VILNIAUS UNIVERSITETO
MATEMATIKOS IR INFORMATIKOS FAKULTETAS

Vilijandas Bagdonavičius

Julius Jonas Kruopis

MATEMATINĖ STATISTIKA

Vadovėlis

IV DALIS

DAUGIAMATĖ STATISTIKA

Vilniaus universiteto leidykla
2015

UDK 519.2(075.8)

Apsvarstė ir rekomendavo spausdinti Vilniaus universiteto Matematikos ir informatikos fakulteto taryba (2015 m. vasario 17 d.; protokolas Nr. 3); vadovėlio statusą suteikė Vilniaus universiteto senatas (2015 m. balandžio 21 d., nutarimas Nr. S – 2015 – 4 –12).

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

ISBN 978–609–459–518–9

Turinys

Pratarmė

Stebimiems objektams aprašyti dažnai nepakanka vieno požymio, o tenka naudoti požymių vektorių. Tada skirtingų objektų požymių vektoriaus matavimus galima traktuoti kaip imčių, gautų stebint tam tikrą daugiamatį atsitiktinį vektorių, realizacijas.

Gana dažnai stebimo atsitiktinio vektoriaus skirstinį patenkinamai galima aprašyti daugiamačiu normaliuoju skirstiniu. Kaip ir vienmačiu atveju normaliojo modelio paplitimas aiškinamas centrine ribine teorema. Daugiamatis normalusis skirstinys aprašo tokio atsitiktinio vektoriaus, kuris gaunamas sumuojant didelį skaičių nepriklausomų ar silpnai priklausomų atsitiktinių vektorių, tarp kurių nėra „dominuojančių“, skirstinį.

Kita vertus, daugiamačio normaliojo skirstinio statistiniai metodai yra labiau išvystyti ir daugeliu atvejų įgiję užbaigtą pavidalą. Gauti daugiamačiai vidurkių ir dispersijų palyginimo, hipotezių tiesiniuose modeliuose tikrinimo ir kt. kriterijų analogai. Tačiau daugiamatėje matematinėje statistikoje nagrinėjami ir specifiniai uždaviniai, kurių nėra vienmatėje teorijoje. Tai uždaviniai, susiję su kovariacinės matricos struktūra, diskriminantinė analizė, stebimo vektoriaus dimensijos sumažinimo problema ir kt.

Knygos paskirtis lėmė medžiagos iš šios plačios matematinės statistikos srities parinkimą. Apsiribojama normaliojo skirstinio teorija, tariant, kad imties elementai turi daugiamatį normalųjį skirstinį. Norint plačiau susipažinti su daugiamatės matematinės statistikos metodais ir rezultatais, rekomenduojame monografijas [2], [9], [10], [13], [14].

Pirmame skyriuje pateikiami normaliojo vektoriaus parametrų įvertiniai ir jų savybės, antrajame – vidurkių vektoriaus reikšmių ir jų palyginimo kriterijai, t. y. Stjudento kriterijaus daugiamačiai analogai. Trečiajame skyriuje aptariama tiesinių modelių analizės metodai daugiamačio normaliojo skirstinio atveju.

Ketvirtas skyrius skiriamas koreliacinei analizei. Sudaromi kriterijai atsitiktinių dydžių ar vektorių nepriklausomumo hipotezėms tikrinti. Penktame skyriuje pateikti kriterijai kovariacinių matricų palyginimo hipotezėms tikrinti. Šeštame skyriuje nagrinėjami klasifikavimo (diskriminantinės analizės) uždaviniai. Septintame skyriuje aptariama stebimų vektorių dimensijos sumažinimo problema apibrėžiant kanoninius kintamuosius (pagrindinės komponentės, kanoninės koreliacijos, faktorinė analizė).

Prieduose pateikiami dažniausiai knygoje naudojami tiesinės algebros faktai, kai kurios atsitiktinių vektorių ir daugiamačio normaliojo skirstinio savybės.

Dėstoma medžiaga iliustruojama konkrečiais pavyzdžiais. Kiekvieno skyriaus pabaigoje pateikiami pratimai savarankiškam darbui. Iliustracijos ir pratimai daugiausia parinkti iš minėtų monografijų.

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);
TPP – taikomieji programų paketai;
$X,\ Y,\ Z, ...$ – atsitiktiniai dydžiai;
$\mathbf{X},\ \boldsymbol{Y},\ \boldsymbol{Z}, ...$ – atsitiktiniai vektoriai;
$\mathbf{X}^T$ – transponuotas vektorius, t. y. vektorius – eilutė;
$x(P)$ – P-asis kvantilis;
x_P – P-oji kritinė reikšmė;
$\mathbf{\Sigma} = [\sigma_{ij}]_{k\times k}$ – kovariacijų matrica;
$\rho = [\rho_{ij}]_{k\times k}$ – 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;
$N(0,\ 1)$ – standartinis normalusis skirstinys;
$N(\mu,\ \sigma^2)$ – normalusis skirstinys su parametrais μ ir σ^2;
$\chi^2(n)$ – 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;

$W_k(n,\ \mathbf{\Sigma})$ – k-matis Višarto skirstinys su n laisvės laipsnių ir parametru (kovariacine matrica) $\mathbf{\Sigma}$;
$W_k(n,\ \mathbf{\Sigma};\ \boldsymbol{M})$ – necentrinis k-matis Višarto skirstinys su n laisvės laipsnių, parametru (kovariacine matrica) $\mathbf{\Sigma}$ ir necentriškumo parametru (matrica) $\boldsymbol{M}$;
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_α(m, n) — Fišerio skirstinio su m ir n laisvės laipsnių α kritinė reikšmė;
$N_k(\boldsymbol{\mu},\ \mathbf{\Sigma})$ – k-matis normalusis skirstinys su vidurkių vektoriumi $\boldsymbol{\mu}$ ir kovariacijų matrica $\mathbf{\Sigma}$;
$X \sim N(\mu,\ \sigma^2)$ – 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;
$X_n \sim Y_n$ – a. d. X_n ir Y_n asimptotiškai $(n \to \infty)$ ekvivalentūs $(X_n - Y_n \xrightarrow{P} 0)$;
$X \overset{d}{\sim} Y$ – a. d. X ir Y tikimybiniai skirstiniai sutampa;
$||\boldsymbol{x}||$ – 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.

1 skyrius

Parametrų įvertiniai ir jų savybės

1.1. Stebėjimo duomenys

Tarkime, $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n$ yra paprastoji atsitiktinė atsitiktinio vektoriaus $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ imtis; čia $\boldsymbol{\mu} = (\mu_1, ..., \mu_k)^T$ vidurkių vektorius, o $\boldsymbol{\Sigma} = [\sigma_{ij}]_{k \times k}$ kovariacijų matrica. Tarsime, kad matrica $\boldsymbol{\Sigma}$ teigiamai apibrėžta $|\boldsymbol{\Sigma}| > 0$. Vektoriaus $\mathbf{X}_i$ koordinates pažymėję $X_{1i}, X_{2i}, ..., X_{ki}$, stebėjimo duomenis surašykime į lentelę

	$\mathbf{X}_1$	$\mathbf{X}_2$	...	$\mathbf{X}_n$	
$\boldsymbol{Y}_1^T$	X_{11}	X_{12}	$\cdots$	X_{1n}	
$\boldsymbol{Y}_2^T$	X_{21}	X_{22}	$\cdots$	X_{2n}	$= \quad \mathcal{X}^T$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
$\boldsymbol{Y}_k^T$	X_{k1}	X_{k2}	$\cdots$	X_{kn}	

Stulpeliuose surašyti vektoriai $\mathbf{X}_1, ..., \mathbf{X}_n$ yra vienodai pasiskirstę nepriklausomi ir normalieji $\mathbf{X}_i \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, todėl a. v. $\boldsymbol{Y}_j = (X_{j1}, X_{j2}, ..., X_{jn})^T$ koordinatės yra paprastoji imtis, gauta stebint vienmatį atsitiktinį dydį $X_j \sim N(\mu_j, \sigma_{jj})$, $j = 1, ..., k$; matricos $\mathcal{X} = [X_{ij}]_{n \times k}$ stulpelius sudaro vektoriai $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_k$.

1.2. Didžiausiojo tikėtinumo įvertiniai

Remdamiesi daugiamačio normaliojo skirstinio tankio funkcijos išraiška (3 priedas, (10.0.6)), gauname paprastosios imties $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n$ tikėtinumo funkciją

$$L = L(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = (2\pi)^{-nk/2} |\boldsymbol{\Sigma}|^{-n/2} \exp\{-\frac{1}{2} \sum_{i=1}^{n} (\mathbf{X}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{X}_i - \boldsymbol{\mu})\}. \quad (1.2.1)$$

1.2.1 teorema. *Jeigu* $|\boldsymbol{\Sigma}| > 0$ *ir* $n > k$, *tai parametrų* $\boldsymbol{\mu}$ *ir* $\boldsymbol{\Sigma}$ *DT įvertiniai yra*

$$\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}}, \quad \hat{\boldsymbol{\Sigma}} = \frac{1}{n}\boldsymbol{S} = [\hat{\sigma}_{ij}]_{k\times k}, \tag{1.2.2}$$

o tikėtinumo funkcijos maksimumas

$$L(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}) = (2\pi)^{-nk/2} n^{nk/2} |\boldsymbol{S}|^{-n/2} \exp\{-nk/2\}. \tag{1.2.3}$$

Čia

$$\bar{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i = (\frac{1}{n}\sum_{i=1}^{n}X_{1i}, ..., \frac{1}{n}\sum_{i=1}^{n}X_{ki})^T = (\bar{X}_1, ..., \bar{X}_k)^T,$$

$$\boldsymbol{S} = \sum_{i=1}^{n}(\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^T = \sum_{i=1}^{n}\mathbf{X}_i\mathbf{X}_i^T - n\bar{\mathbf{X}}\bar{\mathbf{X}}^T = \mathcal{X}^T\mathcal{X} - n\bar{\mathbf{X}}\bar{\mathbf{X}}^T =$$

$$= [S_{ij}]_{k\times k}, \quad S_{ij} = \boldsymbol{Y}_i^T\boldsymbol{Y}_j - n\bar{X}_i\bar{X}_j. \tag{1.2.4}$$

Įrodymas. Remdamiesi matricos pėdsako savybėmis (1 priedas, (8.2.4)), pertvarkykime reiškinį po eksponentės ženklu (1.2.1) lygybėje:

$$\sum_{i=1}^{n}(\mathbf{X}_i - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{X}_i - \boldsymbol{\mu}) = Tr\{\sum_{i=1}^{n}(\mathbf{X}_i - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{X}_i - \boldsymbol{\mu})\} =$$

$$Tr\{\boldsymbol{\Sigma}^{-1}\sum_{i=1}^{n}(\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})^T\}.$$

Kadangi

$$\sum_{i=1}^{n}(\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})^T = \sum_{i=1}^{n}(\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^T + n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})^T$$

$$= \boldsymbol{S} + n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})^T,$$

tai

$$\sum_{i=1}^{n}(\mathbf{X}_i - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{X}_i - \boldsymbol{\mu}) = Tr\{\boldsymbol{\Sigma}^{-1}(\boldsymbol{S} + n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})^T)\}$$

$$= Tr\{\boldsymbol{\Sigma}^{-1}\boldsymbol{S}\} + nTr\{\boldsymbol{\Sigma}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})^T\} = Tr\{\boldsymbol{\Sigma}^{-1}\boldsymbol{S}\} + n(\bar{\mathbf{X}} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}).$$

Įrašę į (1.2.1), gauname

$$L = (2\pi)^{-nk/2}|\boldsymbol{\Sigma}|^{-n/2}\exp\{-\frac{1}{2}Tr\{\boldsymbol{\Sigma}^{-1}\boldsymbol{S}\}\}\exp\{-\frac{n}{2}(\bar{\mathbf{X}} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu})\}. \tag{1.2.5}$$

Paskutinis daugiklis įgyja maksimalią reikšmę, lygią 1, kai vietoje $\boldsymbol{\mu}$ įrašome įvertinį $\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}}$. Pirmasis daugiklis priklauso tik nuo matricos $\boldsymbol{\Sigma}$. Įrodysime, kad

$$|\boldsymbol{\Sigma}|^{-n/2} \exp\{-\frac{1}{2} Tr\{\boldsymbol{\Sigma}^{-1}\boldsymbol{S}\}\} \leq n^{nk/2} |\boldsymbol{S}|^{-n/2} \exp\{-\frac{nk}{2}\}. \tag{1.2.6}$$

Jeigu $|\boldsymbol{\Sigma}| > 0$ ir $n > k$ tai matrica $\boldsymbol{S}$ teigiamai apibrėžta su tikimybe 1 (žr. 1.6 pratimą). Tada egzistuoja tokia simetriška teigiamai apibrėžta matrica $\boldsymbol{S}^{1/2}$, kad $\boldsymbol{S}^{1/2}\boldsymbol{S}^{1/2} = \boldsymbol{S}$, $\boldsymbol{S}^{-1/2}\boldsymbol{S}^{-1/2} = \boldsymbol{S}^{-1}$, $\boldsymbol{S}^{1/2}\boldsymbol{S}^{-1/2} = \boldsymbol{I}$ (1 priedas (8.2.10)). Remiantis 1 priedo (8.2.24) charakteringoji lygtis

$$|\boldsymbol{\Sigma}^{-1} - \lambda \boldsymbol{S}^{-1}| = 0 \quad \Leftrightarrow \quad |\boldsymbol{S}^{1/2}\boldsymbol{\Sigma}^{-1}\boldsymbol{S}^{1/2} - \lambda \boldsymbol{I}| = 0 \tag{1.2.7}$$

turi k teigiamų šaknų $\lambda_1, ..., \lambda_k > 0$, ir

$$|\boldsymbol{S}^{1/2}\boldsymbol{\Sigma}^{-1}\boldsymbol{S}^{1/2}| = |\boldsymbol{\Sigma}^{-1}||\boldsymbol{S}| = \frac{|\boldsymbol{S}|}{|\boldsymbol{\Sigma}|} = \prod_{j=1}^{k} \lambda_j, \tag{1.2.8}$$

$$Tr\{\boldsymbol{\Sigma}^{-1}\boldsymbol{S}\} = Tr\{\boldsymbol{S}^{1/2}\boldsymbol{\Sigma}^{-1}\boldsymbol{S}^{1/2}\} = \sum_{j=1}^{k} \lambda_j. \tag{1.2.9}$$

Įstatę (1.2.8) ir (1.2.9) į (1.2.6), gauname

$$|\boldsymbol{\Sigma}|^{-n/2} \exp\{-\frac{1}{2} Tr\{\boldsymbol{\Sigma}^{-1}\boldsymbol{S}\}\} = |\boldsymbol{S}|^{-n/2} \prod_{j=1}^{k} (\lambda_j^{n/2} \exp\{-\lambda_j/2\}).$$

Kai x neneigiamas, funkcija $f(x) = x^{n/2} \exp\{-x/2\}$ įgyja maksimalią reikšmę taške $x = n$. Iš čia gauname (1.2.6). ▲

Kadangi $(X_{i1}, ..., X_{in})^T$ yra paprastoji atsitiktinė imtis a. d. $X_i \sim N(\mu_i, \sigma_{ii})$, o

$$S_{ii} = \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2,$$

tai remiantis vienmačio normaliojo skirstinio teorija $\mathbf{E}(S_{ii}) = (n-1)\sigma_{ii}$. Taigi DT įvertinys $\hat{\sigma}_{ii}$ yra paslinktasis. Nepaslinktasis dispersijos σ_{ii} įvertinys yra

$$\frac{n}{n-1}\hat{\sigma}_{ii} = \frac{1}{n-1} S_{ii}. \tag{1.2.10}$$

Nagrinėkime sumas $X_{ij} + X_{i'j}$, $j = 1, ..., n$. Tada $(X_{i1} + X_{i'1}, ..., X_{in} + X_{i'n})^T$ yra paprastoji imtis vienmačio normaliojo a. d. $X_i + X_{i'} \sim N(\mu_i + \mu_{i'}, \sigma^2)$, $\sigma^2 = \sigma_{ii} + \sigma_{i'i'} + 2\sigma_{ii'}$. Nepaslinktas dispersijos σ^2 įvertinys yra

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{j=1}^{n} (X_{ij} + X_{i'j} - \bar{X}_i - \bar{X}_{i'})^2 = \frac{1}{n-1} (S_{ii} + S_{i'i'} + 2S_{ii'}).$$

Kadangi $S_{ii}/(n-1)$ ir $S_{i'i'}/(n-1)$ yra nepaslinktieji parametrų σ_{ii} ir $\sigma_{i'i'}$ įvertiniai, tai $S_{ii'}/(n-1)$ yra nepaslinktasis parametro $\sigma_{ii'}$ įvertinys. Taigi kovariacinės matricos $\mathbf{\Sigma}$ nepaslinktasis įvertinys yra

$$\frac{n}{n-1}\hat{\mathbf{\Sigma}} = \frac{1}{n-1}\boldsymbol{S}. \tag{1.2.11}$$

1.2.1 pastaba. Tikėtinumo funkcija (1.2.1) priklauso $k(k+3)/2$ parametrinei eksponetinių skirstinių šeimai, kuriai statistika $(\bar{\mathbf{X}}, \boldsymbol{S})$ yra pilnoji ir pakankamoji (žr. 1 dalies 4.66 pratimą). Todėl statistikos $(\bar{\mathbf{X}}, \boldsymbol{S})$ funkcija yra jos vidurkio NMD įvertinys. Taigi vektoriaus $\bar{\mathbf{X}}$ ir matricos $\boldsymbol{S}/(n-1)$ elementai yra vektoriaus $\boldsymbol{\mu}$ koordinačių ir matricos $\mathbf{\Sigma}$ elementų NMD įvertiniai.

1.2.1 pavyzdys. Lentelėje pateiktos 30 Panevėžio gamykloje „Ekranas“ pagamintų kineskopų trijų spindulių srovės stiprumo X_1, X_2, X_3 reikšmės, pamatuotos technologinės operacijos „II testerių karuselė“ metu.

1.2.1 lentelė. Statistiniai duomenys

i	X_{1i}	X_{2i}	X_{3i}	i	X_{1i}	X_{2i}	X_{3i}	i	X_{1i}	X_{2i}	X_{3i}
1	6,3	6,4	6,3	11	6,6	6,5	6,5	21	6,2	6,2	6,1
2	6,4	6,3	6,3	12	6,4	6,5	6,3	22	6,4	6,3	6,4
3	6,4	6,3	6,4	13	6,4	6,3	6,4	23	6,2	6,1	6,2
4	6,3	6,2	6,2	14	6,2	6,2	6,3	24	6,2	6,2	6,3
5	6,3	6,2	6,3	15	6,3	6,2	6,4	25	6,2	6,1	6,3
6	6,4	6,4	6,3	16	6,2	6,1	6,3	26	6,1	6,3	6,1
7	6,3	6,2	6,3	17	6,3	6,2	6,3	27	6,3	6,3	6,3
8	6,2	6,2	6,4	18	6,2	6,2	6,3	28	6,3	6,2	6,4
9	6,2	6,2	6,3	19	6,1	6,0	6,4	29	6,2	6,2	6,2
10	6,4	6,2	6,3	20	6,2	6,2	6,2	30	6,2	6,1	6,2

Tardami, kad buvo stebėtas trimatis normalusis vektorius

$$\mathbf{X} = (X_1, X_2, X_3)^T \sim N_3(\boldsymbol{\mu}, \mathbf{\Sigma}),$$

rasime parametrų $\boldsymbol{\mu}$ ir $\mathbf{\Sigma}$ ivertinių realizacijas (įverčius).

Randame

$$\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i = \frac{1}{30}\left\{\begin{pmatrix} 6,3 \\ 6,4 \\ 6,3 \end{pmatrix} + ... + \begin{pmatrix} 6,2 \\ 6,1 \\ 6,2 \end{pmatrix}\right\} = \begin{pmatrix} 6,28 \\ 6,24 \\ 6,30 \end{pmatrix}.$$

Apskaičiuojame matricos $\boldsymbol{S}$ realizaciją

$$\boldsymbol{S} = \sum_{i=1}^{n}\mathbf{X}_i\mathbf{X}_i^T - n\bar{\mathbf{X}}\bar{\mathbf{X}}^T = \begin{pmatrix} 6,3 \\ 6,4 \\ 6,3 \end{pmatrix}(6,3; 6,4; 6,3) + ... + \begin{pmatrix} 6,2 \\ 6,1 \\ 6,2 \end{pmatrix}(6,2; 6,1; 6,2) -$$

$$30\begin{pmatrix} 6,28 \\ 6,24 \\ 6,30 \end{pmatrix}(6,28; 6,24; 6,30) = \begin{pmatrix} 0,348 & 0,264 & 0,160 \\ 0,264 & 0,392 & 0,070 \\ 0,160 & 0,070 & 0,240 \end{pmatrix}.$$

Kovariacinės matricos $\mathbf{\Sigma}$ didžiausiojo tikėtinumo įvertis yra $\boldsymbol{S}/30$, o nepaslinktasis įvertis $\boldsymbol{S}/29$.

1.3. Parametrų įvertinių skirstiniai

Stebint vienmatį normalųjį a. d. empirinis vidurkis taip pat turi normalųjį skirstinį ir nepriklauso nuo empirinės dispersijos. Analogiškas rezultatas yra teisingas ir kai stebimas daugiamatis normalusis vektorius.

1.3.1 teorema. *Jeigu paprastoji atsitiktinė imtis* $\mathbf{X}_1, ..., \mathbf{X}_n$ *gauta stebint normalųjį a. v.* $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}), |\boldsymbol{\Sigma}| > 0,\ n > k$, *tai*

$$\bar{\mathbf{X}} = \frac{1}{n}\sum_{j=1}^{n} \mathbf{X}_j \sim N_k(\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{\Sigma}), \tag{1.3.1}$$

$$n(\bar{\mathbf{X}} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}) \sim \chi^2(k). \tag{1.3.2}$$

Empirinis vidurkis $\bar{\mathbf{X}}$ *nepriklauso nuo statistikos* $\boldsymbol{S}$, *kuri yra pasiskirsčiusi kaip matrica*

$$\sum_{i=1}^{n-1} \boldsymbol{Z}_i \boldsymbol{Z}_i^T, \tag{1.3.3}$$

čia $\boldsymbol{Z}_1, ..., \boldsymbol{Z}_{n-1}$ *yra vienodai pasiskirstę nepriklausomi a. v.,* $\boldsymbol{Z}_i \sim N_k(\mathbf{0}, \boldsymbol{\Sigma})$.

Įrodymas. Savybės (1.3.1) ir (1.3.2) tiesiogiai plaukia iš daugiamačio normaliojo skirstinio savybių (3 priedas (10.0.7), (10.0.4)).

Kad įrodytume (1.3.3), atlikime vektorių $\mathbf{X}_1, ..., \mathbf{X}_n$ tiesinę transformaciją naudodami ortogonalią matricą $\boldsymbol{C} = [c_{ij}]_{n\times n}, \boldsymbol{C}\boldsymbol{C}^T = \boldsymbol{I}$:

$$\boldsymbol{Z}_j = c_{j1}\mathbf{X}_1 + ... + c_{jn}\mathbf{X}_n, \quad j = 1, ..., n. \tag{1.3.4}$$

Tegu matricos $\boldsymbol{C}$ paskutinioji eilutė yra $(1/\sqrt{n}, ..., 1/\sqrt{n})$. Tada iš ortogonalumo sąlygos gaunama, kad kitų eilučių elementų sumos yra lygios nuliui:

$$\sum_{i=1}^{n} c_{ji} = 0, \quad j = 1, 2, ..., n-1.$$

Naujai gautų vektorių vidurkiai

$$\mathbf{E}(\boldsymbol{Z}_n) = \sqrt{n}\mathbf{E}(\bar{\mathbf{X}}) = \sqrt{n}\boldsymbol{\mu}, \quad \mathbf{E}(\boldsymbol{Z}_j) = \sum_{i=1}^{n} c_{ji}\boldsymbol{\mu} = \mathbf{0}, \quad j = 1, ..., n-1. \tag{1.3.5}$$

Randame kovariacijas

$$\mathbf{Cov}\,(\boldsymbol{Z}_j, \boldsymbol{Z}_{j'}) = \mathbf{E}\{[\sum_{i=1}^{n} c_{ji}(\mathbf{X}_i - \boldsymbol{\mu})][\sum_{l=1}^{n} c_{j'l}(\mathbf{X}_l - \boldsymbol{\mu})]^T\}$$

$$= \sum_{i,l=1}^{n} c_{ji}c_{j'l}\mathbf{E}[(\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_l - \boldsymbol{\mu})^T] = \sum_{i=1}^{n} c_{ji}c_{j'i}\boldsymbol{\Sigma},$$

nes vektoriai $\mathbf{X}_i$ ir $\mathbf{X}_l$ nekoreliuoti, kai $i \neq l$. Remdamiesi matricos $\boldsymbol{C}$ ortogonalumu, gauname

$$\boldsymbol{V}(\boldsymbol{Z}_j) = \boldsymbol{\Sigma}, \quad \mathbf{Cov}\,(\boldsymbol{Z}_j, \boldsymbol{Z}_{j'}) = \mathbf{0}, \quad j \neq j'. \tag{1.3.6}$$

Naujai gauti vektoriai $\boldsymbol{Z}_1, ..., \boldsymbol{Z}_n$ turi vidurkius (1.3.5), tokias pat kovariacijas $\boldsymbol{\Sigma}$ kaip ir vektoriai $\mathbf{X}_1, ..., \mathbf{X}_n$, yra nepriklausomi ir normalieji.

Kadangi transformacija ortogonalioji, tai

$$\sum_{i=1}^{n} \boldsymbol{Z}_i \boldsymbol{Z}_i^T = \sum_{i=1}^{n} [\sum_{j=1}^{n} c_{ij} \mathbf{X}_i][\sum_{l=1}^{n} c_{il} \mathbf{X}_i]^T$$

$$= \sum_{i=1}^{n} [\sum_{j,l=1}^{n} c_{ij} c_{il}] \mathbf{X}_i \mathbf{X}_i^T = \sum_{i=1}^{n} \mathbf{X}_i \mathbf{X}_i^T.$$

Gauname

$$\boldsymbol{S} = \sum_{i=1}^{n} \mathbf{X}_i \mathbf{X}_i^T - n\bar{\mathbf{X}}\bar{\mathbf{X}}^T = \sum_{i=1}^{n} \boldsymbol{Z}_i \boldsymbol{Z}_i^T - \boldsymbol{Z}_n \boldsymbol{Z}_n^T = \sum_{i=1}^{n-1} \boldsymbol{Z}_i \boldsymbol{Z}_i^T. \tag{1.3.7}$$

Taigi $\boldsymbol{S}$ išraiškoje yra tik vektoriai $\boldsymbol{Z}_1, ..., \boldsymbol{Z}_{n-1}$, kurie nepriklauso nuo $\boldsymbol{Z}_n = \sqrt{n}\bar{\mathbf{X}}$. Todėl $\bar{\mathbf{X}}$ ir $\boldsymbol{S}$ yra nepriklausomi. Gauta išraiška (1.3.7) sutampa su (1.3.3). ▲

Matricos (1.3.3) elementų skirstinys vadinamas *centriniu Višarto skirstiniu su $n-1$ laisvės laipsniu*. Detaliau žr. 1.5 skyrelį.

1.4. Išvados dėl vidurkių vektoriaus, kai kovariacinė matrica žinoma

1.4.1. Vidurkių vektoriaus pasikliovimo sritys

Kai skirstinys vienmatis normalusis $N(\mu, \sigma^2)$, darydami sprendimus apie vidurkį μ naudojome sąryšį $\sqrt{n}(\bar{X} - \mu)/\sigma \sim N(0,\ 1)$. Kai skirstinys daugiamatis normalusis, vietoje šio sąryšio naudojame (1.3.2).

Tegu $\chi^2_\alpha(k)$ yra χ^2 skirstinio su k laisvės laipsnių lygmens α kritinė reikšmė. Apibrėžkime k-matės erdvės poaibį

$$\boldsymbol{C}(\bar{\mathbf{X}}) = \{\boldsymbol{\mu} : n(\bar{\mathbf{X}} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}) \leq \chi^2_\alpha(k)\}.$$

Tada $\boldsymbol{C}(\bar{\mathbf{X}})$ yra vidurkių vektoriaus $\boldsymbol{\mu}$ pasikliovimo sritis, kai pasikliovimo lygmuo $Q = 1 - \alpha$:

$$\mathbf{P}\{\boldsymbol{\mu} \in \boldsymbol{C}(\bar{\mathbf{X}}) | \boldsymbol{\mu}\} = Q = 1 - \alpha. \tag{1.4.1}$$

Matome, kad ši sritis yra elipsoidas pavidalo turintis centrą $\bar{\mathbf{X}}$.

1.4.2. Vidurkių vektoriaus koordinačių pasikliovimo intervalų rinkiniai

Vietoje pasikliovimo srities (1.4.1) kartais pageidautina turėti pasikliovimo intervalų rinkinį, kuris uždengtų visus dominančius parametrus $\mu_1, ..., \mu_k$ su tikimybe, ne mažesne už Q.

Naudojantis pasikliovimo sritimi (1.4.1) galima sudaryti iš karto visų tiesinių funkcijų $\boldsymbol{c}^T\boldsymbol{\mu}$, $\boldsymbol{c} \in \mathbf{R}^k$ pasikliovimo intervalus.

Pažymėkime $\mathcal{L}$ aibę, kuri gaunama imant tiesines vektoriaus $\boldsymbol{\mu}$ funkcijas: $\mathcal{L} = \{\boldsymbol{c}^T\boldsymbol{\mu} : \boldsymbol{c} \in \mathbf{R}^k\}$.

1.4.1 teorema. *Tarkime, kad* $|\boldsymbol{\Sigma}| = |[\sigma_{ij}]_{k\times k}| > 0$. *Tada su tikimybe* $Q = 1-\alpha$ *iš karto visoms funkcijoms* $\boldsymbol{c}^T\boldsymbol{\mu} \in \mathcal{L}$ *galioja nelygybės*

$$\boldsymbol{c}^T\bar{\mathbf{X}} - \sqrt{\boldsymbol{c}^T\boldsymbol{\Sigma}\boldsymbol{c}\chi^2_\alpha(k)/n} \leq \boldsymbol{c}^T\boldsymbol{\mu} \leq \boldsymbol{c}^T\bar{\mathbf{X}} + \sqrt{\boldsymbol{c}^T\boldsymbol{\Sigma}\boldsymbol{c}\chi^2_\alpha(k)/n}. \qquad (1.4.2)$$

Nelygybes (1.4.2) galima traktuoti kaip pasikliovimo intervalus, sudarytus iš karto visoms tiesinėms funkcijoms $\boldsymbol{c}^T\boldsymbol{\mu} \in \mathcal{L}$. *Jeigu imsime vieną atskirą funkciją* $\boldsymbol{c}^T\boldsymbol{\mu}$ *(arba keletą tokio tipo funkcijų), tai intervalų pasikliovimo lygmuo ne mažesnis už* Q.

Įrodymas. Pagal Koši ir Švarco nelygybę (1 priedas (8.1.9.) bet kuriems vienodos dimensijos vektoriams $\boldsymbol{U}$ ir $\boldsymbol{V}$ galioja sąryšiai

$$(\boldsymbol{U}^T\boldsymbol{V})^2 \leq (\boldsymbol{U}^T\boldsymbol{U})(\boldsymbol{V}^T\boldsymbol{V}), \quad \boldsymbol{V}^T\boldsymbol{V} = \sup_{\boldsymbol{U}} \frac{(\boldsymbol{U}^T\boldsymbol{V})^2}{\boldsymbol{U}^T\boldsymbol{U}}.$$

Kadangi $\boldsymbol{\Sigma}$ teigiamai apibrėžta simetriška matrica, tai pagal (1 priedas (8.2.10)) egzistuoja teigiamai apibrėžta kvadratinė matrica $\boldsymbol{B}$, kad $\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}^T$. Pritaikę Koši–Švarco nelygybę vektoriams $\boldsymbol{B}\boldsymbol{U}$ ir $(\boldsymbol{B}^{-1})^T\boldsymbol{V}$ gausime

$$\boldsymbol{V}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{V} \geq \frac{(\boldsymbol{U}^T\boldsymbol{V})^2}{\boldsymbol{U}^T\boldsymbol{\Sigma}\boldsymbol{U}}, \quad \boldsymbol{V}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{V} = \sup_{\boldsymbol{U}} \frac{(\boldsymbol{U}^T\boldsymbol{V})^2}{\boldsymbol{U}^T\boldsymbol{\Sigma}\boldsymbol{U}}.$$

Supremumas pasiekiamas imant $\boldsymbol{U} = \boldsymbol{\Sigma}^{-1}\boldsymbol{V}$.

Imdami $\boldsymbol{V} = \hat{\boldsymbol{\mu}} - \boldsymbol{\mu}$ ir $\boldsymbol{U} = \boldsymbol{c}$ iš (1.4.1) gauname

$$\mathbf{P}\{(\bar{\mathbf{X}} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \leq \chi^2_\alpha(k)/n|\boldsymbol{\mu}\}$$

$$= \mathbf{P}\{\sup_{\boldsymbol{c}} \frac{|\boldsymbol{c}^T(\bar{\mathbf{X}} - \boldsymbol{\mu})|}{\sqrt{\boldsymbol{c}^T\boldsymbol{\Sigma}\boldsymbol{c}}} \leq \sqrt{\chi^2_\alpha(k)/n}|\boldsymbol{\mu}\}$$

$$= \mathbf{P}\{\boldsymbol{c}^T\bar{\mathbf{X}} - \sqrt{\boldsymbol{c}^T\boldsymbol{\Sigma}\boldsymbol{c}\chi^2_\alpha(k)/n} \leq \boldsymbol{c}^T\boldsymbol{\mu} \leq \boldsymbol{c}^T\bar{\mathbf{X}} + \sqrt{\boldsymbol{c}^T\boldsymbol{\Sigma}\boldsymbol{c}\chi^2_\alpha(k)/n},\ \forall\, \boldsymbol{c} \in \mathbf{R}^k|\boldsymbol{\mu}\} = Q.$$

Aišku, kad atskirai funkcijai $\boldsymbol{c}^T\boldsymbol{\mu}$ (arba keletui tokių funkcijų) pasikliovimo lygmuo yra ne mažesnis už Q.

Imdami paeiliui $\boldsymbol{c}_1 = (1,\ 0, ..., 0)^T$, $\boldsymbol{c}_2 = (0,\ 1, ..., 0)^T, ..., \boldsymbol{c}_k = (0,\ 0, ..., 1)^T$ pagal (1.4.2) gauname parametrų $\mu_1, ..., \mu_m$ pasikliovimo intervalų sistemą $(\underline{\mu}'_i, \overline{\mu}'_i)$:

$$\underline{\mu}'_i = \bar{X}_i - \sqrt{\sigma_{ii}\chi^2_\alpha(k)/n}, \quad \overline{\mu}'_i = \bar{X}_i + \sqrt{\sigma_{ii}\chi^2_\alpha(k)/n},$$

kuriai

$$\mathbf{P}\{\underline{\mu}'_i < \mu_i < \overline{\mu}'_i,\ \forall\ i = 1, ..., k\} \geq Q = 1 - \alpha, \qquad (1.4.3)$$

Tikimybė, kad visi intervalai (1.4.3) uždengs tikrąsias parametrų μ_i reikšmes, gali būti kur kas didesnė už $Q = 1-\alpha$, nes vietoje visų vektorių $\boldsymbol{c} \in \mathbf{R}^k$ imame tik k vektorių $\boldsymbol{c}_1, ..., \boldsymbol{c}_k$. ▲

Jeigu sukonstruotume pasikliovimo intervalus kiekvienai koordinatei μ_i remdamiesi vienmate normaliojo skirstinio teorija:

$$\underline{\mu}_i = \bar{X}_i - z_{\alpha/2}\sqrt{\sigma_{ii}/n}, \quad \bar{\mu}_i = \bar{X}_i + z_{\alpha/2}\sqrt{\sigma_{ii}/n}, \quad i = 1, ..., k, \tag{1.4.4}$$

tai toks intervalų rinkinys nebus ieškomasis, nes tikimybė, kad visi intervalai (1.4.4) uždengs tikrąsias visų parametrų reikšmes, gali būti gerokai mažesnė už $Q = 1-\alpha$. Pavyzdžiui, jeigu įvertiniai $\bar{X}_i, i = 1, ..., k$, yra nepriklausomi, tai intervalai (1.4.4) uždengia visas parametrų reikšmes su tikimybe Q^k. Taigi intervalai (1.4.4) yra trumpesni negu reikėtų. Tikimybė, kad visi intervalai (1.4.3) uždengs tikrąsias parametrų reikšmes, gali būti daug didesnė už $Q = 1-\alpha$, t. y. intervalai (1.4.3) yra ilgesni negu reikėtų.

Kitokį negu (1.4.3) intervalų rinkinio variantą galima gauti naudojant Bonferonio nelygybę. Tegu A_i yra įvykis, kuris reiškia, kad i-asis tipo (1.4.4) intervalas uždengia parametrą μ_i ir tegu $\mathbf{P}\{A_i\} = 1-\alpha_i$. Tada

$$\mathbf{P}\{\cap_{i=1}^k A_i\} = 1 - \mathbf{P}\{\cup_{i=1}^k \bar{A}_i\} \geq$$

$$1 - \sum_{i=1}^k \mathbf{P}\{\bar{A}_i\} = 1 - (\alpha_1 + ... + \alpha_k).$$

Jeigu parinksime $\alpha_i = \alpha/k,\ i = 1, ..., k$, tai intervalų rinkinys

$$\underline{\mu}_i'' = \bar{X}_i - z_{\alpha/(2k)}\sqrt{\sigma_{ii}/n}, \quad \bar{\mu}_i'' = \bar{X}_i + z_{\alpha/(2k)}\sqrt{\sigma_{ii}/n}, \quad i = 1, ..., k, \tag{1.4.5}$$

uždengs visus parametrus $\mu_1, ..., \mu_k$ su tikimybe, ne mažesne už $Q = 1-\alpha$.

1.4.1 pavyzdys. (**1.2.1** pavyzdžio tęsinys). Dėl iliustracijos laikinai tarkime, kad **1.2.1** pavyzdyje kovariacinė matrica $\mathbf{\Sigma}$ yra žinoma ir sutampa su gautu nepaslinktuoju įverčiu $\boldsymbol{S}/29$. Kai ši prielaida teisinga, sudarysime vidurkių vektoriaus $\boldsymbol{\mu}$ pasikliovimo sritį ir vidurkių vektoriaus koordinačių pasikliovimo intervalų rinkinius, kai pasikliovimo lygmuo $Q = 0,95$.

Pasikliovimo sritis (1.4.1) turi tokį pavidalą:

$$\boldsymbol{C}(\bar{\mathbf{X}}) = \{\boldsymbol{\mu} : 870(\bar{\mathbf{X}} - \boldsymbol{\mu})^T \boldsymbol{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \leq \chi^2_{0,05}(3)\}.$$

Apskaičiavę

$$\boldsymbol{S}^{-1} = \begin{pmatrix} 10,587 & -7,155 & -4,971 \\ -7,155 & 7,945 & 2,453 \\ -4,971 & 2,453 & 6,765 \end{pmatrix}$$

ir pažymėję $\boldsymbol{Z} = (Z_1; Z_2; Z_3)^T = \boldsymbol{\mu} - \bar{\mathbf{X}} = (\mu_1 - 6,28; \mu_2 - 6,24; \mu_3 - 6.30)^T$, gausime tokio pavidalo pasikliovimo elipsoidą $\boldsymbol{C}(\bar{\mathbf{X}})$:

$$10,587Z_1^2 + 7,945Z_2^2 + 6,765Z_3^2 - 14,310Z_1Z_2 - 9,942Z_1Z_3 + 4,906Z_2Z_3 \leq 0,009.$$

Antrame ir trečiame 1.4.1 lentelės stulpeliuose yra pateikti pasikliovimo intervalų rinkiniai (1.4.3) ir (1.4.5). Palyginti ketvirtame stulpelyje pateikiami pasikliovimo intervalai (1.4.4), kiekvienai vidurkio koordinatei.

1.4.1 lentelė. Pasikliovimo intervalai

i	$(\underline{\mu}_i'; \overline{\mu}_i')$	$(\underline{\mu}_i''; \overline{\mu}_i'')$	$(\underline{\mu}_i; \overline{\mu}_i)$
1	(6,224; 6,336)	(6,232; 6,328)	(6,241; 6,319)
2	(6,181; 6,299)	(6,189; 6,291)	(6,198; 6,282)
3	(6,254; 6,346)	(6,260; 6,340)	(6,267; 6,333)

Matome, kad pasikliovimo intervalai sudaryti remiantis Bonferonio nelygybe, yra apytiksliai 1,16 karto trumpesni už tuos, kurie sudaryti remiantis pasikliovimo sritimi (1.4.1).

1.4.3. Hipotezių dėl vidurkių vektoriaus reikšmės tikrinimas

Tarkime, reikia patikrinti hipotezę $H : \boldsymbol{\mu} = \boldsymbol{\mu}_0$, kai alternatyva $\bar{H} : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$; čia $\boldsymbol{\mu}_0$ žinomas vektorius. Įrašykime į sąryšio (1.3.2) dešiniąją pusę hipotetinę reikšmę $\boldsymbol{\mu}_0$. Gautoji statistika

$$U^2 = n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T \boldsymbol{\Sigma}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0) \tag{1.4.6}$$

turi centrinį χ^2 skirstinį su k laisvės laipsnių, jeigu tikrinama hipotezė H yra teisinga. Kai hipotezė H neteisinga, statistikos U^2 skirstinys yra necentrinis χ^2 skirstinys, turintis k laisvės laipsnių ir necentriškumo parametrą (3 priedas, (10.0.8))

$$\lambda = n(\boldsymbol{\mu} - \boldsymbol{\mu}_0)^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \boldsymbol{\mu}_0).$$

Taigi hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai teisinga nelygybė

$$U^2 = n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T \boldsymbol{\Sigma}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0) > \chi^2_\alpha(k). \tag{1.4.7}$$

Kriterijaus galia išreiškiama necentrinio χ^2 skirstinio pasiskirstymo funkcija

$$\beta(\boldsymbol{\mu}) = \mathbf{P}\{n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T \boldsymbol{\Sigma}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0) > \chi^2_\alpha(k)|\boldsymbol{\mu}\} =$$
$$= \mathbf{P}\{\chi^2_{k;\lambda} > \chi^2_\alpha(k)\}. \tag{1.4.8}$$

1.4.2 pavyzdys. Dėl iliustracijos kaip ir **1.4.1** pavyzdyje tarkime, kad **1.2.1** pavyzdyje kovariacinė matrica $\boldsymbol{\Sigma}$ yra žinoma ir sutampa su gautu nepaslinktuoju įverčiu $\boldsymbol{S}/29$. Kai ši prielaida teisinga tikrinsime hipotezę, kad vidurkių vektorius $\boldsymbol{\mu}$ lygus fiksuotam vektoriui $\boldsymbol{\mu}_0 = (6,25; 6,25; 6,25)^T$.

Randame statistikos (1.4.6) reikšmę

$$U^2 = 870(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T \boldsymbol{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0) = 12,3212.$$

Kadangi P reikšmė $pv = \mathbf{P}\{\chi^2_3 > 12,3212\} = 0,0064$, tai hipotezė atmetama kriterijumi, kurio reikšmingumo lygmuo viršija 0,0064.

Šis skyrelis yra labiau iliustracinis, nes praktiškai, retai susiduriame su situacijomis, kai kovariacinė matrica būna žinoma.

Kai dispersija nežinoma, vienmačiu atveju naudojame sąryšį $\sqrt{n}(\bar{X} - \mu)/s \sim S(n-1)$, t. y. nežinomą dispersiją pakeičiame jos įvertiniu. Analogiškai daroma ir daugiamačiu atveju, pakeičiant sąryšyje (1.3.2) nežinomą kovariacinę matricą $\boldsymbol{\Sigma}$ jos nepaslinktuoju įvertiniu. Norint sudaryti kriterijus ar pasikliovimo sritis, reikia ištirti gauto a. d. skirstinio savybes.

1.5. Višarto skirstinio apibrėžimas

Višarto skirstinį apibrėšime šiek tiek bendresniu atveju. Tegu $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n$ yra nepriklausomi normalieji vektoriai $\mathbf{X}_i \sim N_k(\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$ su vienodomis kovariacinėmis matricomis $\boldsymbol{\Sigma}$ ir galbūt skirtingais vidurkių vektoriais $\boldsymbol{\mu}_1, ..., \boldsymbol{\mu}_n$.

Vektoriaus $\mathbf{X}_i$ koordinates pažymėję, $X_{1i}, X_{2i}, ..., X_{ki}$ surašykime jas į matricą

	$\mathbf{X}_1$	$\mathbf{X}_2$	...	$\mathbf{X}_n$		
$\boldsymbol{Y}_1^T$ $\boldsymbol{Y}_1^T$	X_{11}	X_{12}	$\cdots$	X_{1n}		
$\boldsymbol{Y}_2^T$	X_{21}	X_{22}	$\cdots$	X_{2n}	$=$	$\mathcal{X}^T$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		
$\boldsymbol{Y}_k^T$	X_{k1}	X_{k2}	$\cdots$	X_{kn}		
$\boldsymbol{Y}^T$	$\boldsymbol{L}^T\mathbf{X}_1$	$\boldsymbol{L}^T\mathbf{X}_2$	$\cdots$	$\boldsymbol{L}^T\mathbf{X}_n$	$=$	$\boldsymbol{L}^T\mathcal{X}^T$

kuri papildyta eilute $\boldsymbol{Y}^T = (\boldsymbol{L}^T\mathbf{X}_1, ..., \boldsymbol{L}^T\mathbf{X}_n)$. Vektoriaus $\boldsymbol{Y}$ koordinatės yra tiesinės vektorių $\mathbf{X}_1, ..., \mathbf{X}_n$ funkcijos; čia $\boldsymbol{L} = (L_1, ..., L_k)^T \in \mathbf{R}^k$ yra fiksuotas vektorius. Kadangi $\mathbf{X}_1, ..., \mathbf{X}_n$ nepriklausomi, tai nepriklausomos ir vektoriaus $\boldsymbol{Y}$ koordinatės, jos turi normaliuosius skirstinius su parametrais (2 priedas, (9.3.6))

$$\boldsymbol{L}^T\mathbf{X}_i \sim N(\boldsymbol{L}^T\boldsymbol{\mu}_i, \sigma^2_{\boldsymbol{L}}), \quad \sigma^2_{\boldsymbol{L}} = \boldsymbol{V}(\boldsymbol{L}^T\mathbf{X}_i) = \boldsymbol{L}^T\boldsymbol{\Sigma}\boldsymbol{L}, \quad i = 1, ..., n. \tag{1.5.1}$$

Vektorių $\boldsymbol{Y}$ galima užrašyti ir taip

$$\boldsymbol{Y} = L_1\boldsymbol{Y}_1 + ... + L_k\boldsymbol{Y}_k = \mathcal{X}\boldsymbol{L}. \tag{1.5.2}$$

1.5.1 apibrėžimas. Matricos

$$\boldsymbol{S} = \sum_{i=1}^{n} \mathbf{X}_i\mathbf{X}_i^T = \mathcal{X}^T\mathcal{X} = [S_{ij}]_{k\times k} = [\boldsymbol{Y}_i^T\boldsymbol{Y}_j]_{k\times k} \tag{1.5.3}$$

elementų bendras daugiamatis skirstinys vadinamas Višarto skirstiniu su n laisvės laipsnių. Žymėsime $\boldsymbol{S} \sim W_k(n, \boldsymbol{\Sigma}; \boldsymbol{M})$. Kadangi matrica $\boldsymbol{S}$ simetrinė, t. y. $S_{ij} = S_{ji}$, tai skirstinio dimensija yra $k(k+1)/2$. Skirstinys priklauso nuo kovariacinės matricos $\boldsymbol{\Sigma}$ ir matricos $\boldsymbol{M} = [\mu_{ji}]_{n\times k}$, kurios eilutėse surašyti a. v. $\mathbf{X}_j$ vidurkiai $\boldsymbol{\mu}_j = (\mu_{1j}, ..., \mu_{kj})^T$. Jeigu $\boldsymbol{M} = \mathbf{0}$, tai skirstinys vadinamas centriniu Višarto skirstiniu su n laisvės laipsnių. Sutrumpintai žymėsime $\boldsymbol{S} \sim W_k(n, \boldsymbol{\Sigma})$.

Jeigu $k = 1$, tai matrica $\boldsymbol{S}$ susideda iš vieno elemento

$$S_{11} = \boldsymbol{Y}_1^T\boldsymbol{Y}_1 = \sum_{j=1}^{n} X_{1j}^2, \tag{1.5.4}$$

kuris yra kvadratų suma nepriklausomų normaliųjų a. d. su vienodomis dispersijomis σ_{11}. Tada $S_{11}/\sigma_{11} \sim \chi^2(k; \lambda)$ turi necentrinį χ^2 skirstinį su k

laisvės laipsnių ir necentriškumo parametru $\lambda = \mu_{11}^2 + ... + \mu_{1n}^2$. Jeigu vidurkiai $\mathbf{E}(X_{1j}) = \mu_{1j} = 0,\ j = 1, ..., n$, tai necentriškumo parametras $\lambda = 0$ ir S_{11}/σ_{11} skirstinys yra $\chi^2(k)$, t. y. centrinis χ^2 skirstinys su k laisvės laipsnių. Taigi Višarto skirstinys yra χ^2 skirstinio apibendrinimas į daugiamatį atvejį.

Vienmačiu atveju greta kvadratų sumos (1.5.4) buvo nagrinėjamos ir kvadratinės formos $\boldsymbol{Y}_1^T \boldsymbol{A} \boldsymbol{Y}_1$, kurios turi χ^2 skirstinį, kai matrica $\boldsymbol{A}$ yra idempotentinė (žr. 3 priedą, 8 savybė). Daugiamačiu atveju kvadratų sumos $\boldsymbol{Y}_1^T \boldsymbol{Y}_1$ analogas yra $\mathcal{X}^T\mathcal{X}$, o kvadratinės formos analogas – matrica $\mathcal{X}^T\boldsymbol{A}\mathcal{X}$, kuri tam tikromis sąlygomis taip pat turi Višarto skirstinį.

1.5.1 teorema. *Tegu* $\mathbf{X}_i \sim N_k(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}),\ i = 1, ..., n$, *yra nepriklausomi a. v., o* $\mathcal{X}, \boldsymbol{Y}_i, \boldsymbol{Y}$ *apibrėžti šiame skyrelyje. Būtina ir pakankama sąlyga, kad* $\mathcal{X}^T\boldsymbol{A}\mathcal{X}$ *turėtų Višarto skirstinį su* r *laisvės laipsnių, yra ta, kad a. d.* $\boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{Y}/\sigma_{\boldsymbol{L}}^2$ *turėtų* χ^2 *skirstinį su* r *laisvės laipsnių su bet kuriuo fiksuotu* $\boldsymbol{L} \in \mathbf{R}^k$*; čia* $r = Rang(\boldsymbol{A}) = Tr(\boldsymbol{A})$. *Be to,* $\mathcal{X}^T\boldsymbol{A}\mathcal{X}$ *skirstinys centrinis tada ir tik tada, kai* $\boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{Y}/\sigma_{\boldsymbol{L}}^2$ *yra centrinis* χ^2 *skirstinys.*

Įrodymas. *Būtinumas.* Tegu $\boldsymbol{S} \sim W_k(r, \boldsymbol{\Sigma}, \boldsymbol{M})$, o $\boldsymbol{L} \in \mathbf{R}^k$ – fiksuotas vektorius. Pagal Višarto skirstinio apibrėžimą $\boldsymbol{S} = \sum_{i=1}^r \boldsymbol{Z}_i\boldsymbol{Z}_i^T$; čia $\boldsymbol{Z}_1, ..., \boldsymbol{Z}_r$ yra n. a. v. ir $\boldsymbol{Z}_i \sim N_k(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}), i = 1, ..., r$.

Padauginę iš $\boldsymbol{L}^T$ ir $\boldsymbol{L}$ gauname

$$\boldsymbol{L}^T\boldsymbol{S}\boldsymbol{L} = \sum_{i=1}^r (\boldsymbol{L}^T\boldsymbol{Z}_i)(\boldsymbol{Z}_i^T\boldsymbol{L}) = \sum_{i=1}^r (\boldsymbol{L}^T\boldsymbol{Z}_i)^2$$

sumą kvadratų normaliųjų a. d. $\boldsymbol{L}^T\boldsymbol{Z}_i \sim N(\boldsymbol{L}^T\boldsymbol{\mu}_i, \sigma_{\boldsymbol{L}}^2)$, su vienodomis dispersijomis $\sigma_{\boldsymbol{L}}^2 = \boldsymbol{L}^T\boldsymbol{\Sigma}\boldsymbol{L}$. Taigi

$$\boldsymbol{L}^T\boldsymbol{S}\boldsymbol{L}/\sigma_{\boldsymbol{L}}^2 \sim \chi^2(r; \lambda), \quad \lambda = \sum_{i=1}^r (\boldsymbol{L}^T\boldsymbol{\mu}_i)^2/\sigma_{\boldsymbol{L}}^2.$$

Jeigu Višarto skirstinys centrinis, tai $\boldsymbol{\mu}_i = \mathbf{0},\ i = 1, ..., r$, ir necentriškumo parametras $\lambda = 0$. Tada $\boldsymbol{L}^T\boldsymbol{S}\boldsymbol{L}/\sigma_{\boldsymbol{L}}^2 \sim \chi^2(r)$ ir χ^2 skirstinys yra centrinis.

Pakankamumas. Tegu $\boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{Y} \sim \sigma_{\boldsymbol{L}}^2\chi_{r;\lambda}^2$. Tada iš kvadratinių formų savybių (žr. 3 priedą, 9 savybė) plaukia, kad $\boldsymbol{A}$ yra rango r idempotentinė matrica. Taigi egzistuoja r ortonormuotų vektorių $\boldsymbol{B}_1, ..., \boldsymbol{B}_r$, $\boldsymbol{B}_i^T\boldsymbol{B}_i = 1, \boldsymbol{B}_i^T\boldsymbol{B}_j = 0, i \neq j$, kad galioja spektrinis skaidinys (1 priedas, (8.2.14))

$$\boldsymbol{A} = \boldsymbol{B}_1\boldsymbol{B}_1^T + ... + \boldsymbol{B}_r\boldsymbol{B}_r^T. \tag{1.5.5}$$

Tada

$$\mathcal{X}^T\boldsymbol{A}\mathcal{X} = \mathcal{X}^T\boldsymbol{B}_1\boldsymbol{B}_1^T\mathcal{X} + ... + \mathcal{X}^T\boldsymbol{B}_r\boldsymbol{B}_r^T\mathcal{X} = \boldsymbol{U}_1\boldsymbol{U}_1^T + ... + \boldsymbol{U}_r\boldsymbol{U}_r^T;$$

čia $\boldsymbol{U}_i = \mathcal{X}^T\boldsymbol{B}_i, i = 1, ..., r$. Vektoriai $\boldsymbol{B}_i$ ortonormuoti, todėl vektoriai $\boldsymbol{U}_1, ..., \boldsymbol{U}_r$ nepriklausomi normalieji su kovariacine matrica $\boldsymbol{\Sigma}$ (žr. 1.3.1 teoremos įrodymą). Pagal Višarto skirstinio apibrėžimą $\mathcal{X}^T\boldsymbol{A}\mathcal{X} \sim W_k(r, \boldsymbol{\Sigma}, \boldsymbol{M})$.

Kadangi $\boldsymbol{Y} \sim N_r(\boldsymbol{ML}, \sigma_{\boldsymbol{L}}^2 \boldsymbol{I})$, tai necentriškumo parametras

$$\lambda = \mathbf{E}(\boldsymbol{Y}^T)\boldsymbol{A}\mathbf{E}(\boldsymbol{Y})/\sigma_{\boldsymbol{L}}^2 = \boldsymbol{L}^T\boldsymbol{M}^T\boldsymbol{AML}/(\boldsymbol{L}^T\boldsymbol{\Sigma L}).$$

Jeigu necentriškumo parametras $\lambda = 0$ bet kuriam vektoriui $\boldsymbol{L} \in \mathbf{R}^k$, tai $\boldsymbol{M}^T\boldsymbol{AM} = 0$. Tačiau iš (1.5.5) turime

$$\boldsymbol{M}^T\boldsymbol{AM} = \sum_{i=1}^{r}(\boldsymbol{M}^T\boldsymbol{B}_i)(\boldsymbol{M}^T\boldsymbol{B}_i) = 0 \Rightarrow \boldsymbol{M}^T\boldsymbol{B}_i = 0, \quad i = 1, ..., r.$$

Taigi $\mathbf{E}(\boldsymbol{U}_i) = \boldsymbol{M}^T\boldsymbol{B}_i = \mathbf{0}$ ir $\mathcal{X}^T\boldsymbol{A}\mathcal{X} \sim W_k(r, \boldsymbol{\Sigma})$ yra centrinis Višarto skirstinys. ▲

Naudojantis pateiktu Višarto ir χ^2 skirstinio sąryšiu galima gauti daug svarbių Višarto skirstinio savybių remiantis χ^2 skirstinio savybėmis. Analogiškas principas buvo naudojamas tiriant daugiamačio normaliojo skirstinio savybes.

1.6. Višarto skirstinio savybės

1 savybė. Matricos $\mathcal{X}^T\boldsymbol{A}_1\mathcal{X} \sim W_k(r, \boldsymbol{\Sigma})$ ir $\mathcal{X}^T\boldsymbol{A}_2\mathcal{X} \sim W_k(r, \boldsymbol{\Sigma})$ yra nepriklausomos tada ir tik tada, kai $\boldsymbol{Y}^T\boldsymbol{A}_1\boldsymbol{Y}/\sigma_{\boldsymbol{L}}^2$ ir $\boldsymbol{Y}^T\boldsymbol{A}_2\boldsymbol{Y}/\sigma_{\boldsymbol{L}}^2$ turi nepriklausomus χ^2 skirstinius su bet kuriuo fiksuotu $\boldsymbol{L} \in \mathbf{R}^k$. Vektorius $\mathcal{X}^T\boldsymbol{B}$ ir matrica $\mathcal{X}^T\boldsymbol{A}\mathcal{X}$ yra nepriklausomi ir turi k-matį normalųjį ir Višarto skirstinius, jeigu $\boldsymbol{Y}^T\boldsymbol{B}$ ir $\boldsymbol{Y}^T\boldsymbol{AY}/\sigma_{\boldsymbol{L}}^2$ yra nepriklausomi ir turi vienmatį normalųjį ir χ^2 skirstinius.

Įrodymas. Analogiškas 1.5.1 teoremai. ▲

2 savybė. Tegu $\mathbf{X}_1, ..., \mathbf{X}_n$ nepriklausomi vienodai pasiskirstę pagal $N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ atsitiktiniai vektoriai. Apibrėžkime

$$\bar{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i, \quad \boldsymbol{S} = \sum_{i=1}^{n}\mathbf{X}_i\mathbf{X}_i^T - n\bar{\mathbf{X}}\bar{\mathbf{X}}^T.$$

Tada $\bar{\mathbf{X}}$ ir $\boldsymbol{S}$ yra nepriklausomi. Be to

$$\bar{\mathbf{X}} \sim N(\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{\Sigma}), \quad \boldsymbol{S} \sim W_k(n-1, \boldsymbol{\Sigma}). \tag{1.6.1}$$

Įrodymas. Fiksuotam vektoriui $\boldsymbol{L} \in \mathbf{R}^k$ nagrinėkime nepriklausomus vienodai pasiskirsčiusius a. d. $\boldsymbol{L}^T\mathbf{X}_1, ..., \boldsymbol{L}^T\mathbf{X}_n$, $\boldsymbol{L}^T\mathbf{X}_i \sim N(\boldsymbol{L}^T\boldsymbol{\mu}, \sigma_{\boldsymbol{L}}^2)$, $\sigma_{\boldsymbol{L}}^2 = \boldsymbol{L}^T\boldsymbol{\Sigma L}$. Iš vienmatės teorijos gauname, kad aritmetinis vidurkis

$$\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{L}^T\mathbf{X}_i = \boldsymbol{L}^T\bar{\mathbf{X}} \sim N(\boldsymbol{L}^T\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{L}^T\boldsymbol{\Sigma L}) \tag{1.6.2}$$

nepriklauso nuo nuokrypių kvadratų sumos

$$\sum_{i=1}^{n}(\boldsymbol{L}^T\mathbf{X}_i)^2 - n(\boldsymbol{L}^T\bar{\mathbf{X}})^2 = \boldsymbol{L}^T(\sum_{i=1}^{n}\mathbf{X}_i\mathbf{X}_i^T - n\bar{\mathbf{X}}\bar{\mathbf{X}}^T)\boldsymbol{L} = \boldsymbol{L}^T\boldsymbol{SL} \sim \sigma_{\boldsymbol{L}}^2\chi_{n-1}^2. \tag{1.6.3}$$

Iš daugiamačio normaliojo skirstinio apibrėžimo (3 priedas, 3P.1 apibrėžimas) išplaukia, kad $\bar{\mathbf{X}} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}/n)$, o pagal 1.5.1 teoremą $\boldsymbol{S} \sim W_k(n-1, \boldsymbol{\Sigma})$. Pagal 1 savybę $\bar{\mathbf{X}}$ ir $\boldsymbol{S}$ yra nepriklausomi. ▲

Reikia pažymėti, kad šie faktai buvo įrodyti 1.3.1 teoremoje.

3 savybė. Tegu $\boldsymbol{S}_1 \sim W_k(n_1, \boldsymbol{\Sigma})$ ir $\boldsymbol{S}_2 \sim W_k(n_2, \boldsymbol{\Sigma})$ yra nepriklausomi. Tada

$$\boldsymbol{S}_1 + \boldsymbol{S}_2 \sim W_k(n_1 + n_2, \boldsymbol{\Sigma}). \tag{1.6.4}$$

Įrodymas. Pagal Višarto skirstinio apibrėžimą galima užrašyti

$$\boldsymbol{S}_1 = \sum_{i=1}^{n_1} \mathbf{X}_i \mathbf{X}_i^T, \quad \boldsymbol{S}_2 = \sum_{i=n_1+1}^{n_1+n_2} \mathbf{X}_i \mathbf{X}_i^T;$$

čia $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_{n_1+n_2}$ nepriklausomi vienodai pasiskirstę normalieji $N_k(\mathbf{0}, \boldsymbol{\Sigma})$ atsitiktiniai vektoriai. Tada

$$\boldsymbol{S}_1 + \boldsymbol{S}_2 = \sum_{i=1}^{n_1+n_2} \mathbf{X}_i \mathbf{X}_i^T \sim W_k(n_1 + n_2, \boldsymbol{\Sigma}). \qquad \blacktriangle$$

4 savybė. Tegu $\boldsymbol{S} \sim W_k(n, \boldsymbol{\Sigma})$ ir $\boldsymbol{B} = [b_{ij}]_{r\times k}$ matrica. Tada $\boldsymbol{B}\boldsymbol{S}\boldsymbol{B}^T \sim W_r(n, \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^T)$.

