimu $c_i = \sigma_i^{-2}/(\sigma_1^{-2}+...+\sigma_k^{-2})$ ir $\boldsymbol{V}(c_1\hat{\theta}_1+...+c_k\hat{\theta}_k) = 1/(\sigma_1^{-2} + ... + \sigma_k^{-2})$.

II.1.17. Įvertinys $\hat{\theta} = c_1\hat{\theta}_1 + ... + c_k\hat{\theta}_k$ yra nepaslinktasis $\mathbf{E}\hat{\theta} = \theta$, o jo dispersija $\boldsymbol{V}(\hat{\theta}) = c_1^2\boldsymbol{V}(\hat{\theta}_1) + ... + c_k^2\boldsymbol{V}(\hat{\theta}_k)$ minimali, nes kiekvienas dėmuo įgyja minimalią reikšmę.

II.1.18. $\mathbf{X} = \boldsymbol{A}\mu + \boldsymbol{e}, \quad \boldsymbol{A}^T = (1,\ 1, ..., 1); \quad \mathbf{E}(\boldsymbol{e}) = \mathbf{0}, \quad \boldsymbol{V}(\boldsymbol{e}) = \sigma^2\boldsymbol{I};$

$$\hat{\mu} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\mathbf{X} = \bar{X}.$$

II.1.2 skyrelis

II.1.19. a) Tikėtinumo funkcija

$$L(\boldsymbol{\beta}, \sigma^2) = (2\pi)^{-n/2}\exp\{-\frac{1}{2\sigma^2}(\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{\beta})^T(\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{\beta}) - \frac{n}{2}\ln\sigma^2\} =$$

$$(2\pi)^{-n/2}\exp\{-\frac{1}{2\sigma^2}SS(\boldsymbol{\beta}) - \frac{n}{2}\ln\sigma^2\}.$$

Funkcijos L maksimizavimas $\boldsymbol{\beta}$ atžvilgiu yra ekvivalentus $SS(\boldsymbol{\beta})$ minimizavimui.

b) Perrašykime tikėtinumo funkciją tokiu būdu

$$L(\boldsymbol{\beta}, \sigma^2) = (2\pi)^{-n/2}\exp\{-\frac{1}{2\sigma^2}(\boldsymbol{Y}^T\boldsymbol{Y} - 2\boldsymbol{\beta}^T\boldsymbol{A}^T\boldsymbol{Y}) - \frac{1}{2\sigma^2}\boldsymbol{\beta}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{\beta} - \frac{n}{2}\ln\sigma^2\}.$$

Imties $\boldsymbol{Y}$ tankis priklauso $(m+1)$ parametrei eksponentinio tipo skirstinių šeimai. Remiantis faktorizacijos kriterijumi $\boldsymbol{T}$ yra pakankamoji statistika. Kadangi parametrų kitimo sričiai priklauso vidiniai taškai, tai statistika $\boldsymbol{T}$ pilnoji.

c) Kadangi $\boldsymbol{T}$ pilnoji ir pakankamoji statistika, tai bet kuri $\boldsymbol{T}$ funkcija yra jos vidurkio NMD įvertinys. Įvertiniai $\hat{\beta}_i, \boldsymbol{L}^T\hat{\boldsymbol{\beta}}, s^2$ yra statistikos $\boldsymbol{T}$ funkcijos ir

$$\mathbf{E}\hat{\beta}_i = \beta_i, \quad \mathbf{E}(\boldsymbol{L}\hat{\boldsymbol{\beta}}) = \boldsymbol{L}\boldsymbol{\beta}, \quad \mathbf{E}s^2 = \sigma^2.$$

II.1.20. a) Įvertiniai $\hat{\boldsymbol{\beta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y}$ ir $\hat{\theta} = \boldsymbol{L}^T\hat{\boldsymbol{\beta}}$ yra normaliojo vektoriaus $\boldsymbol{Y} \sim N_n(\boldsymbol{A}\boldsymbol{\beta}, \sigma^2\boldsymbol{I})$ tiesinės funkcijos. Remiantis daugiamačio normaliojo vektoriaus savybėmis ([3], 7.4 skyrelis)

$$\hat{\boldsymbol{\beta}} \sim N_m(\boldsymbol{\beta}, \sigma^2(\boldsymbol{A}^T\boldsymbol{A})^{-1}), \quad \hat{\theta} \sim N(\theta, \sigma^2\boldsymbol{L}^T(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{L}).$$

Turime (žr. **II.1.2** pratimą)

$$\frac{SS(\boldsymbol{\beta})}{\sigma^2} = \frac{SS_E}{\sigma^2} + \frac{(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T \boldsymbol{A}^T \boldsymbol{A}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})}{\sigma^2}.$$

Kairėje lygybės pusėje $SS(\boldsymbol{\beta})/\sigma^2 = (e_1^2 + ... + e_n^2)/\sigma^2 \sim \chi^2(n)$. Remiantis daugiamačio normaliojo vektoriaus savybėmis ([3], 7.4 skyrelis) antrasis dėmuo dešinėje pusėje turi χ^2 skirstinį $\chi^2(m)$. Remiantis Kornišo ir Fišerio teorema ([15], 3b.4 skyrelis) kvadratinė forma $s^2(n-m)/\sigma^2 = SS_E/\sigma^2 \sim \chi^2(n-m)$ ir nepriklauso nuo $\hat{\boldsymbol{\beta}}$.

b) Remiantis p. a)

$$\frac{\hat{\beta}_i - \beta}{s\sqrt{b_{ii}}} \sim S(n-m), \qquad \frac{\hat{\theta} - \theta}{sb} \sim S(n-m),$$

čia b_{ij} yra matricos$(\boldsymbol{A}^T\boldsymbol{A})^{-1} = [b_{ij}]_{m\times m}$ elementai, o $b^2 = \boldsymbol{L}^T(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{L}$. Gauname pasikliovimo lygmens $Q = 1-\alpha$ pasikliovimo intervalus

$$(\underline{\beta}_i; \overline{\beta}_i) = (\hat{\beta}_i - s\sqrt{b_{ii}}t_{\alpha/2}(n-m); \hat{\beta}_i + s\sqrt{b_{ii}}t_{\alpha/2}(n-m)),$$

$$(\underline{\theta}; \overline{\theta}) = (\hat{\theta} - sbt_{\alpha/2}(n-m); \hat{\theta} + sbt_{\alpha/2}(n-m)).$$

Kadangi $s^2(n-m)/\sigma^2 \sim \chi^2(n-m)$, tai gauname

$$(\underline{\sigma^2}_i; \overline{\sigma^2}) = \left(\frac{s^2(n-m)}{\chi^2_{\alpha/2}(n-m)}; \frac{s^2(n-m)}{\chi^2_{1-\alpha/2}(n-m)}\right).$$

c) Kriterijus galime suformuluoti p. b) surastų pasikliovimo intervalų terminais. Arba kriterijai yra tokie: suformuluotos hipotezės, kai alternatyvos dvipusės, atmetamos reikšmingumo lygmens α kriterijais, kai atitinkamai teisingos nelygybės

$$|t_i| = \frac{|\hat{\beta}_i - \beta_i^0|}{s\sqrt{b_{ii}}} > t_{\alpha/2}(n-m), \qquad |t| = \frac{|\hat{\theta} - \theta_0|}{sb} > t_{\alpha/2}(n-m);$$

$$\frac{s^2(n-m)}{\sigma_0^2} < \chi^2_{1-\alpha/2}(n-m) \quad \text{arba} \quad \frac{s^2(n-m)}{\sigma_0^2} < \chi^2_{\alpha/2}(n-m).$$

II.1.21. a) Įvertinys $\hat{\boldsymbol{\theta}} = \boldsymbol{H}\hat{\boldsymbol{\beta}} = \boldsymbol{H}(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y}$ yra tiesinė normaliojo a. v. $\boldsymbol{Y} \sim N_n(\boldsymbol{A\beta}, \sigma^2\boldsymbol{I})$ funkcija. Taigi

$$\hat{\boldsymbol{\theta}} \sim N_k(\boldsymbol{\theta}, \sigma^2\boldsymbol{\Sigma}), \quad \boldsymbol{\Sigma} = \boldsymbol{H}(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{H}^T,$$

$$(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T\boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})/\sigma^2 \sim \chi^2(k).$$

Kadangi $\hat{\boldsymbol{\theta}}$ ir SS_E nepriklausomi, tai

$$\frac{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T\boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})}{ks^2} \sim F(k, n-m)$$

ir

$$C(\hat{\boldsymbol{\theta}}, s^2) = \left\{\boldsymbol{\theta} : \frac{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T\boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})}{ks^2} < F_\alpha(k, n-m)\right\}$$

yra parametro $\boldsymbol{\theta}$ pasikliovimo lygmens $Q = 1 - \alpha$ pasikliovimo sritis.

c) Hipotezė H atmetama, kai

$$F = \frac{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^T \boldsymbol{\Sigma}^{-1} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)}{ks^2} > F_\alpha(k, n-m),$$

t. y. kai taškas $\boldsymbol{\theta}_0$ nepatenka į pasikliovimo sritį $C(\hat{\boldsymbol{\theta}}, s^2)$.

d) Pažymėkime

$$SS_{EH} = \min_{\boldsymbol{\beta}: \boldsymbol{H}\boldsymbol{\beta} = \boldsymbol{\theta}_0} SS(\boldsymbol{\beta}) = \min_{\boldsymbol{\beta}: \boldsymbol{H}\boldsymbol{\beta} = \boldsymbol{\theta}_0} (\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{\beta})^T (\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{\beta}).$$

Tada (žr. [3], 1.3.2 teoremą) $SS_{EH} - SS_E$ ir SS_E yra nepriklausomi; $(SS_{EH} - SS_E)/\sigma^2 \sim \chi^2(k; \lambda)$, necentriškumo parametras $\lambda = (\boldsymbol{H}\boldsymbol{\beta} - \boldsymbol{\theta}_0)^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{H}\boldsymbol{\beta} - \boldsymbol{\theta}_0)$. Jeigu hipotezė H teisinga, tai $\lambda = 0$ ir $(SS_{EH} - SS_E)/\sigma^2 \sim \chi^2(k)$. Hipotezė H atmetama, kai

$$F = \frac{SS_{EH} - SS_E}{ks^2} > F_\alpha(k, n-m).$$

Pastaba. Palyginę p. c) ir d) kriterijus matome, kad $SS_{EH} - SS_E = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^T \boldsymbol{\Sigma}^{-1} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$. Praktiškai dažnai yra paprasčiau rasti sąlyginį minimumą SS_{EH} negu rasti ir apversti matricą $\boldsymbol{\Sigma}$.

II.1.22. a) Remiantis sritimi $C(\hat{\boldsymbol{\theta}}, s^2)$ gaunama, kad visoms tiesinėms funkcijoms $\boldsymbol{c}^T \boldsymbol{\beta}$ su tikimybe $Q = 1 - \alpha$ galioja nelygybės (žr. [3], 1.3.3 teoremą)

$$\boldsymbol{c}^T \hat{\boldsymbol{\beta}} - \delta_\alpha \sqrt{\boldsymbol{c}^T \boldsymbol{\Sigma} \boldsymbol{c}} \le \boldsymbol{c}^T \boldsymbol{\beta} \le \boldsymbol{c}^T \hat{\boldsymbol{\beta}} + \delta_\alpha \sqrt{\boldsymbol{c}^T \boldsymbol{\Sigma} \boldsymbol{c}},$$

čia $\boldsymbol{\Sigma} = (\boldsymbol{A}^T \boldsymbol{A})^{-1}$, $\delta_\alpha = s\sqrt{mF_\alpha(m, n-m)}$. Imdami paeiliui $\boldsymbol{c}_1 = (1,\ 0, ..., 0)^T$, $\boldsymbol{c}_2 = (0,\ 1, ..., 0)^T$,...,$\boldsymbol{c}_m = (0,\ 0, ..., 1)^T$, gausime

$$\mathbf{P}_{\boldsymbol{\beta}}\{\underline{\beta}_i < \beta_i < \bar{\beta}_i, \ \ i = 1, 2, ..., m\} \ge Q = 1 - \alpha;$$

$$(\underline{\beta}_i; \overline{\beta}_i) = (\hat{\beta}_i - s\sqrt{b_{ii}}\sqrt{mF_\alpha(m, n-m)}; \hat{\beta}_i + s\sqrt{b_{ii}}\sqrt{mF_\alpha(m, n-m)}).$$

b) Tegu A_i įvykis, kad intervalas $(\underline{\beta}_i; \overline{\beta}_i)$, surastas **II.1.20** pratime, uždengia parametrą β_i ir tegu $\mathbf{P}\{A_i\} = 1 - \alpha_i$. Tada

$$\mathbf{P}\{\bigcap_{i=1}^m A_i\} = 1 - \mathbf{P}\{\bigcup_{i=1}^m \bar{A}_i\} \ge 1 - \sum_{i=1}^m \mathbf{P}\{\bar{A}_i\} = 1 - (\alpha_1 + ... + \alpha_m).$$

Jeigu parinksime $\alpha_i = \alpha/m, i = 1, ..., m$, tai intervalų rinkinys

$$(\underline{\beta}'_i; \overline{\beta}'_i) = (\hat{\beta}_i - s\sqrt{b_{ii}} t_{\alpha/(2m)}(n-m); \hat{\beta}_i + s\sqrt{b_{ii}} t_{\alpha/(2m)}(n-m))$$

uždengs visus parametrus β_i su tikimybe, ne mažesne už $Q = 1 - \alpha$.

II.1.23. Plano matrica $\boldsymbol{A}^T = (1,\ 1, ..., 1)$, $\hat{\mu} = (\boldsymbol{A}^T \boldsymbol{A})^{-1} \boldsymbol{A}^T \mathbf{X} = \bar{X}$; $SS_E = \sum_i (X_i - \bar{X})^2/\sigma^2 \sim \chi^2_{n-1}$.

II.1.24. Tegu $\boldsymbol{Z} = (X_1, ..., X_n, Y_1, ..., Y_m)^T$. Tada $\boldsymbol{Z} = \boldsymbol{A}\boldsymbol{\beta} + \boldsymbol{e}$, $\boldsymbol{\beta} = (\mu_1, \mu_2)^T$; $\mathbf{E}(\boldsymbol{e}) = 0, \boldsymbol{V}(\boldsymbol{e}) = \sigma^2 \boldsymbol{I}$. Plano matrica $\boldsymbol{A}$ turi $n + m$ eilučių ir du stulpelius; pirmosios

n eilučių turi pavidalą (1, 0), o likusios turi pavidalą (0, 1). Parametro $\boldsymbol{\beta}$ MK įvertinys $\hat{\boldsymbol{\beta}} = (\hat{\mu}_1, \hat{\mu}_2)^T$; $\hat{\mu}_1 = \bar{X}, \hat{\mu}_2 = \bar{Y}$.

II.1.25. $SS_E/\sigma^2 = [\sum_i (X_i - \bar{X})^2 + \sum_j (Y_j - \bar{Y})^2]/\sigma^2 \sim \chi^2_{n+m-2}$.

II.1.26. Tikrinama hipotezė $H : \boldsymbol{H\beta} = \mathbf{0}$; čia $\boldsymbol{H} = (1,\ -1)$. $SS_{EH} = \sum_i (X_i - \bar{Z})^2 + \sum_j (Y_j - \bar{Z})^2$, $\bar{Z} = (n\bar{X} + m\bar{Y})/(m+n)$; $SS_{EH} - SS_E = mn(\bar{X} - \bar{Y})^2/(m+n)$.

II.1.27. $SS_E = \sum_i (Y_i - \tilde{\alpha} - \tilde{\beta}x_i - \tilde{\gamma}x_i^2)^2$, $SS_{EH} = \sum_i (Y_i - \hat{\alpha} - \hat{\beta}x_i)^2$. Hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai $(SS_{EH} - SS_E)(n-3)/SS_E > F_\alpha(1, n-3)$.

II.1.28. Remdamiesi **II.1.20** pratimu gauname

$$(\underline{\sigma^2}; \overline{\sigma^2}) = (0,0005; 2,5456); \quad (\underline{\alpha}; \overline{\alpha}) = (0,7323; 1,3677);$$

$$(\underline{\beta}_1; \overline{\beta}_1) = (1,6323; 2,2677); \quad (\underline{\beta}_2; \overline{\beta}_2) = (2,4758; 3,3742).$$

II.1.29. Gauname intervalus

a) $(\underline{\alpha}; \overline{\alpha}) = (0,4140; 1,6860)$; $(\underline{\beta}_1; \overline{\beta}_1) = (1,3140; 2,5860)$; $(\underline{\beta}_2; \overline{\beta}_2) = (2,0256; 3,8244)$.

b) $(\underline{\alpha}; \overline{\alpha}) = (0,0953; 2,0047)$; $(\underline{\beta}_1; \overline{\beta}_1) = (0,9953; 2,9047)$; $(\underline{\beta}_2; \overline{\beta}_2) = (1,5748; 4,2752)$.

Intervalai gerokai platesni už **II.1.28** pratimo intervalus.

II.1.30. Kai $m = 2$, atitinkamai 1,0905 ir 1,0919 kartų; kai $m = 5$, atitinkamai 1,2925 ir 1,2930 kartų; kai $m = 10$, atitinkamai 1,5328 ir 1,5296 kartų.

II.1.31. a) Parametro $\boldsymbol{\theta}$ įvertinys ir jo kovariacijų matrica surasti **II.1.8** pratime

$$\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{pmatrix} = \begin{pmatrix} 2\hat{\alpha} - \hat{\beta}_1 \\ \hat{\alpha} + \hat{\beta}_1 - \hat{\beta}_2 \end{pmatrix}, \boldsymbol{V}(\hat{\boldsymbol{\theta}}) = \frac{\sigma^2}{4} \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix} = \sigma^2 \boldsymbol{\Sigma}.$$

Remiantis **II.1.21** pratimu lygmens $Q = 0,95$ pasikliovimo sritis

$$C(\hat{\boldsymbol{\theta}}, s^2) = \{\boldsymbol{\theta} : (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T \boldsymbol{\Sigma}^{-1} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})/(2s^2) < F_{0,05}(2,\ 1)\} =$$

$$\{\boldsymbol{\theta} : 2(\hat{\theta}_1 - \theta_1)^2 - 4(\hat{\theta}_1 - \theta_1)(\hat{\theta}_2 - \theta_2) + 5(\hat{\theta}_2 - \theta_2)^2 < 1,496\}.$$

b) Hipotezė atmetama, kai

$$F = \hat{\boldsymbol{\theta}}^T \boldsymbol{\Sigma}^{-1} \hat{\boldsymbol{\theta}}/(2s^2) > F_{0,05}(2,\ 1) = 199,5.$$

Kadangi statistika F įgijo reikšmę 0,75, tai atmesti hipotezę nėra pagrindo.

II.1.32. Tikrinama hipotezė $H : \boldsymbol{H\beta} = \boldsymbol{\theta} = \boldsymbol{\theta}_0$; čia matricos $\boldsymbol{H}$ pirmoji eilutė $(1,\ 1,\ 0,\ 0)$, o antroji eilutė $(0,\ 0,\ 1,\ 1)$; $\boldsymbol{\theta}_0 = (90,\ 90)^T$. Randame

$$\boldsymbol{V}(\boldsymbol{\theta}) = \sigma^2 \boldsymbol{\Sigma} = \sigma^2 \frac{2}{15} \begin{pmatrix} 4 & -1 \\ -1 & 4 \end{pmatrix}, \quad \boldsymbol{\Sigma}^{-1} = \frac{1}{2} \begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix}.$$

Hipotezė atmetama reikšmingumo lygmens $\alpha = 0,05$ kriterijumi, kai

$$F = \frac{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^T \boldsymbol{\Sigma}^{-1} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)}{2s^2} > F_{0,05}(2, 4).$$

Kadangi statistika F įgijo reikšmę 84,96, o $F_{0,05}(2,\ 4) = 6,9443$, tai hipotezė atmetama; P reikšmė $pv = \mathbf{P}\{F_{2,\ 4} > 84,96\} = 0,0005$.

II.1.33. a) $\boldsymbol{V}(\hat{\boldsymbol{\beta}}) = \sigma^2(\boldsymbol{A}^T\boldsymbol{A})^{-1}$. Pažymėkime matricos $(\boldsymbol{A}^T\boldsymbol{A})^{-1}$ elementus a^{ij}, Tada $\boldsymbol{V}\hat{\beta}_i = \sigma^2 a^{ii}$. Tegu

$$\boldsymbol{A}^T\boldsymbol{A} = \begin{pmatrix} a_{11} & \boldsymbol{b}^T \\ \boldsymbol{b} & \boldsymbol{C} \end{pmatrix}, \quad a^{11} = \frac{|\boldsymbol{C}|}{|\boldsymbol{A}^T\boldsymbol{A}|} = \frac{|\boldsymbol{C}|}{|\boldsymbol{C}|(a_{11} - \boldsymbol{b}^T\boldsymbol{C}^{-1}\boldsymbol{b})} \geq \frac{1}{a_{11}}.$$

$$\boldsymbol{V}\hat{\beta}_i \geq \frac{\sigma^2}{a_{ii}}, \quad i = 1, ..., m.$$

b) Išplaukia iš p. a).

II.1.34. Žr. **II.1.33** pratimą.

II.1.35. Keturis kartus.

II.1.36. a) Pagal pirmo eksperimento rezultatus parametrų įverčiai: $\bar{\beta}_1 = (Y_1 + Y_2 + Y_3 + Y_4)/4 = 9,975$, $\bar{\beta}_2 = (Y_1 - Y_2 + Y_3 - Y_4)/4 = 4,975$, $\bar{\beta}_3 = (Y_1 + Y_2 - Y_3 - Y_4)/4 = 4,175$, $\bar{\beta}_4 = (Y_1 - Y_2 - Y_3 + Y_4)/4 = 1,075$. Analogiškai pagal antro eksperimento rezultatus: $\bar{\beta}_1 = 10,050$, $\bar{\beta}_2 = 5,000$, $\bar{\beta}_3 = 4,050$, $\bar{\beta}_4 = 0,800$; negalima.

b) $\bar{\beta}_1 = 10,0125$, $\bar{\beta}_2 = 4,9875$, $\bar{\beta}_3 = 4,1125$, $\bar{\beta}_4 = 0,9375$, $\hat{\sigma}^2 = s^2 = 0,04875$; keturis kartus;

c) $\boldsymbol{V}(\hat{\sigma}^2) = 2\sigma^4/8$; taikant pirmąjį būdą reikėtų 32 svėrimų.

II.1.37. Parinkime kvadratinę matricą $\boldsymbol{B}$, kad $\boldsymbol{\Lambda} = \boldsymbol{B}\boldsymbol{B}^T$, ir atlikime transformaciją $\boldsymbol{Z} = \boldsymbol{B}^{-1}\boldsymbol{Y}$. Tada a.v. $\boldsymbol{Z}$ tenkina tiesinį modelį: $\boldsymbol{Z} = \boldsymbol{C}\boldsymbol{\beta} + \boldsymbol{\theta}$; čia $\boldsymbol{C} = \boldsymbol{B}^{-1}\boldsymbol{A}$, o $\boldsymbol{V}(\boldsymbol{\theta}) = \sigma^2\boldsymbol{I}$. Gauname $\hat{\boldsymbol{\beta}} = (\boldsymbol{C}^T\boldsymbol{C})^{-1}\boldsymbol{C}^T\boldsymbol{Z} = (\boldsymbol{A}^T\boldsymbol{\Lambda}^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{\Lambda}^{-1}\boldsymbol{Y}$, $\boldsymbol{V}(\hat{\boldsymbol{\beta}}) = \sigma^2(\boldsymbol{C}^T\boldsymbol{C})^{-1} = \sigma^2(\boldsymbol{A}^T\boldsymbol{\Lambda}^{-1}\boldsymbol{A})^{-1}$.

II.1.38. **II.1.37** pratime reikia imti $\boldsymbol{\Lambda} = [\lambda_{ij}]_{n\times n}$; $\lambda_{ii} = 1/\omega_i$, $\lambda_{ij} = 0, i \neq j$. Gauname $\hat{\theta} = \sum_i(\omega_i Y_i)/\sum_i \omega_i$, $\boldsymbol{V}(\hat{\theta}) = \sigma^2/\sum_i \omega_i$.

II.1.39. **II.1.37** pratime reikia imti $\boldsymbol{\Lambda} = [\lambda_{ij}]_{n\times n}$; $\lambda_{ii} = i^2$, $\lambda_{ij} = 0, i \neq j$; $\boldsymbol{A}^T = (1,\ 2, ..., n)$. Gauname $\hat{\theta} = [\sum_i(Y_i/i)]/n$, $\boldsymbol{V}(\hat{\theta} = \sigma^2/n$.

II.1.40. $\hat{\boldsymbol{\beta}} \sim N_m(\boldsymbol{\beta}, \sigma^2(\boldsymbol{A}^T\boldsymbol{\Lambda}^{-1}\boldsymbol{A})^{-1}), \hat{\sigma}^2 = s^2 = SS_E/(n-m)$, $SS_E = (\boldsymbol{B}^{-1})^T(\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}})^T(\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}})\boldsymbol{B}^{-1} \sim \sigma^2\chi^2_{n-m}$.

II.1.41. a) $\hat{d} = [\sum_i(X_iY_i)]/\sum_i X_i^2$; b) $\hat{d} = (\sum_i Y_i)/\sum_i X_i$; c) $\hat{d} = [\sum_i(Y_i/X_i)]/n$.

II.1.42. $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}$, $\hat{\beta}_1 = \sum_i Y_i(X_i - \bar{X})/\sum_i(X_i - \bar{X})^2$; $\boldsymbol{V}\hat{\beta}_1 = \sigma^2(1-\rho)/\sum_i(X_i - \bar{X})^2$, $\boldsymbol{V}\hat{\beta}_0 = \sigma^2[(1+2\rho)/3 + \bar{X}^2(1-\rho)/\sum_i(X_i - \bar{X})^2]$. Nagrinėkime tiesinį modelį $Y_i = \alpha + \beta_1(X_i - \bar{X}) + e_i$, $i = 1, 2, 3; \alpha = \beta_0 + \beta_1\bar{X}$. Tada

$$[\boldsymbol{A}^T\boldsymbol{\Lambda}^{-1}\boldsymbol{A}]^{-1} = \begin{pmatrix} (1+2\rho)/3 & 0 \\ 0 & (1-\rho)/\sum_i(X_i - \bar{X})^2 \end{pmatrix},$$

$$\boldsymbol{A}^T\boldsymbol{\Lambda}^{-1}\boldsymbol{Y} = \begin{pmatrix} \sum_i Y_i/(1+2\rho) \\ \sum_i Y_i(X_i - \bar{X})/(1-\rho) \end{pmatrix}$$

ir lieka pasinaudoti **II.1.37** pratimo sprendimu.

II.1.43. Atlikime ortogonalią transformaciją $\boldsymbol{Z} = \boldsymbol{C}\boldsymbol{Y}$, kai matricos $\boldsymbol{C}$ pirmoji eilutė yra $(1/n, ..., 1/n)$. Tada iš ortogonalumo sąlygos

$$\sum_{j=1}^{n} c_{ij} c_{1j} = \sum_{j=1}^{n} c_{ij}/n = 0, \quad i = 2, 3, ..., n.$$

Tarkime, $\boldsymbol{B}$ yra matrica $\boldsymbol{C}$ be pirmosios eilutės. Tada a. v. $\boldsymbol{Z}_2 = (Z_2, ..., Z_n)^T = \boldsymbol{B}\boldsymbol{Y}$ kovariacinė matrica

$$\boldsymbol{V}(\boldsymbol{B}\boldsymbol{Y}) = \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^T = \rho\sigma^2 \boldsymbol{B}\mathbf{E}\boldsymbol{B}^T + (1-\rho)\sigma^2 \boldsymbol{B}\boldsymbol{I}\boldsymbol{B}^T,$$

čia $\mathbf{E}$ – matrica, kurios visi elementai lygūs 1. Gauname, kad pirmasis dėmuo lygus 0, nes $\boldsymbol{B}\mathbf{E} = \mathbf{0}$. Taigi a. v. $\boldsymbol{Z}_2$ kovariacinė matrica yra diagonali $\boldsymbol{V}(\boldsymbol{Z}_2) = (1-\rho)\sigma^2 \boldsymbol{I}$.

II.1.44. a) Pradžioje tarkime, kad $\boldsymbol{\gamma}$ yra žinomas. Gauname tiesinį modelį

$$\boldsymbol{Y} - \boldsymbol{B}\boldsymbol{\gamma} = \boldsymbol{A}\boldsymbol{\beta} + \boldsymbol{e}.$$

Iš čia parametro $\boldsymbol{\beta}$ mažiausiųjų kvadratų įvertinys

$$\tilde{\boldsymbol{\beta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T(\boldsymbol{Y} - \boldsymbol{B}\boldsymbol{\gamma}).$$

Įrašę šį įvertinį į pradinę išraišką ir pažymėję $\boldsymbol{R} = \boldsymbol{I} - \boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T$, gausime tiesinį modelį parametro $\boldsymbol{\gamma}$ atžvilgiu

$$\boldsymbol{R}\boldsymbol{Y} = \boldsymbol{R}\boldsymbol{B}\boldsymbol{\gamma} + \boldsymbol{e}.$$

Parametro $\boldsymbol{\gamma}$ mažiausiųjų kvadratų įvertinys

$$\boldsymbol{\gamma}^* = (\boldsymbol{B}^T\boldsymbol{R}^T\boldsymbol{R}\boldsymbol{B})^{-1}\boldsymbol{B}^T\boldsymbol{R}^T\boldsymbol{R}\boldsymbol{Y} = (\boldsymbol{B}^T\boldsymbol{R}\boldsymbol{B})^{-1}\boldsymbol{B}^T\boldsymbol{R}\boldsymbol{Y}.$$

Įrašę į $\tilde{\boldsymbol{\beta}}$ išraišką, gauname parametro $\boldsymbol{\beta}$ įvertinį

$$\boldsymbol{\beta}^* = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T(\boldsymbol{Y} - \boldsymbol{B}\boldsymbol{\gamma}^*).$$

b) Liekamoji kvadratų suma

$$SS_E = (\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{\beta}^* - \boldsymbol{B}\boldsymbol{\gamma}^*)^T(\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{\beta}^* - \boldsymbol{B}\boldsymbol{\gamma}^*) = (\boldsymbol{Y} - \boldsymbol{B}\boldsymbol{\gamma}^*)^T\boldsymbol{R}(\boldsymbol{Y} - \boldsymbol{B}\boldsymbol{\gamma}^*).$$

II.2. Dispersinė analizė

II.2.1. Vienfaktorė dispersinė analizė

II.2.1. Tarkime, a. d. Y skirstinys gali priklausyti nuo faktoriaus A, kuris gali būti I lygmenų $A_1, A_2, ..., A_I$. Tegu a. d. Y skirstinys, kai faktoriaus A lygmuo yra A_i, yra normalusis $N(\mu_i, \sigma^2)$. Tarkime, kad turime I paprastųjų nepriklausomų imčių; i-ąją imtį sudaro J_i elementų $Y_{i1}, Y_{i2}, ..., Y_{iJ_i}$, gautų, kai faktoriaus A lygmuo yra A_i, $i = 1, ..., I$.

a) Užrašykite jungtinę imtį $\boldsymbol{Y}$ tiesinio modelio pavidalu (žr. **II.1.1** pratimą).

b) Įrodykite, kad $\boldsymbol{T} = (\bar{Y}_{1\cdot}, ..., \bar{Y}_{I\cdot}, \boldsymbol{Y}^T\boldsymbol{Y})^T$, $\bar{Y}_{i\cdot} = \sum_j Y_{ij}/J_i$, yra pilnoji ir pakankamoji parametro $\boldsymbol{\theta} = (\mu_1, ..., \mu_I, \sigma^2)$ statistika. c) Raskite parametro $\boldsymbol{\theta}$ elementų NMD įvertinius.

II.2.2. (**II.2.1** pratimo tęsinys). Raskite parametrų

$$\mu = \frac{1}{n}\sum_{i=1}^{I} J_i\mu_i, \quad n = J_1 + ... + J_I, \quad \alpha_i = \mu_i - \mu, \ \ i = 1, ..., I$$

NMD įvertinius.

II.2.3. (**II.2.1** pratimo tęsinys). Pagrindinė dispersinės analizės hipotezė yra patikrinti, ar priklauso Y skirstinys nuo faktoriaus A. Modelio parametrų terminais reikia patikrinti hipotezę $H_A : \mu_1 = ... = \mu_I$ (arba $H_A : \alpha_1 = ... = \alpha_I = 0$). Raskite kriterijų hipotezei H_A tikrinti.

II.2.4. (**II.2.3** pratimo tęsinys). a) Raskite vidutinių kvadratų sumų vidurkius. b) Raskite **II.2.3** pratime pateikto kriterijaus galią.

II.2.5. (**II.2.3** pratimo tęsinys). Jeigu hipotezė H_A neatmetama, tai analizę galima tuo ir užbaigti. Priešingu atveju natūraliai iškyla klausimas, kaip stebėjimai Y_{ij} priklauso nuo faktoriaus A lygmenų. Kartais remiantis papildoma informacija apie faktoriaus lygmenis galima juos sudalinti į grupes taip, kad hipotezės H_A atmetimą sąlygotų skirtumai tarp grupių, o grupių viduje vidurkiai skirtųsi nežymiai. Raskite kriterijų hipotezei $H'_A : \mu_1 = ... = \mu_r; \mu_{r+1} = ... = \mu_I$ tikrinti.

II.2.6. (**II.2.5** pratimo tęsinys). Jeigu papildomos informacijos apie faktoriaus lygmenis nepakanka, tai galima taikyti statistinius stebėjimų palyginimo metodus esant įvairiems faktoriaus lygmenims, vadinamus kontrastų analize. Parametrų $\mu_1, ..., \mu_I$ *kontrastu* vadinama tiesinė funkcija

$$\psi = \sum_{i=1}^{I} c_i\mu_i, \quad \sum_{i=1}^{I} c_i = 0, \quad c_i \in \mathbf{R}, \quad i = 1, ..., I.$$

Hipotezė $H_A : \mu_1 = ... = \mu_I$ ekvivalenti tvirtinimui, kad visi kontrastai ψ lygūs nuliui. Jei H_A neteisinga, tai atsiras kontrastų, kurie nelygūs nuliui.

Įrodykite, kad hipotezė H_A atmetama reikšmingumo lygmens α kriterijumi (**II.2.3** pratimas) tada ir tik tada, kai egzistuoja kontrastas ψ, kad intervalas

$$(\hat\psi - \sqrt{(I-1)F_\alpha}(I-1, n-I)\sqrt{\hat{\boldsymbol{V}}(\hat\psi)}, \hat\psi + \sqrt{(I-1)F_\alpha}(I-1, n-I)\sqrt{\hat{\boldsymbol{V}}(\hat\psi)})$$

neuždengia 0; čia

$$\hat\psi = \sum_{i=1}^{I} c_i\hat\mu_i = \sum_{i=1}^{I} c_i\bar{Y}_{i.}, \quad \boldsymbol{V}(\hat\psi) = \sigma^2\sum_{i=1}^{I} c_I^2/J_i, \quad \hat{\boldsymbol{V}}(\hat\psi) = s^2\sum_{i=1}^{I} c_I^2/J_i.$$

II.2.7. Tegu $Y_{ij} = \mu + \alpha_i + e_{ij}, i = 1, ..., I, j = 1, ..., J$; čia e_{ij} yra vienodai pasiskirstę n. a. d. ir $\mathbf{E}(e_{ij}) = 0$, o parametrai α_i tenkina sąlygą $\sum_i d_i\alpha_i = 0$, kai d_i žinomi ir $\sum_i d_i \neq 0$. Raskite parametrų μ ir α_i mažiausiųjų kvadratų įvertinius.

II.2.8. Tegu $Y_{ij} = \mu_i + e_{ij}, i = 1, ..., I, j = 1, ..., J$; čia e_{ij} yra nepriklausomi ir normalieji $e_{ij} \sim N(0, \sigma^2)$. a) Raskite hipotezės $H : \mu_1 = 2\mu_2 = 3\mu_3$ tikrinimo kriterijų, kai $I = 4$. b) Įrodykite, kad hipotezės $H : \mu_1 = \mu_2$ tikrinimo F kriterijus, kai $I = 2$, yra ekvivalentus Stjudento kriterijui dėl vidurkių lygybės dviejose normaliosiose imtyse.

II.2.9. Lentelėje pateikti duomenys, apibūdinantys iškvepiamo azoto kiekį Y esant keturioms skirtingoms dietoms (faktoriaus A lygmenys).

A_1	A_2	A_3	A_4
4,079	4,368	4,169	4,928
4,859	5,668	5,709	5,608
3,540	3,752	4,416	4,940
5,047	5,848	5,666	5,291
3,298	3,802	4,123	4,674

A_1	A_2	A_3	A_4
4,679	4,844	5,059	5,038
2,870	3,578	4,403	4,905
4,648	5,393	4,496	5,208
3,847	4,374	4,688	4,806

Atlikite duomenų vienfaktorę dispersinę analizę.

II.2.10. Iš darbininkų, aptarnaujančių didelės įmonės surinkimo konvejerį, buvo atrinkti 4 darbininkai ir kiekvienam iš jų buvo užfiksuotas tam tikros detalės surinkimo laikas.

Darbininkai	Laikas							
1	24,2;	22,2;	24,5;	21,1;	22,0;			
2	19,4;	21,1;	16,2;	21,2;	21,6;	17,8;	19,6;	
3	19,0;	23,1;	23,8;	22,8;				
4	19,9;	15,7;	15,2;	19,8;	18,9;	16,1;	16,2;	18,5

Ar skiriasi darbininkai pagal detalės surinkimo laiką?

II.2.11. Lentelėje pateikti duomenys, apibūdinantys gumos tempiamąjį atsparumą $Y kg/cm^2$ priklausomai nuo vulkanizavimo laiko X min.

X_i	Y_{ij}			
20	152	152	147	152
25	158	155	146	169
30	149	159	115	152
40	143	121	116	156
60	126	165	153	157

Patikrinkite hipotezę, kad gumos tempiamojo atsparumo vidurkis nepriklauso nuo vulkanizavimo laiko.

II.2.12. Tiriant retųjų elementų pasiskirstymą triaso amžiaus nuogulose netoli Birštono, buvo gauti lentelėje pateikti rezultatai (g/t) trijuose šių nuogulų horizontuose ($\bar{X}$ ir s empiriniai vidurkiai ir vidutinio kvadratinio nuokrypio įvertiniai; stebėjimų skaičius atitinkamai I, II ir III horizontuose yra 136, 77, 111). Patikrinkite hipotezes, kad elementų koncentracija visuose trijuose horizontuose yra vienoda.

Elementai	I		II		III	
	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
Varis	43	12	47	12	44	22
Švinas	13	4	16	5	22	8
Titanas	3428	701	3531	776	4255	1071
Manganas	940	182	1022	146	828	296
Chromas	74	21	84	22	110	64
Nikelis	44	13	55	16	60	22

II.2.13. Tarkime, iš pagamintos produkcijos atsitiktinai atrinkta I gaminių. Kiek-

vienam iš jų su paklaida J kartų matuojama tam tikro požymio X reikšmė. Tarkime, kad požymio ir paklaidos skirstiniai yra nepriklausomi normalieji a. d. Paklaidos vidurkis lygus nuliui. Reikia įvertinti požymio parametrus (vidurkį ir dispersiją). Parinkite tinkamą statistinį modelį.

II.2.14. Nagrinėjamas vienfaktorės dispersinės analizės modelis su atsitiktiniu faktoriumi

$$Y_{ij} = \mu + a_i + e_{ij}, \quad i = 1, ..., I, \quad j = 1, ..., J;$$

paklaidos $e_{ij} \sim N(0, \sigma^2)$, a. d. $a_i \sim N(0, \sigma_A^2)$ ir a. d. $\{a_i\}, \{e_{ij}\}$ nepriklausomi. Įrodykite, kad parametro $\boldsymbol{\theta} = (\mu, \sigma_A^2, \sigma^2)^T$ pilnoji ir pakankamoji statistika yra

$$\boldsymbol{T} = (\bar{Y}_{..}, \sum_i \bar{Y}_{i.}^2, \sum_i \sum_j Y_{ij}^2)^T.$$

II.2.15. (**II.2.14** pratimo tęsinys). Raskite parametrų $\mu, \tau^2 = \sigma^2 + J\sigma_A^2, \sigma_A^2, \sigma^2$ NMD įvertinius.

II.2.16. (**II.2.14** pratimo tęsinys). Įrodykite, kad a. d. $\bar{Y}_{..}, SS_A, SS_E$ yra nepriklausomi.

II.2.17. (**II.2.14** pratimo tęsinys). Raskite parametrų $\mu, \tau^2, \sigma^2, \sigma_A^2$ pasikliovimo intervalus ir sudarykite kriterijus hipotezėms dėl šių parametrų reikšmių tikrinti.

II.2.18. (**II.2.14** pratimo tęsinys). Raskite pagrindinės dispersinės analizės hipotezės šioje schemoje analogo $H_A : \sigma^2 = 0$ tikrinimo kriterijų ir jo galios funkciją.

II.2.19. Atsitiktinai parinkus keturis gaminius po 10 kartų buvo išmatuotas jų požymis X:

x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4
6,34	5,95	5,23	4,55	6,85	6,52	5,52	4,73
6,36	6,04	5,27	4,65	6,91	6,60	5,52	4,78
6,41	6,11	5,32	4,68	6,91	6,62	5,53	4,78
6,42	6,31	5,39	4,68	7,02	6,64	5,60	4,84
6,80	6,36	5,40	4,72	7,12	6,71	5,78	4,86

a) atlikite vienfaktorę dispersinę analizę su vienu atsitiktiniu faktoriumi A (jo lygmenys – gaminių numeriai);

b) raskite taškinius ir intervalinius ($Q = 0,95$) parametrų μ, σ_A^2, σ^2 įverčius;

c) raskite parametro σ_A^2/σ^2 pasikliovimo intervalą ($Q = 0,95$).

II.2.20. Konservų fabriko technologiniame procese kiekvienas pjaustantis abrikosus operatorius buvo stebimas penkis dviejų minučių laikotarpius. Trijose skirtingose linijose buvo pjaustomi trijų skirtingų dydžių vaisiai (didesnis numeris reiškia mažesnį vaisių) ir užfiksuojamas supjaustytų vaisių skaičius Y_{ij}; $i = 1, ..., I$ yra operatoriaus numeris, o $j = 1, ..., 5$ žymi laikotarpio numerį. Analizės rezultatai atskirai kiekvieno dydžio vaisiams pateikti lentelėje.

Dydis	I	$Y_{..}$	MS_A	MS_E
2	9	53,17	59,72	1,144
3	17	52,26	68,20	2,537
4	17	47,32	78,96	4,926

Tardami, kad yra teisingas modelis $Y_{ij} = \mu + a_i + e_{ij}$, kai a. d. $\{a_i\}$ ir $\{e_{ij}\}$ yra nepriklausomi ir normalieji $a_i \sim N(0, \sigma_A^2)$, $e_{ij} \sim N(0, \sigma^2)$, raskite taškinius parametrų $\mu, \sigma_A^2, \sigma^2$ įverčius.

II.2.21. Tarkime, kad **II.2.3** pratime faktorius A yra atsitiktinis. Kaip pasikeis dispersinė analizė?

II.2.2. Dvifaktorė dispersinė analizė

II.2.22. Tiriama požymio Y priklausomybė nuo faktoriaus A, kuris yra I lygmenų $A_1, ..., A_I$, ir faktoriaus B, kuris yra J lygmenų $B_1, ..., B_J$. Kiekvienam lygmenų dariniui (A_i, B_j) stebėjimai pakartojami po $K > 1$ kartų. Imties elementai aprašomi dvifaktorės dispersinės analizės modeliu

$$Y_{ijk} = \mu_{ij} + e_{ijk}, \quad i = 1, ..., I, \quad j = 1, ..., J, \quad k = 1, ..., K.$$

Tarsime, kad paklaidos e_{ijk} nepriklausomos ir $e_{ijk} \sim N(0, \sigma^2)$. Modelis visiškai nusakytas IJ parametrais $\mu_{11}, ...\mu_{IJ}$ ir dispersija $\sigma^2 = \boldsymbol{V}(Y_{ijk}) = \boldsymbol{V}(e_{ijk})$. a) Įrodykite, kad $\boldsymbol{T} = (\bar{Y}_{11.}, ..., \bar{Y}_{IJ.}, \sum_i \sum_j \sum_k Y_{ijk}^2)^T$ yra parametro $\boldsymbol{\theta} = (\mu_{11}, ..., \mu_{IJ}, \sigma^2)^T$ pilnoji ir pakankamoji statistika. b) Raskite parametro $\boldsymbol{\theta}$ elementų NMD įvertinius.

II.2.23. (**II.2.22** pratimo tęsinys). Raskite parametrų

$$\mu, \quad \alpha_i = \bar{\mu}_i - \mu, \quad \beta_j = \bar{\mu}_j - \mu, \quad \gamma_{ij} = \mu_{ij} - \bar{\mu}_i - \bar{\mu}_j + \mu;$$

$$\mu = \frac{1}{IJ} \sum_i \sum_j \mu_{ij}, \quad \bar{\mu}_i = \frac{1}{J} \sum_j \mu_{ij}, \quad \bar{\mu}_j = \frac{1}{I} \sum_i \mu_{ij};$$

NMD įvertinius.

II.2.24. (**II.2.22** pratimo tęsinys). Raskite kriterijus pagrindinėms dispersinės analizės hipotezėms $H_A : \alpha_i \equiv 0$, $H_B : \beta_j \equiv 0$, $H_{AB} : \gamma_{ij} \equiv 0$ tikrinti.

II.2.25. (**II.2.24** pratimo tęsinys). a) Įrodykite, kad kvadratų sumos SS_A, SS_B, SS_{AB} ne tik nepriklauso nuo SS_E (remiantis **II.2.21** pratimu), bet nepriklausomos ir tarpusavyje bei nepriklauso nuo $\bar{Y}_{...}$. b) Įrodykite, kad kvadratų sumas galima skaičiuoti tokiu būdu: reikia pakelti kiekvieną daugianarių dėmenų kvadratu paliekant buvusius ženklus ir susumuoti. Pavyzdžiui, $SS_{AB} = K \sum_i \sum_j \sum_k \bar{Y}_{ij.}^2 - JK \sum_i \bar{Y}_{i..} - IK\bar{Y}_{.j.} + IJK\bar{Y}_{...}^2$. c) Patikrinkite lygybę $SS_T = SS_A + SS_B + SS_{AB} + SS_E$, čia $SS_T = \sum_i \sum_j \sum_k (Y_{ijk} - \bar{Y}_{...})$.

II.2.26. (**II.2.24** pratimo tęsinys). a) Raskite vidutinių kvadratų sumų vidurkius. b) Raskite **II.2.24** pratime pateiktų kriterijų galios funkcijas.

II.2.27. Tarkime, kad dvifaktorės dispersinės analizės schemoje turime po vieną stebėjimą langelyje, t. y. $K = 1$. Tada stebėjimų Y_{ij} skaičius $n = IJ$ lygus nežinomų parametrų μ_{ij} skaičiui ir liekamoji kvadratų suma $SS_E = 0$. Parametrų skaičius sumažinamas tariant, kad tarp faktorių nėra sąveikos, t. y. $\gamma_{ij} \equiv 0$. Gauname adityvųjį modelį

$$Y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}, \quad i = 1, ..., I, \quad j = 1, ..., J.$$

Tarsime, kad paklaidos e_{ij} nepriklausomos ir $e_{ij} \sim N(0, \sigma^2)$, o parametrai α_i, β_j tenkina sąlygas

$$\sum_i \alpha_i = 0, \quad \sum_j \beta_j = 0.$$

Modelį visiškai nusako $I + J - 1$ parametras $\mu, \alpha_1, ..., \alpha_{I-1}, \beta_1, ..., \beta_{J-1}$ ir dispersija σ^2.
a) Raskite parametrų α_i, β_j mažiausiųjų kvadratų įvertinius ir liekamąją kvadratų sumą.
b) Įrodykite, kad gautieji įvertiniai yra NMD įvertiniai.

II.2.28. (**II.2.27** pratimo tęsinys). Raskite kriterijus faktorių įtakos nebuvimo hipotezėms $H_A : \alpha_i = 0$, $H_B : \beta_j = 0$ tikrinti.

II.2.29. (Tjukio kriterijus). **II.2.28** pratime pateikti kriterijai nėra korektiški, jeigu prielaida $\gamma_{ij} \equiv 0$ nėra teisinga. Sudarykite kriterijų dėl prielaidos $\gamma_{ij} \equiv 0$ teisingumo remdamiesi stebėjimais, kai $K = 1$, o sąveika tenkina sąlygą $\gamma_{ij} = \gamma \alpha_i \beta_j$, t. y. patikrinkite parametrinę hipotezę $H_\gamma : \gamma = 0$.

II.2.30. Tegu dvifaktorės dispersinės analizės schemoje faktoriaus A lygmenų skaičius $I = 2$, o faktoriaus B lygmenų skaičius $J \geq 2$; kiekviename langelyje turime po vieną stebėjimą $Y_{ij}, i = 1, 2; j = 1, ..., J$. Įrodykite, kad hipotezės H_A tikrinimo F kriterijus yra ekvivalentus Stjudento kriterijui, grindžiamam skirtumais $Z_j = Y_{1j} - Y_{2j}$, $j = 1, ..., J$. Todėl dispersinės analizės prielaidos gali būti susilpnintos: kriterijus nepakis, jeigu tarsime, kad paklaidų vektoriai $(e_{1j}, e_{2j})^T, j = 1, ..., J$ yra nepriklausomi ir turi dvimatį normalųjį skirstinį su nuliniu vidurkių vektoriumi.

II.2.31(**II.2.30** pratimo tęsinys). Lentelėje pateikti miego trukmės pakitimo duomenys Y_{1j}, naudojant pirmo tipo migdomuosius vaistus, ir Y_{2j} – naudojant antro tipo migdomuosius vaistus; čia j žymi paciento numerį, $j = 1, ..., 10$.

$i; j$	1	2	3	4	5	6	7	8	9	10
1	$+0,7$	$-1,6$	$-0,2$	$-1,2$	$-1,0$	$+3,4$	$+3,7$	$+0,8$	$0,0$	$+2,0$
2	$+1,9$	$+0,8$	$+1,1$	$+0,1$	$-0,1$	$+4,4$	$+5,5$	$+1,6$	$+4,6$	$+3,4$

a) Priėmę normalumo prielaidą patikrinkite hipotezę, kad miego trukmės padidėjimo vidurkiai, naudojant pirmo ir antro tipo migdomuosius, nesiskiria; $\alpha = 0,01$.

b) Tarkime, yra žinoma, kad $\boldsymbol{V}(Y_{1j} - Y_{2j}) \leq 1,25$. Kokį skaičių pacientų reikėtų turėti, kad, tikrinant vidurkių lygybės hipotezę, ji būtų atmesta su tikimybe, ne mažesne už 0,95, kai miego trukmės padidėjimo vidurkių skirtumas viršija 1 valandą.

II.2.32. Tarkime, kad **II.2.9** pratime tiriame iškvepiamo azoto kiekio priklausomybę ne tik nuo dietos (faktorius A), bet ir nuo lyties (faktorius B). Gauti stebėjimų duomenys pateikti lentelėje [1].

	A_1	A_2	A_3	A_4
B_1	4,079	4,368	4,169	4,928
	4,859	5,668	5,709	5,608
	3,540	3,752	4,416	4,940
B_2	2,870	3,578	4,403	4,905
	4,648	5,393	4,496	5,208
	3,847	4,374	4,688	4,806

Atlikite dvifaktorę dispersinę analizę. Patikrinkite pagrindines dispersinės analizės hipotezes.

II.2.33. Iš kiekvienos 4 pelių (faktorius A) palikuonių buvo atrinkta po 1 peliuką ir jam buvo taikoma viena iš trijų dietų (faktorius B). Po trijų savaičių išmatuotas svorio prieaugis Y [1].

	A_1	A_2	A_3	A_4
B_1	5,2	11,4	4,2	10,7
B_2	7,4	13,0	9,5	8,8
B_3	9,1	13,8	8,8	13,0

Atlikite dvifaktorę dispersinę analizę su vienu stebėjimu langelyje. Remdamiesi Tjukio kriterijumi patikrinkite sąveikos nebuvimo hipotezę.

II.2.34. Tiriant, kiek aštuonių skirtingų rūšių aliejaus (faktorius A) sugeria spurgos, šešias dienas (faktorius B) buvo gaminamos vienodo didumo spurgų partijos su kiekviena aliejaus rūšimi [16].

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
B_1	164	172	177	178	163	163	150	164
B_2	177	197	184	196	177	193	179	169
B_3	168	167	187	177	144	176	146	155
B_4	156	161	169	181	165	172	141	149
B_5	172	180	179	184	166	176	169	170
B_6	196	190	197	191	178	178	183	167

Užpildykite dispersinės analizės lentelę ir patikrinkite pagrindines hipotezes.

II.2.35 (**II.2.34** pratimo tęsinys). Tarkime, kad skirtingų aliejaus rūšių sugėrimo vidurkiai tenkina sąlygas $\mu_5 = \mu_7 = \mu_8 = \mu$, $\mu_1 = \mu_2 = \mu_6 = \mu + 12$, $\mu_3 = \mu_4 = \mu + 22$. Kokia tikimybė atmesti hipotezę H_B, jeigu paklaidos dispersija lygi **I.2.34** pratime surastam įverčiui (kriterijaus reikšmingumo lygmuo $\alpha = 0,05$)?

II.2.36 (**II.2.34** pratimo tęsinys). Kadangi 5, 7 ir 8 aliejaus rūšys atrodo ekonomiškiausios, tai tolesni eksperimentai bus atliekami tik su šiomis trimis rūšimis. Kiek eksperimentų reikėtų atlikti tikrinant hipotezę $H : \mu_5 = \mu_7 = \mu_8$ kriterijumi, kai reikšmingumo lygmuo $\alpha = 0,05$, kad bet kokį vidurkių skirtumą, viršijantį 10 vienetų, galėtume pastebėti su tikimybe, ne mažesne už 0,8?

II.2.37. Tiriama konservų dėžutės svorio priklausomybė nuo mėsos tiekėjo (faktorius A) ir nuo dėžutes užpildančio automato užpildymo cilindro (faktorius B). Iš kiekvieno tiekėjo ir kiekvieno cilindro konservų dėžučių partijos atsitiktinai atrenkama po 3 dėžutes. Dėžučių svoriai (sąlyginiais vienetais) pateikti lentelėje [16].

	A_1	A_2	A_3	A_4	A_5
B_1	1 1; 2	4; 3; 5	6; 3; 7	3; 1; 3	1; 3; 3
B_1	−1; 3; −1	−2; 1; 0	3; 1; 5	2; 0; 1	1; 0; 1
B_3	1; 1; 1	2; 0; 1	2; 4; 3	1; 3; 3	3; 3; 3
B_4	−2; 3; 0	−2; 0; 1	3; 3; 4	0; 0; 2	0; 1; 1
B_5	1; 1; −1	2; 1; 5	0; 1; 2	1; 0; −1	−2; 3; 1
B_6	0; 1; 1	0; 0; 3	3; 3; 4	3; 0; 2	3; 1; 2

Užpildykite dispersinės analizės lentelę ir patikrinkite pagrindines hipotezes.

II.2.38. Eksperimente hibridinius žiurkiukus maitino hibridinės žiurkių patelės. Lentelėje pateikti žiurkiukų svoriai praėjus 28 dienoms nuo gimimo. Šiame eksperimente faktorius A yra maitinančios žiurkės genotipas, o faktorius B – žiurkiukų vados genotipas.

	A_1	A_2	A_3	A_4		A_1	A_2	A_3	A_4
B_1	61,5	55,0	52,5	42,0	B_3	37,0	56,3	39,6	50,0
	68,2	42,0	61,8	54,0		36,3	69,8	46,0	43,8
	64,0	60,2	49,5	61,0		68,0	67,0	61,3	54,5
	65,0		52,7	48,2				55,3	
	59,7			39,6				55,7	
B_2	60,3	50,8	56,5	51,3	B_4	59,0	59,5	45,2	44,8
	51,7	64,7	59,0	40,5		57,4	52,8	57,0	51,5
	49,3	61,7	47,2			54,0	56,0	61,4	53,0
	48,0	64,0	53,0						42,0
		62,0							54,0

Atlikite dvifaktorę dispersinę analizę, kai stebėjimų skaičiai langeliuose skirtingi. Patikrinkite sąveikos nebuvimo hipotezę. Patikrinkite faktorių A ir B įtakos nebuvimo hipotezes dviem būdais: a) tariant, kad modelis adityvus; b) neatsižvelgiant į sąveiką.

II.2.39. Įrodykite, kad dvifaktorėje dispersinėje analizėje, kai $I = 2$, kvadratų sumas SS_A ir SS_{AB} galima apskaičiuoti šitaip:

$$SS_A = JK(\bar{Y}_{1..} - \bar{Y}_{2..})^2/2,$$

$$SS_{AB} = (K\sum_j (\bar{Y}_{1j.} - \bar{Y}_{2j.})^2 - SS_A)/2,$$

o jei $K = 2$, tai

$$SS_E = \sum_i \sum_j (Y_{ij1} - Y_{ij2})^2/2.$$

II.2.40. Tarkime, kad imties $\boldsymbol{Y}$ skirstinys gali priklausyti nuo dviejų atsitiktinių faktorių. Imties $\boldsymbol{Y}$ elementai turi tokią struktūrą (žr. [3], 2.5.1 skyrelį):

$$Y_{ijk} = \mu + a_i + b_j + c_{ij} + e_{ijk}, \quad i = 1, ..., I, \quad j = 1, ..., J, \quad k = 1, ..., K;$$

čia $\{a_i\}, \{b_j\}, \{c_{ij}\}, \{c_{ij}\}, \{e_{ijk}\}$ yra nepriklausomi a. d. ir $a_i \sim N(0, \sigma_A^2)$, $b_j \sim N(0, \sigma_B^2)$, $c_{ij} \sim N(0, \sigma_{AB}^2)$, $e_{ijk} \sim N(0, \sigma^2)$. Modelį visiškai nusako vidurkis μ ir keturios dispersijos $\sigma_A^2, \sigma_B^2, \sigma_{AB}^2, \sigma^2$. Įrodyta (žr. [15], 4f.3 skyrelį), kad parametro $\boldsymbol{\theta} = (\mu, \sigma_A^2, \sigma_B^2, \sigma_{AB}^2, \sigma^2)^T$ pilnoji ir pakankamoji statistika yra $\boldsymbol{T} = (\bar{Y}_{...}, SS_A, SS_B, SS_{AB}, SS_E)^T$. a) Raskite vidutinių kvadratų sumų vidurkius. b) Raskite parametro $\boldsymbol{\theta}$ elementų NMD įvertinius.

II.2.41. (**II.2.40** pratimo tęsinys). a) Įrodykite, kad kvadratų sumos SS_A, SS_B, SS_{AB}, SS_E yra nepriklausomos. b) Raskite įvertinių $\hat{\sigma}_A^2$, $\hat{\sigma}_B^2$, $\hat{\sigma}_{AB}^2$, $\hat{\sigma}^2$ dispersijas. c) Sudarykite pagrindinių dispersijos analizės hipotezių analogų $H_A : \sigma_A^2 = 0$, $H_B : \sigma_B^2 = 0$, $H_{AB} : \sigma_{AB}^2 = 0$ tikrinimo kriterijus ir raskite jų galios funkcijas.

II.2.42. Tegu $Y_{ijk} = \mu + a_i + b_j + e_{ijk}$ ir a. d. $\{a_i\}$, $\{b_j\}$, $\{e_{ijk}\}$ nepriklausomi, $a_i \sim N(0, \sigma_A^2)$, $b_j \sim N(0, \sigma_B^2)$, $e_{ijk} \sim N(0, \sigma^2)$; $i = 1, ..., I, j = 1, ..., J, k = 1, ..., K$. Parametrai $\mu, \sigma_A^2, \sigma_B^2, \sigma^2$ nežinomi. Raskite TGN kriterijų

a) hipotezei $H : \tau^2 = \sigma_A^2/(\sigma^2 + K\sigma^2) \leq \Delta$, kai alternatyva $\bar{H} : \tau^2 > \Delta$, tikrinti;

b) hipotezei $H : \sigma_{AB}/\sigma^2 \leq \Delta$, kai alternatyva $\bar{H} : \sigma_{AB}^2/\sigma^2 > \Delta$, tikrinti.

II.2.43. Atlikus dvifaktorę dispersinę analizę su dviem atsitiktiniais faktoriais A ir B, gauta tokia dispersinės analizės lentelė

Faktorius	ν	$\overline{SS}$
A	24	3 243
B	3	46 659
$A \times B$	72	459
E	1100	243

a) Patikrinkite hipotezes H_A, H_B, H_{AB}, kai reikšmingumo lygmuo $\alpha = 0,025$.

b) Raskite dispersijos komponenčių σ_A^2, σ_B^2, σ_{AB}^2, σ^2 įverčius ir jų dispersijų įverčius.

c) Apskaičiuokite kiekvienos dispersijos komponentės apytikslį pasikliovimo intervalą, kai pasikliovimo lygmuo $Q = 0,95$.