Įrodymas. Pagal Višarto skirstinio apibrėžimą

$$\boldsymbol{S} = \sum_{i=1}^{n} \mathbf{X}_i \mathbf{X}_i^T;$$

čia $\mathbf{X}_1, ..., \mathbf{X}_n$ nepriklausomi vienodai pasiskirstę normalieji $N_k(\mathbf{0}, \boldsymbol{\Sigma})$ atsitiktiniai vektoriai. Tada

$$\boldsymbol{B}\boldsymbol{S}\boldsymbol{B}^T = \boldsymbol{B}(\sum_{i=1}^{n} \mathbf{X}_i \mathbf{X}_i^T)\boldsymbol{B}^T = \sum_{i=1}^{n} (\boldsymbol{B}\mathbf{X}_i)(\boldsymbol{B}\mathbf{X}_i)^T$$

Kadangi $\boldsymbol{B}\mathbf{X}_i \sim N_r(\mathbf{0}, \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^{\mathbf{T}}), i = 1, ..., n,$ tai remiantis Višarto skirstinio apibrėžimu

$$\boldsymbol{B}\boldsymbol{S}\boldsymbol{B}^T \sim W_r(n, \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^T). \qquad \blacktriangle$$

5 savybė. Tegu egzistuoja matrica $\boldsymbol{\Sigma}^{-1} = [\sigma^{ij}]_{k\times k}$ atvirkštinė matricai $\boldsymbol{\Sigma} = [\sigma_{ij}]_{k\times k}$ ir matrica $\boldsymbol{S}^{-1} = [S^{ij}]_{k\times k}$ atvirkštinė matricai $\boldsymbol{S} = [S_{ij}]_{k\times k}$. Jeigu $\boldsymbol{S} \sim W_k(n, \boldsymbol{\Sigma})$, tai

a) santykis

$$\frac{\sigma^{kk}}{S^{kk}} \sim \chi^2(n - k + 1); \tag{1.6.5}$$

ir nepriklauso nuo a. d. $S_{ij}, \quad i, j = 1, ..., k - 1;$

b) su bet kuriuo fiksuotu $\boldsymbol{L} \in \mathbf{R}^k$ santykis

$$\frac{\boldsymbol{L}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{L}}{\boldsymbol{L}^T \boldsymbol{S}^{-1} \boldsymbol{L}} \sim \chi^2(n-k+1). \tag{1.6.6}$$

Įrodymas. Tegu $\boldsymbol{S} = \sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i^T$; čia $\mathbf{X}_1, ..., \mathbf{X}_n$ nepriklausomi vienodai pasiskirstę normalieji $N_k(\mathbf{0}, \boldsymbol{\Sigma})$ atsitiktiniai vektoriai. Fiksuokime a. v. $\mathbf{X}_i$ koordinates $X_{1i}, X_{2i}, ..., X_{k-1,i}$. Tada sąlyginis a. d. X_{ki} skirstinys, kai kitos koordinatės fiksuotos, yra (žr. 3 priedą, 11 savybė) normalusis

$$N(\beta_1 X_{1i} + ... + \beta_{k-1} X_{k-1,i}, 1/\sigma^{kk}), \quad i = 1, ..., n.$$

Mažiausiųjų kvadratų metodu įvertinę nežinomus parametrus $\beta_1, ..., \beta_{k-1}$, gauname liekamąją kvadratų sumą

$$SS_E = \sum_i (X_{ki} - \hat{\beta}_1 X_{1i} - ... - \hat{\beta}_{k-1} X_{k-1,i})^2,$$

kuri pasiskirsčiusi kaip a. d. χ^2_{k-r}/σ^{kk}; čia r yra matricos $\boldsymbol{A} = [X_{ij}]_{(k-1)\times n}$ rangas. Kvadratų sumos SS_E skirstinys yra sąlyginis, kai fiksuotos $X_{ij}, i = 1, ..., k-1,\ j = 1, ..., n$ reikšmės. Tačiau kadangi šis skirstinys nepriklauso nuo fiksuotųjų reikšmių, tai jį galima interpretuoti kaip besąlyginį, t. y. SS_E nepriklauso nuo $S_{ij}, i, j = 1, ..., k-1$.

Jeigu $n > k-1$, tai matricos $[X_{ij}]_{(k-1)\times n}$ rangas su tikimybe 1 lygus $k-1$. Iš mažiausiųjų kvadratų teorijos turime, kad $SS_E = 1/S^{kk}$. Taigi tvirtinimas a) įrodytas.

Kad įrodytume tvirtinimą b), nagrinėkime transformaciją $\boldsymbol{BSB}^T$; čia $\boldsymbol{B}$ ortogonali matrica $\boldsymbol{BB}^T = \boldsymbol{I}$. Parinkime matricos $\boldsymbol{B}$ paskutinę eilutę, proporcingą vektoriui $\boldsymbol{L}^T$. Pagal 4 savybę $\boldsymbol{BSB}^T \sim W_k(n, \boldsymbol{B\Sigma B}^T)$. Be to

$$(\boldsymbol{BSB}^T)^{-1} = \boldsymbol{BS}^{-1}\boldsymbol{B}^T, \quad (\boldsymbol{B\Sigma B}^T)^{-1} = \boldsymbol{B\Sigma}^{-1}\boldsymbol{B}^T.$$

Matricos $\boldsymbol{BS}^{-1}\boldsymbol{B}^T$ paskutinis diagonalinis elementas proporcingas $\boldsymbol{L}^T\boldsymbol{S}^{-1}\boldsymbol{L}$, o matricos $\boldsymbol{B\Sigma}^{-1}\boldsymbol{B}^T$ – proporcingas $\boldsymbol{L}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{L}$ su tuo pačiu proporcingumo koeficientu. Pritaikę tvirtinimą a) gauname, kad

$$\frac{\boldsymbol{L}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{L}}{\boldsymbol{L}^T \boldsymbol{S}^{-1} \boldsymbol{L}} \sim \chi^2(n-k+1). \qquad \blacktriangle$$

6 savybė. Tegu $\boldsymbol{S} \sim W_k(n, \boldsymbol{\Sigma})$. Suskaidykime matricą $\boldsymbol{S}$ į blokus

$$\boldsymbol{S} = \begin{pmatrix} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} \\ \boldsymbol{S}_{21} & \boldsymbol{S}_{22} \end{pmatrix},$$

čia $\boldsymbol{S}_{11}, \boldsymbol{S}_{12}, \boldsymbol{S}_{22}$ yra eilės $r \times r, r \times s, s \times s$, matricos, $r + s = k$, $\boldsymbol{S}_{21} = \boldsymbol{S}_{12}^T$. Tada

$$\boldsymbol{S}_{22} - \boldsymbol{S}_{21}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12} \sim W_s(n-r, \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}); \tag{1.6.7}$$

čia $\boldsymbol{\Sigma}_{ij}$ atitinkami matricos $\boldsymbol{\Sigma}$ blokai.

Įrodymas. Matricą $\boldsymbol{S}$ galima užrašyti taip:

$$\boldsymbol{S} = \sum_{i=1}^{n} \mathbf{X}_i \mathbf{X}_i^T, \quad \mathbf{X}_i \sim N_k(\mathbf{0}, \boldsymbol{\Sigma}).$$

Tegu $\mathbf{X}_i^T = (\mathbf{X}_{1i}^T \vdots \mathbf{X}_{2i}^T)$ yra vektoriaus $\mathbf{X}_i$ atitinkamas suskaidymas į dimensijos r ir s vektorius $\mathbf{X}_{1i}$ ir $\mathbf{X}_{2i}$. Nagrinėkime $r+1$-mačius vektorius $(\mathbf{X}_{1i}^T \vdots \boldsymbol{L}^T \mathbf{X}_{2i}), i = 1, ..., n$, kai $\boldsymbol{L} \in \mathbf{R}^s$ yra fiksuotas vektorius. Šių vektorių Višarto matrica yra

$$\boldsymbol{S}_{\boldsymbol{L}} = \begin{pmatrix} \boldsymbol{S}_{11} & \boldsymbol{S}_{12}\boldsymbol{L} \\ \boldsymbol{L}^T \boldsymbol{S}_{21} & \boldsymbol{L}^T \boldsymbol{S}_{22} \boldsymbol{L} \end{pmatrix}.$$

Remdamiesi 5 savybės a) rezultatu gauname

$$\frac{|\boldsymbol{S}_{\boldsymbol{L}}|}{|\boldsymbol{S}_{11}|} \sim c\chi^2_{n-r}.$$

Kairėje pusėje esantis reiškinys yra

$$\boldsymbol{L}^T \boldsymbol{S}_{22} \boldsymbol{L} - \boldsymbol{L}^T \boldsymbol{S}_{21} \boldsymbol{S}_{11}^{-1} \boldsymbol{S}_{12} \boldsymbol{L} = \boldsymbol{L}^T (\boldsymbol{S}_{22} - \boldsymbol{S}_{21} \boldsymbol{S}_{11}^{-1} \boldsymbol{S}_{12}) \boldsymbol{L}.$$

Analogiškai gauname, kad proporcingumo koeficientas

$$c = \boldsymbol{L}^T (\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12}) \boldsymbol{L}.$$

Remdamiesi 1.5.1 teorema gauname (1.6.7). ▲

7 savybė. Tegu $\boldsymbol{S} \sim W_k(n, \boldsymbol{\Sigma})$, $n > k$ ir $|\boldsymbol{\Sigma}| > 0$. Tada determinantų santykis $|\boldsymbol{S}|/|\boldsymbol{\Sigma}|$ pasiskirstęs kaip nepriklausomų χ^2 atsitiktinių dydžių su $n-k+1, n-k+2, ..., n-1, n$ laisvės laipsnių sandauga.

Įrodymas. Pažymėkime $|\boldsymbol{S}|_r$ determinantą $|[S_{ij}]_{r\times r}|$, kai $i, j = 1, ..., r$. Tada yra teisingas dėstinys

$$\frac{|\boldsymbol{S}|}{|\boldsymbol{\Sigma}|} = \left(\frac{|\boldsymbol{S}|_k |\boldsymbol{\Sigma}|_{k-1}}{|\boldsymbol{S}|_{k-1} |\boldsymbol{\Sigma}|_k}\right) \left(\frac{|\boldsymbol{S}|_{k-1} |\boldsymbol{\Sigma}|_{k-2}}{|\boldsymbol{S}|_{k-2} |\boldsymbol{\Sigma}|_{k-1}}\right) \cdots \left(\frac{|\boldsymbol{S}|_1}{|\boldsymbol{\Sigma}|_1}\right). \tag{1.6.8}$$

Daugikliai remiantis 5 savybės teiginiu a) yra nepriklausomi a. d., turintys χ^2 skirstinius su nurodytais laisvės laipsnių skaičiais. ▲

1.6.1 apibrėžimas. Tegu $\boldsymbol{S} \sim W_k(n, \boldsymbol{\Sigma})$ ir $\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}/c)$ yra nepriklausomi. *Hotelingo T^2 statistika* apibrėžiama taip:

$$T^2 = cn\boldsymbol{Y}^T \boldsymbol{S}^{-1} \boldsymbol{Y}. \tag{1.6.9}$$

8 savybė. Santykis

$$\frac{n-k+1}{k} \frac{T^2}{n} \sim F(k, n-k+1; \delta) \tag{1.6.10}$$

turi necentrinį Fišerio skirstinį su k ir $n-k+1$ laisvės laipsnių ir necentriškumo parametru $\delta = c\boldsymbol{\mu}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$. Jeigu $\boldsymbol{\mu} = \mathbf{0}$, tai $\delta = 0$ ir

$$\frac{n-k+1}{k}\frac{T^2}{n} \sim F(k, n-k+1). \tag{1.6.11}$$

Įrodymas. Užrašykime T^2 dviejų daugiklių sandauga

$$T^2 = \frac{\boldsymbol{Y}^T\boldsymbol{S}^{-1}\boldsymbol{Y}}{\boldsymbol{Y}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}}(cn\boldsymbol{Y}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}).$$

Remiantis 5 savybės b) tvirtinimu su bet kuriuo fiksuotu $\boldsymbol{Y}$ santykis

$$\frac{\boldsymbol{Y}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}}{\boldsymbol{Y}^T\boldsymbol{S}^{-1}\boldsymbol{Y}} \sim \chi^2(n-k+1) \tag{1.6.12}$$

turi χ^2 skirstinį su $n-k+1$ laisvės laipsniu. Kadangi šis skirstinys nepriklauso nuo fiksuotosios $\boldsymbol{Y}$ reikšmės, tai (1.6.12) santykis nepriklauso nuo $\boldsymbol{Y}$. Remiantis daugiamačio normaliojo skirstinio savybėmis (žr. 3 priedas, 8 savybė) iš sąlygos, kad $\boldsymbol{Y} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}/c)$ išplaukia, kad $c\boldsymbol{Y}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} \sim \chi^2(k;\delta), \delta = c\boldsymbol{\mu}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$. Padaliję T^2 iš n, o skaitiklį ir vardiklį iš atitinkamų laisvės laipsnių skaičių, gausime a. d., turintį necentrinį Fišerio skirstinį (1.6.10). Jeigu vidurkių vektorius $\boldsymbol{\mu} = \mathbf{0}$, tai $\delta = 0$ ir skirstinys tampa centriniu (1.6.11). ▲

9 savybė. Jeigu $\boldsymbol{S} \sim W_k(n, \boldsymbol{\Sigma})$, $|\boldsymbol{\Sigma}| > 0$ ir $n > k$, tai Višarto skirstinio tankio funkcija yra

$$f(\boldsymbol{S}|k,n,\boldsymbol{\Sigma}) = \frac{1}{K}|\boldsymbol{S}|^{\frac{n-k-1}{2}}e^{-\frac{1}{2}Tr(\boldsymbol{S}\boldsymbol{\Sigma}^{-1})}; \tag{1.6.13}$$

čia normuojanti konstanta

$$K = K(k,n,\boldsymbol{\Sigma}) = 2^{nk/2}\pi^{k(k-1)/4}|\boldsymbol{\Sigma}|^{n/2}\prod_{i=1}^{k}\Gamma((n+1-i)/2).$$

Tankis sukoncentruotas $k(k+1)/2$-matės erdvės dalyje, kurioje matrica $\boldsymbol{S}$ teigiamai apibrėžta.

Įrodymas. Taikysime indukciją pagal parametrą k.

Jeigu $k = 1$, tai $\boldsymbol{S} = S_{11}$, yra suma kvadratų nepriklausomų normaliųjų a. d. su vienodomis dispersijomis σ_{11} (žr. (1.5.4)). Taigi

$$\frac{S_{11}}{\sigma_{11}} \sim \chi^2(n)$$

ir a. d. S_{11} tankio funkcija

$$f(S_{11}|1,n,\sigma_{11}) = \frac{1}{K}S_{11}^{\frac{n-2}{2}}e^{-\frac{1}{2}\frac{S_{11}}{\sigma_{11}}}, \tag{1.6.14}$$

$$K = K(1,n,\sigma_{11}) = 2^{n/2}\sigma_{11}^{n/2}\Gamma(n/2), \quad S_{11} > 0,$$

sutampa su (1.6.13).

Tegu $k = 2$ ir pradžioje nagrinėkime atvejį, kai kovariacinė matrica $\boldsymbol{\Sigma} = \boldsymbol{I}$ yra vienetinė. Turime paprastąją imtį $\boldsymbol{Z}_1, \boldsymbol{Z}_2, ..., \boldsymbol{Z}_n$, kurios elementai $\boldsymbol{Z}_i = (Z_{1i}, Z_{2i})^T \sim N_2(\boldsymbol{0}, \boldsymbol{I})$, $i = 1, ..., n$. Atitinkanti Višarto matrica

$$\boldsymbol{S} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \begin{pmatrix} \boldsymbol{Y}_1^T \boldsymbol{Y}_1 & \boldsymbol{Y}_1^T \boldsymbol{Y}_2 \\ \boldsymbol{Y}_2^T \boldsymbol{Y}_1 & \boldsymbol{Y}_2^T \boldsymbol{Y}_2 \end{pmatrix},$$

čia $\boldsymbol{Y}_1 = (Z_{11}, ..., Z_{1n})^T, \quad \boldsymbol{Y}_2 = (Z_{21}, ..., Z_{2n})^T$.

Fiksuokime vektorių $\boldsymbol{Y}_1 = (Z_{11}, ..., Z_{1n})^T$. Tada a. d. $S_{12} = S_{21} = \boldsymbol{Y}_1^T \boldsymbol{Y}_2$ skirstinys yra normalusis $N(0, S_{11})$, jo tankis

$$g_1(S_{12}|\boldsymbol{Y}_1) = (2\pi)^{-1/2} S_{11}^{-1/2} \exp\{-\frac{1}{2}\frac{S_{12}^2}{S_{11}}\}. \tag{1.6.15}$$

Prognozuojant vektoriaus $\boldsymbol{Z}_i$ antrąją koordinatę Z_{2i} pagal pirmąją koordinatę Z_{1i}, liekamoji kvadratų suma

$$SS_E = \min_\beta \sum_{i=1}^n (Z_{2i} - \beta Z_{1i})^2 = S_{22} - \frac{S_{12}^2}{S_{11}} \sim \chi^2(n-1)$$

turi χ^2 skirstinį ir nepriklauso nuo S_{12}.

Todėl a. v. $(S_{12}, SS_E)^T$ sąlyginis tankis gaunamas sudauginant tankį (1.6.15) ir χ^2 su $n-1$ laisvės laipsniu tankį. Gauname

$$g(S_{12}, SS_E|\boldsymbol{Y}_1) = \frac{1}{2^{n/2}\sqrt{\pi}\Gamma((n-1)/2)} \frac{SS_E^{(n-3)/2}}{S_{11}^{1/2}} \exp\{-\frac{S_{22}}{2}\}.$$

Perėję prie a. v. $(S_{12}, S_{22})^T$, gauname jo sąlyginį tankį (pakeitimo jakobianas lygus 1):

$$h(S_{12}, S_{22}|\boldsymbol{Y}_1) = \frac{1}{2^{n/2}\sqrt{\pi}\Gamma(n-1)/2} \frac{(S_{11}S_{22} - S_{12}^2)^{(n-3)/2}}{S_{11}^{(n-2)/2}} \exp\{-\frac{S_{22}}{2}\}.$$

Sudauginę šį tankį su tankiu a. d. S_{11} (formulėje (1.6.14) reikia imti $\sigma_{11} = 1$), gauname besąlyginį trimačio a. v. $(S_{11}, S_{12}, S_{22})^T$ tankį

$$f(S_{11}, S_{12}, S_{22}|2, n, \boldsymbol{I}) = \frac{1}{2^n\sqrt{\pi}\Gamma((n-1)/2)\Gamma(n/2)} |\boldsymbol{S}|^{(n-3)/2} \exp\{-\frac{S_{11}+S_{22}}{2}\}, \tag{1.6.16}$$

kuris sutampa su (1.6.13) tankiu, kai $k = 2$ ir $\boldsymbol{\Sigma} = \boldsymbol{I}$.

Pereiname prie bendro atvejo. Tegu $\boldsymbol{S} = \sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i^T$, kai $\mathbf{X}_1, ..., \mathbf{X}_n$ yra nepriklausomi normalieji a. v. $\mathbf{X}_i \sim N_2(\boldsymbol{0}, \boldsymbol{\Sigma})$, o $\boldsymbol{S}^* = \sum_{i=1}^n \boldsymbol{Z}_i \boldsymbol{Z}_i^T$, kai $\boldsymbol{Z}_1, ..., \boldsymbol{Z}_n$ yra nepriklausomi normalieji a. v. $\boldsymbol{Z}_i \sim N_2(\boldsymbol{0}, \boldsymbol{I})$. Parinkime trikampę matricą $\boldsymbol{C} = [c_{ij}]_{2\times 2}, c_{21} = 0$, kad $\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}^T = \boldsymbol{I}$ (1 priedas (8.2.11)). Matricos $\boldsymbol{C}\boldsymbol{S}\boldsymbol{C}^T = \sum_i (\boldsymbol{C}\mathbf{X}_i)(\boldsymbol{C}\mathbf{X}_i)^T$ elementų skirstinys sutampa su matricos $\boldsymbol{S}^*$

elementų skirstiniu, nes $\boldsymbol{C}\mathbf{X}_i \sim N_2(\mathbf{0}, \boldsymbol{I})$. Taigi matricos $\boldsymbol{S}$ elementų skirstinys gaunamas atliekant transformaciją $\boldsymbol{S}^* = \boldsymbol{C}\boldsymbol{S}\boldsymbol{C}^T$.

Kadangi

$$1 = |\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}^T| = |\boldsymbol{C}||\boldsymbol{\Sigma}||\boldsymbol{C}^T| = |\boldsymbol{\Sigma}||\boldsymbol{C}\boldsymbol{C}^T|,$$

tai

$$|\boldsymbol{C}\boldsymbol{C}^T| = \frac{1}{|\boldsymbol{\Sigma}|}, \quad |\boldsymbol{C}| = \frac{1}{\sqrt{|\boldsymbol{\Sigma}|}}.$$

Turime

$$\boldsymbol{\Sigma} = \boldsymbol{C}^{-1}(\boldsymbol{C}^T)^{-1} = (\boldsymbol{C}^T\boldsymbol{C})^{-1}, \quad Tr(\boldsymbol{C}\boldsymbol{S}\boldsymbol{C}^T) = Tr(\boldsymbol{S}\boldsymbol{C}^T\boldsymbol{C}) = Tr(\boldsymbol{S}\boldsymbol{\Sigma}^{-1}). \tag{1.6.17}$$

Be to

$$|\boldsymbol{S}^*| = |\boldsymbol{C}\boldsymbol{S}\boldsymbol{C}^T| = |\boldsymbol{C}||\boldsymbol{S}||\boldsymbol{C}^T| = |\boldsymbol{S}||\boldsymbol{C}\boldsymbol{C}^T| = \frac{|\boldsymbol{S}|}{|\boldsymbol{\Sigma}|}. \tag{1.6.18}$$

Lieka rasti pakeitimo jakobianą. Atlikta tokia transformacija

$$S_{11}^* = c_{11}^2 S_{11} + 2c_{11}c_{12}S_{12} + c_{12}^2 S_{22},$$
$$S_{12}^* = c_{11}c_{22}S_{12} + c_{12}c_{22}S_{22},$$
$$S_{22}^* = c_{22}^2 S_{22}.$$

Pakeitimo jakobianas

$$\left|\frac{\mathcal{D}(S_{11}^*, S_{12}^*, S_{22}^*)}{\mathcal{D}(S_{11}, S_{12}, S_{22})}\right| = \begin{pmatrix} c_{11}^2 & 2c_{11}c_{12} & c_{12}^2 \\ 0 & c_{11}c_{22} & c_{12}c_{22} \\ 0 & 0 & c_{22}^2 \end{pmatrix} = c_{11}^3 c_{22}^3 = |\boldsymbol{\Sigma}|^{3/2}.$$

Įrašę į (1.6.16) išraišką (1.6.17), (1.6.18) ir padauginę iš jakobiano, gauname tankį (1.6.13) atveju $k = 2$.

Tarkime, kad tankio išraiška (1.6.13) yra teisinga su $k - 1$. Įrodysime, kad ji teisinga ir su k. Įrodymas analogiškas kaip ir atveju $k = 2$.

Pradžioje nagrinėjame atvejį, kai kovariacinė matrica $\boldsymbol{\Sigma} = \boldsymbol{I}$ yra vienetinė.

Suskaidykime matricą $\boldsymbol{S}$ į blokus

$$\boldsymbol{S} = \begin{pmatrix} \tilde{\boldsymbol{S}} & \boldsymbol{S}_{(1)} \\ \boldsymbol{S}_{(1)}^T & \boldsymbol{S}_{kk} \end{pmatrix},$$

čia $\tilde{\boldsymbol{S}} = [S_{ij}]_{(k-1)\times(k-1)}$ matrica, kurios elementų skirstinio tankis pagal indukcijos prielaidą yra (1.6.13) pavidalo, kai vietoje k įrašyta $k-1$, o vietoje $\boldsymbol{\Sigma}$ – vienetinė matrica $\boldsymbol{I}$; $\boldsymbol{S}_{(1)}$ yra $(k-1)$-matis vektorius $\boldsymbol{S}_{(1)} = (\boldsymbol{Y}_1^T\boldsymbol{Y}_k, \boldsymbol{Y}_2^T\boldsymbol{Y}_k, ..., \boldsymbol{Y}_{k-1}^T\boldsymbol{Y}_k)^T = (S_{1k}, S_{2k}, ..., S_{k-1,k})^T$. Naudojame tuos pačius žymenis kaip ir 1.1 skyrelyje pakeitę vektoriaus $\mathbf{X}_i$ koordinates $X_{1i}, ..., X_{ki}$ vektoriaus $\boldsymbol{Z}_i$ koordinatėmis $Z_{1i}, ..., Z_{ki}$.

Fiksuokime vektorius $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_{k-1}$. Tada vektoriaus $\boldsymbol{S}_{(1)}$ skirstinys yra $(k-1)$-matis normalusis $N_{k-1}(\mathbf{0}, \tilde{\boldsymbol{S}})$, kurio tankis

$$(2\pi)^{-(k-1)/2}|\tilde{\boldsymbol{S}}|^{-1/2}\exp\{-\frac{1}{2}\boldsymbol{S}_{(1)}^T\tilde{\boldsymbol{S}}^{-1}\boldsymbol{S}_{(1)}\}. \tag{1.6.19}$$

Prognozuojant vektoriaus $\boldsymbol{Z}_i$ paskutiniąją koordinatę Z_{ki} pagal pirmąsias $Z_{1i}, ..., Z_{k-1,i}$ koordinates, liekamoji kvadratų suma

$$SS_E = \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} (Z_{ki} - \beta_1 Z_{1i} - ... - \beta_{k-1} Z_{k-1,i})^2 =$$

$$= S_{kk} - \boldsymbol{S}_{(1)}^T \tilde{\boldsymbol{S}}^{-1} \boldsymbol{S}_{(1)} \sim \chi^2(n-k+1)$$

turi χ^2 skirstinį ir nepriklauso nuo a. v. $\boldsymbol{S}_{(1)}$.

Todėl a. v. $(S_{1k}, S_{2k}, ..., S_{k-1,k}, SS_E)^T$ sąlyginis tankis gaunamas sudauginant tankį (1.6.19) ir χ^2 su $n-k+1$ laisvės laipsnių tankiu, kuriame argumentas yra SS_E. Perėję prie a. v. $(S_{1k}, S_{2k}, ..., S_{k-1,k}, S_{kk})^T$, gauname jo sąlyginį tankį (pakeitimo jakobianas lygus 1):

$$\frac{(S_{kk} - \boldsymbol{S}_{(1)}^T \tilde{\boldsymbol{S}}^{-1} \boldsymbol{S}_{(1)})^{(n-k-1)/2}}{2^{(n-k+1)/2} \Gamma((n-k+1)/2)(2\pi)^{(k-1)/2} |\tilde{\boldsymbol{S}}|^{1/2}} \exp\{-\frac{1}{2} S_{kk}\}. \qquad (1.6.20)$$

Galutinai matricos $\boldsymbol{S}$ elementų tankį gauname sudauginę besąlyginį matricos $\tilde{\boldsymbol{S}}$ elementų tankį (1.6.13), kuriame vietoje $\boldsymbol{\Sigma}$ įrašyta vienetinė matrica $\boldsymbol{I}$, o vietoje k įrašyta $k-1$, ir tankį (1.6.20).

Normuojanti konstanta

$$K(k-1, n, \boldsymbol{I}) 2^{(n-k+1)/2} \Gamma((n-k+1)/2)(2\pi)^{(k-1)/2} = K(k, n, \boldsymbol{I});$$

po eksponentės ženklu

$$-\frac{1}{2} Tr(\tilde{\boldsymbol{S}}) - \frac{1}{2} S_{kk} = -\frac{1}{2} Tr(\boldsymbol{S});$$

likęs daugiklis

$$|\tilde{\boldsymbol{S}}|^{(n-k-1)/2} (S_{kk} - \boldsymbol{S}_{(1)}^T \tilde{\boldsymbol{S}}^{-1} \boldsymbol{S}_{(1)})^{(n-k-1)/2} = |\boldsymbol{S}|^{(n-k-1)/2},$$

nes

$$|\boldsymbol{S}| = |\tilde{\boldsymbol{S}}| (S_{kk} - \boldsymbol{S}_{(1)}^T \tilde{\boldsymbol{S}}^{-1} \boldsymbol{S}_{(1)}).$$

Gauname, kad matricos $\boldsymbol{S}$ elementų tankio pavidalas, kai kovariacinė matrica $\boldsymbol{\Sigma} = \boldsymbol{I}$, yra

$$f(\boldsymbol{S}|k, n, \boldsymbol{\Sigma}) = \frac{1}{K(k, n, \boldsymbol{I})} |\boldsymbol{S}|^{\frac{n-k-1}{2}} e^{-\frac{1}{2} Tr(\boldsymbol{S})}. \qquad (1.6.21)$$

Pereiname prie bendro atvejo. Tegu $\boldsymbol{S} = \sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i^T$, kai $\mathbf{X}_1, ..., \mathbf{X}_n$ yra nepriklausomi normalieji a. v. $\mathbf{X}_i \sim N_k(\mathbf{0}, \boldsymbol{\Sigma})$, o $\boldsymbol{S}^* = \sum_{i=1}^n \boldsymbol{Z}_i \boldsymbol{Z}_i^T$, kai $\boldsymbol{Z}_1, ..., \boldsymbol{Z}_n$ yra nepriklausomi normalieji a. v. $\boldsymbol{Z}_i \sim N_k(\mathbf{0}, \boldsymbol{I})$. Parinkime trikampę matricą $\boldsymbol{C} = [c_{ij}]_{k \times k}, c_{ij} = 0, i > j$, kad $\boldsymbol{C\Sigma C}^T = \boldsymbol{I}$ (1 priedas, (8.2.11)). Matricos $\boldsymbol{CSC}^T = \sum_i (\boldsymbol{C}\mathbf{X}_i)(\boldsymbol{C}\mathbf{X}_i)^T$ elementų skirstinys sutampa su matricos $\boldsymbol{S}^*$ elementų skirstiniu, nes $\boldsymbol{C}\mathbf{X}_i \sim N_k(\mathbf{0}, \boldsymbol{I})$. Taigi matricos $\boldsymbol{S}$ elementų skirstinys gaunamas atliekant transformaciją $\boldsymbol{S}^* = \boldsymbol{CSC}^T$.

Kadangi

$$1 = |\boldsymbol{C\Sigma C}^T| = |\boldsymbol{C}||\boldsymbol{\Sigma}||\boldsymbol{C}^T| = |\boldsymbol{\Sigma}||\boldsymbol{CC}^T|,$$

tai

$$|\boldsymbol{CC}^T| = \frac{1}{|\boldsymbol{\Sigma}|}, \quad |\boldsymbol{C}| = \frac{1}{\sqrt{|\boldsymbol{\Sigma}|}}.$$

Turime

$$\boldsymbol{\Sigma} = \boldsymbol{C}^{-1}(\boldsymbol{C}^T)^{-1} = (\boldsymbol{C}^T\boldsymbol{C})^{-1}, \quad Tr(\boldsymbol{CSC}^T) = Tr(\boldsymbol{SC}^T\boldsymbol{C}) = Tr(\boldsymbol{S\Sigma}^{-1}). \tag{1.6.22}$$

Be to,

$$|\boldsymbol{S}^*| = |\boldsymbol{CSC}^T| = |\boldsymbol{C}||\boldsymbol{S}||\boldsymbol{C}^T| = |\boldsymbol{S}||\boldsymbol{CC}^T| = \frac{|\boldsymbol{S}|}{|\boldsymbol{\Sigma}|}. \tag{1.6.23}$$

Lieka rasti pakeitimo jakobianą. Atlikta tokia transformacija

$$S^*_{ij} = \sum_{k,l} c_{ik}S_{kl}c_{jl}, \quad i > j.$$

Dalinės išvestinės yra

$$\frac{\partial S^*_{ij}}{\partial S_{kk}} = c_{ik}c_{jk}, \quad \frac{\partial S^*_{ij}}{\partial S_{kl}} = c_{ik}c_{ji} + c_{il}c_{jk}, \quad l \neq k. \tag{1.6.24}$$

Surašykime skirtingus matricos $\boldsymbol{S}$ elementus į vieną bendrą vektorių: $(S_{11}, S_{12}, ..., S_{1k}, S_{22}, S_{23}, ..., S_{2k}, ..., S_{kk})^T$. Analogiška tvarka surašykime matricos $\boldsymbol{S}^*$ elementus. Kadangi matrica $\boldsymbol{C}$ trikampė, tai jakobiane visi elementai žemiau pagrindinės diagonalės lygūs 0. Todėl determinantui apskaičiuoti pakanka sudauginti diagonalinius elementus. Remiantis (1.6.24) gauname, kad jakobiano diagonaliniai elementai yra tokie: $c^2_{11}, c_{11}c_{22}, ..., c_{11}c_{kk}, c^2_{22}, c_{22}c_{33}, ..., c_{22}c_{kk}, ..., c^2_{kk}$. Sudauginę šiuos diagonalinius elementus gauname pakeitimo jakobianą

$$\boldsymbol{J} = c^{k+1}_{11}c^{k+1}_{22}\cdots c^{k+1}_{kk} = (c_{11}c_{22}\cdots c_{kk})^{k+1} = |\boldsymbol{C}|^{k+1} = \frac{1}{|\boldsymbol{\Sigma}|^{(k+1)/2}}. \tag{1.6.25}$$

Įrašę į (1.6.21) išraiškas (1.6.22), (1.6.23) ir padauginę iš jakobiano (1.6.25), gauname tankį (1.6.13). ▲

1.7. Pratimai

1.1. Tegu $\mathbf{X}_1, ..., \mathbf{X}_n$ yra paprastoji imtis a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Įrodykite, kad $\bar{\mathbf{X}}$ ir $\boldsymbol{S}$ yra pilnoji ir pakankamoji parametro $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ statistika.

1.2. Tegu $\mathbf{X}_1, ..., \mathbf{X}_n$ yra nepriklausomi a. v. $\mathbf{X}_j \sim N_k(\boldsymbol{\mu}c_j, \boldsymbol{\Sigma})$; čia $c_1, ..., c_n$ – žinomos konstantos. Įrodykite, kad $\bar{\mathbf{X}} = (1/\sum_j c^2_j)\sum_i c_i\mathbf{X}_i \sim N_k(\boldsymbol{\mu}, (1/\sum_j c^2_j)\boldsymbol{\Sigma})$, o matrica $\sum_j(\mathbf{X}_j - \bar{\mathbf{X}}c_j)(\mathbf{X}_j - \bar{\mathbf{X}}c_j)^T$ yra pasiskirsčiusi taip pat kaip matrica $\sum_{j=1}^{n-1}\boldsymbol{Z}_j\boldsymbol{Z}_j^T$; čia $\boldsymbol{Z}_1, ..., \boldsymbol{Z}_{n-1}$ yra vienodai pasiskirstę n. a. v. $\boldsymbol{Z}_j \sim N_k(\mathbf{0}, \boldsymbol{\Sigma})$.

1.3. Tegu $\boldsymbol{S} \sim W_k(n, \mathbf{\Sigma}), |\mathbf{\Sigma}| \neq 0$. Pažymėkime $\mathbf{\Theta} = [\theta_{ij}]_{k\times k}$ simetrinę kvadratinę matricą. Įrodykite, kad a. v. $(S_{11}, ..., S_{kk}, 2S_{12}, ..., 2S_{1k}, ..., 2S_{k-1,k})^T$ charakteristinė funkcija yra

$$\psi(\mathbf{\Theta}) = \mathbf{E}(\exp\{iTr(\boldsymbol{S}\mathbf{\Theta})\}) = \frac{|\mathbf{\Sigma}^{-1}|^{n/2}}{|\mathbf{\Sigma}^{-1} - 2i\mathbf{\Theta}|^{n/2}}.$$

1.4. Tegu $\mathbf{X}_1, ..., \mathbf{X}_n$ yra paprastoji imtis a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \mathbf{\Sigma})$. Apibrėžkime Hotelingo statistiką $T^2 = n(n-1)(\bar{\mathbf{X}} - \boldsymbol{\mu})^T \boldsymbol{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu})$. a) Įrodykite, kad

$$T^2 = n \max_{\boldsymbol{L}} \frac{[\boldsymbol{L}^T(\bar{\mathbf{X}} - \boldsymbol{\mu})]^2}{\boldsymbol{L}^T \boldsymbol{S} \boldsymbol{L}}.$$

b) Remdamiesi a) rezultatu įrodykite, kad

$$\mathbf{P}\{n[\boldsymbol{L}^T(\bar{\mathbf{X}} - \boldsymbol{\mu})]^2 < \boldsymbol{L}^T \boldsymbol{S} \boldsymbol{L} \mathbf{\Delta}, \quad \forall \boldsymbol{L} \in \mathbf{R}^k\} = 1 - \alpha;$$

čia $\mathbf{\Delta} = k(n-1)F_\alpha(k, n-k)/(n-k)$. c) Remdamiesi b) rezultatu įrodykite, kad su bet kokia tiesine $\boldsymbol{\mu}$ funkcija $\boldsymbol{L}^T\boldsymbol{\mu}, \boldsymbol{L} \in \mathbf{R}^k$

$$\mathbf{P}\{\boldsymbol{L}^T\bar{\mathbf{X}} - \sqrt{\boldsymbol{L}^T \boldsymbol{S}\boldsymbol{L}\mathbf{\Delta}/n} < \boldsymbol{L}^T\boldsymbol{\mu} < \boldsymbol{L}^T\bar{\mathbf{X}} - \sqrt{\boldsymbol{L}^T\boldsymbol{S}\boldsymbol{L}\mathbf{\Delta}/n}\} \geq 1 - \alpha.$$

1.5. Tegu $\boldsymbol{S} = [S_{ij}]_{k\times k} \sim W_k(n, \mathbf{\Sigma})$. Įrodykite, kad matrica $\boldsymbol{S}_r = [S_{ij}]_{r\times r}, r < k$, turi Višarto skirstinį $\boldsymbol{S}_r \sim W_r(n, \mathbf{\Sigma}_r)$; čia $\mathbf{\Sigma}_r = [\sigma_{ij}]_{r\times r}$.

1.6. Tarkime $\mathbf{X} = (X_1, ..., X_n)^T$ yra paprastoji imtis a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \mathbf{\Sigma})$, $|\mathbf{\Sigma}| > 0, n > k$. Įrodykite, kad matrica $\boldsymbol{S} = \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^T$ yra teigiamai apibrėžta su tikimybe 1.

1.7. Raskite parametrų $\boldsymbol{\mu} = \mathbf{E}(\mathbf{X})$ ir $\mathbf{\Sigma} = \boldsymbol{V}(\mathbf{X})$ nepaslinktuosius įvertinius pagal pateikiamas vektoriaus $\mathbf{X} = (X_1, X_2)^T$ nepriklausomas realizacijas

X_{1i}	34	12	33	44	89	59	50	88
X_{2i}	55	29	75	89	62	69	41	67

1.8. Raskite parametrų $\boldsymbol{\mu} = \mathbf{E}(\mathbf{X})$ ir $\mathbf{\Sigma} = \boldsymbol{V}(\mathbf{X})$ nepaslinktuosius įvertinius pagal pateikiamas vektoriaus $\mathbf{X} = (X_1, X_2, X_3, X_4)^T$ nepriklausomas realizacijas; čia X_1 ir X_2 yra pirmojo sūnaus galvos ilgis ir plotis, o X_3 ir X_4 – antrojo sūnaus galvos ilgis ir plotis (žr. [2]).

X_{1i}	X_{2i}	X_{3i}	X_{4i}	X_{1i}	X_{2i}	X_{3i}	X_{4i}
191	155	179	145	190	159	195	157
195	149	201	152	188	151	187	158
181	148	185	149	163	137	161	130
183	153	188	149	195	155	183	158
176	144	171	142	186	153	173	148
208	157	192	152	181	145	182	146
189	150	190	149	175	140	165	137
197	159	189	152	192	154	185	152
188	152	197	159	174	143	178	147
192	150	187	151	176	139	176	143
179	158	186	148	197	167	200	158
183	147	174	147	190	163	187	150
174	150	185	152				

1.9. a) Pamatavus 47 kačių svorį X_{1i} (kilogramais) ir jų širdies svorį X_{2i} (gramais), gauti tokie rezultatai [2]:

$$\sum_i X_{1i} = 110,9; \quad \sum_i X_{2i} = 432,5; \quad \sum_i X_{1i}^2 = 265,13;$$

$$\sum_i X_{2i}^2 = 4064,71; \quad \sum_i X_{1i}X_{2i} = 1029,62.$$

b) Atlikus analogiškus 97 katinų matavimus gauti tokie rezultatai:

$$\sum_i X_{1i} = 281,3; \quad \sum_i X_{2i} = 1098,3; \quad \sum_i X_{1i}^2 = 836,75;$$

$$\sum_i X_{2i}^2 = 13056,17; \quad \sum_i X_{1i}X_{2i} = 3275,55.$$

Raskite vidurkių vektoriaus ir kovariacijų matricos DT įverčius atskirai atveju a) ir atveju b).

1.10. Lentelėje yra pateikta $n = 30$ lentų standumo matavimų, t. y. keturmačio a. v. $\mathbf{X} = (X_1, X_2, X_3, X_4)^T$ realizacijų; čia X_1 – standumas veikiant smūgine banga, X_2 – standumas veikiant vibracijai, X_3 ir X_4 – stacionarios būsenos [9]. Tardami, kad buvo stebėtas normalusis vektorius $\mathbf{X} \sim N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, raskite a) parametrų $\boldsymbol{\mu}$ ir $\boldsymbol{\Sigma}$ DT įverčius; b) a. v. $\boldsymbol{Y} = (X_1-X_4, X_2-X_4, X_3-X_4)^T$ skirstinio parametrų DT įverčius.

X_{1i}	X_{2i}	X_{3i}	X_{4i}	X_{1i}	X_{2i}	X_{3i}	X_{4i}
1889	1651	1561	1778	1954	2149	1180	1281
2403	2048	2087	2197	1325	1170	1002	1176
2119	1700	1815	2222	1419	1371	1252	1308
1645	1627	1110	1533	1828	1634	1602	1755
1976	1916	1614	1883	1725	1594	1313	1646
1712	1712	1439	1546	2276	2189	1547	2111
1943	1685	1271	1671	1899	1614	1422	1477
2104	1820	1717	1874	1633	1513	1290	1516
2983	2794	2412	2581	2061	1867	1646	2037
1745	1600	1384	1508	1856	1493	1356	1533
1710	1591	1518	1667	1727	1412	1238	1469
2046	1907	1627	1898	2168	1896	1701	1834
1840	1841	1595	1741	1655	1675	1414	1597
1867	1685	1493	1678	2326	2301	2065	2234
1859	1649	1389	1714	1490	1382	1214	1284

1.11. (**1.10** pratimo tęsinys). Tarę, kad kovariacinė matrica $\boldsymbol{\Sigma}$ yra žinoma ir sutampa su **1.10** pratime surastu įverčiu, a) raskite vektoriaus $\boldsymbol{Y}$ vidurkių vektoriaus koordinačių pasikliovimo intervalus remdamiesi Bonferonio nelygybe (pasikliovimo lygmuo $Q = 0,95$); b) patikrinkite hipotezę $H : \mathbf{E}\boldsymbol{Y} = \boldsymbol{\nu}_0 = (120; 0; -200)^T$.

1.12. Buvo tiriama $n = 96$ suomių studentų muzikiniai gebėjimai; X_1 – melodija, X_2 – harmonija, X_3 – tempas, X_4 – ritmas, X_5 – frazuotė, X_6 – balansas, X_7 – stilius. Pagal šiuos duomenis gauti vidurkių vektoriaus ir dispersijų vektoriaus NMD įverčiai [9]:

$$(\hat{\mu}_1; ...; \hat{\mu}_7) = (28,1; 26,6; 35,4; 34,2; 23,6; 22,0; 22,7),$$

$$(\hat{\sigma}_{11}^2; ...; \hat{\sigma}_{77}^2) = (33,178; 34,223; 14,592; 26,214; 14,138; 15,445; 16,241).$$

Tarę, kad buvo stebėtas normalusis a. v., kurio koordinačių dispersijos yra žinomos ir sutampa su gautais įverčiais, raskite vidurkių vektoriaus koordinačių pasikliovimo lygmens $Q = 0,95$ pasikliovimo intervalus a) remdamiesi pasikliovimo sritimi (1.4.1); b) remdamiesi Bonferonio nelygybe. Palyginkite gautuosius intervalus su intervalais, sudarytais kiekvienai koordinatei atskirai remiantis vienmate teorija.

1.13. Automobilių gamyboje buvo tikrinama suvirinimo proceso charakteristikos: X_1 – įtampa; X_2 – srovės stiprumas; X_3 – padavimo greitis; X_4 – inertinių dujų srovė. Lentelėje pateikta $n = 40$ keturmačio a. v. $\mathbf{X} = (X_1, X_2, X_3, X_4)^T$ matavimų, atliktų 5 sekundžių intervalu [9].

X_{1i}	X_{2i}	X_{3i}	X_{4i}	X_{1i}	X_{2i}	X_{3i}	X_{4i}
23,0	276	289,6	51,0	22,9	271	288,3	51,0
22,0	281	289,0	51,7	21,3	274	289,0	52,0
22,8	270	288,2	51,3	21,8	280	290,0	52,0
22,1	278	288,0	52,3	22,0	268	288,3	51,0
22,5	275	288,0	53,0	22,8	269	288,7	52,0
22,2	273	288,0	51,0	22,0	264	290,0	51,0
22,0	275	290,0	53,0	22,5	273	288,6	52,0
22,1	268	289,0	54,0	22,2	269	288,2	52,0
22,5	277	289,0	52,0	22,6	273	286,0	52,0
22,5	278	289,0	52,0	21,7	283	290,0	52,7
22,3	269	287,0	54,0	21,9	273	288,7	55,3
21,8	274	287,6	52,0	22,3	264	287,0	52,0
22,3	270	288,4	51,0	22,2	263	288,0	52,0
22,2	273	290,2	51,3	22,3	266	288,6	51,7
22,1	274	286,0	51,0	22,0	263	288,0	51,7
22,1	277	287,0	52,0	22,8	272	289,0	52,3
21,8	277	287,0	51,0	22,0	277	287,7	53,3
22,6	276	290,0	51,0	22,7	272	289,0	52,0
22,3	278	287,0	51,7	22,6	274	287,2	52,7
23,0	266	289,1	51,0	22,7	270	290,0	51,0

Tarę, kad buvo stebėtas normalusis a. v. $\mathbf{X} \sim N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ raskite parametrų $\boldsymbol{\mu}$ ir $\boldsymbol{\Sigma}$ DT įverčius.

1.14. (**1.13** pratimo tęsinys). Tarę, kad kovariacinė matrica $\boldsymbol{\Sigma}$ yra žinoma ir sutampa su **1.13** pratime surastu įverčiu, a) patikrinkite hipotezę $H : \boldsymbol{\mu} = \boldsymbol{\mu}_0 = (22,3; 270; 288,5; 52)^T$ (kriterijaus reikšmingumo lygmuo $\alpha = 0,05$); b) sudarykite nepriklausomo nuo turimų a. v. $\mathbf{X}^* = (X_1^*, X_2^*, X_3^*, X_4^*)^T$ prognozės sritį, kai pasikliovimo lygmuo $Q = 0,95$.

Atsakymai ir nurodymai

1.4. Nurodymas. Remdamiesi tankiu (1.6.13) gauname

$$\mathbf{E}(\exp\{iTr(\boldsymbol{S\Theta})\}) = \int \cdots \int \frac{|\boldsymbol{S}|^{(n-k-1)/2}}{K(k,n,\boldsymbol{\Sigma})} e^{-\frac{1}{2}Tr(\boldsymbol{S}[\boldsymbol{\Sigma}^{-1}-2i\boldsymbol{\Theta}])} d\boldsymbol{S} =$$

$$\frac{K(k,n,\boldsymbol{\Sigma})}{K(k,n,(\boldsymbol{\Sigma}^{-1}-2i\boldsymbol{\Theta}))^{-1}},$$

nes likęs integralas lygus 1. Lieka pasinaudoti normuojančio daugiklio K išraiška iš (1.6.13). **1.7.** $\hat{\boldsymbol{\mu}} = (51,125; 60,875)^T$; $\hat{\sigma}_{11} = 722,9821$; $\hat{\sigma}_{12} = \hat{\sigma}_{21} = 178,0179$; $\hat{\sigma}_{22} = 362,9821$. **1.8.** $\hat{\boldsymbol{\mu}} = (185,72; 151,12; 183,84; 149,24)$; $\hat{\sigma}_{11} = 95,293$; $\hat{\sigma}_{12} = 52,868$; $\hat{\sigma}_{13} = 69,662$; $\hat{\sigma}_{14} = 46,112$; $\hat{\sigma}_{22} = 54,360$; $\hat{\sigma}_{23} = 51,312$; $\hat{\sigma}_{24} = 35,053$; $\hat{\sigma}_{33} = 100,807$; $\hat{\sigma}_{34} = 56,540$; $\hat{\sigma}_{44} = 45,023$. **1.9** a) $\hat{\boldsymbol{\mu}} = (2,360; 9,202)^T$; $\hat{\sigma}_{11} = 0,073$; $\hat{\sigma}_{12} = \hat{\sigma}_{21} = 0,1977$; $\hat{\sigma}_{22} = 1,804$. b) $\hat{\boldsymbol{\mu}} = (2,900; 11,323)^T$; $\hat{\sigma}_{11} = 0,216$; $\hat{\sigma}_{12} = \hat{\sigma}_{21} = 0,933$; $\hat{\sigma}_{22} = 6,397$. **1.10.** a) DT įverčiai:

$$\hat{\boldsymbol{\mu}} = \begin{pmatrix} 1906,1 \\ 1749,5 \\ 1509,1 \\ 1725,0 \end{pmatrix}; \quad \hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 102095,8 & 91459,8 & 84380,1 & 91089,7 \\ 91459,8 & 98126,5 & 73599,2 & 78362,2 \\ 84380,1 & 73599,2 & 88853,2 & 87340,6 \\ 91089,7 & 78362,2 & 87340,6 & 100753,7 \end{pmatrix}.$$

b) a. v. $\boldsymbol{Y}$ taip pat normalusis $N_3(\boldsymbol{\nu}, \boldsymbol{\Sigma}_0)$; parametrų įverčiai:

$$\hat{\boldsymbol{\nu}} = \begin{pmatrix} 181,13 \\ 24,57 \\ -215,83 \end{pmatrix}; \quad \hat{\boldsymbol{\Sigma}}_0 = \begin{pmatrix} 20670,1 & 22761,5 & 6703,4 \\ 22761,5 & 42155,7 & 8650,0 \\ 6703,4 & 8650,0 & 14925,6 \end{pmatrix}.$$

1.11. a) vidurkių vektoriaus $\boldsymbol{\nu} = (\nu_1, \nu_2, \nu_3)^T$ koordinačių pasikliovimo intervalai yra tokie: (118,29; 243,97); (-65,17;114,31); (-269,23;-162,44); b) statistika U^2, kuri esant teisingai hipotezei turi χ^2 skirstinį su 3 laisvės laipsniais, įgijo reikšmę 11,132; kadangi P reikšmė $pv = \mathbf{P}\{\chi^2_3 > 11,132\} = 0,011$, tai hipotezė atmetama, jeigu kriterijaus reikšmingumo lygmuo viršija 0,011. **1.12.** Pažymėkime vidurkių vektoriaus i-osios koordinatės μ_i pasikliovimo intervalo, gauto remiantis elipsoidu (1.4.1), rėžius $\underline{\mu}'_i, \overline{\mu}'_i$, o gauto remiantis Bonferonio nelygybe – $\underline{\mu}''_i, \overline{\mu}''_i$. Gautieji pasikliovimo intervalų rėžiai pateikti lentelėje. Pirmuose stulpeliuose pateikti pasikliovimo intervalai sudaryti kiekvienai koordinatei atskirai.

i	$\underline{\mu}_i$	$\overline{\mu}_i$	$\underline{\mu}''_i$	$\overline{\mu}''_i$	$\underline{\mu}'_i$	$\overline{\mu}'_i$
1	26,95	29,25	26,52	29,68	25,90	30,30
2	25,43	27,77	24,99	28,21	24,36	28,84
3	34,64	36,16	34,35	36,45	33,94	36,86
4	33,18	35,22	32,79	35,61	32,24	36,16
5	22,85	24,35	22,57	24,63	22,16	25,04
6	21,21	22,79	20,92	23,08	20,50	23,50
7	21,89	23,51	21,59	23,81	21,16	24,24

1.13. Vidurkių vektoriaus ir kovariacinės matricos įverčiai yra

$$\hat{\boldsymbol{\mu}} = \begin{pmatrix} 22,2875 \\ 272,575 \\ 288,435 \\ 51,975 \end{pmatrix}; \quad \hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 0,1461 & -0,4003 & 0,0149 & -0,0936 \\ -0,4003 & 24,1444 & 0,4849 & 0,5244 \\ 0,0149 & 0,4849 & 1,2013 & -0,0766 \\ -0,0936 & 0,5244 & -0,0766 & 0,9014 \end{pmatrix}.$$

1.14. a) Statistika U^2 iš (1.4.6) įgijo reikšmę 11,8514; P reikšmė $pv = \mathbf{P}\{\chi^2_4 > 11,8514\} = 0,0185$; hipotezė atmetama, jeigu kriterijaus reikšmingumo lygmuo viršija 0,0185. b) Pažymėkime $Z_1 = X_1^* - 22,2875$, $Z_2 = X_2^* - 272,575$, $Z_3 = X_3^* - 288,435$, $Z_4 = X_4^* - 51,975$. Tada prognozės elipsoidas yra $7,6249Z_1^2 + 0,0440Z_2^2 + 0,8465Z_3^2 + 1,1993Z_4^2 + 0,2254Z_1Z_2 - 0,1886Z_1Z_3 + 1,436Z_1Z_4 - 0,0404Z_2Z_3 - 0,0312Z_2Z_4 + 0,1478Z_3Z_4 < 9,256$. **Nurodymas.** A.d. $n(\mathbf{X}^* - \hat{\boldsymbol{\mu}})^T\boldsymbol{\Sigma}^{-1}(\mathbf{X}^* - \hat{\boldsymbol{\mu}})/(n+1)$ turi χ^2 skirstinį su 4 laisvės laipsniais.

2 skyrius

Hotelingo statistikos taikymai

Minėjome, kad T^2 statistika yra Stjudento t statistikos analogas daugiamačiu atveju. Vienmačiu atveju Stjudento t statistika yra naudojama hipotezėms apie normaliojo skirstinio vidurkio reikšmes arba hipotezei apie vidurkių lygybę, kai dispersija nežinoma,tikrinti, taip T^2 statistika daugiamačiu atveju naudojama hipotezėms apie normaliojo vektoriaus vidurkio vektoriaus reikšmes arba vidurkių vektorių lygybę, kai kovariacinė matrica nežinoma, tikrinti.

2.1. Išvados apie vidurkių vektorių, kai kovariacinė matrica nežinoma

Jeigu turime paprastąją imtį $(X_1, ..., X_n)^T$, gautą stebint a. d. $X \sim N(\mu, \sigma^2)$, ir reikia patikrinti hipotezę $H : \mu = \mu_0$ arba sudaryti parametro μ pasikliovimo intervalą, šiam tikslui buvo naudojama Stjudento statistika

$$t = \sqrt{n}\frac{\bar{X} - \mu}{s}, \quad \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2. \tag{2.1.1}$$

Parametro μ pasikliovimo intervalas, kai pasikliovimo lygmuo $Q = 1 - \alpha$, nusakomas nelygybe

$$|t| = \sqrt{n}\frac{|\bar{X} - \mu|}{s} < t_{\alpha/2}(n-1). \tag{2.1.2}$$

Hipotezė $H : \mu = \mu_0$, kai alternatyva $\bar{H} : \mu \neq \mu_0$, atmetama α lygmens kriterijumi, kai teisinga nelygybė

$$|t| = \sqrt{n}\frac{|\bar{X} - \mu_0|}{s} > t_{\alpha/2}(n-1). \tag{2.1.3}$$

Tarkime, kad paprastoji imtis $\mathbf{X}_1, ..., \mathbf{X}_n$ yra gauta stebint a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $|\boldsymbol{\Sigma}| > 0, n > k$. Vidurkių vektorius $\boldsymbol{\mu}$ ir kovariacinė matrica $\boldsymbol{\Sigma}$ nežinomi. Įrašykime į kairiąją sąryšio (1.3.2) pusę kovariacinės matricos NMD įvertinį

$$T^2 = n(\bar{\mathbf{X}} - \boldsymbol{\mu})^T \hat{\boldsymbol{\Sigma}}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}) = n(n-1)(\bar{\mathbf{X}} - \boldsymbol{\mu})^T \boldsymbol{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}), \tag{2.1.4}$$

$$\bar{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i, \quad \boldsymbol{S} = \sum_{i=1}^{n}(\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^T.$$

Kadangi remiantis 1.3.1 teorema $\bar{\mathbf{X}}$ ir $\boldsymbol{S}$ nepriklausomi ir

$$\bar{\mathbf{X}} \sim N_k(\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{\Sigma}), \quad \boldsymbol{S} \sim W_k(n-1, \boldsymbol{\Sigma}),$$

tai pagal 1.6.1 apibrėžimą (2.1.4) yra Hotelingo T^2 statistika. Remiantis (1.6.10)

$$\frac{n-1-k+1}{k}\frac{T^2}{n-1} = \frac{n(n-k)}{k}(\bar{\mathbf{X}} - \boldsymbol{\mu})^T \boldsymbol{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \sim F(k, n-k). \tag{2.1.5}$$

Atveju $k = 1$ (2.1.5) yra Stjudento statistikos $t = \sqrt{n}(\bar{X} - \mu)/s \sim S(n-1)$ kvadratas, kuris turi Fišerio skirstinį su 1 ir $n-1$ laisvės laipsniu.

2.1.1. Vidurkių vektoriaus pasikliovimo sritys

Tegu $F_\alpha(k, n-k)$ yra Fišerio skirstinio su k ir $n-k$ laisvės laipsnių lygmens α kritinė reikšmė. Apibrėžkime k-matės erdvės poaibį

$$\boldsymbol{C}(\bar{\mathbf{X}}, \boldsymbol{S}) = \{\boldsymbol{\mu} : \frac{n(n-k)}{k}(\bar{\mathbf{X}} - \boldsymbol{\mu})^T \boldsymbol{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \le F_\alpha(k, n-k)\}.$$

Tada $\boldsymbol{C}(\bar{\mathbf{X}}, \boldsymbol{S})$ yra vidurkių vektoriaus $\boldsymbol{\mu}$ pasikliovimo sritis, kai pasikliovimo lygmuo $Q = 1 - \alpha$:

$$\mathbf{P}\{\boldsymbol{\mu} \in \boldsymbol{C}(\bar{\mathbf{X}}, \boldsymbol{S}) | \boldsymbol{\mu}\} = Q = 1 - \alpha. \tag{2.1.6}$$

Matome, kad ši sritis yra elipsoido pavidalo, jos centras taške $\bar{\mathbf{X}}$, o didumas priklauso nuo matricos $\boldsymbol{S}$. Vienmačiu atveju gauname pirmiau pateiktą parametro μ pasikliovimo intervalą (2.1.2).

2.1.2. Vidurkių vektoriaus koordinačių pasikliovimo intervalų rinkiniai

Parametro $\theta = \boldsymbol{L}^T\boldsymbol{\mu}$ pasikliovimo intervalą,kai pasikliovimo lygmuo $Q = 1 - \alpha$, galima rasti naudojant paprastąją imtį $\boldsymbol{Y} = (\boldsymbol{L}^T\mathbf{X}_1, ..., \boldsymbol{L}^T\mathbf{X}_n)^T$, gautą stebint a. d. $\boldsymbol{L}^T\mathbf{X} \sim N(\theta, \sigma_{\boldsymbol{L}}^2)$, $\sigma_{\boldsymbol{L}}^2 = \boldsymbol{L}^T\boldsymbol{\Sigma}\boldsymbol{L}$. Gauname

$$\hat{\theta} = \boldsymbol{L}^T\bar{\mathbf{X}} \sim N(\theta, \frac{1}{n}\sigma_{\boldsymbol{L}}^2), \quad \hat{\sigma}_{\boldsymbol{L}}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\boldsymbol{L}^T\mathbf{X}_i - \boldsymbol{L}^T\bar{\mathbf{X}})^2 = \frac{1}{n-1}\boldsymbol{L}^T\boldsymbol{S}\boldsymbol{L};$$

$$\sqrt{n}\frac{\hat{\theta}-\theta}{\hat{\sigma}_{\boldsymbol{L}}} = \sqrt{n(n-1)}\frac{\hat{\theta}-\theta}{\sqrt{\boldsymbol{L}^T\boldsymbol{S}\boldsymbol{L}}} \sim S(n-1).$$

Analogiškai (2.1.2) gauname parametro θ pasikliovimo intervalo $(\overline{\theta}, \underline{\theta})$ rėžius

$$\hat{\theta} \pm t_{\alpha/2}(n-1)\frac{\sqrt{\boldsymbol{L}^T\boldsymbol{S}\boldsymbol{L}}}{\sqrt{n(n-1)}}. \tag{2.1.7}$$

Atskiru atveju imdami $\boldsymbol{L} = (0, ..., 0, 1, 0, ..., 0)^T$, kai vienetas parašytas j-oje vietoje, gauname vidurkių vektoriaus $\boldsymbol{\mu}$ j-osios koordinatės μ_j pasikliovimo intervalą:

$$\mathbf{P}\{\underline{\mu}_j = \bar{X}_j - t_{\alpha/2}(n-1)\frac{\sqrt{S_{jj}}}{\sqrt{n(n-1)}} < \mu_j < \bar{X}_j + t_{\alpha/2}(n-1)\frac{\sqrt{S_{jj}}}{\sqrt{n(n-1)}} = \overline{\mu}_j|\mu_j\}$$
$$= Q = 1-\alpha. \tag{2.1.8}$$

Tarkime, kad reikia sudaryti pasikliovimo intervalų rinkinį $(\underline{\mu}_j, \overline{\mu}_j), j = 1, ..., k$, tokį, kad

$$\mathbf{P}\{\underline{\mu}_j < \mu_j < \overline{\mu}_j, \forall j = 1, ..., k|\boldsymbol{\mu}\} = Q = 1-\alpha. \tag{2.1.9}$$

Aišku, kad intervalai (2.1.8) šios sąlygos netenkina. Pavyzdžiui, jeigu a. v. $\mathbf{X}$ koordinatės nepriklausomos, tai intervalai (2.1.8) uždengia visus parametrus $\mu_j, j = 1, ..., k$, su tikimybe $(1-\alpha)^k < 1-\alpha$. Norint, kad ši tikimybė būtų lygi $1-\alpha$, reikėtų atskirus intervalus sudaryti su pasikliovimo lygmeniu $(1-\alpha)^{1/k}$.

Atsižvelgdami į a. v. $\mathbf{X}$ koordinačių priklausomumą, pasikliovimo intervalų rinkinį galime gauti remdamiesi pasikliovimo sritimi (2.1.6).

2.1.1 teorema. *Imant bet kokius vektorius $\boldsymbol{L} \in \mathbf{R}^k$ galioja lygybė*

$$\mathbf{P}\{\boldsymbol{L}^T\bar{\mathbf{X}} - \Delta\sqrt{\boldsymbol{L}^T\boldsymbol{S}\boldsymbol{L}} \le \boldsymbol{L}^T\boldsymbol{\mu} \le \boldsymbol{L}^T\bar{\mathbf{X}} + \Delta\sqrt{\boldsymbol{L}^T\boldsymbol{S}\boldsymbol{L}}, \forall \boldsymbol{L} \in \mathbf{R}^k|\boldsymbol{\mu}\} = Q = 1-\alpha, \tag{2.1.10}$$

čia

$$\Delta^2 = \frac{k}{n(n-k)}F_\alpha(k, n-k).$$

Įrodymas. Analogiškai teoremai 1.4.1 remdamiesi (2.1.6) gauname

$$\mathbf{P}\{(\bar{\mathbf{X}}-\boldsymbol{\mu})^T\boldsymbol{S}^{-1}(\bar{\mathbf{X}}-\boldsymbol{\mu}) \le \Delta^2|\boldsymbol{\mu}\} = \mathbf{P}\{\max_{\boldsymbol{L}} \frac{[\boldsymbol{L}^T(\bar{\mathbf{X}}-\boldsymbol{\mu})]^2}{\boldsymbol{L}^T\boldsymbol{S}\boldsymbol{L}} \le \Delta^2|\boldsymbol{\mu}\}$$

$$= \mathbf{P}\{|\boldsymbol{L}^T(\bar{\mathbf{X}}-\boldsymbol{\mu})| \le \Delta\sqrt{\boldsymbol{L}^T\boldsymbol{S}\boldsymbol{L}}, \forall \boldsymbol{L} \in \mathbf{R}^k|\boldsymbol{\mu}\} = Q = 1-\alpha.$$

▲

Lygybę (2.1.10) galima interpretuoti kaip pasikliovimo intervalus, sudarytus iš karto visoms tiesinėms funkcijoms $\boldsymbol{L}^T\boldsymbol{\mu}, \boldsymbol{L} \in \mathbf{R}^k$. Imdami $\boldsymbol{L}_1 = (1, 0, ..., 0)^T$,

$\boldsymbol{L}_2 = (0, 1, ..., 0)^T, ..., \boldsymbol{L}_k = (0, 0, ..., 1)^T$, gausime vidurkių vektoriaus $\boldsymbol{\mu}$ visų koordinačių pasikliovimo intervalų rinkinį:

$$\underline{\mu}_i' = \bar{X}_i - \Delta\sqrt{S_{ii}}, \quad \overline{\mu}_i' = \bar{X}_i + \Delta\sqrt{S_{ii}},$$

kuriam

$$\mathbf{P}\{\underline{\mu}_i' < \mu_i < \overline{\mu}_i', \ \forall \ i = 1, ..., k\} \geq Q = 1 - \alpha, \tag{2.1.11}$$

Kadangi intervalai (2.1.10) sudaryti iš karto visoms tiesinėms funkcijoms $\boldsymbol{L}^T\boldsymbol{\mu}$, tai šiam konkrečiam k tiesinių funkcijų rinkiniui pasikliovimo lygmuo Q gali gerokai viršyti $1 - \alpha$, t. y. intervalai (2.1.11) gali būti ilgesni negu būtina.

Šiek tiek trumpesnius intervalus gauname remdamiesi Bonferonio nelygybe. Tegu A_j reiškia įvykį, kad parametro μ_j tipo (2.1.6) pasikliovimo intervalas uždengia tikrąją reikšmę. Tada

$$\mathbf{P}\{A_1 \cap \cdots \cap A_k\} = 1 - \mathbf{P}\{\bar{A}_1 \cup \cdots \cup \bar{A}_k\} \geq$$
$$\geq 1 - (\mathbf{P}\{\bar{A}_1\} + \cdots + \mathbf{P}\{\bar{A}_k\}). \tag{2.1.12}$$

Jeigu pasikliovimo intervalus (2.1.6) sudarysime taip, kad kiekvieno iš jų pasikliovimo lygmuo būtų $1 - \alpha/k$;

$$\underline{\mu}_i'' = \bar{X}_i - t_{\alpha/(2k)}(n-1)\frac{\sqrt{S_{jj}}}{\sqrt{n(n-1)}},$$

$$\overline{\mu}_i'' = \bar{X}_i + t_{\alpha/(2k)}(n-1)\frac{\sqrt{S_{jj}}}{\sqrt{n(n-1)}}, \tag{2.1.13}$$

tai iš (2.1.12) plaukia, kad tikimybė, jog jie visi uždengs tikrąsias vidurkių reikšmes, bus nemažesnė už $1 - \alpha$. Intervalų (2.1.11) ilgis yra

$$\frac{\Delta\sqrt{n(n-1)}}{t_{\alpha/(2k)}(n-1)} = \frac{\sqrt{k(n-1)}F_\alpha(k, n-k)}{\sqrt{n-k}\,t_{\alpha/(2k)}(n-1)}$$

kartų didesnis už intervalų (2.1.13), sudarytų remiantis Bonferonio nelygybe, ilgį.

2.1.1 pavyzdys. (**1.2.1** pavyzdžio tęsinys.) Tarkime, kad **1.2.1** pavyzdyje stebėjimai yra trimačio normaliojo skirstinio $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ paprastosios imties realizacija, kai vidurkių vektorius $\boldsymbol{\mu}$ ir kovariacinė matrica $\boldsymbol{\Sigma}$ yra nežinomi. Surasime vidurkių vektoriaus $\boldsymbol{\mu}$ pasikliautinąją sritį ir vidurkių vektoriaus koordinačių pasikliovimo intervalų rinkinius, kai pasikliovimo lygmuo $Q = 0,95$.

Pasikliautinoji sritis (2.1.6) turi tokį pavidalą:

$$\boldsymbol{C}(\bar{\mathbf{X}}, \boldsymbol{S}) = \{\boldsymbol{\mu} : 270(\bar{\mathbf{X}} - \boldsymbol{\mu})^T \boldsymbol{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \leq F_{0,05}(3,\ 27)\}.$$

Pasinaudoję parametrų $\boldsymbol{\mu}$ ir $\boldsymbol{\Sigma}$ įverčiais iš **1.2.1** pavyzdžio, **1.4.1** pavyzdyje apskaičiuota matricos $\boldsymbol{S}^{-1}$ reikšme ir pažymėję $\boldsymbol{Z} = (Z_1; Z_2; Z_3)^T = \boldsymbol{\mu} - \bar{\mathbf{X}} = (\mu_1 - 6,28; \mu_2 - 6,24; \mu_3 - 6.30)^T$, gausime pasikliovimo elipsoidą $\boldsymbol{C}(\bar{\mathbf{X}}, \boldsymbol{S})$ tokio pavidalo:

$$10,587Z_1^2 + 7,945Z_2^2 + 6,765Z_3^2 - 14,310Z_1Z_2 - 9,942Z_1Z_3 + 4,906Z_2Z_3 \leq 0,011.$$

Matome, kad šis elipsoidas skiriasi nuo gauto **1.4.1** pavyzdyje tik tuo, kad dešinėje nelygybės pusėje skaičius 0,009 yra pakeistas 0,011. Natūralu, kad nežinant kovariacinės matricos, vidurkių vektoriaus to paties pasikliovimo lygmens pasikliovimo sritis yra didesnė, negu tuo atveju, kai kovariacinė matrica žinoma ir sutampa su įverčiu.

Antrame ir trečiame 2.1.1 lentelės stulpeliuose pateikti pasikliovimo intervalų rinkiniai (2.1.11) ir (2.1.13). Palyginti ketvirtame stulpelyje pateikiami pasikliovimo intervalai (2.1.10), sudaryti kiekvienai vidurkio koordinatei.

2.1.1 lentelė. Pasikliovimo intervalai

i	$(\underline{\mu}'_i; \overline{\mu}'_i)$	$(\underline{\mu}''_i; \overline{\mu}''_i)$	$(\underline{\mu}_i; \overline{\mu}_i)$
1	(6,218; 6,342)	(6,248; 6,312)	(6,251; 6,309)
2	(6,174; 6,306)	(6,206; 6,274)	(6,210; 6,270)
3	(6,249; 6,351)	(6,274; 6,326)	(6,276; 6,324)

Matome, kad pasikliovimo intervalai, sudaryti remiantis Bonferonio nelygybe, yra apytiksliai 1,95 karto trumpesni už tuos, kurie sudaryti remiantis pasikliautinąja sritimi (2.1.6).

2.1.3. Hipotezės dėl vidurkių vektoriaus reikšmės tikrinimas

Tarkime reikia patikrinti hipotezę $H : \boldsymbol{\mu} = \boldsymbol{\mu}_0$, kai alternatyva $\bar{H} : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$; čia $\boldsymbol{\mu}_0$ žinomas vektorius. Įrašykime į sąryšio (2.1.5) kairią pusę hipotetinę reikšmę $\boldsymbol{\mu}_0$. Gautoji statistika

$$F = \frac{n(n-k)}{k}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T \boldsymbol{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0) \tag{2.1.14}$$

turi centrinį Fišerio skirstinį su k ir $n-k$ laisvės laipsnių, jeigu tikrinamoji hipotezė H yra teisinga. Kai hipotezė H neteisinga, tai statistikos F skirstinys yra necentrinis Fišerio skirstinys su k ir $n-k$ laisvės laipsnių ir necentriškumo parametru

$$\lambda = n(\boldsymbol{\mu} - \boldsymbol{\mu}_0)^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \boldsymbol{\mu}_0).$$

Taigi hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai teisinga nelygybė

$$F > F_\alpha(k, n-k). \tag{2.1.15}$$

Kriterijaus galia išreiškiama necentrinio Fišerio skirstinio pasiskirstymo funkcija

$$\begin{aligned}\beta(\boldsymbol{\mu}) &= \mathbf{P}\{F > F_\alpha(k, n-k) | \boldsymbol{\mu}\} \\ &= \mathbf{P}\{F_{k,n-k;\lambda} > F_\alpha(k, n-k)\}.\end{aligned} \tag{2.1.16}$$

Vienmačiu atveju kriterijus yra ekvivalentus Stjudento kriterijui

$$|t| = \sqrt{n}|\bar{X} - \mu_0|/s > t_{\alpha/2}(n-1),$$

nes $t^2_{\alpha/2}(n-1) = F_\alpha(1, n-1)$, o $F = t^2$.

2.1.2 pavyzdys (**1.2.1** pavyzdžio tęsinys). Tarkime, kad **1.2.1** pavyzdyje buvo stebimas trimatis normalusis vektorius su nežinoma kovariacinė matrica $\boldsymbol{\Sigma}$. Tikrinsime hipotezę, kad vidurkių vektorius $\boldsymbol{\mu}$ lygus fiksuotam vektoriui $\boldsymbol{\mu}_0 = (6,25; 6,25; 6,25)^T$.

Pasinaudoję **1.2.1** pavyzdyje surastais parametrų įverčiais ir **1.4.1** pavyzdyje apskaičiuota matricos $\boldsymbol{S}^{-1}$ realizacija, randame statistikos (2.1.14) reikšmę

$$F = 270(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T \boldsymbol{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0) = 3,8238.$$

Kadangi P reikšmė $pv = \mathbf{P}\{F_{3;27} > 3,8238\} = 0,0210$, tai hipotezė atmetama kriterijumi, kurio reikšmingumo lygmuo viršija 0,0210. Matome, kad kai kovariacinė matrica nežinoma, P reikšmė yra gerokai didesnė (palyginkite su **1.4.2** pavyzdžiu).

2.1.2 teorema. *Kriterijus (2.1.15) yra ekvivalentus tikėtinumų santykio kriterijui.*

Įrodymas. Tikėtinumų santykis

$$\Lambda = \frac{\max_{\boldsymbol{\Sigma}} L(\boldsymbol{\mu}_0, \boldsymbol{\Sigma})}{\max_{\boldsymbol{\mu},\boldsymbol{\Sigma}} L(\boldsymbol{\mu}, \boldsymbol{\Sigma})} = \frac{L(\boldsymbol{\mu}_0, \hat{\boldsymbol{\Sigma}}_0)}{L(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})}; \tag{2.1.17}$$

čia $L(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tikėtinumo funkcija (1.2.2); $\hat{\boldsymbol{\Sigma}}_0$ kovariacinės matricos DT įvertinys, surastas esant sąlygai, kad vidurkis $\boldsymbol{\mu} = \boldsymbol{\mu}_0$; $\hat{\boldsymbol{\mu}}$ ir $\hat{\boldsymbol{\Sigma}}$ besąlyginiai parametrų $\boldsymbol{\mu}$ ir $\boldsymbol{\Sigma}$ didžiausiojo tikėtinumo įvertiniai.

Remdamiesi tikėtinumo funkcijos maksimumo išraiška (1.2.4) gauname

$$\Lambda = \left(\frac{|\hat{\boldsymbol{\Sigma}}|}{|\hat{\boldsymbol{\Sigma}}_0|}\right)^{n/2}, \quad \hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^T = \frac{1}{n}\boldsymbol{S},$$

$$\hat{\boldsymbol{\Sigma}}_0 = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T = \frac{1}{n}[\boldsymbol{S} + n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T];$$

$$\Lambda^{2/n} = \frac{|\boldsymbol{S}|}{|\boldsymbol{S} + n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T|}.$$

Remdamiesi determinantų savybe (žr. 1 priedą (8.2.19)), gauname

$$\begin{vmatrix} \boldsymbol{S} & -\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0) \\ \sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0) & 1 \end{vmatrix} =$$

$$1 \cdot |\boldsymbol{S} + n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T| = |\boldsymbol{S}|(1 + n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T \boldsymbol{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)).$$

Taigi

$$\Lambda^{2/n} = \frac{|\boldsymbol{S}|}{|\boldsymbol{S} + n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T|} = \frac{1}{1 + T^2/(n-1)},$$

$$T^2 = n(n-1)(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T \boldsymbol{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0).$$

Tikėtinumų santykis Λ yra monotoniškai mažėjantis Hotelingo statistikos T^2 atžvilgiu. Todėl tikėtinumų santykio kriterijus, apibrėžiamas nelygybe $\Lambda < \Lambda_{1-\alpha}$, yra ekvivalentus nelygybei $T^2 > T^2_\alpha$ arba nelygybei (2.1.15). ▲

2.2. Dviejų imčių vidurkių palyginimo hipotezės

Tarkime, dvi paprastosios imtys $\mathbf{X}_{11}, ..., \mathbf{X}_{1n_1}$ ir $\mathbf{X}_{21}, ..., \mathbf{X}_{2n_2}$ gautos stebint nepriklausomus normaliuosius vektorius $\mathbf{X}_1 \sim N_k(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$ ir $\mathbf{X}_2 \sim N_k(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$. Vidurkių vektoriai $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$ ir vienoda abiem imtims kovariacinė matrica $\boldsymbol{\Sigma}$ nežinomi, $|\boldsymbol{\Sigma}| > 0, n_1 > k, n_2 > k$.

Kadangi

$$\boldsymbol{S}_1 = \sum_{i=1}^{n_1} (\mathbf{X}_{1i} - \bar{\mathbf{X}}_1)(\mathbf{X}_{1i} - \bar{\mathbf{X}}_1)^T \sim W_k(n_1 - 1, \boldsymbol{\Sigma}), \quad \bar{\mathbf{X}}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} \mathbf{X}_{1i},$$

$$\boldsymbol{S}_2 = \sum_{i=1}^{n_2} (\mathbf{X}_{2i} - \bar{\mathbf{X}}_2)(\mathbf{X}_{2i} - \bar{\mathbf{X}}_2)^T \sim W_k(n_2 - 1, \boldsymbol{\Sigma}), \quad \bar{\mathbf{X}}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} \mathbf{X}_{2i},$$

tai pagal trečią Višarto skirstinio savybę

$$\boldsymbol{S} = \boldsymbol{S}_1 + \boldsymbol{S}_2 \sim W_k(n_1 + n_2 - 2, \boldsymbol{\Sigma}). \tag{2.2.1}$$

Nepriklausantis nuo $\boldsymbol{S}$ parametro $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$ įvertinys

$$\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2 \sim N_k \left(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2, \frac{n_1 + n_2}{n_1 n_2} \boldsymbol{\Sigma} \right). \tag{2.2.2}$$

Remdamiesi Hotelingo statistikos apibrėžimu, gauname

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} (n_1 + n_2 - 2)(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2)^T \boldsymbol{S}^{-1} (\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2).$$

Tada santykis

$$\frac{n_1 n_2 (n_1 + n_2 - 1 - k)}{k(n_1 + n_2)} (\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2)^T \boldsymbol{S}^{-1} (\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) \sim F(k, n_1 + n_2 - k - 1; \delta) \tag{2.2.3}$$

turi necentrinį Fišerio skirstinį su k ir $n_1 + n_2 - k - 1$ laisvės laipsnių ir necentriškumo parametru

$$\delta = \frac{n_1 n_2}{n_1 + n_2} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2).$$

Tarkime, reikia patikrinti hipotezę $H : \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \boldsymbol{\beta}_0$; čia $\boldsymbol{\beta}_0$ žinomas fiksuotas vektorius, kai alternatyva $\bar{H} : \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 \neq \boldsymbol{\beta}_0$. Jeigu hipotezė H teisinga, tai statistika

$$F = \frac{n_1 n_2 (n_1 + n_2 - 1 - k)}{k(n_1 + n_2)} (\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2 - \boldsymbol{\beta}_0)^T (\boldsymbol{S}_1 + \boldsymbol{S}_2)^{-1} (\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2 - \boldsymbol{\beta}_0) \sim$$

$$\sim F(k, n_1 + n_2 - k - 1) \tag{2.2.4}$$

turi centrinį Fišerio skirstinį su k ir n_1+n_2-k-2 laisvės laipsnių. Jeigu hipotezė neteisinga, tai statistika F turi necentrinį Fišerio skirstinį su tais pačiais laisvės laipsniais ir necentriškumo parametru

$$\lambda = \frac{n_1 n_2}{n_1+n_2}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2-\boldsymbol{\beta}_0)^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2-\boldsymbol{\beta}_0).$$

Hipotezė H atmetama reikšmingumo lygmens α kriterijumi, kai teisinga nelygybė

$$F > F_\alpha(k, n_1+n_2-k-1). \tag{2.2.5}$$

Kriterijaus galia išreiškiama necentrinio Fišerio skirstinio pasiskirstymo funkcija:

$$\beta(\lambda) = \mathbf{P}\{F_{k,n_1+n_2-k-1;\lambda} > F_\alpha(k, n_1+n_2-k-1)\}.$$

Matome, kad kriterijaus galia priklauso ne tik nuo nuokrypio $\boldsymbol{\mu}_1-\boldsymbol{\mu}_2-\boldsymbol{\beta}_0$ bet ir nuo kovariacinės matricos $\boldsymbol{\Sigma}$.

2.2.1 pastaba. Remiantis (2.2.3) analogiškai 2.1.1 ir 2.1.2 skyreliams galima sudaryti parametro $\boldsymbol{\mu}_1-\boldsymbol{\mu}_2$ pasikliovimo sritį arba jo koordinačių pasikliovimo intervalų rinkinius.

2.2.1 pavyzdys. Buvo pamatuoti tie patys parametrai kaip ir **1.2.1** pavyzdyje, 24 Panevėžio gamyklos kineskopų, pagamintų kitu laikotarpiu. Gauti statistiniai duomenys pateikiami **2.2.1** lentelėje.

2.2.1 lentelė. Statistiniai duomenys

i	X_{1i}	X_{2i}	X_{3i}	i	X_{1i}	X_{2i}	X_{3i}	i	X_{1i}	X_{2i}	X_{3i}
1	6,1	6,1	6,2	9	6,3	6,3	6,5	17	6,3	6,3	6,3
2	6,2	6,0	6,2	10	6,3	6,5	6,5	18	6,3	6,3	6,4
3	6,2	6,2	6,2	11	6,2	6,1	6,2	19	6,6	6,3	6,2
4	6,3	6,0	6,3	12	6,3	6,2	6,4	20	6,3	6,2	6,2
5	6,3	6,3	6,1	13	6,1	6,1	6,3	21	6,2	6,3	6,1
6	6,2	6,2	6,2	14	6,4	6,4	6,3	22	6,4	6,3	6,4
7	6,3	6,3	6,4	15	6,2	6,1	6,2	23	6,1	6,1	6,2
8	6,1	6,1	6,2	16	6,4	6,4	6,4	24	6,3	6,3	6,3

Tarę, kad **1.2.1** ir **2.2.1** lentelių duomenys yra nepriklausomų trimačių normaliųjų vektorių realizacijos, patikrinsime hipotezę, kad vidurkių vektorius nepakito (priimame prielaidą, kad kovariacinė matrica išliko nepakitusi).

Pagal **2.2.1** lentelės duomenis randame

$$\hat{\boldsymbol{\mu}}_2 = \bar{\mathbf{X}}_2 = \frac{1}{n_2}\sum_{i=1}^{n_2}\mathbf{X}_i = \frac{1}{24}\left\{\begin{pmatrix}6,1\\6,1\\6,2\end{pmatrix} + ... + \begin{pmatrix}6,3\\6,3\\6,3\end{pmatrix}\right\} = \begin{pmatrix}6,267\\6,225\\6,279\end{pmatrix}$$

ir apskaičiuojame matricos $\boldsymbol{S}$ realizaciją

$$\boldsymbol{S}_2 = \sum_{i=1}^{n}\mathbf{X}_i\mathbf{X}_i^T - n\bar{\mathbf{X}}\bar{\mathbf{X}}^T = \begin{pmatrix}6,1\\6,1\\6,2\end{pmatrix}(6,1;6,1;6,2) + ... + \begin{pmatrix}6,3\\6,3\\6,3\end{pmatrix}(6,3;6,3;6,3) -$$

$$24\begin{pmatrix}6,267\\6,225\\6,279\end{pmatrix}(6,267;6,225;6,279) = \begin{pmatrix}0,313 & 0,210 & 0,103\\0,210 & 0,385 & 0,163\\0,103 & 0,163 & 0,300\end{pmatrix}.$$

Pasinaudoję **1.2.1** pratime apskaičiuotomis aritmetinio vidurkio ir matricos $\boldsymbol{S}$ realizacijomis (pažymime $\bar{\mathbf{X}}_1$ ir $\boldsymbol{S}_1$), gauname

$$\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2 = \left\{ \begin{pmatrix} 6,28 \\ 6,24 \\ 6,30 \end{pmatrix} - \begin{pmatrix} 6,267 \\ 6,225 \\ 6,279 \end{pmatrix} \right\} = \begin{pmatrix} 0,013 \\ 0,015 \\ 0,021 \end{pmatrix};$$

$$\boldsymbol{S} = \boldsymbol{S}_1 + \boldsymbol{S}_2 = \begin{pmatrix} 0,661 & 0,474 & 0,263 \\ 0,474 & 0,777 & 0,233 \\ 0,263 & 0,233 & 0,540 \end{pmatrix}; \quad \boldsymbol{S}^{-1} = \begin{pmatrix} 2,935 & -1,564 & -0,754 \\ -1,564 & 2,312 & -0,236 \\ 0,754 & -0,236 & 2,321 \end{pmatrix}.$$

Pagaliau randame statistikos F realizaciją (vektorius $\boldsymbol{\beta}_0 = \mathbf{0}$):

$$F = \frac{n_1 n_2 (n_1 + n_2 - 1 - k)}{k(n_1 + n_2)} (\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2)^T \boldsymbol{S}^{-1} (\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) = 0,101.$$

Kadangi P reikšmė $pv = \mathbf{P}\{F_{3;50} > 0,101\} = 0,959$, tai atmesti vidurkių vektorių lygybės hipotezę nėra pagrindo.

2.3. Kelių imčių vidurkių palyginimo hipotezės

Turime m paprastųjų imčių $\mathbf{X}_{11}, ..., \mathbf{X}_{1n_1}; ...; \mathbf{X}_{m1}, ..., \mathbf{X}_{mn_m}$, gautų stebint nepriklausomus normaliuosius vektorius $\mathbf{X}_1 \sim N_k(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}), ..., \mathbf{X}_m \sim N_k(\boldsymbol{\mu}_m, \boldsymbol{\Sigma})$. Vidurkių vektoriai $\boldsymbol{\mu}_1, ..., \boldsymbol{\mu}_m$ ir vienoda visoms imtims kovariacinė matrica $\boldsymbol{\Sigma}$ nežinomi, $|\boldsymbol{\Sigma}| > 0, n_i > k, i = 1, ..., m$.

Reikia patikrinti hipotezę $H : \beta_1 \boldsymbol{\mu}_1 + ... + \beta_m \boldsymbol{\mu}_m = \boldsymbol{\mu}_0$; čia $\beta_1, ..., \beta_m$ žinomos konstantos, o $\boldsymbol{\mu}_0$ – žinomas vektorius. Matricos

$$\boldsymbol{S}_i = \sum_{j=1}^{n_i} (\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)(\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)^T \sim W_k(n_i - 1, \boldsymbol{\Sigma}), \quad i = 1, ..., m,$$

ir vektoriai

$$\bar{\mathbf{X}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{X}_{ij}, \quad i = 1, ..., m,$$

yra nepriklausomi. Pagal trečią Višarto skirstinio savybę suma

$$\boldsymbol{S} = \boldsymbol{S}_1 + ... + \boldsymbol{S}_m \sim W_k(n - m, \boldsymbol{\Sigma}), \quad n = \sum_{i=1}^{m} n_i. \tag{2.3.1}$$

Nepriklausomas nuo $\boldsymbol{S}$ parametro $\boldsymbol{\theta} = \beta_1 \boldsymbol{\mu}_1 + ... + \beta_m \boldsymbol{\mu}_m$ įvertinys

$$\hat{\boldsymbol{\theta}} = \sum_{i=1}^{m} \beta_i \bar{\mathbf{X}}_i \sim N_k \left(\sum_{i=1}^{m} \beta_i \boldsymbol{\mu}_i, \ (\sum_{i=1}^{m} \frac{\beta_i^2}{n_i}) \boldsymbol{\Sigma} \right). \tag{2.3.2}$$

Remdamiesi Hotelingo statistikos apibrėžimu, gauname

$$T^2 = c(n - m)(\hat{\boldsymbol{\theta}} - \boldsymbol{\mu}_0)^T \boldsymbol{S}^{-1} (\hat{\boldsymbol{\theta}} - \boldsymbol{\mu}_0), \quad c = \left(\sum_{i=1}^{m} \frac{\beta_i^2}{n_i} \right)^{-1}.$$

Jeigu hipotezė teisinga, tai statistika

$$F = \frac{c(n-m-k-1)}{k}(\hat{\boldsymbol{\theta}} - \boldsymbol{\mu}_0)^T \boldsymbol{S}^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\mu}_0) \sim F(k, n-m-k+1)$$

turi centrinį Fišerio skirstinį su k ir $n-m-k+1$ laisvės laipsnių. Jeigu hipotezė neteisinga, tai statistika F turi necentrinį Fišerio skirstinį su k ir $n-m-k+1$ laisvės laipsnių ir necentriškumo parametru $\delta = c(\boldsymbol{\theta} - \boldsymbol{\mu}_0)^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta} - \boldsymbol{\mu}_0)$.

Hipotezė atmetama reikšmingumo α lygmens kriterijumi, kai teisinga nelygybė

$$F > F_\alpha(k, n-m-k+1). \tag{2.3.3}$$

Kriterijaus galia išreiškiama necentrinio Fišerio skirstinio pasiskirstymo funkcija.

2.4. Simetriškumo hipotezė

Tarkime, $\mathbf{X}_1, ..., \mathbf{X}_n$ yra paprastoji imtis a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $|\boldsymbol{\Sigma}| > 0, n > k$. Reikia patikrinti hipotezę $H : \mu_1 = ... = \mu_k$, kad vidurkių vektoriaus $\boldsymbol{\mu} = (\mu_1, ..., \mu_k)^T$ koordinatės yra lygios. Tegu $\boldsymbol{C} = [c_{ij}]_{(k-1)\times k}$ yra rango $k-1$ matrica tokia, kad

$$\boldsymbol{C}\boldsymbol{\varepsilon} = \mathbf{0}, \quad \boldsymbol{\varepsilon} = (1, ..., 1)^T. \tag{2.4.1}$$

Atlikime a. v. $\mathbf{X}_1, ..., \mathbf{X}_n$ transformaciją naudodami matricą $\boldsymbol{C}$:

$$\boldsymbol{Z}_j = \boldsymbol{C}\mathbf{X}_j, \quad j = 1, ..., n. \tag{2.4.2}$$

Vektorių $\boldsymbol{Z}_j$ vidurkiai ir kovariacinė matrica yra

$$\mathbf{E}(\boldsymbol{Z}_j) = \boldsymbol{C}\boldsymbol{\mu}, \quad \boldsymbol{V}(\boldsymbol{Z}_j) = \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}^T. \tag{2.4.3}$$

Vektorių $\boldsymbol{Z}_j$ terminais hipotezė H ekvivalenti tvirtinimui $H_{\boldsymbol{C}} : \boldsymbol{C}\boldsymbol{\mu} = \mathbf{0}$. Apibrėžkime statistikas

$$\bar{\boldsymbol{Z}} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{Z}_i \sim N_{k-1}(\boldsymbol{C}\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}^T),$$

$$\boldsymbol{S} = \sum_{i=1}^n (\boldsymbol{Z}_i - \bar{\boldsymbol{Z}})(\boldsymbol{Z}_i - \bar{\boldsymbol{Z}})^T \sim W_{k-1}(n-1, \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}^T). \tag{2.4.4}$$

Remdamiesi šiomis statistikomis sudarome Hotelingio T^2 statistiką

$$T^2 = n(n-1)\bar{\boldsymbol{Z}}^T \boldsymbol{S}^{-1}\bar{\boldsymbol{Z}}$$

ir statistiką

$$F = \frac{(n-k+1)n}{k-1}\bar{\boldsymbol{Z}}^T \boldsymbol{S}^{-1}\bar{\boldsymbol{Z}}, \tag{2.4.5}$$

kuri, kai hipotezė $H_{\boldsymbol{C}}: \boldsymbol{C}\boldsymbol{\mu} = \mathbf{0}$ teisinga, turi Fišerio skirstinį su $k-1$ ir $n-k+1$ laisvės laipsniu. Hipotezė $H_{\boldsymbol{C}}$ atmetama reikšmingumo lygmens α kriterijumi, kai teisinga nelygybė

$$F > F_\alpha(k-1, n-k+1). \tag{2.4.6}$$

Dar reikia įsitikinti, kad surastas kriterijus nepriklauso nuo matricos, tenkinančios (2.4.1) sąlygą, parinkimo.

Tegu $\boldsymbol{B} = [b_{ij}]_{(k-1)\times(k-1)}$ matrica, kurios determinantas $|\boldsymbol{B}| \neq 0$. Tada matrica $\boldsymbol{C}^* = \boldsymbol{B}\boldsymbol{C}$ taip pat tenkina (2.4.1) sąlygą

$$\boldsymbol{C}^*\boldsymbol{\varepsilon} = \boldsymbol{B}(\boldsymbol{C}\boldsymbol{\varepsilon}) = \mathbf{0}.$$

Atlikime a. v. $\mathbf{X}_1, ..., \mathbf{X}_n$ transformaciją naudodami matricą $\boldsymbol{C}^*$:

$$\boldsymbol{Z}_j^* = \boldsymbol{C}^*\mathbf{X}_j = \boldsymbol{B}\boldsymbol{C}\mathbf{X}_j = \boldsymbol{B}\boldsymbol{Z}_j, \quad j = 1, ..., n;$$

$$\mathbf{E}(\boldsymbol{Z}_j^*) = \boldsymbol{B}\boldsymbol{C}\boldsymbol{\mu}, \quad \boldsymbol{V}(\boldsymbol{Z}_j^*) = \boldsymbol{B}\boldsymbol{V}(\boldsymbol{Z}_j)\boldsymbol{B}^T = \boldsymbol{B}(\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}^T)\boldsymbol{B}^T.$$

Apibrėžkime a. v. $\boldsymbol{Z}_1^*, ..., \boldsymbol{Z}_n^*$ statistikas analogiškas (2.2.10)

$$\bar{\boldsymbol{Z}}^* \sim N_k(\boldsymbol{B}\boldsymbol{C}\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{B}(\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}^T)\boldsymbol{B}^T).$$

$$\boldsymbol{S}^* = \sum_{i=1}^n (\boldsymbol{B}\boldsymbol{Z}_i - \boldsymbol{B}\bar{\boldsymbol{Z}})(\boldsymbol{B}\boldsymbol{Z}_i - \boldsymbol{B}\bar{\boldsymbol{Z}})^T = \boldsymbol{B}\boldsymbol{S}\boldsymbol{B}^T$$

ir sudarykime Hotelingo statistiką

$$T^{*2} = n(n-1)\bar{\boldsymbol{Z}}^{*T}\boldsymbol{S}^{*-1}\bar{\boldsymbol{Z}}^* = n(n-1)\bar{\boldsymbol{Z}}^T\boldsymbol{B}^T(\boldsymbol{B}\boldsymbol{S}\boldsymbol{B}^T)^{-1}\boldsymbol{B}\bar{\boldsymbol{Z}}$$

$$= n(n-1)\bar{\boldsymbol{Z}}^T\boldsymbol{B}^T(\boldsymbol{B}^T)^{-1}\boldsymbol{S}^{-1}\boldsymbol{B}^{-1}\boldsymbol{B}\bar{\boldsymbol{Z}} = n(n-1)\bar{\boldsymbol{Z}}^T\boldsymbol{S}^{-1}\bar{\boldsymbol{Z}} = T^2.$$

Matome, kad Hotelingo T^2 statistika yra invariantiška atžvilgiu matricų, tenkinančių (2.4.1) sąlygą. Todėl matricą $\boldsymbol{C}$ galima parinkti, pavyzdžiui, tokio pavidalo: i-oji eilutė yra $(0, ..., 0, 1, 0, ..., -1)$; čia 1 parašytas i-ojoje vietoje, $i = 1, ..., k-1$.

2.2.2 pavyzdys. (**2.2.1** pavyzdžio tęsinys.) Remdamiesi **2.2.1** lentelės duomenimis patikrinsime hipotezę $H: \mu_1 = \mu_2 = \mu_3$, kad vidurkių vektoriaus $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)^T$ koordinatės yra lygios. Parinkę matricą $\boldsymbol{C}$, kurios pirmoji eilutė yra $(1\ 0\ -1)$, o antroji – $(0\ 1\ -1)$, atliekame **2.2.1** lentelės duomenų transformaciją $\boldsymbol{Z}_i = \boldsymbol{C}\mathbf{X}_i$. Transformuoti duomenys pateikti **2.2.2** lentelėje.

2.2.2 lentelė. Transformuoti statistiniai duomenys

i	Z_{1i}	Z_{2i}	i	Z_{1i}	Z_{2i}	i	Z_{1i}	Z_{2i}
1	-0,1	-0,1	9	-0,2	-0,2	17	0,0	0,0
2	0,0	-0,2	10	-0,2	0,0	18	-0,1	-0,1
3	0,0	0,0	11	0,0	-0,1	19	0,4	0,1
4	0,0	-0,3	12	-0,1	-0,2	20	0,1	0,0
5	0,2	0,2	13	-0,2	-0,2	21	0,1	0,2
6	0,0	0,0	14	0,1	0,1	22	0,0	-0,1
7	-0,1	-0,1	15	0,0	-0,1	23	-0,1	-0,1
8	-0,1	-0,1	16	0,0	0,0	24	0,0	0,0

Pagal šios lentelės duomenis apskaičiuojame

$$\bar{\boldsymbol{Z}} = \begin{pmatrix} -0,0125 \\ -0,0542 \end{pmatrix}; \quad \boldsymbol{S} = \begin{pmatrix} 0,406 & 0,244 \\ 0,244 & 0,360 \end{pmatrix}; \quad \boldsymbol{S}^{-1} = \begin{pmatrix} 4,149 & -2,813 \\ -2,813 & 4,688 \end{pmatrix}$$

ir randame statistikos F realizaciją:

$$F = \frac{(n-k+1)n}{k-1} \bar{\boldsymbol{Z}}^T \boldsymbol{S}^{-1} \bar{\boldsymbol{Z}} = 2,797.$$

Kadangi P reikšmė $pv = \mathbf{P}\{F_{2;22} > 2,797\} = 0,083$, tai hipotezė atmetama, jei kriterijaus reikšmingumo lygmuo viršija 0,083.

2.5. Vidurkių palyginimo hipotezės, kai kovariacinės matricos skirtingos

Kaip žinome (žr. 1 dalį, 5.7.2 skyrelį), vienmačiu atveju tikrinant hipotezę apie normaliųjų skirstinių vidurkių lygybę, kai apie dispersijas nieko nežinome, TG ar TGN kriterijai neegzistuoja (Berenso ir Fišerio problema). Natūralu, kad analogiškas efektas matomas ir daugiamačiu atveju.

Didelės imtys. Tarkime, $\mathbf{X}_{11}, ..., \mathbf{X}_{1n_1}$ ir $\mathbf{X}_{21}, ..., \mathbf{X}_{2n_2}$ yra dvi paprastosios imtys, gautos stebint normaliuosius a. v. $\mathbf{X}_1 \sim N_k(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ ir $\mathbf{X}_2 \sim N_k(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$.

Remiantis daugiamačio normaliojo skirstinio savybėmis (3 priedas (10.0.4), (10.0.7))) galima tvirtinti, kad

$$(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2 - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2))^T \left(\frac{1}{n_1} \boldsymbol{\Sigma}_1 + \frac{1}{n_2} \boldsymbol{\Sigma}_2 \right)^{-1} (\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2 - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)) \sim \chi^2_k.$$

Jeigu n_1 ir n_2 yra pakankamai dideli, kai hipotezė $H : \mu_1 - \mu_2 = \boldsymbol{\beta}_0$ teisinga, statistikos

$$\tilde{U}^2 = (\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2 - \boldsymbol{\beta}_0)^T \left(\frac{1}{n_1} \hat{\boldsymbol{\Sigma}}_1 + \frac{1}{n_2} \hat{\boldsymbol{\Sigma}}_2 \right)^{-1} (\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2 - \boldsymbol{\beta}_0)$$

skirstinį galima aproksimuoti χ^2 skirstiniu su k laisvės laipsnių. Gauname asimptotinį dviejų vidurkių vektorių palyginimo kriterijų: hipotezė H atmetama apytiksliu reikšmingumo lygmens α kriterijumi, kai teisinga nelygybė

$$\tilde{U}^2 > \chi^2_\alpha(k). \tag{2.5.1}$$

Jeigu imčių didumai nepakankami, tai vidurkių palyginimo kriterijų galima sudaryti imant dviejų imčių elementų skirtumus. Analogiškas metodas buvo naudojamas vienmačiu atveju sudarant Stjudento priklausomų imčių kriterijų.

Vienodo didumo imtys. Tarkime, $\mathbf{X}_{11}, ..., \mathbf{X}_{1n_1}$ ir $\mathbf{X}_{21}, ..., \mathbf{X}_{2n_2}$, $n_2 = n_1$, yra dvi paprastosios imtys, gautos stebint normaliuosius a. v. $\mathbf{X}_1 \sim N_k(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ ir $\mathbf{X}_2 \sim N_k(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$. Norėdami gauti vienmačio Stjudento kriterijaus priklausomoms imtims analogą, atsisakykime reikalavimo dėl a. v. $\mathbf{X}_1$ ir $\mathbf{X}_2$ nepriklausomumo.

Skirtumai $\boldsymbol{Y}_1 = \mathbf{X}_{11} - \mathbf{X}_{21}, ..., \boldsymbol{Y}_{n_1} = \mathbf{X}_{1n_1} - \mathbf{X}_{2n_1}$ yra paprastoji didumo n_1 imtis, gauta stebint a. v. $\boldsymbol{Y} = \mathbf{X}_1 - \mathbf{X}_2 \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\mu} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$, $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2 + 2\boldsymbol{\Sigma}_{12}$; čia $\boldsymbol{\Sigma}_{12} = \mathbf{Cov}\,(\mathbf{X}_1, \mathbf{X}_2)$. Jeigu imtys nepriklausomos, tai gautume atvejį, kai $\boldsymbol{\Sigma}_{12} = \mathbf{0}$.

Tarkime, kad $n_1 > k$, $|\boldsymbol{\Sigma}_1| > 0, |\boldsymbol{\Sigma}_2| > 0$. Vietoje hipotezės $H : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ tikrinkime hipotezę, kad $\mathbf{E}(\boldsymbol{Y}) = \boldsymbol{\mu} = \mathbf{0}$, kai kovariacinė matrica $\boldsymbol{\Sigma}$ yra nežinoma. Tokį uždavinį sprendėme 2.2.1 skyrelyje. Sudarome statistiką

$$T^2 = n_1(n_1 - 1)\bar{\boldsymbol{Y}}^T \boldsymbol{S}^{-1} \bar{\boldsymbol{Y}},$$

$$\bar{\boldsymbol{Y}} = \frac{1}{n_1} \sum_{i=1}^{n_1} \boldsymbol{Y}_i, \quad \boldsymbol{S} = \sum_{i=1}^{n_1} (\boldsymbol{Y}_i - \bar{\boldsymbol{Y}})(\boldsymbol{Y}_i - \bar{\boldsymbol{Y}})^T.$$

Tada statistika

$$F = \frac{n_1 - k}{k} \frac{T^2}{n_1 - 1},$$

kai $\boldsymbol{\mu} = \mathbf{0}$, turi Fišerio skirstinį su k ir $n_1 - k$ laisvės laipsnių. Hipotezė atmetama reikšmingumo α kriterijumi, kai teisinga nelygybė

$$F > F_\alpha(k, n_1 - k). \tag{2.5.2}$$

Suprantama, jeigu yra pagrindo tvirtinti, kad, kai imtys nepriklausomos ir $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$, tai geriau naudoti kriterijų (2.2.2) iš skyrelio 2.2.2, kai Fišerio statistikos vardiklio laisvės laipsnių skaičius yra $2n_1 - k - 1$. Taigi, naudojant šio skyrelio kriterijų prarandama $n_1 - 1$ laisvės laipsnis.

Skirtingo didumo imtys. Tarkime, $\mathbf{X}_{11}, ..., \mathbf{X}_{1n_1}$ ir $\mathbf{X}_{21}, ..., \mathbf{X}_{2n_2}$, $n_1 \leq n_2$, yra dvi paprastosios imtys, gautos stebint nepriklausomus normaliuosius a. v. $\mathbf{X}_1 \sim N_k(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ ir $\mathbf{X}_2 \sim N_k(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$; $|\boldsymbol{\Sigma}_1| > 0, |\boldsymbol{\Sigma}_2| > 0, n_1 > k$.

Jeigu skirtumas $n_2 - n_1$ lyginant su n_1 yra mažas, tai galime tiesiog atmesti stebėjimus $\mathbf{X}_{2,n_1+1}, ..., \mathbf{X}_{2,n_2}$ ir gauti dvi vienodo didumo n_1 imtis. Tačiau kriterijų galima šiek tiek patikslinti iš dalies panaudojant ir stebėjimus $\mathbf{X}_{2,n_1+1}, ...,$ $\mathbf{X}_{2,n_2}$. Pažymėkime

$$\boldsymbol{Y}_j = \mathbf{X}_{1j} - \sqrt{\frac{n_1}{n_2}} \mathbf{X}_{2j} + \frac{1}{\sqrt{n_1 n_2}} \sum_{l=1}^{n_1} \mathbf{X}_{2l} - \frac{1}{n_2} \sum_{r=1}^{n_2} \mathbf{X}_{2r},$$

$$j = 1, 2, ..., n_1.$$

Atsitiktiniai vektoriai $\boldsymbol{Y}_j$ turi vienodus vidurkius

$$\mathbf{E}(\boldsymbol{Y}_j) = \boldsymbol{\mu}_1 - \sqrt{\frac{n_1}{n_2}} \boldsymbol{\mu}_2 + \frac{n_1}{\sqrt{n_1 n_2}} \boldsymbol{\mu}_2 - \frac{n_2}{n_2} \boldsymbol{\mu}_2 = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$$

ir vienodas kovariacines matricas

$$\boldsymbol{V}(\boldsymbol{Y}_j) = \boldsymbol{\Sigma}_1 + \frac{n_1}{n_2} \boldsymbol{\Sigma}_2.$$

Tai šiek tiek mažiau, negu skirtumo $\mathbf{X}_{1j} - \mathbf{X}_{2j}$ kovariacinė matrica $\boldsymbol{V}(\mathbf{X}_{1j} - \mathbf{X}_{2j}) = \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2$. Kadangi $\mathbf{Cov}\,(\boldsymbol{Y}_i, \boldsymbol{Y}_j) = \mathbf{0}$, kai $i \neq j$, tai, tikrinant hipotezę $H : \mathbf{E}(\boldsymbol{Y}_j) = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \mathbf{0}$, pritaikoma 2.2.1 skyrelio metodika.

Hipotezė atmetama reikšmingumo α kriterijumi, kai teisinga nelygybė

$$F = \frac{(n_1 - k)n_1}{k} \bar{\boldsymbol{Y}}^T \boldsymbol{S}^{-1} \bar{\boldsymbol{Y}} > F_\alpha(k, n_1 - k), \tag{2.5.3}$$

$$\bar{\boldsymbol{Y}} = \frac{1}{n_1} \sum_{i=1}^{n_1} \boldsymbol{Y}_i, \quad \boldsymbol{S} = \sum_{i=1}^{n_1} (\boldsymbol{Y}_i - \bar{\boldsymbol{Y}})(\boldsymbol{Y}_i - \bar{\boldsymbol{Y}})^T.$$

Pažymėjus

$$\boldsymbol{U}_j = \mathbf{X}_{1j} - \sqrt{\frac{n_1}{n_2}} \mathbf{X}_{2j}, \quad j = 1, ..., n_1,$$

matricą $\boldsymbol{S}$ galima užrašyti tokiu pavidalu

$$\boldsymbol{S} = \sum_{i=1}^{n} (\boldsymbol{U}_i - \bar{\boldsymbol{U}})(\boldsymbol{U}_i - \bar{\boldsymbol{U}})^T, \quad \bar{\boldsymbol{U}} = \frac{1}{n_1} \sum_{i=1}^{n_1} \boldsymbol{U}_i. \tag{2.5.4}$$

Keletas skirtingo didumo imčių. Tarkime, kad paprastosios imtys $\mathbf{X}_{11}, ..., \mathbf{X}_{1n_1}; ...; \mathbf{X}_{m1}, ..., \mathbf{X}_{mn_m}$ yra gautos stebint nepriklausomus normaliuosius a. v. $\mathbf{X}_1 \sim N_k(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$; ...; $\mathbf{X}_m \sim N_k(\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m)$, $|\boldsymbol{\Sigma}_i| > 0, i = 1, ..., m$. Nemažinant bendrumo galima tarti, kad $n_1 \leq n_j, j = 2, ..., m,, n_1 > k$.

Reikia patikrinti hipotezę $H : \theta = \sum_{i=1}^m \beta_i \boldsymbol{\mu}_i = \boldsymbol{\mu}_0$; čia $\beta_1, ..., \beta_m$ žinomos konstantos, o $\boldsymbol{\mu}_0$ – žinomas vektorius.

Apibrėžkime atsitiktinius vektorius

$$\boldsymbol{Y}_j = \beta_1 \mathbf{X}_{1j} + \sum_{i=2}^m \beta_i \sqrt{\frac{n_1}{n_i}} (\mathbf{X}_{ij} - \frac{1}{n_1} \sum_{l=1}^{n_1} \mathbf{X}_{il} + \frac{1}{\sqrt{n_1 n_i}} \sum_{r=1}^{n_i} \mathbf{X}_{ir}),$$

$$j = 1, ..., n_1.$$

Šie vektoriai turi vienodus vidurkius

$$\mathbf{E}(\boldsymbol{Y}_j) = \beta_1 \boldsymbol{\mu}_1 + \sum_{i=2}^m \beta_i \sqrt{\frac{n_1}{n_i}} (\boldsymbol{\mu}_i - \frac{n_1}{n_1} \boldsymbol{\mu}_i + \frac{n_i}{\sqrt{n_1 n_i}} \boldsymbol{\mu}_i) = \sum_{i=1}^m \beta_i \boldsymbol{\mu}_i = \theta$$

ir vienodas kovariacines matricas

$$\boldsymbol{V}(\boldsymbol{Y}_j) = \sum_{j=1}^m \beta_i^2 \frac{n_1}{n_i} \boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}.$$

Nesunku patikrinti, kad $\mathbf{Cov}\,(\boldsymbol{Y}_i, \boldsymbol{Y}_j) = \mathbf{0}$, kai $i \neq j$. Todėl a. v. $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_{n_1}$ yra pritaikoma 2.2.1 skyrelio metodika. Apibrėžiame T^2 statistiką

$$T^2 = n_1(n_1 - 1)(\bar{\boldsymbol{Y}} - \boldsymbol{\mu}_0)^T \boldsymbol{S}^{-1} (\bar{\boldsymbol{Y}} - \boldsymbol{\mu}_0),$$

$$\bar{\boldsymbol{Y}} = \frac{1}{n_1}\sum_{i=1}^{n_1}\boldsymbol{Y}_i, \quad \boldsymbol{S} = \sum_{i=1}^{n_1}(\boldsymbol{Y}_i - \bar{\boldsymbol{Y}})(\boldsymbol{Y}_i - \bar{\boldsymbol{Y}})^T.$$

Hipotezė atmetama reikšmingumo α kriterijumi, kai teisinga nelygybė

$$F = \frac{(n_1 - k)n_1}{k(n_1 - 1)}T^2 > F_\alpha(k, n_1 - k). \tag{2.5.5}$$

Apibrėžus atsitiktinius vektorius

$$\boldsymbol{U}_j = \sum_{i=1}^{m}\beta_i\sqrt{\frac{n_1}{n_i}}\mathbf{X}_{ij}, \quad j = 1, ..., n_1,$$

matricą $\boldsymbol{S}$ galima apskaičiuoti taip:

$$\boldsymbol{S} = \sum_{i=1}^{n_1}(\boldsymbol{U}_i - \bar{\boldsymbol{U}})(\boldsymbol{U}_i - \bar{\boldsymbol{U}})^T, \quad \bar{\boldsymbol{U}} = \frac{1}{n_1}\sum_{i=1}^{n_1}\boldsymbol{U}_i.$$

2.5.1 pavyzdys. (**2.2.1** pavyzdžio tęsinys.) Tų pačių 24 kineskopų trijų spindulių srovės stiprumai buvo pamatuoti kitos technologinės operacijos metu. Gauti duomenys pateikti **2.5.1** lentelėje.

2.5.1 lentelė. Statistiniai duomenys

i	X_{1i}	X_{2i}	X_{3i}	i	X_{1i}	X_{2i}	X_{3i}	i	X_{1i}	X_{2i}	X_{3i}
1	6,1	6,2	6,1	9	6,3	6,3	6,1	17	6,2	6,4	6,2
2	6,2	6,1	6,1	10	6,2	6,6	6,5	18	6,3	6,4	6,3
3	6,1	6,2	6,1	11	6,2	6,2	6,1	19	6,0	6,3	6,2
4	5,8	6,3	6,1	12	6,3	6,3	6,2	20	6,1	6,2	6,2
5	6,4	6,3	6,2	13	6,1	6,2	6,2	21	6,1	6,0	6,2
6	6,3	6,1	6,3	14	6,4	6,6	6,3	22	6,4	6,5	6,4
7	6,3	6,3	6,2	15	6,2	6,2	6,2	23	6,2	6,2	6,1
8	5,9	6,0	6,1	16	6,4	6,5	6,3	24	6,3	6,4	6,2

Reikia patikrinti hipotezę, kad vidurkių vektorius nepakito.

Jeigu tartume, kad **2.2.1** lentelėje yra a. v. $\mathbf{X}_1 \sim N_3(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ paprastosios imties realizacijos, o lentelėje **2.5.1** – a. v. $\mathbf{X}_2 \sim N_3(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$ paprastosios imties realizacijos, tai reikėtų pripažinti, kad a. v. $\mathbf{X}_1$ ir $\mathbf{X}_2$ yra galbūt priklausomi (matuojami tų pačių kineskopų parametrai). Sprendžiant tokį uždavinį vienmačiu atveju buvo naudojamas Stjudento priklausomų imčių kriterijus. Daugiamačiu atveju taip pat reikėtų naudoti šiame skyrelyje pateiktą tokio kriterijaus analogą.

Randameme skirtumus $\boldsymbol{Y}_i = \mathbf{X}_{1i} - \mathbf{X}_{2i}, \quad i = 1, ..., 24$, jie pateikti **2.5.2** lentelėje.

2.5.2 lentelė. Skirtumai

i	Y_{1i}	Y_{2i}	Y_{3i}	i	Y_{1i}	Y_{2i}	Y_{3i}	i	Y_{1i}	Y_{2i}	Y_{3i}
1	0,0	-0,1	0,1	9	0,0	0,0	0,4	17	0,1	-0,1	0,1
2	0,0	-0,1	0,1	10	0,1	-0,1	0,0	18	0,0	-0,1	0,1
3	0,1	0,0	0,1	11	0,0	-0,1	0,1	19	0,6	0,0	0,0
4	0,5	-0,3	0,2	12	0,0	-0,1	0,2	20	0,2	0,0	0,0
5	-0,1	0,0	-0,1	13	0,0	-0,1	0,1	21	0,1	0,3	-0,1
6	-0,1	0,1	-0,1	14	0,0	-0,2	0,0	22	0,0	-0,2	0,0
7	0,0	0,0	0,2	15	0,0	-0,1	0,0	23	-0,1	-0,1	0,1
8	0,2	0,1	0,1	16	0,0	-0,1	0,1	24	0,0	-0,1	0,1

Naudodami šiuos duomenis tikriname hipotezę, kad a. v. $\boldsymbol{Y}$ vidurkis $\mathbf{E}(\boldsymbol{Y}) = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$ lygus nuliniam vektoriui. Tokias hipotezes tikrinome 2.1 skyrelyje. Reikia pažymėti, kad nereikia daryti prielaidos dėl pradinių vektorių $\mathbf{X}_1$ ir $\mathbf{X}_2$ kovariacijų matricų $\boldsymbol{\Sigma}_1$ ir $\boldsymbol{\Sigma}_2$ lygybės.

Pagal **2.5.2** lentelės duomenis apskaičiuojame

$$\bar{\boldsymbol{Y}} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{Y}_i = \begin{pmatrix} 0,0667 \\ -0,0583 \\ 0,0750 \end{pmatrix};$$

$$\boldsymbol{S} = \sum_{i=1}^{n}\boldsymbol{Y}_i\boldsymbol{Y}_i^T - n\bar{\boldsymbol{Y}}\bar{\boldsymbol{Y}}^T = \begin{pmatrix} 0,6533 & -0,0267 & 0,0200 \\ -0,0267 & 0,3183 & -0,0950 \\ 0,0200 & -0,0950 & 0,2850 \end{pmatrix};$$

$$\boldsymbol{S}^{-1} = \begin{pmatrix} 1,5373 & 0,1074 & -0,0721 \\ 0,1074 & 3,4963 & 1,1579 \\ -0,0721 & 1,1579 & 3,8998 \end{pmatrix}$$

ir randame statistikos F realizaciją:

$$F = \frac{(n-k)n}{k}\bar{\boldsymbol{Y}}^T\boldsymbol{S}^{-1}\bar{\boldsymbol{Y}} = 4,8681.$$

Kadangi P reikšmė $pv = \mathbf{P}\{F_{3;21} > 4,8681\} = 0,010$, tai hipotezė atmetama, jei kriterijaus reikšmingumo lygmuo viršija 0,01.

2.6. Pratimai

2.1. Tegu $\mathbf{X}_1, ..., \mathbf{X}_n$ yra nepriklausomi a. v. $\mathbf{X}_j \sim N_k(\boldsymbol{\mu} + \boldsymbol{\beta}(z_j - \bar{z}), \boldsymbol{\Sigma})$, $j = 1, ..., n$, $\bar{z} = \sum_j z_j/n$. Įrodykite, kad

$$T^2 = \sum_j (z_j - \bar{z})^2\boldsymbol{b}^T\boldsymbol{S}^{-1}\boldsymbol{b}$$

turi Hotelingo skirstinį su $n-2$ laisvės laipsniais. Čia

$$\boldsymbol{b} = \frac{\sum_j \mathbf{X}_j(z_j - \bar{z})}{\sum_j (z_j - \bar{z})^2}, \quad \boldsymbol{S} = \sum_j (\mathbf{X}_j - \bar{\mathbf{X}} - \boldsymbol{b}(z_j - \bar{z}))(\mathbf{X}_j - \bar{\mathbf{X}} - \boldsymbol{b}(z_j - \bar{z}))^T.$$

2.2. Tegu $\mathbf{X}_1, ..., \mathbf{X}_n$ yra paprastoji imtis a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, pagal kurią gautos statistikos $\bar{\mathbf{X}}$ ir $\boldsymbol{S}$. Raskite tolesnio nepriklausomo stebėjimo $\mathbf{X}_{n+1}$ pasikliovimo sritį.

2.3. Įrodykite, kad Hotelingo statistika $T^2 = n(n-1)\bar{\mathbf{X}}\boldsymbol{S}^{-1}\bar{\mathbf{X}}$, kai hipotezė $H : \boldsymbol{\mu} = \mathbf{0}$ teisinga, asimptotiškai $n \to \infty$ turi χ^2 skirstinį su k laisvės laipsnių.

2.4. Įrodykite, kad Hotelingo statistika $T^2 = n(n-1)\bar{\mathbf{X}}\boldsymbol{S}^{-1}\bar{\mathbf{X}}$ nepakis vidurkių vektorių $\boldsymbol{\mu} = (\mu_1, ..., \mu_k)^T$ pakeitus vektoriumi $(\lambda, 0, ..., 0)$, $\lambda^2 = \boldsymbol{\mu}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$, o $\boldsymbol{\Sigma}$ pakeitus vienetine matrica $\boldsymbol{I}$.

2.5. Įrodykite, kad tikėtinumų santykio kriterijus tikrinant hipotezę dėl dviejų vidurkių vektorių lygybės yra ekvivalentus kriterijui (2.2.5), grindžiamam Hotelingo statistika.

2.6. Įrodykite, kad tikėtinumų santykio kriterijus tikrinant hipotezę dėl kelių vidurkių vektorių palyginimo yra ekvivalentus kriterijui (2.3.3), grindžiamam Hotelingo statistika.

2.7. Lentelėje pateikta 10 pacientų papildomo miego laikas (valandomis) X_1 vartojant vaistą A ir X_2 vartojant vaistą B.

Pacientas	X_1	X_2	Pacientas	X_1	X_2
1	1,9	0,7	6	4,4	3,4
2	0,8	-1,6	7	5,5	3,7
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

a) Tarę, kad buvo stebimas normalusis vektorius, patikrinkite hipotezę, kad a. v. $(X_1, X_2)^T$ vidurkis lygus nuliniam vektoriui; kriterijaus reikšmingumo lygmuo $\alpha = 0,05$. Patikrinkite, kad gautasis kriterijus yra ekvivalentus Stjudento priklausomų imčių kriterijui vienmačiu atveju. b) Sudarykite vidurkių vektoriaus pasikliovimo sritį, kai pasikliovimo lygmuo $Q = 0,95$.

2.8. Tegu X_1 ir X_2 yra taurėlapio ilgis ir plotis, o X_3 ir X_4 – vainiklapio ilgis ir plotis. Atlikta po 50 matavimų *Iris versicolor* ir *Iris setosa* [2]. Pagal šiuos duomenis gautos statistikų realizacijos

$$\bar{\mathbf{X}}_1 = \begin{pmatrix} 5,936 \\ 2,770 \\ 4,260 \\ 1,326 \end{pmatrix}, \quad \bar{\mathbf{X}}_2 = \begin{pmatrix} 5,006 \\ 3,428 \\ 1,462 \\ 0,246 \end{pmatrix};$$

$$\boldsymbol{S} = \boldsymbol{S}_1 + \boldsymbol{S}_2 = \begin{pmatrix} 19,1434 & 9,0356 & 9,7634 & 3,2394 \\ 9,0356 & 11,8658 & 4,6232 & 2,4746 \\ 9,7634 & 4,6232 & 12,2978 & 3,8794 \\ 3,2394 & 2,4746 & 3,8794 & 2,4604 \end{pmatrix}.$$

Tarę, kad buvo stebimas normalusis vektorius, patikrinkite hipotezę $H : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ dėl vidurkių vektorių lygybės kriterijumi, kurio reikšmingumo lygmuo $\alpha = 0,01$.

2.9. Tegu Š, R, P, V yra kamštinio medžio žievės storis keturiomis pasaulio kryptimis. Konkretaus medžio šių keturių dydžių matavimus galime interpretuoti kaip keturmačio a. v. realizaciją. Pažymėkime X_1 =Š-R-V+P, X_2 =P-V, X_3 =Š-P. Atlikus 28 matavimus gautos statistikų realizacijos [14]

$$\bar{\mathbf{X}} = \begin{pmatrix} 8,86 \\ 4,50 \\ 0,86 \end{pmatrix}, \quad \hat{\mathbf{\Sigma}} = \frac{1}{28}\boldsymbol{S} = \begin{pmatrix} 128,72 & 61,41 & -21,02 \\ 61,41 & 56,93 & -28,30 \\ -21,02 & -28,30 & 63,53 \end{pmatrix}.$$

Tarę, kad buvo stebimas normalusis a. v., patikrinkite hipotezę $H : \mu_1 = \mu_2 = \mu_3 = 0$; čia $\mu_i = \mathbf{E}X_i, i = 1,2,3$.

2.10. Lentelėje yra pateikti duomenys, gauti stebint sveikų moterų prakaitavimo rodiklius, t. y. trimatį a. v. $\mathbf{X} = (X_1, X_2, X_3)^T$; čia $\boldsymbol{X}_1$ – prakaito kiekis, X_2 – natrio kiekis, X_3 – kalio kiekis [9].

i	X_{1i}	X_{2i}	X_{3i}	i	X_{1i}	X_{2i}	X_{3i}
1	3,7	48,5	9,3	11	3,9	36,9	12,7
2	5,7	65,1	8,0	12	4,5	58,8	12,3
3	3,8	47,2	10,9	13	3,5	27,8	9,8
4	3,2	53,2	12,0	14	4,5	40,2	8,4
5	3,1	55,5	9,7	15	1,5	17,5	10,1
6	4,6	36,1	7,9	16	8,5	56,4	7,1
7	2,4	24,8	14,0	17	4,5	71,6	8,2
8	7,2	33,1	7,6	18	6,5	52,8	10,9
9	6,7	47,4	8,5	19	4,1	44,1	11,2
10	5,4	54,1	11,3	20	5,5	40,9	9,4

Tarę, kad buvo stebėtas normalusis a. v., raskite vidurkių vektoriaus $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)^T$ pasikliovimo sritį ir jo koordinačių pasikliovimo intervalų rinkinius, kai pasikliovimo lygmuo $Q = 0,95$.

2.11. (**2.10** tęsinys). Tarę, kad buvo stebimas normalusis vektorius, pagal **2.10** pratimo duomenis patikrinkite hipotezę $H : \boldsymbol{\mu} = \boldsymbol{\mu}_0 = (4; 50; 10)^T$, kai kriterijaus reikšmingumo lygmuo $\alpha = 0,05$.

2.12. (**1.12** pratimo tęsinys). Išspręskite **1.12** pratimą tuo atveju, kai a. v. koordinačių dispersijos nėra žinomos.

2.13. (**1.11** pratimo tęsinys). Išspręskite **1.11** pratimą tuo atveju, kai kovariacijų matrica $\mathbf{\Sigma}$ nėra žinoma.

2.14. (**1.8** pratimo tęsinys). Tarkime, kad **1.8** pratimo sąlygomis vidurkių vektorius $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4)^T$. Priėmę normalumo prielaidą patikrinkite hipotezę $H: \mu_1 = \mu_3, \mu_2 = \mu_4$.

2.15. (**1.14** pratimo tęsinys). Išspręskite **1.14** pratimą tuo atveju, kai kovariacijų matrica $\boldsymbol{\Sigma}$ nėra žinoma.