II.2.44. Atlikite dispersinę analizę pagal **II.2.33** pratimo duomenis tarę, kad abu faktoriai yra atsitiktiniai.

II.2.45. Atlikite dispersinę analizę pagal **II.2.37** pratimo duomenis tarę, kad abu faktoriai yra atsitiktiniai.

II.2.46. Tarkime, kad Y skirstinys gali priklausyti nuo pastovaus faktoriaus A ir atsitiktinio faktoriaus B. Imties $\boldsymbol{Y}$ elementai turi tokią struktūrą (žr. [3], 2.6.1 skyrelį):

$$Y_{ijk} = \mu + \alpha_i + b_j + c_{ij} + e_{ijk}, \quad i = 1, ..., I, \quad j = 1, ..., J, \quad k = 1, ..., K,$$

čia μ, α_i konstantos, o b_j, c_{ij}, e_{ijk} atsitiktiniai dydžiai,

$$\sum_i \alpha_i = 0, \quad \sum_i c_{ij} \equiv 0, \quad \mathbf{E} b_j = \mathbf{E} c_{ij} = \mathbf{E} e_{ijk} = 0.$$

Tarsime, kad a. v. $(b_j, c_{1j}, ..., c_{Ij})^T$ skirstinys yra daugiamatis normalusis (su skirtingais j šie vektoriai nepriklausomi), o a. d. $e_{ijk} \sim N(0, \sigma^2)$ nepriklausomi tarpusavyje ir nepriklauso nuo šių vektorių. Žymėsime $\boldsymbol{V} b_j = \sigma_B^2, \boldsymbol{V} c_{ij} = \sigma_{AB;i}^2, i = 1, ..., I$. Analogiškai pirmesnėms schemoms sudarykime kvadratų sumas SS_A, SS_B, SS_{AB}, SS_E. Raskite vidutinių kvadratų sumų vidurkius.

II.2.47. (**II.2.46** pratimo tęsinys). a) Raskite parametrų $\sigma_A^2, \sigma_B^2, \sigma_{AB}^2, \sigma^2$ nepaslinktuosius įvertinius. b) Sudarykite kriterijus pagrindinių dispersinės analizės hipotezių analogams $H_A : \sigma_A^2 = 0$, $H_B : \sigma_B^2 = 0$, $H_{AB} : \sigma_{AB}^2 = 0$ tikrinti.

II.2.48. Lentelėje pateikti duomenys, apibūdinantys kuro ištekėjimo iš trijų skirtingų tipų tūtų greitį; matavimus atliko 5 operatoriai, iš kurių kiekvienas atliko po 3 matavimus kiekvienoje tūtoje.

Tūta	1			2			3		
A	6	6	−15	26	12	5	11	4	4
B	13	6	13	4	4	11	17	10	17
C	10	10	−11	−35	0	−14	11	−10	−17

Tūta	4			5		
A	21	14	7	25	18	25
B	−5	2	−5	15	8	1
C	−12	−2	−16	−4	10	24

a) Atlikite mišraus modelio, kuriame faktorius A (tūtos) yra pastovus, o faktorius B (operatorius) – atsitiktinis, dispersinę analizę.

b) Atlikite analizę laikydami abu faktorius pastoviais. Kuo paaiškinti skirtingus atsakymus, gautus tikrinant hipotezę, kad rezultatai nepriklauso nuo tūtų tipo.

II.2.49. Lentelėje pateikta tam tikros medžiagos nepralaidumo vandeniui charakteristika priklausomai nuo trijų tipų staklių (faktorius A), su kuriomis ji buvo pagaminta, per 9 skirtingas dienas (faktorius B) [16].

	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9
A_1	1,40	1,45	1,91	1,89	1,77	1,66	1,92	1,84	1,54
	1,35	1,57	1,48	1,48	1,73	1,54	1,93	1,79	1,43
	1,62	1,82	1,89	1,39	1,54	1,68	2,13	2,04	1,70
A_2	1,31	1,24	1,51	1,67	1,23	1,40	1,23	1,58	1,64
	1,63	1,18	1,58	1,37	1,40	1,45	1,51	1,63	1,07
	1,41	1,52	1,65	1,11	1,53	1,63	1,44	1.28	1,38
A_3	1,93	1,43	1,38	1,72	1,32	1,63	1,33	1,69	1,70
	1,40	1,86	1,36	1,37	1,34	1,36	1,38	1,80	1,84
	1,62	1,69	1,49	1,43	1,48	1,49	1,29	1,45	1,75

Atlikite duomenų dispersinę analizę, tarę, kad modelis yra mišrusis, kuriame faktorius A yra pastovus, o faktorius B – atsitiktinis.

II.2.50. Tiriant dujų sunaudojimą per devynias savaites (faktorius A) buvo fiksuojamas dujų sunaudojimas per parą kiekvieną savaitės dieną nuo pirmadienio iki šeštadienio imtinai (faktorius B). Gauti rezultatai (sąlyginiais vienetais) pateikiami lentelėje [16].

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9
B_1	5	1	-4	5	-13	-8	-2	-4	-10
B_2	3	6	-10	-2	-7	-2	-4	2	2
B_3	8	4	-14	-3	3	0	5	-11	-12
B_4	8	10	-5	-1	4	-2	4	1	-12
B_5	4	-1	7	-5	5	-3	-7	-3	-6
B_6	3	-9	3	-8	-6	0	-3	8	-1

Atlikite dispersinę analizę, tarę, kad faktorius A atsitiktinis, o faktorius B pastovus. Kaip pasikeistų atsakymai, jeigu tartume, kad abu faktoriai yra pastovūs?

II.2.51. Atlikite **II.2.32** pratimo duomenų analizę, tarę, kad faktorius A yra atsitiktinis, o faktorius B – pastovus.

II.2.52. Atlikite **II.2.34** pratimo duomenų analizę, tarę, kad faktorius A yra pastovus, o faktorius B – atsitiktinis.

II.2.3. Daugiafaktorė dispersinė analizė

II.2.53. Trifaktorės dispersinės analizės skirtingų langelių stebėjimų vidurkiai $\mu_{ijl}, i = 1,2,3;\ j = 1,2,3; l = 1,2$, pateikti lentelėse

C_1	B_1	B_2	B_3	C_2	B_1	B_2	B_3
A_1	5	6	10	A_1	9	7	14
A_2	7	7	1	A_2	9	6	3
A_3	6	5	7	A_3	9	5	10

Įrodykite, kad visų trijų faktorių sąveika $A \times B \times C$ yra lygi 0.

II.2.54. Nagrinėjame tiesinį modelį $Y_{ijk} = \mu_{ijk} + e_{ijk}, i = 1, ..., I; j = 1, ..., J; k = 1, ..., K$, kuriame $\{e_{ijk}\}$ yra nepriklausomi normalieji $e_{ijk} \sim N(0, \sigma^2)$ a. d. Tarkime, vidurkiai μ_{ijk} tenkina sąlygą

$$\mu_{ijk} = \bar{\mu}_{...} + (\bar{\mu}_{i..} - \bar{\mu}_{...}) + (\bar{\mu}_{ij.} - \bar{\mu}_{i..}) + (\bar{\mu}_{..k} - \bar{\mu}_{...}) =$$

$$= \mu + \alpha_i + \beta_{ij} + \gamma_k.$$

Raskite parametrų $\mu, \alpha_i, \beta_{ij}, \gamma_k$ įvertinius. Sukurkite kriterijų hipotezei $H_A : \alpha_i = 0, i = 1, ..., I$.

II.2.55. Sudarykite trifaktorės dispersinės analizės lentelę, kai visi trys faktoriai yra atsitiktiniai. Sukurkite kriterijus pagrindinėms hipotezėms tikrinti.

II.2.56. Tarkime, pilnoje dispersinės analizės schemoje, kai vienodas stebėjimų skaičius langelyje, faktorius A turi I lygmenų. Stebėjimų aritmetinius vidurkius, gautus suvidurkinus pagal visų kitų faktorių lygmenis, pažymėkime $\bar{Y}_1, ..., \bar{Y}_I$. Tegu šie vidurkiai padalyti į dvi didumo I_1 ir I_2, $I_1 + I_2 = I$ aibes $\bar{Y}_1, ..., \bar{Y}_{I_1}$ ir $\bar{Y}_{I_1+1}, ..., \bar{Y}_I$, o $\bar{Y}^{(1)}$ ir $\bar{Y}^{(2)}$ yra šių aibių vidurkiai. Įrodykite, kad

$$\sum_{i=1}^{I} (\bar{Y}_i - \bar{Y}_.)^2 = \sum_{i=1}^{I_1} (\bar{Y}_i - \bar{Y}^{(1)})^2 + \sum_{i=I_1+1}^{I} (\bar{Y}_i - \bar{Y}^{(2)})^2 + I_1 I_2 (\bar{Y}^{(1)} - \bar{Y}^{(2)})^2 / I.$$

II.2.57. Apibendrinkite **II.2.34** pratimą, kai vidurkiai $\bar{Y}_1, ..., \bar{Y}_I$ padalijami į tris nesikertančias aibes.

II.2.58. Lentelėje pateikti duomenys, apibūdinantys betoninį kelią priklausomai nuo padengimo storio (faktorius A), nuo pagrindo storio (faktorius B) ir nuo papildomo (apatinio) pagrindo storio (faktorius C). Atlikta po du matavimus kiekvieno iš 27 kelio variantų ([11]).

	A_1			A_2			A_3		
	B_1	B_2	B_3	B_1	B_2	B_3	B_1	B_2	B_3
C_1	2,8	4,3	5,7	4,1	5,4	6,7	6,0	6,3	7,1
	2,6	4,5	5,3	4,4	5,5	6,9	6,2	6,5	6,9
C_2	4,1	5,7	6,9	5,3	6,5	7,7	6,1	7,2	8,1
	4,4	5,8	7,1	5,1	6,7	7,4	5,8	7,1	8,4
C_3	5,5	7,0	8,1	6,5	7,7	8,8	7,0	8,0	9,1
	5,3	6,8	8,3	6,7	7,5	9,1	7,2	8,3	9,0

Atlikite trifaktorę analizę su trimis pastoviais faktoriais.

II.2.59. Lentelėje pateikiama izoliacijos kokybės charakteristikos priklausomai nuo keturių faktorių: A – padengimo tipas; B – temperatūra; C – slėgis; D – skirtingos plieninės panelės, kurios buvo naudojamos eksperimente [16].

	A_1				A_2			
	B_1	B_1	B_2	B_2	B_1	B_1	B_2	B_2
	C_1	C_2	C_1	C_2	C_1	C_2	C_1	C_2
D_1	0,25	0,16	0,30	0,27	0,41	0,10	0,13	0,06
D_2	0,36	0,002	0,18	0,03	0,28	0,04	0,06	0,03
D_3	0,36	0,06	0,44	0,13	0,33	0,03	0,19	0,04
D_4	0,25	0,10	0,34	0,04	0,21	0,01	0,20	0,01

	A_3				A_4			
	B_1	B_1	B_2	B_2	B_1	B_1	B_2	B_2
	C_1	C_2	C_1	C_2	C_1	C_2	C_1	C_2
D_1	0,44	0,24	0,22	0,18	0,43	0,27	0,26	0,21
D_2	0,65	0,08	0,14	0,36	0,62	0,03	0,51	0,03
D_3	0,42	0,49	0,17	0,25	0,47	0,28	0,21	0,25
D_4	0,47	0,14	0,36	0,19	0,52	0,07	0,32	0,38

Atlikite dispersinę analizę, tarę, kad visų keturių faktorių sąveika ir sąveikos po tris faktorius yra lygios 0.

II.2.4. Nepilni dispersinės analizės planai

II.2.60. Lentelėje pateikta keturių atspaudų charakteristikos (sąlyginiais vienetais) priklausomai nuo keturių galvučių skirtinguose spausdinimo įrenginiuose [16].

A_1				A_2				A_3			
1	2	3	4	5	6	7	8	9	10	11	12
6	13	1	7	10	2	4	0	0	10	8	7
2	3	10	4	9	1	1	3	0	11	5	2
0	9	0	7	7	1	7	4	5	6	0	5
8	8	6	9	12	10	9	1	5	7	7	4

A_4				A_5			
13	14	15	16	17	18	19	20
11	5	1	0	1	6	3	3
0	10	8	8	4	7	0	7
6	8	9	6	7	0	2	4
4	3	4	5	9	3	2	0

Atlikite dispersinę analizę, tarę, kad faktorius B (galvutės) yra sugrupuotas pagal faktorių A (spausdinimo įrenginiai); galvučių numeriai yra antroje lentelės eilutėje. Faktorių B laikyti atsitiktiniu, o skirtingas kopijas interpretuoti kaip matavimo kartotinumą.

II.2.61. Lentelėje pateikti duomenys apie kondensatorinio popieriaus poringumą priklausomai nuo partijos (faktorius A) ir nuo atsitiktinai iš partijos atrinkto rulono (faktorius B). Kiekvieną kartą atlikta po tris matavimus [11].

A_1				A_2				A_3			
1	2	3	4	5	6	7	8	9	10	11	12
1,5	1,5	2,7	3,0	1,9	2,3	1,8	1,9	2,5	3,2	1,4	7,8
1,7	1,6	1,9	2,4	1,5	2,4	2,9	3,5	2,9	5,5	1,5	5,2
1,6	1,7	2,0	2,6	2,1	2,4	4,7	2,8	3,3	7,1	3,4	5,0

Atlikite dispersinę analizę, tarę, kad faktorius B (rulonai) yra sugrupuotas pagal faktorių A (partijos); rulonų numeriai yra antroje lentelės eilutėje. Abu faktorius laikyti atsitiktiniais.

II.2.62. Atliekant eksperimentą 9 jūrų kiaulytės (faktorius B) atsitiktinai buvo suskirstytos į grupes po 3 ir įdėtos į skirtingus narvelius. Kiekvieno narvelio gyvūnai buvo aprūpinami skirtingai NO_2 lygiais (faktorius A); A_1 – kontrolinis; A_2 – dvigubai didesnis už normą; A_3 – trigubai didesnis už normą. Po savaitės buvo atlikta po du kintamojo Y (arterinis PH) matavimus [1].

	B_1		B_2		B_3	
A_1	7,08	7,02	7,04	7,07	7,07	6,98
A_2	7,29	7,18	7,42	7,32	7,08	7,28
A_3	7,74	7,54	7,53	7,50	7,51	7,63

Ištirkite kintamojo Y priklausomybę nuo faktoriaus A, tardami, kad faktorius B sugrupuotas pagal faktorių A.

II.2.63. Keturios pelės turi po 3 peliukus. Pelės atsitiktinai suskirstytos į 3 grupes po dvi. Pirmos grupės peliukams taikoma pirmoji dieta, o antrosios – antroji dieta. Po trijų savaičių užregistruotas peliukų svorio prieaugis Y [1].

Dieta	Pelė	Y			Dieta	Pelė	Y		
Pirmoji	1	11,8	10,5	12,5	Antroji	3	7,4	9,7	8,2
	2	12,3	15,5	11,4		4	7,2	8,6	7,1

Parinkite tinkamą dispersinės analizės schemą ir užpildykite dispersinės analizės lentelę. Patikrinkite pagrindines dispersinės analizės hipotezes.

II.2.64. Penkiolikai pacientų matuotas fermento kiekis iš karto po širdies operacijos (D_0), praėjus vienai dienai (D_1), dviem dienoms (D_2) ir savaitei (D_7) po operacijos.

Pacientas	D_0	D_1	D_2	D_7	Pacientas	D_0	D_1	D_2	D_7
1	108	63	45	42	9	106	65	49	49
2	112	75	56	52	10	110	70	46	47
3	114	75	51	46	11	120	85	60	62
4	129	87	69	69	12	118	78	51	56
5	115	71	52	54	13	110	65	46	47
6	122	80	68	68	14	132	92	73	63
7	105	71	52	54	15	127	90	73	68
8	117	77	54	61					

Patikrinkite, ar fermento kiekio vidurkis kinta po širdies operacijos.

II.2.65. Lentelėje pateikti duomenys apie elektroninių spindulinių vamzdelių stiklo savybes priklausomai nuo cecho (faktorius A) ir nuo pamainos (faktorius B). Matavimai atlikti trijų atsitiktinai parinktų savaičių metu (faktorius C; blokas) [11]. Atlikite blokuotųjų duomenų dispersinę analizę.

	A_1			A_2		
	B_1	B_2	B_3	B_1	B_2	B_3
	3	3	3	6	3	6
C_1	6	4	6	8	9	8
	6	7	7	11	11	13
	14	8	11	4	15	4
C_2	16	8	12	6	15	7
	19	9	17	7	17	10
	2	2	2	2	2	10
C_3	3	3	4	5	4	12
	6	4	6	7	6	13

II.2.66. Tiriant kineskopo elektros srovės stiprumo priklausomybę nuo keturių katodo kaitinimo siūlelio apdorojimo režimų (A, B, C, D), per vieną dieną galima realizuoti tik tris apdorojimo metodus. Todėl eksperimentas atliktas pagal nepilnų subalansuotų blokų schemą, kurioje blokus atitinka dienos [11].

Dienos	A	B	C	D
1	2	–	20	7
2	–	32	14	3
3	4	13	31	–
4	0	23	–	11

Patikrinkite hipotezę, kad srovės stiprumas nepriklauso nuo apdorojimo metodo.

II.2.67. Reikia palyginti televizorių ryškumo įvertinimą, gautą keturių operatorių (A, B, C, D). Per vieną dieną eksperimente gali dalyvauti tik trys operatoriai. Duomenys pateikti lentelėje [11].

Dienos	A	B	C	D
1	780	820	800	–
2	950	–	920	940
3	–	880	880	820
4	840	780	–	820

Atlikite duomenų analizę. Ar galima tvirtinti, kad operatorių įvertinimai skiriasi?

II.2.68. Penkioms skalbimo priemonėms (A, B, C, D, E) palyginti buvo atliktas toks eksperimentas. Skalbimo priemonės buvo lyginamos plaunant specialiai užterštas lėkštes blokuose po tris rezervuarus su skirtingomis skalbimo priemonėmis. Lentelėje pateiktas išplautų vienodu skalbimo priemonės kiekiu lėkščių skaičius [16].

	1	2	3	4	5	6	7	8	9	10
A	27	28	30	31	29	30	–	–	–	–
B	26	26	29	–	–	–	30	21	26	–
C	30	–	–	34	32	–	34	31	–	33
D	–	29	–	33	–	34	31	–	33	31
E	–	–	26	–	24	25	–	23	24	26

Atlikite duomenų analizę. Ar galima tvirtinti, kad skalbimo priemonės skiriasi?

II.2.69. Tiriant penkių tipų (A, B, C, D, E) elektrodus eksperimento metu buvo pradeginta po 5 skylutes 5 metalo juostose. Eksperimentas atliktas pagal lotyniškojo kvadrato schemą, kurioje eilutės atitinka metalo juostas, stulpeliai – skylutės padėtį metalo juostoje, o elektrodo tipas nurodytas skliausteliuose. Registruojama skylutės pradeginimo laikas [11].

3,5 (A)	2,1 (B)	2,5 (C)	3,5 (D)	2,4 (E)
2,6 (E)	3,3 (A)	2,1 (B)	2,5 (C)	2,7 (D)
2,9 (D)	2,6 (E)	3,5 (A)	2,7 (B)	2,9 (C)
2,5 (C)	2,9 (D)	3,0 (E)	3,3 (A)	2,3 (B)
2,1 (B)	2,3 (C)	3,7 (D)	3,2 (E)	3,5 (A)

Atlikite duomenų analizę ir užpildykite dispersinės analizės lentelę. Ar galima tvirtinti, kad elektrodų tipai skiriasi?

II.2.70. Tiriant tam tikros ankštinės kultūros šešių veislių (A, B, C, D, E, F) derlingumą, eksperimentas atliktas pagal lotyniškojo kvadrato schemą. Lentelėje nurodyti gauti derlingumo rodikliai, o kultūros veislė nurodyta skliausteliuose [16].

220 (B)	98 (F)	149 (D)	92 (A)	282 (E)	169 (C)
74 (A)	238 (E)	153 (B)	228 (C)	48 (F)	188 (D)
118 (D)	279 (C)	118 (F)	272 (E)	176 (B)	65 (A)
295 (E)	222 (B)	54 (A)	104 (D)	213 (C)	163 (F)
187 (C)	90 (D)	242 (E)	96 (F)	66 (A)	122 (B)
90 (F)	124 (A)	195 (C)	109 (B)	79 (D)	211 (E)

Atlikite duomenų dispersinę analizę.

II.2.5. Atsakymai, sprendimai, nurodymai

II.2.1 skyrelis

II.2.1. a) Sujungę Y_{ij} į vieną bendrą vektorių

$$\boldsymbol{Y} = (Y_{11}, ..., Y_{1J_1}, Y_{21}, ..., Y_{2J_2}, ..., Y_{I1}, ..., Y_{IJ_I})^T$$

ir pažymėję $\boldsymbol{\beta} = (\mu_1, ..., \mu_I)^T$, imtį $\boldsymbol{Y}$ galime užrašyti kaip tiesinį modelį

$$\boldsymbol{Y} = \boldsymbol{A\beta} + \boldsymbol{e}, \quad \mathbf{E}(\boldsymbol{Y}) = \boldsymbol{A\beta}, \quad \boldsymbol{V}(\boldsymbol{Y}) = \boldsymbol{V}(\boldsymbol{e}) = \sigma^2\boldsymbol{I}, \quad \boldsymbol{Y} \sim N_n(\boldsymbol{A\beta}, \sigma^2\boldsymbol{I}).$$

Plano matrica $\boldsymbol{A}$ turi $n = J_1 + ... + J_I$ eilučių ir I stulpelių. Pirmosios J_1 eilutės turi pavidalą $(1,\ 0, ..., 0)$, paskui J_2 eilučių turi pavidalą $(0,\ 1, ..., 0)$, pagaliau paskutinės J_I eilučių turi pavidalą $(0,\ 0, ..., 1)$.

b) Tikėtinumo funkcija

$$L(\boldsymbol{\theta}) = (2\pi)^{-n/2}\exp\{\frac{-1}{2\sigma^2}\sum_{i=1}^{I}\sum_{j=1}^{J_i}(Y_{ij} - \mu_i)^2 - \frac{n}{2}\ln\sigma^2\} =$$

$$(2\pi)^{-n/2}\exp\{\frac{-1}{2\sigma^2}\boldsymbol{Y}^T\boldsymbol{Y} + \frac{1}{\sigma^2}\sum_{i=1}^{I}J_i\mu_i\bar{Y}_{i\cdot} - \sum_{i=1}^{I}\frac{J_i\mu_i^2}{2\sigma^2} - \frac{n}{2}\ln\sigma^2\}$$

priklauso $(I+1)$-parametrei eksponentinio tipo skirstinių šeimai. Remiantis faktorizacijos kriterijumi $\boldsymbol{T}$ yra parametro $\boldsymbol{\theta}$ pakankamoji statistika. Kadangi parametro kitimo sričiai priklauso vidiniai taškai, tai statistika $\boldsymbol{T}$ yra pilnoji.

c) Kadangi $\boldsymbol{T}$ yra pilnoji ir pakankamoji statistika, tai bet kuri $\boldsymbol{T}$ funkcija yra jos vidurkio NMD įvertinys. Randame $\mathbf{E}\bar{Y}_{i\cdot} = \mu_i$, todėl $\hat{\mu}_i = \bar{Y}_{i\cdot}$ yra parametro μ_i NMD įvertinys. Liekamoji kvadratų suma

$$SS_E/\sigma^2 = \sum_i\sum_j(Y_{ij} - Y_{i\cdot})^2/\sigma^2 \sim \chi^2(n-I).$$

Gauname, kad

$$s^2 = SS_E/(n-I), \quad \mathbf{E}s^2 = \sigma^2,$$

yra parametro σ^2 NMD įvertinys.

II.2.2. Parametrai μ, α_i yra tiesinės parametrų μ_i funkcijos. Nepaslinktieji (kartu ir NMD) įvertiniai gaunami imant tokias pačias tiesines įvertinių $\hat{\mu}_i$ funkcijas

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} J_i\bar{Y}_{i.} = \bar{Y}_{..}, \quad \hat{\alpha}_i = \bar{Y}_{i.} - \bar{Y}_{..}.$$

Šie įvertiniai tenkina sąryšį

$$\sum_{i=1}^{I} J_i\hat{\alpha}_i = \sum_{i=1}^{I} J_i(\bar{Y}_{i.} - \bar{Y}_{..}) = 0.$$

II.2.3. Hipotezę H_A galima suformuluoti tokiu būdu $H_A : \boldsymbol{H\beta} = \boldsymbol{\theta}_0$ (žr. **I.1.21** pratimą), kai matricos $\boldsymbol{H}$ rangas yra $k = I-1$. Pavyzdžiui, matricą $\boldsymbol{H}$ galima imti tokio pavidalo: pirmoji eilutė $(1,\ 0, ..., 0,\ -1)$, antroji $(0,\ 1, ..., 0,\ -1)$ ir t. t., ir $(I-1)$-oji eilutė $(0,\ 0, ..., 1,\ -1)$. Matricos $\boldsymbol{H}$ rangas yra $I-1$, o $\boldsymbol{\theta}_0 = \mathbf{0}$. Remiantis **I.1.21** pratimo p. d) hipotezė H_A atmetama reikšmingumo lygmens α kriterijumi, kai

$$F_A = \frac{(SS_{EH_A} - SS_E)(n-I)}{(I-1)SS_E} = \frac{SS_A(n-I)}{(I-1)SS_E} > F_\alpha(I-1, n-I),$$

čia SS_E – liekamoji kvadratų suma (surasta **II.2.1** pratime), o SS_{EH_A} yra kvadratinės formos $SS(\boldsymbol{\beta})$ sąlyginis minimumas, surastas esant sąlygai, kad H_A teisinga.

Sąlyginę kvadratinę formą lengva surasti tokiu būdu. Tapatybės

$$Y_{ij} - \mu_i = Y_{ij} - \hat{\mu}_i + \hat{\alpha}_i + \hat{\mu} - \mu$$

abi puses pakelkime kvadratu ir susumuokime pagal i ir j. Gauname

$$SS(\boldsymbol{\beta}) = SS_E + \sum_i J_i(\hat{\alpha}_i - \alpha_i)^2 + n(\hat{\mu} - \mu)^2,$$

nes visos mišrios sandaugos lygios nuliui. Tada

$$SS_{EH_A} = \min_{\alpha_i = 0} SS(\boldsymbol{\beta}) = SS_E + \sum_i J_i\hat{\alpha}_i^2,$$

$$SS_A = SS_{EH_A} - SS_E = \sum_i J_i(\bar{Y}_{i.} - \bar{Y}_{..})^2.$$

Pažymėjus vidutines kvadratų sumas

$$MS_A = \frac{SS_A}{I-1}, \quad MS_E = \frac{SS_E}{n-I},$$

kriterijus įgauna tokį pavidalą: hipotezė H_A atmetama reikšmingumo lygmens α kriterijumi, kai

$$F_A = \frac{MS_A}{MS_E} > F_\alpha(I-1, n-I).$$

Skaičiavimo rezultatai analizę atliekant kompiuteriu pateikiami tokioje lentelėje.

Dispersinės analizės lentelė

Faktorius	SS	ν	$MS = SS/\nu$	$\mathbf{E}$(MS)
A	$SS_A = \sum_i J_i(\bar{Y}_{i.} - \bar{Y}_{..})^2$	$I-1$	MS_A	$\sigma^2 + \sigma_A^2$
E	$SS_E = \sum_i \sum_j (Y_{ij} - \bar{Y}_{i.})^2$	$n-I$	MS_E	σ^2
T	$SS_T = \sum_i \sum_j (Y_{ij} - \bar{Y}_{..})^2$			

Šioje lentelėje antrame stulpelyje pateikiamos kvadratų sumos, trečiame stulpelyje – laisvės laipsniai, ketvirtame – vidutinės kvadratų sumos, $\sigma_A^2 = \sum_i J_i\alpha_i^2/(I-1)$.

II.2.4. a)

$$\mathbf{E}(SS_E) = \mathbf{E}(\sum_{i=1}^{I}\sum_{j=1}^{J_i}(e_{ij} - \bar{e}_{i.})^2) = \sigma^2\sum_{i=1}^{I}(J_i - 1) = \sigma^2(n-I);$$

$$\mathbf{E}(MS_E) = \sigma^2.$$

$$\mathbf{E}(SS_A) = \mathbf{E}(\sum_{i=1}^{I} J_i(\alpha_i + \bar{e}_{i.} - \bar{e}_{..})^2) = \sum_{i=1}^{I} J_i\alpha_i^2 + \mathbf{E}\sum_{i=1}^{I} J_i(\bar{e}_{i.} - \bar{e}_{..})^2 =$$

$$\sum_{i=1}^{I} J_i\alpha_i^2 + (I-1)\sigma^2; \quad \mathbf{E}(MS_A) = \sigma^2 + \sum_{i=1}^{I}\frac{J_i\alpha_i^2}{I-1} = \sigma^2 + \sigma_A^2.$$

b) remiantis **I.1.21** pratimu $SS_A/\sigma^2 \sim \chi^2(I-1, \lambda)$, kai necentriškumo parametras $\lambda = (I-1)\sigma_A^2/\sigma^2$. Tada statistika F_A turi necentrinį Fišerio skirstinį $F(I-1, n-I; \lambda)$. Kriterijaus galia

$$\beta(\lambda) = \mathbf{P}\{F_{I-1,n-I;\lambda} > F_\alpha(I-1, n-I)\}.$$

II.2.5. Kai hipotezė H'_A teisinga, tai yra tik du parametrai: vidurkis μ'_1 pirmajai faktoriaus lygmenų grupei ir μ'_2 – antrajai. Šių parametrų DT įvertiniai

$$\hat{\mu}'_1 = \sum_{i=1}^{r}\sum_{j=1}^{J_i} Y_{ij}/n'_1 = \bar{Y}_{..}^{(1)}, \quad \hat{\mu}'_2 = \sum_{i=r+1}^{I}\sum_{j=1}^{J_i} Y_{ij}/n'_2 = \bar{Y}_{..}^{(2)},$$

$$n'_1 = J_1 + ... + J_r, \quad n'_2 = J_{r+1} + ... + J_I$$

ir

$$SS_{EH'_A} - SS_E = SS'_A = \sum_{i=1}^{r} J_i(\bar{Y}_{i.} - \bar{Y}_{..}^{(1)})^2 + \sum_{i=r+1}^{I} J_i(\bar{Y}_{i.} - \bar{Y}_{..}^{(2)})^2 \stackrel{d}{=} \sigma^2\chi^2_{I-2},$$

kai H'_A teisinga. Hipotezė H'_A atmetama, kai

$$F'_A == \frac{SS'_A(n-I)}{(I-2)SS_E} = \frac{MS'_A}{MS_E} > F_\alpha(I-2, n-I).$$

II.2.6. Žr. [3], 2.1.2 teoremą.

II.2.7. Tarkime, kad $d_I \neq 0$. Tada $\alpha_I = -\sum_{i=1}^{I-1}\alpha_i d_i/d_I$.

Pažymėję $\boldsymbol{Y} = (Y_{11}, ..., Y_{1J}, Y_{21}, ..., Y_{IJ})^T$ ir $\boldsymbol{\beta} = (\mu, \alpha_1, ..., \alpha_{I-1})^T$, gauname tiesinį modelį $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{\beta} + \boldsymbol{e}$; $\hat{\boldsymbol{\beta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y}$.

II.2.8. a) Vertindami parametrą $\boldsymbol{\beta} = (\mu_1, \mu_2, \mu_3, \mu_4)^T$ turime tiesinį modelį, kai matrica $\boldsymbol{A}^T\boldsymbol{A}$ diagonali su diagonaliniais elementais J. Taigi parametrų μ_i įvertiniai $\hat{\mu}_i = \bar{Y}_{i.}$ yra nepriklausomi ir $\hat{\mu}_i \sim N(\mu_i, \sigma^2/J)$. Liekamoji kvadratų suma

$$SS_E/\sigma^2 = \sum_{i=1}^{4}\sum_{j=1}^{J}(Y_{ij} - \bar{Y}_{i.})^2/\sigma^2 \sim \chi^2(4(J-1)).$$

Tikrinamąją hipotezę galima užrašyti taip $H : \boldsymbol{\theta} = \mathbf{0}$, kai $\boldsymbol{\theta} = (\theta_1, \theta_2)^T$, $\theta_1 = \mu_1 - 2\mu_2, \theta_2 = \mu_1 - 3\mu_3$. Parametro $\boldsymbol{\theta}$ NMD įvertinys

$$\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{pmatrix} = \begin{pmatrix} \hat{\mu}_1 - 2\hat{\mu}_2 \\ \hat{\mu}_1 - 3\hat{\mu}_3 \end{pmatrix} \sim N_2(\boldsymbol{\theta}, \boldsymbol{\Sigma}),$$

čia $\boldsymbol{\Sigma} = [\sigma_{ij}]_{2\times 2}$, $\sigma_{11} = 5\sigma^2/J$, $\sigma_{12} = \sigma^2/J$, $\sigma_{22} = 10\sigma^2/J$. Jeigu hipotezė H teisinga, tai kvadratinė forma

$$\hat{\boldsymbol{\theta}}^T\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\theta}} = J(10\hat{\theta}_1^2 - 2\hat{\theta}_1\hat{\theta}_2 + 5\hat{\theta}_2^2)/(49\sigma^2) \sim \chi^2(2).$$

Hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai

$$F = \frac{2\hat{\boldsymbol{\theta}}^T\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\theta}}(J-1)}{SS_E/\sigma^2} > F_\alpha(2, 4(J-1)).$$

Uždavinį galima išspręsti kitu būdu remiantis **II.1.21** pratimo p. d). Randame kvadratinės formos $SS(\boldsymbol{\beta})$ sąlyginį minimumą SS_{EH}, kai $\mu_1 = 2\mu_2, \mu_1 = 3\mu_3$. Pažymėję $\mu_2 = \mu_1/2$, $\mu_3 = \mu_1/3$, gausime tiesinį modelį, priklausantį tik nuo dviejų parametrų μ_1 ir μ_4. Plano matrica turi 2 stulpelius ir $4J$ eilučių: pirmosios J eilučių yra (0, 1), antrosios J eilučių yra (1/2, 0), po to J eilučių (1/3, 0) ir paskutinės J eilučių yra (0, 1). Gauname parametrų μ_1 ir μ_4 įvertinius

$$\tilde{\mu}_1 = (\hat{\mu}_1 + \hat{\mu}_2/2 + \hat{\mu}_3/3)/(1 + 1/4 + 1/9) = \frac{36}{49}(\mu)_1 + \frac{\hat{\mu}_2}{2} + \frac{\hat{\mu}_3}{3}, \quad \tilde{\mu}_4 = \hat{\mu}_4$$

ir liekamąją kvadratų sumą

$$SS_{EH} = \sum_i (Y_{1j} - \tilde{\mu}_1)^2 + \sum_i (Y_{2j} - \tilde{\mu}_1/2)^2 + \sum_i (Y_{3j} - \tilde{\mu}_1/3)^2 + \sum_i (Y_{4j} - \hat{\mu}_4)^2.$$

Skirtumas

$$(SS_{EH} - SS_E)/\sigma^2 = [(\hat{\mu}_1 - \tilde{\mu}_1)^2 + (\hat{\mu}_2 - \tilde{\mu}_1/2)^2 + (\hat{\mu}_3 - \tilde{\mu}_1/3)^2]/\sigma^2 \sim \chi^2(2),$$

kai hipotezė teisinga. Hipotezė atmetama, kai

$$\frac{2(SS_{EH} - SS_E)(J-1)}{SS_E} > F_\alpha(2, 4(J-1)).$$

Nesunku patikrinti, kad abu kriterijai ekvivalentūs.

b) Stjudento kriterijus atmeta hipotezę, kai

$$|t| = \sqrt{\frac{J}{2}}\frac{|\bar{Y}_{1.} - \bar{Y}_{2.}|}{s} > t_{\alpha/2}(2(J-1)), \quad s^2 = \frac{SS_E}{2(J-1)}.$$

Kadangi $SS_A = J(\bar{Y}_{1.} - \bar{Y}_{2.})^2/2$, tai F kriterijus atmeta hipotezę, kai

$$F = \frac{J}{2}\frac{(\bar{Y}_{1.} - \bar{Y}_{2.})^2}{s^2} > F_\alpha(1, 2(J-1)).$$

Akivaizdu, kad šie kriterijai ekvivalentūs.

II.2.9. Dispersinės analizės lentelė:

Faktorius	SS	ν	MS	F	$Pr > F$
A	$SS_A = 4,23214$	3	1,41071	3,21143	0,0359
E	$SS_E = 14,05691$	32	0,39047		
T	$SS_T = 18,28905$				

Statistika F_A (**II.2.3** pratimas), kuri esant teisingai hipotezei turi Fišerio skirstinį su 3 ir 32 laisvės laipsniais, įgijo reikšmę 3,21143; P reikšmė $pv = \mathbf{P}\{F_{3,32} > 3,21143\} = 0,0369$. Hipotezė atmetama kriterijumi, kurio reikšmingumo lygmuo viršija 0,0359.

II.2.10. Atlikę analizę analogišką **II.2.9** pratimui, gauname, kad statistika F_A įgijo reikšmę 9,9285; hipotezė atmetama kriterijumi, kurio reikšmingumo lygmuo $\alpha > pv = \mathbf{P}\{F_{3;20} > 9,9285\} = 0,00032$.

II.2.11. Statistika F_A įgijo reikšmę 1,3574; $\mathbf{P}\{F_{3;16} > 1,3574\} = 0,2950$; atmesti hipotezę nėra pagrindo.

II.2.12. Statistika F_A, kuri esant teisingai hipotezei turi Fišerio skirstinį su 2 ir 321 laisvės laipsniu, įgijo atitinkamai reikšmes: 1,538; 72,507; 31,009; 18,312; 23,904; 27,724. Duomenys neprieštarauja prielaidai, kad vario koncentracija yra vienoda; kitų elementų koncentracijų vienodumo hipotezės atmetamos kriterijais su aukštais reikšmingumo lygmenimis.

II.2.13. Tarkime, i-ojo gaminio požymis nusakomas a. d. X_i, o jį matuojant j-ąjį kartą prisideda paklaida e_{ij}. Tarkime, kad paklaidos e_{ij} nepriklausomos ir normaliosios $e_{ij} \sim N(0, \sigma^2)$. Kadangi reikia apibūdinti visą gaminių aibę, o ne tik palyginti tarpusavyje kelis eksperimente dalyvaujančius gaminius, tai tarsime, kad $X_1, ..., X_I$ yra paprastoji imtis, gauta stebint a. d. X. Tarsime, kad X_i nepriklauso nuo e_{ij} ir $X_i \sim N(\mu, \sigma_A^2)$. Pažymėję $a_i = X_i - \mu$ gausime, kad matavimai Y_{ij} turi tokią struktūrą

$$Y_{ij} = X_i + e_{ij} = \mu + a_i + e_{ij}, \quad i = 1, ..., I, \quad j = 1, ..., J.$$

Tai vienfaktorės dispersinės analizės modelis su atsitiktiniu faktoriumi. Modelį visiškai nusako parametrai $\mu, \sigma_A^2, \sigma^2$.

II.2.14. Nagrinėkime nepriklausomus a. v. $\boldsymbol{Y}_i = (Y_{i1}, Y_{i2}, ..., Y_{iJ}) \sim N_J(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $i = 1, ..., I$; čia $\boldsymbol{\mu} = (\mu, \mu, ..., \mu)^T$, $\boldsymbol{\Sigma} = [\sigma_{kl}]_{J \times J}$, $\sigma_{kk} = \sigma_A^2 + \sigma^2, k = 1, ..., J$, $\sigma_{kl} = \sigma_A^2$, kai $k \neq l = 1, ..., J$.

Atlikime vektoriaus $\boldsymbol{Y}_i$ ortonormuotą transformaciją $\boldsymbol{Z}_i = \boldsymbol{C}\boldsymbol{Y}_i = (Z_{i1}, ..., Z_{iJ})^T$ pasinaudodami matrica $\boldsymbol{C} = [c_{kl}]_{J \times J}$. Tegu matricos $\boldsymbol{C}$ pirmoji eilutė $(c_{11}, ..., c_{1J}) = (1/\sqrt{J}, ..., 1/\sqrt{J})$. Tada iš ortogonalumo sąlygos gauname

$$\sum_{j=1}^{J} c_{ij} = 0, \quad i = 2, ..., J; \ \sum_{j=1}^{J} c_{ij}^2 = 1, \quad \sum_{j=1}^{J} c_{ij}c_{i'j} = 0 \quad i \neq i'.$$

Gauname

$$Z_{i1}=\sqrt{J}\bar{Y}_{i.},\quad \mathbf{E}Z_{i1}=\sqrt{J}\mu,\quad \mathbf{E}Z_{il}=0,\quad l=2,...,J;$$

$$\boldsymbol{V}(\boldsymbol{Z}_i)=\boldsymbol{V}(\boldsymbol{C}\boldsymbol{Y}_i)=\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}^T=\sigma_A^2\boldsymbol{C}\boldsymbol{U}\boldsymbol{C}^T+\sigma^2\boldsymbol{C}\boldsymbol{I}\boldsymbol{C}^T,$$

čia $\boldsymbol{U}$ – matrica, kurios visi elementai yra vienetai.

Pirmojo dėmens matricos tik pirmos eilutės pirmasis elementas yra $J\sigma_A^2$, o visi kiti elementai nuliai; antrojo dėmens matrica diagonali su diagonaliniais elementais σ^2. Galutinai turime

$$\boldsymbol{V}(\boldsymbol{Z}_i)=\Lambda=[\lambda_{rs}]_{J\times J},\quad \lambda_{11}=\sigma^2+J\sigma_A^2,\quad \lambda_{rr}=\sigma^2,\quad r=2,...,J,\lambda_{rs}=0,\ r\neq s.$$

Atlikę tą pačią transformaciją su visais vektoriais $\boldsymbol{Y}_1,...,\boldsymbol{Y}_I$, gauname nepriklausomų a. v. sistemą $\boldsymbol{Z}_1,...,\boldsymbol{Z}_I$, kurių ir koordinatės nepriklausomos; pirmosios koordinatės $Z_{i1}\sim N(\sqrt{J}\mu,\sigma^2+J\sigma_A^2)$, visos kitos koordinatės $Z_{il}\sim N(0,\sigma^2)$, $l=2,...,J$.

Tikėtinumo funkcija Z_{ij} terminais yra

$$L(\mu,\sigma_A^2,\sigma^2)=\exp\{-\frac{1}{2(\sigma^2+J\sigma_A^2)}\sum_{i=1}^I(Z_{i1}-\sqrt{J}\mu)^2-\frac{1}{2\sigma^2}\sum_{i=1}^I\sum_{j=2}^J Z_{ij}^2-b(\mu,\sigma_A^2,\sigma^2)\}=$$

$$\exp\{(\frac{1}{2\sigma^2}-\frac{1}{2(\sigma^2+J\sigma_A^2)})\sum_{i=1}^I Z_{i1}^2+\frac{\mu\sqrt{J}}{\sigma^2+\sigma_A^2}\sum_{i=1}^I Z_{i1}-\frac{1}{2\sigma^2}\sum_{i=1}^I\sum_{j=1}^J Z_{ij}^2-B(\mu,\sigma_A^2,\sigma^2)\},$$

priklauso triparametrei eksponentinio tipo skirstinių šeimai. Statistika

$$(\sum_i Z_{i1}^2,\sum_i Z_{i1},\sum_i\sum_j Z_{ij}^2)^T$$

yra pilnoji ir pakankamoji. Grįžę prie pradinių kintamųjų Y_{ij} įsitikiname, kad statistika $\boldsymbol{T}$ irgi yra pilnoji ir pakankamoji.

II.2.15. Statistika $\boldsymbol{T}$ yra pilnoji ir pakankamoji, todėl bet kuri $\boldsymbol{T}$ funkcija yra jos vidurkio NMD įvertinys. Imkime tokias $\boldsymbol{T}$ funkcijas

$$\hat{\mu}=\bar{Y}_{..}=\mu+\bar{a}_.+\bar{e}_{..}\sim N(\mu,\tau^2/(IJ));$$

$$SS_E=\sum_i\sum_j(Y_{ij}-\bar{Y}_{i.})^2=\sum_i\sum_j(e_{ij}-\bar{e}_{i.})^2\sim\sigma^2\chi^2_{I(J-1)},$$

$$\hat{\sigma^2}=s^2=\frac{SS_E}{I(J-1)}=MS_E,\quad \mathbf{E}s^2=\mathbf{E}(MS_E)=\sigma^2;$$

$$SS_A=J\sum_i(\bar{Y}_{i.}-\bar{Y}_{..})^2=J\sum_i(a_i+\bar{e}_{i.}-\bar{a}_.-\bar{e}_{..})^2=J\sum_i(U_i-\bar{U}_.)^2,$$

čia $U_1,...,U_I$ yra vienodai pasiskirstę n. a. d. ir $U_i\sim N(0,\sigma_A^2+\sigma^2/J)$. Taigi

$$SS_A\sim(\sigma^2+J\sigma_A^2)\chi^2_{I-1},\quad \hat{\tau^2}=SS_A/(I-1)=MS_A,$$

$$\mathbf{E}(\hat{\tau^2})=\tau^2=\sigma^2+J\sigma_A^2.$$

Parametro σ_A^2 NMD įvertinys

$$\hat{\sigma}_A^2 = (MS_A - MS_E)/J, \quad \mathbf{E}(\hat{\sigma}_A^2) = \sigma_A^2.$$

II.2.16. Šie a. d. sudaryti iš tokių a. d.$\{\bar{a}_.\}, \{a_i - \bar{a}_.\}$, $\{\bar{e}_{i.} - \bar{e}_{..}\}$, $\{\bar{e}_{ij} - \bar{e}_{i.}\}$. A. d. sistemos, pažymėtos skirtingomis raidėmis, yra nepriklausomos, taigi reikia tik patikrinti, kad nepriklausomos sistemos žymimos vienodomis raidėmis. Kadangi a. d. yra normalieji, tai pakanka patikrinti, kad jie nekoreliuoti. Gauname

$$\mathbf{Cov}\,(\bar{a}_., a_i - \bar{a}_.) = \frac{1}{J}\sigma_A^2 - \frac{J}{J^2}\sigma_A^2 = 0;$$

$$\mathbf{Cov}\,(e_{ij} - \bar{e}_{i.}, \bar{e}_{i.} - \bar{e}_{..}) = \frac{1}{J}\sigma^2 - \frac{J}{J^2}\sigma^2 + \frac{J}{J^2}\sigma^2 - \frac{J}{J^2}\sigma^2 = 0.$$

II.2.17. Remiantis **II.2.13**, **II.2.14** pratimais

$$\sqrt{IJ}\frac{\bar{Y}_{..} - \mu}{\sqrt{MS_A}} \sim S(I-1).$$

Parametro μ lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalas

$$(\underline{\mu}; \overline{\mu}) = (\bar{Y}_{..} - t_\alpha(I-1)\sqrt{MS_A}/\sqrt{IJ}; \bar{Y}_{..} + t_\alpha(I-1)\sqrt{MS_A}/\sqrt{IJ}).$$

Parametrų τ^2 ir σ^2 pasikliovimo intervalai gaunami standartiniu būdu naudojant sąryšius

$$\frac{\hat{\tau^2}(I-1)}{\tau^2} \sim \chi^2(I-1), \quad \frac{\hat{\sigma^2}I(J-1)}{\sigma^2} \sim \chi^2(I(J-1)).$$

Parametro σ_A^2 įvertinio skirstinys tiesiogiai nesuvedamas prie žinomo skirstinio. Remdamiesi SS_A ir SS_E nepriklausomumu gauname

$$\boldsymbol{V}(\hat{\sigma}_A^2) = \frac{2}{J^2}\left[\frac{(\sigma^2 + J\sigma_A^2)^2}{(I-1)^2} + \frac{\sigma^4}{I^2(J-1)^2}\right] \to 0, \quad I \to \infty.$$

Jeigu I pakankamai didelis, tai apytikslį parametro σ_A^2 pasikliovimo intervalą galima gauti aproksimuojant a. d. $(\hat{\sigma}_A^2 - \sigma_A^2)/\sqrt{\hat{\boldsymbol{V}}(\hat{\sigma}_A^2)}$ skirstinį standartiniu normaliuoju. Stogelis virš $\boldsymbol{V}$ reiškia, kad dispersijos išraiškoje parametrai pakeisti jų įvertiniais.

Kriterijai dėl šių parametrų reikšmių sudaromi standartiniu būdu (pavyzdžiui, suformuluojant juos pasikliovimo intervalų terminais).

II.2.18. Kadangi SS_A ir SS_E nepriklausomi ir

$$MS_A/(\sigma^2 + J\sigma_A^2) \sim \chi^2_{I-1}/(I-1), \quad MS_E/\sigma^2 \sim \chi^2_{I(J-1)}/(I(J-1)),$$

tai santykis

$$\frac{MS_A}{MS_E(1 + J\Delta^2)} \sim F(I-1, I(J-1)), \quad \Delta^2 = \frac{\sigma_A^2}{\sigma^2}.$$

Iš čia standartiniu būdu galime surasti parametro Δ^2 pasikliovimo intervalus ar kriterijus tikrinant hipotezes dėl šio parametro reikšmių. Atskiru atveju hipotezė $H : \Delta^2 = 0$ yra ekvivalenti hipotezei $H_A : \sigma_A^2 = 0$. Hipotezė H_A atmetama reikšmingumo lygmens α kriterijumi, kai

$$F_A = \frac{MS_A}{MS_E} > F_\alpha(I-1, I(J-1)).$$

Gauname tokį pat kriterijų, kaip ir schemoje su pastoviu faktoriumi. Tačiau kriterijaus galia išreiškiama centrinio Fišerio skirstinio pasiskirstymo funkcija:

$$\beta(\sigma_A^2/\sigma^2) = \beta(\Delta^2) = \mathbf{P}\{F_{I-1,I(J-1)} > F_\alpha(I-1, I(J-1))/(1+J\Delta^2)\} \to 1, \quad \Delta^2 \to \infty.$$

Dispersinės analizės lentelė yra tokia pati, kaip ir **II.2.3** pratime, tik parametro σ_A^2 prasmė kitokia.

II.2.19. a) Kai faktorius atsitiktinis, gauname tokio pat pavidalo dispersinės analizės lentelę, kaip ir pratime **II.2.9**; $MS_A = 8,156, MS_E = 0,0504$. Statistika $F_A = 161,81$; P reikšmė $pv = \mathbf{P}\{F_{3,36} > 161,81\} = 0,00032$. b) $\hat{\mu} = 5,82$; $(\underline{\mu}; \overline{\mu})(4,38; 7,26)$; $\hat{\sigma}^2 = 0,0504$; $(\underline{\sigma^2}; \overline{\sigma^2}) = (0,033; 0,085)$; $\hat{\sigma}_A^2 = 0,8105$; $(\underline{\sigma^2}_A; \overline{\sigma^2}_A) = (0,257; 11,33)$; c) $(4,5; 227,3)$.

II.2.20. Pagal pirmos eilutės duomenis: $\hat{\mu} = 53,17; \hat{\sigma}^2 = 1,144; \hat{\sigma}_A^2 = 11,715$; pagal antros eilutės duomenis: $\hat{\mu} = 52,26; \hat{\sigma}^2 = 2,537; \hat{\sigma}_A^2 = 13,133$; pagal trečios eilutės duomenis: $\hat{\mu} = 47,32; \hat{\sigma}^2 = 4,926; \hat{\sigma}_A^2 = 14,807$.

II.2.21. Kriterijus išliks toks pat, tačiau kriterijaus galia bus išreiškiama centrinio Fišerio skirstinio pasiskirstymo funkcija.

II.2.2 skyrelis

II.2.22. a) Tikėtinumo funkcija

$$L(\boldsymbol{\theta}) = C\exp\{-\frac{1}{2\sigma^2}\sum_i\sum_j\sum_k (Y_{ijk} - \mu_{ij})^2 - \frac{IJK}{2}\ln\sigma^2\} =$$

$$C\exp\{-\frac{1}{2\sigma^2}\sum_i\sum_j\sum_k Y_{ijk}^2 + \frac{K}{\sigma^2}\sum_i\sum_j \mu_{ij}\bar{Y}_{ij.} - B(\boldsymbol{\theta})\}$$

priklauso $(IJ+1)$-parametrei eksponentinio tipo skirstinių šeimai. Remiantis faktorizacijos kriterijumi $\boldsymbol{T}$ yra pakankamoji statistika. Kadangi parametrų kitimo sritis turi vidinių taškų, tai $\boldsymbol{T}$ yra pilnoji statistika.

b) Bet kuri pilnosios ir pakankamosios statistikos $\boldsymbol{T}$ funkcija yra jos vidurkio NMD įvertinys. Kadangi $\mathbf{E}(\bar{Y}_{ij.}) = \mu_{ij}$, tai $\hat{\mu}_{ij} = \bar{Y}_{ij.}$ yra parametro μ_{ij} NMD įvertinys. Liekamoji kvadratų suma

$$SS_E/\sigma^2 = \sum_i\sum_j\sum_k (Y_{ijk} - \bar{Y}_{ij.})^2/\sigma^2 \sim \chi^2(IJ(K-1)).$$

Taigi

$$s^2 = \frac{SS_E}{IJ(K-1)} = MS_E, \quad \mathbf{E}s^2 = \sigma^2,$$

yra parametro σ^2 NMD įvertinys.

II.2.23. Parametrai $\mu, \alpha_i, \beta_j, \gamma_{ij}$ yra tiesinės parametrų μ_{ij} funkcijos. Nepaslinktieji (kartu pačiu ir NMD) įvertiniai gaunami imant tokias pačias tiesines įvertinių $\hat{\mu}_{ij}$ funkcijas:

$$\hat{\mu} = \bar{Y}_{...}, \quad \hat{\alpha}_i = \bar{Y}_{i..} - \bar{Y}_{...}, \quad \hat{\beta}_j = \bar{Y}_{.j.} - \bar{Y}_{...},$$
$$\hat{\gamma}_{ij} = \bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}.$$

Šie įvertiniai tenkina analogiškas sąlygas, kaip ir parametrai

$$\sum_i \hat{\alpha}_i = 0, \quad \sum_j \hat{\beta}_j = 0, \quad \sum_i \hat{\gamma}_{ij} \equiv 0, \quad \sum_j \hat{\gamma}_{ij} \equiv 0.$$

II.2.24. Hipotezes galima suformuluoti **II.1.21** pratime nurodytu būdu $H : \boldsymbol{H\beta} = \boldsymbol{\theta}_0$, kai matricos $\boldsymbol{H}$ rangas yra atitinkamai $I-1, J-1, (I-1)(J-1)$. Remiantis **II.1.21** pratimo p. d) hipotezė H_A atmetama, kai

$$F_A = \frac{(SS_{EH_A} - SS_E)IJ(K-1)}{(I-1)SS_E} = \frac{SS_A IJ(K-1)}{(I-1)SS_E} = \frac{MS_A}{MS_E} > F_\alpha(I-1, IJ(K-1)),$$

čia SS_E liekamoji kvadratų suma (žr. **II.2.22** pratimą); SS_{EH_A} yra sąlyginis kvadratinės formos $SS(\boldsymbol{\beta})$ minimumas, surastas esant sąlygai, kad H_A teisinga.

Sąlyginį kvadratinės formos minimumą lengva surasti tokiu būdu. Tapatybės

$$Y_{ijk} - \mu_{ij} = (Y_{ijk} - \hat{\mu}_{ij}) + (\hat{\alpha}_i - \alpha) + (\hat{\beta}_j - \beta) + (\hat{\gamma}_{ij} - \gamma_{ij}) + (\hat{\mu} - \mu)$$

abi puses pakelkime kvadratu ir susumuokime pagal i, j, k. Gauname

$$\boldsymbol{S}(\boldsymbol{\beta}) = SS_E + JK\sum_i(\hat{\alpha}_i - \alpha)^2 + IK\sum_j(\hat{\beta}_j - \beta)^2 + K\sum_i\sum_j(\hat{\gamma}_{ij} - \gamma_{ij})^2 + IJK(\hat{\mu} - \mu)^2,$$

nes visos mišrios sandaugos lygios nuliui. Tada

$$SS_{EH_A} = \min_{\alpha_i = 0} SS(\boldsymbol{\beta}) = SS_E + JK\sum_i \hat{\alpha}_i^2,$$

$$SS_A = JK\sum_i \hat{\alpha}_i^2 = JK\sum_i(\bar{Y}_{i..} - \bar{Y}_{...})^2.$$

Analogiškai

$$SS_B = IK\sum_j \hat{\beta}_j^2 = IK\sum_j(\bar{Y}_{.j.} - \bar{Y}_{...})^2,$$

$$SS_{AB} = K\sum_i\sum_j(\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2.$$

Hipotezės H_B, H_{AB} atmetamos, kai atitinkamai

$$F_B = \frac{MS_B}{MS_E} > F_\alpha(J-1, IJ(K-1)), \quad F_{AB} = \frac{MS_{AB}}{MS_E} > F_\alpha((I-1)(J-1), IJ(K-1)).$$

Skaičiavimo rezultatai analizę atliekant kompiuteriu pateikiami tokioje lentelėje.

Dispersinės analizės lentelė

Faktorius	SS	ν	$MS = SS/\nu$	$\mathbf{E}(MS)$
A	SS_A	$I-1$	MS_A	$\sigma^2 + JK\sigma_A^2$
B	SS_B	$J-1$	MS_B	$\sigma^2 + IK\sigma_B^2$
$A \times B$	SS_{AB}	$(I-1)(J-1)$	MS_{AB}	$\sigma^2 + K\sigma_{AB}^2$
E	SS_E	$IJ(K-1)$	MS_E	σ^2
T	SS_T			

Parametrai $\sigma_a^2, \sigma_B^2, \sigma_{AB}^2$ pateikti **II.2.26** pratime.

II.2.25. a) Pakanka patikrinti, kad bet kuris sistemų $\{\bar{Y}_{...}\}$, $\{\bar{Y}_{i..} - \bar{Y}_{...}\}$, $\{\bar{Y}_{.j.} - \bar{Y}_{...}\}$, $\{\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}\}$, $\{Y_{ijk} - \bar{Y}_{ij.}\}$ a. d. yra nekoreliuotas su bet kuriuo kitos sistemos a. d. Pavyzdžiui,

$$\mathbf{Cov}\,(\bar{Y}_{i..} - \bar{Y}_{...}, \bar{Y}_{.j.} - \bar{Y}_{...}) = \frac{K\sigma^2}{JKIK} - \frac{JK\sigma^2}{JKIJK} - \frac{IK\sigma^2}{IJKIK} + \frac{IJK\sigma^2}{(IJK)^2} = 0.$$

b) Reikia daugianarį pakelti kvadratu ir sutraukti panašiuosius narius.

c) Užrašę kvadratų sumas p. b) nurodytu būdu ir sudėję, gausime pateiktą lygybę.

II.2.26. a) Remiantis **II.1.21** pratimu

$$SS_A/\sigma^2 \sim \chi^2(I-1; \lambda_A), \quad \lambda_A = JK\sum_i \alpha_i^2/\sigma^2.$$

Gauname

$$\mathbf{E}(MS_A) = \mathbf{E}(\frac{SS_A}{I-1}) = \frac{\sigma^2}{I-1}(I - 1 + \lambda_A) = \sigma^2 + \frac{JK}{I-1}\sum_i \alpha_i^2,$$

taigi **II.2.24** pratime $\sigma_A^2 = \sum_i \alpha_i^2/(I-1)$. Analogiškai gauname

$$\mathbf{E}(MS_B) = \sigma^2 + \frac{IK}{J-1}\sum_j \beta_j^2, \quad \mathbf{E}(MS_{AB}) = \sigma^2 + \frac{K}{(I-1)(J-1)}\sum_i\sum_j \gamma_{ij}^2.$$

b) Kriterijų galios išreiškiamos necentrinio Fišerio skirstinio pasiskirstymo funkcijomis. Pavyzdžiui, hipotezės H_A tikrinimo kriterijaus galia

$$\beta_A(\lambda_A) = \mathbf{P}\{F_{I-1,IJ(K-1);\lambda_A} > F_\alpha(I-1, IJ(K-1))\}.$$

Analogiškai hipotezių H_B ir H_{AB} atvejais.