2.16. (**1.9** pratimo tęsinys). Tarę, kad **1.9** pratimo sąlygomis atveju a) stebėtas a. v. $\mathbf{X}_1 \sim N_2(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$, o atveju b) a. v. $\mathbf{X}_2 \sim N_2(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$. Patikrinkite hipotezę $H: 1,22\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \mathbf{0}$ a) tarę, kad kovariacinės matricos $\boldsymbol{\Sigma}_1$ ir $\boldsymbol{\Sigma}_2$ yra lygios; b) kovariacinių matricų lygybės prielaida nepriimama.

2.17. Aliaskos natūralistas Haris Robertsas tyrė grizlių populiaciją. Buvo allikti $n = 61$ grizlių charakteristikų X_1 – svoris; X_2 – kūno ilgis; X_3 – kaklo ilgis; X_4 – liemens apimtis; X_5 – galvos ilgis; X_6 – galvos plotis matavimai. Pagal šiuos matavimus gauta vidurkių vektoriaus $\boldsymbol{\mu}$ ir kovariacinės matricos įverčiai [9]:

$$\hat{\boldsymbol{\mu}} = \begin{pmatrix} 95,52 \\ 164,38 \\ 55,69 \\ 93,39 \\ 17,98 \\ 31,13 \end{pmatrix}, \quad \hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 3266,46 & 1343,97 & 731,54 & 1175,50 & 162,68 & 238,37 \\ 1343,97 & 721,91 & 324,25 & 537,35 & 80,17 & 117,73 \\ 731,54 & 324,25 & 179,28 & 281,17 & 39,15 & 56,80 \\ 1175,50 & 537,35 & 281,17 & 474,98 & 63,73 & 94,85 \\ 162,68 & 80,17 & 39,15 & 63,73 & 9,95 & 13,88 \\ 238,37 & 117,73 & 56,80 & 94,85 & 13,88 & 21,26 \end{pmatrix}.$$

Priėmę normalumo prielaidą a) raskite pasikliovimo lygmens $Q = 0,95$ vidurkių vektoriaus $\boldsymbol{\mu}$ pasikliovimo elipsoidą; b) raskite pasikliovimo lygmens $Q = 0,95$ vidurkių vektoriaus $\boldsymbol{\nu} = (\nu_1, \nu_2, \nu_3)^T = (\mu_2, \mu_3 - \mu_4, \mu_5 - \mu_6)^T$ pasikliovimo elipsoidą; c) raskite parametrų ν_1, ν_2, ν_3 pasikliovimo intervalų rinkinius naudodami surastą pasikliovimo elipsoidą, remdamiesi Bonferonio nelygybe ir palyginkite juos su to paties pasikliovimo lygmens koordinačių pasikliovimo intervalais.

2.18. Tirta pieno transportavimo į perdirbimo įmones charakteristikos. Gautos dvi didumo $n_1 = 36$ ir $n_2 = 18$ paprastosios nepriklausomos imtys stebint a. v. $\mathbf{X} = (X_1, X_2, X_3)^T$ ir $\boldsymbol{Y} = (Y_1, Y_2, Y_3)^T$. Čia X_1 – išlaidos kurui; X_2 – remonto išlaidos, X_3 – įmonės kapitalas, kai naudojami benzininiai vilkikai; Y_1, Y_2, Y_3 – analogiški rodikliai, naudojant dyzelinius vilkikus. Duomenys pateikiami lentelėje [9].

X_{1i}	X_{2i}	X_{3i}	X_{1i}	X_{2i}	X_{3i}	Y_{1i}	Y_{2i}	Y_{3i}
16,44	12,43	11,23	12,34	7,73	11,68	8,50	12,26	9,11
7,19	2,70	3,92	8,51	14,02	12,01	7,42	5,13	17,15
9,92	1,35	9,75	26,16	17,44	16,89	10,28	3,32	11,23
4,24	5,78	7,78	12,95	8,24	7,18	10,16	14,72	5,99
11,20	5,05	10,67	16,93	13,37	17,59	12,79	4,17	29,28
14,25	5,78	9,88	14,70	10,78	14,58	9,60	12,72	11,00
13,50	10,98	10,60	10,32	5,16	17,00	6,47	8,89	19,00
13,32	14,27	9,45	8,98	4,49	4,26	11,35	9,95	14,53
29,11	15,09	3,28	9,70	11,59	6,83	9,15	2,94	13,68
12,68	7,61	10,23	12,72	8,63	5,59	9,70	5,06	20,84
7,51	5,80	8,13	9,49	2,16	6,23	9,77	17,86	35,18
9,90	3,63	9,13	8,22	7,95	6,72	11,61	11,75	17,00
10,25	5,07	10,17	13,70	11,22	4,91	9,09	13,25	20,66
11,11	6,15	7,61	8,21	9,85	8,17	8,53	10,14	17,45
12,17	14,26	14,39	15,86	11,42	13,06	8,29	6,22	16,38
10,24	2,59	6,09	9,18	9,18	9,49	15,90	12,90	19,09
10,18	6,05	12,14	12,49	4,67	11,94	11,94	5,69	14,77
8,88	2,70	12,23	17,32	6,86	4,44	9,54	16,77	22,66

Priėmę normalumo prielaidą a) patikrinkite simetriškumo hipotezę atskirai pirmajai ir antrajai imčiai; b) patikrinkite vidurkių vektorių lygybės hipotezę, tarę kad kovariacinės matricos yra vienodos.

Atsakymai ir nurodymai

2.2. $\mathbf{P}\{\mathbf{X}_{n+1} \in \boldsymbol{C}(\bar{\mathbf{X}}, \boldsymbol{S})\} = Q = 1 - \alpha$, kai

$$\boldsymbol{C}(\bar{\mathbf{X}}, \boldsymbol{S}) = \{\mathbf{X}_{n+1} : \frac{n(n-k)}{k(n+1)}(\bar{\mathbf{X}} - \mathbf{X}_{n+1})^T \boldsymbol{S}^{-1}(\bar{\mathbf{X}} - \mathbf{X}_{n+1}) < F_\alpha(k, n-k)\}.$$

2.7 a) Randame

$$\bar{\mathbf{X}} = \begin{pmatrix} 2,33 \\ 0,75 \end{pmatrix}, \quad \boldsymbol{S} = \begin{pmatrix} 36,081 & 25,635 \\ 25,635 & 28,805 \end{pmatrix}, \quad \boldsymbol{S}^{-1} = \begin{pmatrix} 0,0754 & -0,0671 \\ -0,0671 & 0,0944 \end{pmatrix}$$

ir apskaičiuojame statistikos F realizaciją

$$F = \frac{n(n-k)}{k}\bar{\mathbf{X}}^T \boldsymbol{S}^{-1}\bar{\mathbf{X}} = 9,115.$$

Kadangi P reikšmė $pv = \mathbf{P}\{F_{2;8} > 9,115\} = 0,0087$, tai hipotezė atmestina. b) $\boldsymbol{C}(\bar{\mathbf{X}}, \boldsymbol{S}) = \{\boldsymbol{\mu} : \frac{n(n-k)}{k}(\bar{\mathbf{X}} - \boldsymbol{\mu})^T \boldsymbol{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \leq F_\alpha(k, n-k)\} = \{(\mu_1, \mu_2) : 3,016\,(2,33 - \mu_1)^2 - 5,368\,(2,33 - \mu_1)(0,75 - \mu_2) + 3,776\,(0,75 - \mu_2)^2 < 4,459\}$.

2.8. Randame $(\boldsymbol{S}_1 + \boldsymbol{S}_2)^{-1}$ ir apskaičiuojame statistikos F realizaciją

$$F = \frac{n_1 n_2(n_1 + n_2 - k - 1)}{(n_1 + n_2)k}(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2)^T(\boldsymbol{S}_1 + \boldsymbol{S}_2)^{-1}(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) = 625,5.$$

Hipotezė atmestina.

2.9. Randame

$$\boldsymbol{S}^{-1} = \frac{1}{28}\begin{pmatrix} 0,0165 & -0,0194 & -0,0032 \\ -0,0194 & 0,0453 & 0,0138 \\ -0,0032 & 0,0138 & 0,0208 \end{pmatrix}$$

ir apskaičiuojame statistikos F realizaciją

$$F = \frac{n(n-k)}{k}\bar{\mathbf{X}}^T \boldsymbol{S}^{-1}\bar{\mathbf{X}} = 6,1784.$$

Kadangi P reikšmė $pv = \mathbf{P}\{F_{3;25} > 6,1784\} = 0,0027$, tai hipotezė atmestina.

2.10.

$$\bar{\mathbf{X}} = \begin{pmatrix} 4,640 \\ 45,600 \\ 9,965 \end{pmatrix}, \quad \hat{\boldsymbol{\Sigma}} = \frac{1}{19}\boldsymbol{S} = \begin{pmatrix} 2.879 & 9.349 & -1.809 \\ 9.349 & 187.157 & -5.612 \\ -1.809 & -5.612 & 3.628 \end{pmatrix}.$$

$$\mathbf{P}\left\{\frac{340}{3}(\bar{\mathbf{X}} - \boldsymbol{\mu})^T \boldsymbol{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \leq F_{0,05}(3,\ 17)\right\} = 0,95.$$

Iš čia pasikliovimo intervalų rinkinys vektoriaus $\boldsymbol{\mu}$ koordinatėms yra

$$(\underline{\mu}_1; \overline{\mu}_1) = (3,398;\ 5,822); \quad (\underline{\mu}_2; \overline{\mu}_2) = (35,585;\ 55,615); \quad (\underline{\mu}_3; \overline{\mu}_3) = (8,570;\ 11,360).$$

Remdamiesi Bonferonio nelygybe gauname intervalų rinkinį:

$$(\underline{\mu}_1; \overline{\mu}_1) = (3,644;\ 5,636); \quad (\underline{\mu}_2; \overline{\mu}_2) = (37,569;\ 53,631); \quad (\underline{\mu}_3; \overline{\mu}_3) = (8,836;\ 11,094).$$

Jeigu, pavyzdžiui, reikėtų sudaryti tik vienos koordinatės μ_1 pasikliovimo intervalą, kai pasikliovimo lygmuo tas pats, tai pagal (2.1.2) gautume $(\underline{\mu}_1; \overline{\mu}_1) = (3,846; 5,434)$.

2.11. Randame statistikos F reikšmę

$$F = \frac{n(n-k)}{k}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T \boldsymbol{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0) = 2,905.$$

Kadangi $2,905 < F_{0,05}(3,\ 17) = 3,1968$, darome išvadą, kad turimi duomenys neprieštarauja iškeltai hipotezei.

2.12 **1.12** pratimo žymenimis gauname pasikliovimo intervalus

i	$\underline{\mu}_i$	$\overline{\mu}_i$	$\underline{\mu}''_i$	$\overline{\mu}''_i$	$\underline{\mu}'_i$	$\overline{\mu}'_i$
1	26,93	29,27	26,48	29,72	25,76	30,44
2	25,41	27,79	24,96	28,24	24,33	28,97
3	34,63	36,17	34,33	36,47	33,85	36,95
4	33,16	35,24	32,76	35,64	32,12	36,28
5	22,84	24,36	22,54	24,66	22,07	25,13
6	21,20	22,80	20,90	23,10	20,41	23,59
7	21,88	23,52	21,57	23,83	21,07	24,33

2.13. a) vidurkių vektoriaus $\boldsymbol{\nu} = (\nu_1, \nu_2, \nu_3)^T$ koordinačių pasikliovimo intervalai yra tokie: (106,81; 255,46); (-81,58; 130,71); (-278,99; -152,68); b) statistika F, kuri esant teisingai hipotezei turi Fišerio skirstinį su 3 ir 27 laisvės laipsniais, įgijo reikšmę 3,3396; kadangi P reikšmė $pv = \mathbf{P}\{F_{3;27} > 3,3396\} = 0,034$, hipotezė atmetama, jeigu kriterijaus reikšmingumo lygmuo viršija 0,034. **2.14.** Statistika F, kuri esant teisingai hipotezei turi Fišerio skirstinį su 2 ir 23 laisvės laipsniais, įgijo reikšmę 1,7309; kadangi P reikšmė $pv = \mathbf{P}\{F_{2;23} > 1,7309\} = 0,1994$, tai atmesti hipotezę nėra pagrindo. **Nurodymas.** A. v. $\boldsymbol{Z} = (Z_1, Z_2)^T = (X_1 - X_3, X_2 - X_3)^T$ turi dvimatį normalųjį skirstinį su vidurkių vektoriumi $\mathbf{E}(\boldsymbol{Z}) = (\mu_1 - \mu_3, \mu_2 - \mu_4)^T$. Taigi reikia patikrinti hipotezę $H : \mathbf{E}(\boldsymbol{Z}) = \mathbf{0}$, kad dvimačio normaliojo skirstinio vidurkių vektorius lygus nuliniam vektoriui, kai kovariacinė matrica nežinoma, pagal paprastąją imtį $\boldsymbol{Z}_1, ..., \boldsymbol{Z}_{25}$. **2.15.** a) Statistika F iš (2.1.14) įgijo reikšmę 2,7349; P reikšmė $pv = \mathbf{P}\{F_{4;36} > 2,7349\} = 0,0438$; hipotezė atmetama, jeigu kriterijaus reikšmingumo lygmuo viršija 0,0438. b) Pažymėkime $Z_1 = X_1^* - 22,2875$, $Z_2 = X_2^* - 272,575$, $Z_3 = X_3^* - 288,435$, $Z_4 = X_4^* - 51,975$. Tada prognozės elipsoidas yra $7,6249Z_1^2 + 0,0440Z_2^2 + 0,8465Z_3^2 + 1,1993Z_4^2 + 0,2254Z_1Z_2 - 0,1886Z_1Z_3 + 1,436Z_1Z_4 - 0,0404Z_2Z_3 - 0,0312Z_2Z_4 + 0,1478Z_3Z_4 < 11,1336$. **Nurodymas.** A. d. $(n+1)(n-k)(\mathbf{X}^* - \hat{\boldsymbol{\mu}})^T\boldsymbol{\Sigma}^{-1}(\mathbf{X}^* - \hat{\boldsymbol{\mu}})/(n(n-1)k)$ turi Fišerio skirstinį su 4 ir 36 laisvės laipsniais. **2.16.** a) Statistika F, kuri esant teisingai hipotezei turi Fišerio skirstinį su 2 ir 42 laisvės laipsniais, įgijo reikšmę 1,1084; P reikšmė $pv = \mathbf{P}\{F_{2;42} > 1,1084\} = 0,3395$; atmesti hipotezę nėra pagrindo. b) Statistika, kuri esant teisingai hipotezei asimptotiškai turi χ^2 skirstinį su 2 laisvės laipsniais, įgijo reikšmę 0,1557; P reikšmė $pv = \mathbf{P}\{\chi_2^2 > 0,1557\} = 0,9251$; atmesti hipotezę nėra pagrindo. **2.17.** a) Pasikliovimo elipsoidas $E(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}) = \{\boldsymbol{\mu} : 33550(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})^T\hat{\boldsymbol{\Sigma}}^{-1}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}) < 2,2687\}$; b) Pasikliovimo elipsoidas $E(\hat{\boldsymbol{\nu}}, \hat{\Gamma}) = \{\boldsymbol{\nu} : 70760(\hat{\boldsymbol{\nu}} - \boldsymbol{\nu})^T\hat{\Gamma}^{-1}(\hat{\boldsymbol{\nu}} - \boldsymbol{\nu}) < 2,7636\}$,

$$\hat{\boldsymbol{\nu}} = \begin{pmatrix} 164,38 \\ -37,70 \\ -13,15 \end{pmatrix}, \quad \hat{\Gamma} = \begin{pmatrix} 721,91 & -213,10 & -37,56 \\ -213,10 & 91,92 & 13,44 \\ -37,56 & 13,44 & 3,45 \end{pmatrix}.$$

c) parametro ν_1 pasikliovimo intervalų rėžiai yra: naudojant pasikliovimo elipsoidą $164,38 \pm 10,1583$; remiantis Bonferonio nelygybe $164,38 \pm 9,4645$; individualus intervalas $164,38 \pm 7,9745$; analogiškai kitiems parametrams.

2.18. a) Pirmosios imties atveju statistika (2.4.5) įgijo reikšmę 18,9445; P reikšmė $pv = \mathbf{P}\{F_{2;34} > 18,9445\} = 3 \cdot 10^{-6}$; hipotezė atmestina; antrosios imties atveju statistika (2.4.5) įgijo reikšmę 10,5451; P reikšmė $pv = \mathbf{P}\{F_{2;16} > 10,5451\} = 0,0012$; hipotezė atmetama, kai kriterijaus reikšmingumo lygmuo viršija 0,0012; b) statistika (2.2.4) įgijo reikšmę 12,206; P reikšmė $pv = \mathbf{P}\{F_{3;50} > 12,206\} = 4 \cdot 10^{-6}$; hipotezė atmestina.

3 skyrius

Tiesiniai modeliai daugiamačiu atveju

Antros vadovėlio dalies tiesiniuose modeliuose nagrinėjome vienmačio požymio priklausomybę nuo tam tikrų kintamųjų (kovariančių, faktorių). Tarėme, kad tiriamo požymio vidurkis yra tiesinė nežinomų parametrų funkcija, kurios koeficientai nusakomi tam tikru būdu suplanuoto eksperimento plano matrica.

Tačiau praktiškai objektui apibūdinti gali nepakakti vieno požymio. Pavyzdžiui, nagrinėjant skirtingas kviečių veisles, gali dominti ne tik jų derlingumas, bet ir kitos charakteristikos: grūdų krakmolingumas ir glitumas, augalų aukštis ir atsparumas išgulimui, šiaudų kiekis ir kt. Nagrinėjant pacientų sistolinio kraujo spaudimo priklausomybę nuo jų amžius ir svorio, gali būti svarbu kartu nagrinėti ir kitus požymius: kardiogramos kreivės charakteristikas, kraujagyslių sienelių elastingumą, cholesterolio kiekį ir kt.

Apibendrinant vienmatį tiesinį modelį natūralu tarti, kad kiekvieno tiriamo požymio vidurkis yra tiesinė nežinomų parametrų (galbūt skirtingų kiekvienam požymiui) funkcija, kurios koeficientus nusako ta pati eksperimento plano matrica.

3.1. Matematinis modelis

Tarkime, kad n kartų yra stebimas k požymių vektorius $\mathbf{X} = (X_1, ..., X_k)^T$. Stebėjimus galime surašyti į lentelę.

	$\mathbf{X}_1$	$\mathbf{X}_2$	...	$\mathbf{X}_n$	
$\boldsymbol{Y}_1^T$	X_{11}	X_{12}	$\cdots$	X_{1n}	
$\boldsymbol{Y}_2^T$	X_{21}	X_{22}	$\cdots$	X_{2n}	$= \mathcal{Y}^T$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
$\boldsymbol{Y}_k^T$	X_{k1}	X_{k2}	$\cdots$	X_{kn}	

Tariame, kad stulpeliuose parašyti vektoriai $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n$ yra nepriklausomi. O skirtingose eilutėse parašyti vektoriai, kurie reiškia tų pačių objektų įvairių požymių matavimus, gali būti priklausomi. Tarkime, kad kovariacinė matrica $\boldsymbol{V}(\mathbf{X}) = \boldsymbol{V}(\mathbf{X}_i) = \boldsymbol{\Sigma} = [\sigma_{ij}]_{k\times k}$, $i = 1, ..., n$, nekinta, kai požymių vektoriaus $\mathbf{X}$ matavimo numeris kinta nuo 1 iki n. Atsitiktinių vektorių $\mathbf{X}_i, i = 1, ..., n$, vidurkiai gali būti skirtingi. Tarsime, kad vektoriaus $\boldsymbol{Y}_i = (X_{i1}, ..., X_{in})^T$, kuris reiškia i-ojo požymio matavimus, koordinačių vidurkis kinta pagal vienmatį Gauso ir Markovo tiesinį modelį (vadovėlio 2 dalies 1 skyrius). Taigi stebėjimų matematinis modelis yra vienmačių Gauso ir Markovo modelių, sudarytų kiekvienam požymiui, sistema

$$\boldsymbol{Y}_i = \boldsymbol{A}\boldsymbol{\beta}_i + \boldsymbol{e}_i, \quad \mathbf{E}(\boldsymbol{Y}_i) = \boldsymbol{A}\boldsymbol{\beta}_i, \quad \boldsymbol{V}(\boldsymbol{Y}_i) = \sigma_{ii}\boldsymbol{I}, \quad i = 1, ..., k; \tag{3.1.1}$$

čia $\boldsymbol{A} = [a_{jl}]_{n\times m}$ yra eksperimento plano matrica, $\boldsymbol{\beta}_i = (\beta_{i1}, ..., \beta_{im})^T$ – nežinomų parametrų vektorius, apibūdinantis i-ojo požymio vidurkio kitimą, $\boldsymbol{e}_i = (e_{i1}, ..., e_{in})^T$ – paklaidų vektorius, $\mathbf{E}(\boldsymbol{e}_i) = \mathbf{0}, \boldsymbol{V}(\boldsymbol{e}_i) = \sigma_{ii}\boldsymbol{I}$. Tiriamų požymių vektorių priklausomumą nusako kovariacija

$$\mathbf{Cov}\,(\boldsymbol{Y}_i, \boldsymbol{Y}_{i'}) = \mathbf{Cov}\,(\boldsymbol{e}_i, \boldsymbol{e}_{i'}) = \sigma_{ii'}\boldsymbol{I}, \quad i, i' = 1, ..., k. \tag{3.1.2}$$

Naudojant matricinius žymenis modelį (3.1.1) galima užrašyti tokiu pavidalu

$$\mathcal{Y} = \boldsymbol{A}\boldsymbol{B}^T + \boldsymbol{e}; \tag{3.1.3}$$

čia $\boldsymbol{B}^T = [\beta_{ij}]_{m\times k}$ yra nežinomų parametrų matrica, kurios stulpelius sudaro vektoriai $\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_k$; $\boldsymbol{e} = [e_{ji}]_{n\times k}$ – paklaidų matrica, jos stulpelių vektoriai $\boldsymbol{e}_1, ..., \boldsymbol{e}_k$.

Užrašius detaliau, lygybė (3.1.3) yra tokia:

$$(\boldsymbol{Y}_1 \vdots \boldsymbol{Y}_2 \vdots \ldots \vdots \boldsymbol{Y}_k) = \boldsymbol{A}(\boldsymbol{\beta}_1 \vdots \boldsymbol{\beta}_2 \vdots \ldots \vdots \boldsymbol{\beta}_k) + (\boldsymbol{e}_1 \vdots \boldsymbol{e}_2 \vdots \ldots \vdots \boldsymbol{e}_k).$$

Transponavę abi (3.1.3) lygybės puses gauname sąryšį

$$\mathcal{Y}^T = \boldsymbol{B}\boldsymbol{A}^T + \boldsymbol{e}^T,$$

arba

$$(\mathbf{X}_1 \vdots \mathbf{X}_2 \vdots \ldots \vdots \mathbf{X}_n) = \boldsymbol{B}(\boldsymbol{a}_1 \vdots \boldsymbol{a}_2 \vdots \ldots \vdots \boldsymbol{a}_n) + (\boldsymbol{c}_1 \vdots \boldsymbol{c}_2 \vdots \ldots \vdots \boldsymbol{c}_n), \tag{3.1.4}$$

čia $\boldsymbol{a}_j$ – vektorius, kurį sudaro plano matricos $\boldsymbol{A}$ j-osios eilutės elementai, $\boldsymbol{c}_j$ yra vektorius, kurį sudaro matricos $\boldsymbol{e}$ j-osios eilutės elementai. Taigi

$$\mathbf{E}(\mathbf{X}_j) = \boldsymbol{B}\boldsymbol{a}_j, \quad \mathbf{E}(\boldsymbol{c}_j) = \mathbf{0}, \quad \boldsymbol{V}(\mathbf{X}_j) = \boldsymbol{V}(\boldsymbol{c}_j) = \boldsymbol{\Sigma}, \quad j = 1, ..., n;$$

$$\mathbf{Cov}\,(\mathbf{X}_j, \mathbf{X}_{j'}) = \mathbf{Cov}\,(\boldsymbol{c}_j, \boldsymbol{c}_{j'}) = \mathbf{0}, \quad j \neq j' = 1, ..., n.$$

3.1.1 pavyzdys. Tarkime, kad yra lyginamos m skirtingų kviečių veislių. Tuo tikslu atsitiktinai parinktų n_j sklypelių apsėjama j-ąja kviečių veisle; iš viso sklypelių $n = n_1 + ... + n_m$. Kviečių veislės lyginamos atsižvelgiant į dimensijos k skirtingų požymių vektorių $\mathbf{X} = (X_1, ..., X_k)^T$. Pažymėkime $\mathbf{X}_{lj} = (X_{1lj}, ..., X_{klj})^T$ požymio vektoriaus matavimus, gautus j-ajame sklypelyje, kuris apsėtas l-ąja kviečių veisle, $j = 1, ..., n_l$, $l = 1, ..., m$. Tarsime, kad vektoriaus $\mathbf{X}_{lj}$ vidurkis $\mathbf{E}(\mathbf{X}_{lj}) = \boldsymbol{\mu}_l = (\mu_{1l}, ..., \mu_{kl})^T$ nepriklauso nuo sklypelio numerio j, tačiau gali priklausyti nuo kviečių veislės numerio l. Kovariacinė matrica $\boldsymbol{V}(\mathbf{X}_{lj}) = \boldsymbol{\Sigma} = [\sigma_{ii'}]_{k\times k}$ nepriklauso nuo indeksų j ir l, o kovariacijos $\mathbf{Cov}\,(\mathbf{X}_{lj}, \mathbf{X}_{l'j'}) = \mathbf{0}$, kai $l \neq l'$, arba kai $l = l'$ bet $j \neq j'$.

Šiame pavyzdyje vektorių $\boldsymbol{Y}_i$ sudaro i-ojo požymio matavimai

$$\boldsymbol{Y}_i = (X_{i11}, ..., X_{i1n_1}, X_{i21}, , X_{i2n_2}, ..., X_{im1}, ..., X_{imn_m})^T,$$

kuriam galioja vienmatis tiesinis Gauso ir Markovo modelis

$$\boldsymbol{Y}_i = \boldsymbol{A}\boldsymbol{\beta}_i + \boldsymbol{e}_i;$$

čia $\boldsymbol{\beta}_i = (\mu_{i1}, \mu_{i2}, ..., \mu_{im})^T$ yra i-ojo požymio vidurkių vektorius visoms m kviečių veislėms; plano matricos pirmosios n_1 eilučių yra $\boldsymbol{a}_i^T = (1,\ 0, ..., 0), i = 1, ..., n_1$; tolesnės n_2 eilučių yra $\boldsymbol{a}_i^T = (0,\ 1, ..., 0), i = n_1 + 1, ..., n_1 + n_2$, pagaliau paskutinės n_m eilučių yra $\boldsymbol{a}_i^T = (0,\ 0, ..., 1), i = n_1 + ... + n_{m-1} + 1, ..., n$.

Imdami pateiktų vienmačių Gauso ir Markovo modelių sistemą, kai $i = 1, ..., k$, gausime modelį (3.1.3).

Šis modelis yra vienfaktorės dispersinės analizės modelio apibendrinimas daugiamačiu atveju.

3.1.2 pavyzdys. Tarkime, kad vertinama pacientų tam tikro požymių vektoriaus $\mathbf{X} = (X_1, ..., X_k)^T$ skirstinio priklausomybė nuo dviejų kovariančių – svorio (Z_1) ir amžiaus (Z_2). Tuo tikslu atsitiktinai atrenkama n pacientų, kuriems pamatuojamos požymio vektoriaus $\mathbf{X}$ ir kovariančių vektoriaus $(Z_1, Z_2)^T$ reikšmės. Pažymėkime $\mathbf{X}_i = (X_{1i}, ..., X_{ki})^T$ tiriamų požymių vektorių ir $(Z_{1i}, Z_{2i})^T$ kovariančių vektorių i-ajam pacientui. Tarsime, kad požymių vektoriaus vidurkiai tiesiškai priklauso nuo kovariančių, kovariacijų matrica $\boldsymbol{V}\mathbf{X}_i = \boldsymbol{\Sigma} = [\sigma_{ij}]_{k\times k}$ nekinta ir kovariacijos $\mathbf{Cov}\,(\mathbf{X}_i, \mathbf{X}_{i'}) = \mathbf{0}, i \neq i' = 1, ..., n$. Tada fiksavus kovariančių reikšmes, i-ojo požymio vektorius $\boldsymbol{Y}_i = (X_{i1}, ..., X_{in})^T$ tenkina tiesinį Gauso ir Markovo modelį

$$\boldsymbol{Y}_i = \boldsymbol{A}\boldsymbol{\beta}_i + \boldsymbol{e}_i, \quad \mathbf{E}(\boldsymbol{Y}_i) = \boldsymbol{A}\boldsymbol{\beta}_i, \quad \mathbf{E}(\boldsymbol{e}_i) = \mathbf{0}, \quad \boldsymbol{V}(\boldsymbol{Y}_i) = \boldsymbol{V}(\boldsymbol{e}_i) = \sigma_{ii}\boldsymbol{I};$$

čia $\boldsymbol{\beta}_i = (\alpha_i, \beta_{1i}, \beta_{2i})^T$ nežinomų regresijos parametrų vektorius, o matricos $\boldsymbol{A}$ i-oji eilutė yra $\boldsymbol{a}_i^T = (1, Z_{1i}, Z_{2i}), i = 1, ..., n$.

Sujungę šiuos modelius į vieną sistemą, kai $i = 1, ..., k$, gausime modelį (3.1.3). Šis modelis yra tiesinės regresijos modelio apibendrinimas daugiamačiu atveju.

3.2. Parametrų įvertiniai

Tarsime, kad plano matricos $\boldsymbol{A}$ rangas lygus m, t. y. $|\boldsymbol{A}^T\boldsymbol{A}| \neq 0$.

Pažymėkime $\hat{\boldsymbol{\beta}}_i$ mažiausiųjų kvadratų įvertinį, gautą iš i-ojo vienmačio Gauso ir Markovo tiesinio modelio (3.1.1):

$$\hat{\boldsymbol{\beta}}_i = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y}_i, \quad i = 1, ..., k. \tag{3.2.1}$$

3.2.1 teorema. *Su bet kuriuo $\boldsymbol{L} \in \mathbf{R}^k$ parametro $\theta_i = \boldsymbol{L}^T\boldsymbol{\beta}_i$, kuris yra tik nežinomų parametrų vektoriaus $\boldsymbol{\beta}_i = (\beta_{i1}, ..., \beta_{im})^T$ tiesinė funkcija, įvertinys $\hat{\theta}_i = \boldsymbol{L}^T\hat{\boldsymbol{\beta}}_i$ yra vienintelis minimalios dispersijos įvertinys nepaslinktųjų tiesinių visų stebėjimų $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_k$ įvertinių aibėje. Šio įvertinio dispersija ir kovariacija su kitos tiesinės funkcijos $\boldsymbol{K}^T\boldsymbol{\beta}_i$ įvertiniu $\boldsymbol{K}^T\hat{\boldsymbol{\beta}}_i$ yra*

$$\boldsymbol{V}(\boldsymbol{L}^T\hat{\boldsymbol{\beta}}_i) = \sigma_{ii}\boldsymbol{L}^T(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{L}, \quad \mathbf{Cov}\,(\boldsymbol{L}^T\hat{\boldsymbol{\beta}}_i, \boldsymbol{K}^T\hat{\boldsymbol{\beta}}_i) = \sigma_{ii}\boldsymbol{L}^T(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{K}. \tag{3.2.2}$$

Įrodymas Tarkime, kad

$$\tilde{\theta}_i = \boldsymbol{H}_1^T \boldsymbol{Y}_1 + ... + \boldsymbol{H}_k^T \boldsymbol{Y}_k$$

yra kitas nepaslinktas tiesinis visų stebėjimų parametro θ_i įvertinys. Iš nepaslinktumo sąlygos: su visais $\boldsymbol{\beta}_i \in \mathbf{R}^k$

$$\mathbf{E}(\tilde{\theta}_i) = \mathbf{E}(\boldsymbol{H}_1^T \boldsymbol{Y}_1) + ... + \mathbf{E}(\boldsymbol{H}_k^T \boldsymbol{Y}_k) = \boldsymbol{H}_1^T \boldsymbol{A}\boldsymbol{\beta}_1 + ... + \boldsymbol{H}_k^T \boldsymbol{A}\boldsymbol{\beta}_k = \boldsymbol{L}^T \boldsymbol{\beta}_i$$

išplaukia, kad

$$\boldsymbol{H}_i^T \boldsymbol{A} = \boldsymbol{L}^T, \quad \boldsymbol{H}_j^T \boldsymbol{A} = 0, \quad j \neq i. \tag{3.2.3}$$

Turime

$$\boldsymbol{V}(\tilde{\theta}_i) = \boldsymbol{V}(\tilde{\theta}_i - \hat{\theta}_i + \hat{\theta}_i) = \boldsymbol{V}(\hat{\theta}_i) + \boldsymbol{V}(\tilde{\theta}_i - \hat{\theta}_i) + 2\mathbf{Cov}\,(\tilde{\theta}_i - \hat{\theta}_i, \hat{\theta}_i).$$

Kadangi

$$\mathbf{Cov}\,(\tilde{\theta}_i, \hat{\theta}_i) = \sum_{j=1}^{k} \mathbf{Cov}\,(\boldsymbol{H}_j^T \boldsymbol{Y}_j, \boldsymbol{L}^T \hat{\boldsymbol{\beta}}_i) = \sum_{j=1}^{k} \mathbf{Cov}\,(\boldsymbol{H}_j^T \boldsymbol{Y}_j, \boldsymbol{L}^T (\boldsymbol{A}^T \boldsymbol{A})^{-1} \boldsymbol{A}^T \boldsymbol{Y}_i)$$

$$= \sum_{j=1}^{k} \sigma_{ij} \boldsymbol{H}_j^T \boldsymbol{A} (\boldsymbol{A}^T \boldsymbol{A})^{-1} \boldsymbol{L} = \sigma_{ii} \boldsymbol{L}^T (\boldsymbol{A}^T \boldsymbol{A})^{-1} \boldsymbol{L} = \mathbf{Cov}\,(\hat{\theta}_i, \hat{\theta}_i),$$

tai

$$\mathbf{Cov}\,(\tilde{\theta}_i - \hat{\theta}_i, \hat{\theta}_i) = 0, \quad \boldsymbol{V}(\tilde{\theta}_i) \geq \boldsymbol{V}(\hat{\theta}_i).$$

Lygybė pasiekiama tada ir tik tada, kai $\tilde{\theta}_i = \hat{\theta}_i$.

Taigi parametro $\boldsymbol{L}^T \boldsymbol{\beta}_i$ nepaslinktas minimalios dispersijos įvertinys priklauso tik nuo vektoriaus $\boldsymbol{Y}_i$ ir sutampa su įvertiniu, gautu vienmačiu atveju.

Įvertinio $\boldsymbol{L}^T \hat{\boldsymbol{\beta}}_i$ dispersija ir kovariacija su įvertiniu $\boldsymbol{K}^T \hat{\boldsymbol{\beta}}_i$ sutampa su analogiškomis išraiškomis, gautomis vienmačiu atveju. ▲

Kadangi mažiausiųjų kvadratų įvertinių suma taip pat yra MK įvertinys, tai

$$\hat{\theta} = \boldsymbol{L}_1^T \hat{\boldsymbol{\beta}}_1 + ... + \boldsymbol{L}_k^T \hat{\boldsymbol{\beta}}_k \tag{3.2.4}$$

yra parametro $\theta = \boldsymbol{L}_1^T \boldsymbol{\beta}_1 + ... + \boldsymbol{L}_k^T \boldsymbol{\beta}_k$ mažiausiųjų kvadratų įvertinys, kurio dispersija

$$\boldsymbol{V}(\hat{\theta}) = \sum_{i=1}^{k} \sigma_{ii} \boldsymbol{L}_i^T (\boldsymbol{A}^T \boldsymbol{A})^{-1} \boldsymbol{L}_i + \sum_{i \neq j} \sigma_{ij} \boldsymbol{L}_i^T (\boldsymbol{A}^T \boldsymbol{A})^{-1} \boldsymbol{L}_j. \tag{3.2.5}$$

Nežinomų parametrų matricos $\boldsymbol{B}$ mažiausiųjų kvadratų įvertinys

$$\hat{\boldsymbol{B}}^T = (\hat{\boldsymbol{\beta}}_1 \vdots ... \vdots \hat{\boldsymbol{\beta}}_k) = (\boldsymbol{A}^T \boldsymbol{A})^{-1} \boldsymbol{A}^T \mathcal{Y} = (\boldsymbol{A}^T \boldsymbol{A})^{-1} \boldsymbol{A}^T (\boldsymbol{Y}_1 \vdots ... \vdots \boldsymbol{Y}_k), \tag{3.2.6}$$

$$\boldsymbol{V}(\hat{\boldsymbol{\beta}}_i) = \sigma_{ii} (\boldsymbol{A}^T \boldsymbol{A})^{-1}, \quad \mathbf{Cov}\,(\hat{\boldsymbol{\beta}}_i, \hat{\boldsymbol{\beta}}_j) = \sigma_{ij} (\boldsymbol{A}^T \boldsymbol{A})^{-1}.$$

3.2.2 teorema. *Nepaslinktasis kovariacijų matricos* $\boldsymbol{\Sigma}$ *įvertinys yra*

$$\tilde{\boldsymbol{\Sigma}} = \frac{1}{n-m}\boldsymbol{SS}_E = \frac{1}{n-m}[SS_E(i,j)]_{k\times k}. \tag{3.2.7}$$

Čia $SS_E(i,i)$ liekamoji kvadratų sumą, gauta iš i-ojo tiesinio modelio (3.1.1)

$$SS_E(i,i) = (\boldsymbol{Y}_i - \boldsymbol{A}\hat{\boldsymbol{\beta}}_i)^T(\boldsymbol{Y}_i - \boldsymbol{A}\hat{\boldsymbol{\beta}}_i) = \boldsymbol{Y}_i^T\boldsymbol{Y}_i - \boldsymbol{Y}_i^T\boldsymbol{A}\hat{\boldsymbol{\beta}}_i, \tag{3.2.8}$$

o $SS_E(i,j), i \neq j$, vadinamoji *liekamųjų sandaugų suma*

$$SS_E(i,j) = (\boldsymbol{Y}_i - \boldsymbol{A}\hat{\boldsymbol{\beta}}_i)^T(\boldsymbol{Y}_j - \boldsymbol{A}\hat{\boldsymbol{\beta}}_j) = \boldsymbol{Y}_i^T\boldsymbol{Y}_j - \boldsymbol{Y}_i^T\boldsymbol{A}\hat{\boldsymbol{\beta}}_j = \boldsymbol{Y}_i^T\boldsymbol{Y}_j - \boldsymbol{Y}_j^T\boldsymbol{A}\hat{\boldsymbol{\beta}}_i. \tag{3.2.9}$$

Įrodymas. Iš vienamatės tiesinių modelių teorijos (žr. vadovėlio 2 dalį, 1 skyrių) turime, kad dispersijos σ_{ii} nepaslinktasis įvertinys yra $SS_E(i,i)/(n-m)$.

Nagrinėkime atsitiktinį vektorių $\boldsymbol{Y} = \boldsymbol{Y}_i + \boldsymbol{Y}_j$, kurio vidurkis $\mathbf{E}(\boldsymbol{Y}) = \boldsymbol{A}(\boldsymbol{\beta}_i + \boldsymbol{\beta}_j) = \boldsymbol{A}\boldsymbol{\beta}$, o kovariacinė matrica $\boldsymbol{V}(\boldsymbol{Y}) = \sigma^2\boldsymbol{I}$, $\quad \sigma^2 = \sigma_{ii} + \sigma_{jj} + 2\sigma_{ij}$. Šio tiesinio modelio normaliųjų lygčių sistema

$$\boldsymbol{A}^T\boldsymbol{A}\hat{\boldsymbol{\beta}} = \boldsymbol{A}^T\boldsymbol{Y}, \quad \hat{\boldsymbol{\beta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y},$$

o liekamoji kvadratų suma

$$SS_E = (\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}})^T(\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}}) = (\boldsymbol{Y}_i + \boldsymbol{Y}_j - \boldsymbol{A}\hat{\boldsymbol{\beta}}_i - \boldsymbol{A}\hat{\boldsymbol{\beta}}_j)^T(\boldsymbol{Y}_i + \boldsymbol{Y}_j - \boldsymbol{A}\hat{\boldsymbol{\beta}}_i - \boldsymbol{A}\hat{\boldsymbol{\beta}}_j)$$

$$= (\boldsymbol{Y}_i - \boldsymbol{A}\hat{\boldsymbol{\beta}}_i)^T(\boldsymbol{Y}_i - \boldsymbol{A}\hat{\boldsymbol{\beta}}_i) + (\boldsymbol{Y}_j - \boldsymbol{A}\hat{\boldsymbol{\beta}}_j)^T(\boldsymbol{Y}_j - \boldsymbol{A}\hat{\boldsymbol{\beta}}_j) + 2(\boldsymbol{Y}_i - \boldsymbol{A}\hat{\boldsymbol{\beta}}_i)^T(\boldsymbol{Y}_j - \boldsymbol{A}\hat{\boldsymbol{\beta}}_j)$$

$$= SS_E(i,i) + SS_E(j,j) + 2SS_E(i,j).$$

Dispersijos $\sigma^2 = \sigma_{ii} + \sigma_{jj} + 2\sigma_{ij}$ nepaslinktasis įvertinys yra $SS_E/(n-m)$. Kadangi $SS_E(i,i)/(n-m)$ ir $SS_E(j,j)/(n-m)$ yra nepaslinktieji parametrų σ_{ii} ir σ_{jj} įvertiniai, tai $SS_E(i,j)/(n-m)$ yra nepaslinktasis parametro σ_{ij} įvertinys. ▲

3.2.1 pastaba. Tiesiogiai įsitikiname, kad matricą $\boldsymbol{SS}_E = [SS_E(i,j)]_{k\times k}$ galime užrašyti ir tokiu pavidalu:

$$\boldsymbol{SS}_E = \sum_{i=1}^{n} \mathbf{X}_i\mathbf{X}_i^T - \hat{\boldsymbol{B}}\boldsymbol{A}^T\boldsymbol{A}\hat{\boldsymbol{B}}^T. \tag{3.2.10}$$

3.2.1 pavyzdys. (**3.1.1** pavyzdžio tęsinys). Pagal **3.1.1** pavyzdžio duomenis gauname, kad parametro $\boldsymbol{\beta}_i = (\mu_{1i}, ..., \mu_{mi})^T$ MK įvertinys yra

$$\hat{\boldsymbol{\beta}}_i = (\hat{\mu}_{1i}, ..., \hat{\mu}_{mi})^T = (\bar{X}_{1i.}, ..., \bar{X}_{mi.})^T, \quad \bar{X}_{li.} = \frac{1}{n_l}\sum_{j=1}^{n_l} X_{lij},$$

o matricos $\boldsymbol{SS}_E$ elementai

$$SS_E(i,i') = \sum_{l=1}^{m}\sum_{j=1}^{n_l}(X_{lij} - \bar{X}_{li.})(X_{li'j} - \bar{X}_{li'.}), \quad i, i' = 1, ..., k.$$

3.2.2 pavyzdys. (**3.1.2** pavyzdžio tęsinys). Pagal **3.1.2** pavyzdžio duomenis parametro $\boldsymbol{\beta}_i = (\alpha_i, \beta_{1i}, \beta_{2i})^T$ įvertinį gauname spręsdami trijų lygčių sistemą

$$\boldsymbol{A}^T\boldsymbol{A}\hat{\boldsymbol{\beta}}_i = \boldsymbol{A}^T\boldsymbol{Y}_i, \quad \hat{\boldsymbol{\beta}}_i = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y}_i, \quad i = 1, ..., k.$$

Išskleistu pavidalu

$$\begin{pmatrix} n & \sum_j Z_{1j} & \sum_j Z_{2j} \\ \sum_j Z_{1j} & \sum_j Z_{1j}^2 & \sum_j Z_{1j}Z_{2j} \\ \sum_j Z_{2j} & \sum_j Z_{1j}Z_{2j} & \sum_j Z_{2j}^2 \end{pmatrix} \begin{pmatrix} \hat{\alpha}_i \\ \hat{\beta}_{1i} \\ \hat{\beta}_{2i} \end{pmatrix} = \begin{pmatrix} \sum_j X_{ij} \\ \sum_j X_{ij}Z_{1j} \\ \sum_j X_{ij}Z_{2j} \end{pmatrix}.$$

Matricos $\boldsymbol{SS}_E$ elementai yra

$$SS_E(i, i') = \sum_{j=1}^{n}(X_{ij} - \hat{\alpha}_i - \hat{\beta}_{1i}Z_{1i} - \hat{\beta}_{2i}Z_{2i})(X_{i'j} - \hat{\alpha}_{i'} -$$

$$\hat{\beta}_{1i'}Z_{1i'} - \hat{\beta}_{2i'}Z_{2i'}), \quad i, i' = 1, ..., k.$$

3.3. Normaliojo skirstinio atvejis

Norėdami sudaryti kriterijus hipotezėms tikrinti, priimsime prielaidą, kad a. v. $\mathbf{X}_1, ..., \mathbf{X}_n$ turi normaliuosius skirstinius:

$$\mathbf{X}_j \sim N_k(\boldsymbol{B}\boldsymbol{a}_j, \boldsymbol{\Sigma}), \quad j = 1, ..., n; \tag{3.3.1}$$

čia $\boldsymbol{a}_j^T$ yra plano matricos $\boldsymbol{A}$ j-oji eilutė.

3.3.1 teorema. *Jeigu* $|\boldsymbol{\Sigma}| > 0$, $|\boldsymbol{A}^T\boldsymbol{A}| > 0$, $n > k, n > m$, *tai nežinomų parametrų matricos* $\boldsymbol{B}$ *ir kovariacinės matricos* $\boldsymbol{\Sigma}$ *DT įvertiniai yra* $\hat{\boldsymbol{B}}$ *ir* $\boldsymbol{SS}_E/n$. *Čia matrica* $\hat{\boldsymbol{B}}$ *nusakyta (3.2.6), o matricos* $\boldsymbol{SS}_E = [SS_E(i, j)]_{k\times k}$ *elementai apibrėžti (3.2.8), (3.2.9) lygybėmis.*

Įrodymas. Imties $\mathbf{X}_1, ..., \mathbf{X}_n$ tikėtinumo funkcija

$$L = L(\boldsymbol{B}, \boldsymbol{\Sigma}) = [(2\pi)^k|\boldsymbol{\Sigma}|]^{-n/2}\exp\{-\frac{1}{2}\sum_{j=1}^{n}(\mathbf{X}_j - \boldsymbol{B}\boldsymbol{a}_j)^T\boldsymbol{\Sigma}^{-1}(\mathbf{X}_j - \boldsymbol{B}\boldsymbol{a}_j)\}. \tag{3.3.2}$$

Pertvarkome reiškinį po eksponentės ženklu

$$\sum_{j=1}^{n}(\mathbf{X}_j - \boldsymbol{B}\boldsymbol{a}_j)^T\boldsymbol{\Sigma}^{-1}(\mathbf{X}_j - \boldsymbol{B}\boldsymbol{a}_j) = Tr\{\sum_{j=1}^{n}(\mathbf{X}_j - \boldsymbol{B}\boldsymbol{a}_j)^T\boldsymbol{\Sigma}^{-1}(\mathbf{X}_j - \boldsymbol{B}\boldsymbol{a}_j)\}$$

$$= Tr\{\boldsymbol{\Sigma}^{-1}\sum_{j=1}^{n}(\mathbf{X}_j - \boldsymbol{B}\boldsymbol{a}_j)(\mathbf{X}_j - \boldsymbol{B}\boldsymbol{a}_j)^T\}.$$

Panagrinėkime sumą

$$\sum_{j=1}^{n}(\mathbf{X}_j - \boldsymbol{B}\boldsymbol{a}_j)(\mathbf{X}_j - \boldsymbol{B}\boldsymbol{a}_j)^T = \sum_{j=1}^{n}(\mathbf{X}_j - \hat{\boldsymbol{B}}\boldsymbol{a}_j)(\mathbf{X}_j - \hat{\boldsymbol{B}}\boldsymbol{a}_j)^T$$

$$+2\sum_{j=1}^{n}(\mathbf{X}_j-\hat{\boldsymbol{B}}\boldsymbol{a}_j)\boldsymbol{a}_j^T(\hat{\boldsymbol{B}}-\boldsymbol{B})^T+\sum_{j=1}^{n}(\hat{\boldsymbol{B}}-\boldsymbol{B})\boldsymbol{a}_j\boldsymbol{a}_j^T(\hat{\boldsymbol{B}}-\boldsymbol{B})^T.$$

Remdamiesi (3.2.6) ir (3.2.10) išraiškomis įsitikiname, kad pirmas dėmuo sutampa su matrica $\boldsymbol{SS}_E$, o antras dėmuo lygus 0. Trečias dėmuo

$$\sum_{j=1}^{n}(\hat{\boldsymbol{B}}-\boldsymbol{B})\boldsymbol{a}_j\boldsymbol{a}_j^T(\hat{\boldsymbol{B}}-\boldsymbol{B})^T=(\hat{\boldsymbol{B}}-\boldsymbol{B})\boldsymbol{A}^T\boldsymbol{A}(\hat{\boldsymbol{B}}-\boldsymbol{B})^T.$$

Gauname

$$\sum_{j=1}^{n}(\mathbf{X}_j-\boldsymbol{B}\boldsymbol{a}_j)(\mathbf{X}_j-\boldsymbol{B}\boldsymbol{a}_j)^T=\boldsymbol{SS}_E+(\hat{\boldsymbol{B}}-\boldsymbol{B})\boldsymbol{A}^T\boldsymbol{A}(\hat{\boldsymbol{B}}-\boldsymbol{B})^T. \qquad (3.3.3)$$

Įrašę (3.3.4) ir (3.3.5) į tikėtinumo funkcijos išraišką (3.2.3)

$$L=L(\boldsymbol{B},\boldsymbol{\Sigma})=[(2\pi)^k|\boldsymbol{\Sigma}|]^{-n/2}\exp\{-\frac{1}{2}Tr(\boldsymbol{\Sigma}^{-1}\boldsymbol{SS}_E)\}$$

$$\exp\{-\frac{1}{2}Tr(\boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{B}}-\boldsymbol{B})\boldsymbol{A}^T\boldsymbol{A}(\hat{\boldsymbol{B}}-\boldsymbol{B})^T)\} \qquad (3.3.4)$$

matome, kad ji yra dviejų daugiklių sandauga. Pirmas iš jų priklauso tik nuo $\boldsymbol{\Sigma}$, antras įgyja maksimalią reikšmę 1, kai $\boldsymbol{B}=\hat{\boldsymbol{B}}$.

Kaip ir 1.2.1 teoremoje įsitikiname, kad pirmas daugiklis įgyja maksimalią reikšmę, kai $\hat{\boldsymbol{\Sigma}}=\boldsymbol{SS}_E/n$. Tikėtinumo funkcijos maksimumas yra

$$L(\hat{\boldsymbol{B}},\tilde{\boldsymbol{\Sigma}})=(\frac{1}{\sqrt{2\pi}})^{kn}n^{nk/2}(\boldsymbol{SS}_E)^{-n/2}\exp\{-\frac{nk}{2}\}. \qquad (3.3.5)$$

▲

Matome, kad kovariacijų matricos DT įvertinys $\hat{\boldsymbol{\Sigma}}=\boldsymbol{SS}_E/n$ yra paslinktasis. Poslinkį nesunku atitaisyti imant įvertinį $\tilde{\boldsymbol{\Sigma}}=\boldsymbol{SS}_E/(n-m)$.

3.3.2 teorema. *Tegu* $|\boldsymbol{\Sigma}|>0$, $|\boldsymbol{A}^T\boldsymbol{A}|>0$, $n>k$, $n>m$, *tada įvertiniai* $\hat{\boldsymbol{B}}$ *ir* $\hat{\boldsymbol{\Sigma}}$ *yra nepriklausomi. Jeigu matricą* $\hat{\boldsymbol{B}}$ *sudarančius vektorius surašysime į vieną jungtinį vektorių* $(\hat{\boldsymbol{\beta}}_1^T,...,\hat{\boldsymbol{\beta}}_k^T)^T$, *tai šio* mk *dimensijos vektoriaus skirstinys yra normalusis su kovariacine matrica*

$$\begin{pmatrix} \sigma_{11}(\boldsymbol{A}^T\boldsymbol{A})^{-1} & ... & \sigma_{1k}(\boldsymbol{A}^T\boldsymbol{A})^{-1} \\ ... & ... & ... \\ \sigma_{11}(\boldsymbol{A}^T\boldsymbol{A})^{-1} & ... & \sigma_{kk}(\boldsymbol{A}^T\boldsymbol{A})^{-1} \end{pmatrix}.$$

Matrica $(n-m)\hat{\boldsymbol{\Sigma}}=\boldsymbol{SS}_E$ *turi Višarto skirstinį* $W_k(n-m,\boldsymbol{\Sigma})$.

Įrodymas. Kadangi įvertiniai $\hat{\boldsymbol{\beta}}_1,...,\hat{\boldsymbol{\beta}}_k$ yra normaliųjų a. v. tiesinės funkcijos, tai jų skirstiniai yra normalieji, kurių kovariacijų matricų išraiškos pateiktos (3.2.6).

Tegu $\boldsymbol{L} = (L_1, ..., L_k)^T \in \mathbf{R}^k$ yra fiksuotas vektorius. Nagrinėkime a. v. $\boldsymbol{Y} = L_1\boldsymbol{Y}_1 + ... + L_k\boldsymbol{Y}_k$. Jo pirmieji momentai

$$\mathbf{E}(\boldsymbol{Y}) = L_1\boldsymbol{A}\boldsymbol{\beta}_1 + ... + L_k\boldsymbol{A}\boldsymbol{\beta}_k = \boldsymbol{A}\boldsymbol{\beta}, \quad \boldsymbol{\beta} = L_1\boldsymbol{\beta}_1 + ... + L_k\boldsymbol{\beta}_k,$$

$$\boldsymbol{V}(\boldsymbol{Y}) = (\sum_i L_i^2\sigma_{ii} + \sum_{i\neq j} L_iL_j\sigma_{ij})\boldsymbol{I} = \boldsymbol{L}^T\boldsymbol{\Sigma}\boldsymbol{L}\boldsymbol{I} = \sigma_{\boldsymbol{L}}^2\boldsymbol{I} \tag{3.3.6}$$

Turime vienmatį Gauso ir Markovo modelį. Parametro $\boldsymbol{\beta}$ mažiausiųjų kvadratų įvertinys

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}[L_1\boldsymbol{A}^T\boldsymbol{Y}_1 + ... + L_k\boldsymbol{A}^T\boldsymbol{Y}_k] = L_1\hat{\boldsymbol{\beta}}_1 + ... + L_k\hat{\boldsymbol{\beta}}_k \tag{3.3.7}$$

nepriklauso nuo liekamųjų kvadratų sumos

$$(\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}})^T(\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}}) = [\sum_i L_i(\boldsymbol{Y}_i - \boldsymbol{A}\hat{\boldsymbol{\beta}}_i)]^T[\sum_j L_j(\boldsymbol{Y}_j - \boldsymbol{A}\hat{\boldsymbol{\beta}}_j)]$$

$$= \sum_i\sum_j L_iL_j(\boldsymbol{Y}_i - \boldsymbol{A}\hat{\boldsymbol{\beta}}_i)^T(\boldsymbol{Y}_j - \boldsymbol{A}\hat{\boldsymbol{\beta}}_j)$$

$$= \sum_i\sum_j L_iL_jSS_E(i,j) = \boldsymbol{L}^T\boldsymbol{SS}_E\boldsymbol{L}. \tag{3.3.8}$$

Remdamiesi vienmate tiesinių modelių teorija gauname, kad $\hat{\boldsymbol{\beta}}$ ir $\boldsymbol{L}^T\boldsymbol{SS}_E\boldsymbol{L}$ yra nepriklausomi ir

$$\hat{\boldsymbol{\beta}} \sim N_m(\boldsymbol{\beta}, \sigma_{\boldsymbol{L}}^2(\boldsymbol{A}^T\boldsymbol{A})^{-1}), \quad \boldsymbol{L}^T\boldsymbol{SS}_E\boldsymbol{L} \sim \sigma_{\boldsymbol{L}}^2\chi_{n-m}^2. \tag{3.3.9}$$

Remdamiesi Višarto skirstinio apibrėžimu darome išvadą, kad $\boldsymbol{SS}_E \sim W_k(n-m, \boldsymbol{\Sigma})$, o iš pirmos Višarto skirstinio savybės išplaukia, kad $\hat{\boldsymbol{B}}$ ir $\boldsymbol{SS}_E$ yra nepriklausomi. ▲

3.4. Tiesinių hipotezių tikrinimas

Sudarant kriterijus hipotezėms apie parametrų $\boldsymbol{B}$ reikšmes galima išskirti atvejus, kai faktiškai uždavinys tampa vienmačiu. Pavyzdžiui, jeigu tikriname paprastąją hipotezę $\theta = L^T\boldsymbol{\beta}_i = \theta_0$ arba sudėtingąją hipotezę $\boldsymbol{H}\boldsymbol{\beta}_i = \boldsymbol{b}_0$, kurių formuluotėse figūruoja tik parametrų vektorius $\boldsymbol{\beta}_i$, tai pakanka nagrinėti tik i-ąjį Gauso ir Markovo modelį iš (3.1.1). Tokie uždaviniai buvo sprendžiami 2 dalies 2 skyriuje. Analogiškai paprastų ar sudėtinių hipotezių apie parametrų vektorių $\boldsymbol{\beta} = L_1\boldsymbol{\beta}_1 + ... + L_k\boldsymbol{\beta}_k$ tikrinimas taip pat suvedamas į vienmatį atvejį nagrinėjant a. v. $\boldsymbol{Y} = L_1\boldsymbol{Y}_1 + ... + L_k\boldsymbol{Y}_k$ tiesinį modelį (žr. 3.3 skyrelį).

Tarkime, reikia patikrinti prielaidą, kad yra teisingos iš karto visos tiesinės hipotezės:

$$\boldsymbol{H}\boldsymbol{\beta}_i = \boldsymbol{\theta}_i^0, \quad i = 1, ..., k; \tag{3.4.1}$$

čia $\boldsymbol{H}$ dydžio $r \times m$ ir rango r matrica; o $\boldsymbol{\theta}_i^0$ – žinomi vektoriai. Tokio uždavinio sprendimui nepakanka 2 dalies 2 skyriaus vienmatės teorijos. Pavyzdžiui, 3.1.1 pavyzdyje hipotezė $H_i : \mu_{1i} = ... = \mu_{mi}$, arba hipotezė $H : L_1\mu_{11} + ... + L_k\mu_{1k} = ... = L_1\mu_{m1} + ... + L_k\mu_{mk}$ suvedama į viematį atvejį. Tuo tarpu tikrinant hipotezę $H : \boldsymbol{\mu}_1 = ... = \boldsymbol{\mu}_m$, t. y. teisingi visi tvirtinimai $H_1, ..., H_k$, nepakanka vienmatės teorijos. Tiesa, atveju $m = 2$ šis uždavinys buvo išspręstas sudarius kriterijų, grindžiamą Hotelingo T^2 statistika.

Pateiksime analogiją su vienmate normaliojo skirstinio teorija. Tikrinant dviejų vidurkių lygybės hipotezę pagal dvi nepriklausomas normaliųjų a. d. imtis buvo taikomas Stjudento dviejų nepriklausomų imčių kriterijus. Kai imčių skaičius didesnis už 2, vidurkiams palyginti buvo sudarytas Fišerio kriterijus (vienfaktorė dispersinė analizė). Analogiškai daugiamačiu atveju tikrinant dviejų vidurkių vektorių lygybės hipotezę pagal dvi nepriklausomas normaliųjų vektorių imtis taikomas kriterijus, grindžiamas Hotelingo statistika (Stjudento statistikos daugiamatis analogas). Kai nepriklausomų imčių skaičius didesnis už 2 ir reikia tikrinti vidurkių vektorių $\boldsymbol{\mu}_1, ..., \boldsymbol{\mu}_m$ lygybės hipotezę, reikia kriterijaus, kuris būtų vienfaktorės dispersinės analizės kriterijaus daugiamatis analogas.

Parinkę fiksuotą vektorių $\boldsymbol{L} = (L_1, ..., L_k)^T \in \mathbf{R}^k$, vietoje hipotezių (3.4.1) pirmiausia nagrinėkime hipotezę

$$H_{\boldsymbol{L}} : \boldsymbol{H}\boldsymbol{\beta} = \boldsymbol{\theta}^0, \quad \boldsymbol{\beta} = L_1\boldsymbol{\beta}_1 + ... + L_k\boldsymbol{\beta}_k, \quad \boldsymbol{\theta}_0 = L_1\boldsymbol{\theta}_1^0 + ... + L_k\boldsymbol{\theta}_k^0. \quad (3.4.2)$$

remdamiesi a. v. $\boldsymbol{Y} = L_1\boldsymbol{Y}_1 + ... + L_k\boldsymbol{Y}_k$ tiesiniu modeliu:

$$\mathbf{E}(\boldsymbol{Y}) = \boldsymbol{A}\boldsymbol{\beta}, \quad \boldsymbol{V}(\boldsymbol{Y}) = \sigma_{\boldsymbol{L}}^2\boldsymbol{I}. \quad (3.4.3)$$

3.4.1 teorema. *Tegu* $|\boldsymbol{\Sigma}| > 0, |\boldsymbol{A}^T\boldsymbol{A}| > 0$, *Rang*$(\boldsymbol{H}) = r$, $n > k, n > m, m > r$. *Tada hipotezės* $H_{\boldsymbol{L}}$ *tikrinimo kriterijus turi tokį pavidalą: hipotezė atmetama reikšmingumo lygmens* α *kriterijumi, kai teisinga nelygybė*

$$F = \frac{\boldsymbol{L}^T(\boldsymbol{SS}_{EH} - \boldsymbol{SS}_E)\boldsymbol{L}(n-m)}{r\boldsymbol{L}^T\boldsymbol{SS}_E\boldsymbol{L}} > F_\alpha(r, n-m). \quad (3.4.4)$$

Čia matricos $\boldsymbol{SS}_E$ *elementai apibrėžti (3.2.8), (3.2.9) lygybėmis, o matricos* $\boldsymbol{SS}_{EH}$ *elementai turi tokias pačias išraiškas, tik įvertinius* $\hat{\boldsymbol{\beta}}_i$ *reikia pakeisti MK įvertiniais* $\tilde{\boldsymbol{\beta}}_i$, *surastais kai* $H_{\boldsymbol{L}}$ *teisinga.*

Įrodymas. Remiantis 2 dalies 1 skyriumi, kriterijus hipotezei (3.4.2) tikrinti sudaromas tokiu būdu. Randame

$$SS_E = \min_{\boldsymbol{\beta}}(\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{\beta})^T(\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{\beta}) = (\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}})^T(\boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{\beta}}),$$

$$SS_{EH} = \min_{\boldsymbol{H}\boldsymbol{\beta}=\boldsymbol{\theta}_0}(\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{\beta})^T(\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{\beta}) = (\boldsymbol{Y} - \boldsymbol{A}\tilde{\boldsymbol{\beta}})^T(\boldsymbol{Y} - \boldsymbol{A}\tilde{\boldsymbol{\beta}}).$$

Kai hipotezė (3.4.2) teisinga, a. d. SS_E ir $SS_{EH} - SS_E$ yra nepriklausomi ir turi χ^2 skirstinius su $n - m$ ir r laisvės laipsnių. Jeigu hipotezė neteisinga, tai

$SS_{EH} - SS_E$ turi necentrinį χ^2 skirstinį su r laisvės laipsnių ir necentriškumo parametru

$$\lambda = (\boldsymbol{H\beta} - \boldsymbol{\theta}_0)^T[\boldsymbol{H}(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{H}^T]^{-1}(\boldsymbol{H\beta} - \boldsymbol{\theta}_0).$$

Taigi kvadratų suma SS_{EH} išskaidoma į dvi nepriklausomas komponentes:

$$SS_{EH} = SS_E + (SS_{EH} - SS_E). \tag{3.4.5}$$

Pirmosios iš jų skirstinys nepriklauso nuo hipotezės teisingumo, o antroji apibūdina nuokrypį nuo hipotezės. Jų palyginimas leidžia sudaryti kriterijų hipotezei (3.4.2) tikrinti. Hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai

$$F = \frac{(SS_{EH} - SS_E)(n-m)}{rSS_E} > F_\alpha(r, n-m). \tag{3.4.6}$$

Kvadratų sumą SS_E surademe 3.3 skyrelyje

$$SS_E = \boldsymbol{L}^T\boldsymbol{SS}_E\boldsymbol{L};$$

čia $\boldsymbol{SS}_E$ yra liekamoji kvadratų sumų ir sandaugų matrica, apibrėžta (3.2.8) ir (3.2.9) formulėmis. Ji turi Višarto skirstinį $\boldsymbol{SS}_E \sim W_k(n-m, \boldsymbol{\Sigma})$.