II.2.27. Imdami kvadratinės formos

$$SS(\boldsymbol{\beta}) = \sum_i\sum_j (Y_{ij} - \mu - \alpha_i - \beta_j)^2$$

išvestines pagal μ, α_i, β_j prilyginę nuliui, gausime įvertinius

$$\hat{\mu} = \bar{Y}_{..}, \quad \hat{\alpha}_i = \bar{Y}_{i.} - \bar{Y}_{..},\ i = 1, ..., I; \hat{\beta}_j = \bar{Y}_{.j} - \bar{Y}_{..},\ j = 1, ..., J.$$

Gautieji įvertiniai tenkina sąlygas

$$\sum_i \hat{\alpha}_i = 0, \quad \sum_j \hat{\beta}_j = 0.$$

Liekamoji kvadratų suma

$$SS_{AB} = \sum_i \sum_j (Y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j)^2 = \sum_i \sum_j (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2 =$$

$$\sum_i \sum_j (e_{ij} - \bar{e}_{i.} - \bar{e}_{.j} + \bar{e}_{..})^2 \sim \sigma^2 \chi^2_{(I-1)(J-1)}.$$

Dispersijos σ^2 nepaslinktasis įvertinys

$$s^2 = \frac{SS_{AB}}{(I-1)(J-1)} = MS_{AB}, \quad \mathbf{E}s^2 = \sigma^2.$$

b) Išreiškę $\alpha_I = -\alpha_1 - ... - \alpha_{I-1}$ ir $\beta_J = -\beta_1 - ... - \beta_{J-1}$, gausime tiesinį modelį

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{\beta} + \boldsymbol{e},$$

kai nežinomų parametrų vektorius $\boldsymbol{\beta} = (\mu, \alpha_1, ..., \alpha_{i-1}, \beta_1, ..., \beta_{J-1})^T$. Plano matrica $\boldsymbol{A}$ turi IJ eilučių ir $1 + (I-1) + (J-1) = I + J - 1$ stulpelių. Pirmojo matricos $\boldsymbol{A}$ stulpelio elementai lygūs 1. Tolimesni $I-1$ stulpelių yra tokie: pirmosios J eilučių turi pavidalą $(1,\ 0, ..., 0)$, paskui J eilučių turi pavidalą $(0,\ 1, ..., 0)$, pagaliau J eilučių turi pavidalą $(0,\ 0, ..., 1)$; paskutinės J eilučių turi pavidalą $(-1,\ -1, ..., -1)$. Aptarsime paskutinius $J-1$ matricos $\boldsymbol{A}$ stulpelius: pirmoji eilutė yra $(1,\ 0, ..., 0)$, antroji $(0,\ 1, ..., 0)$, $(J-1)$-oji $(0,\ 0, ..., 1)$, o J-oji $(-1,\ -1, ..., -1)$. Po to šios J eilučių pakartojamos dar $I-1$ kartą.

Remiantis **II.1.26** pratimu statistika

$$\boldsymbol{T} = ((\boldsymbol{A}^T\boldsymbol{Y})^T, \boldsymbol{Y}^T\boldsymbol{Y}) = (IJ\bar{Y}_{..}, J(\bar{Y}_{1.} - \bar{Y}_{I.}), ..., J(\bar{Y}_{I-1,.} - \bar{Y}_{I.}),$$

$$I(\bar{Y}_{.1} - \bar{Y}_{.J}), ..., I(\bar{Y}_{.,J-1} - \bar{Y}_{.J}), \boldsymbol{Y}^T\boldsymbol{Y})^T$$

yra pilnoji ir pakankamoji parametro $\boldsymbol{\beta}$ statistika.

Kadangi p. a) surasti įvertiniai yra nepaslinktieji ir yra statistikos T funkcijos, tai jie yra NMD įvertiniai.

II.2.28. Analogiškai kaip **II.2.22** pratime, gausime, kad hipotezės H_A, H_B atmetamos reikšmingumo lygmens α kriterijais, kai atitinkamai

$$F_A = \frac{MS_A}{MS_{AB}} > F_\alpha(I-1, (I-1)(J-1)), \quad F_B = \frac{MS_B}{MS_{AB}} > F_\alpha(J-1, (I-1)(J-1)),$$

čia

$$MS_A = \frac{SS_A}{I-1} = \frac{J}{I-1}\sum_i (\bar{Y}_{i.} - \bar{Y}_{..})^2, \quad MS_B = \frac{SS_B}{J-1} = \frac{I}{J-1}\sum_j (\bar{Y}_{.j} - \bar{Y}_{..})^2,$$

o SS_{AB} liekamoji kvadratų suma iš **II.2.27** pratimo. Lyginant su **II.2.24** pratimo kriterijais skirtumas yra tas, kad statistikų vardiklyje vietoje SS_E yra SS_{AB}.

II.2.29. Įrodyta, kad esant teisingai hipotezei H_γ (žr. [3], 2.2.5 skyrelį) statistika

$$SS_\gamma = \frac{\sum_i \sum_j \hat{\alpha}_i \hat{\beta}_j Y_{ij}}{\sum_i \hat{\alpha}_i^2 \sum_j \hat{\beta}_j^2} \sim \sigma^2 \chi_1^2,$$

čia $\hat{\alpha}_i = \bar{Y}_{i.} - \bar{Y}_{..}$, $\hat{\beta}_j = \bar{Y}_{.j} - \bar{Y}_{..}$, ir nepriklauso nuo

$$SS_L = SS_{AB} - SS_\gamma \sim \sigma^2 \chi^2_{IJ-I-J}.$$

Tjukio kriterijus atmeta hipotezę H_γ reikšmingumo lygmens α kriterijumi, kai

$$F_\gamma = \frac{SS_\gamma(IJ-I-J)}{SS_L} > F_\alpha(1, IJ-I-J).$$

II.2.30. Stjudento kriterijus atmeta hipotezę H_A, kai

$$|t| = \sqrt{J(J-1)}\frac{|\bar{Y}_{1.} - \bar{Y}_{2.}|}{\sqrt{\sum_j (Z_j - \bar{Z}_.)^2}} > T_{\alpha/2}(J-1).$$

Randame

$$SS_A = J\sum_{i=1}^{2}(\bar{Y}_{i.} - \bar{Y}_{..})^2 = J[(\bar{Y}_{1.} - \frac{\bar{Y}_{1.} + \bar{Y}_{2.}}{2})^2 + (\bar{Y}_{2.} - \frac{\bar{Y}_{1.} + \bar{Y}_{2.}}{2})^2] =$$

$$J(\bar{Y}_{1.} - \bar{Y}_{2.})^2/2;$$

$$SS_{AB} = \sum_j (Y_{1j} - \bar{Y}_{1.} - \bar{Y}_{.j} + \bar{Y}_{..})^2 + \sum_j (Y_{2j} - \bar{Y}_{2.} - \bar{Y}_{.j} + \bar{Y}_{..})^2 =$$

$$\sum_j (Z_j - \bar{Z}_.)^2/2.$$

Gauname

$$F_A = \frac{MS_A}{MS_{AB}} = \frac{SS_A(J-1)}{SS_{AB}} = J(J-1)\frac{(\bar{Y}_{1.} - \bar{Y}_{2.})^2}{\sum_j (Z_j - \bar{Z}_.)^2} = t^2.$$

Kriterijus

$$F_A > F_\alpha(1, J-1),$$

akivaizdu, ekvivalentus Stjudento kriterijui.

II.2.31. a) Statistika, kuri esant teisingai hipotezei turi Fišerio skirstinį su 1 ir 9 laisvės laipsniais, įgijo reikšmę 21,824; hipotezė atmetama kriterijumi, kurio reikšmingumo lygmuo $\alpha > \mathbf{P}\{F_{1,9} > 21,824\} = 0,0012$; b) $n \geq 32$.

II.2.32. Atlikę skaičiavimus gauname dispersinės analizės lentelę

Faktorius	SS	ν	MS	F	Pr>F
A	3.649	3	1.216	2.48	0.0980
B	0.331	1	0.331	0.68	0.4228
$A \times B$	0.043	3	0.014	0.03	0.9930
E	7.835	16	0.490		
T	11.859	23			

Gauname statistikų realizacijas $F_A = 2,48, F_B = 0,68, F_{AB} = 0,03$, kurias atitinka P reikšmės 0,0980, 0,4228, 0,9930.

II.2.33. Tjukio statistika F_γ įgijo reikšmę 0,7234; $\mathbf{P}\{F_{1,\,5} > 0,7234\} = 0,4339$; atmesti hipotezę nėra pagrindo. Gauname statistikų realizacijas $F_A = 9,01, F_B = 4,61$, kurias atitinka P reikšmės 0,0122, 0,0613.

II.2.34. Tjukio statistika F_γ įgijo reikšmę 2,2754; $\mathbf{P}\{F_{1,\,34} > 2,2754\} = 0,1407$; atmesti hipotezę nėra pagrindo. Gauname statistikų realizacijas $F_A = 9,11, F_B = 15,31$. Hipotezės atmestinos.

II.2.35. Kvadratų suma

$$SS_B/\sigma^2 \sim \chi^2(J-1;\lambda_B), \quad \lambda_B = I\sum_{i=1}^{8}\beta_j^2.$$

Vidurkis $\bar{\mu} = \mu + 10$; $\beta_6 = \beta_7 = \beta_8 = -10$, $\beta_1 = \beta_2 = \beta_6 = 2$, $\beta_3 = \beta_4 = 12$ ir $\lambda_B = 3600/\sigma^2$ (σ^2 iš **2.34** pratimo). Statistika F_B turi necentrinį Fišerio skirstinį su 7 ir 35 laisvės laipsniais ir necentriškumo parametru λ_B. Randame $\mathbf{P}\{F_{7,35;\lambda_B} > F_{0,05}(7,35)\}$.

II.2.36. Tarkime, $\mu_5 = \mu_7 = \mu, \mu_8 = \mu + 10$. Tada

$$SS_B/\sigma^2 \sim \chi^2(2;\lambda_B), \quad \lambda_B = n\sum_j \beta_j^2.$$

Vidurkis $\bar{\mu} = \mu + 10/3$; $\beta_5 = \beta_7 = 10/3$, $\beta_8 = -20/3$, ir $\lambda_B = 200n/(3\sigma^2)$ (σ^2 iš **II.2.34** pratimo).

Imties didumui rasti turime nelygybę

$$\mathbf{P}\{F_{2,2(n-1);\lambda_B} > F_{0,05}(2,2(n-1))\} \geq 0,8.$$

II.2.37. Gauname statistikų realizacijas $F_A = 8,31, F_B = 5,94, F_{AB} = 1,29$, kurias atitinka P reikšmės 0,000021, 0,0002, 0,2206.

II.2.38. Statistika F_{AB} įgijo reikšmę 1,92; P reikšmė 0,0807; a) Gauname statistikų realizacijas $F_A = 4,54, F_B = 0,42$, kurias atitinka P reikšmės 0,0066, 0,7362; b) Gauname statistikų realizacijas $F_A = 4,40, F_B = 0,26$, kurias atitinka P reikšmės 0,0085, 0,8546.

II.2.39. Randame

$$SS_A = JK[(\bar{Y}_{1..} - \frac{\bar{Y}_{1..}+\bar{Y}_{2..}}{2})^2 + (\bar{Y}_{2..} - \frac{\bar{Y}_{1..}+\bar{Y}_{2..}}{2})^2] =$$
$$JK(\bar{Y}_{1..} - \bar{Y}_{2..})^2/2;$$
$$SS_{AB} = K[\sum_j(\bar{Y}_{1j.} - \bar{Y}_{1..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2 + \sum_j(\bar{Y}_{2j.} - \bar{Y}_{2..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2] =$$
$$K\sum_j(\bar{Y}_{1j.} + \bar{Y}_{2j.})^2/2 - JK(\bar{Y}_{1..} + \bar{Y}_{2..})^2/2.$$

Jeigu $K = 2$, tai

$$SS_E = \sum_i\sum_j[(Y_{ij1} - \bar{Y}_{ij.})^2 + (Y_{ij2} - \bar{Y}_{ij.})^2] = \sum_i\sum_j(Y_{ij1} - Y_{ij2})^2/2.$$

II.2.40. a) Gauname

$$SS_E = \sum_i \sum_j \sum_k (Y_{ijk} - \bar{Y}_{ij.})^2 = \sum_i \sum_j \sum_k (e_{ijk} - \bar{e}_{ij.})^2 \sim \sigma^2 \chi^2_{IJ(K-1)}$$

ir $\mathbf{E}(MS_E) = \sigma^2$.

$$SS_A = JK \sum_i (\bar{Y}_{i..} - \bar{Y}_{...})^2 = JK \sum_i (a_i + \bar{c}_{i.} + \bar{e}_{i..} - (\bar{a}_. + \bar{c}_{..} + \bar{e}_{...}))^2 =$$

$$JK \sum_i (Z_i - \bar{Z}_.)^2 \sim (\sigma^2 + K\sigma^2_{AB} + JK\sigma^2_A)\chi^2_{I-1},$$

nes vienodai pasiskirstę n. a. d. $Z_i = a_i + \bar{c}_{i.} + \bar{e}_{i..} \sim N(0, \sigma^2_A + \sigma^2_{AB}/J + \sigma^2/(JK))$. Taigi $\mathbf{E}(MS_A) = \sigma^2 + K\sigma^2_{AB} + JK\sigma^2_A$.

Analogiškai gauname, kad

$$\mathbf{E}(MS_B) = \sigma^2 + K\sigma^2_{AB} + IK\sigma^2_B, \quad \mathbf{E}(MS_{AB}) = \sigma^2 + K\sigma^2_{AB}.$$

Dispersinės analizės lentelė skirsis nuo **II.2.24** pratimo lentelės tik paskutiniuoju stulpeliu, kuriame reikia įrašyti šiame pratime gautas $\mathbf{E}(MS)$ išraiškas.

b) Parametro μ NMD įvertinys

$$\hat{\mu} = \bar{Y}_{...} = \mu + \bar{a}_. + \bar{b}_. + \bar{c}_{..} + \bar{e}_{...} \sim N(\mu, \sigma^2_A/I + \sigma^2_B/J + \sigma^2_{AB}/(IJ) + \sigma^2/(IJK)).$$

Kitų parametrų NMD įvertiniai randami remiantis gautomis $\mathbf{E}(MS)$ išraiškomis.

$$\hat{\sigma}^2_A = \frac{MS_A - MS_{AB}}{JK}, \quad \hat{\sigma}^2_B = \frac{MS_B - MS_{AB}}{IK},$$

$$\hat{\sigma}^2_{AB} = \frac{MS_{AB} - MS_E}{K}, \quad \hat{\sigma}^2 = MS_E.$$

II.2.41. a) Kvadratų sumos sudarytos iš tokių a. d. sistemų $\{a_i - \bar{a}_i\}$, $\{b_j - \bar{b}_j\}$, $\{c_{ij} - \bar{c}_{i.} - \bar{c}_{.j} + \bar{c}_{..}\}$, $\{\bar{c}_{.j} + \bar{c}_{..}\}$, $\{\bar{c}_{i.} + \bar{c}_{..}\}$, $\{\bar{e}_{i..} + \bar{e}_{...}\}$, $\{\bar{r}_{.j.} + \bar{e}_{...}\}$, $\{\bar{e}_{ij.} - \bar{e}_{i..} - \bar{e}_{.j.} + \bar{e}_{...}\}$, $\{e_{ijk} - \bar{e}_{ij.}\}$. Kadangi a. d., žymimi skirtingomis raidėmis, yra nepriklausomi, tai reikia tik įsitikinti, kad a. d. sistemos, sudarytos iš vienodomis raidėmis žymimų a. d., yra nepriklausomos. Tarkime, kad schemoje su pastoviais faktoriais kintamuosius Y_{ijk} atitinka e_{ijk}. Tada remdamiesi **II.2.25** pratimu gauname, kad a. d., žymimų raidėmis e, sistemos yra nepriklausomos. Analogiškai tarę, kad schemoje su vienu stebėjime langelyje kintamuosius Y_{ij} atitinka c_{ij}, gausime, kad a. d. sistemos, žymimos raidėmis c, yra nepriklausomos.

b) Remdamiesi p. a) ir **II.2.40** pratimu gauname

$$\boldsymbol{V}(\hat{\sigma^2_A}) = \frac{1}{J^2K^2}[\boldsymbol{V}(MS_A) + \boldsymbol{V}(MS_{AB})] = \frac{2}{J^2K^2}\left[\frac{(\mathbf{E}(MS_A))^2}{I-1} + \frac{(\mathbf{E}(MS_{AB}))^2}{(I-1)(J-1)}\right],$$

$$\boldsymbol{V}(\hat{\sigma^2_B}) = \frac{2}{I^2K^2}\left[\frac{(\mathbf{E}(MS_B))^2}{J-1} + \frac{(\mathbf{E}(MS_{AB}))^2}{(I-1)(J-1)}\right], \quad \boldsymbol{V}(\hat{\sigma}^2) = \frac{2\sigma^4}{IJ(K-1)},$$

$$\boldsymbol{V}(\hat{\sigma}^2_{AB}) = \frac{2}{K^2}\left[\frac{(\mathbf{E}(MS_{AB}))^2}{(I-1)(J-1)} + \frac{(\mathbf{E}(MS_E))^2}{IJ(K-1)}\right].$$

c) Remiantis **II.2.40** pratimu santykiai $SS/\mathbf{E}(MS) \sim \chi^2(\nu)$. Pasirinkę bet kurių dviejų skirtingų trupmenų $MS/\mathbf{E}(MS)$ santykius gausime Fišerio skirstinį su atitinkamais laisvės laipsniais. Tikrinant hipotezę $H_A : \sigma_A^2 = 0$, reikia imti santykį trupmenų $MS_A/\mathbf{E}(MS_A)$ ir $MS_{AB}/\mathbf{E}(MS_{AB})$, nes esant teisingai hipotezei H_A vidurkiai $\mathbf{E}(MS_A)$ ir $\mathbf{E}(MS_{AB})$ sutampa. Gauname statistiką $F_A = MS_A/MS_{AB}$. Hipotezė H_A atmetama reikšmingumo lygmens α kriterijumi, kai

$$F_A = \frac{MS_A}{MS_{AB}} > F_\alpha(I-1,(I-1)(J-1)).$$

Analogiškai hipotezės H_B ir H_{AB} atmetamos, kai

$$F_B = \frac{MS_B}{MS_{AB}} > F_\alpha(J-1,(I-1)(J-1)),$$

$$F_{AB} = \frac{MS_{AB}}{MS_E} > F_\alpha((I-1)(J-1), IJ(K-1)).$$

Matome, kad, skirtingai nuo schemos su pastoviais faktoriais, tikrinant hipotezes H_A ir H_B statistikos vardiklyje vietoje MS_E yra MS_{AB}.

Jeigu hipotezės neteisingos, tai statistikos skiriasi nuo Fišerio a. d. tik daugikliu. Todėl galios funkcijos išreiškiamos centrinio Fišerio skirstinio pasiskirstymo funkcija. Pavyzdžiui, tikrinant hipotezę $H_A : \sigma^2 = 0$, kriterijaus galios funkcija

$$\beta(\sigma_A^2) = \mathbf{P}\{\frac{MS_A}{MS_{AB}} > F_\alpha(I-1,(I-1)(J-1))|\sigma_A^2\} =$$

$$\mathbf{P}\{F_{I-1,(I-1)(J-1)} > \frac{\sigma^2 + K\sigma_{AB}^2}{\sigma^2 + K\sigma_{AB}^2 + JK\sigma_A^2} F_\alpha(I-1,(I-1)(J-1))\}$$

artėja prie vieneto, kai santykis $JK\sigma_A^2/(\sigma^2 + K\sigma_{AB}^2)$ didėja.

II.2.42. a) Hipotezė atmetama α lygmens kriterijumi, kai $F_A = MS_A/MS_{AB} > (1 + JK\Delta)F_\alpha(I-1,(I-1)(J-1))$; b) Hipotezė atmetama α lygmens kriterijumi, kai $F_{AB} = MS_{AB}/MS_E > (1 + K\Delta)F_\alpha((I-1)(J-1), IJ(K-1))$.

II.2.43. a) Gauname $F_A = 21{,}196$, $F_B = 2439{,}686$, $F_{AB} = 28{,}846$; visos trys hipotezės atmetamos. b) $\hat\sigma^2 = 0{,}221$; $\hat\sigma_{AB}^2 = 0{,}559$; $\hat\sigma_A^2 = 2{,}926$; $\hat\sigma_B^2 = 56{,}533$; $\hat{\boldsymbol{V}}(\hat\sigma^2) = 0{,}000089$; $\hat{\boldsymbol{V}}(\hat\sigma_{AB}^2) = 0{,}00933$; $\hat{\boldsymbol{V}}(\hat\sigma_A^2) = 0{,}7865$; $\hat{\boldsymbol{V}}(\hat\sigma_B^2) = 2132{,}415$. c) Taikydami normaliąją aproksimaciją gauname parametrų $\sigma_A^2, \sigma_B^2, \sigma_{AB}^2, \sigma^2$ pasikliovimo intervalus: $(1{,}188; 4{,}664)$, $(0; 147{,}042)$, $(0{,}370; 0{,}748)$, $(0{,}203; 0{,}239)$.

II.2.44. Gauname statistikų realizacijas $F_A = 9{,}01, F_B = 4{,}61$, kurias atitinka P reikšmės 0,0122, 0,0613.

II.2.45. Gauname statistikų realizacijas $F_A = 6{,}44, F_B = 4{,}60, F_{AB} = 1{,}29$, kurias atitinka P reikšmės 0,0017, 0,0059, 0,2206.

II.2.46. Kvadratų suma

$$SS_E = \sum_i \sum_j \sum_k (Y_{ijk} - \bar Y_{ij.})^2 = \sum_i \sum_j \sum_k (e_{ijk} - \bar e_{ij.})^2 \sim \sigma^2 \chi^2_{IJ(K-1)}$$

turi tokias pačias savybes, kaip ir ankstesnėse schemose; $\mathbf{E}(MS_E) = \sigma^2$. Kvadratų suma

$$SS_B = IK \sum_j (\bar{Y}_{.j.} - \bar{Y}_{...})^2 = IK \sum_j (b_j + \bar{e}_{.j.} - \bar{b}_{.} - \bar{e}_{...})^2 \sim (\sigma^2 + IK\sigma_B^2)\chi^2_{J-1},$$

$$\mathbf{E}(MS_B) = \sigma^2 + IK\sigma_B^2,$$

nes vienodai pasiskirstę n. a. d. $b_j + \bar{e}_{.j.} \sim N(0, \sigma_B^2 + \sigma^2/(IK))$.

Likusių kvadratų sumų

$$SS_A = JK \sum_i (\bar{Y}_{i..} - \bar{Y}_{...})^2 = JK \sum_i (\alpha_i + \bar{c}_{i.} + \bar{e}_{i..} - \bar{e}_{...})^2,$$

$$SS_{AB} = K \sum_i \sum_j (c_{ij} - \bar{c}_{i.} + \bar{e}_{ij.} - \bar{e}_{i..} - \bar{e}_{.j.} + \bar{e}_{...})^2$$

skirstiniai bendru atveju nesuvedami į χ^2 skirstinius. Adityviuoju atveju ($c_{ij} \equiv 0$) SS_A suvedama į necentrinį χ^2, o SS_{AB} – į centrinį χ^2 skirstinį. Rasime vidurkius

$$\mathbf{E}(SS_A) = JK \sum_i \alpha_i^2 + JK \sum_i \mathbf{E}(\bar{c}_{i.})^2 + JK\mathbf{E}(\sum_i (\bar{e}_{i..} - \bar{e}_{...})^2) =$$

$$JK \sum_i \alpha_i^2 + K \sum_i \sigma^2_{AB;i} + \sigma^2(I-1),$$

$$\mathbf{E}(MS_A) = \sigma^2 + K\sigma^2_{AB} + JK\sigma_A^2; \quad \sigma_A^2 == \frac{1}{I-1} \sum_i \alpha_i^2, \quad \sigma^2_{AB} = \frac{1}{I-1} \sum_i \sigma^2_{AB;i}.$$

$$\mathbf{E}(SS_{AB}) = K\mathbf{E}(\sum_i \sum_j (c_{ij} - \bar{c}_{i.})^2) + K\mathbf{E}(\sum_i \sum_j (\bar{e}_{ij.} - \bar{e}_{i..} - \bar{e}_{.j.} + \bar{e}_{...})^2) =$$

$$K(J-1) \sum_i \sigma^2_{AB;i} + (I-1)(J-1)\sigma^2; \quad \mathbf{E}(MS_{AB}) = \sigma^2 + K\sigma^2_{AB}.$$

Dispersinės analizės lentelė skirsis nuo **II.2.24** pratimo lentelės tik paskutiniu stulpeliu, kuriame reikia įrašyti šiame pratime surastas $\mathbf{E}(MS)$ išraiškas.

II.2.47. a) Naudodami surastas $\mathbf{E}$(MS) išraiškas gauname

$$\hat{\sigma}_A^2 = \frac{MS_A - MS_{AB}}{JK}, \quad \hat{\sigma}_B^2 = \frac{MS_B - MS_{AB}}{IK},$$

$$\hat{\sigma}^2_{AB} = \frac{MS_{AB} - MS_E}{K}, \quad \hat{\sigma}^2 = MS_E.$$

b) Hipotezė H_B atmetama reikšmingumo lygmens α kriterijumi, kai

$$F_B = \frac{MS_B}{MS_E} > F_\alpha(J-1, IJ(K-1));$$

kriterijaus galia išreiškiama centrinio Fišerio skirstinio pasiskirstymo funkcija. Atkreipsime dėmesį, kad tikrinama hipotezė apie atsitiktinio faktoriaus įtakos nebuvimą, tačiau statistikos vardiklyje yra MS_E (kaip schemoje su pastoviais faktoriais).

Jeigu hipotezė H_{AB} teisinga (t. y. $c_{ij} \equiv 0$), tai santykis

$$F_{AB} = \frac{MS_{AB}}{MS_E} \sim F((I-1)(J-1), IJ(K-1)).$$

Hipotezė H_{AB} atmetama, kai

$$F_{AB} > F_\alpha((I-1)(J-1), IJ(K-1)),$$

t. y. kaip ir ankstesnėse schemose. Tačiau kriterijaus galia žinomais skirstiniais neišreiškiama.

Kai teisinga hipotezė $H_A : \alpha_1 = ... = \alpha_I = 0$, statistikos

$$F_A = \frac{MS_A}{MS_{AB}}$$

skaitiklio ir vardiklio vidurkiai vienodi, tačiau jų skirstiniai nėra χ^2 skirstiniai. Todėl kriterijus: atmesti H_A, kai teisinga nelygybė

$$F_A > F_\alpha(I-1, (I-1)(J-1)),$$

yra apytikslis. Jeigu faktorių sąveikos nėra ($c_{ij} \equiv 0$), tai F_A turi centrinį Fišerio skirstinį, kai H_A teisinga, ir necentrinį Fišerio skirstinį, kai H_A neteisinga. Taigi šiuo atveju kriterijus yra tikslus.

II.2.48. a) Gauname statistikų realizacijas $F_A = 4,04, F_B = 2,63, F_{AB} = 2,46$, kurias atitinka P reikšmės 0,0613, 0,0537, 0,0534; b) Gauname statistikų realizacijas $F_A = 9,92, F_B = 2,63, F_{AB} = 2,46$, kurias atitinka P reikšmės 0,0005, 0,0537, 0,0534.

II.2.49. Gauname statistikų realizacijas $F_A = 6,29, F_B = 1,05, F_{AB} = 2,39$, kurias atitinka P reikšmės 0,0097, 0,4116, 0,0087.

II.2.50. Gauname statistikų realizacijas $F_A = 1,94, F_B = 0,49$, kurias atitinka P reikšmės 0,0801, 0,7813.

II.2.51. Gauname statistikų realizacijas $F_A = 2,48, F_B = 23,21, F_{AB} = 0,03$, kurias atitinka P reikšmės 0,0980, 0,0170, 0,9930.

II.2.52. Gauname statistikų realizacijas $F_A = 9,11, F_B = 15,312, F_{AB} = 2,46$. Hipotezes atmetame.

II.2.3 skyrelis

II.2.53. Reikia patikrinti, kad su visais i, j, l galioja lygybės

$$\mu_{ijl} - \bar{\mu}_{ij.} - \bar{\mu}_{i.l} - \bar{\mu}_{.jl} + \bar{\mu}_{i..} + \bar{\mu}_{.j.} + \bar{\mu}_{..l} - \bar{\mu}_{...} = 0.$$

II.2.54. $\hat{\mu} = \bar{Y}_{...}$, $\hat{\alpha}_i = \bar{Y}_{i..} - \bar{Y}_{...}$, $\hat{\beta}_{ij} = \bar{Y}_{ij.} - \bar{Y}_{i..}$, $\hat{\gamma}_k = \bar{Y}_{..k} - \bar{Y}_{...}$; hipotezė H_A atmetama α lygmens kriterijumi, kai $(SS_{EH_A} - SS_E)(IJK - IJ - K + 1)/(I-1)SS_E > F_\alpha(I-1, (IJ-1)(K-1))$; $SS_E = \sum_i \sum_j \sum_k (Y_{ijk} - \bar{Y}_{ij.} + \bar{Y}_{..k} - \bar{Y}_{...})^2$; $SS_{EH_A} = \sum_i \sum_j \sum_k (Y_{ijk} - \bar{Y}_{ij.} - \bar{Y}_{i..} + \bar{Y}_{..k})^2$.

II.2.55. Formalios daugiafaktorės dispersinės analizės lentelės sudarymo taisyklės subalansuotų planų atveju pateiktos [3], 2.7 skyrelyje. Tos lentelės skiriasi paskutiniuoju stulpeliu, kuriame pateikiami vidutinių kvadratų sumų vidurkiai. Trijų atsitiktinių faktorių atveju gauname

$$\mathbf{E}(MS_E) = \sigma^2, \quad \mathbf{E}(MS_{ABC}) = \sigma^2 + K\sigma^2_{ABC}, \quad \mathbf{E}(MS_{AB}) = \sigma^2 + K\sigma^2_{ABC} + LK\sigma^2_{AB},$$

$$\mathbf{E}(MS_{AC}) = \sigma^2 + K\sigma^2_{ABC} + jK\sigma^2_{AC}, \quad \mathbf{E}(MS_{BC}) = \sigma^2 + K\sigma^2_{ABC} + IK\sigma^2_{BC},$$

$$\mathbf{E}(MS_A) = \sigma^2 + K\sigma^2_{ABC} + LK\sigma^2_{AB} + JK\sigma^2_{AC} + JLK\sigma^2_A,$$

$$\mathbf{E}(MS_B) = \sigma^2 + K\sigma^2_{ABC} + LK\sigma^2_{AB} + IK\sigma^2_{BC} + ILK\sigma^2_B,$$

$$\mathbf{E}(MS_C) = \sigma^2 + K\sigma^2_{ABC} + JK\sigma^2_{AC} + IK\sigma^2_{BC} + IJK\sigma^2_C.$$

Hipotezė H_{ABC} atmetama, kai $MS_{ABC}/MS_E > F_\alpha((I-1)(J-1)(L-1), IJL(K-1))$; hipotezė H_{AB} atmetama, kai $MS_{AB}/MS_{ABC} > F_\alpha((I-1)(J-1), (I-1)(J-1)(L-1))$: analogiškai hipotezių H_{AC} ir H_{BC} atvejais; neegzistuoja tokie du MS, kad jų vidurkiai sutaptų esant teisingai hipotezei H_A; remiantis [3], 2.7 skyreliu hipotezė H_A atmetama apytiksliu Fišerio kriterijumi, kai $MS_A/(MS_{AB} - MS_{AC} - MS_{ABC}) > F_\alpha(I-1, \tilde{\nu})$; čia $\tilde{\nu} = (MS_{AB} + MS_{AC} - MS_{ABC})^2/[(MS_{AB})^2/((I-1)(J-1)) + (MS_{AC})^2/((I-1)(L-1)) + (MS_{ABC})^2/((I-1)(J-1)(L-1))]$; analogiškai tikrinamos hipotezės H_B ir H_C.

II.2.56. Gauname

$$\sum_{i=1}^{I}(Y_i - \bar{Y}_\cdot)^2 = \sum_{i=1}^{I_1}(Y_i - \bar{Y}^{(1)} + \bar{Y}^{(1)} - \bar{Y}_\cdot)^2 + \sum_{i=I_1+1}^{I}(Y_i - \bar{Y}^{(2)} + \bar{Y}^{(2)} - \bar{Y}_\cdot)^2 =$$

$$\sum_{i=1}^{I_1}(Y_i - \bar{Y}^{(1)})^2 + \sum_{i=I_1+1}^{I}(Y_i - \bar{Y}^{(2)})^2 + I_1(\bar{Y^{(1)}} - \bar{Y}_\cdot)^2 + I_2(\bar{Y^{(2)}} - \bar{Y}_\cdot)^2.$$

Kadangi $\bar{Y}_\cdot = (I_1\bar{Y}^{(1)} + \bar{Y}^{(2)})/(I_1 + I_2)$, tai paskutinių dviejų dėmenų suma yra

$$\frac{I_1 I_2}{I_1 + I_2}(\bar{Y}^{(1)} - \bar{Y}^{(2)})^2.$$

II.2.57. Tegu $I = I_1 + I_2 + I_3$. Tada analogiškai **II.2.56** pratimui

$$\sum_{i=1}^{I}(Y_i - \bar{Y}_\cdot)^2 = \sum_{i=1}^{I_1}(Y_i - \bar{Y}^{(1)})^2 + \sum_{i=I_1+1}^{I_1+I_2}(Y_i - \bar{Y}^{(2)})^2 + \sum_{i=I_1+I_2+1}^{I}(Y_i - \bar{Y}^{(3)})^2 +$$

$$I_1(\bar{Y^{(1)}} - \bar{Y}_\cdot)^2 + I_2(\bar{Y^{(2)}} - \bar{Y}_\cdot)^2 + I_3(\bar{Y}^{(3)} - \bar{Y}_\cdot)^2, \quad \bar{Y}_\cdot = \sum_{i=1}^{3} I_i\bar{Y}^{(i)}/I.$$

II.2.58. Gauname, kad statistikos F_A, F_B, F_C įgijo reikšmes 480,47, 903,91, 786,43; hipotezės atmetamos. Statistikos F_{AB} ir F_{AC} įgijo reikšmes 17,48 ir 17,91; hipotezės atmetamos. Statistikos F_{BC} ir F_{ABC} įgyja reikšmes 2,62 ir 4,72, kurias atitinka P reikšmės 0,0571 ir 0,0011.

II.2.59. Gauname statistikų realizacijas $F_A = 9,12$, $F_B = 5,72$, $F_C = 46,50$, $F_D = 0,54$, kurias atitinka P reikšmės $0,0002$, $0,0227$, $8,85 \times 10^{-8}$, $0,6560$. Gauname statistikų realizacijas $F_{AB} = 1,44$, $F_{AC} = 0,79$, $F_{AD} = 0,34$, kurias atitinka P reikšmės $0,2475$, $0,5070$, $0,9539$. Gauname statistikų realizacijas $F_{BC} = 10,32$, $F_{BD} = 0,82$, $F_{CD} = 1,82$, kurias atitinka P reikšmės $0,0029$, $0,4897$, $0,1627$.

II.2.4 skyrelis

II.2.60. Statistikos F_A ir $F_{B(A)}$ (žr. [3], 2.9 skyrelį) įgijo reikšmes $0,60$ ir 1,76, kurias atitinka P reikšmės $0,6700$ ir $0,0625$.

II.2.61. Statistikos F_A ir $F_{B(A)}$ (žr. [3], 2.9 skyrelį) įgijo reikšmes $3,39$ ir 4,45, kurias atitinka P reikšmės $0,0798$ ir $0,0016$.

II.2.62. Statistikos F_A ir $F_{B(A)}$ (žr. [3], 2.9 skyrelį) įgijo reikšmes $46,78$ ir 1,27, kurias atitinka P reikšmės $0,0002$ ir $0,3590$.

II.2.63. Hierarchinės klasifikacijos modelis. Faktoriaus B (pelė) lygmenys sugrupuoti pagal faktoriaus A (dieta) lygmenis. Abu faktoriai pastovūs. Statistikos F_A ir $F_{B(A)}$ (žr. [3], 2.9 skyrelį) įgijo reikšmes $28,67$ ir 1,08, kurias atitinka P reikšmės $0,0007$ ir $0,3838$.

II.2.64. Apie blokuotųjų duomenų analizę žr. [3], 2.10.2 skyrelį. Pagal turimus duomenis gauname statistikos F_A realizaciją 1301,7. Hipotezė H_A atmetama.

II.2.65. Gauname statistikų realizacijas $F_A = 157,92$, $F_B = 30,42$. Hipotezės H_A ir H_B atmestinos.

II.2.66. Apie duomenų analizę, kai eksperimentas atliktas pagal nepilnų subalansuotų blokų planą, žr. [3], 2.10.3 skyrelį. Pagal turimus duomenis gauname, kad gumos mišinio įtaka padangų atsparumui yra statistiškai reikšminga (P reikšmė 0,0034).

II.2.67. Gauname statistikų realizacijas $F_A = 0,17$ (operatorius) ir $F_B = 10,61$ (diena), kurias atitinka P reikšmės $0,9131$, $0,0132$.

II.2.68. Gauname statistikos realizaciją $27,24$. Hipotezė, kad skalbimo priemonės vienodos, atmetama kriterijumi su aukštu reikšmingumo lygmeniu.

II.2.69. Apie duomenų analizę, kai eksperimentas atliekamas pagal lotyniškųjų kvadratų schemą, žr. [3], 2.11 skyrelį. Pagal turimus duomenis gauname statistikos F_A realizaciją 15,83. Hipotezė, kad elektrodų tipai vienodi, atmetama kriterijumi su aukštu reikšmingumo lygmeniu.

II.2.70. Gauname statistikos F_A realizaciją 19,63. Hipotezė, kad veislių derlingumas vienodas, atmetama kriterijumi su aukštu reikšmingumo lygmeniu.

II.3. Regresinė analizė

II.3.1. Prognozavimo uždaviniai

II.3.1. Tarkime, reikia prognozuoti a. d. Y remiantis a. v. $\mathbf{X} = (X_1, ..., X_m)^T$. Prognoze vadiname mačiąją funkciją $\hat{Y} = h(\mathbf{X}), \boldsymbol{V}(h(\mathbf{X})) < \infty$, o skirtumą $Y - h(\mathbf{X})$

vadiname prognozės paklaida. Prognozė $\hat{Y} = h(\mathbf{X})$ vadinama optimaliąja, jeigu ji minimizuoja vidutinę kvadratinę paklaidą $\mathbf{E}(Y - h(\mathbf{X}))^2$. a) Įrodykite, kad a. d. Y regresija a. v. $\mathbf{X}$ atžvilgiu $\mu(\mathbf{X}) = \mathbf{E}(Y|\mathbf{X})$ yra optimalioji prognozė. b) Įrodykite, kad koreliacijos koeficientas tarp Y ir prognozės $\hat{Y} = h(\mathbf{X})$ yra maksimalus, kai $\hat{Y} = \mu(\mathbf{X})$.

II.3.2. (**II.3.1** pratimo tęsinys). Koreliacijos koeficiento $\rho(Y, \mu(\mathbf{X}))$ kvadratas vadinamas koreliaciniu santykiu ir žymimas

$$\eta^2_{Y\boldsymbol{X}} = \eta^2_{Y(X_1 \ldots X_m)} = \rho^2(Y, \mu(\mathbf{X})) = \boldsymbol{V}(\mu(\mathbf{X}))/\boldsymbol{V}(Y).$$

Pateikite koreliacinio santykio $\eta^2_{Y\boldsymbol{X}}$ interpretaciją.

II.3.3. (**II.3.1** pratimo tęsinys). Suskaidykime a. v. $\mathbf{X}$ į dvi komponentes $\mathbf{X}^{(1)} = (X_1, ..., X_k)^T$ ir $\mathbf{X}^{(2)} = (X_{k+1}, ..., X_m)^T$. Tada prognozės paklaidos dispersijos santykinis sumažėjimas papildžius vektorių $\mathbf{X}^{(1)}$ komponente $\mathbf{X}^{(2)}$ yra

$$\eta^2_{Y\boldsymbol{X}^{(2)}|\boldsymbol{X}^{(1)}} = \frac{\eta^2_{Y\boldsymbol{X}} - \eta^2_{Y\boldsymbol{X}^{(1)}}}{1 - \eta^2_{Y\boldsymbol{X}^{(1)}}}.$$

Dydis $\eta_{Y\boldsymbol{X}^{(2)}|\boldsymbol{X}^{(1)}} = \sqrt{\eta^2_{Y\boldsymbol{X}^{(2)}|\boldsymbol{X}^{(1)}}} \geq 0$ vadinamas daliniu koreliaciniu santykiu. Įrodykite lygybę

$$1 - \eta^2_{Y(X_1,...,X_m)} = \prod_{j=2}^{m}(1 - \eta^2_{YX_j|(X_1,...,X_{j-1})})(1 - \eta^2_{YX_1}).$$

II.3.4. Prognozuodami a. d. Y apsiribokime tiesinėmis a. v. $\mathbf{X} = (X_1, ..., X_m)^T$ funkcijomis $l(\mathbf{X}) = \alpha + \boldsymbol{\beta}^T\mathbf{X}$. Optimali tiesinė prognozė $l^* = \alpha^* + (\boldsymbol{\beta}^*)^T\mathbf{X}$ minimizuoja vidurkį

$$SS(\alpha, \boldsymbol{\beta}) = \mathbf{E}(Y - \alpha - \boldsymbol{\beta}^T\mathbf{X})^2.$$

a) Įrodykite, kad optimali tiesinė prognozė gaunama, kai

$$\boldsymbol{\beta}^* = \boldsymbol{\Sigma}^{-1}\sigma_{Y\boldsymbol{X}}, \quad \alpha^* = \mathbf{E}Y - (\boldsymbol{\beta}^*)^T\mathbf{E}(\mathbf{X}),$$

čia $\boldsymbol{\Sigma} = \boldsymbol{V}(\mathbf{X})$, o $\sigma_{Y\boldsymbol{X}} = \mathbf{Cov}\,(Y, \mathbf{X})$. b) Įrodykite, kad koreliacijos koeficientas tarp Y ir tiesinės prognozės $l = \alpha + \boldsymbol{\beta}^T\mathbf{X}$ yra maksimalus, kai $l = l^*$.

II.3.5. (**II.3.4** pratimo tęsinys). Pateikite dauginio koreliacinio koeficiento interpretaciją.

II.3.6. (**II.3.4** pratimo tęsinys). Suskaidykime a. v. $\mathbf{X}$ į dvi komponentes $\mathbf{X}^{(1)} = (X_1, ..., X_k)^T$ ir $\mathbf{X}^{(2)} = (X_{k+1}, ..., X_m)^T$. Tada optimalios tiesinės prognozės paklaidos dispersijos santykinis sumažėjimas papildžius vektorių $\mathbf{X}^{(1)}$ komponente $\mathbf{X}^{(2)}$ yra

$$\rho^2_{Y\boldsymbol{X}^{(2)}|\boldsymbol{X}^{(1)}} = \frac{\rho^2_{Y\boldsymbol{X}} - \rho^2_{Y\boldsymbol{X}^{(1)}}}{1 - \rho^2_{Y\boldsymbol{X}^{(1)}}}.$$

Dydis $\rho_{Y\boldsymbol{X}^{(2)}|\boldsymbol{X}^{(1)}} = \sqrt{\rho^2_{Y\boldsymbol{X}^{(2)}|\boldsymbol{X}^{(1)}}} \geq 0$ vadinamas daliniu dauginiu koreliacijos koeficientu. Įrodykite, kad

$$1 - \rho^2_{Y(X_1,...,X_m)} = \prod_{j=2}^{m}(1 - \rho^2_{YX_j|(X_1,...,X_{j-1})})(1 - \rho^2_{YX_1}).$$

II.3.7. Vektorius $(X,\ Y)^T$ pasiskirstęs pagal dvimatį normalųjį skirstinį su nuliniais vidurkiais, vienetinėmis dispersijomis ir koreliacijos koeficientu ρ. Raskite regresijos kreives: a) a. d. X atžvilgiu Y; b) a. d. X atžvilgiu Y^2; c) a. d. X^2 atžvilgiu Y^2; d) a. d. X^2 atžvilgiu Y.

Apskaičiuokite koreliacinius santykius ir palyginkite juos su koreliacijos koeficientų kvadratais.

II.3.8. Atsitiktinis vektorius $(X,\ Y)^T$ turi tolygųjį skirstinį viduje elipsės

$$ax^2 + 2hxy + by^2 = c, \quad h \neq 0, \quad h^2 < a\,b; \quad a,\ b > 0.$$

Įrodykite, kad a. d. X ir Y regresija kito kintamojo atžvilgiu yra tiesinė.

II.3.9. Atsitiktinis vektorius $(X,\ Y)^T$ tolygiai pasiskirstęs lygiagretainyje, apribotame tiesėmis $x = 3(y+1)$, $x = 3(y-1)$, $x = y+1$, $x = y-1$. Įrodykite, kad atsitiktinio dydžio Y regresija X atžvilgiu yra tiesinė, o X regresija Y atžvilgiu yra kreivė, susidedanti iš trijų atkarpų.

II.3.10. Atsitiktinis vektorius $(X, Y)^T$ tolygiai pasiskirstęs ant pusapskritimio $y = +\sqrt{1-x^2}$, $0 \leq x \leq 1$. Įrodykite, kad $\eta^2_{YX} = 1$, o $\rho^2_{YX} = 0$.

II.3.11. Atsitiktinis vektorius $(X, Y)^T$ tolygiai pasiskirstęs skritulyje $(x-a)^2 + (y-b)^2 \leq r^2$. Įrodykite, kad $\eta^2_{YX} = \rho^2_{YX} = 0$.

II.3.12. Tegu $\rho = [\rho_{ij}]_{k\times k}$ yra koreliacijos koeficientų matrica. Įrodykite, kad dauginis koreliacijos koeficientas ir daliniai koreliacijos koeficientai susieti lygybėmis:

$$1 - \rho^2_{1.(2...k)} = (1-\rho^2_{12})(1-\rho^2_{13.2})...(1-\rho^2_{1k.2...k-1});$$

$$\rho_{12.3} = \frac{(\rho_{12} - \rho_{13}\rho_{23})}{(1-\rho^2_{13})^{1/2}(1-\rho^2_{23})^{1/2}};$$

$$\rho_{12} = \frac{(\rho_{12.3} - \rho_{13.2}\rho_{23.1})}{(1-\rho^2_{13.2})^{1/2}(1-\rho^2_{23.1})^{1/2}};$$

$$\rho_{12.(3...k-1)} = \frac{(\rho_{12.(3...k)} - \rho_{1k.(2...k-1)}\rho_{2k.(13...k-1)})}{(1-\rho^2_{1k.(2...k-1)})^{1/2}(1-\rho^2_{2k.(13...k-1)})^{1/2}};$$

$$\rho_{12.(3...k)} = \frac{(\rho_{12.(3...k-1)} - \rho_{1k.(3...k-1)}\rho_{2k.(3...k-1)})}{(1-\rho^2_{1k.(2...k-1)})^{1/2}(1-\rho^2_{2k.(3...k-1)})^{1/2}}.$$

II.3.2. Tiesinė vieno kintamojo regresija

II.3.13. Prognozuojamas a. d. Y remiantis vienmate kovariante X. Fiksavus kovariantės reikšmes $x_1, ..., x_n$ gauta imtis $Y_1, ..., Y_n$. Tarkime, kad Y sąlyginis vidurkis yra x_i tiesinė funkcija, o dispersija nuo fiksuotosios kovariantės reikšmės nepriklauso. Imties elementai turi tokią struktūrą

$$Y_i = \alpha + \beta(x_i - \bar{x}) + e_i, \quad i = 1, ..., n, \quad \bar{x} = \sum_i x_i/n.$$

Turime tiesinį modelį (žr. **II.1.1** pratimą)

$$\boldsymbol{Y} = \boldsymbol{A\beta} + \boldsymbol{e}, \quad \mathbf{E}(\boldsymbol{Y}) = \boldsymbol{A\beta}, \quad \mathbf{E}(\boldsymbol{e}) = \mathbf{0}, \quad \boldsymbol{V}(\boldsymbol{e}) = \sigma^2\boldsymbol{I},$$

čia $\boldsymbol{\beta} = (\alpha, \beta)^T$, $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$, $\boldsymbol{e} = (e_1, ..., e_n)^T$; plano matrica $\boldsymbol{A}$ turi n eilučių ir du stulpelius, i-oji eilutė yra $(1, x_i - \bar{x})$. a) Raskite parametrų α, β mažiausiųjų kvadratų įvertinius ir jų pirmuosius momentus. b) Raskite nepaslinktąjį dispersijos σ^2 įvertinį.

II.3.14. (**II.3.13** pratimo tęsinys). Tarkime, kad paklaidų vektorius $\boldsymbol{e}$ turi normalųjį skirstinį $\boldsymbol{e} \sim N_n(\mathbf{0}, \sigma^2\boldsymbol{I})$. a) Įrodykite, kad **II.3.13** pratime surasti įvertiniai yra NMD įvertiniai. b) Raskite parametrų $\alpha, \beta, \theta = c_1\alpha + c_2\beta, \sigma^2$ įvertinių skirstinius.

II.3.15. (**II.3.14** pratimo tęsinys). a) Raskite parametrų σ^2, α, β, $\mu(x) = \alpha + \beta(x - \bar{x})$ lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalus. b) Raskite kriterijus hipotezėms dėl šių parametrų reikšmių tikrinti.

II.3.16. Tarkime, kad turime n_i a. d. Y stebėjimų, kai kovariantės reikšmė yra $x_i, i = 1, ..., k$. Tegu paklaidų vektorius $\boldsymbol{e} \sim N_n(\mathbf{0}, \sigma^2\boldsymbol{I}), n = n_1 + ... + n_k$. Raskite kriterijų regresijos tiesiškumo hipotezei $H : \mathbf{E}(Y|X = x) = \beta_0 + \beta_1(x - \bar{x})$ tikrinti.

II.3.17. (**II.3.16** pratimo tęsinys). Raskite kriterijų hipotezei $H : \mathbf{E}(Y|X = x) = \alpha + \beta_1(x - \bar{x}) + \beta_2(x - \bar{x})^2 + ... + \beta_r(x - \bar{x})^r$, $r < k$, tikrinti.

II.3.18. (**II.3.14** pratimo tęsinys). Tarkime, pagal **II.3.14** pratimo duomenis įvertinta regresija $\hat{\mu}(x) = \hat{\alpha} + \hat{\beta}(x - \bar{x})$. Tegu gautas tolimesnis nepriklausomas stebėjimas Y_{n+1}, kuris gaunamas, kai kovariantės X reikšmė yra x. Raskite lygmens $Q = 1 - \alpha$ a. d. Y_{n+1} prognozės intervalą.

II.3.19. (**II.3.18** pratimo tęsinys). Tarkime, kad stebėjimas Y_{n+1} žinomas. Raskite kovariantės reikšmės x, kurią atitinka Y_{n+1}, pasikliovimo intervalą.

II.3.20. Pagal didumo n imtį $Y_i = \alpha + \beta(X_i - \bar{X}) + e_i$, $i = 1, ..., n$, kai a. d. $\{e_i\}$ nepriklausomi ir $e_i \sim N(0, \sigma^2)$ gauti įvertiniai $\hat{\alpha}$ ir $\hat{\beta}$. Sukurkite TGN kriterijų hipotezei $H : \theta = a\alpha + b\beta = \theta_0$, kai alternatyva $\bar{H} : \theta \neq \theta_0$, tikrinti; čia a, b, θ_0 – žinomi.

II.3.21. Pagal dvi nepriklausomas didumo n_1 ir n_2 imtis įvertinti tiesinės vieno kintamojo regresijos koeficientai $\hat{\alpha}_1$, $\hat{\beta}_1$ ir $\hat{\alpha}_2$, $\hat{\beta}_2$. Tarkime, kad visi stebėjimai yra nepriklausomi ir normalieji, turi vienodas dispersijas σ^2. Raskite kriterijus: a) hipotezei $H : \beta_1 = \beta_2$ tikrinti; b) hipotezei $H : \alpha_1 = \alpha_2$ tikrinti; c) hipotezei $H : \alpha_1 = \alpha_2$, $\beta_1 = \beta_2$ tikrinti; d) hipotezei, kad regresijos tiesės susikerta taške, kurio abscisė $x = c$, tikrinti.

II.3.22. Raskite parametrų α, β ir σ^2 įverčius naudodami šiuos stebėjimus:

$$0,15 = \alpha - 3\beta + e_1,$$

$$2,07 = \alpha - \beta + e_2,$$

$$4,31 = \alpha + \beta + e_3,$$

$$6,49 = \alpha + 3\beta + e_4,$$

čia $e_1, ..., e_4$ – nepriklausomi vienodai pasiskirstę atsitiktiniai dydžiai $e_i \sim N(0, \sigma^2)$, $i = 1, ..., 4$.

II.3.23. Taškas tolygiai juda tiese. Laiko momentais $t = 0,\ 1,\ 2,\ 3,\ 4$ buvo užfiksuotos tokios jo koordinačių reikšmės: $S_t = 12,98;\ 13,05;\ 13,35;\ 13,97;\ 14,22$. Tegu visų matavimų paklaidos nepriklausomos ir turi normalųjį skirstinį $N(0,\ \sigma^2)$. Raskite greičio v ir dispersijos σ^2 taškinius ir intervalinius įverčius, kai pasikliovimo lygmuo $Q = 0,95$.

II.3.24. Matuojant gauti stebėjimai, kurie yra tiesinės parametro α funkcijos:

$$Y_1 = e_1,$$

$$Y_2 = \alpha + e_1 + e_2,$$

$$Y_3 = 2\alpha + e_1 + e_2 + e_3,$$

$$Y_4 = 3\alpha + e_1 + e_2 + e_3 + e_4,$$

$$Y_5 = 4\alpha + e_1 + e_2 + e_3 + e_4 + e_5,$$

čia $e_1, ..., e_5$ – nepriklausomi vienodai pasiskirstę atsitiktiniai dydžiai $e_i \sim N(0, \sigma^2)$, $i = 1, ..., 5$. Raskite parametro α NMD įvertį, jeigu atsitiktinio vektoriaus $\boldsymbol{Y} = (Y_1, ..., Y_5)^T$ realizacija yra $(-0,42,\ 0,30,\ 1,30,\ 2,56,\ 3,26)^T$. Sudarykite parametrų α ir σ^2 pasikliovimo intervalus, kai pasikliovimo lygmuo $Q = 0,95$.

II.3.25. Atsitiktinis vektorius $\boldsymbol{Y} = (Y_1, ..., Y_5)^T$ turi normalųjį skirstinį su parametrais $\mathbf{E}Y_i = (i-1)\alpha,\ i = 1, ..., 5$; $\boldsymbol{V}Y_i = \mathbf{Cov}\,(Y_i,\ Y_j) = i,\ j > i,\ i = 1, ..., 5$. Raskite parametro α pasikliovimo intervalą, kai pasikliovimo lygmuo $Q = 0,95$, pagal atsitiktinio vektoriaus $\boldsymbol{Y}$ realizaciją $\boldsymbol{y} = (-0,42;\ 0,30;\ 1,30;\ 2,56;\ 3,26)^T$.

II.3.26. Išmatuotas grunto atsparumas poslinkiui Y, veikiant įvairiems slėgiams P kg/cm^2, ir gauti šitokie rezultatai:

P_i	Y_{i1}	Y_{i2}	Y_{i3}	Y_{i4}	Y_{i5}	Y_{i6}	Y_{i7}	Y_{i8}	Y_{i9}
1	0,962	0,612	0,612	0,888	1,038	1,188	0,812	0,988	0,875
2	0,962	0,688	0,662	1,112	1,118	1,325	0,988	1,025	1,075
3	1,125	0,800	0,700	0,900	1,475	1,400	1,150	1,175	1,300
4	1,412	0,850	0,650	1,075	1,312	1,425	1,250	1,175	1,162
5	1,125	0,900	0,675	1,225	1,225	1,650	1,025	1,075	1,500
6	1,138	0,950	0,625	1,500	1,375	1,588	1,025	1,150	1,100

a) patikrinkite Y regresijos P atžvilgiu tiesiškumo hipotezę;
b) laikydami Y regresiją P atžvilgiu tiesine, įvertinkite jos koeficientus;
c) sudarykite regresijos koeficientų pasikliovimo intervalus ($Q = 0,95$).

II.3.27. Sąlyginis Y skirstinys, kai X reikšmės fiksuotos, yra normalusis su dispersija $\sigma^2 = 0,1$. Tarkime, kad taškuose $x = 0,0;\ 0,1; ...; 1,0$ turimos vienodo didumo n nepriklausomos imtys. Koks turi būti n, kad regresijos tiesiškumo hipotezė $H : \mathbf{E}(Y|X = x) = x, x = 0,0,\ 0,1,\ ...\ 1,0$ būtų atmesta su tikimybe, ne mažesne už 0,95, jeigu $\mathbf{E}(Y|X = x) = x^2$ (kriterijaus reikšmingumo lygmuo $P = 0,05$).

II.3.28. Tarkime, kad pagal nepriklausomus stebėjimus $Y_i = \alpha + \beta(x_i - \bar{x}) + e_i$, $i = 1, ..., n$, kai a. d. $\{e_i\}$ nepriklausomi ir normalūs $e_i \sim N(0, \sigma^2)$, įvertinta regresijos tiesė $\hat{\alpha} + \hat{\beta}(x - \bar{x})$. Tegu taške x yra gauta k nepriklausomų Y_{n+1} matavimų $Y_{n+1}^{(1)}, ..., Y_{n+1}^{(k)}$. Sukonstruokite stebėjimo Y_{n+1} prognozės intervalą ir argumento x pasikliovimo intervalą.

II.3.29. Pasiūlykite, kaip modelį

$$y = \frac{\alpha\beta}{\alpha \sin^2 x + \beta \cos^2 x}$$

pertvarkyti į tiesinį pavidalą.

II.3.30. Kalibruojamas pieno rūgšties koncentracijos kraujuje matavimo prietaisas. Tuo tikslu gauti $n = 20$ pavyzdžių matavimai $Y_i, i = 1, 2, ..., 20$, kai koncentracija X tiksliai žinoma [1].

X	1	1	1	1	3	3	3	3	3	5
Y	1,1	0,7	1,8	0,4	3,0	4,4	4,9	4,4	4,5	7,3
X	5	5	10	10	10	10	15	15	15	15
Y	8,2	6,2	12,0	13,1	12,6	13,2	18,7	19,7	17,4	17,1

Tardami, kad paklaidos normaliosios, įvertinkite tiesinės regresijos parametrus. Patikrinkite regresijos tiesiškumo hipotezę. Raskite tolesnių nepriklausomų Y matavimų prognozės intervalus, kai pasikliovimo lygmuo $Q = 0,95$, jei žinoma, kad matavimai bus atliekami taškuose $X = 12$, $X = 18$.

II.3.31. Pamatavus $n = 108$ pacientų arterinį PH (kintamasis Y) ir veninį PH (kintamasis X) gauti vidurkių ir kovariacijų matricos elementų įverčiai: $\bar{X} = 7,373$, $\bar{Y} = 7,413$, $s_X^2 = 0,1253, s_Y^2 = 0,1184, s_{XY} = 0,1101$ [1]. Priėmę prielaidą dėl a. v. $(X, Y)^T$ normalumo, įvertinkite a. d. Y regresijos tiesę a. d. X atžvilgiu. Patikrinkite hipotezę $H : \beta = 0$. Kokią dalį a. d. Y sklaidos paaiškina regresijos lygtis?

II.3.32. Kintamasis Y reiškia tam tikro cheminio proceso savybę priklausomai nuo laiko t [1]:

t	0,0	1,0	1,5	2,0	2,5	3,0	3,5	4,0
Y	0,000	0,025	0,035	0,045	0,055	0,065	0,075	0,082
t	4,5	5,0	5,5	6,0	6,5	7,0	7,5	8,0
Y	0,088	0,094	0,100	0,105	0,110	0,115	0,120	0,125

a) Įvertinkite tiesinės regresijos parametrus. Panagrinėkite prognozės paklaidas ir aptarkite regresijos tiesinio pavidalo priimtinumą. b) Tarkime, kad iš teorinių samprotavimų gauta, jog regresijos kreivė turėtų būti tokio pavidalo: $Y = \alpha(1 - e^{-\beta t})$. Įvertinkite netiesinio modelio parametrus ir raskite jų apytikslius pasikliovimo intervalus.

II.3.33. Farmakinetikoje vertinant vaisto koncentracijos Y priklausomybę nuo laiko t dažnai naudojamas dvikomponentis matematinis modelis $Y = \alpha_1 e^{\alpha_2 t} + \beta_1 e^{\beta_2 t}$. Pagal pateikiamus duomenis [1]

t	0,10	0,25	0,50	1,00	1,50	2,00	4,00	8,00	12,00	24,00	48,00
Y	18,7	16,9	14,5	11,1	8,9	7,5	5,2	3,6	2,6	1,0	0,2

įvertinkite regresijos parametrus. Nagrinėdami prognozės paklaidas aptarkite prognozės tikslumą.

II.3.3. Tiesinė keleto kintamųjų regresija

II.3.34. Tarkime, prognozuojame a. d. Y remdamiesi kovariantėmis $X_1, ..., X_m$. Tegu $\mathbf{X} = (X_0, X_1, ..., X_m)^T$ yra kovariančių vektorius, papildytas koordinate $X_0 \equiv 1$. Fiksavus kovariantės reikšmes $\boldsymbol{x}^{(i)} = (1, x_{1i}, ..., x_{mi})^T$ gauti nepriklausomi stebėjimai $Y_i, i = 1, ..., n$. Tarsime, kad imties $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ elementai turi tokią struktūrą

$$Y_i = \beta_0 + \beta_1 x_{1i} + ... + \beta_m x_{mi} + e_i, \quad i = 1, ..., n.$$

Tegu paklaidos e_i nepriklausomos ir normaliosios $e_i \sim N(0, \sigma^2)$. Pažymėję $\boldsymbol{\beta} = (\beta_1, ..., \beta_m)^T$ ir $\boldsymbol{e} = (e_1, ..., e_n)^T$, gauname tiesinį modelį

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{\beta} + \boldsymbol{e}, \quad \mathbf{E}(\boldsymbol{Y}) = \boldsymbol{A}\boldsymbol{\beta}, \quad \mathbf{E}(\boldsymbol{e}) = \mathbf{0}, \quad \boldsymbol{V}(\boldsymbol{e}) = \sigma^2 \boldsymbol{I},$$

kai plano matricos i-oji eilutė yra $(1, x_{1i}, ..., x_{mi})$.

a) Raskite parametro $\boldsymbol{\theta} = (\beta_0, \beta_i, ..., \beta_m, \sigma^2)^T$ pilnąją ir pakankamąją statistiką.

b) Raskite parametro $\boldsymbol{\theta}$ elementų NMD įvertinius.

II.3.35. (**II.3.34** pratimo tęsinys). Raskite parametrų $\beta_i, \theta = \boldsymbol{L}^T\boldsymbol{\beta}, \sigma^2$ lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalus.

II.3.36. (**II.3.34** pratimo tęsinys). Raskite kriterijus hipotezėms $H_{j_1,...,j_k} : \beta_{j_1} = ... = \beta_{j_k} = 0$ tikrinti. Jei ši hipotezė teisinga, tai kovariantės $X_{j_1}, ..., X_{j_k}$ nėra reikšmingos prognozuojant Y.

II.3.37. (**II.3.34** pratimo tęsinys). Tarkime, kad $\boldsymbol{x}^{(i)} = (1, x_{1i}, ..., x_{mi})^T$, $i = 1, ..., k$, yra skirtingos kovariantės reikšmės ir, kai yra reikšmė $\boldsymbol{x}^{(i)}$, gauta n_i nepriklausomų Y stebėjimų $Y_{ij}, j = 1, ..., n_i$, $n = n_1 + ... + n_k$. Raskite kriterijų hipotezei $H : \mathbf{E}(Y|\mathbf{X} = \boldsymbol{x}) = \boldsymbol{\beta}^T\boldsymbol{x}$ apie regresijos tiesinį pavidalą tikrinti.

II.3.38. (**II.3.34** pratimo tęsinys). Tarkime, pagal **II.3.34** pratimo duomenis įvertinta regresija

$$\hat{\mu}(\boldsymbol{x}) = \hat{\beta}_0 + \hat{\beta}_1 x_1 + ... + \hat{\beta}_m x_m$$

Raskite tolesnio nepriklausomo stebėjimo Y_{n+1} prognozės intervalą, jei žinoma, kad jis bus atliktas taške $\boldsymbol{x} = (1, x_1, ..., x_m)$.

II.3.39. Noradrenalino kiekio Y priklausomybė nuo gaunamo su maistu natrio kiekio X yra arba polinominio, arba eksponentinio tipo. Pagal pateikiamus duomenis

X	2	10	5	6	20	3	14	21
Y	31,66	19,20	45,03	13,30	23,44	20,61	18,46	11,98
X	103	122	136	80	196	196	224	245
Y	5,58	15,21	7,58	9,77	13,60	10,01	3,68	7,03
X	97	86	56	127	171	257	157	
Y	13,90	14,00	14,61	15,26	14,12	7,30	10,32	

a) parinkite polinominę regresiją ir aptarkite prognozės tikslumą; b) parinkite eksponentinio tipo priklausomybę, įvertinkite parametrus ir aptarkite prognozės tikslumą.

II.3.40. Tarkime, kad įvertinome regresiją $\mathbf{E}(Y) = \beta_0 + \beta_1 x$, o tikroji regresijos lygtis yra $\mathbf{E}(Y) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$ pavidalo. Koks bus įvertinių $\hat{\beta}_0$ ir $\hat{\beta}_1$ poslinkis, kai jie įvertinti pagal nepriklausomus stebėjimus taškuose $x = -3, -2, -1, 0, +1, +2, +3$?

II.3.41. Tegu $Y_1 = \theta_1 + \theta_2 + e_1$, $Y_2 = 2\theta_2 + e_2$, $Y_3 = -\theta_1 + \theta_2 + e_3$; čia $\{e_i\}$ nepriklausomi ir $e_i \sim N(0, \sigma^2)$. Sudarykite kriterijų hipotezei $H : \theta_1 = 2\theta_2$ tikrinti.

II.3.42. Tegu Y_1, Y_2, Y_3, Y_4 yra keturkampio kampų $\varphi_1, \varphi_2, \varphi_3, \varphi_4$ matavimai. Tarkime, kad matavimo paklaidos yra normalieji nepriklausomi a. d., turintys nulinius vidurkius ir vienodas dispersijas. Patikrinkite hipotezę, kad keturkampis yra lygiagretainis, t. y. $H : \varphi_1 = \varphi_3, \varphi_2 = \varphi_4$.

II.3.43. Tarkime, $\mathbf{E}(Y_t) = \beta_0 + \beta_1 \cos(2\pi k_1 t/n) + \beta_2 \sin(2\pi k_2 t/n)$; čia $t = 1, ..., n$, o k_1 ir k_2 – žinomos teigiamos konstantos. Raskite parametrų $\beta_0, \beta_1, \beta_2$ mažiausiųjų kvadratų įvertinius.

II.3.44. Nagrinėjamas benzino oktaninis skaičius Y priklausomai nuo priedų A ir B koncentracijų X_1 ir X_2. Pateikiami tokie duomenys [1]

X_1	2	2	2	2	3	3	3	3
X_2	2	3	4	5	2	3	4	5
Y	96,3	95,7	99,9	99,4	95,1	97,8	99,3	104,9
X_1	4	4	4	4	5	5	5	5
X_2	2	3	4	5	2	3	4	5
Y	96,2	100,1	103,2	104,3	97,8	102,2	104,7	108,8

a) Įvertinkite regresijos parametrus tardami, kad a. d. Y regresija X_1 ir X_2 atžvilgiu yra tiesinė. b) Tarę, kad paklaidos normaliosios, patikrinkite regresijos koeficientų lygybės 0 hipotezes. c) Raskite a. d. Y tolesnio nepriklausomo stebėjimo, kuris atliktas taške $(X_1, X_2) = (4,5,\ 3,5)$, prognozės intervalą, kai pasikliovimo lygmuo $Q = 0,95$.

II.3.45. Dviem skirtingais būdais buvo matuojamas 141 paciento arterinis kraujo spaudimas. Pirmu būdu matuojant sistolinį X_1, diastolinį X_2 ir vidutinį X_3 kraujo spaudimą buvo naudojamas kateteris. Antru būdu buvo matuojamas sistolinis X_4 ir diastolinis X_5 kraujo spaudimas naudojant kompresinę manžetę. Atlikus pradinę analizę gauti šie rezultatai.

i	$\bar{X}_i$	s_i	r_{ij}				
1	112,2	28,6	1,000	0,839	0,927	0,871	0,753
2	59,4	17,1		1,000	0,967	0,778	0,828
3	76,8	21,0			1,000	0,845	0,852
4	107,0	28,9				1,000	0,837
5	66,8	19,3					1,000

Šioje lentelėje $\bar{X}_i$ – aritmetiniai vidurkiai, o s_i – vidutinių kvadratinių nuokrypių įvertiniai, $i = 1, ..., 5$; r_{ij} – koreliacijos koeficientų įvertiniai.

Naudodami pažingsninės regresijos metodą, parinkite vektoriaus $(X_1,\ X_2,\ X_3)^T$ koordinates kintamiesiems X_4 ir X_5 prognozuoti ($P' = 0,01,\ P = 0,05$). Raskite regresijos koeficientus ir dauginius koreliacijos koeficientus.

II.3.46. Lentelėje pateikti Velso gyventojų surašymo duomenys Y (mln.)

Metai	Y	Metai	Y	Metai	Y
1811	10,16	1861	20,07	1901	32,53
1821	12,00	1871	22,71	1911	36,07
1831	13,90	1881	25,97	1921	37,89
1841	15,91	1891	29,00	1931	39,95
1851	17,93				

a) Parinkite polinominį regresijos modelį.

b) Apskaičiuokite liekamosios kvadratinės formos sumažėjimą didindami polinomo laipsnį. Kokio mažiausiojo laipsnio polinomas priimtinas?

II.3.47. Reikia parinkti tokio pavidalo regresijos modelį:

$$\mathbf{E}(Y_i) = \beta_0 + \beta_1 X_i + \beta_2 \varphi(X_i), \quad i = 1, 2, 3;$$

čia $\varphi(x)$ – antrojo laipsnio polinomas. Parinkite $\varphi(x)$ taip, kad plano matrica turėtų ortogonalius stulpelius, kai $X_1 = -1, X_2 = 0, X_3 = +1$.

II.3.4. Kovariacinė analizė

II.3.48. Atliekant dispersinę analizę, kai tiriama a. d. Y skirstinio priklausomybė nuo tam tikrų faktorių ar jų sąveikų, Y skirstinys gali priklausyti ir nuo tam tikrų tolydžių kovariančių, kurios gali iškreipti išvadas. Jeigu kartu su a. d. Y stebėjimais galima gauti ir tokių kovariančių stebėjimus, tai atliekant analizę jų įtaką galima eliminuoti. Kovariacinėje analizėje imties $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ vidurkis užrašomas dviejų dėmenų suma (žr. [3], 4.1–4.3 skyrelius)

$$\mathbf{E}(\boldsymbol{Y}) = \boldsymbol{A\beta} + \boldsymbol{C\gamma}.$$

Pirmasis dėmuo apibūdina kokį nors dispersinės analizės modelį, antrasis aprašo tiesinę Y regresiją trukdančiųjų kovariančių atžvilgiu. Rango m dimensijos $n \times m$ matrica $\boldsymbol{A}$ yra dispersinės analizės plano matrica; rango k dimensijos $n \times k$ matricos $\boldsymbol{C}$ i-asis stulpelis susideda iš i-osios kovariantės matavimų. Imties $\boldsymbol{Y}$ elementai turi tokią struktūrą

$$Y_i = a_{i1}\beta_1 + ... + a_{im}\beta_m + \gamma_1 x_{1i} + ... + \gamma_k x_{ki} + e_i, \quad i = 1, ..., n.$$

Įrodykite, kad parametrų mažiausiųjų kvadratų įvertiniai ir liekamoji kvadratų suma gali būti surasta tokiu būdu.

Tegu $\hat{\boldsymbol{\beta}}_0$ ir $\hat{\boldsymbol{\beta}}_i, i = 1, ..., k$, yra lygčių sistemų

$$\boldsymbol{A}^T\boldsymbol{A}\hat{\boldsymbol{\beta}}_0 = \boldsymbol{A}^T\boldsymbol{Y}, \quad \boldsymbol{A}^T\boldsymbol{A}\hat{\boldsymbol{\beta}}_i = \boldsymbol{A}^T\boldsymbol{C}_i, \quad i = 1, .., k,$$

sprendiniai; čia $\boldsymbol{C}_i$ yra matricos $\boldsymbol{C}$ i-asis stulpelis. Tegu

$$R_{yy} = (\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}}_0)^T(\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}}_0) = \boldsymbol{Y}^T\boldsymbol{Y} - \hat{\boldsymbol{\beta}}_0^T\boldsymbol{A}^T\boldsymbol{Y},$$

$$R_{yx_i} = (\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}}_0)^T(\boldsymbol{C}_i - \boldsymbol{A}\hat{\boldsymbol{\beta}}_i), \quad i = 1, ..., k,$$

$$R_{x_ix_j} = (\boldsymbol{C}_i - \boldsymbol{A}\hat{\boldsymbol{\beta}}_i)^T(\boldsymbol{C}_j - \boldsymbol{A}\hat{\boldsymbol{\beta}}_j), \quad i, j = 1, ..., k.$$

Tarkime, $\hat{\boldsymbol{\gamma}} = (\hat{\gamma}_1, ..., \hat{\gamma}_k)$ yra lygčių sistemos

$$R_{x_1x_i}\hat{\gamma}_1 + ... + R_{x_kx_i}\hat{\gamma}_k = R_{yx_i}, \quad i = 1, ..., k,$$

sprendinys ir $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_0 - \hat{\gamma}_1\hat{\boldsymbol{\beta}}_1 - ... - \hat{\gamma}_k\hat{\boldsymbol{\beta}}_k$. Tada $(\hat{\boldsymbol{\gamma}}^T, \hat{\boldsymbol{\beta}}^T)^T$ yra mažiausiųjų kvadratų įvertinys, o liekamoji kvadratų suma

$$SS_E = R_{yy} - \hat{\gamma}_1 R_{yx_1} - ... - \hat{\gamma}_k R_{yx_k} \sim \sigma^2\chi^2_{n-m-k}.$$

II.3.49. (**II.3.48** pratimo tęsinys). a) Raskite kriterijų hipotezei $H_{\boldsymbol{\gamma}} : \boldsymbol{\gamma} = \mathbf{0}$, kad kovariantės $X_1, ..., X_k$ nedaro įtakos Y skirstiniui, tikrinti. b) Raskite kriterijų hipotezei $H_{i_1,...,i_l} : \gamma_{i_1} = ... = \gamma_{i_l} = 0$, kad kovariantės $X_{i_1}, ..., X_{i_l}$ nedaro įtakos Y skirstiniui, tikrinti.

II.3.50. (**II.3.48** pratimo tęsinys). Raskite kriterijų kuriai nors dispersinės analizės hipotezei tikrinti eliminuodami kovariančių $X_1, ..., X_k$ įtaką. Tarkime, tikrinamą hipotezę galima suformuluoti tokiu būdu $H : \beta_{i_1} = ... = \beta_{i_s}$.

II.3.51. Pakartokite **II.3.48** – **II.3.50** pratimus imdami vienfaktorį vieno kintamojo kovariacinės analizės modelį.

II.3.52. Tegu $Y_{ij} = \mu_i + \gamma_1 X_{ij}^{(1)} + \gamma_2 X_{ij}^{(2)} + e_{ij}$, $i = 1, ..., I, j = 1, ..., J$, o a. d. $\{e_{ij}\}$ nepriklausomi ir normalieji $e_{ij} \sim N(0, \sigma^2)$.

a) Raskite parametrų γ_1 ir γ_2 mažiausiųjų kvadratų įvertinius $\hat{\gamma}_1$ ir $\hat{\gamma}_2$.

b) Raskite įvertinių $\hat{\gamma}_1$ ir $\hat{\gamma}_2$ kovariacijų matricą. Kokiomis sąlygomis įvertiniai $\hat{\gamma}_1$ ir $\hat{\gamma}_2$ nekoreliuoti?

II.3.53. Tegu $Y_{ijk} = \mu_{ij} + \gamma_{ij} X_{ijk} + e_{ijk}$, $i = 1, ..., I, j = 1, ..., J, k = 1, ..., K$, o a. d. $\{e_{ijk}\}$ nepriklausomi ir normalieji $e_{ijk} \sim N(0, \sigma^2)$. a) Raskite kriterijų hipotezei $H : \gamma_{ij} = \gamma$ tikrinti. b) Tardami, kad hipotezė H teisinga, sudarykite parametro γ pasikliovimo intervalą.

II.3.54. Tegu $Y_{ijk} = \mu + \alpha_i + \beta_{ij} + \gamma X_{ijk} + e_{ijk}$, $i = 1, ..., I, j = 1, ..., J, k = 1, ..., K$, o a. d. $\{e_{ijk}\}$ nepriklausomi ir normalieji $e_{ijk} \sim N(0, \sigma^2)$; $\sum_i \alpha_i = 0$, $\sum_j \beta_{ij} = 0, i = 1, ..., I$. a) Raskite kriterijų hipotezei $H : \gamma = 0$ tikrinti. b) Raskite kriterijų hipotezei $H_A : \alpha_i = 0, i = 1, ..., I$, tikrinti.

II.3.55. Tegu $Y_{ij} = \mu_i + \gamma_i X_j + e_{ij}$, $i = 1, 2, j = 1, ..., J$, o a. d. $\{e_{ij}\}$ nepriklausomi ir normalieji $e_{ij} \sim N(0, \sigma^2)$. Remdamiesi kovariacine analize raskite kriterijų hipotezei $H : \gamma_1 = \gamma_2$ tikrinti. Įsitikinkite, kad gautasis kriterijus yra ekvivalentus dviejų regresijos tiesių lygiagretumo kriterijui.

II.3.56. Lentelėje pateiktos bandelių, iškeptų iš 100 (g) tešlos, apimtis Y priklausomai nuo 17 miltų rūšių ir nuo juose esančio bromistinio kalio kiekio X (g), kai $X = 0, 1, 2, 3, 4$ (žr. [16]).

X	0	1	2	3	4	X	0	1	2	3	4
A_1	950	1075	1055	975	880	A_{10}	885	1000	1015	960	895
A_2	890	980	955	865	825	A_{11}	895	935	965	950	920
A_3	830	850	820	770	735	A_{12}	685	835	870	875	880
A_4	770	815	765	725	700	A_{13}	615	665	650	680	660
A_5	860	1040	1065	975	945	A_{14}	885	910	890	835	785
A_6	835	960	985	915	845	A_{15}	985	1075	1070	1015	1005
A_7	795	900	905	880	785	A_{16}	710	750	740	725	720
A_8	800	860	870	850	850	A_{17}	785	845	865	825	820
A_9	750	940	1000	960	960						

a) Atlikite vienfaktorę dispersinę analizę neatsižvelgdami į kintamąjį X.

b) Atlikite kovariacinę analizę, tarę, kad X yra kiekybinis kintamasis, ir pasirinkdami pirmojo, antrojo ir trečiojo laipsnio polinomus kintamojo X atžvilgiu.

II.3.57. Lentelėje pateikti duomenys, gauti atlikus eksperimentą pagal keturfaktorės dispersinės analizės pilną kryžminės klasifikacijos schemą. Registruojamas tam tikro maisto produkto drėgnumas priklausomai nuo druskos rūšies (faktorius A), druskos kiekio (faktorius B), rūgšties lygio (faktorius C) ir dviejų skirtingų priemaišų (faktorius D) [16].

		A_1			A_2			A_3		
		B_1	B_2	B_3	B_1	B_2	B_3	B_1	B_2	B_3
C_1	D_1	8	17	22	7	26	34	10	24	39
C_1	D_2	5	11	16	3	17	32	5	14	33
C_2	D_1	8	13	20	10	24	34	9	24	36
C_2	D_2	4	10	15	5	19	29	4	16	34

a) Atlikite keturfaktorę dispersinę analizę. b) Atlikite kovariacinę analizę tarę, kad faktorius B kiekybinis, ir parinkdami priklausomybei nuo jo apibūdinti pirmojo ir antrojo laipsnio polinomus.

II.3.58. Lentelėje pateikti duomenys apie krakmolo plėvelės tvirtumą Y priklausomai nuo krakmolo tipo (faktorius A); A_1 – iš kviečių; A_2 – iš ryžių; A_3 – iš kukurūzų; A_4 – iš bulvių; A_5 – iš saldžiųjų bulvių) ir nuo plėvelės storio X [16].

A_1		A_2		A_3		A_4		A_5	
Y	X	Y	X	Y	X	Y	X	Y	X
263,7	5,0	556,7	7,1	731,0	8,0	983,3	13,0	837,1	9,4
130,8	3,5	552,5	6,7	710,0	7,3	958,8	13,3	901,2	10,6
382,9	4,7	397,5	5,6	604,7	7,2	747,8	10,7	595,7	9,0
302,5	4,3	532,3	8,1	508,8	6,1	866,0	12,2	510,0	7,6
213,3	3,8	587,8	8,7	393,0	6,4	810,8	11,6		
132,1	3,0	520,9	8,3	416,0	6,4	950,0	9,7		
292,0	4,2	574,3	8,4	400,0	6,9	1282,0	10,8		
315,5	4,5	505,0	7,3	335,6	5,8	1233,8	10,1		
262,4	4,3	604,6	8,5	306,4	5,3	1660,0	12,7		
314,4	4,1	522,5	7,8	426,0	6,7	746,0	9,8		
310.8	5,5	555,0	8,0	382,5	5,8	650,0	10,0		
280,8	4,8	561,1	8,4	340,8	5,7	992,5	13,8		
331,7	4,8			436,7	6,1	896,5	13,3		
672,5	8,0			333,3	6,2	873,9	12,4		
496,0	7,4			382,3	6,3	924,4	12,2		
311,9	5,2			397,7	6,0	1050,0	14,1		
276,7	4,7			619,1	6,8	973,3	13,7		
325,7	5,4			857,3	7,9				
310,8	5,4			592,5	7,2				
288,0	5,4								
269,3	4,9								

a) Atlikite kintamojo Y vienfaktorę dispersinę analizę priklausomai nuo faktoriaus A; b) Atlikite kintamojo Y regresinę analizę neatsižvelgdami į faktorių A. c) Atlikite kovariacinę analizę ir palyginkite gautus rezultatus.

II.3.59. Lentelėje pateiktas trejų metų kviečių derlingumas Y (faktorius A) šešiose skirtingose Anglijos žemės ūkio stotyse (faktorius B). Kartu užregistruotas augalo aukštis varpų atsiradimo metu (kintamasis X_1) ir vidutinis augalų iš vieno kelmelio skaičius (kintamasis Z) [16].

Metai	Kintamasis	B_1	B_2	B_3	B_4	B_5	B_6
1933	Y	19,0	22,2	35,3	32,8	25,3	35,8
	X	25,6	25,4	30,8	33,0	28,5	28,0
	Z	14,9	13,3	4,6	14,7	12,8	7,5
1934	Y	32,4	32,2	43,7	35,7	28,3	35,2
	X	25,4	28,3	35,3	32,4	25,9	24,2
	Z	7,2	9,5	6,8	9,7	9,2	7,5
1935	Y	26,2	34,7	40,0	29,6	20,6	47,2
	X	27,2	34,4	32,5	27,5	23,7	32,9
	Z	18,6	22,2	10,0	17,6	14,4	7,9

a) Ar yra derlingumo priklausomybė nuo metų, kuri nėra paaiškinama Y regresija X ir Z atžvilgiu? b) Ar yra derlingumo priklausomybė nuo stoties, kuri nėra paaiškinama Y regresija X ir Z atžvilgiu? c) Patikrinkite hipotezę, kad nėra regresijos Y kintamojo Z atžvilgiu. d) 1934 metais stoties B_5 apylinkėje užregistruotas vidutinis augalų aukštis $X = 27$ ir augalų iš vieno krūmelio skaičius $Z = 10$. Gaukite taškinį numatomo derliaus įvertį.

II.3.60. Lyginami keturi vaistai (faktorius A), mažinantys kraujo spaudimą. Registruojamas kraujo spaudimas po gydymo Y. Kartu užregistruotas kraujo spaudimas prieš gydymą X [1].

A	X	Y	A	X	Y
A_1	194	157	A_3	172	136
	162	136		196	182
	183	145		158	134
	180	153			
A_2	154	124	A_4	158	124
	184	123		165	124
	173	143		186	132
	170	136		182	133

a) Atlikite dispersinę analizę, neatsižvelgdami į kintamąjį X.

b) Atlikite vienfaktorę vieno kintamojo kovariacinę analizę.

c) Palyginkite a) ir b) gautus rezultatus.

II.3.61. Lentelėje pateikta 30 kiaulių svorio prieaugis Y priklausomai nuo fermos (faktorius A), maitinimo tipo (faktorius B), lyties (faktorius C). Eksperimentas suplanuotas pagal trifaktorės dispersinės analizės kryžminės klasifikacijos schemą su vienu stebėjimu langelyje. Daroma prielaida, kad svorio prieaugis gali priklausyti nuo tolydžios kovariantės – pradinio svorio X, kurio reikšmės taip pat pateiktos lentelėje [15].

Reikia įvertinti faktorių A, B, C poveikį svorio prieaugiui eliminuojant kovariantės X įtaką.

A	B	C	X	Y	A	B	C	X	Y
A_1	B_1	C_1	48	9,94	A_3	B_3	C_1	33	7,63
A_1	B_2	C_1	48	10,00	A_3	B_1	C_1	35	9,32
A_1	B_3	C_1	48	9,75	A_3	B_2	C_1	41	9,34
A_1	B_3	C_2	48	9,11	A_3	B_2	C_2	46	8,43
A_1	B_2	C_2	39	8,51	A_3	B_3	C_2	42	8,90
A_1	B_1	C_2	38	9,52	A_3	B_1	C_2	41	9,32
A_2	B_2	C_1	32	9,24	A_4	B_3	C_1	50	10,37
A_2	B_3	C_1	28	8,66	A_4	B_1	C_2	48	10,56
A_2	B_1	C_1	32	9,48	A_4	B_2	C_1	46	9,68
A_2	B_3	C_2	37	8,50	A_4	B_1	C_1	46	10,98
A_2	B_1	C_2	35	8,21	A_4	B_2	C_2	40	8,86
A_2	B_2	C_2	38	9,95	A_4	B_3	C_2	42	9,51
A_5	B_2	C_1	37	9,67	A_5	B_2	C_2	40	9,20
A_5	B_1	C_1	32	8,82	A_5	B_3	C_2	40	8,76
A_5	B_3	C_1	30	8,57	A_5	B_1	C_2	43	10,42

a) Tarę, kad nėra faktorių sąveikos, atlikite stebėjimų trifaktorę dispersinę analizę.
b) Atlikite trifaktorę analizę eliminuodami kintamojo X įtaką.
c) Palyginkite a) ir b) punktuose gautus rezultatus.

II.3.5. Faktoriniai eksperimentai 2^k

II.3.62. Faktoriniame eksperimente 2^2 su trimis stebėjimais langelyje gauti rezultatai pateikti lentelėje eksperimentus žymint kodiniu pavidalu (žr. [3], 4.5.1 skyrelį).

Kodas	(1)	a	b	ab
Y	0; 2; 1	4; 6; 2	−1; −3; 1	−1; −3; −7

Įvertinkite Y tiesinės regresijos parametrus kintamųjų Z_1 ir Z_2 atžvilgiu. Priėmę normalumo prielaidą, patikrinkite regresijos koeficientų lygybės 0 hipotezes.

II.3.63. Tiriamos galingumo sąnaudos Y pjaustant metalą keraminiu instrumentu priklausomai nuo instrumento tipo (kintamasis X_1), rėžiklio briaunelės kampo (kintamasis X_2) ir nuo pjovimo tipo (kintamasis X_3). Atliktas visas faktorinis eksperimentas 2^3 (žr. [3], 4.5.2 skyrelį). Rezultatai (sąlyginiais vienetais), eksperimentus žymint kodiniu pavidalu, pateikti lentelėje [8].

Kodas	(1)	a	b	ab	c	bc	ac	abc
Y	2	−5	15	13	−12	−2	−17	−7

a) Įvertinkite Y tiesinės regresijos parametrus kintamųjų Z_1, Z_2, Z_3 atžvilgiu. Priėmę normalumo prielaidą, patikrinkite šių parametrų lygybės 0 hipotezes. b) Papildomai įvertinkite regresijos koeficientus β_{ij} prie sandaugų Z_iZ_j, $i \neq j$. Patikrinkite šių koeficientų lygybės 0 hipotezes.

II.3.64. Lentelėje pateikti tam tikro cheminio eksperimento duomenys. Atliktas visas faktorinis eksperimentas 2^3 (žr. [3], 4.5.2 skyrelį) su dviem stebėjimais langelyje.

Kodas	(1)	a	b	ab	c	bc	ac	abc
Y	1595	1573	1835	1700	1745	1838	2184	1717
	1578	1592	1823	1815	1689	1614	1538	1806

Atlikite duomenų analizę.

II.3.65. Lentelėje pateikti tam tikro faktorinio eksperimento 2^4 (žr. [3], 4.5.3 skyrelį) su dviem stebėjimais langelyje rezultatai.

Kodas	(1)	a	b	ab	c	bc	ac	abc
Y	1985	1595	1765	1835	1694	1806	2243	1614
	1592	2067	1700	1823	1712	1758	1745	1838
Kodas	d	ad	bd	abd	cd	bcd	acd	$abcd$
Y	2156	1578	1923	1863	2184	1957	1745	1917
	2032	1733	2007	1910	1921	1717	1818	1922

a) Atlikite duomenų analizę, tarę, kad regresijos koeficientai prie trijų ir keturių kovariančių sandaugų lygūs 0. b) Priėmę normalumo prielaidą, patikrinkite regresijos koeficientų lygybės 0 hipotezes. c) Raskite tolesnio nepriklausomo Y stebėjimo prognozės intervalą, jeigu žinoma, kad jis bus atliktas taške $\boldsymbol{z} = (z_1, ..., z_4)^T$, kurio koordinatės tenkina sąlygą

$$\sum_i z_i^2 + \sum_{i \neq j} z_i^2 z_j^2 = \rho^2.$$

II.3.66. Atliekamas visas faktorinis eksperimentas 3^2, kai kiekviena kovariantė Z_i (jeigu reikia atlikus transformavimą) įgyja reikšmes $-1; 0; +1$, t. y. eksperimentas atliktas kvadrato viršūnėse, centre ir kraštinių vidurio taškuose. Patikrinkite, kad regresijos lygties

$$Y_j = \beta_0 + \beta_1 Z_{1j} + \beta_2 Z_{2j} + \beta_{11} U_{1j}^2 + \beta_{22} U_{2j}^2 + \beta_{12} U_{1j}^2 U_{2j}^2 +$$

$$+\beta_{211} Z_{2j} U_{1j}^2 + \beta_{122} Z_{1j} U_{2j}^2 + e_j, \quad U_{ij}^2 = Z_{ij}^2 - \bar{Z}^2{}_{i.}, \quad i = 1,2; \quad j = 1, ..., 9.$$

plano matrica turi ortogonalius stulpelius. Raskite regresijos parametrų įvertinius ir jų dispersijas. Tardami, kad a. d. $\{e_j\}$ yra nepriklausomi ir normalieji $e_j \sim N(0, \sigma^2)$, raskite kriterijus regresijos koeficientų lygybės 0 hipotezėms tikrinti.

II.3.67. (**II.3.66** pratimo tęsinys). Norint įvertinti regresijos koeficientus prie kovariančių sandaugų ir kvadratų, eksperimentas 2^2 atliktas kvadrato viršūnėse, papildomas stebėjimais keturiuose taškuose $(\pm a, 0), (0, \pm a)$. Norėdami įvertinti dispersiją, eksperimentą du kartus pakartojame eksperimento centre, t. y. taške $(0, 0)$. Parinkite a taip, kad regresijos lygties plano matrica turėtų ortogonalius stulpelius.

II.3.68. (**II.3.67** pratimo tęsinys.) Apibendrinkite **II.3.67** pratimą 3 kovariančių atveju.

II.3.69. Lentelėje pateiktos jėgos Y, reikalingos nustumti gaminį nuo konvejerio juostos priklausomai nuo temperatūros (kovariantė X_1) ir nuo drėgnumo (kovariantė X_2). Eksperimentas atliktas pagal visą faktorinio eksperimento 3^2 planą su dviem stebėjimais langelyje. Pateikiamos lentelės pirmoje eilutėje yra transformuotos kovariantės X_1 reikšmės, o pirmajame stulpelyje – kovariantės X_2 reikšmės.

	-1	-1	0	0	$+1$	$+1$
-1	0,8;	2,8	1,5;	3,2	2,5;	4,2
0	1,0;	1,6	1,6;	1,8	1,8;	1,0
$+1$	2,0;	2,2	1,5;	0,8	2,5;	4,0

Įvertinkite regresijos lygties parametrus. Priėmę normalumo prielaidą patikrinkite regresijos koeficientų lygybės 0 hipotezes.

II.3.70. Tarkime, kad vienu metu galime atlikti tik keturis **II.3.63** pratimo eksperimentus. Sudalinkite eksperimentus į dvi replikas (žr. [3], 4.5.4 skyrelį) su tarpblokiniu efektu sumaišydami a) visų trijų faktorių sąveiką; b) faktorių A_1 ir A_3 sąveiką.

II.3.6. Apibendrintieji tiesiniai modeliai

II.3.71. Firmoje užregistruoti klientų telefono skambučių skaičiai per kiekvieną iš 7 darbo valandų (kovariantė X_1) kiekvienai iš 5 savaitės darbo dienų (kovariantė X_2). Toks pat eksperimentas pakartotas kitą savaitę. Skambučių skaičiaus $Y_{ijk}, i = 1, ..., 7, j = 1, ..., 5, k = 1,\ 2$ stebiniai pateikti lentelėje.

$Z_{1i}Z_{2j}$	-2	-1	0	1	2	$\sum$
-3	30 44	30 36	26 30	31 31	18 43	325
-2	29 34	31 36	22 35	18 30	25 31	291
-1	28 41	22 24	23 26	21 29	26 28	268
0	23 24	19 24	23 31	20 25	21 31	241
1	30 30	32 40	26 33	26 34	26 36	313
2	30 38	28 40	36 37	23 25	20 24	301
3	34 39	24 41	25 34	21 26	25 41	310
$\sum$	454	427	407	366	395	2049

Lentelėje pateiktos centruotos kovariančių reikšmės $Z_{1i} = X_{1i} - 4, i = 1, ..., 7$; $Z_{2j} = X_{2j} - 3, j = 1, ..., 5$.

Tarkime, kad skambučių srautas yra puasoninis su pastoviu intensyvumu valandos intervale, t. y. $Y_{ijk} \sim \mathcal{P}(\lambda_{ij})$.

Atlikite imties

$$(Y_{111}, Z_{11}, Z_{21}), (Y_{112}, Z_{11}, Z_{21}), ..., (Y_{771}, Z_{17}, Z_{27}), (Y_{772}, Z_{17}, Z_{27})$$

puasoninę regresinę analizę (žr. [3], 5.2.1 skyrelį). Iš paskutinio lentelės stulpelio matyti, kad skambučių skaičiaus kitimas dienos metu nėra tiesinis, todėl Y_{ijk} priklausomybei nuo kovariančių apibūdinti naudokite tokį modelį:

$$\mathbf{E}Y_{ijk} = \dot{B}(\boldsymbol{\beta}^T \boldsymbol{Z}_{ij}) = e^{\boldsymbol{\beta}^T \boldsymbol{Z}_{ij}} = e^{\beta_0 + \beta_1 Z_{1i} + \beta_2 Z_{1i}^2 + \beta_3 Z_{2j}},$$

$$i = 1, ..., 7, \quad j = 1, ..., 5, \quad k = 1,\ 2.$$

a) Raskite parametro $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3)^T$ DT įvertį $\hat{\boldsymbol{\beta}}$ ir kovariacinės matricos $\boldsymbol{V}(\hat{\boldsymbol{\beta}})$ įvertį $\hat{\boldsymbol{V}}(\hat{\boldsymbol{\beta}})$.

b) Patikrinkite hipotezę $H : \beta_1 = \beta_2 = \beta_3 = 0$ ir hipotezes $H_j : \beta_j = 0, j = 1, 2, 3$.

c) Raskite trečios savaitės darbo dienos pirmos ir ketvirtos valandos vidutinio skambučių skaičiaus taškinius ir intervalinius ($Q = 0,95$) įverčius.

II.3.72. (**II.3.71** pratimo tęsinys). Atlikite **II.3.71** pratimo užduotis a), b), c) tardami, kad teisingas normaliosios teorijos tiesinis regresijos modelis:

$$Y_{ijk} = \alpha_0 + \alpha_1 Z_{1i} + \alpha_2 Z_{1i}^2 + \alpha_3 Z_{2i} + e_{ijk},$$

čia e_{ijk} nepriklausomi normalieji a. d. $e_{ijk} \sim N(0, \sigma^2)$. Palyginkite **II.3.71** ir **II.3.72** pratimų atsakymus.

II.3.73. (**II.3.71** pratimo tęsinys). Atlikite **II.3.71** pratimo stebinių dispersiją stabilizuojančią transformaciją $U_{ijk} = \sqrt{4Y_{ijk} - 1}$. Kai λ_{ij} didelis, a. d. U_{ijk} skirstinys aproksimuojamas normaliuoju su vienetine dispersija. Atlikite **II.3.71** pratimo užduotis a), b), c) tardami, kad teisingas tiesinis regresijos modelis:

$$U_{ijk} = \gamma_0 + \gamma_1 Z_{1i} + \gamma_2 Z_{1i}^2 + \gamma_3 Z_{2i} + e_{ijk},$$

čia e_{ijk} nepriklausomi normalieji a. d. $e_{ijk} \sim N(0,\ 1)$. Palyginkite **II.3.71**, **II.3.72** ir **II.3.73** pratimų atsakymus.

II.3.74. Tiriant gaminių patikimumą iš 4 įmonių (kovariantė X) pagamintos produkcijos atsitiktinai atrinkta ir išbandyta po 20 gaminių. Gaminių darbo laiko iki gedimo Y_{ij} stebiniai pateikti lentelėje.

X	Y_{ij}									
I	65,10	74,80	25,11	69,89	28,73	13,27	49,60	26,96	30,03	16,46
II	7,98	31,27	29,81	51,75	18,61	7,07	13,60	27,30	5,56	17,50
III	27,17	13,64	10,66	15,59	26,56	17,76	26,11	13,32	14,40	20,80
IV	15,55	30,47	14,16	24,02	9,09	13,82	18,68	16,73	19,69	21,54

X	Y_{ij}									
I	38,72	59,69	80,03	86,27	19,77	37,36	36,30	37,98	60,31	51,45
II	19,99	53,88	21,94	32,27	32,34	7,28	17,87	7,42	17,17	11,66
III	20,82	30,93	29,28	15,57	35,78	33,10	32,44	15,84	17,11	16,04
IV	27,46	42,90	11,62	13,66	11,54	16,33	18,75	13,14	36,04	24,92

Tarkime, kad gaminio darbo laikas Y_{ij} turi gama skirstinį $Y_{ij} \sim G(\lambda_j,\ 3)$. Kadangi kovariantė X nominali, tai ją koduokime keisdami vektoriumi $\boldsymbol{Z} = (Z_1, Z_2, Z_3)^T$, kuris įgyja reikšmes $(0,\ 0,\ 0)^T$, $(1,\ 0,\ 0)^T$, $(0,\ 1,\ 0)^T$ ir $(0,\ 0,\ 1)^T$ atitinkamai pirmajai, antrajai, trečiajai ir ketvirtajai įmonei.

Atlikite imties

$$(Y_{11}, \boldsymbol{Z}_{11}), ..., (Y_{20,4}, \boldsymbol{Z}_{20,4})$$

gama regresinę analizę (žr. [3], 5.2.2 skyrelį). Priklausomybei nuo kovariantės X apibūdinti naudokite tokį modelį:

$$\mathbf{E}Y_{ij} = 3e^{\beta_0+\beta_1 Z_{i1}+\beta_2 Z_{i2}+\beta_3 Z_{i3}}, \quad i = 1, ..., 20, \quad j = 1, ..., 4.$$

a) Raskite parametro $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3)^T$ DT įvertį $\hat{\boldsymbol{\beta}}$ ir kovariacinę matricą $\boldsymbol{V}(\hat{\boldsymbol{\beta}})$.

b) Patikrinkite hipotezę $H : \beta_1 = \beta_2 = \beta_3 = 0$, hipotezes $H_j : \beta_j = 0, j = 1, 2, 3$ ir hipotezę $H_{23} : \beta_2 = \beta_3$.

c) Raskite trečios įmonės pagaminto gaminio darbo laiko vidurkio taškinį ir intervalinį $(Q = 0,95)$ įverčius.

II.3.75. Turime 24 studentų įskaitos duomenis. Priklausomas kintamasis $Y = 1$, jeigu studentas gavo įskaitą, ir $Y = 0$ – jeigu negavo. Kiek valandų studentas dirbo pratybų metu, rodo kintamasis X_2. Ar studentas prieš pat sesiją ko nors klausė dėstytojo, rodo kintamasis X_1 ($X_1 = 1$, jeigu klausė, ir $X_1 = 0$, jeigu neklausė). Duomenys pateikti lentelėje ([8], II dalis).

Y	0	0	0	0	0	0	0	0	0	0	0	1
X_1	0	0	0	0	0	1	1	1	1	1	1	0
X_2	19	17	13	15	19	21	17	18	23	15	13	26
Y	1	1	1	1	1	1	1	1	1	1	1	1
X_1	0	0	1	1	1	1	1	1	1	0	0	0
X_2	30	19	22	21	24	28	30	27	21	24	20	16

a) Įvertinkite logistinės regresijos parametrus. b) Įvertinkite kovariančių X_1 ir X_2 šansų santykius. c) Įvertinkite tikimybę gauti įskaitą, kai $X_2 = 20$, $X_1 = 0$ ir kai $X_2 = 20$, $X_1 = 1$. d) Sudarykite klasifikacinę lentelę, kai objektas priskiriamas tai klasei, kurios tikimybės įvertis didesnis.

II.3.76. Gimdymo namuose surinkti duomenys apie gimdyvės amžių X_1, rūkymą ($X_2 = 1$ – rūko, $X_2 = 0$ – nerūko), hipertoniją X_3 ($X_3 = 1$ – serga, $X_3 = 0$ – neserga), moters svorį X_4 ir naujagimio svorį Z. Naujagimis sveria nepakankamai, jeigu jo svoris nesiekia 2500 g ([8], II dalis).

X_1	X_2	X_3	X_4	Z	X_1	X_2	X_3	X_4	Z
24	0	0	64,0	1703	29	0	0	75,0	2922
21	1	1	82,5	1792	26	1	0	84,0	2922
21	0	0	100,0	1930	17	0	0	56,5	2922
19	0	0	51,0	2084	35	1	0	60,5	2950
24	0	0	69,0	2102	33	1	0	54,5	3035
17	1	0	55,0	2227	21	1	0	92,5	3044
18	0	0	74,0	2284	19	0	0	94,5	3064
15	0	0	57,5	2383	21	0	0	80,0	3064
17	0	0	60,0	2440	19	0	0	57,5	3177
20	0	0	52,5	2452	28	0	0	70,0	3236
14	1	0	50,5	2468	16	1	0	67,5	3376
14	0	0	50,0	2497	22	0	0	65,5	3462
21	1	1	65,0	2497	32	0	0	85,0	3475
33	0	0	77,5	2553	19	0	0	52,5	3574
32	0	0	60,5	2837	24	0	0	55,0	3730
28	0	0	83,5	2879	25	0	1	60,0	3985

a) Tarkime, $Y = 1$, jeigu naujagimio svoris mažesnis už 2500 g, ir $Y = 0$ – priešingu atveju. Įvertinkite logistinės regresijos parametrus prognozuodami kintamąjį Y pagal X_1, X_2, X_3, X_4.

b) Įvertinkite šansų santykius.

c) Patikrinkite hipotezes dėl regresijos parametrų reikšmingumo.

d) Įvertinkite tikimybę, kad 30 metų amžiaus motina, kuri rūko, serga hipertonija ir sveria 85 kg, pagimdys nepakankamo svorio kūdikį. Palyginkite šią prognozę su ta, kurią gautume prognozuodami kintamąjį Z tiesine regresija kintamųjų X_1, X_2, X_3, X_4 atžvilgiu.

II.3.77. Ar galima pagal pajamas (kintamasis X_1) ir darbo prestižiškumo indeksą (kintamasis X_2) atpažinti, kad respondentas turi aukštąjį išsilavinimą ($Y = 1$ – jeigu turi ir $Y = 0$ – jeigu neturi)? Duomenys pateikti lentelėje ([8], II dalis).

X_1	3670	1923	3067	3811	3494	2012	1637	1265	2722
X_2	60	65	70	105	70	55	55	35	105
Y	1	0	1	1	1	0	0	0	0
X_1	3125	4050	3458	2219	3781	2736	2568	3408	3298
X_2	95	115	95	95	90	85	135	110	60
Y	0	1	0	0	1	0	0	0	1
X_1	4050	1501	3340	3193	3043	3536	3780	3798	
X_2	135	50	65	60	95	80	94	78	
Y	1	0	1	1	1	1	1	1	

Atlikite duomenų logistinę regresiją.

II.3.78. Lentelėje pateikti dviejų futbolo lygų (kintamasis Z) duomenys: atstumas iki vartų (X), bandymų įmušti įvartį skaičius (N), sėkmių skaičius (M).

a) Nagrinėdami lygas atskirai, parinkite logistinės regresijos modelį, kuriame kovariantė yra atstumas.

b) Parinkite logistinės regresijos modelį, kuriame kovariantės yra atstumas ir lyga.

Lyga Z	Atstumas X	Sėkmių skaičius M	Bandymų skaičius N
0	14,5	68	77
0	24,5	74	95
0	34,5	61	113
0	44,5	38	138
0	52,0	2	38
1	14,5	62	67
1	24,5	49	70
1	34,5	43	79
1	44,5	25	82
1	52,0	7	24

II.3.7. Sprendimai, atsakymai, nurodymai

II.3.1 skyrelis

II.3.1. a) Pažymėję $h = h(\mathbf{X}), \mu = \mu(\mathbf{X})$, gauname

$$\mathbf{E}(Y-h)^2 = \mathbf{E}((Y-\mu)+(\mu-h))^2 = \mathbf{E}(Y-\mu)^2 + \mathbf{E}(\mu-h)^2 \geq \mathbf{E}(Y-\mu)^2,$$

nes

$$\mathbf{E}((Y-\mu)(\mu-h)) = \mathbf{E}\{(\mu-h)\mathbf{E}[(Y-\mu)|\mathbf{X}]\} = 0.$$

b) Randame

$$\mathbf{Cov}\,(Y,h) = \mathbf{E}\{(h-\mathbf{E}h)\mathbf{E}[(Y-\mathbf{E}Y)|\mathbf{X}]\} = \mathbf{Cov}\,(h,\mu);$$

kai $h=\mu$, tai $\mathbf{Cov}\,(Y,\mu) = \mathbf{Cov}\,(\mu,\mu) = \boldsymbol{V}\mu$. Tada

$$\rho^2(Y,h) = \frac{\mathbf{Cov}^{\,2}(Y,h)}{\boldsymbol{V}Y\boldsymbol{V}h}\frac{\boldsymbol{V}\mu}{\boldsymbol{V}\mu} = \rho^2(\mu,h)\rho^2(Y,\mu) \leq \rho^2(Y,\mu).$$

II.3.2. Kadangi

$$\mathbf{Cov}\,(\mu, Y-\mu) = \boldsymbol{V}\mu - \boldsymbol{V}\mu = 0,$$

tai

$$\boldsymbol{V}Y = \boldsymbol{V}\mu + \boldsymbol{V}(Y-\mu) = \boldsymbol{V}Y\eta^2_{Y\boldsymbol{X}} + \boldsymbol{V}Y(1-\eta^2_{Y\boldsymbol{X}}.)$$

Koreliacinis santykis $\eta^2_{Y\boldsymbol{X}}$ rodo, kurią dalį Y dispersijos $\boldsymbol{V}Y$ sudaro optimalios prognozės $\mu(\mathbf{X})$ dispersija. Prognozės paklaidos $Y-\mu(\mathbf{X})$ dispersija sudaro $(1-\eta^2_{Y\boldsymbol{X}})$ dispersijos $\boldsymbol{V}Y$ dalį. Jeigu $\eta^2_{Y\boldsymbol{X}} = 1$, tai Y ir $\mathbf{X}$ susieti funkcine priklausomybe, nes $\boldsymbol{V}(Y-\mu(\mathbf{X})) = 0$; jeigu $\eta^2_{Y\boldsymbol{X}} = 0$, tai prognozuoti Y pagal $\mathbf{X}$ nėra prasmės, nes $\boldsymbol{V}(\mu(\mathbf{X})) = 0$ ir $\mu(\mathbf{X}) = \mathbf{E}Y$ nepriklauso nuo $\mathbf{X}$.

II.3.3. Atėmę $\eta^2_{Y\boldsymbol{X}^{(2)}|\boldsymbol{X}^{(1)}}$ apibrėžime abi lygybės puses iš vieneto ir padauginę iš $1-\eta^2_{Y\boldsymbol{X}^{(1)}}$, gauname

$$1-\eta^2_{Y\boldsymbol{X}} = (1-\eta^2_{Y\boldsymbol{X}^{(1)}})(1-\eta^2_{Y\boldsymbol{X}^{(2)}|\boldsymbol{X}^{(1)}}).$$

Pakartotinai taikydami šią lygybę gausime pratime pateiktą lygybę.

II.3.4. a) Gauname

$$SS(\alpha, \boldsymbol{\beta}) = \mathbf{E}(Y - \alpha - \boldsymbol{\beta}^T \mathbf{X}) =$$

$$= \mathbf{E}((Y - \mathbf{E}Y) + (\mathbf{E}Y + (\mathbf{E}Y - \alpha - \boldsymbol{\beta}^T \mathbf{E}(\mathbf{X}))) - \boldsymbol{\beta}^T (\mathbf{X} - \mathbf{E}(\mathbf{X})))^2 =$$

$$\boldsymbol{V}Y + (\mathbf{E}Y - \alpha - \boldsymbol{\beta}^T \mathbf{E}(\mathbf{X})))^2 + \boldsymbol{\beta}^T \boldsymbol{\Sigma}\boldsymbol{\beta} - 2\boldsymbol{\beta}^T \sigma_{Y\boldsymbol{X}} \geq \boldsymbol{V}Y + \boldsymbol{\beta}^T \boldsymbol{\Sigma}\boldsymbol{\beta} - 2\boldsymbol{\beta}^T \sigma_{Y\boldsymbol{X}}.$$

Imdami dešinės pusės išvestinę pagal $\boldsymbol{\beta}$ ir prilyginę nuliui, gauname lygčių sistemą $\boldsymbol{\Sigma}\boldsymbol{\beta} = \sigma_{Y\boldsymbol{X}}$. Iš čia, jeigu $\boldsymbol{\Sigma}$ teigiamai apibrėžta matrica, gauname

$$\boldsymbol{\beta}^* = \boldsymbol{\Sigma}^{-1}\sigma_{Y\boldsymbol{X}}, \quad \alpha^* = \mathbf{E}Y - (\boldsymbol{\beta}^*)^T \mathbf{E}(\mathbf{X}).$$

b) Pažymėsime, kad $\sigma_{Y\boldsymbol{X}} = \boldsymbol{\Sigma}\boldsymbol{\beta}^*$. Gauname

$$\mathbf{Cov}\,(Y, \boldsymbol{\beta}^T \mathbf{X}) = \boldsymbol{\beta}^T \sigma_{Y\boldsymbol{X}} = \boldsymbol{\beta}^T \boldsymbol{\Sigma}\boldsymbol{\beta}^*, \quad \mathbf{Cov}\,(Y, (\boldsymbol{\beta}^*)^T \mathbf{X}) = (\boldsymbol{\beta}^*)^T \boldsymbol{\Sigma}\boldsymbol{\beta}^*;$$

$$\rho^2(Y, \boldsymbol{\beta}^T \mathbf{X}) = \frac{(\boldsymbol{\beta}^T \boldsymbol{\Sigma}\boldsymbol{\beta}^*)^2}{\boldsymbol{V}Y \boldsymbol{\beta}^T \boldsymbol{\Sigma}\boldsymbol{\beta}} \leq \frac{\boldsymbol{\beta}^T \boldsymbol{\Sigma}\boldsymbol{\beta}(\boldsymbol{\beta}^*)^T \boldsymbol{\Sigma}\boldsymbol{\beta}^*}{\boldsymbol{V}Y \boldsymbol{\beta}^T \boldsymbol{\Sigma}\boldsymbol{\beta}} =$$

$$\frac{((\boldsymbol{\beta}^*)^T \boldsymbol{\Sigma}\boldsymbol{\beta}^*)^2}{\boldsymbol{V}Y (\boldsymbol{\beta}^*)^T \boldsymbol{\Sigma}\boldsymbol{\beta}^*} = \rho^2(Y, (\boldsymbol{\beta}^*)^T \mathbf{X}).$$

Koreliacijos koeficiento $\rho(Y, l^*)$ kvadratas yra

$$\rho^2_{Y\boldsymbol{X}} = \frac{\boldsymbol{V}l^*}{\boldsymbol{V}Y} = \frac{\boldsymbol{\beta}^{*T} \boldsymbol{\Sigma}\boldsymbol{\beta}^*}{\boldsymbol{V}Y} = \frac{\sigma_{Y\boldsymbol{X}} \boldsymbol{\Sigma}^{-1} \sigma_{Y\boldsymbol{X}}}{\boldsymbol{V}Y},$$

o $\rho_{Y\boldsymbol{X}} = \sqrt{\rho^2_{Y\boldsymbol{X}}} \geq 0$ vadinamas dauginiu koreliacijos koeficientu.

II.3.5. Kadangi

$$\mathbf{Cov}\,(l^*, Y - l^*) = 0,$$

tai

$$\boldsymbol{V}Y = \boldsymbol{V}l^* + \boldsymbol{V}(Y_l^*) = \boldsymbol{V}Y\rho^2_{Y\boldsymbol{X}} + \boldsymbol{V}Y(1 - \rho^2_{Y\boldsymbol{X}}.)$$

Dauginio koreliacijos koeficiento kvadratas $\rho^2_{Y\boldsymbol{X}}$ rodo, kurią dalį Y dispersijos $\boldsymbol{V}Y$ sudaro optimalios tiesinės prognozės l^* dispersija. Prognozės paklaidos $Y - l^*$ dispersija sudaro $(1 - \rho^2_{Y\boldsymbol{X}})$ dispersijos $\boldsymbol{V}Y$ dalį. Jeigu $\rho^2_{Y\boldsymbol{X}} = 1$, tai Y ir $\mathbf{X}$ susieti tiesine priklausomybe, nes $\boldsymbol{V}(Y - l^*) = 0$; jeigu $\rho^2_{Y\boldsymbol{X}} = 0$, tai tiesiškai prognozuoti Y pagal $\mathbf{X}$ nėra prasmės.

Pažymėsime, kad minimali tiesinė prognozė minimizuoja paklaidos dispersiją siauresnėje klasėje negu optimali nebūtinai tiesinė prognozė, tai galioja nelygybės

$$0 \leq \rho^2_{Y\boldsymbol{X}} \leq \eta^2_{Y\boldsymbol{X}} \leq 1.$$

Jeigu vektorius $(Y, X_1, ..., X_m)^T$ turi daugiamatį normalųjį skirstinį, tai $\rho^2_{Y\boldsymbol{X}} = \eta^2_{Y\boldsymbol{X}}$ (žr. [3], 7.4 skyrelį).

II.3.6. Atėmę $\rho^2_{Y\boldsymbol{X}^{(2)}|\boldsymbol{X}^{(1)}}$ apibrėžime abi lygybės puses iš vieneto ir padauginę iš $1 - \rho^2_{Y\boldsymbol{X}^{(1)}}$, gauname

$$1 - \rho^2_{Y\boldsymbol{X}} = (1 - \rho^2_{Y\boldsymbol{X}^{(1)}})(1 - \rho^2_{Y\boldsymbol{X}^{(2)}|\boldsymbol{X}^{(1)}}).$$

Pakartotinai taikydami šią lygybę gausime pratime pateiktą lygybę.

II.3.7. a) A. d. X sąlyginis skirstinys, kai $Y = y$ fiksuotas, yra normalusis

$$(X|Y=y) \sim N(\rho y, 1 - rho^2), \quad \mathbf{E}(X|Y=y) = \rho y, \quad \eta^2_{X|Y} = \rho^2.$$

b) A. d. X sąlyginis skirstinys, kai $Y^2 = y$ fiksuotas, yra mišinys normaliųjų skirstinių $N(-rho\sqrt{y}, 1 - rho^2)$ ir $N(\rho\sqrt{y}, 1-\rho^2)$ su vienodais svoriais $1/2$.

$$\mathbf{E}(X|Y^2 = y) = \frac{1}{2}[-rho\sqrt{y} + \rho\sqrt{y}] = 0, \quad \mathbf{E}(X|Y^2) = 0, \quad \eta^2_{X|Y^2} = 0.$$

c) Remdamiesi p. b) gauname

$$\mathbf{E}(X^2|Y^2 = y^2) = 1 - \rho^2 + \rho^2 y^2, \quad \eta^2_{X^2|Y^2} = \frac{\mathbf{Cov}^2(X^2, \rho^2 Y^2)}{\boldsymbol{V} X^2 \boldsymbol{V}(\rho^2 Y^2)} = \rho^4.$$

d) Kadangi $(X^2|Y=y) \sim N(\rho y, 1-\rho^2)$, tai

$$\mathbf{E}(X^2|Y = y) = 1 - \rho^2 + \rho^2 y^2 \quad \eta^2_{X^2|Y^2} = \frac{\mathbf{Cov}^2(X^2, \rho^2 Y^2)}{\boldsymbol{V} X^2 \boldsymbol{V}(\rho^2 Y^2)} = \rho^4.$$

II.3.8. Pradžioje tarkime, kad elipsė užrašyta kanoniniu pavidalu $x^2/a^2 + y^2/b^2 = 1$. Pažymėkime $|S| = \pi ab$ figūros, apribotos elipse, plotą, o S pačią figūrą. Tada tankio funkcija

$$f_{X,Y}(x,y) = \frac{1}{|S|}, \quad (x,y) \in S.$$

A. d. X tankis

$$f_X(x) = \int_{-b\sqrt{1-x^2/a^2}}^{b\sqrt{1-x^2/a^2}} f_{X,Y}(x,y)dy = \frac{2b}{|S|}\sqrt{1 - \frac{x^2}{a^2}}, \quad x \in (-a,\ a).$$

Sąlyginis tankis

$$f_{Y|X=x} = \frac{1}{2b\sqrt{1-x^2/a^2}}, \quad y \in (-b\sqrt{1-x^2/a^2},\ b\sqrt{1-x^2/a^2}).$$

Sąlyginis vidurkis

$$\mathbf{E}(Y|X=x) = \frac{1}{2b\sqrt{1-x^2/a^2}} \int_{-b\sqrt{1-x^2/a^2}}^{b\sqrt{1-x^2/a^2}} y dy = 0.$$

Bendras atvejis. Naudojant tiesines transformacijas

$$y' = c_i x + c_2 y, \quad x' = d_1 x,$$

elipsės lygtis suvedama į kanoninį pavidalą (pakanka išskirti pilną kvadratą). Kanoninio pavidalo elipsei $\mathbf{E}(Y;|X') = 0$, nes tolygusis skirstinys po tiesinės transformacijos išlieka. Išreiškiame

$$X = a_1 X', \quad Y = b_1 X' + b_2 Y'$$

ir gauname, kad sąlyginis vidurkis

$$\mathbf{E}(Y|X) = \mathbf{E}(b_1X' + b_2Y'|a_1x") = \frac{b_1}{a_1}\mathbf{E}(a_1X'|a_1X') +$$

$$\mathbf{E}(g_1Y'|a_1X') = \frac{b_1}{a_1}X$$

yra tiesinė X funkcija.

II.3.9. Pažymėkime lygiagretainio viršūnes $A(-3,-2), B(0,\ 1), C(3,2), D(0,-1)$; tiesės AB lygtis $x = y-1$, tiesės BC lygtis $x = 3(y-1)$, tiesės CD lygtis $x = y+1$ ir tiesės AD lygtis $x = 3(y+1)$.

Fiksuokime tašką x iš intervalo $(-3,\ 0)$. Tada a. d. Y turi tolygųjį skirstinį intervale $y_1 \le y \le y_2$; čia y_1 taškas ant tiesės AB, o y_2 – ant tiesės AD; $y_1 = x+1, y_2 = x/3-1$. Sąlyginis vidurkis $\mathbf{E}(Y|X = x) = (y_1+y_2)/2 = 2x/3$. Taigi kai $x \in (-3,\ 0)$, tai $\mathbf{E}(Y|X = x)$ yra tiesės $y = 2x/3$ atkarpa. Analogiškai, imdami intervalą $0 \le x < 3$, gausime, kad $\mathbf{E}(Y|X = x)$ yra tiesės $y = 2x/3$ atkarpa. A. d. Y regresija X atžvilgiu yra tiesinė.

Fiksuokime tašką y iš intervalo $(-2,-1)$. Tada a. d. X turi tolygųjį skirstinį intervale $x_1 < x < x_2$; čia x_1 taškas ant tiesės AB, o x_2 taškas ant tiesės AD; $x_1 = y-1, x_2 = 3(y-1)$. Sąlyginis vidurkis $\mathbf{E}(X|Y = y) = (x_1+x_2)/2 = 2y+1$. Kai $-2 < y < -1$, tai regresija $\mathbf{E}(X|Y = y)$ yra tiesės $x = 2y+1$ atkarpa. Fiksuokime tašką y iš intervalo $(-1,\ 1)$. Tada a. d. X turi tolygųjį skirstinį intervale $x_1 < x < x_2$; čia x_1 taškas ant tiesės AB, o x_2 taškas ant tiesės CD; $x_1 = y-1, x_2 = y+1$. Sąlyginis vidurkis $\mathbf{E}(X|Y = y) = (x_1+x_2)/2 = y$. Kai $-1 < y < 1$, tai regresija $\mathbf{E}(X|Y = y)$ yra tiesės $x = y$ atkarpa. Analogiškai gausime, kad intervale $1 < y < 2$ sąlyginis vidurkis $\mathbf{E}(X|Y = y)$ yra tiesės $x = 2y-1$ atkarpa.