Tegu $\tilde{\boldsymbol{\beta}}_i$ yra parametro $\boldsymbol{\beta}_i$ mažiausiųjų kvadratų įvertinys, surastas kai hipotezė $\boldsymbol{H\beta}_i = \boldsymbol{\theta}_i^0$ yra teisinga ir matricos $\boldsymbol{SS}_{EH}$ elementai yra

$$SS_{EH}(i,j) = (\boldsymbol{Y}_i - \boldsymbol{A}\tilde{\boldsymbol{\beta}}_i)^T(\boldsymbol{Y}_j - \boldsymbol{A}\tilde{\boldsymbol{\beta}}_j), \quad i,j = 1,...,k. \tag{3.4.7}$$

Pakartoję 3.3 skyrelio samprotavimus gauname, kad

$$SS_{EH} = \min_{\boldsymbol{H\beta}=\boldsymbol{\theta_0}} (\boldsymbol{Y} - \boldsymbol{A\beta})^T(\boldsymbol{Y} - \boldsymbol{A\beta}) = \boldsymbol{L}^T\boldsymbol{SS}_{EH}\boldsymbol{L},$$

o

$$SS_{EH} - SS_E = \boldsymbol{L}^T(\boldsymbol{SS}_{EH} - \boldsymbol{SS}_E)\boldsymbol{L} \sim \sigma^2_{\boldsymbol{L}}\chi^2(r), \tag{3.4.8}$$

jeigu hipotezė (3.4.2) yra teisinga. Pagal Višarto skirstinio apibrėžimą $\boldsymbol{S}_{EH} - \boldsymbol{SS}_E \sim W_k(r, \boldsymbol{\Sigma})$. Kadangi SS_E ir $SS_{EH} - SS_E$ turi nepriklausomus χ^2 skirstinius, tai remdamiesi Višarto skirstinio pirma savybe gauname, kad matricos $\boldsymbol{SS}_E$ ir $\boldsymbol{SS}_{EH} - \boldsymbol{SS}_E$ turi nepriklausomus Višarto skirstinius. Taigi gavome matricos $\boldsymbol{SS}_{EH}$ išskaidymą į dvi nepriklausomas Višarto matricas:

$$\boldsymbol{SS}_{EH} = \boldsymbol{SS}_E + (\boldsymbol{SS}_{EH} - \boldsymbol{SS}_E), \tag{3.4.9}$$

kurį galima laikyti skaidinio (3.4.5) daugiamačiu analogu. Pirmoji matrica $\boldsymbol{SS}_E \sim W_k(n-m, \boldsymbol{\Sigma})$ nepriklauso nuo hipotezių (3.4.1) teisingumo. Antroji apibūdina nuokrypį nuo hipotezių: jei hipotezės teisingos, tai $\boldsymbol{SS}_{EH} - \boldsymbol{SS}_E \sim W_k(r, \boldsymbol{\Sigma})$, o jei neteisingos, tai ji turi necentrinį Višarto skirstinį, nes (3.4.8) turi necentrinį χ^2 skirstinį.

Naujais žymenimis kriterijus (3.4.6) turi tokį pavidalą

$$F = \frac{\boldsymbol{L}^T(\boldsymbol{SS}_{EH} - \boldsymbol{SS}_E)\boldsymbol{L}(n-m)}{r\boldsymbol{L}^T\boldsymbol{SS}_E\boldsymbol{L}} > F_\alpha(r, n-m). \qquad \blacktriangle$$

3.4.1 pastaba. Jeigu (3.4.5) dešinės pusės dėmenis padalinsime iš atitinkamų laisvės laipsnių, tai esant teisingai hipotezei $H_{\boldsymbol{L}}$ gausime du nepriklausomus nepaslinktuosius dispersijos $\sigma^2_{\boldsymbol{L}}$ įvertinius. Statistika F (3.4.6) lygybėje ir yra šių dviejų dispersijos įvertinių santykis (nuo to ir kilęs dispersinės analizės pavadinimas). Jeigu hipotezė teisinga, tai antrojo dėmens indėlis (3.4.5) lygybėje neturėtų būti didelis. Galima kriterijų sudaryti ir kitu būdu, palyginant pirmojo (3.4.5) lygybės dėmens indėlį į bendrą sumą SS_{EH}. Pasinaudokime a. d., turinčio Fišerio skirstinį, ir a. d., turinčio beta skirstinį, sąryšiu: jeigu $F \sim F(r, n-m)$, tai $B = 1/(1 + rF/(n-m)) \sim Be((n-m)/2, r/2)$. Gauname ekvivalentų (3.4.10) kriterijų: hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai teisinga nelygybė:

$$B = \frac{1}{1 + rF/(n-m)} = \frac{\boldsymbol{L}^T \boldsymbol{SS}_E \boldsymbol{L}}{\boldsymbol{L}^T \boldsymbol{SS}_{EH} \boldsymbol{L}} < B_{1-\alpha}\left(\frac{n-m}{2}, \frac{r}{2}\right), \qquad (3.4.10)$$

čia $B_\alpha(\nu_1, \nu_2)$ yra beta skirstinio, kurio parametrai ν_1, ν_2, lygmens α kritinė reikšmė.

Lieka sudaryti kriterijų, grindžiamą dviejų nepriklausomų Višarto matricų iš (3.4.9) palyginimu.

Vietoje hipotezių (3.4.1) rinkinio nagrinėjome hipotezę (3.4.2) parinkę fiksuotą vektorių $\boldsymbol{L} \in \mathbf{R}^k$. Jeigu hipotezės $\boldsymbol{H}\boldsymbol{\beta}_i = \boldsymbol{\theta}_0$, $\forall i = 1, ..., k$, yra teisingos, tai (3.4.2) hipotezė $H_{\boldsymbol{L}}$ yra teisinga su bet kuriuo $\boldsymbol{L} \in \mathbf{R}^k$. Priešingu atveju atsiras toks $\boldsymbol{L} \in \mathbf{R}^k$, kad hipotezė $H_{\boldsymbol{L}}$ bus atmesta.

Parinkime $\boldsymbol{L}$ taip, kad statistika B įgytų kuo mažesnę reikšmę. Tada gauname statistiką

$$\lambda_1 = \min_{\boldsymbol{L}} B = \min_{\boldsymbol{L}} \frac{\boldsymbol{L}^T \boldsymbol{SS}_E \boldsymbol{L}}{\boldsymbol{L}^T \boldsymbol{SS}_{EH} \boldsymbol{L}},$$

t. y. λ_1 yra mažiausioji charakteringosios lygties

$$|\boldsymbol{SS}_E - \lambda \boldsymbol{SS}_{EH}| = 0$$

šaknis. Atrodytų, kad kriterijų reikėtų grįsti statistika λ_1. Tačiau matricų $\boldsymbol{SS}_E$ ir $\boldsymbol{SS}_{EH}$ elementai yra atsitiktiniai, todėl jų palyginimo kriterijus remiantis tik mažiausiąja šaknimi λ_1 gali būti nestabilus. Praktiškai naudojami kriterijai, kurių statistikos yra simetrinės šios charakteringosios lygties šaknų $\lambda_1, ..., \lambda_k$ funkcijos. Pavyzdžiui,

$$U = \prod_{i=1}^k \lambda_i = \frac{|\boldsymbol{SS}_E|}{|\boldsymbol{SS}_{EH}|}, \qquad (3.4.11)$$

$$\sum_{i=1}^k ((1-\lambda_i)/\lambda_i), \quad \prod_{i=1}^k ((1-\lambda_i)/\lambda_i).$$

3.4.2 pastaba. Dažniausiai naudojamas kriterijus, grindžiamas statistika (3.4.11). Argumentu jo naudai yra tai, kad jis ekvivalentus tikėtinumų santykio kriterijui. Iš tikrųjų, analogiškai skyreliui 3.3 įrodoma, kad kovariacinės matricos

DT įvertinys, kai (3.4.1) teisinga, yra $\hat{\mathbf{\Sigma}}^* = \boldsymbol{SS}_{EH}/n$, o tikėtinumo funkcijos maksimumas turi (3.3.7) pavidalą, kai vietoje $\hat{\mathbf{\Sigma}}$ įrašyta $\hat{\mathbf{\Sigma}}^*$. Taigi tikėtinumų santykio statistika

$$\Lambda = \frac{\max_{\boldsymbol{H\beta_i=\theta_0}} L(\boldsymbol{B}, \mathbf{\Sigma})}{\max L(\boldsymbol{B}, \mathbf{\Sigma})} = \left(\frac{|\hat{\mathbf{\Sigma}}|}{|\hat{\mathbf{\Sigma}}^*|}\right)^{n/2} = \left(\frac{|\boldsymbol{SS}_E|}{|\boldsymbol{SS}_{EH}|}\right)^{n/2}$$

yra ekvivalenti statistikai (3.4.11).

3.4.1 pavyzdys. (**3.1.1** pavyzdžio tęsinys). Pagal **3.1.1** pavyzdžio duomenis reikia patikrinti prielaidą, kad nagrinėjamos m kviečių veislių nesiskiria pagal jokius nagrinėjamus požymius. Kitaip tariant, reikia patikrinti prielaidas, kad vidurkiai $\mu_{1i} = ... = \mu_{mi}$ su visais $i = 1, ..., k$. Šias prielaidas galima suformuluoti (3.4.1) hipotezių pavidalu. Tegu matricos $\boldsymbol{H}$ j-oji eilutė yra tokia: $(0, ...0, 1, 0, ..., -1)$; čia 1 parašytas j-oje vietoje, $j = 1, ..., m-1$. Tada reikia patikrinti, kad galioja hipotezės $H_i : \boldsymbol{H\beta}_i = \mathbf{0}$ su visais $i = 1, ..., k$.

Parametro $\boldsymbol{\beta}_i$ MK įvertinys $\tilde{\boldsymbol{\beta}}_i$, kai hipotezės H_i teisingos, yra

$$\tilde{\boldsymbol{\beta}}_i = (\bar{X}_{.i.}, ..., \bar{X}_{.i.})^T, \quad \bar{X}_{.i.} = \frac{1}{n}\sum_{l=1}^{m}\sum_{j=1}^{n_l} X_{lij},$$

o matricos $\boldsymbol{SS}_{EH}$ elementai

$$SS_{EH}(i, i') = \sum_{l=1}^{m}\sum_{j=1}^{n_l}(X_{lij} - \bar{X}_{.i.})(X_{li'j} - \bar{X}_{.i'.}), \quad i, i' = 1, ..., k.$$

Statistiką U iš (3.4.11) gauname padaliję matricos $\boldsymbol{SS}_E$ iš pratimo **3.2.1** determinantą iš matricos $\boldsymbol{SS}_{EH}$ determinanto.

3.4.2 pavyzdys. (**3.1.2** pavyzdžio tęsinys.) Reikia patikrinti prielaidą, kad kovariantės Z_1 ir Z_2 neturi įtakos jokiems tiriamiems pacientų požymių skirstiniams. Kitaip tariant, reikia patikrinti hipotezes, kad regresijų koeficientai $\beta_{1i} = \beta_{2i} = 0$ su visais $i = 1, ..., k$. Kai hipotezės teisingos, tereikia įvertinti parametrus $\alpha_i, i = 1, ..., k$. Šių parametrų MK įvertiniai yra

$$\tilde{\alpha}_i = \bar{X}_{i.} = \sum_{j=1}^{n} X_{ij}, \quad i = 1, ..., k,$$

o matricos $\boldsymbol{SS}_{EH}$ elementai

$$SS_{EH}(i, i') = \sum_{j=1}^{n}(X_{ij} - \bar{X}_{i.})(X_{i'j} - \bar{X}_{i'.}), \quad i, i' = 1, ..., k.$$

Statistiką U iš (3.4.11) gauname padaliję matricos $\boldsymbol{SS}_E$ iš pratimo **3.2.2** determinantą iš matricos $\boldsymbol{SS}_{EH}$ determinanto.

3.5. Tikėtinumų santykio statistikos savybės

3.5.1. Tikėtinumų santykio statistikos momentai

Ieškosime statistikos

$$U = \Lambda^{2/n} = \frac{|\boldsymbol{SS}_E|}{|\boldsymbol{SS}_{EH}|}$$

momento $\mathbf{E}(U^h)$, kai hipotezės (3.4.1) yra teisingos.

3.5.1 teorema. *Jeigu* $|\boldsymbol{\Sigma}| > 0$, $n - m \geq k, m > r$ *ir hipotezė (3.4.1) yra teisinga, tai momentas*

$$\mathbf{E}(U^h) = \prod_{i=1}^{k} \left\{ \frac{\Gamma(\frac{n-m+2h+1-i}{2})\Gamma(\frac{n-m+r+1-i}{2})}{\Gamma(\frac{n-m+1-i}{2})\Gamma(\frac{n-m+r+2h+1-i}{2})} \right\}. \tag{3.5.1}$$

Įrodymas. Kai hipotezė (3.4.1) yra teisinga, statistiką U galima užrašyti šitaip

$$U = \frac{|\boldsymbol{S}_1|}{|\boldsymbol{S}_1 + \boldsymbol{S}_2|}, \quad \boldsymbol{S}_1 = \sum_{i=1}^{n-m} \boldsymbol{Z}_i \boldsymbol{Z}_i^T, \quad \boldsymbol{S}_2 = \sum_{i=1}^{r} \boldsymbol{Y}_i \boldsymbol{Y}_i^T;$$

čia $\boldsymbol{Z}_1, ..., \boldsymbol{Z}_{n-m}, \boldsymbol{Y}_1, ..., \boldsymbol{Y}_r$ yra nepriklausomi vienodai pagal $N_k(\mathbf{0}, \boldsymbol{\Sigma})$ pasiskirstę atsitiktiniai vektoriai. Todėl $\boldsymbol{S}_1$ ir $\boldsymbol{S}_2$ yra nepriklausomi ir turi Višarto skirstinius $\boldsymbol{S}_1 \sim W_k(n-m, \boldsymbol{\Sigma})$, $\boldsymbol{S}_2 \sim W_k(r, \boldsymbol{\Sigma})$.

Naudodami Višarto skirstinio tankį (1.6.13), galime užrašyti

$$\mathbf{E}(U^h) = \int \cdots \int \frac{|\boldsymbol{S}_1|^h}{|\boldsymbol{S}_1 + \boldsymbol{S}_2|^h} f(\boldsymbol{S}_1 | k, n-m, \boldsymbol{\Sigma}) f(\boldsymbol{S}_2 | k, r, \boldsymbol{\Sigma}) d\boldsymbol{S}_1 d\boldsymbol{S}_2$$

$$= \int \cdots \int \frac{|\boldsymbol{S}_1|^h}{|\boldsymbol{S}_1 + \boldsymbol{S}_2|^h} \frac{|\boldsymbol{S}_1|^{(n-m-k-1)/2}}{K(k, n-m, \boldsymbol{\Sigma})} e^{-\frac{1}{2} Tr(\boldsymbol{S}_1 \boldsymbol{\Sigma}^{-1})} f(\boldsymbol{S}_2 | k, r, \boldsymbol{\Sigma}) d\boldsymbol{S}_1 d\boldsymbol{S}_2$$

$$= \frac{K(k, n-m+2h, \boldsymbol{\Sigma})}{K(k, n-m, \boldsymbol{\Sigma})} \int \cdots \int |\boldsymbol{S}_1 + \boldsymbol{S}_2|^{-h} \left\{ \frac{|\boldsymbol{S}_1|^{(n-m+2h-k-1)/2}}{K(k, n-m+2h, \boldsymbol{\Sigma})} \times \right.$$

$$\left. \times\, e^{-\frac{1}{2} Tr(\boldsymbol{S}_1 \boldsymbol{\Sigma}^{-1})} \right\} f(\boldsymbol{S}_2 | k, r, \boldsymbol{\Sigma}) d\boldsymbol{S}_1 d\boldsymbol{S}_2$$

$$= \frac{K(k, n-m+2h, \boldsymbol{\Sigma})}{K(k, n-m, \boldsymbol{\Sigma})} \int \cdots \int |\boldsymbol{S}_1 + \boldsymbol{S}_2|^{-h} f(\boldsymbol{S}_1 | k, n-m+2h, \boldsymbol{\Sigma}) \times$$

$$\times f(\boldsymbol{S}_2 | k, r, \boldsymbol{\Sigma}) d\boldsymbol{S}_1 d\boldsymbol{S}_2.$$

Po integralo ženklu turime Višarto skirstinio $W_k(n-m+2h, \boldsymbol{\Sigma})$ tankį $f(\boldsymbol{S}_1 | k, n-m+2h, \boldsymbol{\Sigma})$, padaugintą iš Višarto skirstinio tankio $f(\boldsymbol{S}_2 | k, r, \boldsymbol{\Sigma})$, ir daugiklį $|\boldsymbol{S}_1 + \boldsymbol{S}_2|^{-h}$. Kitaip tariant, integralas reiškia momentą $\mathbf{E}(|\boldsymbol{S}_1 + \boldsymbol{S}_2|^{-h})$. Tačiau šį momentą galime rasti ir kitaip, remdamiesi tuo, kad $\boldsymbol{S} = \boldsymbol{S}_1 + \boldsymbol{S}_2 \sim W_k(n-m+2h+r, \boldsymbol{\Sigma})$. Gauname

$$\mathbf{E}(U^h) = \frac{K(k, n-m+2h, \boldsymbol{\Sigma})}{K(k, n-m, \boldsymbol{\Sigma})} \int \cdots \int |\boldsymbol{S}|^{-h} \frac{|\boldsymbol{S}|^{(n-m+2h+r-k-1)/2}}{K(k, n-m+r+2h, \boldsymbol{\Sigma})} \times$$

$$\times e^{-\frac{1}{2} Tr(\boldsymbol{S} \boldsymbol{\Sigma}^{-1})} d\boldsymbol{S} = \frac{K(k, n-m+2h, \boldsymbol{\Sigma})}{K(k, n-m, \boldsymbol{\Sigma})} \frac{K(k, n-m+r, \boldsymbol{\Sigma})}{K(k, n-m+r+2h, \boldsymbol{\Sigma})},$$

nes likęs integralas lygus 1.

Pasinaudoję normuojančio daugiklio išraiška iš (1.6.14) gauname (3.5.1). ▲

3.5.2. Tikėtinumų santykio statistikos skirstiniai

Momentas (3.5.1) egzistuoja su bet kuriuo $h > -1/2$. Momentą (3.5.1) galime interpretuoti kaip a. d. $\ln U$ momentus generuojančią funkciją

$$\psi(h) = \mathbf{E}(U^h) = \mathbf{E}(e^{h \ln U}).$$

Kadangi $\psi(h)$ apibrėžta intervale $|h| < 1/2$, apimančiame tašką $h = 0$, tai ji visiškai nusako a. d. $\ln U$ ir a. d. U skirstinį.

3.5.2 teorema. *Atsitiktinio dydžio U skirstinys sutampa su nepriklausomų a. d., turinčių beta skirstinius, sandaugos:*

$$U \overset{d}{\sim} \prod_{i=1}^{k} \xi_i, \quad \ln U \overset{d}{\sim} \sum_{i=1}^{k} \ln \xi_i \tag{3.5.2}$$

skirstiniu; čia $\xi_1, ..., \xi_k$ yra nepriklausomi a. d., turintys beta skirstinius $\xi_i \sim Be((n-m+1-i)/2, r/2), i = 1, ..., k$.

Įrodymas. Tegu a. d. $\eta \sim Be(a/2, b/2)$. Tada momentas $\mathbf{E}(\eta^h)$ yra

$$\mathbf{E}(\eta^h) = \frac{\Gamma((a+b)/2)}{\Gamma(a/2)\Gamma(b/2)} \int_0^1 x^{\frac{a}{2}+h-1}(1-x)^{\frac{b}{2}-1} dx =$$

$$= \frac{\Gamma((a+b)/2)\Gamma((a+2h)/2)}{\Gamma((a+b+2h)/2)\Gamma(a/2)}.$$

Palyginę šią išraišką su (3.5.1) matome, kad skliausteliuose yra momentas $\mathbf{E}(\xi_i^h)$, kai $\xi_i \sim Be((n-m+1-i)/2, r/2)$. Tada

$$\mathbf{E}(U^h) = \prod_{i=1}^{k} \mathbf{E}(\xi_i^h) = \mathbf{E}(\prod_{i=1}^{k} \xi_i)^h. \tag{3.5.3}$$

Kadangi a. d. $U, \xi_1, ..., \xi_k$ momentai visiškai nusako skirstinius, tai iš (3.5.3) išplaukia, kad $\xi_1, ..., \xi_k$ nepriklausomi ir (3.5.2) yra teisinga. ▲

Tam tikrais atvejais (3.5.2) sąryšį galima supaprastinti.

3.5.3 teorema. *Tarkime, kad $\nu = n - m + r$ yra fiksuotas. Tada momento (3.5.1) išraiškoje k ir r galima sukeisti vietomis*

$$\prod_{i=1}^{k} \left\{ \frac{\Gamma(\frac{\nu-r+1-i+2h}{2})\Gamma(\frac{\nu+1-i}{2})}{\Gamma(\frac{\nu-r+1-i}{2})\Gamma(\frac{\nu+2h+1-i}{2})} \right\} = \prod_{i=1}^{r} \left\{ \frac{\Gamma(\frac{\nu-k+2h+1-i}{2})\Gamma(\frac{\nu+1-i}{2})}{\Gamma(\frac{\nu-k+1-i}{2})\Gamma(\frac{\nu+2h+1-i}{2})} \right\}. \tag{3.5.4}$$

Kitaip tariant, atsitiktinio dydžio U skirstinys sutampa su nepriklausomų a. d., turinčių beta skirstinius, sandaugos:

$$U \overset{d}{\sim} \prod_{i=1}^{r} \zeta_i \quad \ln U \overset{d}{\sim} \sum_{i=1}^{r} \ln \zeta_i \tag{3.5.5}$$

skirstiniu; čia $\zeta_1, ..., \zeta_r$ nepriklausomi a. d., turintys beta skirstinius $\zeta_i \sim Be((\nu-k+1-i)/2, k/2)$, $i = 1, ..., r$.

Įrodymas. Tegu $r < k$. Pertvarkome (3.5.4) lygybės kairiąją pusę

$$\prod_{i=1}^{r} \frac{\Gamma(\frac{\nu+1-i}{2})}{\Gamma(\frac{\nu+1+2h-i}{2})} \prod_{i=r+1}^{k} \frac{\Gamma(\frac{\nu+1-i}{2})}{\Gamma(\frac{\nu+1+2h-i}{2})} \prod_{i=1}^{k-r} \frac{\Gamma(\frac{\nu-r+1+2h-i}{2})}{\Gamma(\frac{\nu-r+1-i}{2})} \prod_{i=k-r+1}^{k} \frac{\Gamma(\frac{\nu-r+1+2h-i}{2})}{\Gamma(\frac{\nu-r+1-i}{2})}$$

$$= \prod_{i=1}^{r} \frac{\Gamma(\frac{\nu+1-i}{2})}{\Gamma(\frac{\nu+1+2h-i}{2})} \left\{ \prod_{i=1}^{k-r} \frac{\Gamma(\frac{\nu-r+1-i}{2})}{\Gamma(\frac{\nu-r+1+2h-i}{2})} \prod_{i=1}^{k-r} \frac{\Gamma(\frac{\nu-r+1+2h-i}{2})}{\Gamma(\frac{\nu-r+1-i}{2})} \right\} \prod_{i=1}^{r} \frac{\Gamma(\frac{\nu-k+1+2h-i}{2})}{\Gamma(\frac{\nu-k+1-i}{2})}.$$

Kadangi apskliaustas reiškinys lygus 1, tai gauname (3.5.4) lygybės dešiniąją pusę.

Jeigu $r > k$, tai analogiškai pertvarkydami (3.5.4) lygybės dešiniąją pusę gauname (3.5.4) lygybės kairiąją pusę. ▲

Natūralu naudoti tą iš formulių (3.5.2) ar (3.5.5), kurioje yra mažiau daugiklių.

3.5.4 teorema. *Jeigu $k = 2s$ lyginis, tai a. d. U skirstinys sutampa su skirstiniu nepriklausomų a. d. sandaugos, kurioje yra s daugiklių:*

$$U \stackrel{d}{\sim} \prod_{j=1}^{s} \eta_j^2, \tag{3.5.6}$$

čia $\eta_1, ..., \eta_s$ yra nepriklausomi a. d., turintys beta skirstinius $\eta_j \sim Be(n - m + 1 - 2j, r)$, $j = 1, ..., s$. Jeigu $k = 2s + 1$ nelyginis, tai U skirstinys sutampa su skirstiniu nepriklausomų a. d. sandaugos, kurioje yra $s + 1$ daugiklis:

$$U \stackrel{d}{\sim} \prod_{j=1}^{s} \eta_j^2 \xi_{2s+1}; \tag{3.5.7}$$

čia $\eta_1, ..., \eta_s$ yra nepriklausomi a. d., turintys beta skirstinius $\eta_j \sim Be(n-m+1-2j, r)$, $j = 1, ..., s$, o nepriklausantis nuo jų a. d. $\xi_{2s+1} \sim Be((n-m-2s)/2, r/2)$.

Įrodymas. Naudosime gama funkcijos nuo dvigubo argumento išraišką

$$\Gamma(\alpha + \frac{1}{2})\Gamma(\alpha + 1) = \sqrt{\pi}\Gamma(2\alpha + 1)2^{-2\alpha}. \tag{3.5.8}$$

Sudauginę (3.5.1) trupmenos skaitiklyje daugiklius, kai $i = 2j - 1$ ir $i = 2j$, gauname

$$\Gamma\left(\frac{n - m + 2h - 2j + 2}{2}\right)\Gamma\left(\frac{n - m + 2h - 2j + 1}{2}\right) =$$

$$= \Gamma(n - m + 2h - 2j + 1)\sqrt{\pi}2^{(n-m+2h-2j)}.$$

Analogiškai sudauginę poromis kitus (3.5.1) skaitiklio ir vardiklio daugiklius, gauname

$$\mathbf{E}(U^h) = \prod_{j-1}^{s} \left\{ \frac{\Gamma(n - m + 2h + 1 - 2j)\Gamma(n - m + r + 1 - 2j)}{\Gamma(n - m + 1 - 2j)\Gamma(n - m + r + 1 + 2h - 2j)} \right\}.$$

Kaip ir 3.5.2 teoremoje įsitikiname, kad apskliaustas reiškinys yra momentas $\mathbf{E}(\eta_j^2)^h = \mathbf{E}(\eta_j^{2h})$, kai $\eta_j \sim Be(n-m+1-2j, r)$. Taigi

$$\mathbf{E}(U^h) = \prod_{i=1}^{s} \mathbf{E}(\eta_i^2)^h = \left(\prod_{i=1}^{s} \eta_i^2\right)^h,$$

ir (3.5.6) lygybė įrodyta.

Pertvarkę $2s$ pirmųjų (3.5.1) daugiklių pagal (3.5.6) ir palikę paskutinį daugiklį $i = 2s+1$ pagal (3.5.5), gausime (3.5.7). ▲

Jeigu $r < k$, tai natūralu naudoti (3.5.4) išraišką, kurioje daugiklių skaičių irgi galime sumažinti remiantis (3.5.8).

3.5.5 teorema. *Jeigu $r = 2s$ lyginis, tai a. d. U skirstinys sutampa su skirstiniu nepriklausomų a. d. sandaugos, kurioje yra s daugiklių:*

$$U \overset{d}{\sim} \prod_{j=1}^{s} \theta_j^2, \tag{3.5.9}$$

čia $\theta_1, ..., \theta_s$ yra nepriklausomi a. d., turintys beta skirstinius $\theta_j \sim Be(\nu - k + 1 - 2j, k)$, $j = 1, ..., s$. Jeigu $r = 2s+1$ nelyginis, tai U skirstinys sutampa su skirstiniu nepriklausomų a. d. sandaugos, kurioje yra $s+1$ daugiklis:

$$U \overset{d}{\sim} \prod_{j=1}^{s} \theta_j^2 \zeta_{2s+1}; \tag{3.5.10}$$

čia $\theta_1, ..., \theta_s$ yra nepriklausomi a. d., turintys beta skirstinius $\theta_j \sim Be(\nu-k+1-2j, k)$, $j = 1, ..., s$, o nepriklausantis nuo jų a. d. $\zeta_{2s+1} \sim Be((\nu-k-2s)/2, k/2)$.

Įrodymas. Analogiškas 3.5.4 teoremos įrodymui.

3.5.3. Tikėtinumų santykio statistikos tam tikri atvejai.

1. *Atvejis* $k = 1$. Stebėjimai $\boldsymbol{Y} = (X_1, ..., X_n)^T$ yra nepriklausomi vienmačiai a. d. su vienodomis dispersijomis, o vidurkio kitimą apibūdina m-matis parametrų vektorius $\boldsymbol{\beta} = (\beta_1, ..., \beta_m)^T$, t. y.

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{\beta} + \boldsymbol{e}, \quad \boldsymbol{e} \sim N_n(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$$

Pagal 3.5.2 teoremą $U \overset{d}{\sim} \xi_1 \sim Be((n-m)/2, r/2)$. Tada remdamiesi beta ir Fišerio skirstinių sąryšiu gauname, kad a. d.

$$F = \frac{(1-U)(n-m)}{rU} \overset{d}{\sim} \frac{(1-\xi_1)(n-m)}{r\xi_1} \sim F(r, n-m).$$

Hipotezė $H : \boldsymbol{H}\boldsymbol{\beta} = \boldsymbol{\theta}_0, Rang(\boldsymbol{H}) = r \leq m$ yra atmetama reikšmingumo lygmens α kriterijumi, kai

$$F > F_\alpha(r, n-m). \tag{3.5.11}$$

Kadangi

$$U = \frac{SS_E}{SS_{EH}} = \frac{SS_E}{SS_E + (SS_{EH} - SS_E)},$$

tai

$$F = \frac{(1-U)(n-m)}{rU} = \frac{(SS_{EH} - SS_E)(n-m)}{rSS_E}.$$

Kriterijus, grindžiamas statistika U, yra ekvivalentus kriterijui, kuris buvo gautas nagrinėjant vienmačius Gauso ir Markovo modelius.

2. *Atvejis* $r = 1$. Matricos $\boldsymbol{H}$ rangas (3.4.1) lygybėse yra lygus 1. Remiantis 3.5.3 teorema

$$U \overset{d}{\sim} \zeta \sim Be(\frac{\nu - k}{2}, \frac{k}{2}), \quad \nu = n - m + r.$$

Pereidami prie Fišerio skirstinio gauname kriterijų

$$F = \frac{(1-U)(\nu - k)}{kU} > F_\alpha(k, \nu - k). \tag{3.5.12}$$

Pavyzdžiui, jei $m = 2$ ir $\boldsymbol{H} = (1, -1), r = Rang(\boldsymbol{H}) = 1, \boldsymbol{\theta}_0 = \mathbf{0}$, tai turime dviejų vidurkių vektorių palyginimo uždavinį. Nesunku patikrinti, kad kriterijus (3.5.12) sutampa su kriterijumi (2.2.5), kuris buvo gautas naudojant Hotelingo statistiką (žr. 3.4 pratimą). Jeigu $m > 2$, tai kriterijus (3.5.12) yra ekvivalentus kriterijui, gautam 2.3 skyrelyje.

3. *Atvejis* $k = 2$. Stebimi dvimačiai nepriklausomi normalieji vektoriai, turintys vienodas kovariacines matricas. Remiantis 3.5.4 teorema

$$\sqrt{U} \overset{d}{\sim} \eta_1 \sim Be(n - m - 1, r).$$

Tada perėję prie Fišerio skirstinio gauname

$$F = \frac{(1-\sqrt{U})(n-m-1)}{r\sqrt{U}} \overset{d}{\sim} \frac{(1-\eta_1)(n-m-1)}{r\eta_1} \sim F(2r, 2(n-m-1)). \tag{3.5.13}$$

Hipotezės (3.4.1) atmetamos reikšmingumo lygmens α kriterijumi, kai

$$F > F_\alpha(2r, 2(n-m-1)). \tag{3.5.14}$$

4. *Atvejis* $r = 2$. Matricos $\boldsymbol{H}$ rangas (3.4.1) lygus 2. Remiantis 3.5.5 teorema

$$\sqrt{U} \overset{d}{\sim} \theta_1 \sim Be(\nu - k - 1, k), \quad \nu = n - m + r.$$

Taigi

$$F = \frac{(1-\sqrt{U})(\nu-k-1)}{k\sqrt{U}} \overset{d}{\sim} \frac{(1-\theta_1)(\nu-k-1)}{k\theta_1} \sim F(2k, 2(\nu-k-1)). \tag{3.5.15}$$

Hipotezės (3.4.1) atmetamos reikšmingumo lygmens α kriterijumi, kai

$$F > F_\alpha(2k, 2(n-m+r-k-1)). \tag{3.5.16}$$

5. *Atvejis $k=3$, $k=4$ (arba $r=3$, $r=4$).* Remiantis 3.5.2 ir 3.5.3 teoremomis $k=3$ atveju atsitiktinis dydis $U \overset{d}{\sim} Z_1^2 Z_2$; čia Z_1 ir Z_2 yra nepriklausomi a. d., turintys beta skirstinius $Z_1 \sim Be(n-m-1,r)$, $Z_2 \sim Be((n-m-4)/2, r/2)$. Pažymėję a. d. Z_1 ir Z_2 tankius atitinkamai $f_1(z_1)$ ir $f_2(z_2)$, gausime a. d. U pasiskirstymo funkcijos išraišką

$$G(u) = \mathbf{P}\{U \le u\} = \int_0^u f_1(z_1)dz_1 + \int_u^1 f_2(z_2)[\int_0^{\sqrt{u/z_2}} f_1(z_1)dz_1]dz_2. \tag{3.5.17}$$

Pirmasis dėmuo yra beta skirstinio $Be(n-m-1,r)$ pasiskirstymo funkcijos reikšmė taške u. Antrąjį dėmenį, atlikus integravimą, galima užrašyti išreikštine argumento u funkcija. Turėdami pasiskirstymo funkciją, galima rasti a. d. U kritinę reikšmę $u_{1-\alpha}$. Tada hipotezė (3.4.1) atmetama, kai

$$U < u_{1-\alpha}. \tag{3.5.18}$$

Atveju $k=4$ atsitiktinis dydis $U \overset{d}{\sim} Z_1^2 Z_2^2$; čia Z_1 ir Z_2 yra nepriklausomi a. d., turintys beta skirstinius $Z_1 \sim Be(n-m-1,r)$, $Z_2 \sim Be(n-m-3,r)$. Vėl pažymėję a. d. Z_1 ir Z_2 tankius atitinkamai $f_1(z_1)$ ir $f_2(z_2)$, gausime a. d. U pasiskirstymo funkcijos išraišką

$$G(u) = \mathbf{P}\{U \le u\} = \int_0^{\sqrt{u}} f_1(z_1)dz_1 + \int_{\sqrt{u}}^1 f_2(z_2)[\int_0^{\sqrt{u}/z_2} f_1(z_1)dz_1]dz_2. \tag{3.5.19}$$

Tam tikrais atvejais (pavyzdžiui, kai $k = 3,\ 4$), antrąjį dėmenį, atlikus integravimą, galima užrašyti išreikštine argumento u funkcija. Kitais atvejais (3.5.19) antrojo dėmens integravimą galima atlikti skaitiniais metodais. Tada skaitiniais metodais galima rasti ir kritinę reikšmę $u_{1-\alpha}$ ir kriterijų (3.5.18).

Atvejai $r=3, r=4$ nagrinėjami analogiškai remiantis 3.5.4 ir 3.5.5 teoremomis.

6. *Atvejis $k>4$, $r>4$.* Jeigu $k>4$ ir $r>4$, tai (3.5.17) ir (3.5.18) analogų integravimas skaitiniais metodais gali būti sunkiai realizuojamas. Tada galima naudoti modeliavimo metodą (žr.3.5.5 skyrelį).

3.5.1 pavyzdys (**2.2.1** pavyzdžio tęsinys). Pamatuotas trijų spindulių srovės stiprumas $n_3 = 18$ Panevėžio gamyklos kineskopų, kurie buvo pagaminti kitu laikotarpiu, negu tie, apie kuriuos kalbama **2.2.1** pavyzdyje. Duomenys pateikiami 3.5.1 lentelėje.

3.5.1 lentelė. Statistiniai duomenys

i	X_{1i}	X_{2i}	X_{3i}	i	X_{1i}	X_{2i}	X_{3i}	i	X_{1i}	X_{2i}	X_{3i}
1	6,3	6,1	6,1	7	6,2	6,1	6,2	13	6,1	6,4	6,2
2	6,6	6,4	6,7	8	6,4	6,3	6,4	14	6,2	6,0	6,1
3	6,4	6,4	6,5	9	6,3	6,4	6,3	15	6,2	6,1	6,3
4	6,4	6,7	6,7	10	6,3	6,3	6,4	16	6,2	6,1	6,2
5	6,2	6,3	6,3	11	6,3	6,2	6,3	17	6,2	6,2	6,2
6	6,2	6,2	6,2	12	6,3	6,3	6,3	18	6,3	6,4	6,2

Tarę, kad **1.2.1**, **2.2.1** ir **3.5.1** lentelių duomenys yra nepriklausomų trimačių normaliųjų vektorių realizacijos, patikrinsime hipotezę, kad vidurkių vektorius nepakito (priimame prielaidą, kad kovariacinė matrica išliko nepakitusi).

Pagal **3.5.1** lentelės duomenis randame

$$\hat{\boldsymbol{\mu}}_3 = \bar{\mathbf{X}}_3 = \frac{1}{n_3}\sum_{i=1}^{n_3}\mathbf{X}_i = \frac{1}{18}\left\{\begin{pmatrix} 6,3 \\ 6,1 \\ 6,1 \end{pmatrix} + ... + \begin{pmatrix} 6,3 \\ 6,4 \\ 6,2 \end{pmatrix}\right\} = \begin{pmatrix} 6,28333 \\ 6,27222 \\ 6,31111 \end{pmatrix}$$

ir apskaičiuojame matricos $\boldsymbol{S}$ realizaciją

$$\boldsymbol{S}_3 = \sum_{i=1}^{n}\mathbf{X}_i\mathbf{X}_i^T - n\bar{\mathbf{X}}\bar{\mathbf{X}}^T = \begin{pmatrix} 6,3 \\ 6,1 \\ 6,1 \end{pmatrix}(6,3; 6,1; 6,1) + ... + \begin{pmatrix} 6,3 \\ 6,4 \\ 6,2 \end{pmatrix}(6,3; 6,4; 6,2) -$$

$$18\begin{pmatrix} 6,28333 \\ 6,27222 \\ 6,31111 \end{pmatrix}(6,28333; 6,27222; 6,31111) = \begin{pmatrix} 0,22500 & 0,16167 & 0,27333 \\ 0,16167 & 0,47611 & 0,36556 \\ 0,27333 & 0,36556 & 0,51778 \end{pmatrix}.$$

Matricos $\boldsymbol{SS}_E$ relizacija gaunama (žr. pvz. **3.2.1**) sudedant matricų $\boldsymbol{S}$ realizacijas pagal **1.2.1**, **2.2.1** ir **3.5.1** lentelių duomenis. Gauname

$$\boldsymbol{SS}_E = \boldsymbol{S}_1 + \boldsymbol{S}_2 + \boldsymbol{S}_3 = \begin{pmatrix} 0,88633 & 0,63567 & 0,53666 \\ 0,63567 & 1,25311 & 0,59806 \\ 0,53666 & 0,59806 & 1,05736 \end{pmatrix}.$$

Matrica $\boldsymbol{SS}_{EH}$ (žr. pvz. **3.4.1**) gali būti apskaičiuota taip

$$\boldsymbol{SS}_{EH} = \boldsymbol{SS}_E + \sum_{i=1}^{3} n_i(\bar{X}_i - \bar{X})(\bar{X}_i - \bar{X})^T,$$

čia $\bar{X}$ yra visų stebėjimų aritmetinis vidurkis

$$\bar{X} = \frac{1}{72}(30\bar{X}_1 + 24\bar{X}_2 + 18\bar{X}_3) = \begin{pmatrix} 6,27639 \\ 6,24306 \\ 6,29583 \end{pmatrix}.$$

Atlikę skaičiavimus gauname

$$\boldsymbol{SS}_{EH} = \begin{pmatrix} 0,88983 & 0,64319 & 0,54291 \\ 0,64319 & 1,27652 & 0,61292 \\ 0,54291 & 0,61292 & 1,06875 \end{pmatrix}.$$

Apskaičiuojame determinantus ir statistikos U realizaciją

$$|\boldsymbol{SS}_E| = 0,47725, \quad |\boldsymbol{SS}_{EH}| = 0,48939, \quad U = |\boldsymbol{SS}_E|/|\boldsymbol{SS}_{EH}| = 0,97519.$$

Kadangi tikriname trijų vidurkių vektorių lygybės hipotezę, tai lygybėse (3.4.1) matrica $\boldsymbol{H}$ gali būti parinkta taip: pirmoji eilutė (1; 0; -1), antroji eilutė (0; 1; -1); $\boldsymbol{\theta}_0 = \mathbf{0}$. Kadangi matricos $\boldsymbol{H}$ rangas $r = 2$, tai pritaikomas 4 išnagrinėtas atvejis ir yra galimybė sudaryti tikslų tikėtinumų santykio kriterijų. Randame

$$F = \frac{(1-\sqrt{U})67}{3\sqrt{U}} = 0,28227, \quad pv = \mathbf{P}\{F_{6,\,134} > 0,28227\} = 0,944505.$$

Atmesti vidurkių vektorių lygybės hipotezę nėra pagrindo.

3.5.4. Tikėtinumų santykio statistikos asimptotinis skirstinys

Remiantis tikėtinumų santykio asimptotinėmis savybėmis (žr. 1 dalį, 4.5.4 skyrelis, 4.5.6 teorema), galima tvirtinti, kad a. d.

$$Z = -2\ln\Lambda = -n\ln U \xrightarrow{d} \chi^2_\nu, \quad n \to \infty. \qquad (3.5.20)$$

Laisvės laipsnių skaičius $\nu = rk$, nes hipotezėje (3.4.1) yra po r apribojimų kiekvieno nežinomų parametrų vektorių $\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_k$ koordinatėms.

Apytikslis α lygmens tikėtinumų santykio kriterijus atmeta hipotezę, kai

$$Z = -2\ln\Lambda = -n\ln U > \chi^2_\alpha(kr). \qquad (3.5.21)$$

Kriterijų galima užrašyti ir asimptotinės P reikšmės terminais: hipotezė atmetama apytiksliu α lygmens kriterijumi, kai

$$pv_a = \mathbf{P}\{\chi^2_{kr} > z\} < \alpha, \qquad (3.5.22)$$

čia z yra statistikos Z realizacija.

Asimptotinio sąryšio (3.5.20) teisingumu galima įsitikinti ir tiesiogiai. Atsitiktinio dydžio $Z = -2\ln\Lambda$ charakteristinė funkcija

$$\psi(t) = \mathbf{E}(e^{it(-2\ln\Lambda)}) = \mathbf{E}(e^{-itn\ln U}) = \mathbf{E}(U^{-itn})$$

gaunama momento išraiškoje (3.5.1) arba (3.5.4) vietoje h įrašius $-itn$:

$$\begin{aligned}\psi(t) &= \prod_{j=1}^{k}\left\{\frac{\Gamma(\frac{n-m+1-j}{2} - itn)\Gamma(\frac{n-m+r+1-j}{2})}{\Gamma(\frac{n-m+1-j}{2})\Gamma(\frac{n-m+r+1-j}{2} - itn)}\right\} \\ &= \prod_{i=1}^{r}\left\{\frac{\Gamma(\frac{\nu-k+1-i}{2} - itn)\Gamma(\frac{\nu+1-i}{2})}{\Gamma(\frac{\nu-k+1-i}{2})\Gamma(\frac{\nu+1-i}{2} - itn)}\right\}, \quad \nu = n-m+r.\end{aligned} \qquad (3.5.23)$$

Skleidžiant gama funkcijas pagal Stirlingo formulę galima įsitikinti, kad $n \to \infty$ kiekvienas apskliaustas daugiklis baigtiniuose t kitimo intervaluose konverguoja į $(1-2it)^{-r/2}$, t. y. $\psi(t)$ konverguoja į$(1-2it)^{-kr/2}$ (žr. 3.5.5 skyrelį).

3.5.2 pavyzdys. (**3.5.1** pavyzdžio tęsinys). Palyginti išspręsime tą patį uždavinį kaip ir pavyzdyje **3.5.1**, taikydami asimptotinį tikėtinumų santykio kriterijų. Randame statistikos $Z = -n\ln U$ realizaciją

$$Z = -72\ln 0,97519 = 1,80859.$$

Laisvės laipsnių skaičius $\nu = kr = 6$ ir asimptotinė P reikšmė

$$pv_a = \mathbf{P}\{\chi^2_6 > 1,80859\} = 0,93643.$$

Kadangi P reikšmė didelė, atmesti vidurkių vektorių lygybės hipotezę nėra pagrindo. Matome, kad asimptotinė P reikšmė pv_a yra šiek tiek mažesnė už tikrąją P reikšmę pv, rastą **3.5.1** pavyzdyje.

3.5.5. Asimptotinio skirstinio patikslinimai

Remiantis charakteristinės funkcijos (3.5.23) išraiška aproksimacijos (3.5.20) tikslumą galima padidinti.

3.5.5.1. Vidurkių sutapatinimas

Kriterijus (3.5.21) yra asimptotinis ir jo tikslumas priklauso nuo (3.5.20) konvergavimo greičio. Todėl kai nedideli n, matematinės statistikos knygose siūlomos tam tikros tikėtinumų santykio statistikos modifikacijos, norint patikslinti asimptotinį tikėtinumų santykio kriterijų.

Viena iš tokių rekomendacijų yra taip modifikuoti tikėtinumų santykio statistiką, kad jos vidurkis būtų artimesnis ribinio skirstinio vidurkiui. Pavyzdžiui, jeigu

$$Z = -2\ln\Lambda \xrightarrow{d} \chi^2_\nu, \quad \mathbf{E}Z = \nu(1 + \frac{a}{n} + O(1/n^2)),$$

tai vietoje Z rekomenduojama naudoti statistiką

$$Z^* = Z(1 - \frac{a}{n}), \quad \mathbf{E}Z^* = \nu + O(1/n^2)$$

ir χ^2 skirstiniu aproksimuoti statistiką Z^*.

Tačiau, turint charakteristinės funkcijos išraišką (3.5.22), galima pasiekti, kad vidurkiai visiškai sutaptų. Skleidžiant $\ln\varphi(t)$ gaunamas formalus skleidinys

$$\ln\varphi(t) = \sum_{j=1}^{\infty} \frac{\kappa_j}{j!}(it)^j, \tag{3.5.24}$$

čia κ_j yra statistikos $Z = -2\ln\Lambda$ j-asis semiinvariantas. Pirmieji trys semiinvariantai yra

$$\kappa_1 = \mathbf{E}Z, \quad \kappa_2 = \boldsymbol{V}Z, \quad \kappa_3 = \mathbf{E}(Z - \mathbf{E}Z)^3. \tag{3.5.25}$$

Randame

$$\kappa_s = \sum_{j=1}^{k}(-1)^s n^s \left[\psi_s(\frac{n-m+1-j}{2}) - \psi_s(\frac{n-m+r+1-j}{2})\right] =$$

$$\sum_{j=1}^{r}(-1)^s n^s \left[\psi_s(\frac{\nu-k+1-j}{2}) - \psi_s(\frac{\nu+1-j}{2})\right], \quad s = 1, 2, ...; \tag{3.5.26}$$

čia $\psi_s(x)$ yra funkcijos $\ln\Gamma(x)$ s-oji išvestinė. Pirmosios išvestinės yra tabuliuotos (žr. [1]), arba jų reikšmių skaičiavimas numatytas kai kuriuose kompiuterių programų paketuose (pvz., SAS).

Jeigu vietoje Z imsime statistiką $Z^* = \delta Z$, kai $\delta = kr/\kappa_1$, tai $\mathbf{E}Z^* = kr$ sutampa su ribinio skirstinio iš (3.5.20) vidurkiu.

Hipotezė (3.4.1) atmetama asimptotiniu modifikuotu kriterijumi, kai

$$Z^* = -\delta n \ln U > \chi^2_\alpha(kr),$$

arba P reikšmių terminais, kai

$$pv_a^{(1)} = \mathbf{P}\{\chi^2_{kr} > \delta z\}, \tag{3.5.27}$$

čia z yra statistikos Z realizacija.

3.5.5.2. Dviejų momentų sutapatinimas

Atsitiktinio dydžio, turinčio χ^2 skirstinį, vidurkis ir dispersija santykiauja kaip $1:2$. Parinkime daugiklį δ taip, kad statistikos $Z^* = \delta Z$ vidurkis ir dispersija santykiautų kaip $1:2$. Išspręskime lygčių sistemą

$$\mathbf{E}Z^* = \mathbf{E}(\delta Z) = \delta\kappa_1 = \nu,$$

$$\boldsymbol{V}Z^* = \boldsymbol{V}(\delta Z) = \delta^2\kappa_2 = 2\nu,$$

parametrų δ ir ν atžvilgiu. Gauname

$$\delta = \frac{2\kappa_1}{\kappa_2}, \quad \nu = \frac{2\kappa_1^2}{\kappa_2}. \tag{3.5.28}$$

Vietoje Z imame statistiką $Z^* = \delta Z$ ir jos skirstinį aproksimuojame χ^2 skirstiniu su ν laisvės laipsnių (šis laisvės laipsnių skaičius gali skirtis nuo kr). Hipotezė (3.4.1) atmetama asimptotiniu modifikuotu reikšmingumo lygmens α kriterijumi, kai

$$pv_a^{(2)} = \mathbf{P}\{\chi_\nu^2 > \delta z\} < \alpha, \tag{3.5.29}$$

čia z – statistikos Z realizacija.

3.5.2 pastaba. Laisvės laipsnių skaičius ν nebūtinai bus sveikasis skaičius. Todėl skaičiuojant (3.5.29) reikia prisiminti, kad χ^2 skirstinys yra atskiras gama skirstinio atvejis: $\chi_\nu^2 \sim G(1/2, \nu/2)$.

3.5.5.3. Trijų momentų sutapatinimas

Atsitiktinio dydžio, turinčio χ^2 skirstinį, vidurkis, dispersija ir trečias centrinis momentas santykiauja kaip $1:2:8$. Parinkime parametrus δ ir γ taip, kad statistikos $Z^* = \delta Z + \gamma$ vidurkis, dispersija ir trečiasis centrinis momentas santykiautų kaip $1:2:8$. Išspręskime lygčių sistemą

$$\mathbf{E}Z^* = \mathbf{E}(\delta Z + \gamma) = \delta\kappa_1 + \gamma = \nu,$$

$$\boldsymbol{V}Z^* = \boldsymbol{V}(\delta Z) = \delta^2\kappa_2 = 2\nu,$$

$$\mathbf{E}(Z^* - \mathbf{E}Z^*)^3 = \delta^3\kappa_3 = 8\nu$$

parametrų δ, γ ir ν atžvilgiu. Gauname

$$\delta = \frac{4\kappa_2}{\kappa_3}, \quad \nu = \frac{8\kappa_2^3}{\kappa_3^2}, \quad \gamma = \frac{4\kappa_2[2\kappa_2^2 - \kappa_1\kappa_3]}{\kappa_3^2}. \tag{3.5.30}$$

Vietoje Z imame statistiką $Z^* = \delta Z + \gamma$ ir jos skirstinį aproksimuojame χ^2 skirstiniu su ν laisvės laipsnių. Hipotezė (3.4.1) atmetama asimptotiniu modifikuotu reikšmingumo lygmens α kriterijumi, kai

$$pv_a^{(3)} = \mathbf{P}\{\chi_\nu^2 > \delta z + \gamma\} < \alpha, \tag{3.5.31}$$

čia z yra statistikos Z realizacija.

3.5.3 pavyzdys. (**3.5.1** pavyzdžio tęsinys). Palyginti išspręsime tą patį uždavinį kaip ir pavyzdyje **3.5.1**, taikydami pateiktus asimptotinio skirstinio patikslinimus.

Kadangi $r = 1$, tai, naudodami formulę $\Gamma(x+1) = x\Gamma(x)$, iš (3.5.23) gauname

$$\ln\varphi(t) = \sum_{j=1}^{3}\left(\ln\frac{n-m+1-j}{n-m+1-j-2it}\right) = -\sum_{j=1}^{3}\ln\left(1-\frac{2itn}{n-m+1-j}\right).$$

Skleisdami eilute gauname pirmuosius semiinvariantus:

$$\kappa_1 = 2\left(\frac{n}{n-3}+\frac{n}{n-4}+\frac{n}{n-5}\right) = 6,35386;$$

$$\kappa_2 = 2\left(\left(\frac{n}{n-3}\right)^2+\left(\frac{n}{n-4}\right)^2+\left(\frac{n}{n-5}\right)^2\right) = 13,45911;$$

$$\kappa_3 = 2\left(\left(\frac{n}{n-3}\right)^3+\left(\frac{n}{n-4}\right)^3+\left(\frac{n}{n-5}\right)^3\right) = 57,02794.$$

1) Randame $\delta = kr/\kappa_1 = 0,94417$, $\delta z = 1,70787$ ir

$$pv_a^{(1)} = \mathbf{P}\{\chi^2_{kr} > \delta z\} = \mathbf{P}\{\chi^2_6 > 1,70787\} = 0,94451;$$

2) Pagal (3.5.28) randame $\delta = 0,94417$, $\nu = 5,99914$, $\delta z = 1,70762$ ir

$$pv_a^{(2)} = \mathbf{P}\{\chi^2_\nu > \delta z\} = \mathbf{P}\{Y > 1,70762\} = 0,94450,$$

čia $Y \sim G(1/2, \nu/2)$;

2) Pagal (3.5.30) randame $\delta = 0,94404$, $\nu = 5,99741$, $\gamma = -0,00086$. $\delta z + \gamma = 1,70652$ ir

$$pv_a^{(3)} = \mathbf{P}\{\chi^2_\nu > \delta z + \gamma\} = \mathbf{P}\{Y > 1,70652\} = 0,94451,$$

čia $Y \sim G(1/2, \nu/2)$.

Visais trimis atvejais P reikšmės yra didelės ir atmesti vidurkių vektorių lygybės hipotezę nėra pagrindo. Matome, kad jau vidurkių sutapatinimas duoda asimptotinę P reikšmę, kuri daug artimesnė tikrajai pv iš **3.5.1** pratimo, negu asimptotinė P reikšmė, gauta **3.5.2** pratime neatliekant asimptotinio skirstinio patikslinimo.

3.5.5.4. Asimptotiniai skleidiniai

Knygoje [2] siūloma patikslinti konvergavimą (3.5.20) imant asimptotinius skleidinius. Jie gaunami skleidžiant funkciją $\ln\varphi(\delta t)$ ne it, o $(1-2it)$ laipsniais.

3.5.6 teorema. *Teoremos 3.5.1 sąlygomis, kai $n \to \infty$, o k, m, r fiksuoti, gauname tokį P reikšmės patikslinimą:*

$$pv_a^{(4)} = \mathbf{P}\{\chi^2_{kr} > \delta z\} + \frac{\omega_2}{n^2\delta^2}[\mathbf{P}\{\chi^2_{kr+4} > \delta z\} - \mathbf{P}\{\chi^2_{kr} > \delta z\}] + O\left(\frac{1}{n^3}\right), \tag{3.5.32}$$

čia z yra statistikos Z realizacija, o parametrai

$$\delta = \frac{n-m-(k-r+1)/2}{n}, \quad \omega_2 = \frac{kr(k^2+r^2-5)}{48}. \tag{3.5.33}$$

Įrodymas. Naudosime funkcijos $\ln\Gamma(x+h)$ skleidimą, kai $x\to\infty$, o h yra fiksuotas (žr. [1])

$$\ln\Gamma(x+h) = \ln\sqrt{2\pi} + (x+h-\frac{1}{2})\ln x - x +$$

$$\sum_{s=1}^{l}(-1)^{s+1}\frac{B_{s+1}(h)}{s(s+1)x^s} + O\left(\frac{1}{x^{l+1}}\right), \tag{3.5.34}$$

čia $B_s(h)$ yra laipsnio s Bernulio polinomas. Bernulio polinomai galima rasti iš tapatybės

$$\tau e^{h\tau} = (e^\tau - 1)\sum_{s=0}^{\infty}\frac{\tau^s}{s!}B_s(h).$$

Pirmieji trys Bernulio polinomai yra

$$B_1(h) = h - \frac{1}{2},\quad B_2(h) = h^2 - h + \frac{1}{6},\quad B_3(h) = h^3 - \frac{3}{2}h^2 + \frac{1}{2}h. \tag{3.5.35}$$

Pertvarkome statistikos $\delta Z = -\delta n\ln U$ charakteristinę funkciją (žr. (3.5.23)):

$$\ln\varphi(\delta t) = \sum_{j=1}^{k}\{\ln\Gamma(x(1-2it)+h_{1j}) - \ln\Gamma(x(1-2it)+h_{2j})$$

$$-[\ln\Gamma(x+h_{1j}) - \ln\Gamma(x+h_{2j})]\}, \tag{3.5.36}$$

čia

$$x = \frac{\delta n}{2},\quad h_{1j} = \frac{n(1-\delta)-m+1-j}{2},\quad h_{2j} = \frac{n(1-\delta)-m+r+1-j}{2}.$$

Jeigu $x\to\infty$, o h_{1j} ir h_{2j} aprėžti, galima taikyti (3.5.34) skleidimą. Imdami pirmuosius tris (3.5.34) skleidinio narius gauname

$$\ln\varphi(\delta t) = -\frac{rk}{2}\ln(1-2it) + O(1/n).$$

Taigi tiesiogiai įsitikiname statistikos Z skirstinio konvergavimu (3.5.20). Imdami tolesnius (3.5.34) narius gauname skleidinį

$$\ln\varphi(\delta t) = \exp\{-\frac{rk}{2}\ln(1-2it) + \sum_{s=1}^{l}\frac{\omega_s}{(n\delta)^s}\left[\frac{1}{(1-2it)^s} - 1\right] + O(\frac{1}{n^{l+1}})\}, \tag{3.5.37}$$

čia

$$\omega_s = (-1)^{s+1}2^s\sum_{j=1}^{k}\frac{B_{s+1}(h_{1j}) - B_{s+1}(h_{2j})}{s(s+1)}.$$

Imdami $s=1$ randame

$$\omega_1 = \sum_{j=1}^{k}(h_{1j}-h_{2j})(h_{1j}+h_{2j}-1) = -\frac{rk}{2}\left(n(1-\delta) - m - \frac{k+1-r}{2}\right).$$

Jeigu parinksime parametrą δ taip, kad $\omega_1 = 0$, tai skleidinys prasidės nuo nario, kurio eilė n^{-2}. Gauname

$$\delta = \frac{n - m - (k + 1 - r)/2}{n}. \tag{3.5.38}$$

Tada koeficientas prie antrojo skleidinio nario yra

$$\omega_2 = -\frac{2}{3}\sum_{j=1}^{k}[B_3(h_{1j}) - B_3(h_{2j})] = \frac{rk}{48}(r^2 + k^2 - 5).$$

Apsiribojęę šiuo nariu gauname charakteristinės funkcijos skleidinį

$$\varphi(\delta t) = \left(\frac{1}{1 - 2it}\right)^{kr/2}\left\{1 + \frac{\omega_2}{(n\delta)^2}\left(\frac{1}{(1 - 2it)^2} - 1\right) + O\left(\frac{1}{n^4}\right)\right\},$$

nes $\omega_3 = 0$. Pritaikę atvertimo formulę gauname (3.5.32).▲

3.5.3 pastaba. Naudojant daugiau (3.5.37) skleidinio narių galima gauti tolesnius tikimybės (3.5.32) patikslinimus. Knygoje [2] pateiktas toks papildomas narys

$$pv_a^{(5)} = pv_a^{(4)} + \frac{\gamma_4}{(\delta n)^4}(\mathbf{P}\{\chi^2_{kr+8} \le z\} - \mathbf{P}\{\chi^2_{kr} \le z\})$$

$$-\frac{\omega_2^2}{(\delta n)^4}(\mathbf{P}\{\chi^2_{kr+4} \le z\} - \mathbf{P}\{\chi^2_{kr} \le z\}) + O(1/n^6), \tag{3.5.39}$$

čia

$$\gamma_4 = \frac{\omega_2^2}{2} + \omega_4, \quad \omega_4 = \frac{kr}{1920}(3k^4 + 3r^4 + 10k^2r^2 - 50(k^2 + r^2) + 159).$$

3.5.4 pavyzdys.(**3.5.1** pavyzdžio tęsinys). Palyginti vėl išspręsime tą patį uždavinį, kaip ir pavyzdyje **3.5.1** naudodami asimptotinio skirstinio (3.5.22) patikslinimą.

Kadangi $r = 1$, tai formulėje (3.5.36) $h_{2j} = h_{1j}+1$ ir, naudodami išraišką $\Gamma(x+1) = x\Gamma(x)$, gauname

$$\ln\varphi(\delta t) = \exp\{-k\ln(1 - 2it)\}\exp\{\sum_{j=1}^{k}[\ln(1 + h_{1j}/x) - \ln(1 + h_{1j}/(x(1 - 2it)))]\}.$$

Skleisdami eilute gauname (3.5.37) analogą

$$\ln\varphi(\delta t) = \exp\{-k\ln(1 - 2it)\}\exp\left\{\sum_{s=1}^{\infty}\frac{\omega_s}{(n\delta)^s}\left[\frac{1}{(1 - 2it)^s} - 1\right]\right\},$$

čia

$$\omega_s = (-1)^s\frac{2^s}{s}\sum_{j=1}^{k}h_{1j}^s, \quad s = 1, 2, \ldots.$$

Jeigu parinksime δ taip, kad $\omega_1 = 0$, t. y. pagal (3.5.38), tai $h_{1j} = (k+1-2j)/4$ ir lengvai gauname, kad $\omega_{2s+1} = 0, \quad \omega_{2s} = 1/s, \quad s = 1, 2, \ldots$. Randame

$$\delta = \frac{68}{72} = 0,94444, \quad n\delta = 68, \quad \omega_2 = 1, \quad \delta z = 1,7081$$

ir asimptotinę P reikšmę

$$pv_a^{(4)} = \mathbf{P}\{\chi_6^2 > \delta z\} + \frac{1}{68^2}\{\mathbf{P}\{\chi_{10}^2 > \delta z\} - \mathbf{P}\{\chi_6^2 > \delta z\}\} = 0,94449 + 0,00001 = 0,94450.$$

Patikslinimas (3.5.39) visai mažas $pv_a^{(5)} - pv_a^{(4)} < 10^{-8}$. Atmesti hipotezę nėra pagrindo.