A. d. X regresija Y atžvilgiu yra kreivė, susidedanti iš trijų atkarpų.

II.3.2 skyrelis

II.3.13. a) Kadangi

$$\boldsymbol{A}^T\boldsymbol{A} = \begin{pmatrix} n & 0 \\ 0 & \sum_i (x_i-\bar{x})^2 \end{pmatrix},$$

tai parametro $\boldsymbol{\beta} = (\alpha, \beta)^T$ mažiausiųjų kvadratų įvertinys

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y} = (\bar{Y}, \sum_i Y_i(x_i-\bar{x}) / \sum_i (x_i-\bar{x})^2)^T;$$

$$\boldsymbol{V}\hat{\alpha} = \sigma^2/n, \quad \boldsymbol{V}\hat{\beta} = \sigma^2 / \sum_i (x_i-\bar{x})^2, \quad \mathbf{Cov}\,(\hat{\alpha}, \hat{\beta}) = 0.$$

b) Liekamoji kvadratų suma

$$SS_E = \sum_i (Y_i - \hat{\alpha} - \hat{\beta}(x_i-\bar{x}))^2, \quad \mathbf{E}(SS_E) = \sigma^2(n-2),$$

ir nepaslinktasis dispersijos σ^2 įvertinys

$$s^2 = \frac{SS_E}{n-2}, \quad \mathbf{E}(s^2) = \sigma^2.$$

II.3.14. a) Tikėtinumo funkcija

$$L(\alpha,\beta,\sigma^2) = (2\pi)^{-n/2}\exp\{-\frac{1}{2\sigma^2}\sum_i (Y_i-\alpha-\beta(x_i-\bar{x}))^2 - \frac{n}{2}\ln\sigma^2\} =$$

$$(2\pi)^{-n/2}\exp\{-\frac{1}{2\sigma^2}\sum_i Y_i^2 + \frac{\alpha}{\sigma^2}\sum_i Y_i + \frac{\beta}{\sigma^2}\sum_i Y_i(x_i-\bar{x}) - B(\alpha,\beta,\sigma^2)\}$$

priklauso triparametrei eksponentinio tipo skirstinių šeimai. Statistika

$$\boldsymbol{T} = (\sum_i Y_i,\ \sum_i Y_i^2,\ \sum_i Y_i(x_i-\bar{x}))$$

yra pilnoji ir pakankamoji. Todėl bet kuri $\boldsymbol{T}$ funkcija yra jos vidurkio NMD įvertinys. Kadangi $\hat\alpha, \hat\beta,\ \hat\theta = c_1\hat\alpha + c_2\hat\beta, s^2$ yra $\boldsymbol{T}$ funkcijos ir nepaslinktieji įvertiniai, tai jie yra ir NMD įvertiniai.

b) Įvertiniai $\hat\alpha, \hat\beta, \hat\theta$ yra normaliųjų a. d. $Y_i \sim N(\alpha+\beta(x_i-\bar{x}),\sigma^2)$ tiesinės funkcijos, tai jie turi normaliuosius skirstinius. Centravę ir normavę gauname

$$\sqrt{n}\frac{\hat\alpha-\alpha}{\sigma} \sim N(0,\ 1), \qquad \frac{\hat\beta-\beta}{\sigma\sqrt{1/\sum_i (x_i-\bar{x})^2}} \sim N(0,\ 1),$$

$$\frac{\hat\theta-\theta}{\sigma\sqrt{1/n+1/\sum_i (x_i-\bar{x})^2}} \sim N(0,\ 1).$$

Dispersijos σ^2 įvertinys s^2 nepriklauso nuo $\hat\alpha$ ir $\hat\beta$ ir

$$s^2 = SS_E/(n-2), \quad s^2(n-2)/\sigma^2 \sim \chi^2(n-2).$$

II.3.15. Remdamiesi sąryšiu $s^2(n-2)/\sigma^2 \sim \chi^2(n-2)$ gauname dispersijos σ^2 lygmens $Q = 1-2\alpha$ pasikliovimo intervalą

$$(\underline{\sigma^2};\overline{\sigma^2}) = \left(\frac{s^2(n-2)}{\chi^2_\alpha(n-2)}; \frac{s^2(n-2)}{\chi^2_{1-\alpha}(n-2)}\right).$$

Remdamiesi **II.3.14** pratimu gauname sąryšius

$$\sqrt{n}\frac{\hat\alpha-\alpha}{s} \sim S(n-2), \qquad \frac{\hat\beta-\beta}{s\sqrt{1/\sum_i (x_i-\bar{x})^2}} \sim S(n-2),$$

$$\frac{\hat\mu(x)-\mu()x}{s\sqrt{1/n+(x-\bar{x})^2/\sum_i (x_i-\bar{x})^2}} \sim S(n-2),$$

iš kurių standartiniu būdu gauname pasikliovimo intervalus.

b) Kriterijus dėl parametrų reikšmių sudarome standartiniu būdu (pavyzdžiui, suformulavus juos pasikliovimo intervalų terminais). Pateiksime kriterijų hipotezei dėl prognozavimo tikslingumo, t. y. hipotezei $H : \beta = 0$, tikrinti. Hipotezė H atmetama reikšmingumo lygmens α kriterijumi, kai

$$|t| = \frac{|\hat\beta-0|}{s\sqrt{1/\sum_i (x_1-\bar{x})^2}} > t_{\alpha/2}(n-2).$$

II.3.16. Tegu regresija $\mu(x)$ yra bet kokio pavidalo. Pažymėkime $\mu_i = \mathbf{E}(Y|x_i), i = 1, ..., k$; $\boldsymbol{\mu} = (\mu_1, ..., \mu_k)^T$. Turime tiesinį modelį

$$Y_{ij} = \mu_i + e_{ij}, \quad j = 1, ..., n_i, \quad i = 1, ..., k.$$

Parametrų μ_i DT įvertiniai $\hat{\mu}_i = \bar{Y}_{i.} = \sum_{j=1}^{n_i} Y_{ij}/n_i$, o liekamoji kvadratų suma

$$SS_E = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_{i.})^2 \stackrel{d}{=} \sigma^2\chi^2_{n-k}.$$

Tikrinamąją hipotezę galima suformuluoti pavidalu $H : \boldsymbol{H\mu} = \boldsymbol{\theta}_0$ (žr. **II.1.21** pratimą), kai matricos $\boldsymbol{H}$ rangas yra $k-2$. Iš tikrųjų

$$\mu_i = \beta_0 + \beta_1 x_i, \ \ i = 1, ..., k, \quad \Rightarrow \mu_1 - \mu_i = \beta_1(x_1 - x_i), i = 2, ..., k, \ \ \Rightarrow$$

$$\frac{\mu_1 - \mu_i}{x_1 - x_i} = \beta_1, \ \ i = 2, ..., k, \ \ \frac{\mu_1 - \mu_i}{x_1 - x_i} - \frac{\mu_1 - \mu_2}{x_1 - x_2} = 0, \ \ i = 3, ..., k.$$

Gauname

$$\frac{x_i - x_2}{x_1 - x_2}\mu_1 + \frac{x_1 - x_i}{x_1 - x_2}\mu_2 - \mu_i = 0, \quad i = 3, ..., k$$

Galime imti matricą $\boldsymbol{H}$, turinčią $k-2$ eilutes ir k stulpelių; i-oji matricos eilutė

$$\left(\frac{x_i - x_2}{x_1 - x_2}, \frac{x_1 - x_i}{x_1 - x_i}, \ 0, ..., 0 \ -1, \ 0, ..., 0\right);$$

čia -1 yra i-oje pozicijoje, $i = 3, ..., k$. Hipotezė užrašoma pavidalu $\boldsymbol{H\mu} = \mathbf{0}$; matricos $\boldsymbol{H}$ rangas yra $k-2$. Hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai

$$F = \frac{(SS_{EH} - SS_E)(n-k)}{(k-2)SS_E} > F_\alpha(k-2, n-k),$$

čia

$$SS_{EH} = \min_{\mu_i = \beta_0 + \beta_1 x_i} SS(\boldsymbol{\mu}) = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \hat{\beta}_0 - \hat{\beta}_1(x_i - \bar{x}))^2,$$

$$SS_{EH} - SS_E = \sum_{i=1}^{k} n_i(\bar{Y}_{i.} - \hat{\beta}_0 - \hat{\beta}_1(x_i - \bar{x}))^2.$$

II.3.17. Hipotezė H atmetama, kai

$$F = \frac{(SS_{EH} - SS_E)(n-k)}{(k-r)SS_E} > F_\alpha(k-r, n-k),$$

čia SS_E surasta **II.3.16** pratime, o

$$SS_{EH} - SS_E = \sum_{i=1}^{k} n_i(\bar{Y}_{i.} - \hat{\beta}_0 - \hat{\beta}_1 x_i - ... - \hat{\beta}_r x_i)^2,$$

parametro $\boldsymbol{\beta} = (\beta_0, \beta_1, ..., \beta_r)^T$ įvertinys yra

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y},$$

kai plano matricos $\boldsymbol{A}$ i-oji eilutė turi pavidalą $(1, x_i, ..., x_i^r)$, o jos rangas lygus r.

II.3.18. Kadangi

$$Y_{n+1} - \hat{\mu}(x) \sim N(0, \sigma^2(1 + b^2(x))), \quad b^2(x) = \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_i (x_i - \bar{x})^2},$$

tai

$$\frac{Y_{n+1} - \hat{\mu}(x)}{s\sqrt{1 + b^2 x}} \sim S(n-2).$$

Gauname prognozės intervalą

$$(\underline{Y}_{n+1}; \overline{Y}_{n+1}) = (\hat{\mu}(x) - st_\alpha(n-2)\sqrt{1 + b^2(x)}; \hat{\mu}(x) + st_\alpha(n-2)\sqrt{1 + b^2(x)}).$$

II.3.19. Naudojame sąryšį

$$\mathbf{P}\{\frac{|Y_{n+1} - \hat{\mu}(x)|}{s\sqrt{1 + b^2 x}} < t_\alpha(n-2)\} = Q = 1 - 2\alpha.$$

Skliaustuose turime antro laipsnio nelygybę x atžvilgiu. Išsprendę nelygybę gausime pasikliovimo intervalo rėžius

$$\bar{x} + (B \mp \sqrt{B^2 - AC})/A, \quad B = \hat{\beta}(Y_{n+1} - \hat{\alpha}),$$

$$A = \hat{\beta}^2 - \frac{s^2 t_\alpha^2(n-2)}{(n-1)\sum_i (x_i - \bar{x})^2}, \quad C = (Y_{n+1}^2 - \hat{\alpha})^2 - \frac{n+1}{n} s^2 t_\alpha^2(n-2).$$

II.3.20. Hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai $|\hat{\theta} - \theta_0|/(sc) > t_{\alpha/2}(n-2)$; $\hat{\theta} = a\hat{\alpha} + b\hat{\beta}$, $s^2 = \sum_i (Y_i - \hat{\alpha} - \hat{\beta}(x_i - \bar{x}))^2/(n-2)$; $c^2 = a^2/n + b^2/\sum_i (x_i - \bar{x})^2$.

II.3.21. Tegu pirmosios imties elementai yra (Y_{1i}, X_{1i}), $i = 1, ..., n_1$, o antrosios – (Y_{2i}, X_{2i}), $i = 1, ..., n_2$. Vertinamos regresijos tiesės pavidalo: $\alpha_1 + \beta_1(X_{1i} - \bar{X}_{1.})$ ir $\alpha_2 + \beta_2(X_{2i} - \bar{X}_{2.})$. a) Hipotezė atmetama α lygmens kriterijumi, kai $|\hat{\beta}_1 - \hat{\beta}_2|/(sc) > t_{\alpha/2}(n_1 + n_2 - 4)$; $s^2 = [\sum_i (Y_{1i} - \hat{\alpha}_1 - \hat{\beta}_1(X_{1i} - \bar{X}_{1.}))^2 + \sum_i (Y_{2i} - \hat{\alpha}_2 - \hat{\beta}_2(X_{2i} - \bar{X}_{2.}))^2]/(n_1 + n_2 - 4)$, $c^2 = 1/\sum (X_{1i} - \bar{X}_{1.})^2 + 1/\sum (X_{2i} - \bar{X}_{2.})^2$. b) Hipotezė atmetama α lygmens kriterijumi, kai $|\hat{\alpha}_1 - \hat{\alpha}_2|/(sb) > t_{\alpha/2}(n_1 + n_2 - 4)$; $b^2 = 1/n_1 + 1/n_2$. c) Hipotezė atmetama α lygmens kriterijumi, kai $(SS_{EH} - SS_E)(n_1 + n_2 - 4)/(2SS_E) > F_\alpha(2, n_1 + n_2 - 4)$; $SS_E = s^2(n_1 + n_2 - 4)$; $SS_{EH} = [\sum_i (Y_{1i} - \hat{\alpha} - \hat{\beta}(X_{1i} - \bar{X}_{..}))^2 + \sum_i (Y_{2i} - \hat{\alpha} - \hat{\beta}(X_{2i} - \bar{X}_{..}))^2]$, čia $\bar{X}_{..} = (n_1\bar{X}_{1.} + n_2\bar{X}_{2.})/(n_1 + n_2)$, o $\hat{\alpha}$ ir $\hat{\beta}$ yra regresijos tiesės $\alpha + \beta(x - \bar{X}_{..})$ parametrų įvertiniai, gauti sujungus visus $n_1 + n_2$ stebėjimus. d) perkelkime koordinačių pradžią: $Z_{1i} = X_{1i} - c$, $i = 1, ..., n_1$ ir $Z_{2j} = X_{2j} - c$, $j = 1, ..., n_2$. Kintamųjų (Y_{1i}, Z_{1i}) ir (Y_{2j}, Z_{2j}) terminais tikrinamoji hipotezė ekvivalenti p. b) hipotezei.

II.3.22. Plano matrica

$$\boldsymbol{A} = \begin{pmatrix} 1 & -3 \\ 1 & -1 \\ 1 & 1 \\ 1 & 3 \end{pmatrix}, \quad (\boldsymbol{A}^T\boldsymbol{A})^{-1} = \begin{pmatrix} 1/4 & 0 \\ 0 & 1/20 \end{pmatrix},$$

$$\boldsymbol{A}^T\boldsymbol{Y} = \begin{pmatrix} Y - 1 + Y_2 + Y_3 + Y_4 \\ -3Y_1 - Y_2 + Y_3 + 3Y_4 \end{pmatrix}.$$

Gauname parametrų įvertinius $\hat{\alpha} = (Y_1 + Y_2 + Y_3 + Y_4)/4$; $\hat{\beta} = (-3Y_1 - Y_2 + Y_3 + 3Y_4)/20$ ir jų realizacijas $\hat{\alpha} = 3,255$, $\hat{\beta} = 1,063$, $s^2 = SS_E/2 = 0,0121$.

II.3.23. Turime tiesinį modelį $S_t = S_0 + vt + \varepsilon_t, t = 0,\ 1,\ 2,\ 3,\ 4$; $\varepsilon_0, ..., \varepsilon_4$ nepriklausomi ir normalieji $N(0, \sigma^2)$. Plano matrica

$$\boldsymbol{A}^T = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \end{pmatrix}, \quad (\boldsymbol{A}^T\boldsymbol{A})^{-1} = \frac{1}{10}\begin{pmatrix} 6 & -2 \\ -2 & 1 \end{pmatrix}, \quad \boldsymbol{A}^T\boldsymbol{Y} = \begin{pmatrix} \sum_t S_t \\ \sum_t tS_t \end{pmatrix}.$$

Gauname parametrų įvertinių realizacijas $\hat{S}_0 = 12,834$, $\hat{v} = 0,340$, $s^2 = SS_E/3 = 0,0259$ ir pasikliovimo intervalus
$(\underline{v}; \overline{v}) = (0,178; 0,502)$; $(\underline{\sigma}^2; \overline{\sigma}^2) = (0,0083; 0,3601)$.

II.3.24. Parametro α įvertis $\hat{\alpha} = 0,92$; $s^2 = 0,0967$; $(\underline{\alpha}; \overline{\alpha}) = (0,5338; 1,3062)$; $(\underline{\sigma}^2; \overline{\sigma}^2) = (0,0347; 0,7992)$. *Nurodymas.* A. d. $Z_1 = Y_1, Z_j = Y_j - Y_{j-1}, j = 2,3,4,5$ yra nepriklausomi su dispersijomis σ^2 ir vidurkiais $\mathbf{E}Z_1 = 0, \mathbf{E}Z_2 = ... = \mathbf{E}Z_5 = \alpha$.

II.3.25. Gauname taškinį įvertį $\hat{\alpha} = 0,92$ ir pasikliovimo intervalą

$$(\underline{\alpha};\ \overline{\alpha}) = (0,5338;\ 1,3062).$$

A. d. $Z_1 = Y_1, Z_j = Y_j - Y_{j-1}, j = 2,3,4,5$ yra nepriklausomi; jų dispersijos σ^2, o vidurkiai $\mathbf{E}Z_1 = 0, \mathbf{E}Z_2 = ... = \mathbf{E}Z_5 = \alpha$. Gauname situaciją, analogišką **II.3.24** pratimui.

II.3.26. a) Statistika F (**II.3.17** pratimas) įgijo reikšmę 0,3985; kadangi

$$pv = \mathbf{P}\{F_{4,48} > 0,3985\} = 0,8087,$$

tai atmesti tiesiškumo hipotezę nėra pagrindo; b) $\hat{\alpha} = 1,0762$, $\underline{\hat{\beta}} = 0,0540$, $\hat{\sigma}^2 = 0,06205$; c) $(\underline{\alpha}; \overline{\alpha}) = (1,0082; 1,1443)$; $(\underline{\beta}; \overline{\beta}) = (-1,4982; 1,6062)$; $(\underline{\sigma^2}; \overline{\sigma^2}) = (0,0437; 0,0950)$.

II.3.27. Modelyje

$$Y_{ij} = \mu_i + e_{ij}, \quad j = 1, ..., n, \quad i = 1, ..., 11,$$

liekamoji kvadratų suma

$$SS_E/\sigma^2 = \sum_{i=1}^{11}\sum_{j=1}^{n}(Y_{ij} - \bar{Y}_{i.})^2/\sigma^2 \sim \chi^2(11(n-1)).$$

Jeigu hipotezė teisinga, tai

$$SS_{EH} = \sum_{i=1}^{11}\sum_{j=1}^{n}(Y_{ij} - xi)^2 = SS_E + n\sum_{i=1}^{11}(\bar{Y}_{i.} - x_i)^2;$$

$$(SS_{EH} - SS_E)/\sigma^2 = n\sum_{i=1}^{11}(\bar{Y}_{i.} - x_i)^2/\sigma^2 \sim \chi^2(11).$$

Hipotezė H atmetama reikšmingumo lygmens α kriterijumi, kai

$$F + \frac{(SS_{EH} - SS_E)11(n-1)}{11SS_E} > F_\alpha(11, 11(n-1)).$$

Jeigu $\mathbf{E}(Y|X = x_i) = x_i^2$, tai

$$\sqrt{n}\frac{\bar{Y}_{i.} - x_i}{\sigma} \sim N(\lambda_i,\ 1), \quad \lambda_i = \sqrt{n}\frac{x_i^2 - x_i}{\sigma}.$$

Todėl

$$(SS_{EH} - SS_E)/\sigma^2 \sim \chi^2(11; \lambda), \quad \lambda = \sum_{i=1}^{11} \lambda_i^2 = 3,333n.$$

Imties didumui n rasti turime nelygybę

$$\mathbf{P}\{F_{11,11(n-1);\lambda} > F_{0,05}(11, 11(n-1))\} \geq 0,95.$$

Iš šios nelygybės gauname, kad $n \geq 32$.

II.3.28. Pratime **3.18** reikia $b^2(x)$ apibrėžti tokiu būdu:

$$b^2(x) = k/n + (x - \bar{x})^2/((n-1)s_x^2).$$

II.3.29. Pažymėkime $u = 1/y, v = \sin^2 x$. Tada

$$u = \gamma + \delta v; \quad \gamma = 1/\alpha, \quad \delta = (\alpha - \beta)/(\alpha\beta).$$

II.3.30. $\hat{\alpha} = 8,535$, $\hat{\beta} = 1,2066$, $\hat{\sigma}^2 = 0,8115$; statistika F (**II.3.16** pratimas) įgijo reikšmę 1,9864; kadangi

$$pv = \mathbf{P}\{F_{3,15} > 1,9864\} = 0,1594,$$

tai atmesti tiesiškumo hipotezę nėra pagrindo; tolesnio matavimo prognozės intervalas yra (14,3215; 15,5385), kai $x = 12$, ir (21,1458; 23,1934), kai $x = 18$.

II.3.31. $\hat{\mu} = 0,9343 + 0,8787x$; hipotezė $H : \beta = 0$ atmetama: a. d., turintis Stjudento skirstinį su 106 laisvės laipsniais, kai hipotezė teisinga, įgijo reikšmę 21,7608; paaiškina 0,817 sklaidos dalį.

II.3.32. a) $\hat{\beta}_0 = 0,0144, \hat{\beta}_1 = 0,0149, \hat{\sigma}^2 = 0,00004$. Determinacijos koeficientas 0,9697. Sklaidos diagramoje prognozės paklaidos išsidėsčiusios neatsitiktinai, todėl regresijos tiesinis pavidalas nepriimtinas. b) $\hat{\alpha} = 0,1758, \hat{\beta}_1 = 0,1531$. Esant pasikliovimo lygmeniui $Q = 0,95$ gauname $(\underline{\alpha}; \overline{\alpha}) = (0,1690; 0,1827)$, $(\underline{\beta}; \overline{\beta}) = (0,1437; 0,1626)$.

II.3.33. $\hat{\alpha}_1 = 13,2255, \hat{\alpha}_2 = 1,0158, \hat{\beta}_1 = 6,7978, \hat{\beta}_2 = 0,0798$. Esant pasikliovimo lygmeniui $Q = 0,95$ gauname $(\underline{\alpha}_1; \overline{\alpha}_1) = (13,0724; 13,3787)$, $(\underline{\alpha}_2; \overline{\alpha}_2) = (0,9915; 1,0402)$, $(\underline{\beta}_1; \overline{\beta}_1) = (6,6336; 6,9620)$, $(\underline{\beta}_2; \overline{\beta}_2) = (0,0767; 0,0825)$.

II.3.3 skyrelis

II.3.34. a) Remiantis **II.1.19** pratimu parametro $\boldsymbol{\theta}$ pilnoji ir pakankamoji statistika yra $\boldsymbol{T} = ((\boldsymbol{A}^T\boldsymbol{Y})^T, \boldsymbol{Y}^T\boldsymbol{Y})^T = (n\bar{Y}_{.}, \sum_i Y_i x_{1i}, ..., \sum_i Y_i x_{mi}, \boldsymbol{Y}^T\boldsymbol{Y})^T$.

b) Jeigu plano matricos $\boldsymbol{A}$ rangas lygus $m + 1$, tai parametro $\boldsymbol{\beta}$ NMD įvertinys yra

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y}.$$

Kadangi šis įvertinys yra normaliojo a. v. $\boldsymbol{Y}$ tiesinė funkcija, tai

$$\hat{\boldsymbol{\beta}} \sim N_{m+1}(\boldsymbol{\beta}, \sigma^2(\boldsymbol{A}^T\boldsymbol{A})^{-1}).$$

Liekamoji kvadratų suma

$$SS_E = \sum_i (Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - ... - \hat{\beta}_m x_{mi})^2 \stackrel{d}{=} \sigma^2 \chi^2_{n-m-1}$$

ir dispersijos σ^2 NMD įvertinys

$$s^2 = \frac{SS_E}{n-m-1}, \quad \mathbf{E}s^2 = \sigma^2, \quad \frac{s^2(n-m-1)}{\sigma^2} \sim \chi^2(n-m-1).$$

II.3.35. Remdamiesi sąryšiu $s^2(n-m-1)/\sigma^2 \sim \chi^2(n-m-1)$ gauname parametro σ^2 lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą

$$(\underline{\sigma^2}; \overline{\sigma^2}) = \left(\frac{s^2(n-m-1)}{\chi^2_\alpha(n-m-1)}; \frac{s^2(n-m-1)}{\chi^2_{1-\alpha}(n-m-1)} \right).$$

Kadangi $\hat{\beta}_i \sim N(\beta_i, \sigma^2 c_{ii})$, čia c_{ij} yra matricos $(\boldsymbol{A}^T\boldsymbol{A})^{-1}$ elementai, tai

$$\frac{\hat{\beta}_i - \beta_i}{s\sqrt{c_{ii}}} \sim S(n-m-1), \quad \frac{\hat{\theta} - \theta}{sb(\boldsymbol{L})} \sim S(n-m-1),$$

čia $b^2(\boldsymbol{L}) = \boldsymbol{L}^T(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{L}$, tai standartiniu būdu galime gauti parametrų pasikliovimo intervalus. Pavyzdžiui,

$$(\underline{\theta}; \overline{\theta}) = (\hat{\theta} - sb(\boldsymbol{L})t_\alpha(n-m-1); \hat{\theta} + sb(\boldsymbol{L})t_\alpha(n-m-1)).$$

II.3.36. Hipotezę $H_{j_1,...,j_k}$ galima suformuluoti pavidalu $H_{j_1,...,j_k} : \boldsymbol{H\beta} = \boldsymbol{\theta}_0$, kai matricos $\boldsymbol{H}$ rangas lygus k. Hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai

$$F_{j_1,...,j_k} = \frac{(SS_E^{(m-k)} - SS_E)(n-m-1)}{kSS_E} > F_\alpha(k, n-m-1),$$

čia $SS_E^{(m-k)}$ yra liekamoji kvadratų suma modelyje be kovariančių $(x_{j_1}, ..., x_{j_k})$.

Tikrinant hipotezę $H_j : \beta_j = 0$ apie atskiros kovariantės x_j įtakos prognozei nebuvimą, galima remiantis **II.3.36** pratimu kriterijus suformuluoti taip: hipotezė H_j atmetama reikšmingumo lygmens kriterijumi, kai

$$|t_j| = \frac{|\hat{\beta}_j|}{s\sqrt{c_{jj}}} > t_{\alpha/2}(n-m-1).$$

Hipotezės $H_{1,...,m}$ atveju

$$SS_E^{(0)} = \min_{\beta_0} \sum_i (Y_i - \beta_0)^2 = \sum_i (Y_i - \bar{Y})^2 = SS_T,$$

$SS_E^{(0)} - SS_E = SS_R$, ir $R^2 = SS_R/SS_T$ vadinamas determinacijos koeficientu. Jis yra dauginio koreliacijos koeficiento kvadrato $\rho^2_{Y(X_1,...,X_m)}$ įvertinys. Hipotezė $H_{1,...,m}$ atmetama reikšmingumo lygmens α kriterijumi, kai

$$F_{1,...,m} = \frac{SS_R(n-m-1)}{mSS_E} > F_\alpha(m, n-m-1).$$

II.3.37. Tarkime, kad regresija yra bet kokio pavidalo. Pažymėkime $\mu_i = \mathbf{E}(Y|\mathbf{X} = \boldsymbol{x}^{(i)}), i = 1, ..., k$, $\boldsymbol{\mu} = (\mu_1, ..., \mu_k)^T$. Turime tiesinį modelį

$$Y_{ij} = \mu_i + e_{ij}, \quad j = 1, ..., n_i, \quad i = 1, ..., k.$$

Parametrų μ_i DT įvertiniai $\hat{\mu}_i = \bar{Y}_{i.} = \sum_{j=1}^{n_i} Y_{ij}/n_i$, o liekamoji kvadratų suma

$$SS_E = \min_{\boldsymbol{\mu}} SS(\boldsymbol{\mu}) = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_{i.})^2 \sim \sigma^2\chi^2_{n-k}.$$

Tikrinamąją hipotezę galima suformuluoti pavidalu $H : \boldsymbol{H\mu} = \boldsymbol{\theta}_0$. Rasime matricos $\boldsymbol{H}$ išraišką. Jeigu H teisinga, tai

$$\mu_i = \beta_0 + \beta_1 x_{1i} + ... + \beta_m x_{mi}, \quad i = 1, ..., k.$$

Imdami skirtumus

$$\mu_i - \mu_1 = \beta_1(x_{1i} - x_{11}) + ... + \beta_m(x_{mi} - x_{m1}), \quad i = 2, ..., k,$$

matome, kad parametras β_0 eliminuotas ir $\mu_i, i = 2, ..., k$ tiesiškai priklauso nuo μ_1 ir $\beta_1, ..., \beta_m$. Padalinę iš $x_{1i} - x_{11}$

$$\frac{\mu_i - \mu_1}{x_{1i} - x_{11}} = \beta_1 + \beta_2\frac{x_{2i} - x_{21}}{x_{1i} - x_{11}} + ... + \beta_m\frac{x_{mi} - x_{m1}}{x_{1i} - x_{11}}, \quad i = 2, ..., k,$$

ir imdami skirtumus $(\mu_i - \mu_1)/(x_{1i} - x_{11}) - (\mu_2 - \mu_1)/(x_{1i} - x_{11})$, $i = 3, ..., k$, eliminuojame parametrą β_1 ir $\mu_i, i = 3, ..., k$ tiesiškai priklausys nuo μ_1, μ_2 ir $\beta_2, ..., \beta_m$. Tęsiant šią procedūrą po $(m+1)$-ojo žingsnio bus eliminuoti visi parametrai $\beta_0, \beta_1, ..., \beta_m$ ir $\mu_i, i = m+2, ..., k$ tiesiškai priklausys nuo parametrų $\mu_1, ..., \mu_{m+1}$:

$$\mu_i = c_{i1}\mu_1 + c_{i2}\mu_2 + ... + c_{i,m+1}\mu_{m+1}, \quad i = m+2, ..., k.$$

Imkime matricos $\boldsymbol{H}$ i-ąją eilutę tokio pavidalo:

$$(c_{i1} ... c_{i,m+1}\ 0...0\ -1\ 0...0), i = m+2, ..., k,$$

čia -1 yra i-oje pozicijoje. Matricos $\boldsymbol{H}$ rangas yra $k - m - 1$. Tikrinamoji hipotezė užrašoma pavidalu $H : \boldsymbol{H\mu} = \mathbf{0}$.

Remiantis **II.1.21** pratimu hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai

$$F = \frac{(SS_{EH} - SS_E)(n-k)}{(k-m-1)SS_E} > F_\alpha(k-m-1, n-k),$$

čia SS_{EH} yra kvadratinės formos $SS(\boldsymbol{\mu})$ sąlyginis minimumas, kai teisinga hipotezė (jis surastas **II.3.34** pratime)

$$SS_{EH} - SS_E = \sum_{i=1}^{k} n_i(\bar{Y}_{i.} - \hat{\boldsymbol{\beta}}^T \boldsymbol{x}^{(i)})^2.$$

II.3.38. Atsitiktiniai dydžiai Y_{n+1} ir $\hat{\boldsymbol{\mu}}(\boldsymbol{x})$ nepriklausomi ir normalieji. Tada

$$Y_{n+1} - \hat{\boldsymbol{\mu}}(\boldsymbol{x}) \sim N(0, \sigma^2(1 + b^2(\boldsymbol{x}))), \quad b^2(\boldsymbol{x}) = \boldsymbol{x}^T(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{x}$$

ir

$$\frac{Y_{n+1} - \hat{\boldsymbol{\mu}}(\boldsymbol{x})}{s\sqrt{1 + b^2(\boldsymbol{x})}} \sim S(n - m - 1).$$

A. d Y_{n+1} pasikliovimo lygmens $Q = 1 - 2\alpha$ prognozės intervalo rėžiai yra

$$\hat{\boldsymbol{\mu}}(\boldsymbol{x}) \mp s\sqrt{1 + b^2(\boldsymbol{x})}t_\alpha(n - m - 1).$$

II.3.39. a) Atlikus polinominę regresiją (pirmo, antro trečio, laipsnio) gauta, kad netenkinamos tiesinės regresijos prielaidos.

b) Imkime eksponentinio tipo priklausomybę $\mathbf{E}(Y|X) = \alpha e^{-\beta X}$. Pažymėję $U = \ln Y$ gausime tiesinį modelį $U_i = \beta_0 - \beta X_i + e_i, \beta_0 = \ln\alpha$. Tardami, kad paklaidos tenkina įprastines regresinės analizės sąlygas, gauname parametrų įvertinius $\hat{\beta}_0 = 4,29799; \hat{\beta} = -0,01824$. Determinacijos koeficientas $R^2 = 0,3489$, todėl gautas modelis netinkamas prognozavimui.

II.3.40. $\mathbf{E}\hat{\beta}_0 - \beta_0 = 4\beta_2$, $\mathbf{E}\hat{\beta}_1 - \beta_1 = 7\beta_3$.

II.3.41. Parametrų θ_1, θ_2 MK įvertiniai $\hat{\theta}_1 = (Y_1 - Y_3)/2$, $\hat{\theta}_2 = (Y_1 + 2Y_2 + Y_3)/6$; $\boldsymbol{V}(\hat{\theta_1}) = \sigma^2/2$, $\boldsymbol{V}(\hat{\theta_2}) = \sigma^2/6$, $\mathbf{Cov}\,(\hat{\theta_1}, \hat{\theta_2}) = 0$; $SS_E = ((Y_1 - Y_2)^2 + 2Y_3^2)^2/3 \sim \sigma^2\chi_1^2$. Tikrinama hipotezė $H : \beta = \theta_1 - 2\theta_2 = 0$; $\hat{\beta} = (Y_1 - 4Y_2 - 5Y_3)/6$, $\boldsymbol{V}$
$hat\beta = 7\sigma^2/6$; $6\hat{\beta}^2/7 \sim \sigma^2\chi_1^2$. Hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai

$$F = \frac{(Y_1 - 4Y_2 - 5Y_3)^2}{14[(Y_1 - Y_2)^2 + 2Y_3^2)]} > F_\alpha(1,\ 1).$$

II.3.42. Kadangi keturkampio kampų suma lygi 2π, tai išreiškę $\varphi_4 = 2\pi - \varphi_1 - \varphi_2 - \varphi_3$ gauname, kad modelis priklauso nuo 3 nežinomų parametrų $\varphi_1, \varphi_2, \varphi_3$. Stebėjimai turi tokią struktūrą

$$\begin{cases} Y_1 = \varphi_1 + e_1, \\ Y_2 = \varphi_2 + e_2, \\ Y_3 = \varphi_3 + e_3, \\ 2\pi - Y_4 = \varphi_1 + \varphi_2 + \varphi_3 + e_4; \end{cases}$$

$$(\boldsymbol{A}^T\boldsymbol{A})^{-1} = \frac{1}{4}\begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix}, \boldsymbol{A}^T\boldsymbol{Y} = \begin{pmatrix} Y_1 - Y_4 + 2\pi \\ Y_2 - Y_4 + 2\pi \\ Y_3 - Y_4 + 2\pi \end{pmatrix}.$$

Gauname parametrų įvertinius $\hat{\varphi}_1 = 3Y_1 - Y_2 - Y_3 - Y_4 + 2\pi$, $\hat{\varphi}_2 = -Y_1 + 3Y_2 - Y_3 - Y_4 + 2\pi$, $\hat{\varphi}_3 = -Y_1 - Y_2 + 3Y_3 - Y_4 + 2\pi$ ir liekamąją kvadratų sumą

$$SS_E = \frac{1}{4}\left(\sum_i Y_i - 2\pi\right)^2 \sim \sigma^2\chi_1^2.$$

Kai hipotezė teisinga yra, du nežinomi parametrai φ_1, φ_2, kurie susieti lygybe $2\varphi_1 + 2\varphi_2 = 2\pi$. Išreiškę $\varphi_2 = \pi - \varphi_1$ gauname, kad modelis priklauso tik nuo vieno parametro φ_1. Stebėjimai turi tokią struktūrą

$$Y_1 = \varphi_1 + e_1, \ \ \pi - Y_2 = \varphi_1 + e_2, \ \ Y_3 = \varphi_1 + e_3, \ \ \pi - Y_4 = \varphi_1 + e_4.$$

Parametro φ_1 MK įvertinys $\tilde{\varphi}_1 = (Y_1 + Y_3 - Y_2 - Y_4 + 2\pi)/4$ ir sąlyginė liekamoji kvadratų suma

$$SS_{EH} = SS_E + \frac{1}{2}((Y_1 - Y_3)^2 + (Y_2 - Y_4)^2);$$

$$SS_{EH} - SS_E = \frac{1}{2}((Y_1 - Y_3)^2 + (Y_2 - Y_4)^2) \sim \sigma^2\chi_2^2.$$

Hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai $F = [(Y_1 - Y_3)^2 + (Y_2 - Y_4)^2]/(\sum_i Y_i - 2\pi)^2 > F_\alpha(2,\ 1)$.

II.3.43. Pažymėkime $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ ir $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$. Tada

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y};$$

čia matrica $\boldsymbol{A}^T$ turi 3 eilutes ir n stulpelių; pirmoji eilutė yra $(1, 1, ..., 1)$;
antroji $(\cos(2\pi k_1/n), \cos(4\pi k_1/n), ..., \cos(2n\pi k_1/n))$; trečioji $(\sin(2\pi k_1/n), \sin(4\pi k_1/n), ..., \sin(2n\pi k_1/n))$.

II.3.44. a) $\hat{\beta}_0 = 84,55$, $\hat{\beta}_1 = 1,83$, $\hat{\beta}_2 = 2,68$; b) Gauname $F_1 = 5,89$ ir $F_2 = 8,62$; regresijos koeficientų lygybės nuliui hipotezės atmetamos; c) $99,02; 105,36$.

II.3.45. Pažingsninės regresijos metodu (žr. [3], 3.3.11 skyrelį) gauname, kad prognozuojant X_4 reikia naudotis visais kintamaisiais X_1, X_2, X_3; regresijos įvertis $\hat{X}_4 = 8,29 + 0,60X_1 - 0,14X_2 + 0,52X_3$; $r_{X_4(X_1,X_2,X_3)} = 0,8770$.

II.3.46. Parinkę pirmojo, antrojo, trečiojo ir ketvirtojo laipsnio polinomus gauname, kad netenkinamos tiesinės regresijos prielaidos.

II.3.47. $\varphi(x) = 2 - 3x^2$.

II.3.4 skyrelis

II.3.48. Žr. [3], 4.3.1 skyrelį.

II.3.49. a) Sąlyginis kvadratinės formos minimumas

$$SS_{EH_{\boldsymbol{\gamma}}} = \min_{\boldsymbol{\beta},\boldsymbol{\gamma}=\boldsymbol{0}}(\boldsymbol{Y} - \boldsymbol{A\beta} - \boldsymbol{C\gamma})^T(\boldsymbol{Y} - \boldsymbol{A\beta} - \boldsymbol{C\gamma}) = (\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}}_0)^T(\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}}_0) = R_{yy};$$

$$SS_{\boldsymbol{\gamma}} = SS_{EH_{\boldsymbol{\gamma}}} - SS_E = \hat{\gamma}_1 R_{yx_1} + ... + \hat{\gamma}_k R_{yx_k}.$$

Hipotezė $H_{\boldsymbol{\gamma}}$ atmetama reikšmingumo lygmens α kriterijumi, kai

$$F_{\boldsymbol{\gamma}} = \frac{SS_{\boldsymbol{\gamma}}(n - m - k)}{kSS_E} > F_\alpha(k, n - m - k).$$

b) Sąlyginis kvadratinės formos minimumas

$$SS_{EH_{\gamma_1,...,\gamma_l}} = \min_{\boldsymbol{\beta},\boldsymbol{\gamma}:\gamma_{i_1}=...=\gamma_{i_l}=0}(\boldsymbol{Y} - \boldsymbol{A\beta} - \boldsymbol{C\gamma})^T(\boldsymbol{Y} - \boldsymbol{A\beta} - \boldsymbol{C\gamma})$$

randamas kaip ir **II.3.48** pratime praleidus modelyje kovariantes $X_{i_1}, ..., X_{i_l}$. Tada hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai

$$F_{\gamma_1,...,\gamma_l} = \frac{(SS_{EH_{\gamma_1,...,\gamma_l}} - SS_E)(n-m-k)}{lSS_E} > F_\alpha(l, n-m-k).$$

II.3.50. Remiantis bendra tiesinių modelių teorija reikia rasti sąlyginį minimumą

$$SS_{EH} = \min_{\beta_{i_1}=...=\beta_{i_s}} (\boldsymbol{Y} - \boldsymbol{A\beta} - \boldsymbol{C\gamma})^T(\boldsymbol{Y} - \boldsymbol{A\beta} - \boldsymbol{C\gamma}).$$

Sąlyginį minimumą galima rasti kaip **II.3.48** pratime vietoje parametro $\boldsymbol{\beta}$ imant parametrą $\boldsymbol{\theta} = (\theta_1, ..., \theta_{m-s+1})$. Šis vektorius susideda iš likusių vektoriaus $\boldsymbol{\beta}$ koordinačių ir parametro θ_{m-s+1}, įrašyto vietoje $\beta_{i_1}, ..., \beta_{i_s}$.

Tada hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai

$$F_H = \frac{(SS_{EH} - SS_E)(n-m-k)}{sSS_E} > F_\alpha(s, n-m-k).$$

II.3.51. Gauname (žr. **II.3.48** pratimą)

$$R_{yy} = \sum_{i=1}^{I}\sum_{j=1}^{J_i}(Y_{ij} - \bar{Y}_{i.})^2, \quad R_{xx} = \sum_{i=1}^{I}\sum_{j=1}^{J_i}(X_{ij} - \bar{X}_{i.})^2,$$

$$R_{xy} = \sum_{i=1}^{I}\sum_{j=1}^{J_i}(Y_{ij} - \bar{Y}_{i.})(X_{ij} - \bar{X}_{i.});$$

$$\hat{\gamma} = \frac{R_{xy}}{R_{xx}}, \quad SS_E = R_{yy} - \hat{\gamma}R_{xy} = R_{yy} - \frac{R_{xy}^2}{R_{xx}} \sim \sigma^2\chi^2_{n-I-1}.$$

Tikrinant $H_\gamma : \gamma = 0$ (žr. **II.3.49** pratimą)

$$SS_\gamma = SS_{EH_\gamma} - SS_E = \frac{R_{xy}^2}{R_{xx}}.$$

Hipotezė H_γ atmetama, kai

$$F_\gamma = \frac{SS_\gamma(n-I-1)}{SS_E} = \frac{R_{xy}^2(n-I-1)}{R_{xx}R_{yy} - R_{xy}^2} > F_\alpha(1, n-I-1).$$

Tikrinant hipotezę $H_A : \alpha_1 = ... = \alpha_I = 0$ (žr. **II.3.50** pratimą) sąlyginis kvadratinės formos minimumas

$$SS_{EH_A} = \min_{\mu,\gamma}\sum_i\sum_j(Y_{ij} - \mu - \gamma x_{ij})^2.$$

Kartodami **II.3.48** pratimą (vietoje raidžių R imdami T) gausime

$$T_{yy} = \sum_i\sum_j(Y_{ij} - \bar{Y}_{..})^2 = R_{yy} + \sum_i J_i(\bar{Y}_{i.} - \bar{Y}_{..})^2 = R_{yy} + SS_A,$$

$$T_{xx} = \sum_i\sum_j(X_{ij} - \bar{X}_{..})^2, \quad T_{xy} = \sum_i\sum_j(Y_{ij} - \bar{Y}_{..})(X_{ij} - \bar{X}_{..});$$

$$\tilde{\gamma} = \frac{T_{xy}}{T_{xx}}, \quad SS_{EH_A} = T_{yy} - \tilde{\gamma}\frac{T_{xy}}{T_{xx}} = T_{yy} - \frac{T_{xy}^2}{T_{xx}};$$

$$SS_{EH_A} - SS_E = SS_A - \frac{T_{xy}^2}{T_{xx}} + \frac{R_{xy}^2}{R_{xx}}.$$

Hipotezė H_A atmetama reikšmingumo lygmens α kriterijumi, kai

$$F_A = \frac{(SS_{EH_A} - SS_E)(n-I-1)}{(I-1)SS_E} =$$

$$= \frac{(SS_A - \frac{T_{xy}^2}{T_{xx}} + \frac{R_{xy}^2}{R_{xx}})(n-I-1)}{(I-1)(R_{yy} - \frac{R_{xy}^2}{R_{xx}})} > F_\alpha(I-1, n-I-1).$$

Norėdami palyginti pateiksime hipotezės H_A tikrinimo kriterijų dispersinės analizės schemoje (neatsižvelgiant į kovariantę X): hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai

$$\frac{SS_A(n-I)}{(I-1)R_{yy}} > F_\alpha(I-1, n-I).$$

II.3.52. Pažymėkime

$$R_{0r} = \sum_i \sum_j (Y_{ij} - \bar{Y}_{i.})(X_{ij}^{(r)} - \bar{X}_{i.}^{(r)}), \; r = 1, 2;$$

$$R_{rs} = \sum_i \sum_j (X_{ij}^{(r)} - \bar{X}_{i.}^{(r)})(X_{ij}^{(s)} - \bar{X}_{i.}^{(s)}), \; r, s = 1, 2.$$

Tada a) $\hat{\gamma}_1 = (R_{01}R_{22} - R_{02}R_{12})/\Delta$, $\hat{\gamma}_2 = (R_{02}R_{11} - R_{01}R_{12})/\Delta$, $\Delta = R_{11}R_{22} - R_{12}^2$; b) $\boldsymbol{V}\hat{\gamma}_1 = \sigma^2 I(J-1)R_{22}/\Delta$, $\boldsymbol{V}\hat{\gamma}_2 = \sigma^2 I(J-1)R_{11}/\Delta$, $\mathbf{Cov}\,(\hat{\gamma}_1, \hat{\gamma}_2) = -\sigma^2 I(J-1)R_{12}/\Delta$; $\hat{\gamma}_1$ ir $\hat{\gamma}_2$ nekoreliuoti, kai $R_{12} = 0$.

II.3.53. a) Randame

$$R_{yy}(i,j) = \sum_k (Y_{ijk} - \bar{Y}_{ij.})^2, \quad R_{xx}(i,j) = \sum_k (X_{ijk} - \bar{X}_{ij.})^2,$$

$$R_{yx}(i,j) = \sum_k (Y_{ijk} - Y_{ij.})(X_{ijk} - \bar{X}_{ij.});$$

$$SS_E(i,j) = R_{yy}(i,j) - R_{yx}^2(i,j)/R_{xx}(i,j), \quad SS_E = \sum_i \sum_j SS_E(i,j) \sim \sigma^2 \chi^2_{IJ(K-2)}.$$

Kai hipotezė H teisinga, tai

$$\hat{\gamma} = R_{yx}/R_{xx}, \quad SS_{EH} = R_{yy} - R_{yx}^2/R_{xx},$$

$$R_{yy} = \sum_i \sum_j R_{yy}(i,j), \quad R_{xx} = \sum_i \sum_j R_{xx}(i,j), \quad R_{yx} = \sum_i \sum_j R_{yx}(i,j);$$

hipotezė H atmetama reikšmingumo lygmens α kriterijumi, kai

$$(SS_{EH} - SS_E)(IJ(K-2))/(SS_E(IJ-1)) > F_\alpha(IJ-1, IJ(K-2)).$$

b) Kai H teisinga, tai a. d.

$$\sqrt{R_{xx}}(\hat{\gamma}-\gamma)/\sqrt{SS_{EH}/(IJK-IJ-1)} \sim S(IJK-IJ-1).$$

II.3.54. Randame $R_{yy}=\sum_i\sum_j\sum_k(Y_{ijk}-\bar{Y}_{ij.})^2$, $R_{xx}=\sum_i\sum_j\sum_k(X_{ijk}-\bar{X}_{ij.})^2$, $R_{yx}=\sum_i\sum_j\sum_k(Y_{ijk}-Y_{ij.})(X_{ijk}-\bar{X}_{ij.})$; $\hat{\gamma}=R_{yx}/R_{xx}$, $SS_E=R_{yy}-\hat{\gamma}R_{yx}\sim \sigma^2\chi^2_{IJK-IJ-1}$.

a) Hipotezė H atmetama reikšmingumo lygmens α kriterijumi, kai $\hat{\gamma}R_{xx}(IJK-IJ-1)/SS_E>F_\alpha(1,IJK-IJ-1)$.

b) Randame $\tilde{R}_{yy}=R_{yy}+JK\sum_i(\bar{Y}_{i..}-\bar{Y}_{...})^2$, $\tilde{R}_{xx}=R_{xx}+JK\sum_i(\bar{X}_{i..}-\bar{X}_{...})^2$, $\tilde{R}_{yx}=R_{yx}+JK\sum_i(\bar{Y}_{i..}-\bar{Y}_{...})(\bar{X}_{i..}-\bar{X}_{...})$, $SS_{EH}=\tilde{R}_{yy}-\hat{\gamma}\tilde{R}_{yx}$; hipotezė H_A atmetama reikšmingumo lygmens α kriterijumi, kai $(SS_{EH}-SS_E)(IJK-IJ-1)/((I-1)SS_E)>F_\alpha(I-1,IJK-IJ-1)$.

II.3.55. Remdamiesi kovariacine analize randame $SS_E=\sum_i[R^{(i)}_{yy}-(R^{(i)}_{yx})^2/R_{xx}]$, $R^{(i)}_{yy}=\sum_j(Y_{ij}-\bar{Y}_{i.})^2$, $R^{(i)}_{yx}=\sum_j(Y_{ij}-\bar{Y}_{i.})(X_j-\bar{X}_.)$, $i=1,\,2$, $R_{xx}=\sum_j(X_j-\bar{X}_.)^2$; $SS_{EH}=R^{(1)}_{yy}+R^{(2)}_{yy}-(R^{(1)}_{yx}+R^{(2)}_{yx})^2/(2R_{xx}$. Hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai $F=(SS_{EH}-SS_E)2(J-1)/SS_E=[R^{(1)}_{yx}-R^{(2)}_{yx}]^2 2(J-1)/(2R_{xx}SS_E)>F_\alpha(1,2(J-1))$. Tikrindami dviejų regresijos tiesių lygiagretumo hipotezę remiamės statistika $T=(\hat{\gamma}_1-\hat{\gamma}_2)\sqrt{R_{xx}}/\sqrt{2SS_E/(2(J-1))}$. Nesunku patikrinti, kad $F=T^2$.

II.3.56. a) Kadangi statistikos reikšmė $F_A=14,6$, tai hipotezė atmetama kriterijumi su gana aukštu reikšmingumo lygmeniu. b) Imdami antro laipsnio polinomą gauname, kad kovariantė ir jos kvadratas yra reikšmingi, nes statistikų reikšmės atitinkamai yra 48,83 ir 53,75. Imdami trečio laipsnio polinomą gauname, kad kovariantė, jos kvadratas ir trečiasis laipsnis yra reikšmingi, nes statistikų reikšmės atitinkamai yra 35,23, 17,89 ir 8,87. Kadangi statistikos reikšmė yra 28,78, tai hipotezė apie faktoriaus įtakos nebuvimą atmetama kriterijumi su gana aukštu reikšmingumo lygmeniu.

II.3.57. a) Atlikę dispersinę analizę gauname, kad faktorius C ($F_C=1,17$; atitinkama P reikšmė yra 0,2947) ir kitų faktorių sąveikos su faktoriumi C ($F_{AC}=1,35, F_{BC}=1,09, F_{CD}=1,17$; atitinkamos P reikšmės 0,2878; 0,3609; 0,2947) nereikšmingos. Atlikus dispersinę analizę neįtraukiant faktoriaus C ir sąveikų su faktoriumi C gautos tokios statistikų realizacijos: $F_A=124,60; F_B=728,98; F_D=118,78; F_{AB}=41,75$, todėl hipotezės atmetamos su gana aukštu reikšmingumo lygmeniu. Statistikos $F_{AD}=0,87, F_{BD}=3,09$; atitinkamos P reikšmės 0,4348, 0,0657. b) Atlikus analizę, kai eliminuota kovariantės B įtaka, gauta: faktorių A ir C, A ir D, C ir D sąveikos nereikšmingos (statistikų reikšmės yra 0,17; 0,11; 0,15; atitinkamos P reikšmės 0,8469; 0,8933; 0,7059). Faktorių A ir D įtakos nebuvimo hipotezė atmetama aukšto reikšmingumo lygmens kriterijais. Faktorius C nereikšmingas (statistikos reikšmė 0,15; atitinkama P reikšmė 0,7059).

II.3.58. a) Kadangi $F_A=43,67$, tai hipotezė atmetama. b) $\hat{\beta}_0=-127,17; \hat{\beta}_1=90,827$. Hipotezė $H:\beta_1=0$ atmetama. c) Tikrinant hipotezę H_A eliminavus kovariantės X įtaką, hipotezė atmetama kriterijumi, kurio reikšmingumo lygmuo $\alpha>0,0098$.

II.3.59. a) $F_A=2,72, F_B=5,50$; atitinkamos P reikšmės 0,1139, 0,0108. b) Eliminavę kintamųjų X ir Z įtaką, gauname $F_A=13,96, F_B=5,96$; atitinkamos P reikšmės 0,0025, 0,0137. Hipotezė, kad kovariantė X nereikšminga (statistikos reikšmė

57,58) atmetama kriterijumi su aukštu reikšmingumo lygmeniu. Hipotezė, kad kovariantė Z nereikšminga (statistikos reikšmė 8,30) atmetama kriterijumi kai reikšmingumo lygmuo $\alpha > 0,0205$.

II.3.60. a) $F_A = 2,46$; P reikšmė 0,1173. b) Tikrinant hipotezę H_A eliminavus kovariantės X įtaką statistikos reikšmė yra 2,88; hipotezė atmetama kriterijumi, kurio reikšmingumo lygmuo $\alpha > 0,0896$.

II.3.61. a) $F_A = 3,08, F_B = 2,88, F_C = 1,13$; atitinkamos P reikšmės $0,0373$, $0,0773$, $0,3001$. b) $F_A = 2,58, F_B = 5,08, F_C = 5,63$; atitinkamos P reikšmės $0,0667$, $0,0159$, $0,0273$. Hipotezė, kad kovariantė X nereikšminga, atmetama kriterijumi su reikšmingumo lygmeniu $\alpha > 0,0006$.

II.3.5 skyrelis

II.3.62. $\hat\beta_0 = 1/12; \hat\beta_1 = -1/12$, $\hat\beta_2 = -29/12$; $\hat\sigma^2 = s^2 = 6,75$. Statistikos, kurios esant teisingoms hipotezėms $H_i : \beta_i = 0$ turi Fišerio skirstinius su 1 ir 9 laisvės laipsniais, įgijo reikšmes 0,012; 0,012, 10,383; atmesti parametrų β_0 ir β_1 lygybės 0 hipotezes nėra pagrindo; hipotezė $H_2 : \beta_2 = 0$ atmetama, kai kriterijaus reikšmingumo lygmuo viršija 0,01.

II.3.63. a) $\hat\beta_0 = -13/8$, $\hat\beta_1 = -19/8$, $\hat\beta_2 = 51/8$, $\hat\beta_3 = -63/8$, $\hat\sigma = s = 2,318$. Statistikos, kurios esant teisingoms hipotezėms $H_i : \beta_i = 0$ turi Fišerio skirstinius su 1 ir 4 laisvės laipsniais, įgijo reikšmes 3,930; 8,395; 60,488; 92,302; atitinkamos P reikšmės yra 0,118; 0,044; 0,0015; 0,0007. b) $\hat\beta_{12} = 5/8$, $\hat\beta_{13} = -1/8$, $\hat\beta_{23} = -11/8$, $\hat\sigma = s = 1,768$; atmesti hipotezes nėra pagrindo.

II.3.64. $\hat\beta_0 = 1727,63$, $\hat\beta_1 = 13$, $\hat\beta_2 = 40,875$, $\hat\beta_3 = 38,75$, $\hat\sigma = s = 167,138$. Hipotezės $H_i : \beta_i = 0, i = 1,2,3$, neatmetamos. $\hat\beta_{12} = -22$, $\hat\beta_{13} = 31,875$, $\hat\beta_{23} = -63,5$, $\hat\sigma = s = 165,564$; atmesti hipotezes $H_{ij} : \beta_{ij} = 0, i \neq j = 1,2,3$, nėra pagrindo.

II.3.65. a) Parametrų įverčiai: $\hat\beta_0 = 1848,59$, $\hat\beta_1 = -20,7188$, $\hat\beta_2 = -13,9063$, $\hat\beta_3 = 0,8438$, $\hat\beta_4 = 50,3438$; $\hat\beta_{12} = 26,2813$, $\hat\beta_{13} = 26,5313$, $\hat\beta_{14} = -67,4688$, $\hat\beta_{23} = -19,4063$, $\hat\beta_{24} = 16,9688$, $\hat\beta_{34} = -2,1563$, $\hat\sigma = 169,803$. b) Hipotezė $H_0 : \beta_0 = 0$ atmetama; hipotezė $H_{14} : \beta_{14} = 0$ atmetama, kai kriterijaus reikšmingumo lygmuo viršija 0,0355; kitos hipotezės neatmetamos. c) Prognozės intervalas su pasikliovimo lygmeniu $Q = 1-\alpha$ yra $\hat Y \pm \hat\sigma t_{\alpha/2}(21)\sqrt{(33+\rho^2)/32}$.

II.3.66. $\hat\beta_0 = \sum_i Y_i/9$, $\hat\beta_1 = (-Y_1 + Y_3 - Y_4 + Y_6 - Y_7 + Y_9)/6$, $\hat\beta_2 = (Y_1 + Y_2 + Y_3 - Y_7 - Y_8 - Y_9)/6$; $\hat\beta_{11} = \sum_i Y_i/3 - (Y_2 + Y_5 + Y_7)$; $\hat\beta_{11} = \sum_i Y_i/3 - (Y_4 + Y_5 + Y_6)$; $\hat\beta_{12} = \sum_i Y_i/9 - (Y_2 + Y_4 - Y_5 + Y_6 + Y_8)/3$; $\hat\beta_{211} = (Y_1 - 2Y_2 + Y_3 - Y_7 + 2Y_8 - Y_9)/3$; $\hat\beta_{122} = (-Y_1 + Y_3 + 2Y_4 - 2Y_6 - Y_7 + Y_8)/3$; $\boldsymbol{V}\hat\beta_0 = \sigma^2/9$; $\boldsymbol{V}\hat\beta_1 = \boldsymbol{V}\hat\beta_2 = \sigma^2/6$; $\boldsymbol{V}\hat\beta_{11} = \boldsymbol{V}\hat\beta_{22} = \sigma^2/2$; $\boldsymbol{V}\hat\beta_{12} = 9\sigma_2/4$; $\boldsymbol{V}\hat\beta_{211} = \boldsymbol{V}\hat\beta_{122} = 3\sigma^2/4$.

II.3.67. $a = \sqrt{\sqrt{10} - 2}$.

II.3.68. $a = 2\sqrt{\sqrt{2} - 1}$.

II.3.69. Parametrų įverčiai: $\hat\beta_0 = 2,0444$, $\hat\beta_1 = 0,4667$, $\hat\beta_2 = -0,1667$, $\hat\beta_{11} = 0,4667$, $\hat\beta_{22} = 0,8667$, $\hat\beta_{12} = 0,1361$, $\hat\beta_{211} = 0,650$, $\hat\beta_{122} = 0,625$, $\hat\sigma^2 = s^2 = 0,8066$. Hipotezė $H_0 : \beta_0 = 0$ atmetama: P reikšmė $pv = 2 \times 10^{-6}$. Hipotezei $H_{22} : \beta_{22} = 0$ P reikšmė

$pv = 0,082$; atmesti kitas hipotezes nėra pagrindo.

II.3.70. Abiem atvejais pirmoji replika ((1), ab, ac, bc).

II.3.6 skyrelis

II.3.71. a) $\hat{\beta}_0 = 3,2975, \hat{\beta}_1 = 0,0023, \hat{\beta}_2 = 0,0188, \hat{\beta}_3 = -0,0437$;

$$\hat{\boldsymbol{V}}(\hat{\boldsymbol{\beta}}) = 10^{-3} \begin{pmatrix} 1,198 & 0 & -0,168 & 0,021 \\ 0 & 0,115 & 0 & 0 \\ -0,168 & 0 & 0,040 & 0 \\ 0,021 & 0 & 0 & 0,245 \end{pmatrix}.$$

b) Tikrinant hipotezes $H_j : \beta_j = 0, j = 1,2,3$, statistikos (žr. [3], 5.2.1 skyrelį) įgijo reikšmes 0,05; 8,89; 7,81; atitinkamos P reikšmės yra 0,8299; 0,0029; 0,0052.

c) $\hat{\mu}_{13} = e^{\hat{\beta}_0 - 3\hat{\beta}_1 + 9\hat{\beta}_2} = 31,80$; $\hat{\mu}_{43} = e^{\hat{\beta}_0} = 27,04$;

$$(\underline{\mu}_{13}; \overline{\mu}_{13}) = (28,86; 35,04); \quad (\underline{\mu}_{43}; \overline{\mu}_{43}) = (25,27; 28,94).$$

Taškinis įvertinys $\hat{\mu}_{ij} = e^{\hat{\theta}_{ij}}, \hat{\theta}_{ij} = \hat{\boldsymbol{\beta}}^T \boldsymbol{Z}_{ij}$. Tiesinio darinio $\theta_{ij} = \boldsymbol{\beta}^T \boldsymbol{Z}_{ij}$ aproksimacinį lygmens $Q = 1 - 2\alpha$ pasikliovimo intervalą gauname aproksimuodami statistikos $(\hat{\theta}_{ij} - \theta_{ij})/\sqrt{\hat{\boldsymbol{V}}(\hat{\theta}_{ij})}$, $\hat{\boldsymbol{V}}(\hat{\theta}_{ij}) = \boldsymbol{Z}_{ij}^T \hat{\boldsymbol{V}}(\hat{\boldsymbol{\beta}}) \boldsymbol{Z}_{ij}$, skirstinį standartiniu normaliuoju skirstiniu:

$$(\underline{\theta}_{ij}; \overline{\theta}_{ij}) = (\hat{\theta}_{ij} - z_\alpha \sqrt{\hat{\boldsymbol{V}}(\hat{\theta}_{ij})}; \hat{\theta}_{ij} + z_\alpha \sqrt{\hat{\boldsymbol{V}}(\hat{\theta}_{ij})}).$$

Tada $\underline{\mu}_{ij} = e^{\underline{\theta}_{ij}}, \quad \overline{\mu}_{ij} = e^{\overline{\theta}_{ij}}$.