3.5.5. pavyzdys. Kad būtų geriau matomas siūlomų patikslinimų indėlis, panagrinėsime iliustracinį pavyzdį, kuriame imtis yra mažesnė, o P reikšmė nėra artima vienetui. Tarkime, kad, tikrinant trijų vidurkių vektorių lygybės hipotezę pagal tris nepriklausomas didumo $n_1 = n_2 = n_3 = 6$ imtis, gautas stebint n. a. v. $\mathbf{X}_j \sim N_3(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}), j = 1, 2, 3$, apskaičiuota statistikos $Z = -n \ln U$ reikšmė yra $12, 6$. Lentelėje **3.5.2** yra pateikta tiksli P reikšmė pv ir apytikslės P reikšmės pv_a, $pv_a^{(j)}, j = 1, ..., 5$ naudojant pateiktus patikslinimus, ir jų nuokrypiai $pv_a - pv$ ir $pv_a^{(j)} - pv$.

3.5.2 lentelė. Asimptotinės P reikšmės.

pv	pv_a	$pv_a^{(1)}$	$pv_a^{(2)}$	$pv_a^{(3)}$	$pv_a^{(4)}$	$pv_a^{(5)}$
0,134997	0,049846	0,134831	0,135101	0,135061	0,134989	0,134989
	-0,08515	-0,00017	0,00010	0,00006	-0,00001	-0,00001

Matome, kad palyginti nedidelėms imtims asimptotinis tikėtinumų santykio kriterijus gali būti netikslus. Jeigu šiame iliustraciniame pavyzdyje būtume parinkę reikšmingumo lygmenį $\alpha = 0, 05$, tai asimptotinis tikėtinumų santykio kriterijus hipotezę atmestų (asimptotinė P reikšmė yra 0,049846). O tiksli P reikšmė lygi 0,134997 ir atmesti hipotezę nėra pagrindo. Jau paprasčiausias patikslinimas tiksliai sutapatinant vidurkius ($pv_a^{(1)}$), arba iš dalies juos sutapatinant (pirmasis dėmuo (3.5.32) formulėje lygus 0,133331) duoda esminį asimptotinės P reikšmės patikslinimą. Tolesnių patikslinimo narių įtaka kur kas mažesnė (žr. $pv_a^{(2)}$, $pv_a^{(3)}$, $pv_a^{(4)}$), o papildomas narys (3.5.39) formulėje didelės svarbos neturi, nes $|pv_a^{(5)} - pv_a^{(4)}| < 10^{-7}$.

3.5.5.5. Kompiuterinis modeliavimas

Turint (3.5.2) ar kitus tokio tipo sąryšius, tikėtinumų santykio kriterijaus kritines reikšmes ar P reikšmes reikiamu tikslumu galima rasti naudojant kompiuterinį modeliavimą.

Pavyzdžiui, naudojant (3.5.2) sąryšį galima realizuoti tokį P reikšmės įvertinimo algoritmą.

0. Apskaičiuojame sprendžiamo uždavinio statistikos $Z = -n \ln U$ realizaciją z.

1. Modeliuojame nepriklausomus a. d. $\xi_j \sim Be((n - m + 1 - j)/2, r/2), j = 1, ..., k$.

2. Pagal (3.5.2) formulę randame statistikos Z modeliuotą realizaciją

$$z^* = -\sum_{j=1}^{k} \ln(\xi_j).$$

3. Patikriname nelygybę $z^* > z$; jeigu ši nelygybė teisinga, tai sumatoriuje esantį skaičių M padidiname vienetu (pradinė M reikšmė lygi 0).

4. Realizuojame aprašytą procedūrą pradedant nuo 1 punkto N kartų.

5. Įvertiname P reikšmę:

$$\hat{pv} = \frac{M}{N}.$$

Tikrinamoji hipotezė (3.4.1) atmetama apytiksliu reikšmingumo lygmens α kriterijumi, kai

$$\hat{pv} < \alpha. \tag{3.5.40}$$

Reikia pažymėti, kad šio kriterijaus tikslumas iš esmės priklauso tik nuo skaičiaus N. Šiuolaikiniai kompiuteriai leidžia gana greit atlikti tokio tipo procedūras reikiamą skaičių kartų.

3.6. Pratimai

3.1. Stebint a. v. $\mathbf{X} = (X_1, ..., X_k)^T$ eksperimentas atliktas pagal dvifaktorės dispersinės analizės schemą su vienodu stebėjimų skaičiumi langelyje. Raskite hipotezių dėl sąveikos ir faktorių įtakos tikrinimo kriterijų daugiamačius analogus.

3.2. Stebint a. v. $\mathbf{X} = (X_1, ..., X_k)^T$ eksperimentas atliktas pagal lotynų kvadratų schemą su vienu stebėjimu langelyje. Raskite hipotezių dėl faktorių įtakos tikrinimo kriterijų daugiamačius analogus.

3.3. Įrodykite, kad $\hat{\boldsymbol{B}}$ minimizuoja apibendrintąją dispersiją

$$|\sum_{j=1}^{n}(\mathbf{X}_j - \boldsymbol{B}\boldsymbol{a}_j)(\mathbf{X}_j - \boldsymbol{B}\boldsymbol{a}_j)^T|.$$

3.4. Įrodykite, kad kriterijus (3.5.12) yra ekvivalentus kriterijui, grindžiamam Hotelingo statistika.

3.5. Aštuoniuose vienodo dydžio sklypeliuose gauti tokie duomenys

Sklypelio numeris	1	2	3	4	5	6	7	8
Grūdų kiekis (X_{1i})	40	17	9	15	6	12	5	9
Šiaudų kiekis (X_{2i})	53	19	10	29	13	27	19	30
Trąšų kiekis (Z_i)	24	11	5	12	7	14	11	18

Patikrinkite hipotezę, kad vektoriaus $(X_1, X_2)^T$ skirstinys nepriklauso nuo įterptų trąšų kiekio Z.

3.6. Pagal didumo $n = 25$ a. v. $(X_1, X_2, Z_1, ..., Z_7)^T$ paprastąją imtį (čia X_1 – papiroso degimo greitis; X_2 – nikotino kiekis; Z_1 – azoto, Z_2 – chloro, Z_3 – kalio, Z_4 – fosforo, Z_5 – kalcio, Z_6 – magnio procentas; $Z_7 = 1$) gauti tokie rezultatai (žr. [2]):

$$\sum_j \mathbf{X}_j = \begin{pmatrix} 42,20 \\ 54,03 \end{pmatrix}, \quad \sum_j (\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})^T = \begin{pmatrix} 0,6690 & 0,4527 \\ 0,4527 & 6,5921 \end{pmatrix};$$

$$\sum_j \boldsymbol{Z}_j = \begin{pmatrix} 53,92 \\ 62,02 \\ 56,00 \\ 12,25 \\ 89,79 \\ 24,10 \\ 25 \end{pmatrix}, \quad \sum_j (\boldsymbol{Z}_j - \bar{\boldsymbol{Z}})(\mathbf{X}_j - \bar{\mathbf{X}})^T = \begin{pmatrix} 0,2501 & 2,6691 \\ -1,5136 & -2,0617 \\ 0,5007 & -0,9503 \\ -0,0421 & -0,0187 \\ -0,1914 & 3,4020 \\ -0,1586 & 1,1663 \\ 0 & 0 \end{pmatrix};$$

$$\sum_j (\boldsymbol{Z}_j - \bar{\boldsymbol{Z}})(\boldsymbol{Z}_j - \bar{\boldsymbol{Z}})^T =$$

$$= \begin{pmatrix} 1,8311 & -0,3589 & -0,01252 & -0,0244 & 1,6379 & 0,5057 & 0 \\ -0,3589 & 8,8102 & -0,3469 & 0,0352 & 0,7920 & 0,2173 & 0 \\ -0,0125 & -0,3469 & 1,5818 & -0,0415 & -1,4278 & -0,4753 & 0 \\ -0,0244 & 0,0352 & -0,0415 & 0,0258 & 0,0043 & 0,0154 & 0 \\ 1,6379 & 0,7920 & -1,4278 & 0,0043 & 3,7248 & 0,9120 & 0 \\ 0,5057 & 0,2173 & -0,4753 & 0,0154 & 0,9120 & 0,3828 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

a) Įvertinkite a. v. $(X_1, X_2)^T$ regresiją vektoriaus $(Z_1, ..., Z_7)^T$ atžvilgiu.
b) Patikrinkite hipotezę, kad regresijos koeficientai prie Z_2, Z_3, Z_4 lygūs 0.
c) Įvertinkite a. v. $(X_1, X_2)^T$ regresiją vektoriaus $(Z_1, Z_5, Z_6, Z_7)^T$ atžvilgiu.

3.7. Buvo atrinkta po 140 vienodo amžiaus moksleivių iš 6 skirtingų Indijos mokyklų ir pamatuoti tokie jų požymiai: $X^{(1)}$ – galvos ilgis; $X^{(2)}$ – galvos plotis; $X^{(3)}$ – svoris. Pažymėkime i-osios mokyklos j-ojo mokinio vektoriaus $\mathbf{X} = (X^{(1)}, X^{(2)}, X^{(3)})^T$ realizaciją $\mathbf{X}_{ij} = (X_{ij}^{(1)}, X_{ij}^{(2)}, X_{ij}^{(3)})^T$, $i = 1, ..., 6$, $j = 1, ..., 140$. Pagal gautus rezultatus apskaičiuotos matricos

$$\left[\sum_{i=1}^{6}\sum_{j=1}^{140}(X_{ij}^{(l)} - \bar{X}_{i.}^{(l)})(X_{ij}^{(l')} - \bar{X}_{i.}^{(l')})\right]_{3\times 3} = \begin{pmatrix} 12809,3 & 1003,7 & 2671,2 \\ 1003,7 & 1499,6 & 4123,6 \\ 2671,2 & 4123,6 & 21009,6 \end{pmatrix},$$

$$\left[140\sum_{i=1}^{6}(\bar{X}_{i.}^{(l)} - \bar{X}_{..}^{(l)})(\bar{X}_{i.}^{(l')} - \bar{X}_{..}^{(l')})\right]_{3\times 3} = \begin{pmatrix} 752,0 & 214,2 & 521,3 \\ 214,2 & 151,3 & 401,2 \\ 521,3 & 401,2 & 1612,7 \end{pmatrix}$$

Tarę, kad buvo stebimi nepriklausomi normalieji a. v. $\mathbf{X}_{ij} \sim N_3(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}), i = 1, ..., 6$, patikrinkite hipotezę $H : \boldsymbol{\mu}_1 = ... = \boldsymbol{\mu}_6$.

3.8. (**1.8** pratimo tęsinys). Sudarykite antrojo sūnaus galvos ilgio ir pločio (X_3, X_4) prognozę pagal pirmojo sūnaus galvos ilgį ir plotį (X_1, X_2) naudodami tiesinės regresijos modelį.

3.9. (**1.10** pratimo tęsinys). Sudarykite a. v. $(X_1, X_2)^T$ prognozę pagal a. v. $(X_3, X_4)^T$ naudodami tiesinės regresijos modelį.

3.10. (**1.13** pratimo tęsinys). a) Raskite a. v. $(X_2, X_3)^T$ prognozę pagal a. v. $(X_1, X_4)^T$ naudodami tiesinės regresijos modelį. b) patikrinkite hipotezę H, kad visi regresijos koeficientai (išskyrus laisvuosius narius) lygūs nuliui.

3.11. (**2.10** pratimo tęsinys). Sudarykite a. v. $(X_1, X_3)^T$ prognozę pagal a. d. X_2 naudodami tiesinės regresijos modelį.

3.12. (**2.17** pratimo tęsinys). Raskite a. v. $(X_1, X_2)^T$ prognozę pagal a. v. $(X_3, X_4, X_5, X_6)^T$ naudodami tiesinės regresijos modelį.

3.13. (**2.18** pratimo tęsinys). Raskite a. v. $(X_1, X_2)^T$ prognozę pagal a. d. X_3 naudodami tiesinės regresijos modelį pirmosios imties atveju.

3.14. Tirtos trijų tipų (privati, ne pelno, valstybinė) socialinės rūpybos organizacijos. Gautos didumo $n_1 = 271$, $n_2 = 138$ ir $n_3 = 107$ imtys, stebint keturmatį a. v. $\mathbf{X} = (X_1, X_2, X_3, X_4)^T$; čia X_1 – slaugymo kaina; X_2 – maisto kaina; X_3 – įrangos eksploatavimo kaina; X_4 – namų ruošos ir skalbimo kaina. Pagal šias imtis gauti vidurkių vektorių ir kovariacijų matricų nepaslinktieji įverčiai [9]:

$$\hat{\boldsymbol{\mu}}_1 = \bar{\mathbf{X}}_{1.} = \begin{pmatrix} 2,066 \\ 0,480 \\ 0,082 \\ 0,360 \end{pmatrix}, \quad \hat{\boldsymbol{\mu}}_2 = \bar{\mathbf{X}}_{2.} = \begin{pmatrix} 2,167 \\ 0,596 \\ 0,124 \\ 0,418 \end{pmatrix}, \quad \hat{\boldsymbol{\mu}}_3 = \bar{\mathbf{X}}_{3.} = \begin{pmatrix} 2,273 \\ 0,521 \\ 0,125 \\ 0,383 \end{pmatrix};$$

$$\hat{\mathbf{\Sigma}}_1 = \begin{pmatrix} 0,291 & -0,001 & 0,002 & 0,010 \\ -0,001 & 0,011 & 0,000 & 0,003 \\ 0,002 & 0,000 & 0,001 & 0,000 \\ 0,010 & 0,003 & 0,000 & 0,010 \end{pmatrix}, \quad \hat{\mathbf{\Sigma}}_2 = \begin{pmatrix} 0,561 & 0,011 & 0,001 & 0,037 \\ 0,011 & 0,025 & 0,004 & 0,007 \\ 0,001 & 0,004 & 0,005 & 0,002 \\ 0,037 & 0,007 & 0,002 & 0,019 \end{pmatrix},$$

$$\hat{\mathbf{\Sigma}}_3 = \begin{pmatrix} 0,261 & 0,030 & 0,003 & 0,018 \\ 0,030 & 0,017 & 0,000 & 0,006 \\ 0,003 & 0,000 & 0,004 & 0,001 \\ 0,018 & 0,006 & 0,001 & 0,013 \end{pmatrix}.$$

Priėmę normalumo prielaidą patikrinkite hipotezę $H : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \boldsymbol{\mu}_3$ dėl vidurkių vektorių lygybės tarę, kad kovariacinės matricos yra lygios.

3.15. (**3.14** pratimo tęsinys). Tegu vidurkių vektoriaus $\boldsymbol{\mu}_i$ koordinatės yra μ_{i1}, μ_{i2}, μ_{i3}, μ_{i4}, $i = 1,2,3$. Pažymėkime $\nu_{ii'j} = \mu_{ij} - \mu_{i'j}$, $i \neq i' = 1,2,3$, $j = 1,2,3,4$ vidurkių skirtumus. a) Raskite pasikliovimo lygmens $Q = 0,95$ parametrų $\nu_{ii'j}$ pasikliovimo intervalus; b) remdamiesi Bonferonio nelygybe raskite pasikliovimo intervalų rinkinį, kad tikimybė, jog jie uždengs visus parametrus $\nu_{ii'j}$, būtų ne mažesnė už 0,95.

3.16. Stebimas dvimatis a. v. $(X,Y)^T$, kurio skirstinys gali priklausyti nuo faktoriaus A, turinčio tris lygmenis A_1, A_2, A_3, ir nuo faktoriaus B, turinčio keturis lygmenis B_1, B_2, B_3, B_4. Gauti stebėjimų rezultatai pateikti lentelėje.

	B_1		B_2		B_3		B_4	
A_1	6	14	4	6	8	8	2	16
	8	8	6	2	12	2	6	-4
A_2	3	1	-3	5	4	0	-4	2
	8	6	2	12	3	15	3	7
A_3	-3	3	-4	-2	3	-11	-4	-6
	2	-2	-5	7	-3	1	-6	6

Priėmę normalumo prielaidą atlikite daugiamatę dvifaktorę dispersinę analizę.

3.17. Tiriama plastikinių plėvelių charakteristikų (X_1 – atsparumas trūkiams, X_2 – blizgumas, X_3 – skaidrumas) priklausomybė nuo dviejų faktorių A – štampavimo greičio, B – priemaišų kiekio, kurie gali būti dviejų lygmenų. Kiekvienam faktorių lygmenų rinkiniui stebėjimai pakartoti po 5 kartus. Duomenys pateikti lentelėje [9].

	B_1					B_2				
	6,5	6,2	5,8	6,5	6,5	6,9	7,2	6,9	6,1	6,3
A_1	9,5	9,9	9,6	9,6	9,2	9,1	10,0	9,9	9,5	9,4
	4,4	6,4	3,0	4,1	0,8	5,7	2,0	3,9	1,9	5,7
	6,7	6,6	7,2	7,1	6,8	7,1	7,0	7,2	7,5	7,6
A_2	9,1	9,3	8,3	8,4	8,5	9,2	8,8	9,7	10,1	9,2
	2,8	4,1	3,8	1,6	3,4	8,4	5,2	6,9	2,7	1,9

Priėmę normalumo prielaidą atlikite daugiamatę dvifaktorę dispersinę analizę.

3.18. Tiriama a. v. $\mathbf{X} = (X_1, X_2,)^T$, priklausomybė nuo dviejų faktorių A ir B, kurie gali būti trijų lygmenų. Kiekvienai faktorių lygmenų rinkiniui stebėjimai pakartoti po 3 kartus. Duomenys pateikti lentelėje [9].

	B_1			B_2			B_3		
A_1	9,33	8,74	9,31	10,22	10,13	10,42	15,25	16,22	17,24
	19,14	19,55	19,24	25,00	25,32	27,12	38,89	36,67	40,74
A_2	12,07	11,03	12,48	15,38	14,21	9,69	38,71	44,74	36,67
	33,03	32,37	31,31	40,00	40,48	33,90	77,14	78,57	71,43
A_3	8,73	7,94	8,37	8,45	6,79	8,34	14,04	13,51	13,33
	23,27	20,87	22,16	26,32	22,73	26,67	44,44	37,93	37,93

Priėmę normalumo prielaidą atlikite daugiamatę dvifaktorę dispersinę analizę.

3.19. Tiriama $n = 17$ pacientų kraujo rodiklių Y_1, Y_2 priklausomybė nuo Z_1 – lytis, Z_2 – pavartotų vaistų kiekis, Z_4 – diastolinis kraujo spaudimas, Z_3, Z_5 – kardiogramos charakteristikos [9].

Y_{1i}	Y_{2i}	Z_{1i}	Z_{2i}	Z_{3i}	Z_{4i}	Z_{5i}
3389	3149	1	7500	220	0	140
1101	653	1	1975	200	0	100
1131	810	0	3600	205	60	111
596	448	1	675	160	60	120
896	844	1	750	185	70	83
1767	1450	1	2500	180	60	80
807	493	1	350	154	80	98
1111	941	0	1500	200	70	93
645	547	1	375	137	60	105
628	392	1	1050	167	60	74
1360	1283	1	3000	180	60	80
652	458	1	450	160	64	60
860	722	1	1750	135	90	79
500	384	0	2000	160	60	80
781	501	0	4500	180	0	100
1070	405	0	1500	170	90	120
1754	1520	1	3000	180	0	129

Sudarykite a. v. $(Y_1, Y_2)^T$ prognozę pagal vektorių $\boldsymbol{Z} = (Z_1, Z_2, Z_3, Z_4, Z_5)^T$ naudodami daugiamatį tiesinės regresijos modelį. Raskite a. v. $(Y_1, Y_2)^T$ pasikliovimo lygmens $Q = 0,95$ prognozės elipsoidą taške $\boldsymbol{Z} = (1; 1200; 140; 70; 85)^T$.

Atsakymai ir nurodymai

3.1. Imties elementus žymėkime $X_{ijm}^{(l)}, l = 1, ..., k$; čia i ir j faktorių A ir B lygmenų numeriai $i = 1, ..., I, j = 1, ..., J$, m – kartotinumas, $m = 1, ..., M$. Tikėtinumų santykio statistikos yra:

$$\Lambda_A = (|\boldsymbol{SS}_E|/|\boldsymbol{SS}_{EH_A}|)^{n/2}, \quad \Lambda_B = (|\boldsymbol{SS}_E|/|\boldsymbol{SS}_{EH_B}|)^{n/2} \quad \Lambda_{AB} = (|\boldsymbol{SS}_E|/|\boldsymbol{SS}_{EH_{AB}}|)^{n/2};$$

$$\boldsymbol{SS}_E = [\boldsymbol{S}_E(r,s)]_{k\times k}, \quad SS_E(r,s) = \sum_i\sum_j\sum_m (X_{ijm}^{(r)} - \bar X_{ij.}^{(r)})(X_{ijm}^{(s)} - \bar X_{ij.}^{(s)}),$$

$$\boldsymbol{SS}_{EH_A} = [\boldsymbol{S}_{EH_A}(r,s)]_{k\times k}, \quad SS_{EH_A}(r,s) = SS_E(r,s) + JM\sum_i (\bar X_{i..}^{(r)} - \bar X_{...}^{(r)})(\bar X_{i..}^{(s)} - \bar X_{...}^{(s)}),$$

$$\boldsymbol{SS}_{EH_B} = [\boldsymbol{S}_{EH_B}(r,s)]_{k\times k}, \quad SS_{EH_B} = SS_E(r,s) + IM\sum_j (\bar X_{.j.}^{(r)} - \bar X_{...}^{(r)})(\bar X_{.j.}^{(s)} - \bar X_{...}^{(s)}),$$

$$\boldsymbol{SS}_{EH_{AB}} = [SS_{EH_{AB}}(r,s)]_{k\times k}, \quad SS_{EH_{AB}}(r,s) = SS_E(r,s) +$$
$$M\sum_i\sum_j (\bar X_{ij.}^{(r)} - \bar X_{i..}^{(r)} - \bar X_{.j.}^{(r)} + \bar X_{...}^{(r)})(\bar X_{ij.}^{(s)} - \bar X_{i..}^{(s)} - \bar X_{.j.}^{(s)} + \bar X_{...}^{(s)}).$$

Remiantis asimptotiniu tikėtinumų santykio kriterijumi pagrindinės dispersinės analizės hipotezės atmetamos reikšmingumo lygmens α kriterijais, kai teisingos nelygybės $-2\ln\Lambda_A > \chi^2_\alpha(I-1)$, $-2\ln\Lambda_B > \chi^2_\alpha(J-1)$, $-2\ln\Lambda_{AB} > \chi^2_\alpha((I-1)(J-1))$. **3.5.** Tarkime X_1 ir X_2 priklausomybę nuo Z aprašo tiesiniai regresijos modeliai: $X_{1i} = \alpha_1 + \beta_1 Z_i + e_{1i}$, $X_{2i} = \alpha_2 + \beta_2 Z_i + e_{2i}$, $i = 1, ..., 8$. Turime daugiamatį (dvimatį) tiesinės vieno kintamojo regresijos modelį. Priėmę normalumo prielaidas tikrinsime hipotezę $H : \beta_1 = 0, \beta_2 = 0$.

Randame parametrų α_1, β_1 ir α_2, β_2 įverčius ir matricą $\boldsymbol{SS}_E = [SS_E(i,j)]_{2\times 2}$,

$$SS_E(1,1) = \sum_{i=1}^8 [X_{1i} - \hat\alpha_1 - \hat\beta_1 Z_i]^2 = 382,5538, \quad SS_E(2,2) = \sum_{i=1}^8 [X_{2i} - \hat\alpha_2 - \hat\beta_2 Z_i]^2 = 98,9276,$$

$$SS_E(1,2) = \sum_{i=1}^8 [X_{1i} - \hat\alpha_1 - \hat\beta_1 Z_i][X_{2i} - \hat\alpha_2 - \hat\beta_2 Z_i] = 143,0225.$$

Kai hipotezė H teisinga, tai matricos $\boldsymbol{SS}_{EH} = [SS_{EH}(i,j)]_{2\times 2}$ elementai yra

$$SS_{EH}(1,1) = \sum_{i=1}^{8}(X_{1i} - \bar{X}_{1\cdot})^2 = 884,875, \quad SS_{EH}(2,2) = \sum_{i=1}^{8}(X_{2i} - \bar{X}_{2\cdot})^2 = 1270,$$

$$SS_{EH}(1,2) = \sum_{i=1}^{8}(X_{1i} - \bar{X}_{1\cdot})(X_{2i} - \bar{X}_{2\cdot}) = 910.$$

Statistikos U realizacija lygi determinantų santykiui

$$U = \frac{|\boldsymbol{SS}_E|}{|\boldsymbol{SS}_{EH}|} = 0,0588.$$

Šiame uždavinyje $n = 8, m = 2, r = 1, k = 2$. Kadangi $r = 1, k = 2$, tai turime išnagrinėtus atskirus atvejus, kuriems galime užrašyti tikslius kriterijus.

1. Kadangi $r = 1$, tai remiantis 3.5.3 teorema, kai teisinga hipotezė

$$U \sim Be\left(\frac{n-m+r-k}{2}, \frac{k}{2}\right) \sim Be(5/2,1) \Leftrightarrow F = \frac{(1-U)5}{2U} \sim F(2,\ 5).$$

Randame $F = 40,018$ ir $pv = \mathbf{P}\{F_{2;5} > 40,018\} = 0,000838$.

2. Kadangi $k = 2$, tai remiantis 3.5.4 teorema, kai teisinga hipotezė

$$\sqrt{U} \sim Be(n-m-1,r) \sim Be(5,1) \Leftrightarrow F = \frac{(1-\sqrt{U})5}{\sqrt{U}} \sim F(2,\ 10).$$

Randame $F = 15,61965$ ir $pv = \mathbf{P}\{F_{2;10} > 16,61965\} = 0,000838$.

Hipotezė atmestina.

Jeigu šiame uždavinyje būtume taikę asimptotinį kriterijų, t. y. statistikos $Z = -n\ln U$ skirstinį aproksimuodami χ^2 skirstiniu su 2 laisvės laipsniais, tai statistikos Z realizacija lygi 22,6689 ir $pv_a = \mathbf{P}\{\chi^2_2 > 22,6689\} = 0,000019$. Matome, kad asimptotinė P reikšmė yra daug mažesnė už tikrąją.

3.6. a) Turime daugiamatį (dvimatį) tiesinės kelių kintamųjų regresijos modelį. Tarkime, kad a. d. X_1 tiesinės regresijos modelis kintamųjų $Z_1, ..., Z_6, Z_7$ atžvilgiu yra

$$\boldsymbol{Y}_1 = (\boldsymbol{A}\vdots\mathbf{1})\begin{pmatrix}\boldsymbol{\beta}_1\\ \cdots\\ \alpha_1\end{pmatrix} + \boldsymbol{e}_1;$$

čia $\boldsymbol{Y}_1 = (X_{11}, ..., X_{1n})^T$, $\boldsymbol{\beta}_1 = (\beta_{11}, ..., \beta_{16})^T$, $\mathbf{1} = (1, ..., 1)^T$, plano matricos $\boldsymbol{A}$ i-oji eilutė yra $(Z_{1i} - \bar{Z}_{1\cdot}, ..., Z_{6i} - \bar{Z}_{6\cdot})$, $i = 1, ..., n$; vektorius $\boldsymbol{e}_1 = (e_{11}, ..., e_{1n})^T \sim N_n(\mathbf{0}, \sigma_{11}\boldsymbol{I})$.

Gauname parametrų įverčius

$$\hat{\alpha}_1 = \bar{X}_{1\cdot} = 1,688; \quad \hat{\boldsymbol{\beta}}_1 = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y}_1 = (0,062;\ -0,160;\ 0,292;\ -0,658;\ 0,173;\ -0,428)^T;$$

čia $\boldsymbol{A}^T\boldsymbol{A}$ yra paskutinioji matrica (išbraukus 7 eilutę ir stulpelį), o $\boldsymbol{A}^T\boldsymbol{Y}_1$ – priešpaskutinės matricos pirmasis stulpelis (be paskutinio elemento). Liekamoji kvadratų suma

$$SS_E(1,1) = \sum_j (X_{1j} - \bar{X}_{1\cdot})^2 - \hat{\boldsymbol{\beta}}_1^T\boldsymbol{A}^T\boldsymbol{Y}_1 = 0,2023.$$

Analogiškai gauname a. d. X_2 tiesinės regresijos modelio

$$\boldsymbol{Y}_2 = (\boldsymbol{A}\vdots\mathbf{1})\begin{pmatrix}\boldsymbol{\beta}_2\\ \cdots\\ \alpha_2\end{pmatrix} + \boldsymbol{e}_2;$$

parametrų įverčius

$$\hat{\alpha}_2 = \bar{X}_{2\cdot} = 2,161; \quad \hat{\boldsymbol{\beta}}_2 = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y}_2 = (0,552;\ -0,278;\ 0,218;\ -0,723;\ 0,323;\ 2,005)^T;$$

čia $\boldsymbol{A}^T\boldsymbol{Y}_2$ yra priešpaskutinės matricos 2 stulpelis (be paskutinio elemento). Liekamoji kvadratų suma

$$SS_E(2,2) = \sum_j (X_{2j} - \bar{X}_{2\cdot})^2 - \hat{\boldsymbol{\beta}}_2^T \boldsymbol{A}^T\boldsymbol{Y}_2 = 1,2997.$$

b) Apskaičiuojme sumą

$$SS_E(1,2) = \sum_j (X_{1j} - \bar{X}_{1\cdot})(X_{2j} - \bar{X}_{2\cdot}) - \hat{\boldsymbol{\beta}}_1^T \boldsymbol{A}^T\boldsymbol{Y}_2 = 0,1334.$$

ir matricos $\boldsymbol{SS}_E = [SS_E(i,j)]_{2\times 2}$ determinantą $|\boldsymbol{SS}_E| = 0,2452$.

Matricos $\boldsymbol{SS}_{EH}$ elementai gaunami analogiškai p. a), jeigu atliekant skaičiavimus išbrauksime paskutinės matricos 2, 3, 4 bei 7 eilutes ir stulpelius. Gauname

$$\boldsymbol{S}_{EH}(1,1) = 0,4215; \quad \boldsymbol{S}_{EH}(2,2) = 1,9143; \quad \boldsymbol{S}_{EH}(1,2) = 0,4878;$$

$$|\boldsymbol{SS}_{EH}| = |[SS_{EH}(i,j)]_{2\times 2}| = 0,5689.$$

ir statistikos $U = |\boldsymbol{SS}_E|/|\boldsymbol{SS}_{EH}|$ realizaciją $U = 0,43097$. Jeigu hipotezė teisinga, tai statistika $Z = -n\ln U$ apytiksliai turi χ^2 skirstinį su $kr = 6$ laisvės laipsniais. Gauname $Z = 21,0429$ ir asimptotinė P reikšmė $pv_a = \mathbf{P}\{\chi^2_6 > 21,0429\} = 0,0018$. Hipotezė atmetama.

c) Regresijos parametrų įverčiai surasti p. b):

$$(\hat{\beta}_{11};\hat{\beta}_{15};\hat{\beta}_{16}) = (0,3911; 0,0110; -0,9572); (\hat{\beta}_{21};\hat{\beta}_{25};\hat{\beta}_{26}) = (0,9158; 0,1461; 1,4889).$$

3.7. Pirmoji pateikta matrica yra matricos $\boldsymbol{SS}_E$ realizacija. Matricos $\boldsymbol{SS}_{EH}$ realizacija lygi abiejų pateiktų matricų sumai. Randame statistikos $U = |\boldsymbol{SS}_E|/|\boldsymbol{SS}_{EH}|$ relizaciją $U = 0,82388$. Jeigu hipotezė teisinga, tai statistika $Z = -n\ln U$ asimptotiškai turi χ^2 skirstinį su $kr = 15$ laisvės laipsnių. Randame statistikos Z realizaciją $Z = 162,731$. Kadangi P reikšmė $pv = \mathbf{P}\{\chi^2_{15} > 162,731\} = 7\cdot 10^{-27}$, tai hipotezę atmetame.

3.8. A. v. $(X_3, X_4)^T$ prognozė remiantis a. v. $(X_1, X_2)^T$ yra

$$X_3 = 183,84 + 0,4503(X_1 - 185,72) + 0,5060(X_2 - 151,12),$$

$$X_4 = 149,24 + 0,2740(X_1 - 185,72) + 0,3784(X_2 - 151,12).$$

A. v. $(X_3, X_4)^T$ sąlyginio skirstinio, kai a. v. $(X_1, X_2)^T$ yra fiksuotas, kovariacinės matricos $\boldsymbol{\Sigma}_{22\cdot 1}$ elementų įverčiai yra tokie: $\sigma_{33\cdot 1} = 41,4114$, $\sigma_{44\cdot 1} = 31,0508$ $\sigma_{34\cdot 1} = 15,6714$. **3.9.** A. v. $(X_1, X_2)^T$ prognozė remiantis a. v. $(X_3, X_4)^T$ yra

$$X_1 = 1906,1 + 0,4122(X_3 - 1509,1) + 0,5467(X_4 - 1725,0),$$

$$X_2 = 1749,5 + 0,4314(X_3 - 1509,1) + 0,4038(X_4 - 1725,0).$$

A. v. $(X_1, X_2)^T$ sąlyginio skirstinio, kai a. v. $(X_3, X_4)^T$ yra fiksuotas, kovariacinės matricos $\boldsymbol{\Sigma}_{11\cdot 2}$ elementų įverčiai yra tokie: $\sigma_{11\cdot 2} = 19455,6$, $\sigma_{22\cdot 2} = 38592,9$, $\sigma_{12\cdot 2} = 38592,9$. **3.10.** a) A. v. $(X_2, X_3)^T$ prognozė remiantis a. v. $(X_1, X_4)^T$ yra

$$X_2 = 272,575 - 2,5359(X_1 - 22,2875) + 0,3184(X_4 - 51,975),$$

$$X_3 = 288,435 + 0,0509(X_1 - 22,2875) - 0,0797(X_4 - 51,975).$$

b) Statistika $U = |\boldsymbol{SS}_E|/|\boldsymbol{SS}_{EH}|$ įgijo reikšmę 0,9269; statistika $-n\ln U$ – reikšmę 3,036; atmesti hipotezę nėra pagrindo. **3.11.** Regresijos tiesių įverčiai:

$$X_1 = 4,640 + 0,0501(X_2 - 45,4), \quad X_3 = 9,965 - 0,0282(X_2 - 45,4).$$

3.12. Tegu $\hat{\boldsymbol{\Sigma}}_{11}, \hat{\boldsymbol{\Sigma}}_{22}\hat{\boldsymbol{\Sigma}}_{1.2}\hat{\boldsymbol{\Sigma}}_{2.2}$ yra matricos $\hat{\boldsymbol{\Sigma}}$ blokai; $\hat{\boldsymbol{\Sigma}}_{11} = [\hat{\sigma}_{ij}]_{2\times 2}, i,j = 1,2$; $\hat{\boldsymbol{\Sigma}}_{22} = [\hat{\sigma}_{ij}]_{4\times 4}, i,j = 3,4,5,6$; $\hat{\boldsymbol{\Sigma}}_{1\cdot 1} = [\hat{\sigma}_{1j}]_{1\times 4}, j = 3,4,5,6$; $\hat{\boldsymbol{\Sigma}}_{2\cdot 1} = [\hat{\sigma}_{1j}]_{1\times 4}, j = 3,4,5,6$. Tada regresijos lygčių įvertiniai yra

$$X_1 = \hat{\mu}_1 + \sum_{j=3}^{6} \hat{\beta}_{1j}(X_j - \hat{\mu}_j), \quad X_2 = \hat{\mu}_2 + \sum_{j=3}^{6} \hat{\beta}_{2j}(X_j - \hat{\mu}_j);$$

vektoriai $\hat{\boldsymbol{\beta}}_1 = (\hat{\beta}_{13}, \hat{\beta}_{14}, \hat{\beta}_{15}, \hat{\beta}_{16})^T$, $\hat{\boldsymbol{\beta}}_2 = (\hat{\beta}_{23}, \hat{\beta}_{24}, \hat{\beta}_{25}, \hat{\beta}_{26})^T$ randami iš lygčių sistemų

$$\hat{\boldsymbol{\beta}}_1 = (\hat{\boldsymbol{\Sigma}}_{22})^{-1}\hat{\boldsymbol{\Sigma}}_{1\cdot 2}, \quad \hat{\boldsymbol{\beta}}_2 = (\hat{\boldsymbol{\Sigma}}_{22})^{-1}\hat{\boldsymbol{\Sigma}}_{2\cdot 2}.$$

3.13. Regresijos lygčių įverčiai:

$$X_1 = 12,2186 + 0,2082(X_3 - 9,5903), \quad X_2 = 8,1125 - 0,3421(X_3 - 9,5903).$$

3.14. Randame

$$\boldsymbol{SS}_E = (n_1-1)\hat{\boldsymbol{\Sigma}}_1 + (n_2-1)\hat{\boldsymbol{\Sigma}}_2 + (n_3-1)\hat{\boldsymbol{\Sigma}}_3 = \begin{pmatrix} 183,093 & 4,417 & 0,995 & 9,677 \\ 4,417 & 8,197 & 0,548 & 2,405 \\ 0,995 & 0,548 & 1,379 & 0,380 \\ 9,677 & 2,405 & 0,380 & 6,681 \end{pmatrix},$$

ir matricą

$$\boldsymbol{B} = \sum_{i=1}^{3} n_i(\bar{\mathbf{X}}_{i.} - \bar{\mathbf{X}}_{..})(\bar{\mathbf{X}}_{i.} - \bar{\mathbf{X}}_{..})^T = \begin{pmatrix} 3,469 & 1,099 & 0,811 & 0,586 \\ 1,099 & 1,231 & 0,450 & 0,616 \\ 0,811 & 0,450 & 0,232 & 0,231 \\ 0,586 & 0,616 & 0,231 & 0,309 \end{pmatrix},$$

čia $\bar{\mathbf{X}}_{..} = (n_1\bar{\mathbf{X}}_{1.} + n_2\bar{\mathbf{X}}_{2.} + n_3\bar{\mathbf{X}}_{3.})/(n_1+n_2+n_3)$. Tada $\boldsymbol{SS}_{EH} = \boldsymbol{SS}_E + \boldsymbol{B}$. Statistika U iš (3.4.11) įgijo reikšmę 0,7628.

Kadangi $Rang(H) = 2$, turime išnagrinėtą atskirą atvejį. Statistika (3.5.15) įgijo reikšmę 18,49, hipotezė atmetama kriterijumi su gana aukštu reikšmingumo lygmeniu.

3.15. a) parametro $\nu_{ii'j}$ pasikliovimo intervalo rėžiai yra

$$\bar{X}_{ij.} - \bar{X}_{i'j.} \mp t_{0,025}(513)\sqrt{SS_E(jj)(n_i+n_{i'})/(n_i n_{i'}(n_1+n_2+n_3-3))};$$

b) intervalo rėžiai gaunami p. a) vietoje $t_{0,025}(513)$ imant $t_{0,025/12}(513)$. Pavyzdžiui, parametrui ν_{121} atveju a) gauname pasikliovimo intervalą $(-0,2414; 0,0394)$, o atveju b) pasikliovimo intervalą $(-0,2942; 0,0922)$.

3.16 $F_A = 5,62$, $F_B = 0,45$, $F_{AB} = 0,21$. Atitinkamos P reikšmės yra 0,0028, 0,8400, 0,9965.

3.17. Randame (žr. **3.1** pratimą)

$$\boldsymbol{SS}_{EH_A} = \begin{pmatrix} 1,7405 & -1,5045 & 0,8555 \\ -1,5045 & 1,3005 & -0,7395 \\ 0,8555 & -0,7395 & 0,4205 \end{pmatrix}, \quad \boldsymbol{SS}_{EH_B} = \begin{pmatrix} 0,7605 & 0,6825 & 1,9305 \\ 0,6825 & 0,6125 & 1,7325 \\ 1,9305 & 1,7325 & 4,9005 \end{pmatrix},$$

$$\boldsymbol{SS}_{EH_{AB}} = \begin{pmatrix} 0,0005 & 0,0165 & 0,0445 \\ 0,0165 & 0,5445 & 1,4685 \\ 0,0445 & 1,4685 & 3,9605 \end{pmatrix}, \boldsymbol{SS}_E = \begin{pmatrix} 1,7640 & 0,0200 & -3,0700 \\ 0,0200 & 2,6280 & -0,5520 \\ -3,0700 & -0,5520 & 64,9240 \end{pmatrix}.$$

Apskaičiuojame determinantų santykių reikšmes

$$U_A = \frac{|\boldsymbol{SS}_E|}{|\boldsymbol{SS}_E + \boldsymbol{SS}_{EH_A}|} = 0,3819, \quad U_B = \frac{|\boldsymbol{SS}_E|}{|\boldsymbol{SS}_E + \boldsymbol{SS}_{EH_B}|} = 0,5230,$$

$$U_{AB} = \frac{|\boldsymbol{SS}_E|}{|\boldsymbol{SS}_E + \boldsymbol{SS}_{EH_{AB}}|} = 0,7771.$$

Kadangi $r = Rang(H)$ visų trijų hipotezių atveju lygus 1, tai turime aptartą atskirą atvejį. Statistikos (3.5.12) įgijo reikšmes $F_A = 7,5543$, $F_B = 4,2556$, $F_{AB} = 1,3385$. Atitinkamos P reikšmės yra 0,0030, 0,0247, 0,3018. Hipotezės dėl faktorių A ir B įtakos nebuvimo atmetamos; hipotezę dėl sąveikos nebuvimo atmesti nėra pagrindo.

3.18. $F_A = 49,40$, $F_B = 56,95$, $F_{AB} = 10,87$. Hipotezės atmetamos kriterijais su gana aukštu reikšmingumo lygmeniu.

3.19. Nagrinėdami Y_1 regresijos modelį atžvilgiu kovariančių $Z_1, \cdots, Z_5$ ir atlikę pažingsninę regresiją, gauname, esant reikšmingumo lygmeniui 0,05, statistiškai reikšmingos yra tik kovariantės Z_1 ir Z_2. Gauname tokius parametrų įverčius $\hat{\alpha}_1 = 56,72$; $\hat{\boldsymbol{\beta}}_1 = (507,07; 0,329)^T$. Nagrinėdami Y_2 regresijos modelį atžvilgiu kovariančių $Z_1, \cdots, Z_5$ ir atlikę pažingsninę regresiją, gauname, esant reikšmingumo lygmeniui 0,05, statistiškai reikšmingos yra tik kovariantės Z_1 ir Z_2. Gauname tokius parametrų įverčius $\hat{\alpha}_2 = -241,35$; $\hat{\boldsymbol{\beta}}_2 = (606,31; 0,324)^T$. Statistika, patikrinti hipotezę, kad regresijos koeficientai prie Z_3, Z_4, Z_5 lygūs 0, įgijo reikšmę 1,69, atitinkama P reikšmė 0,1755.

4 skyrius

Koreliacinė analizė

Remiantis daugiamačio normaliojo skirstinio savybėmis(3 priedas, 2 savybė) galima tvirtinti, kad normalieji a. d. X_1 ir X_2 yra nepriklausomi tada ir tik tada, kai koreliacijos koeficientas

$$\rho = \mathbf{Cov}\,(X_1, X_2)/\sqrt{\boldsymbol{V}X_1\boldsymbol{V}X_2} = 0.$$

Analogiškai, jeigu vektoriaus $\mathbf{X} = (X_1, ..., X_k)^T \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ yra fiksuotos koordinatės $X_{m+1}, ..., X_k$, tai koordinatės X_i ir X_j, $i \neq j$, $i, j = 1, ..., m$, yra nepriklausomos tada ir tik tada, kai koreliacijos koeficientas $\rho_{ij \cdot m+1,...,k}$, apskaičiuotas tarus, kad koordinatės $X_{m+1}, ..., X_k$ fiksuotos (dalinis koreliacijos koeficientas), yra lygus 0. Trečioje vadovėlio dalyje, aptariant teorinius regresijos uždavinius, buvo įvestas dauginis koreliacijos koeficientas, kuris apibūdina vieno normaliojo a. d. X_1 ir normaliųjų a. d. vektoriaus $(X_2, ..., X_k)^T$ priklausomybę.

Šiame skyriuje aptarsime hipotezių dėl koreliacijos koeficiento, dalinio koreliacijos koeficiento ir dauginio koreliacijos koeficiento reikšmių tikrinimo kriterijus. Taip pat bus nagrinėjamos bendresnio tipo normaliųjų atsitiktinių vektorių nepriklausomumo hipotezės.

4.1. Empirinio koreliacijos koeficiento skirstinys

Tegu $\mathbf{X}_1, ..., \mathbf{X}_n$ yra paprastoji imtis a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\mu} = (\mu_1, ..., \mu_k)^T$, $\boldsymbol{\Sigma} = [\sigma_{ij}]_{k\times k}$. Skyrelyje 1.2.1 buvo gautas kovariacinės matricos $\boldsymbol{\Sigma}$ didžiausiojo tikėtinumo įvertinys

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i\mathbf{X}_i^T - \bar{\mathbf{X}}\bar{\mathbf{X}}^T = \frac{1}{n}\boldsymbol{S} = \frac{1}{n}[S_{ij}]_{k\times k} \tag{4.1.1}$$

$$\bar{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i.$$

Kadangi tarp parametrų $\sigma_{ii}, i = 1,...,k,\ \sigma_{ij}, i < j,\ i,j = 1,...,k,$ ir tarp parametrų $\sigma_{ii}, i = 1,...,k,\ \rho_{ij} = \sigma_{ij}/\sqrt{\sigma_{ii}\sigma_{jj}}, i < j,\ i,j = 1,...,k$ yra bijekcija, tai parametro ρ_{ij} DT įvertinys

$$\hat{\rho}_{ij} = r_{ij} = \frac{S_{ij}}{\sqrt{S_{ii}S_{jj}}}, \quad i < j, \quad i,j = 1,...,k.$$

Įvertinys r_{ij} priklauso tik nuo i-osios ir j-osios vektoriaus $\mathbf{X}$ koordinačių matavimų. Todėl, nagrinėjant r_{ij} savybes pakanka apsiriboti dvimačiu normaliuoju vektoriumi, pavyzdžiui, vektoriumi $(X_1, X_2)^T \sim N_2(\boldsymbol{\mu}, \mathbf{\Sigma})$, $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$, $\mathbf{\Sigma} = [\sigma_{ij}]_{2\times 2}$, $\sigma_{11} = \sigma_1^2$, $\sigma_{22} = \sigma_2^2$, $\sigma_{12} = \sigma_{21} = \rho\sigma_1\sigma_2$. Koreliacijos koeficiento įvertinys

$$\hat{\rho} = r = \frac{S_{12}}{\sqrt{S_{11}S_{22}}}, \quad S_{ij} = \sum_{l=1}^{n}(X_{il} - \bar{X}_i)(X_{jl} - \bar{X}_j), \tag{4.1.2}$$

$$\bar{X}_i = \frac{1}{n}\sum_{l=1}^{n} X_{il}, \quad i,j = 1,\ 2.$$

4.1.1 teorema. *Empirinio koreliacijos koeficiento r tankio funkcija, kai $n > 3$, yra*

$$f(r|\rho) = \frac{2^{n-3}(1-\rho^2)^{(n-1)/2}(1-r^2)^{(n-4)/2}}{(n-3)!\pi}\sum_{j=0}^{\infty}\frac{(2\rho r)^j}{j!}\Gamma^2(\frac{n-1+j}{2}). \tag{4.1.3}$$

Įrodymas. Atsitiktinio vektoriaus $(S_{11}, S_{22}, S_{12})^T$ tankio funkcija yra Višarto skirstinio $W_2(n-1, \mathbf{\Sigma})$ tankio funkcija (1.6.13):

$$f(S_{11}, S_{22}, S_{12}|2, n-1, \mathbf{\Sigma}) = \frac{1}{K}|\boldsymbol{S}|^{(n-4)/2}\exp\left\{-\frac{1}{2}Tr(\boldsymbol{S}\mathbf{\Sigma}^{-1})\right\}$$

$$= \frac{1}{K}(S_{11}S_{22} - S_{12})^{(n-4)/2}\exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{S_{11}}{\sigma_1^2} - 2\rho\frac{S_{12}}{\sigma_1\sigma_2} + \frac{S_{22}}{\sigma_2^2}\right)\right\}, \tag{4.1.4}$$

$$K = K(2, n-1, \mathbf{\Sigma}) = 2^{n-1}\sqrt{\pi}(\sigma_1^2\sigma_2^2(1-\rho^2))^{(n-1)/2}\Gamma((n-1)/2)\Gamma((n-2)/2).$$

Tankis apibrėžtas trimatės erdvės srityje, kurioje $S_{11}S_{22} - S_{12}^2 > 0$.

Atsitiktinio dydžio r tankį galime gauti pereidami nuo a. v. $(S_{11}, S_{22}, S_{12})^T$ prie a. v. $(S_{11}, S_{22}, r)^T$, ir integruodami pastarojo tankį pagal S_{11} ir S_{22}. Gauname (pakeitimo jakobianas lygus $\sqrt{S_{11}S_{22}}$) a. v. $(S_{11}, S_{22}, r)^T$ tankį

$$g(S_{11}, S_{22}, r) = \frac{1}{K}(1-r^2)^{\frac{n-4}{2}}(S_{11}S_{22})^{\frac{n-3}{2}}\times$$

$$\times\exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{S_{11}}{\sigma_1^2} - 2\rho\frac{S_{12}}{\sigma_1\sigma_2} + \frac{S_{22}}{\sigma_2^2}\right)\right\}$$

$$= \frac{1}{K}(1-r^2)^{\frac{n-4}{2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{S_{11}}{\sigma_1^2} + \frac{S_{22}}{\sigma_2^2}\right)\right\}\sum_{j=0}^{\infty}\frac{(\rho r)^j}{j!}\frac{(S_{11}S_{22})^{(n-3+j)/2}}{(\sigma_1\sigma_2(1-\rho^2))^j}.$$

Kadangi

$$\int_0^\infty S_{ii}^{(n-3+j)/2}\exp\left\{-\frac{S_{ii}}{2(1-\rho^2)\sigma_i^2}\right\}dS_{ii}=(2\sigma_i^2(1-\rho^2))^{(n-1+j)/2}\Gamma((n-1+j)/2),$$

tai integruodami $g(S_{11},S_{22},r)$ pagal S_{11} ir S_{22} rėžiuose nuo 0 iki ∞ ir, įstatę normuojančio daugiklio K išraišką iš (4.1.4), gauname a. d. r tankio funkciją

$$f(r|\rho)=\frac{(1-\rho^2)^{(n-1)/2}(1-r^2)^{(n-4)/2}}{\sqrt{\pi}\Gamma(\frac{n-1}{2})\Gamma(\frac{n-2}{2})}\sum_{j=0}^{\infty}\frac{(2\rho r)^j}{j!}\Gamma^2\left(\frac{n-1+j}{2}\right). \tag{4.1.5}$$

Remdamiesi gama funkcijos nuo dvigubo argumento išraiška

$$\Gamma(z)\Gamma(z+\frac{1}{2})=\sqrt{\pi}\Gamma(2z)/2^{2z-1},$$

gauname

$$\Gamma((n-1)/2)\Gamma((n-2)/2)=\sqrt{\pi}\Gamma(n-2)/2^{n-3}=\sqrt{\pi}(n-3)!/2^{n-3}.$$

Įstatę šią išraišką į (4.1.5), gauname (4.1.3). ▲

4.1.1 išvada. Jeigu $\rho=0$, tai empirinio koreliacijos koeficiento tankio funkcija

$$f(r|0)=\frac{\Gamma((n-1)/2)}{\sqrt{\pi}\Gamma((n-2)/2)}(1-r^2)^{(n-4)/2},\quad -1<r<1. \tag{4.1.6}$$

Įrodymas. Tankio išraiškoje (4.1.5) įrašome $\rho=0$. ▲

4.1.2 išvada. Jeigu $\rho=0$, tai a. d.

$$t=\sqrt{n-2}\frac{r}{\sqrt{1-r^2}}\sim S(n-2) \tag{4.1.7}$$

turi Stjudento skirstinį su $n-2$ laisvės laipsniais.

Įrodymas. Atlikime kintamųjų keitimą

$$t=\sqrt{n-2}\frac{r}{\sqrt{1-r^2}},\quad r^2=\frac{t^2}{n-2+t^2},\quad 1-r^2=\frac{n-2}{n-2+t^2}.$$

Pakeitimo jakobianas

$$J=\left|\frac{dr}{dt}\right|=\frac{n-2}{(n-2+t^2)^{3/2}}.$$

Įrašę į (4.1.6) gauname a. d. t tankį

$$g(t)=\frac{1}{\sqrt{\pi}\sqrt{n-2}}\frac{\Gamma((n-1)/2)}{\Gamma((n-2)/2)}\left(1+\frac{t^2}{n-2}\right)^{-(n-1)/2},\quad -\infty<t<\infty.$$

Palyginę su pirmos vadovėlio dalies 1 P lentele, matome, kad tai Stjudento skirstinio su $n-2$ laisvės laipsniais tankio funkcija. ▲

4.2. Hipotezių apie koreliacijos koeficiento reikšmes tikrinimas

4.2.1. Nepriklausomumo hipotezės tikrinimas

Tikrinant hipotezę $H : \rho = 0$, kai alternatyvos yra $\bar{H}_1 : \rho > 0$, $\bar{H}_2 : \rho < 0$, arba dvipusė $\bar{H}_3 : \rho \neq 0$, kriterijai gaunami remiantis 4.1.2 išvada. Hipotezė atmetama, kai atitinkamai

$$t > t_\alpha(n-2), \quad t < -t_\alpha(n-2), \quad |t| > t_{\alpha/2}(n-2); \tag{4.2.1}$$

čia $t_\alpha(\nu)$ – Stjudento skirstinio su ν laisvės laipsnių α kritinė reikšmė. Skaičiuojant šių kriterijų galios funkcijas tenka naudoti statistikos r tankio išraišką (4.1.3). Pavyzdžiui,

$$\beta_1(\rho) = \mathbf{P}\{t > t_\alpha(n-2)|\rho\} = \mathbf{P}\{r > t_\alpha(n-2)/\sqrt{n-2+t_\alpha^2(n-2)}|\rho\}$$
$$= 1 - F(t_\alpha(n-2)/\sqrt{n-2+t_\alpha^2(n-2)}|\rho); \tag{4.2.2}$$

čia $F(x|\rho)$ yra statistikos r, kai tikroji koreliacijos koeficiento reikšmė yra ρ, pasiskirstymo funkcija

$$F(x|\rho) = \mathbf{P}\{r \leq x|\rho\} = \int_{-1}^{x} f(u|\rho)du.$$

4.1.1 pavyzdys. (**1.2.1** pavyzdžio tęsinys). Tarę, kad **1.2.1** lentelėje yra trimačio normaliojo vektoriaus realizacijos, patikrinsime hipotezes, kad trijų spindulių srovės stiprumai yra poromis nepriklausomi, kai alternatyva yra jų teigiama priklausomybė.

Naudodamiesi **1.2.1** pavyzdyje surasta matricos $\boldsymbol{S}$ realizacija gauname empirinių koreliacijos koeficientų matricos $\boldsymbol{R}$ realizaciją

$$\boldsymbol{S} = \begin{pmatrix} 0,348 & 0,264 & 0,160 \\ 0,264 & 0,392 & 0,070 \\ 0,160 & 0,070 & 0,240 \end{pmatrix}, \quad \boldsymbol{R} = \begin{pmatrix} 1 & 0,715 & 0,554 \\ 0,715 & 1 & 0,228 \\ 0,554 & 0,228 & 1 \end{pmatrix}.$$

Randame statistikų (4.1.7)

$$t_{ij} = \sqrt{n-2}\frac{r_{ij}}{\sqrt{1-r_{ij}^2}}, \quad i \neq j = 1,2,3,$$

realizacijas

$$t_{12} = 5,408, \quad t_{13} = 3,518, \quad t_{23} = 1,240.$$

Atitinkamos P reikšmės yra 0,000005; 0,0008; 0,113. Taigi a.d. X_1 ir X_2, bei a.d. X_1 ir X_3 nepriklausomumo hipotezės atmetamos. O atmesti a.d. X_2 ir X_3 nepriklausomumo hipotezę nėra pagrindo.

4.2.2. Hipotezės apie koreliacijos koeficiento reikšmes

Tikrinant hipotezę dėl koreliacijos koeficiento reikšmės $H : \rho = \rho_0 \neq 0$, kai alternatyvos yra $\bar{H}_1 : \rho > \rho_0$, $\bar{H}_2 : \rho < \rho_0$, arba dvipusė $\bar{H}_3 : \rho \neq \rho_0$, tikslūs kriterijai gali būti gaunami remiantis empirinio koreliacijos koeficiento tankio išraiška

(4.1.3). P reikšmių terminais kriterijus galime suformuluoti taip: hipotezės atmetamos, jeigu atitinkamai

$$pv = 1 - F(r|\rho_0) < \alpha, \quad pv = F(r|\rho_0) < \alpha, \tag{4.2.3}$$

kai alternatyvos vienpusės ir

$$pv = 2\min(1 - F(r|\rho_0), F(r|\rho_0)) < \alpha,$$

kai alternatyv $\bar{H}_3$ dvipusė; čia r yra gautoji empirinio koreliacijos koeficiento realizacija. Kriterijų galias galime gauti analogiškai (4.2.2).

Remiantis bendra pasikliovimo intervalų sudarymo teorija (žr. I dalį, 4.6 skyrelį), parametro ρ pasikliovimo intervalo $(\underline{\rho}, \overline{\rho})$ su pasikliovimo lygmeniu $Q = 1 - 2P$ rėžiai gaunami sprendžiant ρ atžvilgiu lygtis

$$F(r|\rho) = 1 - P, \quad F(r|\rho) = P. \tag{4.2.4}$$

4.2.3. Apytikslūs kriterijai

Nagrinėjant empirinių momentų funkcijas (1 dalis, 3.5.3 skyrelis) gauta, kad normaliojo skirstinio atveju $n \to \infty$ statistikos r skirstinys artėja į normalųjį

$$\sqrt{n}(r - \rho) \xrightarrow{d} Y \sim N(0, (1 - \rho^2)^2).$$

Tačiau ši aproksimacija gali būti netiksli, kai ρ artimas ± 1 ir a. d. r skirstinys yra labai asimetriškas. Todėl praktiškai dažniau naudojama Fišerio pasiūlyta a. d. r dispersiją stabilizuojanti transformacija

$$th(r) = \frac{1}{2} \ln \frac{1 + r}{1 - r}.$$

Aproksimacija

$$\sqrt{n - 3}(th(r) - th(\rho)) \xrightarrow{d} Z \sim N(0,\ 1) \tag{4.2.5}$$

gana tiksli ir palyginti nedideliems n.

Tarkime, kad reikia patikrinti hipotezę dėl koreliacijos koeficiento reikšmės $H : \rho = \rho_0 \neq 0$, kai alternatyvos yra $\bar{H}_1 : \rho > \rho_0$, $\bar{H}_2 : \rho < \rho_0$, arba dvipusė $\bar{H}_3 : \rho \neq \rho_0$. Įrašykime į (4.2.5) kairiąją pusę hipotetinę reikšmę ρ_0. Kai hipotezė teisinga, gautoji statistika

$$Y = \sqrt{n - 3}(th(r) - th(\rho_0))$$

apytiksliai turi standartinį normalųjį skirstinį. Hipotezė atmetama prieš išvardytas alternatyvas apytiksliu reikšmingumo lygmens α kriterijumi, kai atitinkamai

$$Y > z_\alpha, \quad Y < -z_\alpha, \quad |Y| > z_{\alpha/2}, \tag{4.2.6}$$

čia z_α – standartinio normaliojo skirstinio α kritinė reikšmė. Kriterijaus galia apytiksliai išreiškiama standartinio normaliojo skirstinio pasiskirstymo funkcija. Pavyzdžiui, pirmosios alternatyvos atveju gauname

$$\beta_1(\rho) = \mathbf{P}\{Y > z_\alpha | \rho\} = \mathbf{P}\{\sqrt{n - 3}(th(r) - th(\rho))$$

$$> z_\alpha - \sqrt{n-3}(th(\rho) - th(\rho_0))|\rho\} \approx 1 - \Phi(z_\alpha - \sqrt{n-3}(th(\rho) - th(\rho_0))). \quad (4.2.7)$$

Remiantis aproksimacija (4.2.5) galima užrašyti

$$\mathbf{P}\{-z_P < \sqrt{n-3}\left(\frac{1}{2}th(r) - \frac{1}{2}th(\rho)\right) < z_P|\rho\} \approx Q = 1 - 2P.$$

Išsprendę skliausteliuose parašytas nelygybes ρ atžvilgiu, gauname parametro ρ apytikslio pasikliautinojo intervalo rėžius:

$$\underline{\rho} = \frac{e^{2V_1} - 1}{e^{2V_1} + 1}, \quad \overline{\rho} = \frac{e^{2V_2} - 1}{e^{2V_2} + 1}, \tag{4.2.8}$$

čia

$$V_1 = \frac{1}{2}th(r) - z_P\frac{1}{\sqrt{n-3}}, \quad V_2 = \frac{1}{2}th(r) + z_P\frac{1}{\sqrt{n-3}}.$$

4.2.1 pavyzdys. (**4.1.1** pavyzdžio tęsinys). **4.1.1** pavyzdžio sąlygomis patikrinsime hipotezes $H_1 : \rho_{12} = 1/2$ ir $H_2 : \rho_{13} = 1/2$, kai alternatyvos yra $\bar{H}_1 : \rho_{12} > 1/2$ ir $\bar{H}_2 : \rho_{13} > 1/2$.

Apskaičiuojame

$$th(r_{12}) = \frac{1}{2}\ln\frac{1 + r_{12}}{1 - r_{12}} = 0,897, \quad th(r_{13}) = 0,624, \quad th(0,5) = 0,549,$$

ir atitinkamas statistikos Y reikšmes: 1,808 ir 0,389. Jeigu kriterijaus reikšmingumo lygmuo $\alpha = 0,05$, tai $z_\alpha = 1,645$ ir pirmoji hipotezė atmetama, o antroji priimama. Suradę P reikšmes: $1 - \Phi(1,808) = 0,035$ ir $1 - \Phi(0,389) = 0,349$, darome išvadą, kad pirmoji hipotezė atmetama, jei kriterijaus reikšmingumo lygmuo viršija 0,035, o antrąją hipotezę atmesti nėra pagrindo.

Iliustracijai **4.2.1** lentelėje yra pateiktos apytikslės kriterijaus galios reikšmės taškuose $\rho = 0,5(0,05)0,9$, kai kriterijaus reikšmingumo lygmuo $\alpha = 0,05$.

4.2.1 lentelė. Kriterijaus galios reikšmės

ρ	0,50	0,55	0,60	0,65	0,70	0,75	0,80	0,85	0,90
$\beta_1(\rho)$	0,05	0,178	0,296	0,457	0,644	0,821	0,942	0,992	1,000

Apskaičiuoti pagal (4.2.8) parametrų ρ_{12} ir ρ_{13} asimptotiniai pasikliovimo intervalai, kai pasikliovimo lygmuo $Q = 0,9$, yra tokie:

$$(\underline{\rho}_{12}; \overline{\rho}_{12}) = (0,572; 0,816), \quad (\underline{\rho}_{13}; \overline{\rho}_{13}) = (0,360; 0,702).$$

4.3. Daliniai koreliacijos koeficientai

Daliniai koreliacijos koeficientai – tai sąlyginių skirstinių koreliacijos koeficientai.