II.3.72. a) $\hat{\alpha}_0 = 27,043, \hat{\alpha}_1 = 0,0714, \hat{\alpha}_2 = 0,5571, \hat{\alpha}_3 = -1,279$; $\hat{\sigma}^2 = s^2 = SS_E/(n-4) = 37,891$;

$$\hat{\boldsymbol{V}}(\hat{\boldsymbol{\beta}}) = \begin{pmatrix} 1,2630 & 0 & -0,1804 & 0 \\ 0 & 0,1353 & 0 & 0 \\ -0,1804 & 0 & 0,0451 & 0 \\ 0 & 0 & 0 & 0,2707 \end{pmatrix}.$$

b) Tikrinant hipotezes $H_j : \alpha_j = 0, j = 1,2,3$, statistika įgijo reikšmes $0,19$; $2,62$; $-2,46$; atitinkamos P reikšmės yra 0,8466; 0,0108; 0,0166.

c) $\hat{\mu}_{13} = \hat{\alpha}_0 - 3\hat{\alpha}_1 + 9\hat{\alpha}_2 = 31,84$; $\hat{\mu}_{43} = \hat{\alpha}_0 = 27,04$;

$$(\underline{\mu}_{13}; \overline{\mu}_{13}) = (28,45; 35,24); \quad (\underline{\mu}_{43}; \overline{\mu}_{43}) = (24,80; 29,29).$$

II.3.73. a) $\hat{\gamma}_0 = 10,31, \hat{\gamma}_1 = 0,0141, \hat{\gamma}_2 = 0,0998, \hat{\gamma}_3 = -0,2409; \hat{\sigma}^2 = s^2 = SS_E/(n-4) = 1,295$;

$$\hat{\boldsymbol{V}}(\hat{\boldsymbol{\beta}}) = \begin{pmatrix} 0,04317 & 0 & -0,00617 & 0 \\ 0 & 0,00463 & 0 & 0 \\ -0,00617 & 0 & 0,00154 & 0 \\ 0 & 0 & 0 & 0,00925 \end{pmatrix}.$$

b) Tikrinant hipotezes $H_j : \gamma_j = 0, j = 1,2,3$, statistika įgijo reikšmes $0,21$; $2,54$; $-2,50$; atitinkamos P reikšmės yra 0,8362; 0,0134; 0,0147.

c) $\hat{\mu}_{13} = 31,14$; $\hat{\mu}_{43} = 26,57$;

$$(\underline{\mu}_{13}; \overline{\mu}_{13}) = (27,77; 34,75); \quad (\underline{\mu}_{43}; \overline{\mu}_{43}) = (24,45; 28,73).$$

II.3.74. a) $\hat{\beta}_0 = 1,897$, $\hat{\beta}_1 = 0,8193$, $\hat{\beta}_2 = 0,0773$, $\hat{\beta}_3 = 0,0788$; $\boldsymbol{V}(\hat{\beta}_0) = 1/60$, $\boldsymbol{V}(\hat{\beta}_j) = 1/30$, $j = 1,2,3$; $\mathbf{Cov}\,(\hat{\beta}_0,\ \hat{\beta}_j) = -1/60$, $j = 1,\ 2,\ 3$; $\mathbf{Cov}\,(\hat{\beta}_j,\ \hat{\beta}_l) = 1/60$, $j \neq l = 1,\ 2,\ 3$.

b) Tikrinant hipotezę $H : \beta_1 = \beta_2 = \beta_3 = 0$, tikėtinumų santykio statistika

$$D_R = -120(8\ln 2 + \ln(T_1T_2T_3T_4/T^4)),$$

čia $T_j = \sum_i Y_{ij}, j = 1,2,3,4$; $T = T_1 + T_2 + T_3 + T_4 = 907,83 + 432,27 + 432,92 + 400,11 = 2173,13$; įgijo reikšmę 29,77; P reikšmė yra $1,54 \times 10^{-6}$. Tikrinant hipotezes $H_j : \beta_j = 0$, statistikos $30\hat{\beta}_j^2$ įgijo reikšmes 20,14, 0,179, 0,186; atitinkamos P reikšmės $7,20 \times 10^{-6}$, $0,6720$, $0,6660$. Tikrinant hipotezę $H_{23} : \beta_2 = \beta_3$ parametro $\theta = \beta_2 - \beta_3$ įvertinio dispersija $\boldsymbol{V}(\hat{\theta}) = \boldsymbol{V}(\hat{\beta}_2 - \hat{\beta}_3) = 1/30$; statistika $30\hat{\theta}^2 = 6,8 \times 10^{-5}$; P reikšmė 0,9934.

c) $\hat{\mu}_3 = T_3/20 = 21,65$; $(\underline{\mu_3}; \bar{\mu_3}) = (17,07;\ 28,37)$.

II.3.75. a) $\hat{\beta}_0 = -12,371, \hat{\beta}_1 = 1,459, \hat{\beta}_2 = 0,590$. Tikrinant hipotezes $H_j : \beta_j = 0,\ \ j = 1,\ 2$ statistikų W_j (žr. [3], 5.3.6 skyrelį) kvadratai įgijo reikšmes 1,125, 4,978; atitinkamos P reikšmės 0,2889, 0,0257.

b) $\exp(\hat{\beta}_1) = 4,302$; $\exp(\hat{\beta}_2) = 1,803$.

c) $\pi(X_2 = 20, X_1 = 0) = 0,3593$, $\pi(X_2 = 20, X_1 = 1) = 0,7070$.

d) $V_{11}(0,5) = 9$, $V_{01}(0,5) = V_{10}(0,5) = 4$, $V_{00}(0,5) = 7$.

II.3.76. a) $\hat{\beta}_0 = 6,549, \hat{\beta}_1 = -0,303, \hat{\beta}_2 = -0,327, \hat{\beta}_3 = 1,741, \hat{\beta}_4 = -0,009$.

b) Tikrinant hipotezes $H_j : \beta_j = 0,\ \ j = 1,\ 2,\ 3,\ 4$ statistikų W_j (žr. [3], 5.3.6 skyrelį) kvadratai įgijo reikšmes 4,987, 0,082, 1,372, 0,071; atitinkamos P reikšmės 0,0255, 0,7742, 0,2415, 0,7899.

II.3.77. $\hat{\beta}_0 = -1,721, \hat{\beta}_1 = 0,00117, \hat{\beta}_2 = -0,0199$. Tikrinant hipotezes $H_j : \beta_j = 0,\ \ j = 1,\ 2$ statistikų W_j (žr. [3], 5.3.6 skyrelį) kvadratai įgijo reikšmes 3,211, 0,921; atitinkamos P reikšmės 0,0731, 0,3372.

II.3.78. a) Kai $Z = 0$, gauname $\hat{\beta}_0 = 3,95, \hat{\beta}_1 = -0,112; \exp(\hat{\beta}_1) = 0,894$; tikrinant hipotezę $H_1 : \beta_1 = 0$ statistikos W_1 (žr. [3], 5.3.6 skyrelį) kvadratas įgijo reikšmę 102,6; hipotezė atmetama. Kai $Z = 1$, gauname $\hat{\beta}_0 = 3,33, \hat{\beta}_1 = -0,091; \exp(\hat{\beta}_1) = 0,913$; tikrinant hipotezę $H_1 : \beta_1 = 0$ statistikos W_1 kvadratas įgijo reikšmę 58,8; hipotezė atmetama.

b) $\hat{\beta}_0 = 3,63, \hat{\beta}_1 = -0,103$ (atstumas), $\hat{\beta}_2 = 0,1036$ (lyga). Tikrinant hipotezę $H_1 : \beta_1 = 0$ statistikos W_1 (žr. [3], 5.3.6 skyrelį) kvadratas įgijo reikšmę 161,6; hipotezė atmetama. Tikrinant hipotezę $H_2 : \beta_2 = 0$ statistikos W_2 kvadratas įgijo reikšmę 0,372; atitinkama P reikšmė 0,5419; atmesti hipotezę nėra pagrindo.

III. Neparametrinė statistika

III.1. Chi kvadrato kriterijus

III.1.1. Paprastoji hipotezė

III.1.1. Mendelis stebėjo, kokios žirnių sėklos gaunamos kryžminant augalus, kurių sėklos geltonos ir apvalios, su augalais, kurių sėklos žalios ir raukšlėtos. Rezultatai pateikti lentelėje kartu su teorinėmis tikimybėmis, apskaičiuotomis remiantis Mendelio paveldimumo teorija.

Sėklos	Dažnumai	Tikimybės
Geltonos ir apvalios	315	9/16
Geltonos ir raukšlėtos	101	3/16
Žalios ir apvalios	108	3/16
Žalios ir raukšlėtos	32	1/16
Σ	556	1

Ar stebėjimo duomenys neprieštarauja Mendelio paveldimumo teorijai?

III.1.2. Kryžminant du kukurūzų tipus gauti keturi skirtingi augalų tipai. Pagal paprastąją Mendelio paveldimumo teoriją šie tipai turėtų pasirodyti su tikimybėmis $9/16, 3/16, 3/16$ ir $1/16$. Stebint 1301 augalą gauti tokie dažniai 773, 231, 238 ir 59. Su kokiu reikšmingumo lygmeniu χ^2 kriterijus neprieštarauja Mendelio modeliui?

III.1.3. Skaitmenys $0,\ 1,\ 2, ..., 9$ tarp pirmųjų 800 skaičiaus π ženklų kartojasi atitinkamai 74, 92, 83, 79, 80, 73, 77, 75, 76, 91 kartą. Ar galima šiuos duomenis interpretuoti kaip a. v. $\boldsymbol{U} \sim \mathcal{P}_{10}(800, \boldsymbol{\pi})$, $\boldsymbol{\pi} = (1/10, ..., 1/10)^T$ realizaciją?

III.1.4. Tikrinama hipotezė apie atsitiktinių skaičių lentelės korektiškumą, t. y. hipotezė, kad lentelėje skaitmenys 0, 1, 2, ..., 9 pasitaiko su vienodomis tikimybėmis $p = 0,1$. Hipotezė tikrinama χ^2 kriterijumi. Koks turi būti imties didumas, kad ta hipotezė būtų atmesta su tikimybe, ne mažesne už 0,95, jei žinoma, kad 5 skaitmenys lentelėje pasirodo su tikimybėmis 0,11, o kiti 5 – su tikimybėmis 0,09 (kriterijaus reikšmingumo lygmuo yra 0,05)?

III.1.5. Nuskaitydami prietaiso skalės parodymus, kai paskutinis skaitmuo įvertinamas iš akies, stebėtojai kartais nesąmoningai suteikia pirmenybę kai kuriems skaičiams. Lentelėje pateikti paskutiniųjų skaitmenų dažniai tam tikram stebėtojui atlikus 200 stebėjimų.

Skaitmuo	0	1	2	3	4	5	6	7	8	9
Dažnis	35	16	15	17	17	19	11	16	30	24

Iš lentelės matome, kad skaitmenys 0 ir 8 pasirodo šiek tiek dažniau, palyginti su kitais. Ar galima daryti išvadą, kad stebėtojas suteikia pirmenybę kai kuriems skaičiams?

III.1.6. Įrodykite, kad tikrinant hipotezę apie polinominio skirstinio $(U_1, ..., U_k)^T \sim \mathcal{P}_k(n, \boldsymbol{\pi})$ tikimybių vektoriaus $\boldsymbol{\pi} = (\pi_1, ..., \pi_k)^T$ reikšmę: $H_0 : \boldsymbol{\pi} = \boldsymbol{\pi}_0$; čia $\boldsymbol{\pi}_0 = (\pi_{10}, ..., \pi_{k0})^T$, $0 < \pi_{j0} < 1$, $\pi_{10} + ... + \pi_{k0} = 1$ žinomas vektorius, statistika

$$\tilde{X}_n^2 = n \sum_{i=1}^{k} (1 - \pi_{i0})[H(U_i/n) - H(\pi_{i0})]^2$$

asimptotiškai ($n \to \infty$) pasiskirsčiusi pagal χ kvadrato skirstinį su $k - 1$ laisvės laipsnių; čia $H(x) = \arcsin(2x - 1)$.

III.1.7. Raskite Pirsono statistikos

$$X_n^2 = \sum_{i=1}^{k} \frac{(U_i - n\pi_{i0})^2}{n\pi_{i0}} = \sum_{i=1}^{k} \frac{U_i^2}{n\pi_{i0} - n}$$

pirmuosius du momentus, kai tikrinamoji hipotezė: a) teisinga; b) neteisinga.

III.1.8. (**III.1.7** pratimo tęsinys). Įrodykite, kad jei tikimybės $\pi_{i0} = 1/k$, tai X_n^2 dispersija yra $\boldsymbol{V}(X_n^2) = 2k^2\{2(n-2)\sum_i \pi_i^3 - (2n-3)(\sum_i \pi_i^2)^2 + \sum_i \pi_i^2\}$, o jeigu ir $\pi_i = \pi_{i0} = 1/k$, tai $\boldsymbol{V}(X_n^2) = 2(k-1)$.

III.1.9. Kompiuteriu sugeneruota $n = 80$ atsitiktinių skaičių. Gautieji rezultatai:

0,0100 0,0150 0,0155 0,0310 0,0419 0,0456 0,0880 0,1200 0,1229 0,1279 0,1444 0,1456 0,1621 0,1672 0,1809 0,1855 0,1882 0,1917 0,2277 0,2442 0,2456 0,2476 0,2538 0,2552 0,2681 0,3041 0,3128 0,3810 0,3832 0,3969 0,4050 0,4182 0,4259 0,4365 0,4378 0,4434 0,4482 0,4515 0,4628 0,4637 0,4668 0,4773 0,4799 0,5100 0,5309 0,5391 0,6033 0,6283 0,6468 0,6519 0,6686 0,6689 0,6865 0,6961 0,7058 0,7305 0,7337 0,7339 0,7440 0,7485 0,7516 0,7607 0,7679 0,7765 0,7846 0,8153 0,8445 0,8654 0,8700 0,8732 0,8847 0,8935 0,8987 0,9070 0,9284 0,9308 0,9464 0,9658 0,9728 0,9872

Ar šie duomenys neprieštarauja prielaidai, kad tai yra paprastosios imties, gautos stebint a. d. $X \sim U(0, 1)$, realizacija?

III.1.2. Sudėtinės hipotezės

III.1.10. Per 8000 bandymų nesutaikomi, sudarantys pilną įvykių grupę įvykiai A, B ir C pasirodė 2014, 5012 ir 974 kartus. Tarkime, šių įvykių pasirodymo tikimybės yra $\pi_1, \pi_2, \pi_3, \pi_1 + \pi_2 + \pi_3 = 1$. χ^2 kriterijumi patikrinkite hipotezę, kad įvykių pasirodymo tikimybės yra $\pi_1 = 0,5 - 2\alpha$, $\pi_2 = 0,5 + \alpha$, $\pi_3 = \alpha$, $0 < \alpha < 0,25$.

III.1.11. Tarp 2020 šeimų buvo užregistruota 527 šeimos, kuriose abu vaikai berniukai, 476 šeimos, kuriose abu vaikai mergaitės, ir 1017 šeimų, kuriose vaikai skirtingų lyčių. Ar galima tvirtinti, kad berniukų skaičius šeimose su dviem vaikais yra a) binominis a. d.; b) binominis a. d., kai berniuko ir mergaitės gimimo tikimybės vienodos.

III.1.12. Diskretaus a. d. penkios nepriklausomos realizacijos yra 47, 46, 49, 53 ir 50. Ar galima tvirtinti, kad buvo stebimas Puasono a. d.?

III.1.13. Rezerfordo ir Geigerio bandymuose buvo registruojamas radioaktyvios medžiagos per 2608 ilgio 7,5 sek. periodus išspinduliuotų α dalelių skaičius. Rezultatai pateikti lentelėje (i – išspinduliuotų dalelių skaičius, V_i – periodų, per kuriuos buvo stebima i dalelių, skaičius).

i	0	1	2	3	4	5	6	7	8	9	10	11	12
V_i	57	203	383	525	532	408	273	139	45	27	10	4	2

Ar neprieštarauja gauti duomenys prielaidai, kad per vieną periodą išspinduliuotų dalelių skaičius turi Puasono skirstinį?

III.1.14. Kontroliniu prietaisu buvo išmatuotas atstumas r (mikronais) nuo detalės svorio centro iki jos išorinio cilindro ašies. Matavimo rezultatai pateikti lentelėje (r_i – reikšmės, n_i – dažniai).

r_i	n_i	r_i	n_i
0 – 16	40	80 – 96	45
16 – 32	129	96 – 112	19
32 – 48	140	112 – 128	8
48 – 64	126	128 – 144	3
64 – 80	91	144 – 160	1

Remdamiesi χ^2 kriterijumi, patikrinkite, ar stebėjimo rezultatai neprieštarauja prielaidai, kad stebimi atstumai pasiskirstę pagal Relėjaus dėsnį.

III.1.15. Nustatant 200 elektros lempučių degimo laiką T, gauti rezultatai pateikti lentelėje ($(a_{i-1}, a_i]$ – degimo laiko intervalai, n_i – dažniai).

$a_{i-1} - -a_i$	n_i	$a_{i-1} - -a_i$	n_i
0 – 300	53	1800 – 2100	9
300 – 600	41	2100 – 2400	7
600 – 900	30	2400 – 2700	5
900 – 1200	22	2700 – 3000	3
1200 – 1500	16	3000 – 3300	2
1500 – 1800	12	3300 – 3600	0

Remdamiesi χ^2 kriterijumi, patikrinkite, ar stebėjimo rezultatai neprieštarauja prielaidai, kad lemputės degimo laikas pasiskirstęs pagal eksponentinį dėsnį.

III.1.16. Lentelėje pateikti prapuolimo kampai 209 pašto balandžių, kai atliekant bandymą buvo bandoma paveikti jų „vidinį laikrodį“ (žr. [14]).

Kryptis	Dažnis	Kryptis	Dažnis
0°–	26	180°–	14
30°–	22	210°–	11
60°–	26	240°–	12
90°–	30	270°–	5
120°–	29	300°–	5
150°–	18	330°–	11

Duomenys sugrupuoti į 30° ilgio intervalus. Lentelėje nurodyti kampai φ_i, atitinkantys i-ojo intervalo pradžią, ir patekusių į i-ąjį intervalą dažniai n_i, $i = 1, ...12$. Patikrinkite hipotezę, kad turimi duomenys neprieštarauja prielaidai, jog prapuolimo kampas turi Mizeso skirstinį $M(\mu, \theta)$.

III.1.17. Lentelėje pateikta smėlio grūdelių orientacija plokštumoje (žr. [14]).

Kampas	Kiekis	Kampas	Kiekis	Kampas	Kiekis
$0°-$	244	$60°-$	326	$120°-$	322
$10°-$	262	$70°-$	340	$130°-$	295
$20°-$	246	$80°-$	371	$140°-$	230
$30°-$	290	$90°-$	401	$150°-$	256
$40°-$	284	$100°-$	382	$160°-$	263
$50°-$	314	$110°-$	332	$170°-$	281

Kampai sugrupuoti į $10°$ ilgio intervalus (nurodoma grupavimo intervalo pradžia). Gretimuose stulpeliuose nurodomi smėlio grūdelių, kurių orientacija patenka į atitinkamus intervalus, skaičiai.

Padvigubinę kampus perveskite duomenis į intervalą $[0°-360°]$. Patikrinkite hipotezę, kad stebėtas atsitiktinis kampas turi Mizeso skirstinį $M(\mu, \theta)$.

III.1.18. Didumo $n = 100$ imties realizacija pateikta lentelėje.

338	336	312	322	381	302	296	360	342	334
348	304	323	310	368	341	298	312	322	350
304	302	336	334	304	292	324	331	324	334
314	338	324	292	298	342	338	331	325	324
326	314	312	362	368	321	352	304	302	332
314	304	312	381	290	322	326	316	328	340
324	320	364	304	340	290	318	332	354	324
304	321	356	366	328	332	304	282	330	314
342	322	362	298	316	298	332	342	316	326
308	321	302	304	322	296	322	338	324	323

Modifikuotuoju χ^2 kriterijumi (grupavimo intervalų skaičius $k = 8$) patikrinkite hipotezę, kad buvo stebimas normalusis atsitiktinis dydis.

III.1.19. Lentelėje pateikti duomenys, apibūdinantys tam tikro elemento koncentraciją nesureagavusiame likutyje pasibaigus cheminiam procesui.

10	51	8	47	8	5	56	12	4	5	4	4	7	6	9
30	25	12	3	22	5	15	4	4	29	15	4	2	18	41
3	5	54	110	24	16	2	37	20	2	6	7	16	2	14
68	10	16	11	78	6	17	7	11	21	15	24	6	32	8
11	4	14	45	17	10	15	20	4	65	10	3	5	11	13
35	11	34	3	4	12	7	6	62	13	36	26	6	11	6
13	1	4	36	18	10	37	28	4	12	31	14	3	11	6
4	10	38	6	11	24	9	4	5	8	135	22	6	18	49
17	9	32	27	2	12	8	93	3	9	10	3	14	33	72
14	4	9	10	19	2	5	21	8	25	30	20	12	19	16

Modifikuotuoju χ^2 kriterijumi (grupavimo intervalų skaičius $k = 10$) patikrinkite hipotezę, kad buvo stebimas lognormalusis atsitiktinis dydis.

III.1.20. Modifikuotuoju chi kvadrato kriterijumi patikrinkite hipotezę, kad pateikti $n = 100$ skaičių yra normaliojo a. d. realizacija.

24 41 30 37 25 32 28 35 28 51 36 26 43 25 27 39 21 45 39 25 29 43 66 25 24 56 29 31 41 41 36 57 36 48 25 36 48 24 48 22 40 7 31 24 32 53 33 46 22 33 25 37 34 32 41 36 19 32 25 19 19 37 20 21 48 44 35 19 44 34 29 48 38 43 48 35 42 37 35 36 58 45 34 40 37 21 41 11 41 27 50 24 37 39 33 45 39 43 21 34

Pakoreguokite kriterijų atsižvelgdami į tai, kad duomenys suapvalinti.

III.1.21. Patikrinkite hipotezę, kad tam tikro dalyko pažymiai (penkiabalėje sistemoje) atestate ir per stojamuosius egzaminus yra nepriklausomi. Duomenys pateikti lentelėje (x_i – pažymys atestate, y_j – pažymys per stojamuosius egzaminus).

$x_i y_j$	5	4	3	2	$\sum$
4 – 5	110	70	60	10	250
3	0	10	10	30	50
$\sum$	110	80	70	40	300

III.1.22. Paleidus raketą 87 kartus, buvo gauti tokie duomenys apie atstumą X(m) ir nukrypimą Y(kampo minutės).

$x_i \backslash y_j$	$(-250, -50)$	$(-50, 50)$	$(50, 250)$	Σ
0 – 1200	5	9	7	21
1200 – 1800	7	5	5	21
1800 – 2700	8	21	16	45
Σ	20	35	32	87

Ar šie požymiai nepriklausomi?

III.1.23. Viename sraute iš 300 stojančiųjų pažymius „nepatenkinamai“, „patenkinamai“, „gerai“ ir „labai gerai“ gavo atitinkamai 33, 43, 80 ir 144; kito srauto stojantieji atitinkamai 39, 35, 72 ir 154. Ar galima laikyti, kad abiejų srautų stojantieji pasirengę vienodai?

III.1.24. Tiriant granulometrinę kvarco sudėtį Anykščių ir Afrikos smėlio pavyzdžiuose, buvo gauta duomenų apie jo grūdelių didžiosios ašies ilgį. Remiantis pavyzdžių granulometrine sudėtimi daromos tam tikros išvados apie geologines smėlio susidarymo sąlygas. Pateikiami duomenys sugrupuoti vienodo ilgio intervalais (X_i – i-ojo intervalo vidurys).

X_i	9	13	17	21	25	29	33	37	41	45	49	Σ
Anykščių smėlis	4	12	35	61	52	23	7	4	2	1	0	201
Afrikos smėlis	0	6	10	12	13	12	15	12	11	7	4	102

Remdamiesi χ^2 kriterijumi patikrinkite hipotezę, kad grūdelių didžiosios ašies ilgio skirstinys vienodas Anykščių ir Afrikos smėlio pavyzdžiuose.

III.1.25. Dviejose nepriklausomose didumo 500 imtyse buvo užregistruota laikrodžių, išstatytų įvairių taisyklų vitrinose, rodmenys.

Duomenys sugrupuoti į 12 intervalų (0 reiškia intervalą nuo 0 h iki 1 h; 1 – intervalą nuo 1 h iki 2 h ir t. t.) ir surašyti į lentelę.

Imtis	1	2	3	4	5	6	7	8	9	10	11	12	Σ
1	41	34	54	39	49	45	41	33	37	41	47	39	500
2	36	47	41	47	49	45	32	37	40	41	37	48	500

Remdamiesi χ^2 kriterijumi patikrinkite hipotezę, kad abiejose imtyse laikrodžių rodmenų patekimo į visus intervalus tikimybės yra vienodos.

III.1.26. Ląsteles veikiant rentgeno spinduliais, jose keičiasi kai kurios chromosomos. Lentelėje pateikti kelių nepriklausomų bandymų serijų duomenys (i – chromosomų pasikeitimų skaičius, n_{ik} – ląstelių su i pasikeitimų k-ajame eksperimente skaičius).

i	0	1	2	≥ 3	$\sum n_{ik}$
n_{i1}	280	75	12	1	368
n_{i2}	593	143	20	3	759
n_{i3}	639	141	13	0	793
n_{14}	359	109	13	1	482

Patikrinkite hipotezę, kad visos 4 imtys gautos stebint atsitiktinius dydžius, kurių skirstiniai yra a) Puasono; b) tie patys Puasono.

III.1.3. Sprendimai, atsakymai, nurodymai

III.1.1 skyrelis

III.1.1. Reikia patikrinti hipotezę apie polinominio skirstinio

$$(U_1, ..., U_k)^T \sim \mathcal{P}_k(n, \boldsymbol{\pi})$$

tikimybių vektoriaus $\boldsymbol{\pi} = (\pi_1, ..., \pi_k)^T$ reikšmę: $H_0 : \boldsymbol{\pi} = \boldsymbol{\pi}_0$; čia $\boldsymbol{\pi}_0 = (\pi_{10}, ..., \pi_{k0})^T$, $0 < \pi_{j0} < 1$, $\pi_{10} + ... + \pi_{k0} = 1$ žinomas vektorius.

Tikėtinumų santykio statistika

$$\Lambda_n = \frac{\max_{\boldsymbol{\pi}=\boldsymbol{\pi}_0} L(\boldsymbol{\pi})}{\max_{\boldsymbol{\pi}} L(\boldsymbol{\pi})} = \frac{L(\boldsymbol{\pi}_0)}{L(\hat{\boldsymbol{\pi}})} = \prod_{i=1}^{k} (n\pi_{i0}/U_i)^{U_i}.$$

Kai hipotezė H_0 teisinga ir $n \to \infty$, tai (žr. [4], 7.1.2 pastabą)

$$R_n = -2\ln(\Lambda_n) = 2\sum_{i=1}^{k} U_i \ln(U_i/(n\pi_{i0})) \xrightarrow{d} V \sim \chi^2(k-1).$$

Pirsono statistika X_n^2 irgi turi tą patį asimptotinį skirstinį (žr. [4], 2.1.1 teoremą)

$$X_n^2 = \sum_{i=1}^{k} \frac{(U_i - n\pi_{i0})^2}{n\pi_{i0}} = \sum_{i=1}^{k} \frac{U_i^2}{n\pi_{i0}} - n \xrightarrow{d} \chi^2_{k-1}.$$

Remdamiesi šiais sąryšiais gauname tokius asimptotinius kriterijus. Hipotezė H_0 atmetama asimptotiniu reikšmingumo lygmens α tikėtinumo santykio kriterijumi, kai

$$R_n > \chi^2_\alpha(k-1).$$

Hipotezė H_0 atmetama asimptotiniu reikšmingumo lygmens α Pirsono χ kvadrato kriterijumi, kai

$$X_n^2 > \chi^2_\alpha(k-1).$$

Asimptotinių P reikšmių terminais hipotezė atmetama, kai atitinkamai

$$pv_a = \mathbf{P}\{\chi^2_{k-1} > r_n\}, \quad \text{arba} \quad pv_a = \mathbf{P}\{\chi^2_{k-1} > x^2_n\},$$

čia r_n ir x^2_n yra statistikų R_n ir X^2_n realizacijos. Pagal turimus duomenis statistikos R_n ir X^2_n įgijo reikšmes 0,475 ir 0,47; atitinkamos asimptotinės P reikšmės yra 0,9243 ir 0,9254; duomenys neprieštarauja iškeltajai hipotezei.

III.1.2. Statistika X^2_n įgijo reikšmę 9,2714 ir $pv_a = \mathbf{P}\{\chi^2_3 > 9,2714\} = 0,0259$; hipotezė atmetama, jei kriterijaus reikšmingumo lygmuo viršija 0,0259.

III.1.3. Statistika X^2_n įgijo reikšmę 5,125 ir $pv_a = \mathbf{P}\{\chi^2_9 > 5,125\} = 0,8233$; duomenys neprieštarauja suformuluotai prielaidai.

III.1.4. Jeigu hipotezė H_0 neteisinga, tai Pirsono statistikos X^2_n iš **III.1.1** pratimo skirstinys aproksimuojamas necentriniu χ^2 skirstiniu su $k-1$ laisvės laipsnių ir necentriškumo parametru

$$\delta_n = n \sum_{i=1}^{k} \frac{(\pi_i - \pi_{i0})^2}{\pi_{i0}}$$

(žr. [4], 2.1.5 pastabą).

Pagal pratimo sąlygą $\delta_n = 0,01\, n$. Apytiksliam imties didumui rasti gauname nelygybę

$$\mathbf{P}\{\chi^2_{9;\delta_n} > \chi^2_{0,05}(9)\} \geq 0,95 \Leftrightarrow n \geq 881.$$

III.1.5. a) Tikrinant hipotezę $H : \pi_i = 1/10,\ i = 1, ..., 10$ statistika X^2_n įgijo reikšmę 24,9 ir $pv_a = \mathbf{P}\{\chi^2_9 > 24,9\} = 0,0031$; hipotezė atmestina. b) Atlikdami tolesnę analizę patikrinkime hipotezę, kad skaitmenų 0 arba 8 pasirodymo tikimybė 0,2. Esant teisingai hipotezei $S = U_1 + U_8 \sim B(n, 0,2)$ ir įgijo reikšmę 65. Taigi $pv = \mathbf{P}\{S \geq 65\} = 0,00002$. Išvada: stebėtojas suteikia pirmenybę skaitmenims 0 ir 8.

III.1.6. Remdamiesi delta metodu įrodykite, kad a. v.

$$\sqrt{n}(\sqrt{1-\pi_{10}}(H(U_1/n) - H(\pi_{10})), ..., \sqrt{1-\pi_{k0}}(H(U_k/n) - H(\pi_{k0})))^T$$

yra asimptotiškai normalusis $N_k(\mathbf{0}, \mathbf{\Sigma})$ su ta pačia kovariacine matrica $\mathbf{\Sigma}$, kaip ir [4], 2.1.1 teoremoje.

III.1.7. a) Gauname

$$\mathbf{E}(X^2_n) = \sum_{i=1}^{k} \frac{\mathbf{E}(U_i^2)}{n\pi_{i0}} - n = \sum_{i=1}^{k} \frac{n\pi_{i0} + n(n-1)\pi^2_{i0}}{n\pi_{i0}} - n = k - 1;$$

$$\boldsymbol{V}(X^2_n) = \mathbf{E}(X^4_n) - (\mathbf{E}(X^2_n)) = \mathbf{E}\left(\sum_{i=1}^{k} \frac{U_i^2}{n\pi_{i0}}\right)^2 - (k-1)^2.$$

Randame

$$\mathbf{E}\left(\sum_{i=1}^{k} \frac{U_i^2}{n\pi_{i0}}\right)^2 = \sum_{i=1}^{k} \frac{\mathbf{E}U_i^4}{n^2\pi^2_{i0}} + \sum_{i\neq j} \frac{\mathbf{E}(U_i^2 U_j^2)}{n^2\pi_{i0}\pi_{j0}}.$$

Pasinaudoję sąryšiais tarp pradinių $\alpha_k = \mathbf{E}U_i^k$ ir faktorialinių $\nu_k = \mathbf{E}(U_i(U_i-1)\cdot\ldots\cdot(U_i-k+1))$ momentų

$$\alpha_4 = \nu_4 + 5\nu_3 + 7\nu_2 + \nu_1,$$

$$\alpha_{22} = \nu_{22} + \nu_{21} + \nu_{12} + \nu_{11},$$

randame

$$\mathbf{E}\left(\sum_{i-1}^{k} \frac{U_i^2}{n\pi_{i0}}\right)^2 = \frac{1}{n^2}\sum_{i=1}^{k}(n^{[4]}\pi_{i0}^2 + 6n^{[3]}\pi_{i0} + 7n^{[2]}) +$$

$$\frac{1}{n^2}\sum_{i\neq j}(n^{[4]}\pi_{i0}\pi_{j0} + 6n^{[3]}(\pi_{i0}+\pi_{j0}) + 7n^{[2]}) + \sum_{i=1}^{k}\frac{1}{\pi_{i0}},$$

čia $n^{[k]} = n(n-1)\cdot\ldots\cdot(n-k+1)$.

Susumavę ir sutraukę panašiuosius narius, gausime

$$\boldsymbol{V}(X_n^2) = 2(k-1) + \frac{1}{n}\sum_{i=1}^{k}\frac{1}{\pi_{i0}} - \frac{k^2}{n}.$$

b)

$$\mathbf{E}(X_n^2) = \sum_{i=1}^{k}\frac{\pi_i(1-\pi_i)}{\pi_{i0}} + n\sum_{i=1}^{k}\frac{\pi_i^2}{\pi_{i0}} - n =$$

$$k - 1 + \frac{n-1}{n}\delta_n + \sum_{i=1}^{k}(\pi_i/\pi_{i0}) - k,$$

necentriškumo parametras δ_n apibrėžtas **III.1.4** pratime.

Analogiškai p. a) gauname dispersijos $\boldsymbol{V}(X_n^2)$ išraišką, kai hipotezė nėra teisinga

$$\boldsymbol{V}(X_n^2) = 2\{(2n-4)\sum_i(pi_i^3/\pi_{i0}^2) - (2n-3)(\sum_i(\pi_i^2/\pi_{i0}))^2 -$$

$$2\sum_i(\pi_i^2/\pi_{i0})\sum_i(\pi_i/\pi_{i0}) + 3\sum_i(\pi_i^2/\pi_{i0}^2)\} + [\sum_i(\pi_i/\pi_{i0}^2) - (\sum_i(\pi_i/\pi_{i0}))^2]/n.$$

III.1.8. Pasiremkime **III.1.7** pratimu.

III.1.9. Turime paprastąją didumo n imtį $\mathbf{X} = (X_1, \ldots, X_n)^T$ ir tikriname suderinamumo hipotezę $H_0 : X_i \sim F_0(x)$; čia $F_0(x)$ yra žinoma pasiskirstymo funkcija. Taikant χ^2 kriterijų abscisių ašis sudalinama į intervalus $-\infty = a_0 < a_1 < \ldots < a_k = \infty$. Pažymėję U_i imties elementų, patekusių į i-ąjį intervalą, skaičių, nuo pradinės imties $\mathbf{X}$ pereiname prie mažiau informatyvios grupuotosios imties $\boldsymbol{U} = (U_1, \ldots, U_k)^T \sim \mathcal{P}_k(n, \boldsymbol{\pi})$.

Jeigu hipotezė H_0 teisinga, tai $\boldsymbol{\pi} = \boldsymbol{\pi}_0 = (\pi_{10}, \ldots, \pi_{k0})^T, \pi_{j0} = F_0(a_j) - F_0(a_{j-1}), j = 1, \ldots, k$. Vietoje hipotezės $H_0 : X_i \sim F_0(x)$ tikriname bendresnę hipotezę $H_0' : \pi_j = \pi_{j0}, j = 1, \ldots, k$ (žr. **III.1.1** pratimą). Atmetus H_0' natūralu atmesti ir H_0.

Pagal turimus duomenis sudalinkime intervalą $(0, 1)$ į 5 vienodo ilgio intervalus: $[0; 0,2], (0,2; 0,4], (0,4; 0,6], (0,6; 0,8], (0,8; 1]$. Gauname vektoriaus $\boldsymbol{U}$ realizaciją:

18; 12; 16; 19; 15.

Tikriname hipotezę $H_0' : \pi_i = 0,2,\ i = 1, ..., 5$. Gauname

$$X_n^2 = \frac{1}{n}\sum_{i=1}^{k}\frac{U_i^2}{\pi_{i0}} - n = \frac{18^2 + 12^2 + 16^2 + 19^2 + 15^2}{80 \cdot 0,2} - 80 = 1,875.$$

Asimptotinė P reikšmė $pv_a = \mathbf{P}\{\chi_4^2 > 1,875\} = 0,7587$. Atmesti hipotezę H_0' nėra pagrindo. Tikėtinumų santykio kriterijus duoda tą patį atsakymą, nes statistikos R_n realizacija yra 1,93 ir $pv_a = \mathbf{P}\{\chi_4^2 > 1,93\} = 0,7486$.

III.1.2 skyrelis

III.1.10. Kai hipotezėje H_0 tvirtinama, kad polinomino skirstinio tikimybės π_i yra dimensijos $s < k-1$ parametro $\boldsymbol{\theta}$ funkcijos $\pi_i = \pi_i(\boldsymbol{\theta}), i = 1, ..., k$, tai Pirsono statistika

$$X_n^2(\boldsymbol{\theta}) = \sum_{i=1}^{k}\frac{(U_i - n\pi_i(\boldsymbol{\theta}))^2}{n\pi_i(\boldsymbol{\theta})}$$

priklauso nuo nežinomo parametro $\boldsymbol{\theta}$. Natūralu nežinomą parametrą pakeisti kokiu nors įvertiniu ir išnagrinėti gautosios statistikos savybes.

1. Kai H_0 teisinga, imties $\boldsymbol{U}$ tikėtinumo funkcija ir jos logaritmas yra

$$L(\boldsymbol{\theta}) = \frac{n!}{U_1!...U_k!}\prod_{i=1}^{k}\pi_i^{U_i}(\boldsymbol{\theta}), \quad \ell(\boldsymbol{\theta}) = \sum_{i=1}^{k}U_i\ln(\pi_i(\boldsymbol{\theta})) + C.$$

Pažymėkime $\boldsymbol{\theta}_n^*$ DT įvertinį, gautą maksimizuojant $L(\boldsymbol{\theta})$ arba $\ell(\boldsymbol{\theta})$, ir tegu

$$X_n^2(\boldsymbol{\theta}^*) = \sum_{i=1}^{k}\frac{(U_i - n\pi_i(\boldsymbol{\theta}^*))^2}{n\pi_i(\boldsymbol{\theta}^*)}.$$

2. Parametro $\boldsymbol{\theta}$ įvertinį $\tilde{\boldsymbol{\theta}}$ raskime χ^2 minimumo metodu, t. y. iš sąlygos

$$X_n^2(\tilde{\boldsymbol{\theta}}) = \inf_{\boldsymbol{\theta}}\sum_{i=1}^{k}\frac{(U_i - n\pi_i(\boldsymbol{\theta}))^2}{n\pi_i(\boldsymbol{\theta})}.$$

3. Kartais parametro $\boldsymbol{\theta}$ įvertinį $\bar{\boldsymbol{\theta}}$ randame modifikuotuoju chi^2 minimumo metodu, t. y. iš sąlygos

$$X_n^2(\bar{\boldsymbol{\theta}}) = \inf_{\boldsymbol{\theta}}\sum_{i=1}^{k}\frac{(U_i - n\pi_i(\boldsymbol{\theta}))^2}{U_i}.$$

4. Pagaliau, tegu $R_n(\boldsymbol{\theta}^*)$ žymi tikėtinumų santykio statistiką

$$R_n(\boldsymbol{\theta}^*) = -2\ln\Lambda_n = 2\sum_{i=1}^{k}U_i\ln(U_i/(n\pi_i(\boldsymbol{\theta}^*))).$$

Jeigu hipotezė H_0 teisinga, $n \to \infty$, o funkcijos $\pi_i(\boldsymbol{\theta})$ tenkina gana bendras reguliarumo sąlygas (žr. [4], 2.2.1 teoremą), tai visos keturios statistikos $X_n^2(\boldsymbol{\theta}^*)$, $X_n^2(\tilde{\boldsymbol{\theta}})$,

$X_n^2(\bar{\boldsymbol{\theta}})$, $R_n(\boldsymbol{\theta}^*)$ yra asimptotiškai ekvivalenčios ir turi χ^2 skirstinius su $k-1-s$ laisvės laipsnių.

Pagal pratimo sąlygą gauname, kad parametro α DT įvertinio realizacija $\alpha^* = 0,1235$. Tada statistikų $X_n^2(\boldsymbol{\theta}^*), R_n(\alpha^*)$ yra 0,3633, 0,3641. Įvertinių $\tilde{\alpha}$ ir $\bar{\alpha}$ realizacijos irgi yra 0,1235 (keturių ženklų po kablelio tikslumu); $X_n^2(\tilde{\alpha}) = 0,3633$, $X_n^2(\bar{\alpha}) = 0,3658$. Visų keturių statistikų atveju asimptotinė P reikšmė ne mažesnė už 0,5453; atmesti hipotezę nėra pagrindo.

III.1.11. a) Tegu $X_i = 1$, jei i-asis vaikas yra berniukas, ir $X_i = 0$, jei i-asis vaikas yra mergaitė, $i = 1,\ 2$. Tarkime, kad $\mathbf{P}\{X_1 = 1\} = \mathbf{P}\{X_2 = 1\} = p$ ir a. d. X_1 ir X_2 nepriklausomi. Tada $\mathbf{P}\{X_1 = 1, X_2 = 1\} = p^2$, $\mathbf{P}\{X_1 = 0, X_2 = 0\} = (1-p)^2$, $\mathbf{P}\{X_1 = 1, X_2 = 0\} + \mathbf{P}\{X_1 = 0, X_2 = 1\} = 2p(1-p)$. Tikrinama hipotezė $H : \pi_1 = p^2, \pi_2 = (1-p)^2, \pi_3 = 2p(1-p)$. Šiuo atveju $k = 3, s = 1$. Tikimybės p DT įvertinio realizacija $\hat{p} = 0,5126$. Statistika $X_n^2(\hat{p})$ įgijo reikšmę 0,1159 ir $pv_a = \mathbf{P}\{\chi_1^2 > 0,1159\} = 0,7335$; duomenys neprieštarauja iškeltai hipotezei. b) Jeigu tartume, kad berniuko ir mergaitės gimimo tikimybė vienoda ir lygi 1/2, tai jokių parametrų vertinti nereikia. Statistika X_n^2 įgijo reikšmę 2,6723 ir $pv_a = \mathbf{P}\{\chi_2^2 > 2,6723\} = 0,2629$; duomenys neprieštarauja ir šiai hipotezei.

III.1.12. Statistika X_n^2, turinti asimptotinį chi kvadrato skirstinį su 4 laisvės laipsniais, įgijo reikšmę 0,6122 ir $pv_a = \mathbf{P}\{\chi_4^2 > 0,6122\} = 0,9617$; duomenys neprieštarauja iškeltai hipotezei. *Nurodymas.* Įrodykite ir pasiremkite tokiu faktu: esant teisingai hipotezei imties $(X_1, ..., X_n)^T$ sąlyginis skirstinys, kai suma $S = X_1 + ... + X_n$ fiksuota, yra polinominis $\mathcal{P}_n(S, \boldsymbol{\pi}_0)$, $\boldsymbol{\pi}_0 = (1/n, ..., 1/n)^T$.

III.1.13. Parametro λ DT įvertinio realizacija yra $\hat{\lambda} = 3,8666$. Statistika $X_n^2(\hat{\lambda})$ įgijo reikšmę 13,0146 ir $pv_a = \mathbf{P}\{\chi_{10}^2 > 13,0146\} = 0,2229$; duomenys neprieštarauja iškeltai hipotezei. Skaičiuojant statistikos reikšmę du paskutinieji intervalai buvo sujungti.

III.1.14. Parametro σ^2 DT įvertis yra $\hat{\sigma}^2 = 1581,65$. Statistika $X_n^2(\hat{\sigma}^2)$ įgijo reikšmę 2,6931 ir $pv_a = \mathbf{P}\{\chi_7^2 > 2,6931\} = 0,9119$; duomenys neprieštarauja iškeltai hipotezei. Skaičiuojant statistikos reikšmę du paskutinieji intervalai buvo sujungti.

III.1.15. Parametro θ DT įvertis yra $\hat{\theta} = 878,4$. Statistika $X_n^2(\hat{\theta})$ įgijo reikšmę 4,0477 ir $pv_a = \mathbf{P}\{\chi_8^2 > 4,0477\} = 0,8528$; duomenys neprieštarauja iškeltai hipotezei. Skaičiuojant statistikos reikšmę trys paskutinieji intervalai buvo sujungti.

III.1.16. Parametrų įverčiai $\hat{\mu} = 96,16°, \hat{\theta} = 0,6854$; statistikos $R_n(\hat{\mu}, \hat{\theta})$ ir $X_n^2(\hat{\mu}, \hat{\theta})$ įgijo reikšmes 9,960 ir 10,175; atitinkamos asimptotinės P reikšmės $pv_a = 0,354$ ir $pv_a = 0,337$. Hipotezė neatmetama.

III.1.17. Parametrų įverčiai $\hat{\mu} = 180,8°, \hat{\theta} = 0,2047$; statistikos $R_n(\hat{\mu}, \hat{\theta})$ ir $X_n^2(\hat{\mu}, \hat{\theta})$ įgijo reikšmes 24,858 ir 24,641; atitinkamos asimptotinės P reikšmės $pv_a = 0,0519$ ir $pv_a = 0,0550$. Reikšmingumo lygmens $\alpha = 0,05$ kriterijumi hipotezė neatmetama. Turint omenyje tokį didelį stebėjimų skaičių, matyt, galima daryti išvadą, kad tokio tipo duomenims aprašyti Mizeso modelis yra tinkamas.

III.1.18. Tikriname sudėtinę suderinamumo hipotezę

$$H_0 : F(x) \in \mathcal{F}_0 = \{F(x; \boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta\} \in \mathcal{F},$$

kad paprastosios imties $\mathbf{X} = (X_1, ..., X_n)^T$ elemento X_j pasiskirstymo funkcija priklauso aibei $\mathcal{F}_0$, sudarytai iš žinomo pavidalo pasiskirstymo funkcijų $F(x;\boldsymbol{\theta})$ priklausančių nuo nežinomo parametro $\boldsymbol{\theta} = (\theta_1, ..., \theta_s)^T$. Taikant χ^2 suderinamumo kriterijų abscisių ašis sudalijama į intervalus $-\infty = a_0 < a_1 < ... < a_k = \infty$, $k > s+1$, ir gaunama mažiau informatyvi grupuotoji imtis $\boldsymbol{U} = (U_1, ..., U_k)^T \sim \mathcal{P}_k(n, \boldsymbol{\pi})$. Vietoje hipotezės H_0 tikrinama hipotezė

$$H_0' : \boldsymbol{\pi} = \boldsymbol{\pi}(\boldsymbol{\theta}), \quad \boldsymbol{\pi}(\boldsymbol{\theta}) = (\pi_1(\boldsymbol{\theta}), ..., \pi_k(\boldsymbol{\theta}))^T, \quad \pi_i(\boldsymbol{\theta}) = F_0(a_i;\boldsymbol{\theta}) - F_0(a_{i-1};\boldsymbol{\theta}).$$

Tikrinant šią hipotezę Pirsono χ^2 kriterijus (žr. **III.1.10** pratimą) turi tam tikrų trūkumų. Pirma, nagrinėjant statistikos asimptotiką [4], 2.2.1 teoremoje buvo laikoma, kad grupavimo intervalų galai nepriklauso nuo imties. Tačiau praktiškai grupavimo intervalai parenkami atsižvelgiant į imties rezultatus. Antra, parametro $\boldsymbol{\theta}$ įvertinį reikia rasti DT ar χ^2 minimumo metodu pagal grupuotus duomenis. Gautieji įvertiniai nėra optimalūs, nes naudoja mažiau informatyvią grupuotąją imtį. Naudotis asimptotiškai optimaliais įvertiniais, gautais pagal pradinę negrupuotą imtį, negalima, nes tada statistikos asimptotinis skirstinys priklauso ir nuo skirstinio pavidalo, ir nuo parametro.

Šių trūkumų neturi modifikuotasis χ^2 kriterijus. Sudarant jo statistiką naudojami parametro $\boldsymbol{\theta}$ DT įvertiniai, gauti pagal pradinę imtį, o grupavimo intervalo galai gali tam tikru būdu priklausyti nuo imties.

Kriterijaus statistika Y_n^2, kai $n \to \infty$ ir H_0' teisinga, asimptotiškai turi χ^2 skirstinį (žr. [4], 2.3 skyrelį)

$$Y_n^2 = X_n^2(\hat{\boldsymbol{\theta}}) + Q_n \xrightarrow{d} \chi^2_{k-1}.$$

Atkreipsime dėmesį, kad laisvės laipsnių skaičius nėra mažinamas įvertintų parametrų skaičiumi s. Statistika Y_n^2 yra suma dviejų kvadratinių formų: pirmoji $X_n^2(\hat{\boldsymbol{\theta}})$ yra Pirsono statistika, o kvadratinė forma Q_n parinkta taip, kad Y_n^2 asimptotinis skirstinys būtų χ^2 skirstinys su $k-1$ laisvės laipsnių.

Modifikuoto χ^2 kriterijaus taikymas tikrinant hipotezę apie stebimo a. d. normalumą detaliai iliustruojamas [4], 2.3.3 skyrelio 2.3.2 ir 2.3.3 pavyzdžiuose.

Pagal turimus duomenis parametrų μ ir σ DT įverčiai yra $\hat{\mu} = \bar{X} = 324,57$ ir $\hat{\sigma} = 20,8342$. Parenkame $k = 8$ intervalus. Tada $X_n^2 = 8,0$, $Q_n = 2,8302$, $Y_n^2 = 10,8302$ ir $pv_a = \mathbf{P}\{\chi_7^2 > 10,8302\} = 0,1462$. Hipotezė neatmetama.

III.1.19. Perėję prie logaritmų $Y_i = \ln(X_i)$ gauname DT įverčius $\hat{\mu} = \bar{Y} = 2,4589$ ir $\hat{\sigma} = 0,9529$. Parenkame $k = 10$ intervalų. Tada $X_n^2 = 4,1333$, $Q_n = 1,0668$, $Y_n^2 = 5,2001$ ir $pv_a = \mathbf{P}\{\chi_9^2 > 5,2001\} = 0,8165$. Duomenys neprieštarauja iškeltai hipotezei.

III.1.20. Neatsižvelgiant į duomenų apvalinimą gaunama $X_n^2 = 4,160$, $Q_n = 0,172$, $Y_n^2 = 4,332$ ir $pv_a = \mathbf{P}\{\chi_7^2 > 4,332\} = 0,741$. Duomenys neprieštarauja iškeltai hipotezei. Atlikę korekciją atsižvelgdami į duomenų apvalinimą, gauname $X_n^{'2} = 3,731$, $Q_n' = 0,952$, $Y_n^{'2} = 4,683$ ir $pv_a' = \mathbf{P}\{\chi_7^2 > 4,683\} = 0,699$. Duomenys neprieštarauja iškeltai hipotezei. Reikia pažymėti, kad P reikšmės pv ir pv' gerokai skiriasi. *Nurodymas.* Kadangi duomenys suapvalinti iki sveikųjų skaičių, tai gautuosius intervalų galus a_i reikia pastumti iki artimiausių $m \pm 0,5$ pavidalo rėžių (m – sveikasis skaičius) ir apskaičiuoti statistikos reikšmę naudojant naujai gautus rėžius a_i'.

III.1.21. Tegu turime dvi sistemas $\{A_1, ..., A_s\}$ ir $\{B_1, ..., B_r\}$ nesutaikomų, sudarančių pilnas įvykių grupes atsitiktinių įvykių. Tegu U_{ij} yra įvykio $A_i \bigcap B_j$ pasirodymų skaičius. Atsitiktinis vektorius

$$\boldsymbol{U} = (U_{11}, ..., U_{1r}, U_{21}, ..., U_{2r}, ..., U_{s1}, ..., U_{sr})^T$$

turi polinominį skirstinį

$$\boldsymbol{U} \sim \mathcal{P}_{s\times r}(n, \boldsymbol{\pi}), \quad \boldsymbol{\pi} = (\pi_{11}, ..., \pi_{sr})^T, \quad \sum_i \sum_j \pi_{ij} = 1.$$

Dviejų atsitiktinių įvykių sistemų nepriklausomumo hipotezė

$$H_0' : \pi_{ij} = \mathbf{P}\{A_i \bigcap B_j\} = \mathbf{P}\{A_i\}\mathbf{P}\{B_j\} = \pi_{i.}\pi_{.j}, \ i = 1, ..., s, \ j = 1, ..., r.$$

Jei hipotezė H_0' teisinga, tai turime sudėtinės hipotezės atvejį (žr. **III.1.10** pratimą), kai polinominio skirstinio tikimybės π_{ij} yra dimensijos $s + r - 2$ parametro

$$\boldsymbol{\theta} = (\pi_{1.}, ..., \pi_{s-1.}, \pi_{.1}, ..., \pi_{.r-1})^T$$

funkcijos. Parametro $\boldsymbol{\theta}$ elementų DT įvertiniai yra

$$\hat{\pi}_{i.} = \frac{1}{n}\sum_{j=1}^{r} U_{ij} = \frac{U_{i.}}{n}, \quad \hat{\pi}_{.j} = \frac{1}{n}\sum_{i=1}^{s} U_{ij} = \frac{U_{.j}}{n}.$$

Gauname, kad jei H_0' teisinga ir $n \to \infty$, tai

$$X_n^2(\hat{\boldsymbol{\theta}}_n) = \sum_i \sum_j \frac{(U_{ij} - n\hat{\pi}_{i.}\hat{\pi}_{.j})^2}{n\hat{\pi}_{i.}\hat{\pi}_{.j}} = n\left(\sum_i \sum_j \frac{U_{ij}^2}{U_{i.}U_{.j}} - 1\right) \xrightarrow{d} \chi^2_{(r-1)(s-1)}.$$

Nepriklausomumo hipotezė atmetama asimptotiniu reikšmingumo lygmens α kriterijumi, kai

$$X_n^2(\hat{\boldsymbol{\theta}}_n) > \chi^2_\alpha((r-1)(s-1)).$$

Pagal turimus duomenis gauname, kad statistika $X_n^2(\hat{\boldsymbol{\theta}}_n)$ įgijo reikšmę 121,286. Hipotezė atmetama.

III.1.22. Statistika $X_n^2(\hat{\boldsymbol{\theta}}_n)$ įgijo reikšmę 3,719 ir $pv_a = \mathbf{P}\{\chi^2_4 > 3,719\} = 0,4454$; duomenys neprieštarauja iškeltai hipotezei.

III.1.23. Tegu yra s nepriklausomų objektų grupių; i-osios grupės objektų skaičius $n_i, i = 1, ..., s$. Tegu $\{B_1, ..., B_r\}$ yra pilna nesutaikomų atsitiktinių įvykių aibė. Stebint bet kurį objektų žinoma, kuris iš įvykių $B_1, ..., B_r$ įvyko. Pažymėkime U_{ij} skaičių i-osios grupės objektų, kuriuos stebint įvyko įvykis B_j. Tada atsitiktinis vektorius $\boldsymbol{U}_i = (U_{i1}, U_{i2}, ..., U_{ir}) \sim \mathcal{P}_r(n_i, (\pi_{i1}), ..., \pi_{ir}), i = 1, ..., s$; čia π_{ij} tikimybė įvykti įvykiui B_j, kai objektas yra iš i-osios grupės.

Homogeniškumo hipotezė

$$H_0' : \pi_{1j} = \pi_{2j} = ... = \pi_{sj} = \pi_j, \quad j = 1, ..., r,$$

reiškia, kad įvykio B_j tikimybė yra ta pati visų grupių objektams.

Turime sudėtinės hipotezės atvejį (žr. **III.1.10** pratimą), kai polinominių skirstinių tikimybės π_{ij} yra parametro $\boldsymbol{\theta} = (\pi_1, ..., \pi_{r-1})^T$ funkcijos. DT įvertiniai yra $\hat{\pi}_j = U_{.j} = \sum_i U_{ij/n}$, $n = n_1 + ... + n_s$. Jeigu H_0' teisinga ir $n_i \to \infty$, $i = 1, ..., s$, tai

$$X_n^2(\hat{\boldsymbol{\theta}}_n) = \sum_i \sum_j \frac{(U_{ij} - n_i\hat{\pi}_j)^2}{n_i\hat{\pi}_j} = n\left(\sum_i \sum_j \frac{U_{ij}^2}{n_i U_{.j}} - 1\right) \xrightarrow{d} \chi^2_{(r-1)(s-1)}.$$

Hipotezė atmetama asimptotiniu reikšmingumo lygmens α kriterijumi, kai

$$X_n^2(\hat{\boldsymbol{\theta}}_n) > \chi_\alpha^2((r-1)(s-1)).$$

Atkreipsime dėmesį, kad statistika ir kritinė sritis yra tokia pat, kaip ir **III.1.21** pratime, nors sprendžiamas visai kitas uždavinys.

Pagal turimus duomenis statistika $X_n^2(\hat{\boldsymbol{\theta}}_n)$ įgijo reikšmę 2,0771 ir $pv_a = \mathbf{P}\{\chi_3^2 > 2,0771\} = 0,5566$; duomenys neprieštarauja iškeltajai hipotezei.

III.1.24. Tegu $(X_{i1}, ..., X_{in_i})^T$, $i = 1, ..., s$, yra s paprastųjų nepriklausomų imčių ir $F_i(x)$ yra a. d. X_{ij} pasiskirstymo funkcija. Tada homogeniškumo hipotezė:

$$H_0 : F_1(x) \equiv F_2(x) \equiv ... \equiv F_s(x).$$

Sudalinę abscisių ašį į intervalus $-\infty = a_0 < a_1 < ... < a_r = \infty$, gauname situaciją, aprašytą **III.1.23** pratime: objektus atitinka imčių elementai, o įvykis B_j reiškia patekimą į j-ąjį intervalą. Vietoje hipotezės H_0 tikriname hipotezę $H_0' : \pi_{1j} = ... = \pi_{sj} = \pi_j$, $j = 1, ..., r$. Atmetus hipotezę H_0' natūralu atmesti ir hipotezę H_0.

Pagal turimus duomenis gauname, kad statistika $X_n^2(\hat{\boldsymbol{\theta}})$ įgijo reikšmę 75,035 (trys paskutinieji intervalai sujungti) ir $pv_a = \mathbf{P}\{\chi_7^2 > 75,035\} < 10^{-12}$; hipotezė atmetama.

III.1.25. Tarkime, kad **III.1.24** pratimo sąlygomis tikrinama siauresnė homogeniškumo hipotezė

$$H_0 : F_1(x) \equiv F_2(x) \equiv ... \equiv F_s(x) \equiv F_0(x),$$

kurioje tvirtinama ne tik kad pasiskirstymo funkcijos $F(x), ..., F_s(x)$ yra vienodos, bet ir kad jos sutampa su žinoma pasiskirstymo funkcija $F_0(x)$. Perėję prie grupuotųjų imčių vietoje hipotezės H_0 tikriname paprastąją hipotezę

$$H_0' : \pi_{1j} = ... = \pi_{sj} = \pi_j^{(0)}, \quad \pi_j^{(0)} = F_0(a_j) - F_0(a_{j-1}).$$

Kriterijaus statistika

$$X_n^2 = \sum_{i=1}^{s} \sum_{j=1}^{r} \frac{(U_{ij} - n_i \pi_j^{(0)})^2}{n_i \pi_j^{(0)}} = \sum_{i=1}^{s} \left[\sum_{j=1}^{r} \frac{(U_{ij})^2}{n_i \pi_j^{(0)}} - n_i \right].$$

Jeigu H_0' teisinga ir $n_i \to \infty$, tai vidinė suma artėja į χ^2 skirstinį su $r-1$ laisvės laipsnių, o $X_n^2 \xrightarrow{d} \chi^2_{s(r-1)}$. Hipotezė H_0' atmetama asimptotiniu reikšmingumo lygmens α kriterijumi, kai

$$X_n^2 > \chi_\alpha^2(s(r-1)).$$

Pagal turimus duomenis gauname, kad statistika X_n^2 įgijo reikšmę 18,032 ir $pv_a = \mathbf{P}\{\chi_{22}^2 > 18,032\} = 0,704$.

III.1.26. a) Kiekvienoje imtyje įvertiname parametrą λ, apskaičiuojame statistikų $X_{n_i}^2(\hat{\lambda}_i)$ reikšmes ir jas sudedame. Gauname statistikos, kuri esant teisingai hipotezei asimptotiškai turi chi kvadrato skirstinį su 4 laisvės laipsniais, realizaciją. Gautoji reikšmė yra 2,5659 ir $pv_a = \mathbf{P}\{\chi_4^2 > 2,5659\} = 0,6329$. Hipotezę atmesti nėra pagrindo; b) įvertiname parametrą λ pagal jungtinę imtį ir gauname $\hat{\lambda} = 0,2494$. Apskaičiuojame statistikų $X_{n_i}(\hat{\lambda})$ reikšmes ir jas sudedame; gauname 10,2317. Kadangi $pv_a = \mathbf{P}\{\chi_7^2 > 10,2317\} = 0,1758$, tai ir ši hipotezė neatmetama. Skaičiuojant du paskutinieji intervalai buvo sujungti.

III.2. Glodūs Neimano ir Bartono kriterijai

III.2.1. Paprastoji suderinamumo hipotezė

III.2.1. Remiantis paprastąja imtimi $\mathbf{X} = (X_1, ..., X_n)^T$ tikrinama paprastoji suderinamumo hipotezė

$$H_0 : X_i \sim F_0(x);$$

čia $F_0(x)$ žinoma absoliučiai tolydi pasiskirstymo funkcija su tankiu $f_0(x) = F_0'(x)$, kai sudėtinė alternatyva yra $H : X_i \sim F(x) \in \mathcal{F}$; čia $\mathcal{F}$ yra absoliučiai tolydžių skirstinių aibė su tankiais $f(x) = F'(x)$.

a) Atlikus integralinę transformaciją $Y_i = F_0(X_i), i = 1, ..., n$, hipotezė H_0 pereina į tokią. Remiantis paprastąja imtimi $Y_1, ..., Y_n$ tikrinama hipotezė $H_0 : Y_i \sim U(0,\ 1)$. Kaip transformuojama alternatyvų aibė $\mathcal{F}$?

b) Tegu tikrinama paprastoji hipotezė $H_0 : Y_i \sim U(0,\ 1)$, kai alternatyvų aibė yra $H : Y_i \sim g(y) \in \mathcal{G}$; čia $\mathcal{G}$ yra intervale $(0,\ 1)$ apibrėžtų tankių $g(y), 0 < y < 1$, aibė. Atlikime transformaciją $X_i = F_0^{-1}(Y_i)$. Tada hipotezė H_0 virsta paprastąja hipoteze $H_0 : X_i \sim F_0(x)$. Kaip tokiu atveju transformuojasi alternatyvų aibė $\mathcal{G}$?

III.2.2. Pagal paprastąją didumo n imtį $\mathbf{X} = (X_1, ..., X_n)^T$ tikrinama paprastoji suderinamumo hipotezė $H_0 : X_i \sim \mathcal{E}(1)$, kai alternatyvų aibė yra $\{\mathcal{E}(\lambda),\ \lambda \neq 1,\ \lambda > 0\}$. Raskite alternatyvų aibę $\mathcal{G}$, kai atlikta a. d. $X_1, ..., X_n$ transformacija $Y_i = 1 - e^{-X_i}$.

III.2.3. (**III.2.1** pratimo tęsinys). Tarkime, kad atlikę transformaciją $Y_i = F_0(X_i)$ gavome alternatyvų aibę $\{Be(\lambda, 1),\ \lambda \neq 1,\ \lambda > 0\}$. Kokia yra pradinio uždavinio alternatyvų aibė $\mathcal{F}$, jei $F_0(x) = 1 - e^{-x}$?

III.2.4. Pagal paprastąją didumo n imtį $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ tikrinama paprastoji suderinamumo hipotezė $H_0 : Y_i \sim U(0,\ 1)$, kai Neimano tipo alternatyvių tankių aibė yra $\mathcal{G} = \{g(y|\theta) = \frac{1}{c(\theta)} \exp\{\theta(y - 1/2)\}, 0 < y < 1, \theta > 0\}$. Raskite TG kriterijų hipotezei H_0, kai alternatyva yra $H : Y_i \sim g \in \mathcal{G}$, tikrinti. Naudodami normaliąją aproksimaciją suformuluokite asimptotinį kriterijų.

III.2.5. (**III.2.4** pratimo tęsinys). Raskite **III.2.4** pratime surasto kriterijaus statistikos asimptotinį skirstinį, kai teisinga alternatyva. Naudodami normaliąją aproksimaciją raskite asimptotinio reikšmingumo lygmens α kriterijaus galios funkciją.

III.2.6. (**III.2.5** pratimo tęsinys). Apskaičiuokite asimptotinę reikšmingumo lygmens $\alpha = 0,05$ kriterijaus galią, kai a) $n = 50;\ \theta = 1,2;\ 1,5;\ 2;\ 3;$ b) $\theta = 0,5; n = 50;\ 100;\ 200.$

III.2.7. Pagal paprastąją didumo n imtį $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ tikrinama paprastoji suderinamumo hipotezė $H_0 : Y_i \sim U(0,\ 1)$, kai Neimano tipo alternatyvių tankių aibė yra $\mathcal{G} = \{g(y|\theta) = \frac{1}{c(\theta)} \exp\{\theta(y - 1/2)^2\}, 0 < y < 1, \theta > 0\}$. Raskite TG kriterijų hipotezei H_0, kai alternatyva yra $H : Y_i \sim g \in \mathcal{G}$, tikrinti. Naudodami normaliąją aproksimaciją suformuluokite asimptotinį kriterijų.

III.2.8. (**III.2.7** pratimo tęsinys). Raskite **III.2.7** pratime surasto kriterijaus statistikos asimptotinį skirstinį, kai teisinga alternatyva. Naudodami normaliąją aproksimaciją raskite asimptotinio reikšmingumo lygmens α kriterijaus galios funkciją.

III.2.9. (**III.2.8** pratimo tęsinys). Apskaičiuokite asimptotinę reikšmingumo lygmens $\alpha = 0,05$ kriterijaus galią, kai a) $n = 50;\ \theta = 1,2;\ 1,5;\ 2;\ 3;$ b) $\theta = 1,0; n = 50;\ 100;\ 200.$

III.2.10. Pagal paprastąją didumo n imtį $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ tikrinama paprastoji suderinamumo hipotezė $H_0 : Y_i \sim U(0,\ 1)$, kai alternatyvių tankių aibė yra $\mathcal{G} = \{g(y|\gamma) = \gamma(1-y)^{\gamma-1}, 0 < y < 1, \gamma > 1\}$. Raskite TG kriterijų hipotezei H_0, kai alternatyva yra $H : Y_i \sim g \in \mathcal{G}$, tikrinti ir jo galios funkciją.

III.2.11. Pagal paprastąją didumo n imtį $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ tikrinama paprastoji suderinamumo hipotezė $H_0 : X_i \sim U(0,\ 1)$, kai alternatyvių tankių aibė yra $\mathcal{G} = \{g(y|\gamma) = \gamma(1-y)^{\gamma-1}, 0 < y < 1,\ 0 < \gamma < 1\}$. Raskite TG kriterijų hipotezei H_0, kai alternatyva yra $H : Y_i \sim g \in \mathcal{G}$, tikrinti ir jo galios funkciją.

III.2.12. Pagal paprastąją didumo n imtį $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ tikrinama paprastoji suderinamumo hipotezė $H_0 : X_i \sim U(0,\ 1)$, kai alternatyvių tankių aibė yra 1) $\mathcal{G} = \{g(y|\gamma) = \gamma y^{\gamma-1}, 0 < y < 1,\ 0 < \gamma < 1\}$; 2) $\mathcal{G} = \{g(y|\gamma) = \gamma y^{\gamma-1}, 0 < y < 1,\ 1 < \gamma\}$. Raskite TG kriterijus ir jų galios funkcijas.

III.2.2. Sudėtinės suderinamumo hipotezės

III.2.13. Raskite modifikuotojo Neimano ir Bartono tipo (2 parametrai) asimptotinį kriterijų eksponentiškumo hipotezei $H_0 : X_i \sim \mathcal{E}(1/\lambda), \lambda > 0$ tikrinti.

III.2.14. Raskite modifikuotąjį asimptotinį kriterijų, grindžiamą beta skirstiniu, eksponentiškumo hipotezei $H_0 : X_i \sim \mathcal{E}(1/\lambda), \lambda > 0$ tikrinti.

III.2.15. (**III.2.13** ir **III.2.14** pratimų tęsinys). Remdamiesi **III.2.13** ir **III.2.14** pratimuose rastais kriterijais patikrinkite hipotezę, kad pateiktieji duomenys gauti stebint eksponentinį a. d.

5,017 0,146 6,474 13,291 5,126 8,934 10,971 7,863 5,492 13,930 12,708 7,329 5,408 6,808
0,923 4,679 2,242 4,120 12,080 2,502 16,182 6,592 2,653 4,252 8,609 10,419 2,173 3,321
4,086 11,667 19,474 11,067 11,503 2,284 0,926 2,065 4,703 3,744 5,286 5,497 4,881 0,529
10,397 30,621 5,193 7,901 10,220 16,806 10,672 4,209 5,699 20,952 12,542 7,316 0,272
4,380 9,699 9,466 7,928 13,086 8,871 13,000 16,132 9,950 8,449 8,301 16,127 22,698 4,335

III.2.16. Modifikuotuoju Neimano ir Bartono tipo ir beta skirstiniu grindžiamais kriterijais patikrinkite hipotezę, kad **II.1.18** pratimo duomenys gauti stebint normalųjį a. d.

III.2.17. Modifikuotuoju Neimano ir Bartono tipo ir beta skirstiniu grindžiamais kriterijais patikrinkite hipotezę, kad **II.1.20** pratimo duomenys gauti stebint normalųjį a. d.

III.2.18. Sumodeliuokite didumo $n = 50$ paprastąją imtį, gautą stebint normalųjį a. d., ir, taikydami [4], 3.5 skyrelio kriterijus, patikrinkite hipotezę, kad buvo sumodeliuotas a) normalusis a. d.; b) logistinis a. d.; c) Koši a. d.

III.2.19. Sumodeliuokite didumo $n = 50$ paprastąją imtį, gautą stebint lognormalųjį a. d., ir, taikydami [4], 3.5 skyrelio kriterijus, patikrinkite hipotezę, kad buvo sumodeliuotas a) lognormalusis a. d.; b) loglogistinis a. d.

III.2.20. Sumodeliuokite didumo $n = 50$ paprastąją imtį, gautą stebint Koši a. d., ir, taikydami [4], 3.5 skyrelio kriterijus, patikrinkite hipotezę, kad buvo sumodeliuotas a) normalusis a. d.; b) logistinis a. d.; c) Koši a. d.

III.2.21. Sumodeliuokite didumo $n = 50$ paprastąją imtį, gautą stebint Veibulo a. d., ir, taikydami [4], 3.5 skyrelio kriterijus, patikrinkite hipotezę, kad buvo sumodeliuotas

a) Veibulo a. d.; b) lognormalusis a. d.; c) maksimalių reikšmių a. d.

III.2.3. Sprendimai, atsakymai, nurodymai

III.2.1 skyrelis

III.2.1. a) Tegu $X_i \sim F(x)$. Tada remdamiesi transformacija $Y_i = F_0(X_i)$ ir atvirkštine transformacija $X_i = F_0^{-1}(Y_i)$ gauname a. d. Y_i pasiskirstymo funkciją

$$G(y) = \mathbf{P}\{Y_i < y\} = \mathbf{P}\{F_0(X_i) < y\} = \mathbf{P}\{X_i < F_0^{-1}(y)\} = F(F_0^{-1}(y))$$

ir tankio funkciją

$$g(y) = G'(y) = f(F_0^{-1}(y))[F_0^{-1}(y)]' = \frac{f(F_0^{-1}(y))}{f_0(F_0^{-1}(y))}, \quad 0 < y < 1.$$

b) Remdamiesi transformacija $x = F_0^{-1}(y)$ ir atvirkštine transformacija $y = F_0(x)$, gauname tankio funkciją

$$f(x) = g(F_0(x))d(F_0(x)) = g(F_0(x))f_0(x), \quad g \in \mathcal{G}.$$

III.2.2. Remdamiesi **III.2.1** pratimu gauname tankio funkciją

$$g(y) = \lambda(1 - y)^{\lambda-1}, \; 0 < y < 1, \; \lambda \neq 1, \; \lambda > 0.$$

Taigi $\mathcal{G} = \{\mathcal{B}\rceil(\infty, \lambda), \lambda \neq \infty, \; \lambda > /\}$.

III.2.3. Remdamiesi **III.2.1** pratimu gauname, kad alternatyvų aibę $\mathcal{F}$ sudaro tankiai $f(x) = g(F_0(x))f_0(x) = \lambda e^{-x}(1-e^{-x})^{\lambda-1}, x > 0, \lambda \neq 1$. Jeigu a. d. Y_i pakeistume a. d $Z_i = 1 - Y_i$, tai alternatyvų aibė būtų kaip **III.2.2** pratime $\{\mathcal{E}(\lambda), \lambda \neq 1, \lambda > 0\}$.

III.2.4. Parinkime konkrečią alternatyvą $g(y|\theta)$ iš alternatyvų aibės $\mathcal{G}$. Tikrinkime paprastąją hipotezę $H_0 : Y_i \sim U(0,\ 1)$, kai paprastoji alternatyva yra $\bar{H} : Y_i \sim g(y|\theta), i = 1, ..., n$. Remiantis Neimano ir Pirsono lema hipotezė atmetama, kai

$$\frac{\prod_{i=1}^n g(Y_i|\theta)}{1} = \frac{1}{[c(\theta)]^n} \exp\{\theta \sum_{i=1}^n (Y_i - 1/2)\} > b \;\Leftrightarrow\; T = \sum_{i=1}^n (Y_i - 1/2) > d.$$

Konstanta d randama iš sąlygos, kad esant teisingai hipotezei $\mathbf{P}\{T > d\} = \alpha$; čia α kriterijaus reikšmingumo lygmuo. Kadangi surastas kriterijus nepriklauso nuo pasirinktos alternatyvos, tai jis yra TG su visomis alternatyvomis iš $\mathcal{G}$.

Randame $\mathbf{E}Y_i = \mathbf{E}T = 0$, $\boldsymbol{V}Y_i = 1/12, \boldsymbol{V}T = n/12$ ir

$$\frac{2\sqrt{3}}{\sqrt{n}} T = \frac{2\sqrt{3}}{\sqrt{n}} \sum_{i=1}^n (Y_i - \frac{1}{2}) \xrightarrow{d} Z \sim N(0,\ 1).$$

Asimptotiniu reikšmingumo lygmens α kriterijumi hipotezė H_0 atmetama, kai

$$\frac{2\sqrt{3}}{\sqrt{n}} \sum_{i=1}^n (Y_i - \frac{1}{2}) > z_\alpha.$$

III.2.5. Pažymėkime $\mu(\theta) = \mathbf{E}_\theta(Y_i - 1/2)$ ir $\sigma^2(\theta) = \boldsymbol{V}_\theta(Y_i - 1/2)$. Jeigu θ tikroji parametro reikšmė, tai

$$T(\theta) = (1/(\sigma(\theta)\sqrt{n}))\sum_{i=1}^{n}(Y_i - 1/2 - \mu(\theta)) \xrightarrow{d} Z \sim N(0,\ 1).$$

Asimptotinė reikšmingumo lygmens α kriterijaus galia

$$\beta(\theta) = \mathbf{P}_\theta\{\frac{2\sqrt{3}}{\sqrt{n}}\sum_{i=1}^{n}(Y_i - \frac{1}{2}) > z_\alpha\} =$$

$$\mathbf{P}_\theta\{T(\theta) > \frac{z_\alpha}{2\sqrt{3}\sigma(\theta)} - \frac{\mu(\theta)\sqrt{n}}{\sigma(\theta)}\} = \Phi(\frac{\sqrt{n}\mu(\theta)}{\sigma(\theta)} - \frac{z_\alpha}{\sqrt{12}\sigma(\theta)}).$$

Reikia dar rasti $\mu(\theta)$ ir $\sigma(\theta)$. Randame normuojančią konstantą $c(\theta) = (e^{\theta/2} - e^{-\theta/2})/\theta$. Tada

$$\mu(\theta) = [\ln c(\theta)]'_\theta = [e^\theta(\theta - 2) + \theta + 2]/(2\theta(e^\theta - 1));$$

$$\sigma^2(\theta) = [\ln c(\theta)]''_{\theta^2} = e^\theta(e^\theta + e^{-\theta} - \theta^2 - 2)/(\theta(e^\theta - 1))^2.$$

III.2.6. a) 0,7807; 0,9164; 0,9919; 1,000; b) 0,6451; 0,8897; 0,9925.

III.2.7. Parinkime konkrečią alternatyvą $g(y|\theta)$ iš alternatyvų aibės $\mathcal{G}$. Tikrinkime paprastąją hipotezę $H_0 : Y_i \sim U(0,\ 1)$, kai paprastoji alternatyva yra $\bar{H} : Y_i \sim g(y|\theta), i = 1, ..., n$. Remiantis Neimano ir Pirsono lema hipotezė atmetama, kai

$$\frac{\prod_{i=1}^{n} g(Y_i|\theta)}{1} = \frac{1}{[c(\theta)]^n}\exp\{\theta\sum_{i=1}^{n}(Y_i - 1/2)^2\} > b \quad \Leftrightarrow \quad T = \sum_{i=1}^{n}(Y_i - 1/2)^2 > d.$$

Konstanta d randama iš sąlygos, kad esant teisingai hipotezei $\mathbf{P}\{T > d\} = \alpha$; čia α kriterijaus reikšmingumo lygmuo. Kadangi surastas kriterijus nepriklauso nuo pasirinktos alternatyvos, tai jis yra TG su visomis alternatyvomis iš $\mathcal{G}$.

Kai hipotezė teisinga, tai $\mathbf{E}(Y_i - 1/2)^2 = 1/12$, $\boldsymbol{V}(Y_i - 1/2)^2 = 1/180$ ir

$$\frac{6\sqrt{5}}{\sqrt{n}}\sum_{i=1}^{n}[(Y_i - \frac{1}{2})^2 - \frac{1}{12}] \xrightarrow{d} Z \sim N(0,\ 1).$$

Asimptotiniu reikšmingumo lygmens α kriterijumi hipotezė H_0 atmetama, kai

$$\frac{6\sqrt{5}}{\sqrt{n}}\sum_{i=1}^{n}[(Y_i - \frac{1}{2})^2 - \frac{1}{12}] > z_\alpha.$$

III.2.8. Pažymėkime $\mu(\theta) = \mathbf{E}_\theta(Y_i - 1/2)^2$ ir $\sigma^2(\theta) = \boldsymbol{V}_\theta(Y_i - 1/2)^2$. Jeigu θ tikroji parametro reikšmė, tai

$$T(\theta) = \frac{1}{\sqrt{n}\sigma(\theta)}\sum_{i=1}^{n}[(Y_i - 1/2)^2 - \mu(\theta)] \xrightarrow{d} Z \sim N(0,\ 1).$$

Asimptotinė kriterijaus galia

$$\beta(\theta) = \Phi\left(\frac{\sqrt{n}(\mu(\theta) - 1/12)}{\sigma(\theta)} - \frac{z_\alpha}{(6\sqrt{5}\sigma(\theta))}\right).$$

Normuojanti konstanta $c(\theta) = \int_0^1 \exp\{\theta(x-1/2)^2\}dx$, $c'(\theta) = \int_0^1 (x-1/2)^2 \exp\{\theta(x-1/2)^2\}dx$, $c''(\theta) = \int_0^1 (x-1/2)^4 \exp\{\theta(x-1/2)^2\}dx$. Tada $\mu(\theta) = c'(\theta)/c(\theta)$, $\sigma^2(\theta) = c''(\theta)/c(\theta) - [c'(\theta)/c(\theta)]^2$. Kai θ žinomas, integralus galima apskaičiuoti skaitiniais metodais.

III.2.9. a) 0,1668; 0,2122; 0,3016; 0,5148; b) 0,1403; 0,1950; 0,2912.

III.2.10. Parinkime konkrečią alternatyvą $g(y|\theta)$ iš alternatyvų aibės $\mathcal{G}$. Tikrinkime paprastąją hipotezę $H_0 : Y_i \sim U(0,\ 1)$, kai paprastoji alternatyva yra $\bar{H} : Y_i \sim g(y|\theta), i = 1, ..., n$. Remiantis Neimano ir Pirsono lema hipotezė atmetama, kai

$$\frac{\prod_{i=1}^n g(Y_i|\gamma)}{1} = \gamma^n \prod_{i=1}^n (1-Y_i)^{\gamma-1} = \exp\{(\gamma-1)\sum_{i=1}^n \ln(1-Y_i) + n\ln\gamma\} > c \quad \Leftrightarrow$$

$$T = \sum_{i=1}^n \ln(1-Y_i) > d.$$

Konstanta d randama iš sąlygos, kad esant teisingai hipotezei $\mathbf{P}\{T > d\} = \alpha$; čia α kriterijaus reikšmingumo lygmuo. Kadangi surastas kriterijus nepriklauso nuo pasirinktos alternatyvos, tai jis yra TG su visomis alternatyvomis iš $\mathcal{G}$.

Esant teisingai hipotezei $-\ln(1-Y_i) \sim G(1,\ 1), -2\ln(1-Y_i) \sim \chi^2(2)$, todėl surastąjį TG kriterijų galima suformuluoti taip: hipotezė atmetama, kai

$$-2T = -2\sum_{i=1}^n \ln(1-Y_i) < \chi^2_{1-\alpha}(2n).$$

Kai teisinga alternatyva $g(y|\gamma)$, tai $-2\ln(1-Y_i) \sim G(\gamma, 1), -2\gamma\ln(1-Y_i) \sim \chi^2(2)$. Kriterijaus galios funkcija

$$\beta(\gamma) = \mathbf{P}_\gamma\{-2\sum_{i=1}^n \ln(1-Y_i) < \chi^2_{1-\alpha}(2n)\} = \mathbf{P}\{\chi^2_{2n} < \gamma\chi^2_{1-\alpha}(2n)\}$$

artėja prie vieneto, kai $\gamma \to \infty$.

III.2.11. Analogiškai **III.2.10** pratimui reikšmingumo lygmens α TG kriterijus atmeta hipotezę H_0, kai $T = -2\sum_{i=1}^n \ln(1-Y_i) > \chi^2_\alpha(2n)$. Galios funkcija $\beta(\gamma) = \mathbf{P}\{\chi^2_{2n} > \gamma\chi^2_\alpha(2n)\} \to 1$, kai $\gamma \to 0$.

III.2.12. Atlikę keitimą $Z_i = 1-Y_i$ gauname **III.2.10**, **III.2.11** pratimų alternatyvų šeimas.

III.2.2 skyrelis

III.2.13. Hipotezė atmetama asimptotiniu reikšmingumo lygmens α kriterijumi, kai $T = (T_1^2 - 2\rho T_1 T_2 + T_2^2)/(1-\rho^2) > \chi^2_\alpha(2)$, čia $T_1 = 4\sqrt{3}\sum_{i=1}^n (Y_i - 1/2)/\sqrt{n}$,, $T_2 =$

$6\sqrt{5}\sum_{i=1}^{n}[6(Y_1-1/2)^2-1/2]/\sqrt{31n}$, $Y_i=1-\exp(-X_i/\bar{X})$, $\rho=-0,6956$. *Nurodymas.* Pakartokite [4], 3.4.2 ir 3.4.3 teoremų įrodymus, kai yra eksponentinis skirstinys (vienas mastelio parametras).

III.2.14. Statistikos $\hat{T}_2=\sum_i(\ln(1-Y_i)+1)/\sqrt{n}$ skirstinys išsigimęs. Todėl kriterijų sudarome naudodami tik statistiką $\hat{T}_1=\sum_i \ln Y_i+1/\sqrt{n\sigma_{11}}$, $\sigma_{11}=\pi^2(12-\pi^2)/36$. Hipotezė atmetama asimptotiniu reikšmingumo lygmens α kriterijumi, kai $\hat{T}_1^2>\chi^2_\alpha(1)$.

III.2.15. Statistikos T ir $\hat{T}_1^2$ įgijo reikšmes 11,80085 ir 7,1079; atitinkamos P reikšmės 0,0027 ir 0,0077. Hipotezė atmestina.

III.2.16. Statistikos T ir $\tilde{T}$ įgijo reikšmes 0,2714 ir 0,4506; atitinkamos P reikšmės 0,8731 ir 0,7983. Hipotezė neatmetama.

III.2.17. Statistikos T ir $\tilde{T}$ įgijo reikšmes 0,2714 ir 0,4506; atitinkamos P reikšmės 0,8731 ir 0,7983. Hipotezė neatmetama.

III.3. Kriterijai, grindžiami empiriniais procesais

III.3.1. Pratimai

III.3.1. Remiantis paprastąja imtimi $\mathbf{X}=(X_1,...,Z_n)^T$, gauta stebint a. d. X, kurio pasiskirstymo funkcija priklauso tolydžiųjų skirstinių šeimai $\mathcal{F}$, tikrinama paprastoji hipotezė

$$H_0: F(x)\equiv F_0(x);$$

čia $F_0(x)$ žinoma šeimos $\mathcal{F}$ pasiskirstymo funkcija. Tegu $\hat{F}_n(x)$ yra empirinė pasiskirstymo funkcija.

Įrodykite, kad jei H_0 teisinga, tai Kolmogorovo ir Smirnovo statistikos

$$D_n=\sup_{-\infty<x<\infty}|\hat{F}_n(x)-F_0(x)|,$$

Kramerio ir Mizeso statistikos

$$C_n=\int_{-\infty}^{\infty}(\hat{F}_n(x)-F_0(x))^2 d(F_0(x)),$$

Anderseno ir Darlingo statistikos

$$A_n=\int_{-\infty}^{\infty}\frac{(\hat{F}_n(x)-F_0(x))^2}{F_0(x)(1-F_0(x))}d(F_0(x))$$

skirstiniai nepriklauso nuo $F_0(x)$, o priklauso tik nuo imties didumo n.

III.3.2. (**III.3.1** pratimo tęsinys). Tegu $Y_i=F_0(X_i), i=1,...,n$. Įrodykite, kad Kolmogorovo ir Smirnovo statistikos realizacija gali būti surasta tokiu būdu

$$D_n=\max(D_n^+,D_n^-),$$

$$D_n^+=\max 1\le i\le n(\frac{i}{n}-Y_{(i)}),\quad D_n^-=\max 1\le i\le n(Y_{(i)}-\frac{i-1}{n}).$$

III.3.3. (**III.3.1** pratimo tęsinys). Įrodykite, kad Kramerio ir Mizeso statistikos C_n bei Anderseno ir Darlingo statistikos A_n realizacijos gali būti surastos tokiu būdu

$$nC_n = \frac{1}{12n} + \sum_{i=1}^{n}(Y_{(i)} - \frac{2i-1}{2n})^2,$$

$$nA_n = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)[\ln(Y_{(i)}) + \ln(1 - Y_{(n-i+1)})].$$

III.3.4. Jeigu hipotezė H_0 (žr. **III.3.1** pratimą) teisinga ir $n \to \infty$, tai

$$nC_n \xrightarrow{d} C = \int_0^1 B^2(t)dt, \quad nA_n \xrightarrow{d} A = \int_0^1 \frac{B^2(t)}{t(1-t)}dt,$$

čia $B(t)$ yra intervalo $[0,\ 1]$ Brauno tiltas (žr. [4], 8.2.3 skyrelį). Raskite a. d. C ir A pirmuosius du momentus.

III.3.5. (**III.3.4** pratimo tęsinys). Raskite statistikų nC_n ir nA_n pirmuosius du momentus.

III.3.6. Turime didumo n imtį, gautą stebint tolydųjį a. d. su pasiskirstymo funkcija $F(x)$. Nurodykite sritį, į kurią pasiskirstymo funkcija $F(x)$ patenka su tikimybe, ne mažesne už Q.

III.3.7. Tarkime, X yra diskretus a. d., kurio galimos reikšmės 0, 1, 2...., o jų įgijimo tikimybės $p_k = \mathbf{P}\{X = k\}, k = 0, 1, ...$ Įrodykite, kad a. d.

$$Z = \sum_{k=0}^{X-1} p_k + p_X Y, \quad \sum_{i=1}^{-1} = 0,$$

yra tolygiai pasiskirstęs intervale $(0,\ 1)$, kai $Y \sim U(0,\ 1)$ ir nepriklauso nuo X.

III.3.8. (**III.3.7** pratimo tęsinys). Tarkime, $X_1, ..., X_n$ yra paprastoji imtis diskrečiojo a. d. X, o $\hat{F}_n(x)$ – empirinė pasiskirstymo funkcija. Reikia patikrinti hipotezę $H : \mathbf{E}(\hat{F}_n(x)) = F_0(x), |x| < \infty$; čia $F_0(x)$ – visiškai nusakyta diskrečioji pasiskirstymo funkcija. Remdamiesi **III.3.7** pratimu sukonstruokite randomizuotus Kolmogorovo ir Smirnovo, Kramerio ir Mizeso, Anderseno ir Darlingo kriterijų analogus hipotezei H_0 tikrinti.

III.3.9. (**III.3.8** pratimo tęsinys). Remdamiesi **III.3.8** pratime aptartu kriterijumi atlikite **III.1.3** pratimo užduotį.

III.3.10. (**III.3.8** pratimo tęsinys). Remdamiesi **III.3.8** pratime aptartu kriterijumi atlikite **III.1.1** pratimo užduotį.

III.3.11. Remdamiesi Kolmogorovo ir Smirnovo, Kramerio ir Mizeso bei Anderseno ir Darlingo kriterijais patikrinkite hipotezę, kad **II.1.18** pratimo imtis gauta stebint a) normalųjį a. d., kai jo parametrais imame DT įvertinių $\hat{\mu}$ ir $\hat{\sigma}^2$ realizacijas, b) normalųjį a. d.

III.3.12. Remdamiesi Kolmogorovo ir Smirnovo, Kramerio ir Mizeso bei Anderseno ir Darlingo kriterijais patikrinkite hipotezę, kad **II.1.19** pratimo imtis gauta stebint a) lognormalųjį a. d., kai jo parametrais imame DT įvertinių $\hat{\mu}$ ir $\hat{\sigma}$ realizacijas, b) lognormalųjį a. d.

III.3.13. Kontroliuojant staklių stabilumą, kiekvieną valandą paimama 20 gaminių ir remiantis jų tam tikro parametro matavimo rezultatais apskaičiuojamas nepaslinktasis dispersijos įvertinys s^2. Lentelėje pateikta 47 įvertinių realizacijos.

0,1225	0,1764	0,1024	0,1681	0,0841	0,0729	0,1444	0,0900
0,0961	0,1369	0,1521	0,1089	0,1296	0,1225	0,1156	0,1681
0,0676	0,0784	0,1024	0,1156	0,1024	0,0676	0,1225	0,1521
0,1369	0,1444	0,1521	0,1024	0,1089	0,1600	0,0961	0,1600
0,1024	0,1369	0,1089	0,1681	0,1296	0,1521	0,1600	0,0576
0,0784	0,1089	0,1056	0,1444	0,1296	0,1024	0,1369	

Remdamiesi Kolmogorovo ir Smirnovo, Kramerio ir Mizeso bei Anderseno ir Darlingo kriterijais, patikrinkite hipotezę, kad prietaisas buvo stabilus (pagal matuojamo parametro reikšmių nukrypimus). Laikykite, kad tokiu atveju matuojamasis parametras pasiskirstęs pagal normalųjį dėsnį su dispersija $\sigma^2 = 0,1090$.

III.3.14. Lentelėje pateikti dviejų eksperimentų su musėmis rezultatai. Pirmame eksperimente tam tikrais nuodais musės veikiamos 30 sekundžių, antrame – 60 sekundžių. Paralyžiuojantį nuodų poveikį apibūdina reakcijos laikas (X_{1i} pirmame ir X_{2i} antrame eksperimente), praėjęs nuo musės sąlyčio su nuodais iki to momento, kai musė nebegali stovėti.

i	X_{1i}	i	X_{1i}	i	X_{2i}	i	X_{2i}
1	3,1	9	53,1	1	3,3	9	56,7
2	9,4	10	59,4	2	10,0	10	63,3
3	15,6	11	65,6	3	10,7	11	70,0
4	21,9	12	71,9	4	23,3	12	76,7
5	28,1	13	78,1	5	30,0	13	83,3
6	34,4	14	84,4	6	36,7	14	90,0
7	40,6	15	90,6	7	43,3	15	96,7
8	46,9	16	96,9	8	50,0		

Remdamiesi Kolmogorovo ir Smirnovo bei Kramerio ir Mizeso dviejų imčių kriterijais, patikrinkite homogeniškumo hipotezę.

III.3.15. Sudalinkite **II.1.18** pratimo duomenis į dvi imtis (pirmieji 5 ir likusieji 5 stulpeliai). Remdamiesi Kolmogorovo ir Smirnovo bei Kramerio ir Mizeso dviejų imčių kriterijais, patikrinkite homogeniškumo hipotezę.

III.3.16. Sudalinkite **II.1.19** pratimo duomenis į dvi imtis (pirmieji 5 ir likusieji 10 stulpelių). Remdamiesi Kolmogorovo ir Smirnovo bei Kramerio ir Mizeso dviejų imčių kriterijais patikrinkite homogeniškumo hipotezę.

III.3.2. Sprendimai, atsakymai, nurodymai

III.3.1. Tegu $Y_i = (F_0(X_i))$. Tada $\boldsymbol{Y} = (Y_1, ..., Y_n)^T$ yra paprastoji imtis a. d. $Y \sim U(0,\ 1)$. Pažymėję $\hat{G}_n(y)$ imties $\boldsymbol{Y}$ empirinę pasiskirstymo funkciją, gauname

$$D_n = \sup_{0\le y\le 1} |\hat{G}_n(y) - y|, \quad C_n = \int_0^1 (\hat{G}_n(y) - y)^2 dy,$$

$$A_n = \int_0^1 \frac{(\hat{G}_n(y) - y)^2}{y(1-y)} dy.$$

Taigi statistikų D_n, C_n, A_n skirstiniai nuo F_0 nepriklauso.

III.3.2. Žr. [4], 4.2.1 teoremą.

III.3.3. Žr. [4], 4.3.1 teoremą.

III.3.4. $\mathbf{E}C = 1/6, \mathbf{V}C = 1/45$; $\mathbf{E}A = 1, \mathbf{V}A = 2\pi^2/3 - 6$. *Nurodymas.* $\mathbf{E}C = \int_0^1 \mathbf{E}(B^2(t))dt$, $\mathbf{E}(C^2) = 2\int_0^1\int_0^t \mathbf{E}(B^2(t)B^2(s))dsdt$. Pasinaudokite tuo, kad $B(t) \sim N(0, t(1-t)), 0 \le t \le 1; (B(s), B(t))^T \sim N_2(\mathbf{0}, \mathbf{\Sigma}), \sigma_{11} = s(1-s), \sigma_{22} = t(1-t), \sigma_{12} = s(1-t), 0 \le s \le t \le 1$.

III.3.5. $\mathbf{E}(nC_n) = 1/6, \mathbf{V}(nC_n) = 1/45 - 1/(60n)$; $\mathbf{E}(nA_n) = 1, \mathbf{V}(nA_n) = 2\pi^2/3 - 6 + (10 - \pi^2)/n$.

Nurodymas. $\mathbf{E}(nC_n) = \int_0^1 n\mathbf{E}(\hat{G}_n(y) - y)^2 dy$, $\mathbf{E}(nC_n)^2) = 2\int_0^1\int_0^y n^2\mathbf{E}[(\hat{G}_n(x) - x)^2(\hat{G}_n(y) - y)^2]dxdy$. Atsitiktinis dydis $n\hat{G}_n(y) \sim B(n, y), 0 < y < 1$; atsitiktinis vektorius $(n\hat{G}_n(x), n(\hat{G}_n(y) - \hat{G}_n(x)), n(1 - \hat{G}_n(y)))^T \sim \mathcal{P}_3(n, (x, y - x, 1 - y)), 0 \le x \le y \le 1$. Randame po integralų ženklais parašytų momentų išraiškas ir jas integruojame.

III.3.6. Tegu $\hat{F}_n(x)$ yra empirinė pasiskirstymo funkcija, $D_n = \sup_{-\infty<x<\infty} |\hat{F}_n(x) - F(x)|$ Kolmogorovo ir Smirnovo statistika ir $D_\alpha(n)$ statistikos D_n lygmens α kritinė reikšmė. Tada

$$\mathbf{P}\{\sup_{-\infty<x<\infty} |\hat{F}_n(x) - F(x)| < D_\alpha(n)\} = Q = 1 - \alpha,$$

arba

$$\mathbf{P}\{\hat{F}_n(x) - D_\alpha(n) < F(x) < \hat{F}_n(x) + D_\alpha(n), \ \ \forall x \in \mathbf{R}\} = Q.$$

Kadangi $F(x)$ įgyja reikšmes iš intervalo $[0,\ 1]$, tai tikimybė nepakis, jeigu apatinę ribą pakeisime $\max(0, \hat{F}_n(x) - D_\alpha(n))$, o viršutinę ribą į $\min(\hat{F}_n(x) + D_\alpha(n),\ 1)$.

III.3.7. Pažymėkime $P_k = p_0 + ... + p_k, k = 0, 1, ...$ Pagal pilnosios tikimybės formulę

$$G(z) = \mathbf{P}\{Z \le z\} = \sum_{k=0}^{\infty} \mathbf{P}\{Z \le z | X = k\} p_k;$$

$$\mathbf{P}\{Z \le z | X = k\} = \mathbf{P}\{Y \le (z - P_{k-1})/p_k\};$$

pastaroji tikimybė lygi 0, kai $z < P_{k-1}$, lygi $(z - P_{k-1})/p_k$, kai $P_{k-1} \le z \le P_k$, ir lygi 1, kai $z > P_k$. Tankio funkcija

$$g(z) = G'(z) = 1, \quad \text{kai} \quad 0 \le z \le 1.$$

III.3.8. Hipotezę H pakeiskime hipoteze $H' : \mathbf{E}(\hat{G}_n(z)) = G(z), o < z < 1$. Čia $\hat{G}_n(z)$ – empirinė pasiskirstymo funkcija imties $Z_i = F_0(X_i) + p_{X_i}Y_k,\ i = 1, ..., n,\ k = 1, ..., n$; $Y_1, ..., Y_n$ – nepriklausanti nuo $X_1, ..., X_n$ paprastoji a. d. $Y \sim U(0,\ 1)$ imtis, o $G(z)$ – tolygiojo skirstinio $U(0,\ 1)$ pasiskirstymo funkcija. Taip gauname randomizuotą, pavyzdžiui, Kolmogorovo ir Smirnovo kriterijaus analogą diskretiesiems skirstiniams.

III.3.9. Atlikę **III.3.8** pratime nurodytą randomizaciją, gauname statistikų realizacijas $D_n = 0,0196, C_n = 0,0603, A_n = 0,3901$. Atitinkamos P reikšmės $> 0,25$. Hipotezė neatmetama.

III.3.10. Atlikę **III.3.8** pratime nurodytą randomizaciją, gauname statistikų realizacijas $D_n = 0,0305, C_n = 0,0814, A_n = 0,5955$. Atitinkamos P reikšmės $> 0,25$. Hipotezė neatmetama.

III.3.11. a) Kolmogorovo ir Smirnovo, Kramerio ir Mizeso bei Anderseno ir Darlingo statistikų realizacijos yra 0,0828, 0,1139, 0,7948, o atitinkamos P reikšmės 0,5018, 0,5244, 0,4834. Hipotezė neatmetama. b) Taikydami modifikuotuosius kriterijus gauname tas pačias statistikų realizacijas, o P reikšmės yra 0,0902, 0,0762, 0,0399. Hipotezės teisingumas kelia abejonių.

III.3.12. a) Kolmogorovo ir Smirnovo, Kramerio ir Mizeso bei Anderseno ir Darlingo statistikų realizacijos yra 0,0573, 0,0464, 0,3411, o atitinkamos P reikšmės 0,6927, 0,8911, 0,8978. Hipotezė neatmetama. b) Taikydami modifikuotuosius kriterijus gauname, kad P reikšmė pirmu atveju $> 0,15$, o kitais dviem atvejais $> 0,25$. Atsakymas nepakinta.

III.3.13. Kolmogorovo ir Smirnovo, Kramerio ir Mizeso bei Anderseno ir Darlingo statistikų realizacijos yra 0,2547, 0,8637, 4,3106, o atitinkamos P reikšmės 0,0041, 0,0049, 0,0065. Hipotezė atmestina.

III.3.14. Statistika $D_{m,n}$ įgijo reikšmę 0,075. Kritinė reikšmė $D_{0,05}(15,\ 16) = 0,475$. Asimptotinė P reikšmė $pv_a = 1 - K(0,2087) \approx 1$. Hipotezė neatmetama. Kramerio ir Mizeso statistikos reikšmė 0,0032 ir $pv_a = 1 - a_1(0,0032) \approx 1$. Hipotezė neatmetama.

III.3.15. Statistika $D_{m,n}$ įgijo reikšmę 0,16, $pv = 0,471$ ir $pv_a = 1 - K(1,1) = 0,5441$. Kramerio ir Mizeso statistikos reikšmė 0,2511 ir $pv_a = 1 - a_1(0,2511) = 0,187$. Hipotezė neatmetama.

III.3.16. Statistika $D_{m,n}$ įgijo reikšmę 0,13, $pv = 0,5227$ ir $pv_a = 0,6262$. Kramerio ir Mizeso statistikos reikšmė 0,1668 ir $pv_a = 1 - a_1(0,1668) = 0,3425$. Hipotezė neatmetama.

III.4. Ranginiai kriterijai

III.4.1. Nepriklausomumo hipotezių tikrinimas

III.4.1. Tegu $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis absoliučiai tolydaus a. d. X, o $X_{(1)} \leq ... \leq X_{(n)}$ yra pozicinės statistikos. Imties elemento X_i rangu R_i vadiname to elemento eilės numerį variacinėje eilutėje $(X_{(1)}, ..., X_{(n)})$, t. y.

$$R_i = \text{rangas}(X_i) = j, \quad \text{jeigu} \quad X_i = X_{(j)}.$$

Raskite a. d. R_i vidurkį ir dispersiją.

III.4.2. Krakmolo kiekis bulvėse nustatomas dviem būdais. Norint palyginti tuos būdus, buvo paimta 16 bulvių ir kiekvienos iš jų krakmolo kiekis nustatytas abiem būdais. Gauti rezultatai surašyti lentelėje (X_i – pirmu būdu, o Y_i – antruoju).

i	X_i	Y_i	i	X_i	Y_i
1	21,7	21,5	9	14,0	13,9
2	18,7	18,7	10	17,2	17,0
3	18,3	18,3	11	21,7	21,4
4	17,5	17,4	12	18,6	18,6
5	18,5	18,3	13	17,9	18,0
6	15,6	15,4	14	17,7	17,6
7	17,0	16,7	15	18,3	18,5
8	16,6	16,9	16	15,6	15,5

Patikrinkite hipotezę apie a. d. X ir Y nepriklausomumą, remdamiesi Spirmeno ir Kendalo ranginiais koreliacijos koeficientais.

III.4.3. Tiriant specialios sėjamosios efektyvumą, 10 sklypelių buvo sėjama paprasta sėjamąja ir 10 sklypelių – specialia sėjamąja, paskui buvo lyginamas derlingumas. Norint eliminuoti dirvožemio įtaką, 20 vienodo ploto sklypelių buvo taip sugrupuota poromis, kad jie būtų greta vienas kito. Metant monetą, buvo nusprendžiama, kuriame iš dviejų gretimų sklypelių sėti specialia sėjamąja. Rezultatai pateikti lentelėje (X_i – derlingumas sėjant specialia sėjamąja, Y_i – paprasta sėjamąja).

i	X_i	Y_i	i	X_i	Y_i
1	8,0	5,6	6	7,7	6,1
2	8,4	7,4	7	7,7	6,6
3	8,0	7,3	8	5,6	6,0
4	6,4	6,4	9	5,6	5,5
5	8,6	7,5	10	6,2	5,5

Patikrinkite hipotezę apie a. d. X ir Y nepriklausomumą, remdamiesi Spirmeno ir Kendalo ranginiais koreliacijos koeficientais.

III.4.4. Patikrinkite atsitiktinumo hipotezę pagal **II.1.18** pratimo duomenis.

III.4.5. Patikrinkite atsitiktinumo hipotezę pagal **II.1.19** pratimo duomenis.

III.4.2. Homogeniškumo hipotezių tikrinimas

III.4.6. Remdamiesi Vilkoksono ir Van der Vardeno kriterijais patikrinkite hipotezę apie nuodų poveikio vienodumą pagal **III.3.15** pratimo duomenis.

III.4.7. Įrodykite, kad gama skirstinio $G(1,\eta)$ atveju Vilkoksono kriterijaus ASE, palyginti su Stjudento kriterijumi, kai alternatyvos poslinkio, yra

$$e(W,t) = A(\eta) = \frac{3\eta}{2^{4(\eta-1)}((2\eta-1)B(\eta,\eta))^2}.$$

Patikrinkite, kad $A(\eta) > 1,25$, kai $\eta \le 3$; $A(\eta) \to \infty$, kai $\eta \to 1/2$; $A(\eta) \to 3/\pi$, kai $\eta \to \infty$.

III.4.8. Tikrinama hipotezė, kad impulso atpažinimo paklaida nepriklauso nuo jo intensyvumo. Buvo atlikti du nepriklausomi eksperimentai. Impulsas, kurio intensyvumas 10 sąlyginių vienetų, buvo įvertintas 9, 9, 8, 10, 12, 13, 10, 11 vienetų; impulsas, kurio intensyvumas 20 sąlyginių vienetų, – 15, 16, 17, 23, 22, 20, 21, 24, 27. Remdamiesi ranginiais kriterijais, kai alternatyvos yra mastelio, patikrinkite, ar gauti duomenys neprieštarauja iškeltai hipotezei.

III.4.9. Suskaidykime **II.1.18** pratimo duomenis į 10 vienodo didumo imčių (duomenys surašyti skirtingose eilutėse). Remdamiesi Kruskalo ir Voliso ir inversijų skaičiumi grindžiamais kriterijais, patikrinkite hipotezę, kad visais atvejais buvo stebimas tas pats atsitiktinis dydis.

III.4.10. Trijose gamyklose buvo testuojami kineskopai. Jų darbo trukmė (mėnesiais iki pirmo gedimo) surašyti pateikiamoje lentelėje. Ar galima tvirtinti, kad visose trijose gamyklose gaminamų kineskopų darbo trukmė yra vienoda?

1 gam.	41	70	26	89	62	54	46	77	34	51		
2 gam.	30	69	42	60	44	74	32	47	45	37	52	81
3 gam.	23	35	29	38	21	53	31	25	36	50	61	

III.4.11. Lentelėje pateikti avarijų Lietuvos keliuose 1990–1999 metų duomenys.

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
X	5135	6067	4049	4319	3902	4144	4579	5319	6445	6356
Y	933	1093	779	893	765	672	667	725	829	748
Z	5491	6638	4251	4555	4146	4508	5223	6198	7669	7696

Apskaičiuokite konkordancijos koeficientą ir patikrinkite hipotezę dėl a. d. X, Y ir Z priklausomybės. Remdamiesi Spirmeno ir Kendalo ranginiais koreliacijos koeficientais patikrinkite hipotezes dėl a. d. X ir Y priklausomybės; dėl a. d. X ir Z priklausomybės.

III.4.12. Remdamiesi Vilkoksono ženklų kriterijumi priklausomoms imtims, patikrinkite hipotezę dėl a. d. X ir Y skirstinių vienodumo pagal **III.4.2** pratimo duomenis.

III.4.13. Remdamiesi Vilkoksono ženklų kriterijumi priklausomoms imtims, patikrinkite hipotezę dėl a. d. X ir Y skirstinių vienodumo pagal **III.4.3** pratimo duomenis.

III.4.14. Lentelėje pateikti duomenys apie 3 tiekėjų siūlomų 12 skirtingų tipų spausdintuvų kainas.

Tipas	1 tiek.	2 tiek.	3 tiek.	Tipas	1 tiek.	2 tiek.	3 tiek.
1	660	673	658	7	1980	1950	1970
2	790	799	785	8	2300	2295	2310
3	590	580	599	9	2500	2480	2490
4	950	945	960	10	2190	2190	2210
5	1290	1280	1295	11	5590	5500	5550
6	1550	1500	1499	12	6000	6100	6090

Remdamiesi Frydmano kriterijumi patikrinkite hipotezę, kad skirtingų tiekėjų spausdintuvų kainos nesiskiria.

III.4.3. Sprendimai, atsakymai, nurodymai

III.4.1 skyrelis

III.4.1. A. d. R_i įgyja reikšmes $1, 2, ..., n$ su vienodomis tikimybėmis $1/n$. Randame

$$\mathbf{E}R_i = \sum_{j=1}^{n} j\mathbf{P}\{R_i = j\} = (1 + ... + n)/n = \frac{n+1}{2};$$

$$\boldsymbol{V}R_i = \mathbf{E}R_i^2 - (\mathbf{E}R_i)^2 = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12}.$$

III.4.2. Spirmeno ir Kendalo koreliacijos koeficientai (žr. [4], 5.3.1 ir 5.3.2 skyrelius) įgijo reikšmes $r_S = 0,9889$ ir $\tau_b = 0,9422$. Abiejų kriterijų atveju $pv < 0,0001$. Hipotezė atmetama.

III.4.3. Spirmeno ir Kendalo koreliacijos koeficientai įgijo reikšmes $r_S = 0,7822$ ir $\tau_b = 0,6746$. Atitinkamos P reikšmės $pv = 0,0075$ ir $pv = 0,0084$. Hipotezė atmetama.

III.4.4. Spirmeno ir Kendalo koreliacijos koeficientai įgijo reikšmes $r_S = 0,0935$ ir $\tau_b = 0,0621$. Atitinkamos P reikšmės $pv = 0,3547$ ir $pv = 0,3663$. Atmesti hipotezę nėra pagrindo.

III.4.5. Spirmeno ir Kendalo koreliacijos koeficientai įgijo reikšmes $r_S = 0,0643$ ir $\tau_b = 0,0390$. Atitinkamos P reikšmės $pv = 0,4343$ ir $pv = 0,4852$. Atmesti hipotezę nėra pagrindo.

III.4.6. Vilkoksono statistika (žr. [4], 5.5.1 skyrelį) įgijo reikšmę $W = 239$, $Z_{m,n} = -0,0198$; asimptotinė P reikšmė yra $pv_a = 2\Phi(-0,0198) = 0,9842$. Van der Vardeno statistikos (žr. [4], 5.5.4 skyrelį) reikšmė $V = -0,1228$ ir $pv_a = 0,9617$. Atmesti hipotezę nėra pagrindo.

III.4.7. Vilkoksono kriterijaus ASE Stjudento kriterijaus atžvilgiu, kai imties elemento tankio funkcija $f(x)$ yra (žr. [4], 5.5.3 skyrelį)

$$e(W,t) = 12\tau^2 [\int_{-\infty}^{\infty} f^2(x)dx]^2, \quad \tau^2 = \boldsymbol{V}X_i.$$

Skirstinio $G(1,\eta)$ atveju gauname

$$e(W,t) = \frac{3\eta}{2^{4(\eta-1)}((2\eta-1)B(\eta,\eta))^2}.$$

III.4.8. Reikia naudoti homogeniškumo kriterijų, kai alternatyva yra mastelio (žr. [4], 5.5.5 skyrelį). Pagal turimus duomenis Zygelio ir Tjukio, Ansario ir Bredlio, Mūdo, Klotso kriterijų statistikų reikšmės: $S_{ZT} = 91,1667$; $S_{AB} = 47,5$; $S_K = 2,2603$; $S_M = 8,9333$ ir atitinkamos P reikšmės 0,0623; 0,0713; 0,0320; 0,0342. Asimptotinės P reikšmės: 0,0673; 0,0669; 0,0371; 0,0382. Homogeniškumo hipotezė atmestina.

III.4.9. Kruskalo ir Voliso statistika (žr. [4], 5.8 skyrelį) įgijo reikšmę $F_{KW} = 3,9139$ ir $pv_a = 0,9170$. Atmesti hipotezę nėra pagrindo.

III.4.10. Kruskalo ir Voliso statistika įgijo reikšmę $F_{KW} = 6,5490$ ir $pv_a = 0,0378$. Hipotezė atmetama, jei kriterijaus reikšmingumo lygmuo viršija 0,0378.

III.4.11. Kendalo konkordancijos koeficiento (žr. [4], 5.10 skyrelį) reikšmė 0,6444 ir $pv_a = 0,0428$; nepriklausomumo hipotezė atmetama, kai reikšmingumo lygmuo viršija 0,0428. a) Spirmeno ir Kendalo koreliacijos koeficientai įgijo reikšmes $r_S = 0,2242$ ir $\tau_b = 0,2$; P reikšmės yra 0,5334 ir 0,4208. Nepriklausomumo hipotezė neatmetama. b) Spirmeno ir Kendalo koreliacijos koeficientai įgijo reikšmes $r_S = 0,9879$ ir $\tau_b = 0,9556$; P reikšmės abiem atvejais yra $pv < 0,00001$. Nepriklausomumo hipotezė atmetama.

III.4.12. Ranginio ženklų kriterijaus statistikos (žr. [4], 5.6 skyrelį) reikšmė yra $T^+ = 69,5$ ir $pv = 0,0989$. Homogeniškumo hipotezė atmetama, jei reikšmingumo lygmuo viršija 0,0989.

III.4.13. Ranginio ženklų kriterijaus statistikos reikšmė yra $T^+ = 43$ ir $pv = 0,0117$. Homogeniškumo hipotezė atmetama, jei reikšmingumo lygmuo viršija 0,0117.

III.4.14. Frydmano statistika (žr. [4], 5.9 skyrelį) įgijo reikšmę $S_F = 2,5957$; asimptotinė P reikšmė $pv_a = 0,2731$. Duomenys neprieštarauja iškeltai hipotezei.

III.5. Kiti neparametriniai kriterijai

III.5.1. Pratimai

III.5.1. Remdamiesi ženklų kriterijumi patikrinkite hipotezę apie dviejų krakmolo kiekio nustatymo būdų ekvivalentiškumą pagal **III.4.2** pratimo duomenis.

III.5.2. Remdamiesi ženklų kriterijumi patikrinkite hipotezę apie dviejų sėjamųjų vienodą efektyvumą pagal **III.4.3** pratimo duomenis.

III.5.3. Remdamiesi serijų skaičiumi grindžiamu kriterijumi, patikrinkite hipotezę apie nuodų poveikio vienodumą pagal **III.3.15** pratimo duomenis.

III.5.4. Naudodami serijų kriterijų patikrinkite atsitiktinumo hipotezę pagal **II.1.18** pratimo duomenis.

III.5.5. Naudodami serijų kriterijų patikrinkite atsitiktinumo hipotezę pagal **II.1.19** pratimo duomenis.

III.5.6. Visuomenės nuomonės apklausoje tų pačių 3000 rinkėjų buvo klausiama dėl jų nuomonės apie konkrečią parlamentinę partiją prieš rinkimus ir po jų. Prieš rinkimus neigiamą nuomonę išsakė 300 rinkėjų, o praėjus metams po rinkimų – 350. Be to, 150 rinkėjų nepakeitė savo neigiamos nuomonės, 150 rinkėjų neigiama nuomonė pasikeitė į teigiamą, o 200 rinkėjų teigiama nuomonė pasikeitė į neigiamą. Ar pakito partijos reitingas?

III.5.7. (**III.5.6** pratimo tęsinys). Tarkime, kad tie patys 3000 rinkėjų buvo apklausti prieš kitus rinkimus. 270 rinkėjų nuomonė buvo neigiama. Iš jų 180 rinkėjų buvo tokių, kurie pirmose dviejose apklausose turėjo teigiamą nuomonę, ir 70 rinkėjų, kurių nuomonė buvo teigiama vienoje ir neigiama kitoje iš pirmiau buvusių apklausų. Ar pakito partijos reitingas?

III.5.8. Tegu $X_1, ..., X_n$ yra paprastoji imtis a. d. $X \sim N(\mu, \sigma^2)$ ir $Y_i = X_{i+1} - X_1/(1+\sqrt{n}) - n\bar{X}/(n+\sqrt{n})$. Įrodykite, kad $Y_1, ..., Y_{n-1}$ yra vienodai pasiskirstę n. a. d. ir $Y_j \sim N(0, \sigma^2)$.

III.5.9. (**III.5.8** pratimo tęsinys). Įrodykite, kad $Z_1, ..., Z_{n-2}$ yra nepriklausomi a. d., turintys Stjudento skirstinius $Z_j \sim S(n-j-1)$, jei $Z_j = Y_j\sqrt{n-j-1}/(Y_{j+1}^2 + ... + Y_{n-1}^2)$.

III.5.10. Patikrinkite normalumo hipotezę naudodami [4], 6.5.1 skyrelio kriterijus pagal **II.1.18** pratimo duomenis.

III.5.11. Patikrinkite lognormalumo hipotezę naudodami [4], 6.5.1 skyrelio kriterijus pagal **II.1.19** pratimo duomenis.

III.5.12. Patikrinkite puasoniškumo hipotezę naudodami tuščių dėžių kriterijų pagal **II.1.26** pratimo duomenis.

III.5.2. Sprendimai, atsakymai, nurodymai

III.5.1. Tarp 13 skirtumų $S = 3$ yra neigiami ir $pv = 2\mathbf{P}\{S \le 3\} = 0,0923$. Hipotezė atmetama, jei kriterijaus reikšmingumo lygmuo viršija 0,0923.

III.5.2. Tarp 9 skirtumų $S = 1$ yra neigiamas ir $pv = 2\mathbf{P}\{S \le 1\} = 0,0391$. Hipotezė atmetama, jei kriterijaus reikšmingumo lygmuo viršija 0,0195.

III.5.3. Tikrinant homogeniškumo hipotezę remiantis serijų skaičiumi dažniausiai hipotezė atmetama, kai serijų skaičius mažas. Tačiau per didelis serijų skaičius irgi liudija apie nukrypimą nuo atsitiktinio stebinių išsidėstymo. Šiame pavyzdyje serijų skaičius $V = 29$ aiškiai per daug didelis. Natūralu hipotezę atmesti, kai $V \ge 29$; P reikšmė $pv = \mathbf{P}\{V \ge 29\} = 1,2 \cdot 10^{-6}$. Hipotezė atmetama.

III.5.4. Serijų skaičius $V = 54$, $k_1 = 50$, $k_2 = 50$. Gauname $Z^*_{k_1,k_2} = 0,5025$ ir $pv_a = 2(1 - \Phi(0,5025)) = 0,6153$. Atmesti hipotezę nėra pagrindo.

III.5.5. Serijų skaičius $V = 90$, $k_1 = 76$, $k_2 = 74$. Gauname $Z^*_{k_1,k_2} = 2,4906$ ir $pv_a = 2(1 - \Phi(2,4906)) = 0,0128$. Hipotezė atmetama, jei kriterijaus reikšmingumo lygmuo viršija 0,0128.

III.5.6. Naudojame Maknemaros kriterijų. Rinkėjų, kurių nuomonė pakito iš teigiamos į neigiamą, skaičius yra 200, o iš neigiamos į teigiamą – 150. Tikrinama hipotezė $H : p = 1/2$ apie binominio skirstinio tikimybę, kai Bernulio eksperimentų skaičius yra $U_{01} + U_{10} = 350$, o sėkmių skaičius yra $U_{10} = 150$. Gauname $pv = 2\min(\mathbf{P}\{U_{10} \le 150\}, \mathbf{P}\{U_{10} \ge 150\}) = 0,0087$. Hipotezė atmestina.

III.5.7. Duomenų pakanka, kad būtų galima apskaičiuoti Kochreno statistikos (žr. [4], 6.4 skyrelį) reikšmę. Žymėkime ženklu $+$ teigiamą rinkėjo nuomonę ir ženklu $-$ neigiamą nuomonę. Tada rinkėjų, kurių nuomonė buvo $++-$ yra 180; $+--$ arba $-+-$ yra 70, tai $---$ yra 20, o $--+$ yra 130; rinkėjų $+++$ skaičius yra $2500 - 180 = 2320$, $+-+$ arba $-++$ yra 280. Iš čia gauname $x_{01} = 300$, $x_{02} = 350$, $x_{03} = 270$. Be to, yra 20 rinkėjų, kuriems $x_{i.} = 3$, 200 rinkėjų, kuriems $x_{i.} = 2$, 460 rinkėjų, kuriems $x_{i.} = 1$, ir 2320 rinkėjų, kuriems $x_{i.} = 0$. Gauname $3\bar{X}^2 = 920^2/3$. Statistikos Q skaitiklio reikšmė yra $3 \cdot 2(300^2 + 350^2 + 270^2 - 920^2/3)$, o vardiklio reikšmė $3(3 \cdot 20 + 2 \cdot 200 + 460) - (9 \cdot 20 + 4 \cdot 200 + 460)$. Gauname statistikos Q realizaciją $Q = 14,8485$ ir $pv_a = \mathbf{P}\{\chi^2_2 > 14,8485\} = 0,0006$. Hipotezė atmestina.

III.5.8. Reikia patikrinti, kad $\mathbf{V}Y_j = \sigma^2$, $\mathbf{Cov}\,(Y_i, Y_j) = 0, i \ne j$.

III.5.9. Reikia pasiremti Stjudento skirstinio apibrėžimu ir tuo, kad $Y_1, ..., Y_{n-1}$ yra n. a. d. ir $Y_i \sim N(0, \sigma^2)$.