Tarę, kad a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $|\boldsymbol{\Sigma}| > 0$, suskaidykime jį į du vektorius $\mathbf{X}^{(1)} = (X_1, ..., X_m)^T$ ir $\mathbf{X}^{(2)} = (X_{m+1}, ..., X_k)^T$, ir tegu $\mathbf{E}(\mathbf{X}^{(1)}) = \boldsymbol{\mu}^{(1)}$, $\mathbf{E}(\mathbf{X}^{(2)}) = \boldsymbol{\mu}^{(2)}$,

$$\boldsymbol{V}(\mathbf{X}) = \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix};$$

čia $\boldsymbol{\Sigma}_{11} = \boldsymbol{V}(\mathbf{X}^{(1)})$, $\boldsymbol{\Sigma}_{22} = \boldsymbol{V}(\mathbf{X}^{(2)})$, $\boldsymbol{\Sigma}_{12} = \mathbf{Cov}\,(\mathbf{X}^{(1)}, \mathbf{X}^{(2)})$, $\boldsymbol{\Sigma}_{21} = \boldsymbol{\Sigma}_{12}^T$.

Remiantis daugiamačio normaliojo skirstinio savybėmis (3 priedas, 10 savybė), sąlyginis a. v. $\mathbf{X}^{(1)}$ skirstinys, kai $\mathbf{X}^{(2)} = \boldsymbol{x}^{(2)}$ fiksuotas, yra m-matis normalusis

$$(\mathbf{X}^{(1)}|\mathbf{X}^{(2)} = \boldsymbol{x}^{(2)}) \sim N_m(\boldsymbol{\mu}^{(1)} + \boldsymbol{B}(\boldsymbol{x}^{(2)} - \boldsymbol{\mu}^{(2)}), \boldsymbol{\Sigma}_{11\cdot 2}), \tag{4.3.1}$$

$$\boldsymbol{B} = \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}, \quad \boldsymbol{\Sigma}_{11\cdot 2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}.$$

Pažymėkime $\sigma_{ij\cdot m+1,\dots,k}$, $i, j = 1, ..., m$, matricos $\boldsymbol{\Sigma}_{11\cdot 2}$ elementus:

$$\boldsymbol{\Sigma}_{11\cdot 2} = [\sigma_{ij\cdot m+1,\dots,k}]_{m\times m}.$$

4.3.1 apibrėžimas. Atsitiktinių dydžių X_i ir X_j *daliniu koreliacijos koeficientu* vektoriaus $(X_{m+1}, ..., X_k)^T$ atžvilgiu vadiname koeficientą

$$\rho_{ij\cdot m+1,\dots,k} = \frac{\sigma_{ij\cdot m+1,\dots,k}}{\sqrt{\sigma_{ii\cdot m+1,\dots,k}\sigma_{jj\cdot m+1,\dots,k}}}, \quad i, j = 1, ..., m. \tag{4.3.2}$$

Koeficientai (4.3.2) apibrėžiami matricos $\boldsymbol{\Sigma}_{11\cdot 2}$ elementais lygiai taip pat kaip ir įprastiniai koreliacijos ρ_{ij} apibrėžiami matricos $\boldsymbol{\Sigma} = [\sigma_{ij}]_{k\times k}$ elementais. Jie apibūdina a. d. $X_1, ..., X_m$ priklausomybę, kai vektorius $\mathbf{X}^{(2)}$ yra fiksuotas, t. y. kai eliminuota a. d. $X_{m+1}, ..., X_k$ įtaka.

Sudarysime koeficientų (4.3.2) empirinius analogus (įvertinius).

Tarkime, $\mathbf{X}_1, ..., \mathbf{X}_n$ yra paprastoji imtis a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $|\boldsymbol{\Sigma}| > 0, n > k$. Tada kovariacinės matricos $\boldsymbol{\Sigma}$ didžiausiojo tikėtinumo įvertinys yra

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^T = \frac{1}{n}\boldsymbol{S} = \frac{1}{n}\begin{pmatrix} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} \\ \boldsymbol{S}_{21} & \boldsymbol{S}_{22} \end{pmatrix}, \quad \bar{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i,$$

čia $\boldsymbol{S}_{ij}$, $i, j = 1, 2$, yra matricos $\boldsymbol{S}$ blokai, atitinkantys matricos $\boldsymbol{\Sigma}$ blokus $\boldsymbol{\Sigma}_{ij}$, $i, j = 1, 2$.

Kadangi tarp $\boldsymbol{\Sigma}$ ir $\boldsymbol{\Sigma}_{11\cdot 2}, \boldsymbol{B}, \boldsymbol{\Sigma}_{22}$ yra abipusė vienareikšmė priklausomybė:

$$\boldsymbol{\Sigma}_{12} = \boldsymbol{B}\boldsymbol{\Sigma}_{22}, \quad \boldsymbol{\Sigma}_{11} = \boldsymbol{\Sigma}_{11\cdot 2} + \boldsymbol{B}\boldsymbol{\Sigma}_{22}\boldsymbol{B}^T,$$

tai parametrų $\boldsymbol{\Sigma}_{11\cdot 2}, \boldsymbol{B}, \boldsymbol{\Sigma}_{22}$ didžiausiojo tikėtinumo įvertiniai gaunami tiesiog iš $\hat{\boldsymbol{\Sigma}}$ įvertinio

$$\begin{aligned}\hat{\boldsymbol{\Sigma}}_{11\cdot 2} &= \hat{\boldsymbol{\Sigma}}_{11} - \hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1}\hat{\boldsymbol{\Sigma}}_{21} = \frac{1}{n}\boldsymbol{S}_{11\cdot 2} = \frac{1}{n}(\boldsymbol{S}_{11} - \boldsymbol{S}_{12}\boldsymbol{S}_{22}^{-1}\boldsymbol{S}_{21}) \\ &= \frac{1}{n}[S_{ij\cdot m+1,\dots,k}]_{m\times m},\end{aligned} \tag{4.3.3}$$

$$\hat{\boldsymbol{B}} = \hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1} = \boldsymbol{S}_{12}\boldsymbol{S}_{22}^{-1}, \quad \hat{\boldsymbol{\Sigma}}_{22} = \frac{1}{n}\boldsymbol{S}_{22}.$$

Dalinių koreliacijos koeficientų DT įvertiniai yra

$$\hat{\rho}_{ij\cdot m+1,\dots,k} = r_{ij\cdot m+1,\dots,k} = \frac{S_{ij\cdot m+1,\dots,k}}{\sqrt{S_{ii\cdot m+1,\dots,k}S_{jj\cdot m+1,\dots,k}}}, \quad i, j = 1, ..., m. \tag{4.3.4}$$

4.3.1 teorema. *Daliniai empiriniai koreliacijos koeficientai (4.3.4) turi tokias pat savybes kaip ir įprastiniai empiriniai koreliacijos koeficientai r_{ij}, jeigu atliksime tokius pakeitimus: vietoje matricos $\mathbf{\Sigma}$ imsime matricą $\mathbf{\Sigma}_{11\cdot 2}$, vietoje matricos $\boldsymbol{S}$ matricą $\boldsymbol{S}_{11\cdot 2}$ ir imties dydį sumažinsime $k-m$ vienetų, t. y. vietoje n imsime $n-k+m$.*

Įrodymas. Remiantis antrąja Višarto skirstinio savybe, matricos $\boldsymbol{S}$ elementų skirstinys yra Višarto skirstinys $W_k(n-1,\mathbf{\Sigma})$, t. y. matricą $\boldsymbol{S}$ galima užrašyti suma

$$\boldsymbol{S}=\sum_{i=1}^{n-1}\boldsymbol{Z}_i\boldsymbol{Z}_i^T,$$

čia $\boldsymbol{Z}_1,...,\boldsymbol{Z}_{n-1}$ yra nepriklausomi vienodai pasikirstę a. v. $\boldsymbol{Z}_i\sim N_k(\mathbf{0},\mathbf{\Sigma})$.

Visos empirinių koreliacijos koeficientų r_{ij} savybės, pateiktos 4.1, 4.2 skyreliuose, buvo gautos remiantis vien tik Višarto matricos $\boldsymbol{S}$ ir jos elementų savybėmis.

Pagal šeštą Višarto skirstinio savybę matrica $\boldsymbol{S}_{11\cdot 2}=\boldsymbol{S}_{11}-\boldsymbol{S}_{12}\boldsymbol{S}_{22}^{-1}\boldsymbol{S}_{21}\sim W_m(n-1-k+m,\mathbf{\Sigma}_{11\cdot 2})$, t. y.

$$\boldsymbol{S}_{11\cdot 2}=\sum_{i=1}^{n-1-k+m}\boldsymbol{Y}_i\boldsymbol{Y}_i^T,$$

čia $\boldsymbol{Y}_1,...,\boldsymbol{Y}_{n-1-k+m}$ nepriklausomi vienodai pasikirstę a. v. $\boldsymbol{Y}_i\sim N_m(\mathbf{0},\mathbf{\Sigma}_{11\cdot 2})$.

Todėl matricos $\boldsymbol{S}_{11\cdot 2}$ elementų ar jų funkcijų savybės yra analogiškos matricos $\boldsymbol{S}$ elementų ar jų funkcijų savybėms. ▲

Darant išvadas apie dalinius koreliacijos koeficientus $\rho_{ij\cdot m+1,...,k}$ tinka 4.1 ir 4.2 skyrelių metodai ir rezultatai, atlikus tokius pakeitimus: vietoje įprastinio koreliacijos koeficiento ρ_{ij} reikia imti dalinį koreliacijos koeficientą $\rho_{ij\cdot m+1,...,k}$; o vietoje empirinio koreliacijos koeficiento r_{ij} – dalinį empirinį koreliacijos koeficientą $r_{ij\cdot m+1,...,k}$; imties didumą n reikia sumažinti iki $n-k+m$. Pavyzdžiui, galima tvirtinti, kad

$$t_{ij}=\sqrt{n-2-k+m}\frac{r_{ij\cdot m+1,...,k}}{\sqrt{1-r^2_{ij\cdot m+1,...,k}}}\sim S(n-2-k+m), \qquad (4.3.5)$$

turi Stjudento skirstinį su $n-2-k+m$ laisvės laipsnių, jeigu $\rho_{ij\cdot m+1,...,k}=0$.

6.4.3 pastaba. Dalinius koreliacijos koeficientus apibrėžėme kaip sąlyginius, kai vektorius $\mathbf{X}^{(2)}=\boldsymbol{x}^{(2)}$ yra fiksuotas. Tačiau matricos $\boldsymbol{S}_{11\cdot 2}$ skirstinys nepriklauso nuo fiksuotos reikšmės $\boldsymbol{x}^{(2)}$. Todėl jį galime traktuoti kaip besąlyginį. Taigi ir empirinių dalinių koreliacijos koeficientų $r_{ij\cdot m+1,...,k}$ skirstiniai nepriklauso nuo fiksuotos reikšmės $\boldsymbol{x}^{(2)}$ ir juos galime traktuoti kaip besąlyginius.

4.3.1 pavyzdys. (**4.1.1** pavyzdžio tęsinys). **4.1.1** pavyzdžio sąlygomis įvertinsime dalinį a. d. X_2 ir X_3 koreliacijos koeficientą, kai a. d. X_1 yra fiksuotas, t. y. koeficientą $\rho_{23\cdot 1}$, ir patikrinsime hipotezę $H:\rho_{23\cdot 1}=0$, kai alternatyva $\bar{H}:\rho_{23\cdot 1}\neq 0$.

Šiuo atveju matrica

$$\boldsymbol{S}_{23\cdot 1} = \boldsymbol{S}_{22} - \frac{1}{S_{11}}\boldsymbol{S}_{21}\boldsymbol{S}_{12} = \begin{pmatrix} S_{22} - S_{12}^2/S_{11} & S_{23} - S_{12}S_{13}/S_{11} \\ S_{23} - S_{12}S_{13}/S_{11} & S_{33} - S_{13}^2/S_{11} \end{pmatrix}$$

ir empirinis dalinis koreliacijos koeficientas $r_{23\cdot 1}$ išreiškiamas empiriniais koreliacijos koeficientais:

$$r_{23\cdot 1} = \frac{r_{23} - r_{12}r_{13}}{\sqrt{(1-r_{12}^2)(1-r_{13}^2)}}.$$

Naudodamiesi **4.1.1** pratime surasta empirinių koreliacijos koeficientų matrica $\boldsymbol{R}$, randame $\rho_{23\cdot 1}$ įvertinio realizaciją $r_{23\cdot 1} = -0,240$.

Reikia pažymėti, kad empirinis koreliacijos koeficientas $r_{23} = +0,228$, t. y. jis rodo teigiamą priklausomybę, o dalinis empirinis koreliacijos koeficientas $r_{23\cdot 1} = -0,240$ atspindi neigiamą priklausomybę. Tai gali būti paaiškinta tuo, kad teigiama X_2 ir X_3 priklausomybė, gal būt, yra padarinys to, kad a. d. X_2 ir X_3 gana stipriai teigiamai priklausomi nuo a. d. X_1. Todėl, eliminavus a. d. X_1 įtaką, vaizdas iš esmės pasikeičia ir teigiama priklausomybė gali virsti neigiama. Šis pavyzdys iliustruoja tą faktą, kad tiriant a. d. sąryšius nepakanka apsiriboti koreliacijos koeficientais ρ_{ij}.

Pereiname prie hipotezės H tikrinimo. Randame statistikos t_{23} iš (4.3.5) realizaciją

$$t_{23} = \sqrt{n-2-k+m}\,\frac{r_{23\cdot 1}}{\sqrt{1-r_{23\cdot 1}^2}} = \sqrt{27}\,\frac{-0,24}{\sqrt{1-0,24^2}} = -1,285,$$

ir P reikšmę $pv = 2\mathbf{P}\{t > 1,285\} = 0,210$, čia t yra a. d., turintis Stjudento skirstinį su 27 laisvės laipsniais. Atmesti hipotezę nėra pagrindo.

4.4. Dauginis koreliacijos koeficientas

Tarkime reikia prognozuoti pirmąją a. v. $\mathbf{X} = (X_1, ..., X_k)^T \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}), |\boldsymbol{\Sigma}| > 0$, koordinatę X_1 remiantis vektoriumi $\mathbf{X}^{(2)} = (X_2, ..., X_k)^T$. Tada geriausioji prognozė (žr. vadovėlio 3 dalies 3.1 skyrelį) yra tiesinė regresija

$$\mu(\boldsymbol{x}^{(2)}) = \mathbf{E}(X_1|\mathbf{X}^{(2)} = \boldsymbol{x}^{(2)}) = \mu_1 + \boldsymbol{\beta}^T(\boldsymbol{x}^{(2)} - \boldsymbol{\mu}^{(2)}), \tag{4.4.1}$$

čia $\mu_1 = \mathbf{E}X_1$, $\boldsymbol{\mu}^{(2)} = \mathbf{E}(\mathbf{X}^{(2)})$, $\boldsymbol{\beta} = \boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\sigma}_{(1)}$; $\boldsymbol{V}X_1 = \sigma_{11}$, $\boldsymbol{V}(\mathbf{X}^{(2)}) = \boldsymbol{\Sigma}_{22}$, $\boldsymbol{\sigma}_{(1)} = \mathbf{Cov}\,(X_1, \mathbf{X}^{(2)})$.

4.4.1 apibrėžimas. Koreliacijos koeficientas tarp a. d. X_1 ir geriausiai parinktos prognozės $\hat{X}_1 = \mu(\mathbf{X}^{(2)})$ vadinamas *dauginiu koreliacijos koeficientu*

$$R = \frac{\mathbf{Cov}\,(X_1, \hat{X}_1)}{\sqrt{\boldsymbol{V}(X_1)\boldsymbol{V}(\hat{X}_1)}} = \sqrt{\frac{\boldsymbol{\beta}^T\boldsymbol{\Sigma}_{22}\boldsymbol{\beta}}{\sigma_{11}}} = \sqrt{\frac{\boldsymbol{\sigma}_{(1)}^T\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\sigma}_{(1)}}{\sigma_{11}}}. \tag{4.4.2}$$

Buvo įrodyta (žr. vadovėlio 3 dalies 3.1 skyrelį), kad a. d. X_1 ir tiesinės a. v. $\mathbf{X}^{(2)}$ funkcijos koreliacijos koeficientas yra maksimalus, kai tiesinė funkcija sutampa su $\mu(\mathbf{X}^{(2)})$. Kadangi $\mu(\mathbf{X}^{(2)})$ ir $(X_1 - \mu(\mathbf{X}^{(2)}))$ nekoreliuoti, tai galioja išdėstymas

$$\boldsymbol{V}X_1 = \boldsymbol{V}(\mu(\mathbf{X}^{(2)})) + \boldsymbol{V}(X_1 - \mu(\mathbf{X}^{(2)})),$$
$$\sigma_{11} = \sigma_{11}R^2 + \sigma_{11}(1 - R^2).$$

Iš čia gauname dauginio koreliacijos koeficiento interpretaciją. Dauginio koreliacijos koeficiento kvadratas R^2 parodo, kuri a. d. X_1 dispersijos σ_{11} dalis tenka geriausiai prognozei $\hat{X}_1$. Likusios, nekoreliuotos su pirmąja komponentės $X_1 - \hat{X}_1$, vadinamos prognozės paklaida, dispersija sudaro $1 - R^2$ a. d. X_1 dispersijos σ_{11} dalį. Jeigu $R^2 = 1$, tai reiškia, kad X_1 ir $\mathbf{X}^{(2)}$ susieti tiesine priklausomybe. Jeigu $R^2 = 0$, tai X_1 ir vektorius $\mathbf{X}^{(2)}$ yra nepriklausomi, ir X_1 prognozė remiantis $\mathbf{X}^{(2)}$ neturi prasmės.

Rasime dauginio koreliacijos koeficiento empirinį analogą.

Tegu $\mathbf{X}_1, ..., \mathbf{X}_n$ yra paprastoji a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \mathbf{\Sigma}), |\mathbf{\Sigma}| > 0, n > k$, imtis. Tada analogiškai kaip pirmesniame skyrelyje kovariacinės matricos $\mathbf{\Sigma}$ didžiausiojo tikėtinumo įvertinys

$$\hat{\mathbf{\Sigma}} = \frac{1}{n}\boldsymbol{S} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^T = \frac{1}{n}\begin{pmatrix} S_{11} & \boldsymbol{S}_{(1)}^T \\ \boldsymbol{S}_{(1)} & \boldsymbol{S}_{22} \end{pmatrix};$$

čia $\boldsymbol{S}_{(1)} = (S_{12}, ..., S_{1k})^T$, o matricos $\mathbf{\Sigma}_{22}$ įvertinys yra $\boldsymbol{S}_{22}/n$.

Regresijos koeficientų vektoriaus $\boldsymbol{\beta} = (\beta_2, ..., \beta_k)^T$ ir dauginio koreliacijos koeficiento DT įvertiniai

$$\hat{\boldsymbol{\beta}} = \boldsymbol{S}_{22}^{-1}\boldsymbol{S}_{(1)}, \quad \hat{R} = \sqrt{\frac{\hat{\boldsymbol{\beta}}^T\boldsymbol{S}_{22}\hat{\boldsymbol{\beta}}}{S_{11}}} = \sqrt{\frac{\boldsymbol{S}_{(1)}^T\boldsymbol{S}_{22}^{-1}\boldsymbol{S}_{(1)}}{S_{11}}}. \tag{4.4.3}$$

Pasinaudoję lygybe

$$|\boldsymbol{S}| = |\boldsymbol{S}_{22}|(S_{11} - \boldsymbol{S}_{(1)}^T\boldsymbol{S}_{22}^{-1}\boldsymbol{S}_{(1)}),$$

skirtumą $1 - \hat{R}^2$ galime užrašyti tokiu pavidalu

$$1 - \hat{R}^2 = \frac{|\boldsymbol{S}|}{S_{11}|\boldsymbol{S}_{22}|}. \tag{4.4.4}$$

4.4.1 teorema. *Jeigu dauginis koreliacijos koeficientas $R = 0$, $|\mathbf{\Sigma}| > 0, n > k$, tai santykis*

$$F = \frac{\hat{R}^2(n-k)}{(1-\hat{R}^2)(k-1)} \sim F(k-1, n-k) \tag{4.4.5}$$

turi Fišerio skirstinį su $k-1$ ir $n-k$ laisvės laipsnių.

Įrodymas. Matrica $\boldsymbol{S} \sim W_k(n-1, \mathbf{\Sigma})$, todėl ją galima užrašyti tokiu pavidalu

$$\boldsymbol{S} = \sum_{i=1}^{n-1}\boldsymbol{Z}_i\boldsymbol{Z}_i^T;$$

čia $\boldsymbol{Z}_1, ..., \boldsymbol{Z}_{n-1}$ nepriklausomi vienodai pasiskirstę a. v. $\boldsymbol{Z}_i \sim N_k(\mathbf{0}, \mathbf{\Sigma})$. Pažymėkime vektoriaus $\boldsymbol{Z}_i$ koordinates $Z_{1i}, ..., Z_{ki}$, $\boldsymbol{Z}_i = (Z_{1i}, ..., Z_{ki})^T$, $i = 1, ..., n-1$.

Prognozuodami vektorių $\boldsymbol{Y}_1 = (Z_{11}, ..., Z_{1,n-1})^T$ pagal $\boldsymbol{Y}_j = (Z_{j1}, ..., Z_{j,n-1})$, $j = 2, ..., k$, turime tiesinį Gauso ir Markovo modelį

$$\boldsymbol{Y}_1 = \boldsymbol{A}\boldsymbol{\beta} + \boldsymbol{e},$$

čia $\boldsymbol{\beta} = (\beta_2, ..., \beta_k)^T$ – nežinomų parametrų vektorius, $\boldsymbol{A}$ plano matrica, kurios i-oji eilutė yra $(Z_{2i}, ..., Z_{ki})^T$, paklaidų vektorius $\boldsymbol{e} \sim N_{n-1}(\boldsymbol{0}, \sigma^2\boldsymbol{I})$. Dispersija $\sigma^2 = 1/\sigma^{11}$, čia σ^{11} yra matricos $\boldsymbol{\Sigma}^{-1}$ elementas, atitinkantis matricos $\boldsymbol{\Sigma}$ elementą σ_{11}.

Liekamoji kvadratų suma (žr. 3 dalį 3.1 skyrelį)

$$SS_E = \min_{\boldsymbol{\beta}}(\boldsymbol{Y}_1 - \boldsymbol{A}\boldsymbol{\beta})^T(\boldsymbol{Y}_1 - \boldsymbol{A}\boldsymbol{\beta}) = \frac{1}{S^{11}} = \frac{|\boldsymbol{S}|}{|\boldsymbol{S}_{22}|} \sim \sigma^2\chi^2_{n-k}.$$

Tikriname hipotezę $H : \boldsymbol{\beta} = \boldsymbol{0}$ (tai ekvivalentu, kad dauginis koreliacijos koeficientas $R = 0$). Sąlyginis kvadratinės formos minimumas

$$SS_{EH} = \min_{\boldsymbol{\beta}=\boldsymbol{0}}(\boldsymbol{Y}_1 - \boldsymbol{A}\boldsymbol{\beta})^T(\boldsymbol{Y}_1 - \boldsymbol{A}\boldsymbol{\beta}) = \boldsymbol{Y}_1^T\boldsymbol{Y}_1 = S_{11}.$$

Remiantis 3 dalies 1.3.2 teorema, kai hipotezė H teisinga, $SS_{EH} - SS_E$ ir SS_E yra nepriklausomi ir $SS_{EH} - SS_E \sim \sigma^2\chi^2_{k-1}$. Taigi santykis

$$\frac{(SS_{EH} - SS_E)(n-k)}{(k-1)SS_E} \sim F(k-1, n-k). \tag{4.4.6}$$

Remdamiesi (4.4.4) įsitikiname, kad

$$\frac{SS_{EH} - SS_E}{SS_E} = \frac{\hat{R}^2}{1 - \hat{R}^2}. \qquad \blacktriangle$$

4.4.2 teorema. *Jeigu $R \neq 0$, tai 4.4.1 teoremos sąlygomis statistikos F, apibrėžtos (4.4.5) formule, tankio funkcija*

$$f(F|R) = \frac{k-1}{(n-k)\Gamma(\frac{n-k}{2})}\sum_{j=0}^{\infty}\frac{(\frac{(k-1)F}{n-k})^{(k-1)/2+j-1}\Gamma(\frac{n-1}{2}+j)}{j!\Gamma(\frac{k-1}{2}+j)(1+\frac{(k-1)F}{n-k})^{(n-1)/2+j}}\times$$

$$\times\frac{\varphi^j\Gamma(\frac{n-1}{2}+j)}{(1+\varphi)^{(n-1)/2+j}\Gamma(\frac{n-1}{2})}, \quad \varphi = \frac{R^2}{1-R^2}. \tag{4.4.7}$$

Įrodymas. Jeigu hipotezė $H : \boldsymbol{\beta} = \boldsymbol{0}$ neteisinga, tai santykio (4.4.6) skirstinys yra necentrinis Fišerio skirstinys su $k-1$ ir $n-k$ laisvės laipsnių ir necentriškumo parametru

$$\delta = \frac{\boldsymbol{\beta}^T\boldsymbol{S}_{22}\boldsymbol{\beta}}{\sigma^2} = \frac{\boldsymbol{\beta}^T\sum_{i=1}^{n-1}\boldsymbol{U}_i\boldsymbol{U}_i^T\boldsymbol{\beta}}{\sigma^2}, \tag{4.4.8}$$

čia $\boldsymbol{U}_1, ..., \boldsymbol{U}_{n-1}$ yra nepriklausomi vienodai pasiskirstę a. v. $\boldsymbol{U}_i \sim N_{k-1}(\boldsymbol{0}, \boldsymbol{\Sigma}_{22})$.

Necentrinio Fišerio skirstinio tankio funkcija (žr. 1 dalį, 1 P lentelę) yra

$$f(F|k-1,n-k,\delta) = \frac{(k-1)e^{-\delta/2}}{(n-k)\Gamma(\frac{n-k}{2})}\sum_{j=0}^{\infty}\frac{(\frac{\delta}{2})^j(\frac{(k-1)F}{n-k})^{(k-1)/2+j-1}\Gamma(\frac{n-1}{2}+j)}{j!\Gamma(\frac{k-1}{2}+j)(1+\frac{(k-1)F}{n-k})^{(n-1)/2+j}} \tag{4.4.9}$$

Taigi statistikos (4.4.5) skirstinys priklauso nuo fiksuotųjų reikšmių Z_{ij}, $i = 2,...,k$, $j = 1,...,n-1$. Norint rasti besąlyginį tankį, reikia (4.4.9) dešiniąją pusę suvidurkinti pagal fiksuotųjų a. v. bendrą skirstinį. Tačiau ta priklausomybė yra tik per necentriškumo parametrą δ, kurio skirstinys (žr. Višarto skirstinio apibrėžimą)

$$\delta = \frac{\boldsymbol{\beta}^T\boldsymbol{S}_{22}\boldsymbol{\beta}}{\sigma^2} \sim \frac{\boldsymbol{\beta}^T\boldsymbol{\Sigma}_{22}\boldsymbol{\beta}}{\sigma^2}\chi^2_{n-1} \sim \frac{R^2}{1-R^2}\chi^2_{n-1}.$$

Pažymėjus $\varphi = R^2/(1-R^2)$, tankis

$$f(F|R) = \int_0^\infty f(F|k-1,n-k,\varphi x)g(x|n-1)dx,$$

čia $g(x|n-1)$ yra χ^2 skirstinio su $n-1$ laisvės laipsniu tankio funkcija. Iš (4.4.9) matome, kad, norint rasti $f(F|R)$, pakanka apskaičiuoti vidurkius

$$\kappa_j = \mathbf{E}\left(\left(\frac{\varphi\chi^2_{n-1}}{2}\right)^j\exp\left\{-\frac{\varphi\chi^2_{n-1}}{2}\right\}\right), \quad j = 0,1,2,...,$$

ir įrašyti juos į (4.4.10). Gauname

$$\begin{aligned}\kappa_j &= \frac{1}{2^{(n-1)/2}\Gamma(\frac{n-1}{2})}\int_0^\infty x^{\frac{n-3}{2}}(x\varphi/2)^j e^{-\frac{x}{2}(1+\varphi)}dx \\ &= \frac{\Gamma(\frac{n-1}{2}+j)}{\Gamma(\frac{n-1}{2})}\frac{\varphi^j}{(1+\varphi)^{(n-1)/2+j}}, \quad j = 0,1,2,.... \qquad \blacktriangle\end{aligned}$$

Tikrindami hipotezę $H: R = 0$, kuri ekvivalenti hipotezei, kad visi regresijos koeficientai $\beta_2 = \beta_3 = ... = \beta_k = 0$, kai alternatyva yra $\bar{H}: R \neq 0$, naudojame sąryšį (4.4.5). Hipotezė atmetama, kai

$$F = \frac{\hat{R}^2(n-k)}{(1-\hat{R}^2)(k-1)} > F_\alpha(k-1,n-k). \tag{4.4.10}$$

Kiekviename matematinės statistikos TPP, kuriame yra modulis, skirtas regresinei analizei, pagrindiniu regresijos prasmingumo matu imamas empirinis dauginis koreliacijos koeficientas. Yra pateikiama statistikos $\hat{R}^2$ įgytoji reikšmė ir P reikšmė

$$pv = \mathbf{P}\left\{F_{k-1,n-k} > \frac{\hat{R}^2(n-k)}{(1-\hat{R}^2)(k-1)}\right\},$$

čia $\hat{R}^2$ yra empirinio dauginio koreliacijos koeficiento realizacija. Hipotezė atmetama, kai teisinga nelygybė

$$pv < \alpha,$$

kuri yra ekvivalenti nelygybei (4.4.10).

Kriterijaus galia taškuose $R \neq 0$ išreiškiama integralu nuo tankio (4.4.7).

4.4.1 pavyzdys.(**4.1.1** pavyzdžio tęsinys). Tarkime, kad **4.1.1** pavyzdžio sąlygomis prognozuojame pirmąją koordinatę X_1 remiantis vektoriumi $(X_2, X_3)^T$. Rasime dauginio koreliacijos koeficiento kvadrato įvertinio $\hat{R}^2$ realizaciją ir patikrinsime hipotezę $H : R^2 = 0$.

Naudodamiesi matricos $\boldsymbol{S}$ realizacija, surasta **4.1.1** pavyzdyje, apskaičiuojame statistikos $\hat{R}^2$ realizaciją

$$\hat{R}^2 = \frac{\boldsymbol{S}_{(1)}^T \boldsymbol{S}_{22}^{-1} \boldsymbol{S}_{(1)}}{S_{11}} = 0,672.$$

Tada statistikos F apibrėžtos (4.4.10) realizacija ir P reikšmė yra

$$F = \frac{\hat{R}^2(n-k)}{(1-\hat{R}^2)(k-1)} = 26,84, \quad pv = \mathbf{P}\{F_{2,27} > 26,84\} = 4 \cdot 10^{-7}.$$

Hipotezė H atmetama aukšto reikšmingumo lygmens kriterijumi. Kitaip tariant, a.d. X_1 prognozavimas pagal $(X_2, X_3)^T$ yra prasmingas.

4.5. Atsitiktinių vektorių nepriklausomumo hipotezės

Sudarysime kriterijus bendresnėms nepriklausomumo hipotezėms tikrinti.

4.5.1. Nepriklausomumo hipotezių formulavimas

Tegu a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Suskaidykime vektorių $\mathbf{X}$ į m mažesnės dimensijos k_j, $k_1 + ... + k_m = k$, vektorius $\mathbf{X}^{(1)}, ..., \mathbf{X}^{(m)}$, sudarytus iš skirtingų pradinio vektoriaus $\mathbf{X}$ koordinačių $\mathbf{X} = ((\mathbf{X}^{(1)})^T, ..., (\mathbf{X}^{(m)})^T)^T$. Tada vidurkių vektorius $\boldsymbol{\mu}$ taip pat bus suskaidytas į mažesnės dimensijos vektorius $\boldsymbol{\mu}^{(1)}, ..., \boldsymbol{\mu}^{(m)}$, o kovariacinė matrica $\boldsymbol{\Sigma}$ bus sudalyta į blokus

$$\boldsymbol{\Sigma} = \boldsymbol{V}(\mathbf{X}) = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \cdots & \boldsymbol{\Sigma}_{1m} \\ \cdots & \cdots & \cdots \\ \boldsymbol{\Sigma}_{m1} & \cdots & \boldsymbol{\Sigma}_{mm} \end{pmatrix}. \tag{4.5.1}$$

Remiantis daugiamačio normaliojo vektoriaus savybėmis (3 priedas, 3 savybė), atsitiktiniai vektoriai $\mathbf{X}^{(1)}, ..., \mathbf{X}^{(m)}$ nepriklausomi tada ir tik tada, kai kovariacijų matricos $\boldsymbol{\Sigma}_{ij} = \mathbf{0}, i \neq j,\ i, j = 1, ..., m$. Taigi a. v. $\mathbf{X}^{(1)}, ..., \mathbf{X}^{(m)}$ nepriklausomumo hipotezę H galime suformuluoti taip $H : \boldsymbol{\Sigma}_{ij} = \mathbf{0}, i \neq j,\ i, j = 1, ..., m$, arba $H : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$

$$\boldsymbol{\Sigma}_0 = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots \\ \mathbf{0} & \cdots & \boldsymbol{\Sigma}_{mm} \end{pmatrix}. \tag{4.5.2}$$

4.5.2. Tikėtinumų santykio statistika

Tarkime, kad $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n$ yra paprastoji imtis a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

4.5.1 teorema. *Tarkime, kad $|\boldsymbol{\Sigma}| > 0, n > k$. Tikėtinumų santykio statistika hipotezei H tikrinti turi tokį pavidalą:*

$$\Lambda = \left(\frac{|\hat{\boldsymbol{\Sigma}}|}{|\hat{\boldsymbol{\Sigma}}_{11}| \cdots |\hat{\boldsymbol{\Sigma}}_{mm}|}\right)^{n/2} = \left(\frac{|\boldsymbol{S}|}{|\boldsymbol{S}_{11}| \cdots |\boldsymbol{S}_{mm}|}\right)^{n/2}, \tag{4.5.3}$$

čia $\hat{\boldsymbol{\Sigma}}$ ir $\hat{\boldsymbol{\Sigma}}_{ii}, i = 1, ..., m$, yra kovariacinės matricos ir jos blokų DT įvertiniai.

Įrodymas. Tikėtinumų santykis Λ hipotezei H tikrinti yra

$$\Lambda = \frac{\max_{\boldsymbol{\mu}, \boldsymbol{\Sigma}=\boldsymbol{\Sigma}_0} L(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\max_{\boldsymbol{\mu}, \boldsymbol{\Sigma}} L(\boldsymbol{\mu}, \boldsymbol{\Sigma})} = \frac{L(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}})}{L(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})},$$

čia $\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}$ yra parametrų $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ didžiausiojo tikėtinumo įvertiniai, o $\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}$ yra šių parametrų DT įvertiniai surasti esant sąlygai, kad hipotezė H yra teisinga.

Remiantis 1.1 skyreliu parametrų $\boldsymbol{\mu}$ ir $\boldsymbol{\Sigma}$ didžiausiojo tikėtinumo įvertiniai yra

$$\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i, \quad \hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n} (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^T = \frac{1}{n}\boldsymbol{S},$$

o tikėtinumo funkcijos maksimumas

$$L(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}) = (2\pi)^{-nk/2}(|\hat{\boldsymbol{\Sigma}}|)^{-n/2} \exp\{-nk/2\}.$$

Analogiškai skaitiklyje gauname

$$L(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}) = \max_{\boldsymbol{\mu}^{(i)}, \boldsymbol{\Sigma}_{ii}} \prod_{i=1}^{m} L(\boldsymbol{\mu}^{(i)}, \boldsymbol{\Sigma}_{ii}) = \prod_{i=1}^{m} L(\hat{\boldsymbol{\mu}}^{(i)}, \hat{\boldsymbol{\Sigma}}_{ii}),$$

$$\hat{\boldsymbol{\mu}}_i = \bar{\mathbf{X}}^{(i)} = \frac{1}{n}\sum_{j=1}^{n} \mathbf{X}_j^{(i)}, \quad \hat{\boldsymbol{\Sigma}}_{ii} = \frac{1}{n}\sum_{j=1}^{n} (\mathbf{X}_j^{(i)} - \bar{\mathbf{X}}^{(i)})(\mathbf{X}_j^{(i)} - \bar{\mathbf{X}}^{(i)})^T = \frac{1}{n}\boldsymbol{S}_{ii},$$

$$L(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}) = \prod_{i=1}^{m} (2\pi)^{-nk_i/2}(|\hat{\boldsymbol{\Sigma}}_{ii}|)^{-n/2} \exp\{-nk_i/2\} =$$

$$= (2\pi)^{-nk/2}(|\hat{\boldsymbol{\Sigma}}_{11}| \cdots |\hat{\boldsymbol{\Sigma}}_{mm}|)^{-n/2} \exp\{-nk/2\}.$$

Padaliję gauname tikėtinumų santykio statistiką

$$\Lambda = \left(\frac{|\hat{\boldsymbol{\Sigma}}|}{|\hat{\boldsymbol{\Sigma}}_{11}| \cdots |\hat{\boldsymbol{\Sigma}}_{mm}|}\right)^{n/2} = \left(\frac{|\boldsymbol{S}|}{|\boldsymbol{S}_{11}| \cdots |\boldsymbol{S}_{mm}|}\right)^{n/2}. \tag{4.5.4}$$

▲

Tikėtinumų santykio reikšmingumo lygmens α kriterijus atmeta hipotezę H, kai

$$\Lambda < \Lambda_{1-\alpha}, \tag{4.5.5}$$

čia $\Lambda_{1-\alpha}$ yra statistikos Λ lygmens $1-\alpha$ kritinė reikšmė. Norint rasti šią kritinę reikšmę, reikia ištirti statistikos Λ savybes.

4.5.3. Tikėtinumų santykio statistikos momentai

Ieškosime statistikos $U = \Lambda^{2/n}$ momento $\mathbf{E}(U^h)$, kai hipotezė H teisinga.

4.5.2 teorema. *Jeigu hipotezė H teisinga, $|\mathbf{\Sigma}| > 0$, $n > k$, tai statistikos U eilės h momentas turi tokią išraišką:*

$$\mathbf{E}(U^h) = \prod_{i=2}^{m} \left\{ \prod_{j=1}^{k_i} \left[\frac{\Gamma(\frac{n-k_i^*-j}{2}+h)\Gamma(\frac{n-j}{2})}{\Gamma(\frac{n-k_i^*-j}{2})\Gamma(\frac{n-j}{2}+h)} \right] \right\}, \tag{4.5.6}$$

čia $k_i^* = k_1 + ... + k_{i-1}$, $i = 2, ..., m$.

Įrodymas. Remiantis antra Višarto skirstinio savybe matrica $\boldsymbol{S} \sim W_k(n-1, \mathbf{\Sigma})$. Kai hipotezė H teisinga, tai matricos $\boldsymbol{S}_{11}, ..., \boldsymbol{S}_{mm}$ nepriklausomos ir turi Višarto skirstinius $\boldsymbol{S}_{ii} \sim W_{k_i}(n-1, \mathbf{\Sigma}_{ii})$, $i = 1, ..., m$. Naudodami Višarto skirstinio tankį (1.6.13), gauname

$$\mathbf{E}(U^h) = \int \left(\frac{|\boldsymbol{S}|}{|\boldsymbol{S}_{11}| \cdots |\boldsymbol{S}_{mm}|} \right)^h \frac{|\boldsymbol{S}|^{(n-k-2)/2}}{K(k, n-1, \mathbf{\Sigma})} e^{-\frac{1}{2}Tr(\boldsymbol{S}\mathbf{\Sigma}^{-1})} d\boldsymbol{S}$$

$$= \frac{K(k, n-1+2h, \mathbf{\Sigma})}{K(k, n-1, \mathbf{\Sigma})} \int (|\boldsymbol{S}_{11}| \cdots |\boldsymbol{S}_{mm}|)^{-h} \frac{|\boldsymbol{S}|^{(n-k-2+2h)/2}}{K(k, n+2h-1, \mathbf{\Sigma})} e^{-\frac{1}{2}Tr(\boldsymbol{S}\mathbf{\Sigma}^{-1})} d\boldsymbol{S}.$$

Integralas reiškia vidurkį $\mathbf{E}(|\boldsymbol{S}_{11}| \cdots |\boldsymbol{S}_{mm}|)^{-h}$, kai vidurkinama pagal Višarto skirstinio $W_k(n+2h-1, \mathbf{\Sigma})$ tankį. Kadangi funkcijos, kurios vidurkį skaičiuojame, išraiškoje nėra matricų $\boldsymbol{S}_{ij}, i \neq j$ elementų, tai pradžioje galima suintegruoti pagal visus matricų $\boldsymbol{S}_{ij}, i \neq j$ elementus. Tada gausime matricų $\boldsymbol{S}_{11}, ..., \boldsymbol{S}_{mm}$ elementų tankį, kuris yra Višarto skirstinių $W_{k_i}(n+2h-1, \mathbf{\Sigma}_{ii})$ tankių sandauga. Taigi integralas pavirsta integralų sandauga

$$\mathbf{E}(U^h) = \frac{K(k, n-1+2h, \mathbf{\Sigma})}{K(k, n-1, \mathbf{\Sigma})} \prod_{i=1}^{m} \int |\boldsymbol{S}_{ii}|^{-h} \frac{|\boldsymbol{S}_{ii}|^{(n-k_i-2+2h)/2}}{K(k_i, n+2h-1, \mathbf{\Sigma}_{ii})} \times$$

$$\times e^{-\frac{1}{2}Tr(\boldsymbol{S}_{ii}\mathbf{\Sigma}_{ii}^{-1})} d\boldsymbol{S}_{ii} = \frac{K(k, n-1+2h, \mathbf{\Sigma})}{K(k, n-1, \mathbf{\Sigma})} \prod_{i=1}^{m} \frac{K(k_i, n-1, \mathbf{\Sigma}_{ii})}{K(k, n+2h-1, \mathbf{\Sigma}_{ii})}, \tag{4.5.7}$$

nes likusieji integralai lygūs 1, kaip integralai nuo tankio.

Įrašę į (4.5.7) normuojančių daugiklių išraiškas iš (1.6.13), gauname

$$\mathbf{E}(U^h) = \prod_{r=1}^{k} \left\{ \frac{\Gamma(\frac{n-r}{2}+h)}{\Gamma(\frac{n-r}{2})} \right\} \prod_{i=1}^{m} \left\{ \prod_{j=1}^{k_i} \frac{\Gamma(\frac{n-j}{2})}{\Gamma(\frac{n-j}{2}+h)} \right\}.$$

Šį reiškinį galima supaprastinti. Pirmos ir paskutinės sandaugos daugikliai, kai r kinta nuo 1 iki k_1 ir j kinta nuo 1 iki k_1, susiprastina. Todėl

$$\mathbf{E}(U^h) = \prod_{r=k_1+1}^{k} \frac{\Gamma(\frac{n-r}{2}+h)}{\Gamma(\frac{n-r}{2})} \prod_{i=2}^{m} \left\{ \prod_{j=1}^{k_i} \frac{\Gamma(\frac{n-j}{2}+h)}{\Gamma(\frac{n-j}{2}+h)} \right\}$$

$$= \prod_{i=2}^{m} \left\{ \prod_{j=1}^{k_i} \left[\frac{\Gamma(\frac{n-k_i^*-j}{2}+h)\Gamma(\frac{n-j}{2})}{\Gamma(\frac{n-k_i^*-j}{2})\Gamma(\frac{n-j}{2}+h)} \right] \right\},$$

jeigu pažymėsime $k_i^* = k_1 + ... + k_{i-1}, i = 2, ..., m.$▲

4.5.4. Tikėtinumų santykio statistikos skirstiniai

Momentus (4.5.6) galima traktuoti kaip a. d. $\ln U$ momentus generuojančią funkciją

$$\psi(h) = \mathbf{E}(U^h) = \mathbf{E}(e^{h \ln U}),$$

kuri, esant išpildytai sąlygai $n > k$, egzistuoja h kitimo intervale $|h| < 1/2$. Todėl momentai (4.5.6) visiškai nusako a. d. $\ln U$, o kartu ir a. d. U tikimybinį skirstinį.

4.5.3 teorema. *Teoremos 4.5.2 sąlygomis a. d. U skirstinys sutampa su nepriklausomų a. d. sandaugos*

$$U \overset{d}{\sim} \prod_{i=2}^{m} \left(\prod_{j=1}^{k_i} \xi_{ij} \right) \tag{4.5.8}$$

skirstiniu; čia $\xi_{ij}, j = 1, ..., k_i, i = 2, ..., m$, yra nepriklausomi a. d., turintys beta skirstinius

$$\xi_{ij} \sim Be\left(\frac{n - k_i^* - j}{2}, \frac{k_i^*}{2} \right). \tag{4.5.9}$$

Jeigu $k_i = 2r_i$ yra lyginis, tai daugiklių skaičių galima sumažinti:

$$\prod_{j=1}^{2r_i} \xi_{ij} \overset{d}{\sim} \prod_{j=1}^{r_i} \eta_{ij}^2 \tag{4.5.10}$$

čia $\eta_{ij}, j = 1, ..., r_i$, yra nepriklausomi a. d., turintys beta skirstinius

$$\eta_{ij} \sim Be(n - k_i^* - 2j, \;\; k_i^*). \tag{4.5.11}$$

Įrodymas. Formulės (4.5.6) laužtiniuose skliaustuose esantis daugiklis yra a. d. $\xi_{ij} \sim Be((n - k_i^* - j)/2, k_i^*/2)$ momentas $\mathbf{E}(\xi_{ij}^h)$. Taigi

$$\mathbf{E}(U^h) = \prod_{i=2}^{m} \prod_{j=1}^{k_i} \mathbf{E}(\xi_{ij}^h) = \mathbf{E}(\prod_{i=2}^{m} \prod_{j=1}^{k_i} \xi_{ij})^h.$$

Kadangi momentai $\mathbf{E}(U^h), \mathbf{E}(\xi_{ij}^h)$ visiškai nusako skirstinius, tai gauname (4.5.8).

Norėdami įrodyti (4.5.9), analogiškai teoremai 3.5.4 pasinaudojame gama funkcijos nuo dvigubo argumento išraiška (3.5.8). ▲

4.5.5. Tikėtinumų santykio statistikos tam tikri atvejai

1. *Atvejis* $m = 2$. Tada (4.5.8) dešinėje pusėje yra vienguba sandauga

$$U \overset{d}{\sim} \prod_{j=1}^{k_2} \xi_{2j}, \quad \xi_{2j} \sim Be(\frac{n-k_1-j}{2}, \frac{k_1}{2}), \quad j = 1, .., k_2. \tag{4.5.12}$$

a) Tarkime, $k_1 = k_2 = 1$. Tikrinama a. d. X_1 ir X_2 nepriklausomumo hipotezė pagal dvimačio normaliojo a. v. $(X_1, X_2)^T \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ paprastąją didumo n imtį. Statistika

$$U = \frac{|\boldsymbol{S}|}{S_{11}S_{22}} = \frac{S_{11}S_{22} - r^2 S_{11}S_{22}}{S_{11}S_{22}} = 1 - r^2,$$

čia r yra empirinis koreliacijos koeficientas. Pagal (4.5.12) statistika $U \overset{d}{\sim} \xi_{21} \sim Be((n-2)/2, 1/2)$. Remdamiesi beta ir Fišerio skirstinių sąryšiu, gauname

$$\frac{(n-2)(1-U)}{U} = (n-2)\frac{r^2}{1-r^2} \overset{d}{\sim} \frac{(n-2)(1-\xi_{21})}{\xi_{21}} \sim F(1, n-2).$$

Hipotezė atmetama, kai $(n-2)r^2/(1-r^2) > F_\alpha(1, n-2)$. Tai sutampa su kriterijumi (4.2.1), kai alternatyva yra dvipusė.

b) Tarkime, $k_1 = 1, k_2 = k - 1$. Tikrinama hipotezė, kad pirmoji a. v. $\mathbf{X}$ koordinatė X_1 nepriklauso nuo a. v. $(X_2, ..., X_k)^T$. Statistika (žr. (4.1.4))

$$U = \frac{|\boldsymbol{S}|}{S_{11}|\boldsymbol{S}_{22}|} = 1 - \hat{R}^2,$$

čia $\hat{R}^2$ yra dauginio koreliacijos koeficiento kvadrato R^2 empirinis analogas.

Pagal (4.5.12)

$$U \overset{d}{\sim} \prod_{j=1}^{k-1} \xi_{2j}, \quad \xi_{2j} \sim Be\left(\frac{n-1-j}{2}, \frac{1}{2}\right). \tag{4.5.13}$$

Pirmąją vektoriaus $\mathbf{X}$ koordinatę galime perkelti į vektoriaus paskutinę poziciją. Tada $k_1 = k - 1, k_2 = 1$ ir sandaugoje (4.5.12) yra tik vienas daugiklis

$$U \overset{d}{\sim} \zeta \quad \zeta \sim Be\left(\frac{n-k}{2}, \frac{k-1}{2}\right). \tag{4.5.14}$$

Taigi tarp beta skirstinių turėtų galioti toks sąryšis:

$$\zeta = \prod_{j-1}^{k-1} \xi_{2j} \sim Be\left(\frac{n-k}{2}, \frac{k-1}{2}\right), \tag{4.5.15}$$

kai $\xi_{21}, ..., \xi_{2,k-1}$ yra nepriklausomi a. d., turintys beta skirstinius $\xi_{2j} \sim Be((n-1-j)/2, 1/2), j-1, ..., k-1$.

Remdamiesi beta ir Fišerio skirstinių sąryšiu, gausime

$$\frac{(1-U)(n-k)}{(k-1)U} = \frac{\hat{R}^2(n-k)}{(k-1)(1-\hat{R}^2)} \overset{d}{\sim} \frac{(1-\zeta)(n-k)}{(k-1)\zeta} \sim F(k-1, n-k).$$

Todėl tikėtinumų santykio kriterijus sutampa su kriterijumi (4.4.6), kuris buvo sudarytas remiantis empiriniu dauginiu koreliacijos koeficientu.

c) Tegu $\min(k_1, k_2) = 2$. Perstačius koordinates galima pasiekti, kad $k_2 = 2$. Tikrinama hipotezė, kad bet kokios dimensijos k_1 normalusis a. v. nepriklauso nuo dvimačio normaliojo vektoriaus. Remiantis 4.5.2 teorema galima tvirtinti, kad

$$U \overset{d}{\sim} \eta_{21}^2, \quad \eta_{21} \sim Be(n-k_1-2,\ k_1).$$

Perėję prie Fišerio skirstinio gauname

$$F = \frac{(1-\sqrt{U})(n-k_1-2)}{k_1\sqrt{U}} \overset{d}{\sim} \frac{(1-\eta_{21})(n-k_1-2)}{k_1\eta_{21}} \sim F(2k_1, 2(n-k_1-2)).$$

Nepriklausomumo hipotezė atmetama reikšmingumo lygmens α kriterijumi, kai

$$F > F_\alpha(2k_1, 2(n-k_1-2)). \tag{4.5.16}$$

d) Bendresniu atveju tegu $k_1 > 2,\ k_2 > 2$. Pertvarkius koordinates galima pasiekti, kad $k_2 \le k_1$. Tada (4.5.12) yra vienguba nepriklausomų a. d., turinčių beta skirstinius, sandauga. Tokio pat tipo sandaugas gavome 3.4 skyrelyje tikrindami bendras tiesines hipotezes (3.4.1).

Palyginę (4.5.12) ir (3.5.2) matome, kad sandaugos sutaps, jeigu atliksime tokius pakeitimus: vietoje parametrų $k,\ m,\ r$ formulėje (3.5.2) įrašysime parametrus $k_2,\ k_1{+}1,\ k_1$. Taigi dviejų dimensijos k_1 ir k_2 normaliųjų vektorių nepriklausomumo hipotezės tikrinimo kriterijus yra ekvivalentus tiesinių modelių hipotezės (3.4.1) tikrinimo kriterijui, kai tiesiniame modelyje vektorių dimensija yra k_2, parametrų $\boldsymbol{\beta}_i$ dimensija yra $k_1 + 1$, o matricos $\boldsymbol{H}$ rangas lygus k_1.

Analogiškai, palyginę (4.5.12) su (3.5.5) matome, kad sandaugos sutaps, jeigu atliksime tokius pakeitimus: vietoje parametrų $k,\ m,\ r$ formulėje (3.5.2) įrašysime parametrus $k_1,\ k_2{+}1,\ k_2$. Taigi dviejų dimensijos k_1 ir k_2 normaliųjų vektorių nepriklausomumo hipotezės tikrinimo kriterijus yra ekvivalentus tiesinių modelių hipotezės (3.4.1) tikrinimo kriterijui, kai tiesiniame modelyje vektorių dimensija yra k_1, parametrų $\boldsymbol{\beta}_i$ dimensija yra $k_2{+}1$, o matricos $\boldsymbol{H}$ rangas lygus k_2.

Atveju $m = 2$ tiriant statistikos U savybes pritaikomi visi rezultatai, kurie buvo gauti 3.5 skyrelyje, jeigu atliksime minėtus pakeitimus.

4.5.1 pavyzdys. (**1.2.1** pavyzdžio tęsinys). Buvo pamatuota tų pačių n = 30 kineskopų dviejų spindulių srovės stiprumai X_3 ir X_4 kitos technologinės operacijos (I testerių karuselė) metu. Gauti matavimų rezultatai pateikti **4.5.1** lentelėje.

4.5.1 lentelė. Statistiniai duomenys

i	X_{4i}	X_{5i}	i	X_{4i}	X_{5i}	i	X_{4i}	X_{5i}
1	6,7	6,7	11	6,6	6,5	21	6,7	6,5
2	6,8	6,6	12	6,7	6,7	22	6,8	6,8
3	6,6	6,3	13	6,7	6,7	23	6,7	6,5
4	6,8	6,8	14	6,8	6,8	24	6,8	6,7
5	6,6	6,6	15	6,5	6,5	25	6,6	6,5
6	6,8	6,7	16	6,6	6,5	26	6,7	6,7
7	6,6	6,5	17	6,6	6,5	27	6,7	6,6
8	6,6	6,5	18	6,7	6,7	28	6,5	6,6
9	6,7	6,6	19	6,6	6,5	29	6,7	6,6
10	6,7	6,6	20	6,7	6,7	30	6,7	6,6

Sujungę **1.2.1** ir **4.5.1** lentelių duomenis juos galėsime interpretuoti kaip penkiamačio a. v. $\mathbf{X} = (X_1, X_2, X_3, X_4, X_5)^T$ paprastosios didumo $n = 30$ imties realizaciją. Tardami, kad a. v. $\mathbf{X} \sim N_5(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ yra normalusis, patikrinsime hipotezę, kad a. v. $\mathbf{X}_1 = (X_1, X_2, X_3)^T$ ir $\mathbf{X}_2 = (X_4, X_5)^T$ yra nepriklausomi.

Randame parametrų $\boldsymbol{\mu}$ ir $\boldsymbol{\Sigma}$ DT įverčius

$$\hat{\boldsymbol{\mu}}^T = \bar{\mathbf{X}}^T = \begin{pmatrix} 6,2800 & 6,2333 & 6,3000 & 6,6767 & 6,6033 \end{pmatrix},$$

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{30}\boldsymbol{S} = \frac{1}{30}\begin{pmatrix} 0,3480 & 0,2600 & 0,1600 & 0,0060 & -0,0080 \\ 0,2600 & 0,3667 & 0,0700 & 0,0633 & 0,0967 \\ 0,1600 & 0,0700 & 0,2400 & -0,0900 & -0,0700 \\ 0,0060 & 0,0633 & -0,0900 & 0,2137 & 0,2023 \\ -0,0080 & 0,0967 & -0,0700 & 0,2023 & 0,3897 \end{pmatrix}.$$

Apskaičiavę determinantus gauname statistikos U realizaciją

$$U = \frac{|\boldsymbol{S}|}{|\boldsymbol{S}_{11}||\boldsymbol{S}_{22}|} = \frac{0,000251}{0,009132 \cdot 0,042320} = 0,64999.$$

Kadangi $k_2 = 2$, tai remiantis 4.5.3 teoremos (4.5.10) formule galima tvirtinti, kad

$$U \overset{d}{\sim} \eta_{21}^2, \quad \eta_{21} \sim Be(n - k_1 - 2,\ k_1), \quad k_1 = 3.$$

Pereidami prie Fišerio skirstinio gauname

$$F = \frac{(1 - \sqrt{U})25}{3\sqrt{U}} \overset{d}{\sim} \frac{(1 - \eta_{21})25}{3\eta_{21}} \sim F(6,\ 50).$$

Randame statistikos F realizaciją $F = 2,00296$ ir P reikšmę

$$pv = \mathbf{P}\{F_{6,50} > 2,00296\} = 0,08275.$$

Hipotezė atmetama, kai kriterijaus reikšmingumo lygmuo viršija 0,08275.

2. *Atvejis* $m = 3$. Tikrinama trijų normaliųjų vektorių, kurių dimensijos k_1, k_2 ir k_3, nepriklausomumo hipotezė. Nemažinant bendrumo galime tarti, kad $k_2 \leq k_1, k_3 \leq k_1$. Tada

$$U \overset{d}{\sim} \prod_{j=1}^{k_2} \xi_{2j} \prod_{l=1}^{k_3} \xi_{3l}, \quad \xi_{2j} \sim Be\left(\frac{n - k_i - j}{2}, \frac{k_1}{2}\right), \quad j = 1, ..., k_2,$$

$$\xi_{3l} \sim Be\left(\frac{n - k_1 - k_2 - l}{2}, \frac{k_1 + k_2}{2}\right), \quad l = 1, ..., k_3. \tag{4.5.17}$$

Remiantis sąryšiu (4.5.10) daugiklių skaičių formulėje galima sumažinti. Jeigu k_i lyginis, tai daugiklių skaičius sumažėja du kartus. Jeigu k_i nelyginis, tai daugiklių skaičių galima padaryti lygų $(k_i+1)/2$ imant $(k_i-1)/2$ pagal (4.5.10), o likusį paskutinį daugiklį paliekant pagal (4.5.8).

a) Tegu $k_1 = k_2 = k_3 = 1$. Tada

$$U \overset{d}{\sim} \xi_{21}\xi_{31}, \quad \xi_{21} \sim Be((n-2)/2,\ 1/2), \quad \xi_{31} \sim Be((n-3)/2,\ 1)$$

Integruodami šių nepriklausomų beta skirstinių tankių sandaugą srityje $\{(x,y) : xy < u,\ 0 < x, y < 1\}$ gausime P reikšmę

$$pv = \mathbf{P}\{\xi_{21} < u\} + \frac{2\Gamma((n-1)/2)}{\sqrt{\pi}\Gamma((n-2/2))} u^{\frac{n-3}{2}} \arcsin\sqrt{1-u},$$

čia u yra statistikos U realizacija.

b) Tegu $k_1 = k_2 = k_3 = 2$. Tada remiantis (4.5.10)

$$U \overset{d}{\sim} \eta_{21}^2\eta_{31}^2, \quad \eta_{21} \sim Be(n-4,\ 2), \quad \eta_{31} \sim Be(n-6,\ 4)$$

Integruodami šių nepriklausomų beta skirstinių tankių sandaugą srityje $\{(x,y) : xy < \sqrt{u},\ 0 < x, y < 1\}$ gausime P reikšmę

$$pv = \mathbf{P}\{\eta_{31} < u\} + \frac{\Gamma(n-2)}{\Gamma(n-6)\Gamma(4)} u^{\frac{n-5}{2}} \{\frac{n-1}{6} - \frac{3(n-2)\sqrt{u}}{2} -,$$

$$\frac{3(n-5)u}{2} + \frac{17n-45}{6}u^{3/2} - \frac{3(n-3)}{2}u\ln u - \frac{n-4}{2}u^{3/2}\ln u\},$$

čia u yra statistikos U realizacija.

3. *Atvejis* $m = k \geq 3$. Šiuo atveju $k_1 = k_2 = ... = k_m = 1$. Tikrinama hipotezė, kad a. v. $\mathbf{X} = (X_1, ..., X_k)^T$ visos koordinatės yra nepriklausomos, t. y. kovariacinė matrica turi diagonalinį pavidalą. Remiantis (4.5.8) statistika

$$U \overset{d}{\sim} \prod_{i=2}^{k} \xi_{i1}, \quad \xi_{i1} \sim Be(\frac{n-i}{2}, \frac{i}{2}).$$

Jeigu $k \leq 4$, tai U galima užrašyti ne daugiau kaip dviejų daugiklių sandauga. Tada pasiskirstymo funkcijos $G(u) = \mathbf{P}\{U \leq u\}$ reikšmes galima rasti skaitinio integravimo metodais.

4. Kai m ir $k_i, i = 1, ..., m$, yra didesni, tai ieškant $G(u)$ tektų integruoti daugialypį integralą. Nors pointegralinė funkcija palyginti paprasta, tačiau esant dideliam kartotinumui integralo reikšmių radimas gali būti sunkiai realizuojamas. Tada, matyt, realiau yra naudoti skaitinio modeliavimo metodą (žr. 3.5.5.5 skyrelį).

4.5.6. Asimptotinis tikėtinumų santykio kriterijus

Remiantis tikėtinumų santykio asimptotinėmis savybėmis (žr. 1 dalį,4.5.4 skyrelį, 4.5.6 teoremą), galima tvirtinti, kad a. d.

$$Z = -2\ln\Lambda = -n\ln U \xrightarrow{d} \chi^2_\nu, \quad n \to \infty. \qquad (4.5.18)$$

Laisvės laipsnių skaičius

$$\nu = \frac{k(k+1)}{2} - \sum_{i=1}^{m} \frac{k_i(k_i+1)}{2} = \frac{1}{2}(k^2 - \sum_{i=1}^{m} k_i^2),$$

nes pradinėje matricoje $\mathbf{\Sigma}$ yra $k(k+1)/2$ nežinomų parametrų, o jei hipotezė teisinga, tai matricoje $\mathbf{\Sigma}_0$ lieka $\sum_{i=1}^m (k_i+1)k_i/2$ nežinomų parametrų.

Apytikslis α lygmens tikėtinumų santykio kriterijus atmeta hipotezę, kai

$$Z = -2\ln\Lambda = -n\ln U > \chi^2_\alpha(\nu). \qquad (4.5.19)$$

Kriterijų galima užrašyti ir asimptotinės P reikšmės terminais: hipotezė atmetama apytiksliu α lygmens kriterijumi, kai

$$pv_a = \mathbf{P}\{\chi^2_\nu > z\} < \alpha, \qquad (4.5.20)$$

čia z yra statistikos Z realizacija.