III.5.10. Gauname statistikų realizacijas: $\bar{g}_1 = 2,1598$, $\bar{g}_2 = 0,1524$, $\bar{g}_3 = -1,0849$ ir jas atitinkančias asimptotines P reikšmes 0,0308, 0,8788, 0,2779. Remdamiesi empiriniu asimetrijos koeficientu hipotezę atmetame, jei kriterijaus reikšmingumo lygmuo viršija 0,0308. Atlikę Sarkadi transformaciją gauname paprastąją a. d. $Z \sim U(0, 1)$ imtį $Z_1, ..., Z_{n-2}$. Taikydami Pirsono chi kvadrato kriterijų hipotezei $H : Z \sim U(0, 1)$ tikrinti ir parinkę $k = 8$ vienodų tikimybių intervalus, gauname $X^2_n = 5,5102$ ir $pv_a = 0,5980$.

Kolmogorovo ir Smirnovo, Kramerio ir Mizeso, Anderseno ir Darlingo kriterijų statistikos įgijo reikšmes 0,1071, 0,2629, 1,6185. Normalumo hipotezė neatmetama. Šapiro ir Vilkso statistika įgijo reikšmę 0,9716 ir ją atitinkanti P reikšmė yra 0,0293. Šapiro ir Vilkso kriterijus atmeta normalumo hipotezę, jei kriterijaus reikšmingumo lygmuo viršija 0,0293.

III.5.11. Perėję prie logaritmų gauname statistikų realizacijas: $\bar{g}_1 = 0,5901$, $\bar{g}_2 = 0,8785$, $\bar{g}_3 = 0,3765$ ir jas atitinkančias asimptotines P reikšmes 0,5551, 0,3797,//0,7065. Kriterijai, grindžiami empirinių momentų funkcijomis, lognormalumo hipotezės neatmeta. Atlikę Sarkadi transformaciją gauname paprastąją a. d. $Z \sim U(0,\,1)$ imtį $Z_1, ..., Z_{n-2}$. Taikydami Pirsono chi kvadrato kriterijų hipotezei $H : Z \sim U(0,\,1)$ tikrinti ir parinkę $k = 10$ vienodų tikimybių intervalų, gauname $X_n^2 = 6,1892$ ir $pv_a = 0,7208$. Kolmogorovo ir Smirnovo, Kramerio ir Mizeso ir Anderseno ir Darlingo kriterijų statistikos įgijo reikšmes 0,0548, 0,0752, 0,5594. Atitinkamos P reikšmės viršija 0,25. Lognormalumo hipotezė neatmetama. Šapiro ir Vilkso statistika įgijo reikšmę 0,9920 ir ją atitinkanti P reikšmė yra 0,5644. Šapiro ir Vilkso kriterijus lognormalumo hipotezės taip pat neatmeta.

III.5.12. a) Tuščių dėžių skaičius $Z_0^{(i)}, i = 1, 2, 3, 4$, yra 280, 593, 639, 359. Apskaičiavę $\mathbf{E}(Z_0^{(i)}), \boldsymbol{V}(Z_0^{(i)})$ randame $\bar{Z}_0^{(i)}$ realizacijas: $0,399$; $0,937$; $-0,947$; $-0,826$ ir jas atitinkančias asimptotines P reikšmes 0,345; 0,174; 0,828; 0,796. Atmesti hipotezes nėra pagrindo. b) Tuščių dėžių skaičius Z_0 jungtinėje imtyje 1871. Randame $\bar{Z}_0 = -0,107$ ir $pv_a = 0,543$. Hipotezė neatmetama.

1 priedas. Pagalbinės lentelės

1.P.1 lentelė. Pateikiama pagrindinių žinių apie dažnai naudojamus diskrečiuosius tikimybinius skirstinius.

1.P.2 lentelė. Pateikiama pagrindinių žinių apie dažnai naudojamus absoliučiai tolydžiuosius skirstinius.

Naudojami trumpiniai: nat. sk. – natūralusis skaičius; sv. nen. sk. – sveikasis neneigiamas skaičius; brūkšnys parodo, kad funkcija neturi paprastos išraiškos.

1.P.3 lentelė. Pateikiama žinių apie tikimybinių skirstinių sąryšius.

1.P.1 lentelė. Diskretieji tikimybiniai skirstiniai

Skirstinys	Parametrai	$\mathbf{P}\{X = k\}$	Vidurkis	Dispersija	Generuojančioji funkcija
Binominis	$0 < p < 1$, n-nat.sk.	$C_n^k p^k (1-p)^{n-k}$, $k = 0, ..., n$	np	$np(1-p)$	$(1-p+ps)^n$
Geometrinis	$0 < p < 1$	$p(1-p)^{k-1}$, $k = 1, 2, ...$	$\frac{1}{p}$	$\frac{q}{p^2}$	$\frac{ps}{1-qs}$
Paskalio	$0 < p < 1$, n-nat.sk.	$C_{n+k-1}^k p^n (1-p)^k$ $k = 0, 1, ...$	$\frac{n(1-p)}{p}$	$\frac{n(1-p)}{p^2}$	$\left(\frac{p}{1-qs}\right)^n$
Neigiamasis binominis	$0 < p < 1$, $n > 0$	$\frac{\Gamma(n+k)}{k!\Gamma(n)} p^n (1-p)^k$ $k = 0, 1, ...$	$\frac{n(1-p)}{p}$	$\frac{n(1-p)}{p^2}$	$\left(\frac{p}{1-qs}\right)^n$
Apibendrintas binominis	$\gamma,\ \eta > 0$, n – nat.sk.	$C_n^k \frac{\Gamma(\gamma+\eta)\Gamma(k+\gamma)\Gamma(n-k+\eta)}{\Gamma(\eta)\Gamma(\gamma)\Gamma(n+\gamma+\eta)}$ $k = 0, 1, ..., n$	$n\frac{\gamma}{\gamma+\eta}$	$n\frac{\gamma\eta(\gamma+\eta+n)}{(\gamma+\eta)^2(\gamma+\eta+1)}$	—
Hipergeometrinis	N, M, n–nat.sk. $M \leq N,\ n \leq N$	$\frac{C_M^k C_{N-M}^{n-k}}{C_N^n},\ k \leq \min(n,\ M)$ $\max(0,\ n-N+m) \leq k$	$\frac{nM}{N}$	$np(1-p)\frac{(N-n)}{(N-1)}$ $p = M/N$	—
Neigiamasis hipergeometr.	N, M, n–nat.sk. $M \leq N,\ n \leq N$	$\frac{C_{n+k-1}^{n-1} C_{N-n-k}^{M-n}}{C_N^k}$ $k = 0, 1, ..., N-M$	$\frac{n(N-M)}{M+1}$	žr. (2.25)	—
Pojos	$b,\ r,\ c > 0$	žr. (2.27)	$\frac{nb}{b+r}$	$\frac{nbr(b+r+nc)}{(b+r)^2(b+r+c)}$	—
Puasono	$\lambda > 0$	$\frac{\lambda^k}{k!}e^{-\lambda}$, $k = 0, 1, 2, ...$	λ	λ	$e^{-\lambda(1-s)}$
Sudėtingasis Puasono	$t > 0$, $\lambda > 0$	—	$\lambda t\psi'(1)$	$\lambda t(\psi'(1) + \psi''(1))$	$e^{-\lambda t(1-\psi(s))}$
Logaritminis	$0 < p < 1$	$-\frac{1}{\ln p}\frac{q^k}{k}$, $k = 1, 2, ...$	$-\frac{q}{p\ln p}$	$-\frac{q}{p^2 \ln p}(1 + \frac{q}{\ln p})$	$\frac{\ln(1-sq)}{\ln p}$
Polinominis	$0 < \pi_i < 1$, n -nat.sk. $\pi_1 + ... + \pi_k = 1$	$\frac{n!}{m_1!...m_k!}\pi_1^{m_1}...\pi_k^{m_k}$ $0 \leq m_i \leq n,\ m_i + ... + m_k = n$	$n\boldsymbol{\pi}$	$\sigma_{ii} = n\pi_i(1-\pi_i)$, $\sigma_{ij} = -n\pi_i\pi_j, i \neq j$	$(\pi_1 s_1 + ...+$ $+\pi_k s_k)^n$
Daugiamatis hipergeometr.	$N, M_1, ..., M_k$, n--nat.sk; $\sum_i M_i = N$	$\frac{C_{M_1}^{m_1}...C_{M_k}^{m_k}}{C_N^n}$	$\frac{n}{N}\boldsymbol{M}$	$\sigma_{ii} = np_i(1-p_i)(N-n)/(N-1)$, $\sigma_{ij} = np_ip_j(N-n)/(N-1),\ i \neq j$	—

1.P.2 lentelė. Tolydieji tikimybiniai skirstiniai

Skirstinys	Parametrai	Tankis	Vidurkis	Dispersija	Charakteristinė funkcija
Normalusis	$-\infty < \mu < \infty$, $0 < \sigma < \infty$	$\frac{1}{\sqrt{2\pi}\sigma}\exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}$ $-\infty < x < \infty$	μ	σ^2	$\exp\{i\mu t - \frac{t^2\sigma^2}{2}\}$
Pusiau normalusis	$0 < \sigma < \infty$	$\frac{2}{\sqrt{2\pi}\sigma}\exp\{-\frac{x^2}{2\sigma^2}\},\ \ x > 0$	$\sigma\sqrt{\frac{2}{\pi}}$	$\sigma^2(1 - \frac{2}{\pi})$	$2e^{-\frac{t^2\sigma^2}{2}}\Phi(it\sigma)\}$
Relėjaus	$0 < \sigma < \infty$	$\frac{x}{\sigma^2}\exp\{-\frac{x^2}{2\sigma^2}\},\ \ x > 0$	$\sigma\sqrt{\frac{\pi}{2}}$	$\sigma^2(2 - \frac{\pi}{2})$	—
Maksvelo	$0 < \sigma < \infty$	$\sqrt{\frac{2}{\pi}}\frac{x^2}{\sigma^3}\exp\{-\frac{x^2}{2\sigma^2}\},\ \ x > 0$	$2\sqrt{\frac{2}{\pi}}$	$\frac{3\pi-8}{\pi}$	—
Lognormalusis	$-\infty < \mu < \infty$ $0 < \sigma < \infty$	$\frac{1}{\sqrt{2\pi}\sigma x}\exp\{-\frac{(\ln x-\mu)^2}{2\sigma^2}\}$ $x > 0$	$e^{\mu+\frac{\sigma^2}{2}}$	$e^{2\mu+\sigma}(e^{\sigma^2} - 1)$	—
Koši	$-\infty < \mu < \infty$ $0 < \sigma < \infty$	$\frac{1}{\pi}\frac{\sigma}{\sigma^2+(x-\mu)^2}$ $-\infty < x < \infty$	neegzistuoja	neegzistuoja	$e^{i\mu t-\sigma\vert t\vert}$
Chi kvadrato	n – nat.sk.	$\frac{1}{2^{n/2}\Gamma(n/2)}x^{\frac{n}{2}-1}e^{-\frac{x}{2}},\ \ x > 0$	n	$2n$	$\frac{1}{(1-2it)^{n/2}}$
Necentrinis chi kvadrato	n – nat.sk. $\delta > 0$	žr. (3.18)	$n + \delta$	$2(n + 2\delta)$	$\frac{\exp\{it\sqrt{\delta}-t^2/2\}}{(1-2it)^{(n-1)/2}}$
Stjudento	n – nat.sk.	$\frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)}\left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$ $-\infty < x < \infty$	0, $n > 1$	$\frac{n}{n-2}$, $n > 2$	—
Necentrinis Stjudento	n – nat.sk. $-\infty < \mu < \infty$	žr. (3.23)	$\mu M_n,\ M_n = \sqrt{\frac{n}{2}}\times$ $\times\frac{\Gamma((n-1)/2)}{\Gamma(n/2)},\ n > 1$	$\frac{n}{n-2}+$ $+\mu^2(\frac{n}{n-2} - M_n^2),\ n > 2$	—
Fišerio	m – nat.sk. n – nat.sk.	$\frac{\Gamma((m+n)/2)}{\Gamma(m/2)\Gamma(n/2)}m^{\frac{m}{2}}n^{\frac{n}{2}}\times$ $\times x^{\frac{m}{2}-1}(n + mx)^{-\frac{m+n}{2}}$	$\frac{nm}{n-2}$, $n > 2$	$\frac{2n^2(n+m-2)}{m(n-2)^2(n-4)}$, $n > 4$	—
Necentrinis Fišerio	$m,\ n$ – nat.sk. $\delta > 0$	žr. (3.27)	$\frac{n(m+\delta)}{n-2}$, $n > 2$	$2n^2(\frac{m+2\delta}{(n-2)(n-4)}+$ $+\frac{2(m+\delta)^2}{(n-2)^2(n-4)})$, n>4	—
Gama	$\lambda > 0,\quad \eta > 0$	$\frac{\lambda^\eta}{\Gamma(\eta)}x^{\eta-1}e^{-\lambda x},\quad x > 0$	$\frac{\eta}{\lambda}$	$\frac{\eta}{\lambda^2}$	$\left(\frac{\lambda}{\lambda-it}\right)^\eta$
Eksponentinis	$\lambda > 0$	$\lambda e^{-\lambda x},\ \ x > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda-it}$
Paslinktasis eksponentinis	$\lambda > 0$, $-\infty < \mu < \infty$	$\lambda e^{-\lambda(x-\mu)}$, $\mu < x < \infty$	$\mu + \frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$e^{i\mu t}\frac{\lambda}{\lambda-it}$

1.P.2 lentelės tęsinys. Tolydieji tikimybiniai skirstiniai

Skirstinys	Parametrai	Tankis	Vidurkis	Dispersija	Charakteristinė funkcija
Laplaso	$-\infty < \mu < \infty$ $0 < \lambda < \infty$	$\frac{\lambda}{2}e^{-\lambda\|x-\mu\|}$, $-\infty < x < \infty$	μ	$\frac{2}{\lambda^2}$	$e^{it\mu}\frac{\lambda^2}{\lambda^2+t^2}$
Logistinis	$\theta > 0$, $-\infty < \mu < \infty$	$\frac{1}{\theta}\frac{\exp\{-(x-\mu)/\theta\}}{(1+\exp\{-(x-\mu)/\theta\})^2}$, $-\infty < x < \infty$	μ	$\frac{\theta^2\pi^2}{3}$	$e^{i\mu t}\Gamma(1+it\theta)\times$ $\times\Gamma(1-it\theta)$
Pareto	$\theta > 0,\ \ \alpha > 0$	$\frac{\theta\alpha^\theta}{x^{\theta+1}},\ \ \alpha < x < \infty$	$\frac{\theta\alpha^2}{\theta-1},\ \ \theta > 1$	$\frac{\theta\alpha^2}{(\theta-1)^2(\theta-2)},\ \ \theta > 2$	—
Eksponentinių skirstinių mišinys	$0 < \theta_1,\ \theta_2 < \infty$, $0 < p_1 < 1$, $p_2 = 1 - p_1$	$\frac{p_1}{\theta_1}\exp\{-\frac{x}{\theta_1}\}+$ $+\frac{p_2}{\theta_2}\exp\{-\frac{x}{\theta_2}\}$, $0 < x < \infty$	$\theta_1 p_1 + \theta_2 p_2$	$p_1\theta_1^2 + p_2\theta_2^2+$ $+p_1p_2(\theta_1-\theta_2)^2$	$\frac{p_1}{1-it\theta_1} + \frac{p_2}{1-it\theta_2}$
Loglogistinis	$\theta > 0,\ \ \nu > 0$	$\frac{\nu}{\theta}\left(\frac{x}{\theta}\right)^{\nu-1}\times$ $\times(1+(x/\theta)^\nu)^{-2}$ $0 < x < \infty$	$\theta\Gamma(1+1/\nu)\times$ $\times\Gamma(1-1/\nu)$	$\theta^2(\Gamma(1+2/\nu)\Gamma(1-$ $-2/\nu) - \Gamma^2(1+1/\nu)\times$ $\times\Gamma^2(1-1/\nu))$	—
Veibulo	$\eta > 0$ $\sigma > 0$	$\frac{\eta}{\sigma}\left(\frac{x}{\sigma}\right)^{\eta-1}\exp\{-(\frac{x}{\sigma})^\eta\}$ $0 < x < \infty$	$\sigma\Gamma(\frac{1+\eta}{\eta})$	$\sigma^2\left(\Gamma(\frac{2+\eta}{\eta}) - \Gamma^2(\frac{1+\eta}{\eta})\right)$	—
Apibendrintasis Veibulo	$\theta > 0$, $\sigma > 0$, $\gamma > 0$	$\frac{\eta}{\sigma\gamma}\frac{x}{\sigma}[1+(\frac{x}{\sigma})^\eta]^{(1-\gamma/\gamma)}\times$ $\times\exp\{(1-(1+\frac{x}{\sigma})^\eta)^{1/\gamma}\}$, $0 < x < \infty$	—	—	—
Minimaliųjų reikšmių	$-\infty < \mu < \infty$, $0 < \sigma < \infty$	$\frac{1}{\sigma}\exp\{-\frac{x-\mu}{\sigma} - e^{-\frac{x-\mu}{\sigma}}\}$ $-\infty < x < \infty$	$\mu + \gamma\sigma$, $\gamma = \Gamma'(1)$	$\frac{(\pi\sigma)^2}{6}$	—
Maksimaliųjų reikšmių	$-\infty < \mu < \infty$, $0 < \sigma < \infty$	$\frac{1}{\sigma}\exp\{\frac{x-\mu}{\sigma} - e^{\frac{x-\mu}{\sigma}}\}$ $-\infty < x < \infty$	$\mu - \gamma\sigma$, $\gamma = 0.5772156...$	$\frac{(\pi\sigma)^2}{6}$	—
Beta	$\gamma > 0$, $\eta > 0$	$\frac{\Gamma(\gamma+\eta)}{\Gamma(\gamma)\Gamma(\eta)}x^{\gamma-1}(1-x)^{\eta-1}$, $0 < x < 1$	$\frac{\gamma}{\gamma+\eta}$	$\frac{\gamma\eta}{(\gamma+\eta)^2(\gamma+\eta+1)}$	—
Tolygusis	$\infty < \mu_0 < \mu_1 < \infty$	$\frac{1}{\mu_1-\mu_0},\ \ \mu_0 < x < \mu_1$	$\frac{\mu_1+\mu_0}{2}$	$\frac{(\mu_1-\mu_0)^2}{12}$	$\frac{e^{it\mu_1}-e^{it\mu_0}}{it(\mu_1-\mu_0)}$
k-matis normalusis	$\boldsymbol{\mu} = (\mu_1, ..., \mu_k)^T$, $\boldsymbol{\Sigma} = [\sigma_{ij}]_{k\times k}$	$\frac{1}{(\sqrt{2\pi})^k\sqrt{\|\boldsymbol{\Sigma}\|}}\exp\{-\frac{1}{2}\times$ $\times(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\}$	$\boldsymbol{\mu}$	$\boldsymbol{\Sigma}$	$e^{i(\boldsymbol{t}^T\boldsymbol{\mu})-\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}/2}$, $\boldsymbol{t} \in \mathbf{R}^k$

1.P.3 lentelė. Tikimybinių skirstinių sąryšiai

Atsitiktinis dydis	Funkcija	Funkcijos skirstinys
$X \sim N(\mu,\ \sigma^2)$	$Y = \frac{X-\mu}{\sigma}$	$Y \sim N(0,\ 1)$
$X \sim N(0,\ 1)$	$Y = X^2$	$Y \sim \chi^2(1);\ Y \sim F(1,\ \infty)$
$X_i \sim N(0,\ 1),\ i = 1, ..., n$	$Y = X_1^2 + ... + X_n^2$	$Y \sim \chi^2(n)$
$Z \sim N(0,\ 1),\ X_i \sim N(0,\ 1)$ $i = 1, ..., n$	$Y = \frac{Z}{\sqrt{(X_1^2+...+X_n^2)/n}}$	$Y \sim S(n)$
$Z_i \sim N(0,\ 1),\ i = 1, ..., m,$ $X_j \sim N(0,\ 1),\ j = 1, ..., n$	$Y = \frac{(Z_1^2+...+Z_m^2)/m}{(X_1^2+...+X_n^2)/n}$	$Y \sim F(m,\ n)$
$X_1 \sim G(\lambda,\ \eta_1),$ $X_2 \sim G(\lambda,\ \eta_2)$	$Y = \frac{X_1}{X_1+X_2}$	$Y \sim Be(\eta_1,\ \eta_2)$
$X \sim G(\lambda,\ \eta)$	$Y = 2\lambda X$	$Y \sim \chi^2(2\eta)$
$X_1 \sim \chi^2(n_1),$ $X_2 \sim \chi^2(n_2)$	$Y = \frac{X_1}{X_1+X_2}$	$Y \sim Be(\frac{n_1}{2},\ \frac{n_2}{2})$
$X_1 \sim G(\lambda,\ \eta_1),$ $X_2 \sim G(\lambda,\ \eta_2)$	$Y = \frac{\eta_2 X_1}{\eta_1 X_2}$	$Y \sim F(2\eta_1,\ 2\eta_2)$
$X \sim F(m,\ n)$	$Y = \frac{mX}{n+mX}$	$Y \sim Be(\frac{m}{2},\ \frac{n}{2})$
$X \sim S(n)$	$Y = X^2$	$Y \sim F(1,\ n)$
$X \sim U(0,\ 1)$	$Y = -\frac{\ln X}{\lambda}$	$Y \sim \mathcal{E}(\lambda)$
$X \sim S(n)$	$Y = \frac{n}{n+X^2}$	$Y \sim Be(\frac{n}{2},\ \frac{1}{2})$
$X_i \sim N(\mu_i,\ \sigma_i^2),$ $i = 1, ..., n$	$Y = c_1 X_1 + ... + c_n X_n$	$Y \sim N(\mu,\ \sigma),$ $\mu = \sum_i c_i \mu_i,\ \sigma^2 = \sum_i c_i^2 \sigma_i^2$
$X_i \sim \chi^2(n_i),$ $i = 1, ..., k$	$Y = X_1 + ... + X_k$	$Y \sim \chi^2(n),$ $n = n_1 + ... + n_k$
$X_i \sim \mathcal{P}(\lambda_i),$ $i = 1, ..., n$	$Y = X_1 + ... + X_n$	$Y \sim \mathcal{P}(\lambda),$ $\lambda = \lambda_1 + ... + \lambda_n$
$X_i \sim B(n_i,\ p),$ $i = 1, ..., k$	$Y = X_1 + ... + X_k$	$Y \sim B(n,\ p),$ $n = n_1 + ... + n_k$
$X_i \sim B^-(n_i,\ p),$ $i = 1, ..., k$	$Y = X_1 + ... + X_k - k + 1$	$Y \sim B^-(n,\ p),$ $n = n_1 + ... + n_k - k + 1$
$X_i \sim \mathcal{E}(\lambda),$ $i = 1, ..., n$	$Y = X_1 + ... + X_n$	$Y \sim G(\lambda,\ n),$
$X_i \sim G(\lambda,\ \eta_i),$ $i = 1, ..., n$	$Y = X_1 + ... + X_n$	$Y \sim G(\lambda,\ \eta),$ $\eta = \eta_1 + ... + \eta_n$
$X_i \sim K(\mu_i,\ \sigma_i),$ $i = 1, ..., n$	$Y = X_1 + ... + X_n$	$Y \sim K(\mu,\ \sigma),$ $\mu = \sum \mu_i,\ \sigma = \sum_i \sigma_i$
$X_i \sim K(\mu,\ \sigma),$ $i = 1, ..., n$	$Y = \bar{X} = \frac{X_1+...+X_n}{n}$	$Y \sim K(\mu,\ \sigma),$

Pastaba. Bet kurios funkcijos išraiškoje a. d. laikomi nepriklausomais.

2 priedas. Matematinės statistikos lentelės

2.P.1 lentelė. Nurodytos funkcijos $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\{-\frac{t^2}{2}\}dt$ reikšmės, kai $x = 0,00$ $(0,01)$ $4(0,1)4,9$. Dėl kompaktiškumo devynetai, esantys po kablelio, nerašomi, o tik nurodomas jų skaičius. Pavyzdžiui, $\Phi(3,76) = 0,9^415 = 0,999915$. Kadangi skirstinys $N(0,\ 1)$ simetrinis taško $x = 0$ atžvilgiu, tai $\Phi(-x) = 1 - \Phi(x)$.

2.P.2 lentelė. Pateikiamos standartinio normaliojo skirstinio $N(0,\ 1)$ P-osios kritinės reikšmės z_P (abscisių ašies taškai, tenkinantys lygtį $\Phi(z_P) = 1 - P$, t.y. $z_P = \Phi^{-1}(1 - P)$), kai $P = 0,0010$ $(0,0001)$ $0,003$ $(0,001)$ $0,10$ $(0,01)$ $0,49$. Jeigu $Q > 0,5$, tai iš pasiskirstymo simetrijos gauname, kad $z_Q = -z_{1-Q}$.

2.P.3 lentelė. Pateikiamos $\chi^2(n)$ skirstinio P-osios kritinės reikšmės $\chi^2_P(n)$, t.y. abscisių ašies taškai, tenkinantys lygtį

$$\int_{\chi^2_P(n)}^{\infty} f(x;\ n)dx = P;$$

čia $f(x;\ n)$ – skirstinio $\chi^2(n)$ tikimybinio tankio funkcija. Argumentų reikšmės yra šios: $n = 1(1)\ 20(2)$ $40\ (5)\ 90(10)\ 100$; $P = 0,0005,\ 0,0001,\ 0,01,\ 0,025,\ 0,05,\ 0,10,\ 0,20,\ 0,80,\ 0,90,\ 0,95,\ 0,975,\ 0,99,$ $0,995,\ 0,999,\ 0,9995$.

2.P.4 lentelė. Pateikiamos Stjudento skirstinio $S(n)$ P-osios kritinės reikšmės $t_P = t_P(n)$, t.y. lygties

$$\mathbf{P}\{T_n > x\} = \int_{t_P}^{\infty} f(x|n)dx = P$$

sprendiniai; čia $f(x|n)$ – Stjudento skirstinio su n laisvės laipsnių tikimybinio tankio funkcija. Argumentų reikšmės yra šios: $n = 1(1)20(2)30(4)50(10)100(100)500$; $P = 0,0005,\ 0,001,\ 0,0025,\ 0,005,\ 0,01,\ 0,025,$ $0,05,\ 0,1,\ 0,25$. Paskutinėje eilutėje $(n = \infty)$ yra standartinio normaliojo skirstinio kritinės reikšmės z_P. Naudojantis 2.P.4 lentele, reikia prisiminti, kad dėl simetrijos $t_P(n) = -t_{1-P}(n)$.

2.P.5 lentelė. Pateikiamos Fišerio skirstinio $F(m,\ n)$ P-osios kritinės reikšmės, kai $P = 0,05$, t. y. lygties

$$\int_{F_P(m,n)}^{\infty} f(x|m,n)dx = 0,05$$

sprendiniai; čia $f(x|m,n)$ – Fišerio skirstinio su m ir n laisvės laipsnių tikimybinio tankio funkcija. Argumentų reikšmės yra šios: $m = 1(1)10;\ 12;\ 15;\ 20;\ 24;\ 30;\ 40;\ 60;\ 120;\ \infty$; $n = 1(1)30(10)80;$ $100;\ 150;\ 200;\ 500;\ \infty$. Kai $n = \infty$, tai $F_P(m,n) = \chi^2_P(m)/m$, žr. 2.P.3 lentelę. Pasinaudojus lygybe $F_P(m,\ n)F_{1-P}(n,\ m) = 1$, iš 2.P.2 lentelės galima rasti Fišerio skirstinio kritines reikšmes, kai $P = 0,95$.

2.P.1 lentelė. Standartinio normaliojo skirstinio pasiskirstymo funkcijos reikšmės

x	0	1	2	3	4	5	6	7	8	9
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0.9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	$0,9^303$	$0,9^306$	$0,9^310$	$0,9^313$	$0,9^316$	$0,9^318$	$0,9^321$	$0,9^324$	$0,9^326$	$0,9^329$
3,2	$0,9^331$	$0,9^334$	$0,9^336$	$0,9^338$	$0,9^340$	$0,9^342$	$0,9^344$	$0,9^346$	$0,9^348$	$0,9^350$
3,3	$0,9^352$	$0,9^353$	$0,9^355$	$0,9^357$	$0,9^358$	$0,9^360$	$0,9^361$	$0,9^362$	$0,9^364$	$0,9^365$
3,4	$0,9^366$	$0,9^368$	$0,9^369$	$0,9^370$	$0,9^371$	$0,9^372$	$0,9^373$	$0,9^374$	$0,9^375$	$0,9^376$
3,5	$0,9^377$	$0,9^378$	$0,9^378$	$0,9^379$	$0,9^380$	$0,9^381$	$0,9^381$	$0,9^382$	$0,9^383$	$0,9^383$
3,6	$0,9^384$	$0,9^385$	$0,9^385$	$0,9^386$	$0,9^386$	$0,9^387$	$0,9^387$	$0,9^388$	$0,9^388$	$0,9^389$
3,7	$0,9^389$	$0,9^390$	$0,9^400$	$0,9^404$	$0,9^408$	$0,9^412$	$0,9^415$	$0,9^418$	$0,9^422$	$0,9^425$
3,8	$0,9^428$	$0,9^431$	$0,9^433$	$0,9^436$	$0,9^438$	$0,9^441$	$0,9^443$	$0,9^446$	$0,9^448$	$0,9^450$
3,9	$0,9^452$	$0,9^454$	$0,9^456$	$0,9^458$	$0,9^459$	$0,9^461$	$0,9^463$	$0,9^464$	$0,9^466$	$0,9^467$
4,0	$0,9^468$	$0,9^479$	$0,9^487$	$0,9^515$	$0,9^546$	$0,9^566$	$0,9^579$	$0,9^587$	$0,9^621$	$0,9^652$

2.P.2 lentelė. Standartinio normaliojo skirstinio P-osios kritinės reikšmės

P	0	1	2	3	4	5	6	7	8	9
0,001	3,0902	3,0618	3,0357	3,0115	2,9889	2,9677	2,9478	2,9291	2,9112	2,8943
0,002	2,8782	2,8627	2,8480	2,8338	2,8202	2,8070	2,7943	2,7822	2,7703	2,7589
0,003	2,7478	2,7370	2,7266	2,7164	2,7065	2,6968	2,6874	2,6783	2,6693	2,6606
0,004	2,6521	2,6437	2,6356	2,6276	2,6197	2,6121	2,6045	2,5972	2,5899	2,5828
0,005	2,5758	2,5690	2,5622	2,5556	2,5491	2,5427	2,5364	2,5302	2,5241	2,5181
0,006	2,5121	2,5063	2,5006	2,4949	2,4893	2,4838	2,4783	2,4730	2,4677	2,4624
0,007	2,4573	2,4522	2,4471	2,4422	2,4372	2,4324	2,4276	2,4228	2,4181	2,4135
0,008	2,4089	2,4044	2,3999	2,3955	2,3911	2,3867	2,3824	2,3781	2,3739	2,3698
0,009	2,3656	2,3615	2,3575	2,3535	2,3495	2,3455	2,3416	2,3378	2,3339	2,3301
0,010	2,3263	2,3226	2,3189	2,3152	2,3116	2,3080	2,3044	2,3009	2,2973	2,2938
0,011	2,2904	2,2869	2,2835	2,2801	2,2768	2,2734	2,2701	2,2668	2,2636	2,2603
0,012	2,2571	2,2539	2,2508	2,2476	2,2445	2,2414	2,2383	2,2353	2,2322	2,2292
0,013	2,2262	2,2232	2,2203	2,2173	2,2144	2,2115	2,2086	2,2058	2,2029	2,2001
0,014	2,1973	2,1945	2,1917	2,1890	2,1862	2,1835	2,1808	2,1781	2,1754	2,1727
0,015	2,1701	2,1675	2,1648	2,1622	2,1596	2,1571	2,1545	2,1520	2,1494	2,1469
0,016	2,1444	2,1419	2,1394	2,1370	2,1345	2,1321	2,1297	2,1272	2,1248	2,1225
0,017	2,1201	2,1177	2,1154	2,1130	2,1107	2,1084	2,1061	2,1038	2,1015	2,0992
0,018	2,0969	2,0947	2,0924	2,0902	2,0880	2,0858	2,0836	2,0814	2,0792	2,0770
0,019	2,0749	2,0727	2,0706	2,0684	2,0663	2,0642	2,0621	2,0600	2,0579	2,0558
0,020	2,0537	2,0517	2,0496	2,0476	2,0456	2,0435	2,0415	2,0395	2,0375	2,0355
0,021	2,0335	2,0315	2,0296	2,0276	2,0257	2,0237	2,0218	2,0198	2,0179	2,0160
0,022	2,0141	2,0122	2,0103	2,0084	2,0065	2,0047	2,0028	2,0009	1,9991	1,9972
0,023	1,9954	1,9936	1,9917	1,9899	1,9881	1,9863	1,9845	1,9827	1,9809	1,9791
0,024	1,9774	1,9756	1,9738	1,9721	1,9703	1,9686	1,9669	1,9651	1,9634	1,9617
0,025	1,9600	1,9583	1,9566	1,9549	1,9532	1,9515	1,9498	1,9481	1,9465	1,9448
0,026	1,9431	1,9415	1,9398	1,9382	1,9366	1,9349	1,9333	1,9317	1,9301	1,9284
0,027	1,9268	1,9252	1,9236	1,9220	1,9205	1,9189	1,9173	1,9157	1,9142	1,9126
0,028	1,9110	1,9095	1,9079	1,9064	1,9048	1,9033	1,9018	1,9003	1,8987	1,8972
0,029	1,8957	1,8942	1,8927	1,8912	1,8897	1,8882	1,8867	1,8852	1,8837	1,8823
0,03	1,8808	1,8663	1,8522	1,8384	1,8250	1,8119	1,7991	1,7866	1,7744	1,7624
0,04	1,7507	1,7392	1,7279	1,7169	1,7060	1,6954	1,6849	1,6747	1,6646	1,6546
0,05	1,6449	1,6352	1,6258	1,6164	1,6072	1,5982	1,5893	1,5805	1,5718	1,5632
0,06	1,5548	1,5464	1,5382	1,5301	1,5220	1,5141	1,5063	1,4985	1,4909	1,4833
0,07	1,4758	1,4684	1,4611	1,4538	1,4466	1,4395	1,4325	1,4255	1,4187	1,4118
0,08	1,4051	1,3984	1,3917	1,3852	1,3787	1,3722	1,3658	1,3595	1,3532	1,3469
0,09	1,3408	1,3346	1,3285	1,3225	1,3165	1,3106	1,3047	1,2988	1,2930	1,2873
0,1	1,2816	1,2265	1,1750	1,1264	1,0803	1,0364	0,9945	0,9542	0,9154	0,8779
0,2	0,8416	0,8064	0,7722	0,7388	0,7063	0,6745	0,6433	0,6128	0,5828	0,5534
0,3	0,5224	0,4959	0,4677	0,4399	0,4125	0,3853	0,3585	0,3319	0,3055	0,2793
0,4	0,2533	0,2275	0,2019	0,1764	0,1510	0,1257	0,1004	0,0753	0,0502	0,0251

2.P.3 lentelė. Chi kvadrato skirstinio P-osios kritinės reikšmės

$n\diagdown P$	0,9995	0,999	0,995	0,99	0,975	0,95	0,90	0,80
1	$0,0^6393$	$0,0^5157$	$0,0^4393$	$0,0^3157$	$0,0^3982$	0,00393	0,0158	0,0642
2	0,00100	0,00200	0,0100	0,0201	0,0506	0,103	0,211	0,446
3	0,0153	0,0243	0,0717	0,115	0,216	0,352	0,584	1,005
4	0,0639	0,0908	0,207	0,297	0,484	0,711	1,064	1,649
5	0,158	0,210	0,412	0,554	0,831	1,145	1,610	2,343
6	0,299	0,381	0,676	0,872	1,237	1,635	2,204	3,070
7	0,485	0,598	0,989	1,239	1,690	2,167	2,833	3,822
8	0,710	0,857	1,344	1,646	2,180	2,733	3,490	4,594
9	0,972	1,152	1,535	2,088	2,700	3,325	4,168	5,380
10	1,265	1,479	2,156	2,558	3,247	3,940	4,865	6,179
11	1,587	1,834	2,603	3,053	3,816	4,575	5,578	6,989
12	1,934	2,214	3,074	3,571	4,404	5,226	6,304	7,807
13	2,305	2,617	3,565	4,107	5,009	5,892	7,042	8,634
14	2,697	3,041	4,075	4,660	5,629	6,571	7,790	9,467
15	3,108	3,483	4,601	5,229	6,262	7,261	8,547	10,307
16	3,536	3,942	5,142	5,812	6,908	7,962	9,312	11,152
17	3,980	4,416	5,697	6,408	7,564	8,672	10,085	12,002
18	4,439	4,905	6,265	7,015	8,231	9,390	10,865	12,857
19	4,912	5,406	6,844	7,633	8,907	10,117	11,651	13,716
20	5,398	5,921	7,434	8,260	9,591	10,851	12,443	14,578
22	6,404	6,983	8,643	9,542	10,982	12,338	14,041	16,314
24	7,453	8,085	9,886	10,856	12,401	13,848	15,659	18,062
26	8,538	9,222	11,160	12,198	13,844	15,379	17,292	19,820
28	9,656	10,391	12,461	13,565	15,308	16,928	18,939	21,588
30	10,804	11,588	13,787	14,953	16,791	18,493	20,599	23,364
32	11,979	12,811	15,134	16,362	18,291	20,072	22,271	25,148
34	13,179	14,057	16,501	17,789	19,806	21,664	23,952	26,938
36	14,401	15,324	17,887	19,233	21,336	23,269	25,643	28,735
38	15,644	16,611	19,289	20,691	22,878	24,884	27,343	30,537
40	16,906	17,916	20,707	22,164	24,433	26,509	29,051	32,345
45	20,137	21,251	24,311	25,901	28,366	30,612	33,350	36,884
50	23,461	24,674	27,991	29,707	32,357	34,764	37,689	41,449
55	26,866	28,173	31,735	33,570	36,398	38,958	42,060	46,036
60	30,340	31,738	35,534	37,485	40,482	43,188	46,459	50,641
65	33,877	35,362	39,383	41,444	44,603	47,450	50,883	55,262
70	37,467	39,036	43,275	45,442	48,758	51,739	55,329	59,898
75	41,107	42,757	47,206	49,475	52,942	56,054	59,795	64,547
80	44,791	46,520	51,172	53,540	57,153	60,391	64,278	69,207
85	48,515	50,320	55,170	57,634	61,389	64,749	68,777	73,878
90	52,276	54,155	59,196	61,754	65,647	69,126	73,291	78,558
100	59,896	61,918	67,328	70,065	74,222	77,929	82,358	87,945

2.P.3 lentelės tęsinys. Chi kvadrato skirstinio P-osios kritinės reikšmės

0,20	0,10	0,05	0,025	0,01	0,005	0,001	0,0005	P/n
1,642	2,706	3,841	5,024	6,635	7,879	10,828	12,116	1
3,219	4,605	5,991	7,378	9,210	10,597	13,816	15,202	2
4,642	6,251	7,815	9,348	11,345	12,838	16,266	17,730	3
5,989	7,779	9,488	11,143	13,277	14,860	18,467	19,997	4
7,289	9,236	11,070	12,833	15,086	16,750	20,515	22,105	5
8,558	10,645	12,592	14,449	16,812	18,548	22,458	24,103	6
9,803	12,017	14,067	16,013	18,475	20,278	24,322	26,018	7
11,030	13,362	15,507	17,535	20,090	21,955	26,124	27,868	8
12,242	14,684	16,919	19,023	21,666	23,589	27,877	29,666	9
13,442	15,987	18,307	20,483	23,209	25,188	29,588	31,420	10
14,631	17,275	19,675	21,920	24,725	26,757	31,264	33,137	11
15,812	18,549	21,026	23,337	26,217	28,300	32,909	34,821	12
16,985	19,812	22,362	24,736	27,688	29,819	34,528	36,478	13
18,151	21,064	23,685	26,119	29,141	31,319	36,123	38,109	14
19,311	22,307	24,996	27,488	30,578	32,801	37,697	39,719	15
20,465	23,542	26,296	28,845	32,000	34,267	39,252	41,308	16
21,615	24,769	27,587	30,191	33,409	35,718	40,790	42,879	17
22,760	25,989	28,869	31,526	34,805	37,156	42,312	44,434	18
23,900	27,204	30,144	32,852	36,191	38,582	43,820	45,973	19
25,038	28,412	31,410	34,170	37,566	39,997	45,315	47,498	20
27,301	30,813	33,924	36,781	40,289	42,796	48,268	50,511	22
29,553	33,196	36,415	39,364	42,980	45,559	51,179	53,479	24
31,795	35,563	38,885	41,923	45,642	48,290	54,052	56,407	26
34,027	37,916	41,337	44,461	48,278	50,993	56,892	59,300	28
36,250	40,256	43,773	46,979	50,892	53,672	59,703	62,162	30
38,466	42,585	46,194	49,480	53,486	56,328	62,487	64,995	32
40,676	44,903	48,602	51,966	56,061	58,964	65,247	67,803	34
42,879	47,212	50,998	54,437	58,619	61,581	67,985	70,588	36
45,076	49,513	53,384	56,896	61,162	64,181	70,703	73,351	38
47,269	51,805	55,758	59,342	63,691	66,766	73,402	76,095	40
52,729	57,505	61,656	65,410	69,957	73,166	80,077	82,876	45
58,164	63,167	67,505	71,420	76,154	79,490	86,661	89,561	50
63,577	68,796	73,311	77,380	82,292	85,749	93,168	96,163	55
68,972	74,397	79,082	83,298	88,379	91,952	99,607	102,695	60
74,351	79,973	84,821	89,177	94,422	98,105	105,988	109,164	65
79,715	85,527	90,531	95,023	100,425	104,215	112,317	115,578	70
85,066	91,061	96,217	100,839	106,393	110,286	118,599	121,942	75
90,405	96,578	101,879	106,629	112,329	116,321	124,839	128,261	80
101,054	107,565	113,145	118,136	124,116	128,299	137,208	140,782	90
106,364	113,038	118,752	123,858	129,973	134,247	143,344	146,990	95
111,667	118,498	124,342	129,561	135,807	140,169	149,449	153,167	100

2.P.4 lentelė. Stjudento skirstinio P-osios kritinės reikšmės

$n \diagdown P$	0,25	0,10	0,05	0,025	0,01	0,005	0,0025	0,001	0,0005
1	1,0000	3,0777	6,3138	12,7062	31,8205	63,6567	127,321	318,309	636,619
2	0,8165	1,8856	2,9200	4,3027	6,9646	9,9248	14,0890	22,3271	31,5991
3	0,7649	1,6377	2,3534	3,1824	4,5407	5,8409	7,4533	10,2145	12,9240
4	0,7407	1,5332	2,1318	2,7764	3,7469	4,6041	5,5976	7,1732	8,6103
5	0,7267	1,4759	2,0150	2,5706	3,3649	4,0321	4,7733	5,8934	6,8688
6	0,7176	1,4398	1,9432	2,4469	3,1427	3,7074	4,3168	5,2076	5,9588
7	0,7111	1,4149	1,8946	2,3646	2,9980	3,4995	4,0293	4,7853	5,4079
8	0,7064	1,3968	1,8595	2,3060	2,8965	3,3554	3,8325	4,5008	5,0413
9	0,7027	1,3830	1,8331	2,2622	2,8214	3,2498	3,6897	4,2968	4,7809
10	0,6998	1,3722	1,8125	2,2281	2,7638	3,1693	3,5814	4,1437	4,5869
11	0,6974	1,3634	1,7959	2,2010	2,7181	3,1058	3,4966	4,0247	4,4370
12	0,6955	1,3562	1,7823	2,1788	2,6810	3,0545	3,4284	3,9296	4,3178
13	0,6938	1,3502	1,7709	2,1604	2,6503	3,0123	3,3725	3,8520	4,2208
14	0,6924	1,3450	1,7613	2,1448	2,6245	2,9768	3,3257	3,7874	4,1405
15	0,6912	1,3406	1,7531	2,1314	2,6025	2,9467	3,2860	3,7328	4,0728
16	0,6901	1,3368	1,7459	2,1199	2,5835	2,9208	3,2520	3,6862	4,0150
17	0,6892	1,3334	1,7396	2,1098	2,5669	2,8982	3,2224	3,6458	3,9651
18	0,6884	1,3304	1,7341	2,1009	2,5524	2,8784	3,1966	3,6105	3,9216
19	0,6876	1,3277	1,7291	2,0930	2,5395	2,8609	3,1737	3,5794	3,8834
20	0,6870	1,3253	1,7247	2,0860	2,5280	2,8453	3,1534	3,5518	3,8495
22	0,6858	1,3212	1,7171	2,0739	2,5083	2,8188	3,1188	3,5050	3,7921
24	0,6848	1,3178	1,7109	2,0639	2,4922	2,7969	3,0905	3,4668	3,7454
26	0,6840	1,3150	1,7056	2,0555	2,4786	2,7787	3,0669	3,4350	3,7066
28	0,6834	1,3125	1,7011	2,0484	2,4671	2,7633	3,0469	3,4082	3,6739
30	0,6828	1,3104	1,6973	2,0423	2,4573	2,7500	3,0298	3,3852	3,6460
34	0,6818	1,3070	1,6909	2,0322	2,4411	2,7284	3,0020	3,3479	3,6007
38	0,6810	1,3042	1,6860	2,0244	2,4286	2,7116	2,9803	3,3190	3,5657
42	0,6804	1,3020	1,6820	2,0181	2,4185	2,6981	2,9630	3,2960	3,5377
46	0,6799	1,3002	1,6787	2,0129	2,4102	2,6870	2,9488	3,2771	3,5150
50	0,6794	1,2987	1,6759	2,0086	2,4033	2,6778	2,9370	3,2614	3,4960
60	0,6786	1,2958	1,6706	2,0003	2,3901	2,6603	2,9146	3,2317	3,4602
70	0,6780	1,2938	1,6669	1,9944	2,3808	2,6479	2,8987	3,2108	3,4350
80	0,6776	1,2922	1,6641	1,9901	2,3739	2,6387	2,8870	3,1953	3,4163
90	0,6772	1,2910	1,6620	1,9867	2,3685	2,6316	2,8779	3,1833	3,4019
100	0,6770	1,2901	1,6602	1,9840	2,3642	2,6259	2,8707	3,1737	3,3905
200	0,6757	1,2858	1,6525	1,9719	2,3451	2,6006	2,8385	3,1315	3,3398
300	0,6753	1,2844	1,6499	1,9679	2,3388	2,5923	2,8279	3,1176	3,3233
400	0,6751	1,2837	1,6487	1,9659	2,3357	2,5882	2,8227	3,1107	3,3150
500	0,6750	1,2832	1,6479	1,9647	2,3338	2,5857	2,8195	3,1066	3,3101
∞	0,6745	1,2816	1,6449	1,9600	2,3263	2,5758	2,8070	3,0902	3,2905

2.P.5 lentelė. Fišerio skirstinio su m ir n laisvės laipsnių P-osios kritinės reikšmės

$n\diagdown m$	1	2	3	4	5	6	7	8	9
1	161,45	199,50	215,71	224,58	230,16	233,99	236,77	238,88	240,54
2	18,513	19,000	19,164	19,247	19,296	19,330	19,353	19,371	19,385
3	10,128	9,5521	9,2766	9,1172	9,0135	8,9406	8,8867	8,8452	8,8123
4	7,7086	6,9443	6,5914	6,3882	6,2561	6,1631	6,0942	6,0410	5,9988
5	6,6079	5,7861	5,4095	5,1922	5,0503	4,9503	4,8759	4,8183	4,7725
6	5,9874	5,1433	4,7571	4,5337	4,3874	4,2839	4,2067	4,1468	4,0990
7	5,5914	4,7374	4,3468	4,1203	3,9715	3,8660	3,7870	3,7257	3,6767
8	5,3177	4,4590	4,0662	3,8379	3,6875	3,5806	3,5005	3,4381	3,3881
9	5,1174	4,2565	3,8625	3,6331	3,4817	3,3738	3,2927	3,2296	3,1789
10	4,9646	4,1028	3,7083	3,4780	3,3258	3,2172	3,1355	3,0717	3,0204
11	4,8443	3,9823	3,5874	3,3567	3,2039	3,0946	3,0123	2,9480	2,8962
12	4,7472	3,8853	3,4903	3,2592	3,1059	2,9961	2,9134	2,8486	2,7964
13	4,6672	3,8056	3,4105	3,1791	3,0254	2,9153	2,8321	2,7669	2,7144
14	4,6001	3,7389	3,3439	3,1122	2,9582	2,8477	2,7642	2,6987	2,6458
15	4,5431	3,6823	3,2874	3,0556	2,9013	2,7905	2,7066	2,6408	2,5876
16	4,4940	3,6337	3,2389	3,0069	2,8524	2,7413	2,6572	2,5911	2,5377
17	4,4513	3,5915	3,1968	2,9647	2,8100	2,6987	2,6143	2,5480	2,4943
18	4,4139	3,5546	3,1599	2,9277	2,7729	2,6613	2,5767	2,5102	2,4563
19	4,3807	3,5219	3,1274	2,8951	2,7401	2,6283	2,5435	2,4768	2,4227
20	4,3512	3,4928	3,0984	2,8661	2,7109	2,5990	2,5140	2,4471	2,3928
21	4,3248	3,4668	3,0725	2,8401	2,6848	2,5727	2,4876	2,4205	2,3660
22	4,3009	3,4434	3,0491	2,8167	2,6613	2,5491	2,4638	2,3965	2,3419
23	4,2793	3,4221	3,0280	2,7955	2,6400	2,5277	2,4422	2,3748	2,3201
24	4,2597	3,4028	3,0088	2,7763	2,6207	2,5082	2,4226	2,3551	2,3002
25	4,2417	3,3852	2,9912	2,7587	2,6030	2,4904	2,4047	2,3371	2,2821
26	4,2252	3,3690	2,9752	2,7426	2,5868	2,4741	2,3883	2,3205	2,2655
27	4,2100	3,3541	2,9604	2,7278	2,5719	2,4591	2,3732	2,3053	2,2501
28	4,1960	3,3404	2,9467	2,7141	2,5581	2,4453	2,3593	2,2913	2,2360
29	4,1830	3,3277	2,9340	2,7014	2,5454	2,4324	2,3463	2,2783	2,2229
30	4,1709	3,3158	2,9223	2,6896	2,5336	2,4205	2,3343	2,2662	2,2107
40	4,0847	3,2317	2,8387	2,6060	2,4495	2,3359	2,2490	2,1802	2,1240
50	4,0343	3,1826	2,7900	2,5572	2,4004	2,2864	2,1992	2,1299	2,0734
60	4,0012	3,1504	2,7581	2,5252	2,3683	2,2541	2,1665	2,0970	2,0401
70	3,9778	3,1277	2,7355	2,5027	2,3456	2,2312	2,1435	2,0737	2,0166
80	3,9604	3,1108	2,7188	2,4859	2,3287	2,2142	2,1263	2,0564	1,9991
100	3,9361	3,0873	2,6955	2,4626	2,3053	2,1906	2,1025	2,0323	1,9748
150	3,9042	3,0564	2,6649	2,4320	2,2745	2,1595	2,0711	2,0006	1,9428
200	3,8884	3,0411	2,6498	2,4168	2,2592	2,1441	2,0556	1,9849	1,9269
500	3,8601	3,0138	2,6227	2,3898	2,2320	2,1167	2,0279	1,9569	1,8986
∞	3,8415	2,9957	2,6049	2,3719	2,2141	2,0986	2,0096	1,9384	1,8799

2.P.5 lentelės tęsinys. Fišerio skirstinio su m ir n laisvės laipsnių P-osios kritinės reikšmės

$n \setminus m$	10	12	15	20	24	30	40	60	120	∞
1	241,88	243,91	245,95	248,01	249,05	250,09	251 14	252,20	253,25	254,31
2	19,396	19,413	19,429	19,446	19,454	19,462	19,471	19,479	19,487	19,496
3	8,7855	8,7446	8,7029	8,6602	8,6385	8,6166	8,5944	8,5720	8,5494	8,5265
4	5,9644	5,9117	5,8578	5,8025	5,7744	5,7459	5,7170	5,6877	5,6581	5,6281
5	4,7351	4,6777	4,6188	4,5581	4,5272	4,4957	4,4638	4,4314	4,3985	4,3650
6	4,0600	3,9999	3,9381	3,8742	3,8415	3,8082	3,7743	3,7398	3,7047	3,6689
7	3,6365	3,5747	3,5108	3,4445	3,4105	3,3758	3,3404	3,3043	3,2674	3,2298
8	3,3472	3,2839	3,2184	3,1503	3,1152	3,0794	3,0428	3,0053	2,9669	2,9276
9	3,1373	3,0729	3,0061	2,9365	2,9005	2,8637	2,8259	2,7872	2,7475	2,7067
10	2,9782	2,9130	2,8450	2,7740	2,7372	2,6996	2,6609	2,6211	2,5801	2,5379
11	2,8536	2,7876	2,7186	2,6464	2,6090	2,5705	2,5309	2,4901	2,4480	2,4045
12	2,7534	2,6866	2,6169	2,5436	2,5055	2,4663	2,4259	2,3842	2,3410	2,2962
13	2,6710	2,6037	2,5331	2,4589	2,4202	2,3803	2,3392	2,2966	2,2524	2,2064
14	2,6022	2,5342	2,4630	2,3879	2,3487	2,3082	2,2664	2,2230	2,1778	2,1307
15	2,5437	2,4753	2,4034	2,3275	2,2878	2,2468	2,2043	2,1601	2,1141	2,0658
16	2,4935	2,4247	2,3522	2,2756	2,2354	2,1938	2,1507	2,1058	2,0589	2,0096
17	2,4499	2,3807	2,3077	2,2304	2,1898	2,1477	2,1040	2,0584	2,0107	1,9604
18	2,4117	2,3421	2,2686	2,1906	2,1497	2,1071	2,0629	2,0166	1,9681	1,9168
19	2,3779	2,3080	2,2341	2,1555	2,1141	2,0712	2,0264	1,9795	1,9302	1,8780
20	2,3479	2,2776	2,2033	2,1242	2,0825	2,0391	1,9938	1,9464	1,8963	1,8432
21	2,3210	2,2504	2,1757	2,0960	2,0540	2,0102	1,9645	1,9165	1,8657	1,8117
22	2,2967	2,2258	2,1508	2,0707	2,0283	1,9842	1,9380	1,8894	1,8380	1,7831
23	2,2747	2,2036	2,1282	2,0476	2,0050	1,9605	1,9139	1,8648	1,8128	1,7570
24	2,2547	2,1834	2,1077	2,0267	1,9838	1,9390	1,8920	1,8424	1,7896	1,7330
25	2,2365	2,1649	2,0889	2,0075	1,9643	1,9192	1,8718	1,8217	1,7684	1,7110
26	2,2197	2,1479	2,0716	1,9898	1,9464	1,9010	1,8533	1,8027	1,7488	1,6906
27	2,2043	2,1323	2,0558	1,9736	1,9299	1,8842	1,8361	1,7851	1,7307	1,6717
28	2,1900	2,1179	2,0411	1,9586	1,9147	1,8687	1,8203	1,7689	1,7138	1,6541
29	2,1768	2,1045	2,0275	1,9446	1,9005	1,8543	1,8055	1,7537	1,6981	1,6377
30	2,1646	2,0921	2,0148	1,9317	1,8874	1,8409	1,7918	1,7396	1,6835	1,6223
40	2,0772	2,0035	1,9245	1,8389	1,7929	1,7444	1,6928	1,6373	1,5766	1,5089
50	2,0261	1,9515	1,8714	1,7841	1,7371	1,6872	1,6337	1,5756	1,5115	1,4383
60	1,9926	1,9174	1,8364	1,7480	1,7001	1,6491	1,5943	1,5343	1,4673	1,3893
70	1,9689	1,8932	1,8117	1,7223	1,6738	1,6220	1,5661	1,5046	1,4351	1,3529
80	1,9512	1,8753	1,7932	1,7032	1,6542	1,6017	1,5449	1,4821	1,4107	1,3247
100	1,9267	1,8503	1,7675	1,6764	1,6267	1,5733	1,5151	1,4504	1,3757	1,2832
150	1,8943	1,8172	1,7335	1,6410	1,5902	1,5354	1,4752	1,4074	1,3275	1,2226
200	1,8783	1,8008	1,7166	1,6233	1,5720	1,5164	1,4551	1,3856	1,3024	1,1885
500	1,8496	1,7716	1,6864	1,5916	1,5392	1,4821	1,4186	1,3455	1,2551	1,1132
∞	1,8307	1,7522	1,6664	1,5705	1,5173	1,4591	1,3940	1,3180	1,2214	1,0000

Literatūra

[1] Afifi A. A., Azen S. P. *Statistical Analysis. A Computer Oriented Approach.* Vertimas į rusų kalbą. – Maskva: „Mir“, 1982.

[2] Bagdonavičius V., Kruopis J. J. *Matematinė statistika. Vadovėlis. I dalis. Parametrinė statistika.* Vilnius: VU leidykla, 2015 (www.statistika.mif.vu.lt).

[3] Bagdonavičius V., Kruopis J. J. *Matematinė statistika. Vadovėlis. II dalis. Tiesiniai modeliai.* Vilnius: VU leidykla, 2015 (www.statistika.mif.vu.lt).

[4] Bagdonavičius V., Kruopis J. J. *Matematinė statistika. Vadovėlis. III dalis. Neparametrinė statistika.* Vilnius: VU leidykla, 2015 (www.statistika.mif.vu.lt).

[5] Bagdonavičius V., Kruopis J. J. *Matematinė statistika. Vadovėlis. IV dalis. Daugiamatė statistika.* Vilnius: VU leidykla, 2015 (www.statistika.mif.vu.lt).

[6] Cox D. R., Hinkley D. V. *Problems and Solutions in Theoretical Statistics.* John Wiley Sans, New York, 1978.

[7] Bickel P. J., Doksum K. A. *Mathematical statistics: basic ideas and selected topics.* Prentice and Hall, v. I, II, 2000.

[8] Čekanavičius V., Murauskas G. *Statistika ir jos taikymai.* Vilnius: TEV, I dalis – 2000. II dalis – 2002.

[9] Čibisov D. M., Pagurova V. I. *Matematinės statistikos uždaviniai.* (rusų kalba). Maskva: Maskvos u-to leidykla, 1990.

[10] Härdle W. K., Spokoiny V., Panov V., Wang W. *Basics of Modern Mathematical Statistics. Exercises and Solutions.* Springer - Verlag Berlin Heidelberg, 2014.

[11] Hicks Ch. R., Turner K. V. *Fundamental Concepts in the Design of Experiments.* Oxford University Press, 5th ed., 1999.

[12] Lehmann E. L, Romano J. P. *Testing Statistical Hypotheses.* Springer, 3rd ed., 2008.

[13] Levulienė R. *Statistikos taikymai naudojant SAS.* Vilnius: VU leidykla, 2009.

[14] Mardia K. V. *Statistics of Directional Data.* Academic Press, 1972.

[15] Rao C. R. *Linear Statistical Inference and Its Applications.* Wiley, 2nd ed., 2002.

[16] Scheffe H. *The Analysis of Variance*. Wiley, 1999.

[17] Shao J. *Mathematical Statistics*. Springer, 1999.

[18] Spokoiny V., Dickhaus T. *Basics of Modern Mathematical Statistics*. Springer - Verlag Berlin Heidelberg, 2015.

Vilijandas Bagdonavičius, Julius Jonas Kruopis, Rūta Levulienė
Matematinės statistikos uždavinynas su sprendimais. – Vilnius: Vilniaus universiteto leidykla, 2019. – 288 p.

ISBN 978-609-07-0247-5

Uždavinyne pateikiami teorinio ir praktinio pobūdžio matematinės statistikos uždaviniai su sprendimais. Jį sudaro trys dalys: parametrinė statistika, tiesiniai modeliai, neparametrinė statistika. Knyga skirta universitetų statistikos ir matematikos studijų krypčių studentams, studijuojantiems matematinės statistikos kursą, taip pat ir kitų mokslinės ar praktinės sričių specialistams, taikantiems matematinės statistikos metodus.

Vilijandas Bagdonavičius, Julius Jonas Kruopis, Rūta Levulienė
Matematinės statistikos uždavinynas su sprendimais

Korektūrą skaitė *Dalia Blažinskaitė*
Viršelio dailininkė *Jurga Tėvelienė*
Maketuotoja *Rūta Levulienė*

Vilniaus universiteto leidykla
Saulėtekio al. 9, LT-10222 Vilnius
info@leidykla.vu.lt, www.leidykla.vu.lt

Uždavinyne pateikiami teorinio ir praktinio pobūdžio matematinės statistikos uždaviniai su sprendimais. Jį sudaro trys dalys: parametrinė statistika, tiesiniai modeliai, neparametrinė statistika. Knyga skirta universitetų statistikos ir matematikos studijų krypčių studentams, studijuojantiems matematinės statistikos kursą, taip pat ir kitų mokslinės ar praktinės sričių specialistams, taikantiems matematinės statistikos metodus.

VADOVĖLIAI

9786090702475