Asimptotinio sąryšio (4.5.18) teisingumu galima įsitikinti ir tiesiogiai. Atsitiktinio dydžio $Z = -2\ln\Lambda$ charakteristinė funkcija

$$\psi(t) = \mathbf{E}(e^{it(-2\ln\Lambda)}) = \mathbf{E}(e^{-itn\ln U}) = \mathbf{E}(U^{-itn})$$

gaunama momento išraiškoje (4.5.6) vietoje h įrašius $-itn$:

$$\psi(t) = \mathbf{E}(U^h) = \prod_{i=2}^{m} \left\{ \prod_{j=1}^{k_i} \left[\frac{\Gamma(\frac{n-k_i^*-j}{2} - itn)\Gamma(\frac{n-j}{2})}{\Gamma(\frac{n-k_i^*-j}{2})\Gamma(\frac{n-j}{2} - itn)} \right] \right\}, \qquad (4.5.21)$$

$$k_i^* = k_1 + ... + k_{i-1}, \quad i = 2, ..., m.$$

Skleidžiant gama funkcijas pagal Stirlingo formulę galima įsitikinti, kad $n \to \infty$ charakteristinė funkcija baigtiniuose t kitimo intervaluose konverguoja į $(1 - 2it)^{-\nu/2}$.

4.5.2 pavyzdys. ((**4.5.1** pavyzdžio tęsinys). Išspręsime tą patį uždavinį, kaip ir **4.5.1** pavyzdyje, taikydami asimptotinį tikėtinumų santykio kriterijų.

Randame statistikos $Z = -n\ln U$ realizaciją $Z = 12,13304$ ir asimptotinę P reikšmę

$$pv_a = \mathbf{P}\{\chi^2_6 > 12,13304\} = 0,05907.$$

Hipotezė atmetama, kai kriterijaus reikšmingumo lygmuo viršija 0,05907. Atsakymas gerokai skiriasi nuo atsakymo, gauto **4.5.1** pavyzdyje. Taigi prie palyginti nedidelio imties didumo asimptotinis tikėtinumų santykio kriterijus gali būti netikslus ir privesti prie klaidingos išvados.

4.5.7. Asimptotinio skirstinio patikslinimai

Analogiškai kaip ir 3.5.5 skyrelyje, turint charakteristinę funkciją (4.5.21), tikėtinumų santykio kriterijų galima patikslinti skleidžiant $\ln\psi(t)$ eilute it laipsniais:

$$\ln\psi(t) = \sum_j \frac{\kappa_j}{j!}(it)^j.$$

Iš čia gauname semiinvariantų išraiškas

$$\kappa_s = \sum_{i=2}^{m}\left\{\sum_{j=1}^{k_i}(-n)^s\left[\varphi_s(\frac{n-k_i^*-j}{2}) - \varphi_s(\frac{n-j}{2})\right]\right\}, \quad s = 1, 2, ..., \quad (4.5.22)$$

čia $\varphi_s(x)$ yra funkcijos $\ln\Gamma(x)$ s-oji išvestinė.

Vietoje statistikos $Z = -n\ln U$ imant statistiką $Z^* = \delta Z$, kai $\delta = \nu/\kappa_1$, jos vidurkis $\mathbf{E}Z^* = \nu$ sutampa su ribinio skirstinio vidurkiu. Aproksimuodami a. d. Z^* skirstinį χ^2 skirstiniu su ν laisvės laipsnių, gauname patikslintą kriterijų: hipotezė atmetama asimptotiniu α lygmens kriterijumi, kai teisinga nelygybė

$$Z^* = -\delta n\ln U > \chi^2_\alpha(\nu),$$

arba P reikšmių terminais, kai

$$pv_a^{(1)} = \mathbf{P}\{\chi^2_\nu > \delta z\} < \alpha,$$

čia z yra statistikos Z realizacija.

Analogiškai galima surasti patikslintus kriterijus sutapatinant 2 ar 3 momentus (žr. 3.5.5.2, 3.5.5.3 skyrelius).

Kaip ir skyrelyje 3.5.5.4, aproksimacijos (4.5.20) tikslumą galima padidinti skleidžiant funkciją $\ln\varphi(\delta t)$ ne it, o $(1-2it)$ laipsniais:

$$\ln\varphi(\delta t) = -\frac{\nu}{2}\ln(1-2it) + \sum_{s=1}^{l}\frac{\omega_s}{(n\delta)^s}\left[\frac{1}{(1-2it)^s} - 1\right] + O(1/n^{l+1}). \quad (4.5.23)$$

Tikslinga kaip ir skyrelyje 3.5.5.4 parinkti δ taip, kad ω_1 skleidinyje (4.5.23) virstų 0. Gauname, kad δ reikia parinkti taip:

$$\delta = 1 - \frac{\beta_3 + 9\nu}{3n\nu}, \quad \beta_3 = k^3 - \sum_{i=1}^{m} k_i^3. \quad (4.5.24)$$

Apsiriboję nariu $s = 1$ (dalinis vidurkio sutapatinimas) gausime patikslintą kriterijų P reikšmių terminais: hipotezė atmetama apytiksliu reikšmingumo lygmens α kriterijumi, kai

$$pv_a^{(2)} = \mathbf{P}\{\chi^2_\nu > \delta z\} < \alpha, \quad (4.5.25)$$

čia z yra statistikos Z realizacija.

Imdami ir narį $s = 2$, gauname tokį P reikšmės patikslinimą (analogišką teoremai 3.5.6):

$$pv_a^{(3)} = \mathbf{P}\{\chi^2_\nu > \delta z\} + \frac{\omega_2}{(n\delta)^2}[\mathbf{P}\{\chi^2_{\nu+4} > \delta z\} - \mathbf{P}\{\chi^2_\nu > \delta z\}] + O(1/n^3), \quad (4.5.26)$$

$$\omega_2 = \frac{\beta_4 - 5\nu}{48} - \frac{\beta_3^2}{144\nu}, \quad \beta_4 = k^4 - \sum_{i=1}^m k_i^4.$$

4.5.3 pavyzdys. (**4.5.1** pavyzdžio tęsinys). Patikslinkime asimptotinį tikėtinumų santykio kriterijų remdamiesi pateiktais aproksimacijų patikslinimais ir palyginkime gautus kriterijus su tiksliu kriterijumi, gautu **4.5.1** pavyzdye.

Randame pirmojo semiinvarianto (vidurkio) reikšmę $\kappa_1 = 6,92991$, $\delta z = 10,50493$, ir asimptotinę patikslintą (sutapatinant vidurkius) P reikšmę

$$pv_a^{(1)} = \mathbf{P}\{\chi^2_6 > 10,50493\} = 0,10494.$$

Apsiribodami pirmuoju asimptotinio skleidinio nariu pagal (4.5.24) gauname $\delta = 0,86667$, $\delta z = 10,51530$ ir asimptotinę patikslintą P reikšmę

$$pv_a^{(2)} = \mathbf{P}\{\chi^2_6 > 10,51530\} = 0,10456.$$

Prijungdami antrąjį asimptotinio skleidinio narį pagal (4.5.26) gausime $pv_a^{(3)} = pv_a^{(2)} + 0,00043 = 0,10499$.

Tiksli P reikšmė, rasta **4.5.1** pavyzdyje, yra $0,10499$.

Taigi, gauname tokią išvadą. Esant nedidelėms imtims asimptotinį tikėtinumų santykio kriterijų reikia koreguoti atliekant, pavyzdžiui, pateiktus patikslinimus. Kaip ir 3 skyriuje reikšmingiausią indėlį mažinant paklaidą sudaro tikslus ar apytikslis vidurkio sutapatinimas. Tolesnių pataisų įtaka kur kas mažesnė.

4.6. Pratimai

4.1. Įrodykite, kad koreliacijos koeficiento ρ įvertinys r yra invariantiškas poslinkio ir mastelio atžvilgiu.

4.2. Tegu $\mathbf{X} \sim N_3(\mathbf{0}, \mathbf{\Sigma})$, kai

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix}.$$

Sudarykite parametro ρ pasikliovimo intervalą remdamiesi vienu a. v. $\mathbf{X}$ stebėjimu $(X_1, X_2, X_3)^T$.

4.3. Įrodykite, kad empirinis koreliacijos koeficientas r turi monotoninį tikėtinumo santykį. Įrodykite, kad tarp visų kriterijų, grindžiamų empiriniu koreliacijos koeficientu r, tikrinant hipotezę $H : \rho = \rho_0$, kai alternatyva $\bar{H} : \rho > \rho_0$, kriterijus pavidalo $r > c$ yra tolygiai galingiausias.

4.4. Įrodykite, kad jei normaliojo skirstinio $N_k(\boldsymbol{\mu}, \Sigma)$ kovariacijų matrica $\mathbf{\Sigma}$ yra diagonalioji, $\mathbf{\Sigma} = [\sigma_{ij}]_{k\times k}$, $\sigma_{ii} = \sigma_i^2, \sigma_{ij} = 0, i \neq j = 1, ..., k$, tai empiriniai koreliacijos koeficientai $r_{ij}, i \neq j = 1, ..., k$ nepriklauso nuo diagonalinių matricos $\boldsymbol{S}$ elementų S_{ii}.

4.5. Tegu $\mathbf{X}_1, ..., \mathbf{X}_n$ yra paprastoji imtis a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{I})$. a) Įrodykite, kad koreliacijos koeficientų $r_{ij}, i < j = 1, ..., k$, daugiamačio skirstinio tankio funkcija yra

$$f(\boldsymbol{r}|k, n-1) = \frac{\Gamma^k((n-1)/2)|\boldsymbol{r}|^{(n-k-2)/2}}{\pi^{k(k-1)/4}\prod_{j=1}^k \Gamma((n-j)/2)}, \quad \boldsymbol{r} = [r_{ij}]_{k\times k}.$$

b) Įrodykite, kad determinanto $|\boldsymbol{r}|$ momentas $\mathbf{E}(|\boldsymbol{r}|^h)$ yra

$$\mathbf{E}(|\boldsymbol{r}|^h) = \prod_{j=2}^{k} \left\{ \frac{\Gamma((n-j)/2+h)\Gamma((n-1)/2)}{\Gamma((n-i)/2)\Gamma((n-1)/2+h)} \right\}.$$

c) Įrodykite, kad $|\boldsymbol{r}|$ turi tokį pat skirstinį kaip ir sandauga $\prod_{j=2}^{k} \xi_j$; čia $\xi_2, ..., \xi_k$ yra n. a. d. ir $\xi_j \sim Be((n-1)/2, (j-1)/2), j = 2, ..., k$.

4.6. Koreliacijos koeficiento įvertis pagal dvimačio normaliojo skirstinio paprastąją didumo $n = 10$ imtį $r = 0,65$. Patikrinkite nepriklausomumo hipotezę, kai reikšmingumo lygmuo $\alpha = 0,05$, o alternatyvioji hipotezė tvirtina, kad a. d. yra teigiamai koreliuoti.

4.7. Koreliacijos koeficiento įvertis pagal dvimačio normaliojo skirstinio paprastąją didumo $n = 20$ imtį $r = 0,65$. Patikrinkite hipotezę $H : \rho = 0,4$, kai alternatyva yra $\bar{H} : \rho > 0,4$, o kriterijaus reikšmingumo lygmuo $\alpha = 0,05$.

4.8. Pagal dvimačio normaliojo skirstinio paprastąją didumo $n = 15$ imtį tikrinama nepriklausomumo hipotezė $H : \rho = 0$, kai alternatyvos yra $\bar{H}_1 : \rho > 0$, $\bar{H}_2 : \rho < 0$ ir dvipusė $\bar{H}_3 : \rho \neq 0$, o kriterijaus reikšmingumo lygmuo $\alpha = 0,01$. Sudarykite kriterijų galios funkcijų reikšmių lenteles, kai $\rho = -1(0,2)1$, ir nubraižykite jų grafikus.

4.9. Pagal dvimačio normaliojo skirstinio paprastąją didumo $n = 20$ imtį tikrinama hipotezė $H : \rho = 0,6$, kai alternatyvos yra $\bar{H}_1 : \rho > 0,6$, $\bar{H}_2 : \rho < 0,6$ ir dvipusė $\bar{H}_3 : \rho \neq 0,6$, o kriterijaus reikšmingumo lygmuo $\alpha = 0,05$. Sudarykite kriterijų galios funkcijų reikšmių lenteles, kai $\rho = -1(0,2)1$, ir nubraižykite jų grafikus.

4.10. Koreliacijos koeficiento įvertis pagal dvimačio normaliojo skirstinio paprastąją didumo $n = 50$ imtį $r = -0,7$. Rakite parametro ρ pasikliovimo intervalą, kai pasikliovimo lygmuo $Q = 0,95$.

4.11. Koreliacijos koeficiento įverčiai pagal dvi nepriklausomas dvimačio normaliojo skirstinio paprastąsias didumo $n_1 = 40$ ir $n_2 = 50$ imtis yra 0,5 ir 0,6. Patikrinkite koreliacijos koeficientų lygybės hipotezę, kai alternatyva dvipusė, o kriterijaus reikšmingumo lygmuo $\alpha = 0,01$.

4.12. Tarę, kad buvo stebimi normalieji vektoriai, raskite koreliacijos koeficientų taškinius ir intervalinius įvertinius (pasikliovimo lygmuo $Q = 0,95$) pagal **1.9** pratimo a) ir b) duomenis.

4.13. Tarę, kad buvo stebimas normalusis vektorius pagal **1.8** pratimo duomenis a) a. v. $(X_3, X_4)^T$ sąlyginio skirstinio, kai a. v. $(X_1, X_2)^T$ fiksuotas, parametrų įverčius; b) raskite dalinio koreliacijos koeficiento $\rho_{34.12}$ įvertį $r_{34.12}$; c) raskite parametro $\rho_{34.12}$ pasikliovimo intervalą su pasikliovimo lygmeniu $Q = 0,95$; d) raskite dauginių koreliacijos koeficientų tarp X_3 ir $(X_1, X_2)^T$ ir tarp X_4 ir $(X_1, X_2)^T$ įverčius; e) patikrinkite hipotezes, kad X_3 nepriklauso nuo $(X_1, X_2)^T$ ir X_4 nepriklauso nuo $(X_1, X_2)^T$; f) patikrinkite hipotezę, kad a. v. $(X_1, X_2)^T$ nepriklauso nuo a. v. $(X_3, X_4)^T$.

4.14. Tarę, kad buvo stebimas normalusis vektorius pagal **2.7** pratimo duomenis, patikrinkite nepriklausomumo hipotezę.

4.15. Tarę, kad buvo stebimas normalusis vektorius, pagal **2.9** pratimo duomenis patikrinkite hipotezę, kad a. d. X_1 nepriklauso nuo vektoriaus $(X_2, X_3)^T$.

4.16. Tarę, kad buvo stebimas normalusis vektorius, pagal **3.6** pratimo duomenis a) patikrinkite hipotezę, kad a. v. $(X_1, X_2)^T$ nepriklauso nuo $(Z_1, ..., Z_6)^T$; b) patikrinkite hipotezę, kad a. v. $(X_1, X_2)^T$ nepriklauso nuo $(Z_2, Z_3, Z_4)^T$.

4.17. Pagal $n = 20$ metų duomenis apie vidutinį šieno derlių (X_1), pavasarinių kritulių kiekį (X_2) ir skaičių pavasario dienų, kai temperatūra buvo aukšta (X_3), gautas koreliacijos koeficientų matricos ivertis

$$\begin{pmatrix} 1,00 & 0,80 & -0,40 \\ 0,80 & 1,00 & -0,56 \\ -0,40 & -0,56 & 1,00 \end{pmatrix}.$$

Matome, kad šieno derliaus X_1 ir temperatūros X_3 koreliacija yra neigiama. Paaiškinkite šį faktą remdamiesi dalinio koreliacijos koeficiento $\rho_{13.2}$ įvertiniu.

4.18. (**4.17** tęsinys.) Raskite a. d. X_1 ir a. v. $(X_2, X_3)^T$ dauginio koreliacijos koeficiento įvertį. Tarę, kad buvo stebimas normalusis vektorius, patikrinkite a. d. X_1 ir a. v. $(X_2, X_3)^T$ nepriklausomumo hipotezę.

4.19. Buvo tiriamos darbo sąnaudos lyginant drabužius: X_1 – drabužio uždėjimas ant lyginimo lentos; X_2 – trumpų siūlių lyginimas; X_3 – drabužio perdėjimas ant lentos; X_4 – ilgų siūlių lyginimas; X_5 – ilgų siūlių lyginimo užbaigimas; X_6 – drabužio pakabinimas ant kabyklos. Gauti $n = 76$ darbuotojų vektoriaus $\mathbf{X} = (X_1, ..., X_6)^T$ matavimai, pagal kuriuos apskaičiuota

$$\bar{\mathbf{X}} = \begin{pmatrix} 9,47 \\ 25,56 \\ 13,25 \\ 31,44 \\ 27,29 \\ 8,80 \end{pmatrix}, \quad \boldsymbol{S} = \begin{pmatrix} 2,57 & 0,85 & 1,56 & 1,79 & 1,33 & 0,42 \\ 0,85 & 37,00 & 3,34 & 13,47 & 7,59 & 0,52 \\ 1,56 & 3,34 & 8,44 & 5,77 & 2,00 & 0,50 \\ 1,79 & 13,47 & 5,77 & 34,01 & 10,50 & 1,77 \\ 1,33 & 7,59 & 2,00 & 10,50 & 23,01 & 3,43 \\ 0,42 & 0,52 & 0,50 & 1,77 & 3,43 & 4,59 \end{pmatrix}.$$

Patikrinkite hipotezę, kad a. d. $X_1, ..., X_6$ yra nepriklausomi.

4.20. Vektoriaus $\mathbf{X} = (X_1, ..., X_5)^T$ koordinatės reiškia: X_1 – skaičiavimo greitumas, X_2 – gabumai atliekant skaičiavimus, X_3 – atmintis žodžiams, X_4 – atmintis prasmingiems simboliams, X_5 – atmintis beprasmiams simboliams. Pagal $n = 140$ vektoriaus $\mathbf{X}$ nepriklausomas realizacijas gautas koreliacijų matricos įvertis [2]:

$$\hat{\mathbf{\Sigma}} = \begin{pmatrix} 1,0000 & 0,4248 & 0,0420 & 0,0215 & 0,0573 \\ 0,4248 & 1,0000 & 0,1487 & 0,2489 & 0,2843 \\ 0,0420 & 0,1487 & 1,0000 & 0,6693 & 0,4662 \\ 0,0215 & 0,2489 & 0,6693 & 1,0000 & 0,6915 \\ 0,0573 & 0,2843 & 0,4662 & 0,6915 & 1,0000 \end{pmatrix}.$$

a) Raskite dalinio koreliacijos koeficiento $\rho_{45.3}$ įvertį.
b) Raskite dalinio koreliacijos koeficiento $\rho_{12.345}$ įvertį.
c) Raskite dauginio koreliacijos koeficiento tarp X_1 ir vektoriaus $(X_3, X_4, X_5)^T$ įvertį.
d) Patikrinkite hipotezę, kad a. v. $(X_1, X_2)^T$ ir $(X_3, X_4, X_5)^T$ yra nepriklausomi.

4.21. Tarę, kad buvo stebėtas normalusis a. v., pagal **1.10** pratimo duomenis a) raskite a. v. $\mathbf{X} = (X_1, X_2, X_3, X_4)^T$ koordinačių koreliacijos koeficientų taškinius ir intervalinius $(Q = 0,95)$ įverčius; b) raskite a. v. $\boldsymbol{Y} = (Y_1, Y_2, Y_3)^T$ koordinačių koreliacijos koeficientų taškinius ir intervalinius $(Q = 0,95)$ įverčius; c) patikrinkite a. d. Y_1, Y_2, Y_3 nepriklausomumo hipotezę.

4.22. (**1.13** pratimo tęsinys). Priėmę normalumo prielaidą patikrinkite hipotezę, kad a. v. $\mathbf{X} = (X_1, X_2, X_3, X_4)^T$ koordinatės yra poromis nepriklausomos.

4.23. (**2.10** pratimo tęsinys). Raskite a. d. X_1 ir X_2koreliacijos koeficiento ir a. d. X_1 ir X_2 dalinio koreliacijos koeficiento, kai X_3 fiksuotas, taškinius įverčius.

4.24. (**2.17** pratimo tęsinys). Priėmę normalumo prielaidą **2.17** pratimo sąlygomis a) raskite a. d. X_1, X_2, X_3 empirinius koreliacijos koeficientus; b) raskite a. d. X_1, X_2, X_3 dalinius empirinius koreliacijos koeficientus, kai a. d. X_4, X_5, X_6 fiksuoti, taškinius ir intervalinius $(Q = 0,95)$ įverčius; c) patikrinkite hipotezę, kad a. d. X_2 ir a. v. $(X_1, X_3)^T$ yra nepriklausomi, jeigu a. v. $(X_4, X_5, X_6)^T$ yra fiksuotas.

Atsakymai ir nurodymai

4.2. Parametro ρ pasikliovimo intervalo su pasikliovimo lygmeniu Q rėžiai yra lygties $A\rho^2 - 2B\rho - C = 0$ šaknys; čia $A = X_2^2 + \chi^2_\alpha(3)$, $B = X_2(X_1 + X_3)$, $C = \chi^2_\alpha(3) - X_i^2 - X_2^2 - X_3^2$. **Nurodymas.** Remiantis 3 priedo normaliojo skirstinio 2 savybe $\mathbf{X}^T\mathbf{\Sigma}^{-1}\mathbf{X} \sim \chi^2_3$, $\mathbf{P}\{\mathbf{X}^T\mathbf{\Sigma}^{-1}\mathbf{X} < \chi^2_\alpha(3)\} = Q = 1-\alpha$. Skliausteliuose parašytoji nelygybė ekvivalenti nelygybei $A\rho^2 - 2B\rho - C < 0$. **4.4. Nurodymas.** Naudodamiesi Višarto matricos $\boldsymbol{S}$ elementų $S_{i,j}, j \geq i = 1,...,k$ skirstiniu (1.6.13) gauname vektoriaus $(S_{11},...,S_{kk}, r_{ij}, j > i = 1,...,k)$ skirstinio tankį (pakeitimo jakobianas $|\boldsymbol{J}| = (S_{11}\cdot ... \cdot S_{kk})^{(k-1)/2}$). Įsitikiname, kad gautasis tankis yra sandauga k tankių a.d. $S_{ii} \sim \sigma^2\chi^2_{n-1}$ ir tankio a.v. $(r_{ij}, j > i = 1,...,k)$, kurio išraiška pateikta **4.5.** pratime. **4.5. Nurodymai.** a) žr. **4.4.** pratimą; b) tiesiogiai integruojame naudodami p. a) pateiktą tankio išraišką; c) įsitikiname, kad p.b) gautoje momento išraiškoje apskliaustas daugiklis yra momentas $\mathbf{E}(\xi_j^h)$. **4.6.** Statistikos (4.1.7) realizacija $t = 2,419$. Kadangi $t_{0,05}(8) = 1,8595$, tai nepriklausomumo hipotezė atmetama. **4.7.** Statistikos (4.2.4) realizacija yra 1,4499. Kadangi $z_{0,05} = 1,645$, tai hipotezė neatmetama. **4.8.** Taikant Fišerio aproksimaciją, kai alternatyva $\bar{H}_1$, kriterijaus galia apytiksliai yra

$$\beta_1(\rho) = 1 - \Phi\left(z_{0,01} - \sqrt{n-3}\frac{1}{2}\ln\frac{1+\rho}{1-\rho}\right), \quad \rho > 0.$$

Gauname $\beta_1(0,2) = 0,0522$, $\beta_1(0,4) = 0,1952$, $\beta_1(0,6) = 0,5298$, $\beta_1(0,8) = 0,9305$. **4.9.** Taikant Fišerio aproksimaciją dvipusės alternatyvos $\bar{H}_3$ atveju kriterijaus galia apytiksliai yra

$$\beta_3(\rho) = 1-\Phi\left(z_{\alpha/2} - \sqrt{n-3}\frac{1}{2}\ln\frac{(1+\rho)(1-\rho_0)}{(1-\rho)(1+\rho_0)}\right)+\Phi\left(-z_{\alpha/2} - \sqrt{n-3}\frac{1}{2}\ln\frac{(1+\rho)(1-\rho_0)}{(1-\rho)(1+\rho_0)}\right).$$

Gauname

ρ	-0,8	-0,6	-0,4	-0,2	0,0	0,2	0,4	0,6	0,8
$\beta_3(\rho)$	1	0,99991	0,9959	0,9585	0,8154	0,5248	0,1991	0,05	0,3867

4.10. Naudodami Fišerio aproksimaciją gauname apytikslio pasikliovimo intervalo realizaciją $(\underline{\rho}; \overline{\rho}) = (-0,8188; -0,5237)$. **4.11.** Remiantis Fišerio aproksimacija statistika $U = \sqrt{(n_1-3)(n_2-3)/(n_1+n_2-6)}(th(r_1) - th(r_2))$, kai hipotezė teisinga, turi apytiksliai standartinį normalųjį skirstinį. Gauname statistikos U realizaciją $U = -0,7907$. Kadangi $|U| < z_{0,005} = 2,5758$, tai hipotezė neatmetama. **4.12.** a) $r = 0,5477$; $(\underline{\rho}; \overline{\rho}) = (0,3053; 0,7193)$; b) $r = 0,7937$; $(\underline{\rho}; \overline{\rho}) = (0,7060; 0,8574)$. **4.13.** a) A.v. $\mathbf{X}^{(2)} = (X_3, X_4)^T$ sąlyginis skirstinys, kai a.v. $\mathbf{X}^{(1)} = (X_1, X_2)^T = (x_1, x_2)^T = \boldsymbol{x}^{(1)}$ yra fiksuotas yra dvimatis normalusis su vidurkių vektoriumi $(\mu_3(x_1, x_2), \mu_4(x_1, x_2))^T = \boldsymbol{\mu}^{(2)} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}(\boldsymbol{x}^{(1)} - \boldsymbol{\mu}^{(1)})$ ir kovariacijų matrica $\mathbf{\Sigma}_{22\cdot 1} = \mathbf{\Sigma}_{22} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}$. Naudodami parametrų įverčius, surastus **1.7.** pratime, gauname $\hat{\mu}_3(x_1, x_2) = 23,799 + 0,450x_1 + 0,506x_2$, $\hat{\mu}_4(x_1, x_2) = 41,229 + 0,274x_1 + 0,378x_2$; $\hat{\sigma}_{33\cdot 12} = 43,474$, $\hat{\sigma}_{34\cdot 12} = 18,039$, $\hat{\sigma}_{44\cdot 12} = 19,126$; b) $r_{34\cdot 12} = 0,468$; c) $(\underline{\rho}_{34\cdot 12}; \overline{\rho}_{34\cdot 12}) = (0,0798; 0,9352)$; d) pirmu atveju $\hat{R}^2 = 0,5687$, antruoju – $\hat{R}^2 = 0,5752$; e) statistika F, kuri esant teisingai hipotezei apytiksliai turi Fišerio skirstinį su 2 ir 22 laisvės laipsniais pirmu atveju, įgijo reikšmę 14,506, antruoju – 14,894; atitinkamos asimptotinės P reikšmės yra 0,000096 ir 0,000081; hipotezės atmetamos; f) statistika $F = 21(1 - \sqrt{U}/(2\sqrt{U})$ įgijo reikšmę 6,5760; kadangi P reikšmė $pv = \mathbf{P}\{F_{4;42} > 6,576\} = 0,0003$, tai hipotezė atmestina. **4.14.** Koreliacijos koeficiento įvertis $r = 0,8680$. Statistika $t = \sqrt{n-2}r/\sqrt{1-r^2}$ įgijo reikšmę 4,9430; P reikšmė $pv = 0,0006$. Hipotezė atmestina. **4.15.** Statistikos $\hat{R}^2$ realizacija yra 0,5288. Statistika F, kuri esant teisingai hipotezei turi Fišerio skirstinį su 2 ir 25 laisvės laipsniais, įgijo reikšmę 14,0291; P reikšmė $pv = \mathbf{P}\{F_{2;25} > 14,0291\} = 0,00008$. Hipotezė atmestina. **4.16.** a) statistika $F = 17(1 - \sqrt{U}/(6\sqrt{U})$ įgijo reikšmę 8,7058; kadangi P reikšmė $pv = \mathbf{P}\{F_{12;34} > 8,7058\} = 3\cdot 10^{-7}$, tai hipotezė atmetama; b) statistika $F = 20(1-\sqrt{U}/(3\sqrt{U})$ įgijo reikšmę 5,44503; kadangi P reikšmė $pv = \mathbf{P}\{F_{6;40} > 5,44503\} = 0,00034$, tai hipotezė atmestina. **4.17.** Dalinio koreliacijos koeficiento $\rho_{13\cdot 2}$ įvertis $r_{13\cdot 2} = +0,1$ yra teigiamas. Tai kad koreliacijos koeficiento ρ_{13} įvertis $r_{13} = -0,4$ yra neigiamas, matyt, gali būti paaiškintas aukšta neigiama X_2 ir X_3 priklausomybe (lietingu oru temperatūra turi tendenciją įgyti mažesnes reikšmes). **4.18.** Dauginio koreliacijos koeficiento įvertis $\hat{R} = 0,802$. Statistika F iš (4.1.5) įgijo reikšmę 15,3306. Kadangi $pv = \mathbf{P}\{F_{2;17} > 15,3306\} = 0,00016$, tai nepriklausomumo hipotezė atmetama. **4.19.** Statistika $Z = -n\ln U$ įgijo reikšmę 54,7939. Kadangi $pv_a = \mathbf{P}\{\chi^2_{15} > 54,7939\} = 1,9\cdot 10^{-6}$, tai nepriklausomumo hipotezė atmetama. **4.20.** a) $r_{45\cdot 3} = 0,5773$; b) $r_{12\cdot 345} = 0,4315$; c) $\hat{R} = 0,0726$; d) statistika $F = 135(1-\sqrt{U}/(3\sqrt{U})$ įgijo reikšmę 2,3771; kadangi P reikšmė $pv = \mathbf{P}\{F_{6;270} > 2,3771\} = 0,0296$. Hipotezė atmetama, kai kriterijaus reikšmingumo lygmuo viršija 0,0296. **4.21.** a) $r_{12} = 0,9138$; $(\underline{\rho}_{12};\overline{\rho}_{12}) = (0,8251;0,9585)$; $r_{13} = 0,8859$; $(\underline{\rho}_{13};\overline{\rho}_{13}) = (0,7721;0,9447)$; $r_{14} = 0,8981$; $(\underline{\rho}_{14};\overline{\rho}_{14}) = (0,7951;0,9508)$; $r_{23} = 0,7882$; $(\underline{\rho}_{23};\overline{\rho}_{23}) = (0,5977;0,8945)$; $r_{24} = 0,7881$; $(\underline{\rho}_{24};\overline{\rho}_{24}) = (0,5975;0,8944)$; $r_{34} = 0,9231$; $(\underline{\rho}_{34};\overline{\rho}_{34}) = (0,8433;0,9631)$; b) $r_{12} = 0,7711$; $(\underline{\rho}_{12};\overline{\rho}_{12}) = (0,5688;0,8854)$; $r_{13} = 0,3816$; $(\underline{\rho}_{13};\overline{\rho}_{13}) = (0,0248;0,6522)$; $r_{23} = 0,3448$; $(\underline{\rho}_{23};\overline{\rho}_{23}) = (-0,0176;0,6272)$; c) statistika $Z = -2\ln\Lambda$ įgijo reikšmę 32,0292; asimptotinė P reikšmė $pv_a = \mathbf{P}\{\chi^2_3 > 32,0292\} = 5.2\cdot 10^{-7}$; hipotezė atmetama. **4.22.** Statistika $t_{ij} = \sqrt{n-2}r_{ij}/\sqrt{1-r^2_{ij}}, i \neq j = 1,2,3,$ esant teisingoms tikrinamoms hipotezėms turinti Stjudento skirstinį su 38 laisvės laipsniais įgijo reikšmes -1,3447; 0,2194; -1,6456; 0,4473; 0,6973; -0,4550; atmesti tikrinamas hipotezes nėra pagrindo. **4.23.** $r_{23} = -0,2095$; $r_{23\cdot 1} = 0,0323$. **4.24.** a) $r_{12} = 0,8752$; $r_{13} = 0,9559$; $r_{23} = 0,9013$; b) $r_{12\cdot 456} = -0,0998$; $r_{13\cdot 456} = 0,4920$; $r_{23\cdot 456} = -0,0303$; $(\underline{\rho}_{12\cdot 456};\overline{\rho}_{12\cdot 456}) = (-0,3491;0,1641)$; $(\underline{\rho}_{13\cdot 456};\overline{\rho}_{13\cdot 456}) = (0,2744;0,8030)$; $(\underline{\rho}_{23\cdot 456};\overline{\rho}_{23\cdot 456}) = (-0,2864;0,2298)$; c) statistika (4.4.10) įgijo reikšmę 0,289; P reikšmė $pv = \mathbf{P}\{F_{2;55} > 0,289\} = 0,7501$; hipotezė neatmetama.

5 skyrius

Hipotezės apie kovariacijų matricas

5.1. Kovariacijų matricų lygybės hipotezės

Tikrinant vidurkių vektorių lygybės hipotezę 2 skyriuje ar bendresnes tiesines hipotezes 3 skyriuje, buvo tariama, kad stebimų nepriklausomų vektorių kovariacinės matricos yra lygios. Jeigu yra pagrįstų abejonių dėl šios prielaidos teisingumo, reikėtų turėti kriterijus, kurie leistų patikrinti tokios prielaidos teisingumą.

Tarkime, $\mathbf{X}_{11},...,\mathbf{X}_{1n_1}; \mathbf{X}_{21},...,\mathbf{X}_{2n_2}; \cdots; \mathbf{X}_{m1},...,\mathbf{X}_{mn_m}$, yra didumo $n_1, n_2, ..., n_m$, $n_1+n_2+...+n_m = n$, paprastosios imtys, gautos stebint nepriklausomus a. v. $\mathbf{X}_1 \sim N_k(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$; $\mathbf{X}_2 \sim N_k(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$; $\cdots$; $\mathbf{X}_m \sim N_k(\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m)$. Remiantis šiomis imtimis reikia patikrinti hipotezę

$$H : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = ... = \boldsymbol{\Sigma}_m, \tag{5.1.1}$$

kad šių m nepriklausomų imčių kovariacinės matricos yra lygios.

5.1.1. Tikėtinumų santykio statistika

Tikėtinumų santykio statistika Λ hipotezei H tikrinti yra

$$\Lambda = \frac{\max_{\boldsymbol{\mu}_1,...,\boldsymbol{\mu}_m,\boldsymbol{\Sigma}} L(\boldsymbol{\mu}_1, ..., \boldsymbol{\mu}_m, \boldsymbol{\Sigma})}{\max_{\boldsymbol{\mu}_1,...,\boldsymbol{\mu}_m,\boldsymbol{\Sigma}_1,...,\boldsymbol{\Sigma}_m} L(\boldsymbol{\mu}_1, ..., \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_1, ..., \boldsymbol{\Sigma}_m)} = \frac{L(\tilde{\boldsymbol{\mu}}_1, ..., \tilde{\boldsymbol{\mu}}_m, \tilde{\boldsymbol{\Sigma}})}{L(\hat{\boldsymbol{\mu}}_1, ..., \hat{\boldsymbol{\mu}}_m, \hat{\boldsymbol{\Sigma}}_1, ..., \hat{\boldsymbol{\Sigma}}_m)}, \tag{5.1.2}$$

čia $\hat{\boldsymbol{\mu}}_1, ..., \hat{\boldsymbol{\mu}}_m, \hat{\boldsymbol{\Sigma}}_1, ..., \hat{\boldsymbol{\Sigma}}_m$ yra parametrų DT įvertiniai, o $\tilde{\boldsymbol{\mu}}_1, ..., \tilde{\boldsymbol{\mu}}_m, \tilde{\boldsymbol{\Sigma}}$ – parametrų DT įvertiniai, surasti tarus, kad hipotezė (5.1.1) teisinga.

Remiantis 1.1 skyreliu parametrų DT įvertiniai yra

$$\hat{\boldsymbol{\mu}}_i = \bar{\mathbf{X}}_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\mathbf{X}_{ij}, \quad \hat{\boldsymbol{\Sigma}}_i = \frac{1}{n_i}\boldsymbol{S}_i = \frac{1}{n_i}\sum_{j=1}^{n_i}(\mathbf{X}_{ij}-\bar{\mathbf{X}}_i)(\mathbf{X}_{ij}-\bar{\mathbf{X}}_i)^T, \quad (5.1.3)$$

o tikėtinumo funkcijos maksimumas

$$L(\hat{\boldsymbol{\mu}}_1, ..., \hat{\boldsymbol{\mu}}_m, \hat{\boldsymbol{\Sigma}}_1, ..., \hat{\boldsymbol{\Sigma}}_m) = \prod_{i=1}^{m}\{(2\pi)^{-kn_i/2}(|\hat{\boldsymbol{\Sigma}}_i|)^{-n_i/2}\exp\{-\frac{kn_i}{2}\}\}.$$

Kai hipotezė H teisinga, vidurkių $\boldsymbol{\mu}_1, ..., \boldsymbol{\mu}_m$ įvertiniai yra tokie patys $\hat{\boldsymbol{\mu}}_i = \tilde{\boldsymbol{\mu}}_i, i = 1, ..., m$, o bendros kovariacinės matricos $\boldsymbol{\Sigma}$ įvertinys

$$\tilde{\boldsymbol{\Sigma}} = \frac{1}{n}\boldsymbol{S} = \frac{1}{n}(\boldsymbol{S}_1 + ... + \boldsymbol{S}_m). \quad (5.1.4)$$

Tikėtinumo funkcijos sąlyginis maksimumas

$$L(\hat{\boldsymbol{\mu}}_1, ..., \hat{\boldsymbol{\mu}}_m, \tilde{\boldsymbol{\Sigma}}) = (2\pi)^{-kn/2}(|\tilde{\boldsymbol{\Sigma}}|)^{-n/2}\exp\{-\frac{kn}{2}\}.$$

Taigi, tikėtinumų santykio statistika

$$\Lambda = \frac{|\hat{\boldsymbol{\Sigma}}_1|^{n_1/2}\cdots|\hat{\boldsymbol{\Sigma}}_m|^{n_m/2}}{|\tilde{\boldsymbol{\Sigma}}|^{n/2}} = \frac{n^{kn/2}}{n_1^{kn_1/2}\cdots n_m^{kn_m/2}}\frac{|\boldsymbol{S}_1|^{n_1/2}\cdots|\boldsymbol{S}_m|^{n_m/2}}{|\boldsymbol{S}|^{n/2}}. \quad (5.1.5)$$

Hipotezė H atmetama, kai

$$\Lambda < \Lambda_{1-\alpha}, \quad (5.1.6)$$

čia $\Lambda_{1-\alpha}$ yra statistikos Λ lygmens $1-\alpha$ kritinė reikšmė.

Bartletas pasiūlė modifikuoti statistiką Λ naudojant nepaslinktuosius kovariacinių matricų įvertinius, t. y. (5.1.5) vietoje n_i imti $\nu_i = n_i - 1$, o vietoje $n = n_1 + ... + n_m$ imti $\nu = n - m$. Tada modifikuotoji tikėtinumų santykio statistika yra

$$\tilde{\Lambda} = \frac{\nu^{k\nu/2}}{\nu_1^{k\nu_1/2}\cdots\nu_m^{k\nu_m/2}}\frac{|\boldsymbol{S}_1|^{\nu_1/2}\cdots|\boldsymbol{S}_m|^{\nu_m/2}}{|\boldsymbol{S}|^{\nu/2}}. \quad (5.1.7)$$

Hipotezė H atmetama, kai

$$\tilde{\Lambda} < \tilde{\Lambda}_{1-\alpha} \quad (5.1.8)$$

čia $\tilde{\Lambda}_{1-\alpha}$ yra statistikos $\tilde{\Lambda}$ lygmens $1-\alpha$ kritinė reikšmė.

Kriterijus (5.1.8), grindžiamas modifikuotąja tikėtinumų santykio statistika, yra ekvivalentus kriterijui, grindžiamam statistika (atmetame konstantą)

$$V = \frac{|\boldsymbol{S}_1|^{\nu_1/2}\cdots|\boldsymbol{S}_m|^{\nu_m/2}}{|\boldsymbol{S}|^{\nu/2}}. \quad (5.1.9)$$

Hipotezė atmetama, kai

$$V < V_{1-\alpha}, \tag{5.1.10}$$

čia $V_{1-\alpha}$ yra statistikos V lygmens $1-\alpha$ kritinė reikšmė. Norint rasti kritinę reikšmę $V_{1-\alpha}$, reikia ištirti statistikos V savybes.

5.1.1 pavyzdys. Atveju $k=1, m=2$ nelygybė (5.1.10) turi tokį pavidalą:

$$V = \frac{(s_1^2\nu_1)^{\nu_1/2}(s_2^2\nu_2)^{\nu_2/2}}{(s_1^2\nu_1 + s_2^2\nu_2)^{(\nu_1+\nu_2)/2}} = \frac{\nu_1^{\nu_1/2}\nu_2^{\nu_2/2}F^{\nu_1/2}}{(\nu_1 F + \nu_2)^{(\nu_1+\nu_2)/2}} < V_{1-\alpha}, \tag{5.1.11}$$

čia $F = s_1^2/s_2^2$, o s_1^2 ir s_2^2 yra dispersijų σ_1^2 ir σ_2^2 NMD įvertiniai pagal nepriklausomas didumo n_1 ir n_2 normaliųjų a.d. imtis. Kai $\sigma_1^2 = \sigma_2^2$, tai $F \sim F(\nu_1, \nu_2)$. Nelygybė (5.1.11) ekvivalenti tokioms dviem nelygybėms

$$F < F_{1-\alpha_1}(\nu_1,\nu_2), \quad F > F_{\alpha_2}(\nu_1,\nu_2), \quad \alpha_1 + \alpha_2 = \alpha.$$

Gauname kriterijų, kuris sutampa su TGN kriterijumi hipotezei $H : \sigma_1^2 = \sigma_2^2$ tikrinti, gautu 1 dalies 5.7.2 skyrelyje.

Vienmačiu atveju, kai imčių skaičius $m > 2$, (5.1.8) tipo kriterijus buvo nagrinėtas 1 dalies 5.6.2 skyrelyje, 5.6.3 pavyzdyje.

5.1.2. Tikėtinumų santykio statistikos momentai

Ieškosime statistikos V, apibrėžtos (5.1.9) lygybe, momento $\mathbf{E}(V^h)$.

5.1.1 teorema. *Kai hipotezė (5.1.1) teisinga,* $|\boldsymbol{\Sigma}_j| > 0$, $\nu_j \geq k$, $j = 1, ..., m$, *momentas* $\mathbf{E}(V^h)$ *turi tokį pavidalą*

$$\mathbf{E}(V^h) = \prod_{i=1}^{k} \left\{ \prod_{j=1}^{m} \left[\frac{\Gamma((\nu_j + h\nu_j + 1 - i)/2)}{\Gamma((\nu_j + 1 - i)/2)} \right] \frac{\Gamma((\nu + 1 - i)/2)}{\Gamma((\nu + h\nu + 1 - i)/2)} \right\}. \tag{5.1.12}$$

Jeigu $k = 2r$ *lyginis, tai*

$$\mathbf{E}(V^h) = \prod_{i=1}^{r} \left\{ \prod_{j=1}^{m} \left[\frac{\Gamma(\nu_j + h\nu_j + 1 - 2i)}{\Gamma(\nu_j + 1 - 2i)} \right] \frac{\Gamma(\nu + 1 - 2i)}{\Gamma(\nu + h\nu + 1 - 2i)} \right\}. \tag{5.1.13}$$

Įrodymas. Remiantis 1.2 skyreliu matricų $\boldsymbol{S}_1, ..., \boldsymbol{S}_m$ elementų skirstiniai yra nepriklausomi Višarto skirstiniai $\boldsymbol{S}_i \sim W_k(\nu_i, \boldsymbol{\Sigma}), i = 1, ..., m$. Naudodami tankio išraišką (1.6.13), gauname

$$\mathbf{E}(V^h) = \int |\boldsymbol{S}|^{-h\nu} \prod_{i=1}^{m} \left\{ \frac{|\boldsymbol{S}_i|^{(\nu_i + 2h\nu_i - k - 1)/2} \exp\{-\frac{1}{2} Tr(\boldsymbol{S}_i \boldsymbol{\Sigma}^{-1})\}}{K(k, \nu_i, \boldsymbol{\Sigma})} d\boldsymbol{S}_i \right\}$$

$$= \prod_{i=1}^{m} \left\{ \frac{K(k, \nu_i + h\nu_i, \boldsymbol{\Sigma})}{K(k, \nu_i, \boldsymbol{\Sigma})} \right\} \int |\boldsymbol{S}|^{-h\nu} \times$$

$$\times \prod_{i=1}^{m} \left\{ \frac{|\boldsymbol{S}_i|^{(\nu_i + 2h\nu_i - k - 1)/2} \exp\{-\frac{1}{2} Tr(\boldsymbol{S}_i \boldsymbol{\Sigma}^{-1})\}}{K(k, \nu_i + h\nu_i, \boldsymbol{\Sigma})} d\boldsymbol{S}_i \right\}.$$

Integralas reiškia funkcijos $|\boldsymbol{S}| = |\boldsymbol{S}_1 + ... + \boldsymbol{S}_m|$ momentą $\mathbf{E}(|\boldsymbol{S}|^{-h\nu})$, kai vidurkinama pagal Višarto skirstinių $W_k(\nu_i + h\nu_i, \boldsymbol{\Sigma})$ tankių sandaugą. Remiantis Višarto skirstinių trečia savybe $\boldsymbol{S} = \boldsymbol{S}_1 + ... + \boldsymbol{S}_m \sim W_k(\nu + h\nu, \boldsymbol{\Sigma})$. Todėl momentą $\mathbf{E}(|\boldsymbol{S}|^{-h\nu})$ galime rasti vidurkindami pagal skirstinio $W_k(\nu + h\nu, \boldsymbol{\Sigma})$ tankį. Gauname

$$\mathbf{E}(V^h) = \prod_{i=1}^{m} \left\{ \frac{K(k, \nu_i + h\nu_i, \boldsymbol{\Sigma})}{K(k, \nu_i, \boldsymbol{\Sigma})} \right\} \int |\boldsymbol{S}|^{-h\nu} \times$$

$$\times \frac{|\boldsymbol{S}|^{(\nu+2h\nu-k-1)/2} \exp\{-\frac{1}{2} Tr(\boldsymbol{S}\boldsymbol{\Sigma}^{-1})\}}{K(k, \nu + h\nu, \boldsymbol{\Sigma})} d\boldsymbol{S}$$

$$= \prod_{i=1}^{m} \left\{ \frac{K(k, \nu_i + h\nu_i, \boldsymbol{\Sigma})}{K(k, \nu_i, \boldsymbol{\Sigma})} \right\} \frac{K(k, \nu, \boldsymbol{\Sigma})}{K(k, \nu + h\nu, \boldsymbol{\Sigma})}.$$

Įrašę normuojančių daugiklių išraiškas iš (1.6.13), gauname (5.1.12).

Formulė (5.1.13) gaunama naudojant dvigubo argumento gama funkcijos išraišką (3.5.8) analogiškai kaip teoremoje 3.5.4. ▲

5.1.3. Tikėtinumų santykio statistikos skirstiniai

Remdamiesi (5.1.12) momentų išraiškomis parodysime, kad statistikos V skirstinys sutampa su tam tikro polinomo nuo nepriklausomų a. d., turinčių beta skirstinius, skirstiniu.

5.1.2 teorema. *Kai hipotezė (5.1.1) teisinga, $|\boldsymbol{\Sigma}_j| > 0$, $\nu_j \geq k$, $j = 1, ..., m$, statistikos V skirstinys sutampa su skirstiniu sandaugos:*

$$V \overset{d}{\sim} \prod_{i=1}^{k} \left\{ \prod_{j=1}^{m-1} [\xi_{ij}^{\nu_j^*/2} (1 - \xi_{ij})^{\nu_{j+1}/2}] \xi_i^{\nu/2} \right\}, \tag{5.1.14}$$

čia ξ_{ij}, ξ_i, $i = 1, ..., k$, $j = 1, ..., m-1$, yra nepriklausomi a. d., turintys beta skirstinius:

$$\xi_{ij} \sim Be((\nu_j^* + j(1-i))/2, (\nu_{j+1} + 1 - i)/2),$$
$$\xi_i \sim Be((\nu + m(1-i))/2, (m-1)(i-1)/2),$$
$$\xi_1 = 1, \quad \nu_j^* = \nu_1 + ... + \nu_j.$$

Jeigu $k = 2r$ lyginis, tai

$$V \overset{d}{\sim} \prod_{i=1}^{r} \left\{ \prod_{j=1}^{m-1} [\eta_{ij}^{\nu_j^*} (1 - \eta_{ij})^{\nu_{j+1}}] \eta_i^{\nu} \right\}, \tag{5.1.15}$$

čia $\eta_{ij}, \eta_i, i = 1, ..., r, j = 1, ..., m-1$, yra n. a. d., turintys beta skirstinius

$$\eta_{ij} \sim Be(\nu_j^* + j(1-2i), \nu_{j+1} + 1 - 2i), \quad \eta_i \sim Be(\nu + m(1-2i), (2i-1)(m-1)).$$

Įrodymas. Tegu $Z \sim Be(\gamma, \eta)$. Tada lengva patikrinti, kad momentas

$$\mathbf{E}[Z^a(1-Z)^b]^h = \frac{\Gamma(\gamma+ah)\Gamma(\eta+bh)\Gamma(\gamma+\eta)}{\Gamma(\gamma)\Gamma(\eta)\Gamma(\gamma+\eta+ah+bh)}. \tag{5.1.16}$$

Kai fiksuotas i, imkime du (5.1.12) daugiklius, $j=1$ ir $j=2$, ir papildykime daugikliu, kad gautume (5.1.16) momentą:

$$\frac{\Gamma(\frac{\nu_1+h\nu_1+1-i}{2})\Gamma(\frac{\nu_2+h\nu_2+1-i}{2})\Gamma(\frac{\nu_1+\nu_2+2-2i}{2})}{\Gamma(\frac{\nu_1+1-i}{2})\Gamma(\frac{\nu_2+1-i}{2})\Gamma(\frac{\nu_1+\nu_2+2-2i+h\nu_1+h\nu_2}{2})}$$

$$= \mathbf{E}(\xi_{i1}^{\nu_1/2}(1-\xi_{i1})^{\nu_2/2})^h, \quad \xi_{i1} \sim Be((\nu_1+1-i)/2,\ (\nu_2+1-i)/2).$$

Padalinkime iš daugiklio, kuriuo papildėme sandaugą, prijunkime sekantį daugiklį iš (5.1.12), kai $j=3$, ir papildykime tokiu daugikliu, kad vėl gautume momentą (5.1.16)

$$\frac{\Gamma(\frac{\nu_2^*+h\nu_2^*+2-2i}{2})\Gamma(\frac{\nu_3+h\nu_3+1-i}{2})\Gamma(\frac{\nu_3^*+3-3i}{2})}{\Gamma(\frac{\nu_2^*+2-2i}{2})\Gamma(\frac{\nu_3+1-i}{2})\Gamma(\frac{\nu_3^*+3-3i+h\nu_3^*}{2})}$$

$$= \mathbf{E}(\xi_{i2}^{\nu_2^*/2}(1-\xi_{i2})^{\nu_3/2})^h, \quad \xi_{i2} \sim Be((\nu_2^*+2-2i,\ (\nu_3+1-i)/2)).$$

Tęsdami šią procedūrą po $m-1$ žingsnio gausime visų daugiklių iš (5.1.14), kai j kinta nuo 1 iki $m-1$, eilės h momentus. Prie likusiojo daugiklio prijungę paskutinį daugiklį iš (5.1.12), gausime

$$\frac{\Gamma((\nu+h\nu_1+m-mi)/2)\Gamma((\nu+1-i)/2)}{\Gamma((\nu+m-mi)/2)\Gamma((\nu+h\nu+1-i)/2)},$$

o tai yra momentas $\mathbf{E}(\xi_i^{\nu h})$, kai $\xi_i \sim Be((\nu+m(1-i))/2, (m-1)(i-1)/2)$, $\xi_1 = 1$.

Kadangi momentas

$$\mathbf{E}(V^h) = \mathbf{E}\left(\prod_{i=1}^{k}\left\{\prod_{j=1}^{m-1}[\xi_{ij}^{\nu_j^*/2}(1-\xi_{ij})^{\nu_{j+1}/2}]\xi_i^{\nu/2}\right\}\right)^h$$

$$\prod_{i=1}^{k}\left\{\prod_{j=1}^{m-1}\mathbf{E}[\xi_{ij}^{\nu_j^*/2}(1-\xi_{ij})^{\nu_{j+1}/2}]^h\mathbf{E}(\xi_i^{\nu/2})^h\right\},$$

yra lygus momentų sandaugai, o jie visiškai nusako skirstinius, tai gauname (5.1.14) lygybę, kurios dešinėje pusėje visi a. d., turintys skirtingus indeksus, yra nepriklausomi ir turi beta skirstinius.

Lygybė (5.1.15) įrodoma analogiškai pertvarkant (5.1.13) lygybės dešiniąją pusę. ▲

5.1.2 pavyzdys. Jeigu $k=1$ ir $m=2$, tai iš (5.1.14) gauname

$$V \stackrel{d}{\sim} \xi_{11}^{\nu_1/2}(1-\xi_{11})^{\nu_2/2}, \quad \xi_{11} \sim Be(\nu_1/2, \nu_2/2). \tag{5.1.17}$$

Prisiminę beta ir Fišerio skirstinių sąryšį (jei $F \sim F(\nu_1, \nu_2)$, tai $\nu_1 F/(\nu_1 F + \nu_2) \sim Be(\nu_1/2, \nu_2/2)$) matome, kad dešinioji (5.1.11) pusė lygi (5.1.17).

5.1.1 pastaba. Gautieji tikėtinumų santykio skirstiniai kur kas sudėtingesni už tuos, kuriuos gavome 3 ir 4 skyriuose. Tačiau modeliuojant, kai yra (5.1.13), (5.1.14) sąryšiai, rasti apytiksles kvantilių reikšmes reikiamo tikslumo su šiuolaikiniais kompiuteriais nebus sudėtinga.

5.1.4. Tikėtinumų santykio asimptotinis skirstinys

Jeigu imčių didumai $n_i = c_i n$, $0 < c_i < 1$, c_i fiksuoti, o $n \to \infty$, tai remiantis tikėtinumų santykio asimptotinėmis savybėmis (1 dalis, 4.5.4 skyrelis)

$$-2 \ln \Lambda \xrightarrow{d} \chi^2_f, \tag{5.1.18}$$

čia laisvės laipsnių skaičius

$$f = mk(k+1)/2 - k(k+1)/2 = (m-1)k(k+1)/2, \tag{5.1.19}$$

nes iš pradžių buvo $mk(k+1)/2$ nežinomų parametrų (m kovariacinių matricų elementai), o esant teisingai hipotezei lieka $k(k+1)/2$ nežinomų parametrų (vienos kovariacinės matricos elementai). Pagal asimptotinį tikėtinumų santykio reikšmingumo lygmens α kriterijų hipotezė (5.1.1) atmetama, kai

$$Z = -2 \ln \Lambda > \chi^2_\alpha(f), \tag{5.1.20}$$

arba P reikšmių terminais, kai

$$pv_a = \mathbf{P}\{\chi^2_f > z\} < \alpha, \tag{5.1.21}$$

čia z yra statistikos Z realizacija.

Naudojant Bartleto modifikuotąją tikėtinumų santykio statistiką $\tilde{\Lambda}$, kuri gaunama imant nepaslinktuosius kovariacinių matricų įvertinius, asimptotinis skirstinys išliks nepakitęs:

$$-2 \ln \tilde{\Lambda} = -2 \ln V + k(\nu_1 \ln \nu_1 + ... + \nu_m \ln \nu_m - \nu \ln \nu) \xrightarrow{d} \chi^2_f. \tag{5.1.22}$$

Pagal asimptotinį modifikuotą Bartleto tikėtinumų santykio reikšmingumo lygmens α kriterijų hipotezė (5.1.1) atmetama, kai

$$\tilde{Z} = -2 \ln \tilde{\Lambda} > \chi^2_\alpha(f), \tag{5.1.23}$$

arba P reikšmių terminais, kai

$$\tilde{pv}_a = \mathbf{P}\{\chi^2_f > \tilde{z}\} < \alpha, \tag{5.1.24}$$

čia $\tilde{z}$ yra statistikos $\tilde{Z}$ realizacija.

5.1.3 pavyzdys. (**2.2.1** pavyzdžio tęsinys). **2.2.1** pavyzdyje buvo tikrinama dviejų trimačių vidurkių vektorių lygybės hipotezė, tariant, kad turimos didumo $n_1 = 30$ ir $n_2 = 24$ paprastosios imtys gautos stebint nepriklausomus trimačius normaliuosius vektorius. Buvo priimta prielaida, kad abiejų imčių kovariacinės matricos yra vienodos. Bandysime pagrįsti šios prielaidos teisingumą tikrindami kovariacinių matricų lygybės hipotezę $H : \mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$.

1.2.1 ir **2.2.1** pavyzdžiuose surastos matricų $\boldsymbol{S}_1$ ir $\boldsymbol{S}_2$ realizacijos

$$\boldsymbol{S}_1 = \begin{pmatrix} 0,348 & 0,248 & 0,160 \\ 0,264 & 0,392 & 0,070 \\ 0,160 & 0,070 & 0,240 \end{pmatrix}, \quad \boldsymbol{S}_2 = \begin{pmatrix} 0,313 & 0,210 & 0,103 \\ 0,210 & 0,385 & 0,163 \\ 0,103 & 0,163 & 0,300 \end{pmatrix}.$$

Randame sumą

$$\boldsymbol{S} = \boldsymbol{S}_1 + \boldsymbol{S}_2 = \begin{pmatrix} 0,661 & 0,474 & 0,263 \\ 0,474 & 0,777 & 0,233 \\ 0,263 & 0,233 & 0,540 \end{pmatrix}$$

ir determinantus

$$|\boldsymbol{S}_1| = 0,0101, \quad |\boldsymbol{S}_2| = 0,0176, \quad |\boldsymbol{S}| = 0,1245.$$

Apskaičiuojame

$$z = -2\ln\Lambda = k(n_1\ln n_1 + n_2\ln n_2 - n\ln n) - n_1\ln|\boldsymbol{S}_1| - n_2\ln|\boldsymbol{S}_2| + n\ln|\boldsymbol{S}| = 11,0191$$

ir randame **P** reikšmę $pv_a = \mathbf{P}\{\chi^2_6 > 11,0191\} = 0,0878$. Hipotezė atmetama, jeigu kriterijaus reikšmingumo lygmuo viršija 0,0878.

Taikydami modifikuotąjį Bartleto kriterijų randame $\tilde{z} = 10,7485$ ir $\tilde{pv}_a = \mathbf{P}\{\chi^2_6 > 10,7485\} = 0,0965$. Hipotezė atmetama, jeigu kriterijaus reikšmingumo lygmuo viršija 0,0965.

5.1.4 pavyzdys. (**3.5.1** pavyzdžio tęsinys). **3.5.1** pavyzdyje buvo tikrinama trijų trimačių vidurkių vektorių lygybės hipotezė, tariant, kad turimos didumo $n_1 = 30$, $n_2 = 24$ ir $n_3 = 18$ paprastosios imtys gautos stebint nepriklausomus trimačius normaliuosius vektorius. Buvo priimta prielaida, kad trijų imčių kovariacinės matricos yra vienodos. Tikrinsime kovariacinių matricų lygybės hipotezę $H : \mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}_3$.

Pirmųjų dviejų imčių matricų $\boldsymbol{S}_1$ ir $\boldsymbol{S}_2$ realizacijos pateiktos **5.1.3** pavyzdyje. Trečiosios imties matricos $\boldsymbol{S}_3$ realizacija surasta **3.5.1** pavyzdyje.

$$\boldsymbol{S}_3 = \begin{pmatrix} 0,225 & 0,162 & 0,273 \\ 0,162 & 0,476 & 0,366 \\ 0,273 & 0,366 & 0,518 \end{pmatrix}, \quad \boldsymbol{S} = \boldsymbol{S}_1 + \boldsymbol{S}_2 + \boldsymbol{S}_3 = \begin{pmatrix} 0,886 & 0,636 & 0,536 \\ 0,636 & 1,253 & 0,599 \\ 0,536 & 0,599 & 1,058 \end{pmatrix}.$$

Randame determinantus $|\boldsymbol{S}_3| = 0,0086$ ir $|\boldsymbol{S}| = 0,4771$ ir statistikų $Z = -2\ln\Lambda$ ir $\tilde{Z} = -2\ln\tilde{\Lambda}$ realizacijas: $z = 34,3868$ ir $\tilde{z} = 33,3056$. Atitinkamos P reikšmės yra $pv_a = 0,00059$ ir $\tilde{pv}_a = 0,00088$. Hipotezė atmestina.

5.1.5. Asimptotinio skirstinio patikslinimas

Statistikos (5.1.7) eilės h momentas yra

$$\mathbf{E}(\tilde{\Lambda}^h) = \mathbf{E}\left(e^{h\ln\tilde{\Lambda}}\right) = \left(\frac{\nu^{k\nu/2}}{\nu_1^{k\nu_1/2}\cdots\nu_m^{\nu_m/2}}\right)^h \mathbf{E}(V^h). \tag{5.1.25}$$

Šioje lygybėje vietoje h įrašius $-2it$, gausime statistikos $\tilde{Z} = -2\ln\tilde{\Lambda}$ charakteristinę funkciją. Remiantis (5.1.12) gaunama

$$\psi(t) = \left(\frac{\nu^{-k\nu it}}{\nu_1^{-k\nu_1 it}\cdots\nu_m^{-k\nu_m it}}\right)\prod_{r=1}^{k}\left\{\prod_{j=1}^{m}\left[\frac{\Gamma(\frac{\nu_j+1-r}{2}) - it\nu_j}{\Gamma(\frac{\nu_j+1-r}{2})}\right]\frac{\Gamma(\frac{\nu+1-r}{2})}{\Gamma(\frac{\nu+1-r}{2} - it\nu)}\right\}. \tag{5.1.26}$$

Skleisdami funkciją $\ln\psi(t)$ Teiloro eilute (it) laipsniais, gauname

$$\ln\psi(t) = \sum_s \kappa_s \frac{(it)^s}{s!}, \tag{5.1.27}$$

čia $\kappa_1, \kappa_2, ...,$ yra a. d. $\tilde{Z} = -2\ln\Lambda$ semiinvariantai.

Pirmasis semiinvariantas

$$\kappa_1 = k(\sum_{j=1}^m \nu_j \ln\nu_j - \nu\ln\nu) - \sum_{r=1}^k \left\{ \sum_{j=1}^m \left[\nu_j\varphi_1\left(\frac{\nu_j+1-r}{2}\right)\right] - \nu\varphi_1\left(\frac{\nu+1-r}{2}\right)\right\},$$

čia $\varphi_1(x)$ yra funkcijos $\ln\Gamma(x)$ pirmoji išvestinė (digama funkcija).

Analogiškai gauname kitus semiinvariantus

$$\kappa_s = \sum_{r=1}^k \left\{ \sum_{j=1}^m \left[(-\nu_j)^s\varphi_s\left(\frac{\nu_j+1-r}{2}\right)\right] - (-\nu)^s\varphi_s\left(\frac{\nu+1-r}{2}\right)\right\},$$

čia $\varphi_s(x)$ yra funkcijos $\ln\Gamma(x)$ s-oji išvestinė.

Jeigu vietoje statistikos $\tilde{Z} = -2\ln\tilde{\Lambda}$ imsime statistiką $Z^* = \delta\tilde{Z}$, kai $\delta = f/\kappa_1$, tai a. d. Z^* vidurkis $\mathbf{E}Z^* = f$ tiksliai sutampa su asimptotinio skirstinio (5.1.18) vidurkiu f. Aproksimuodami statistikos Z^* skirstinį χ^2 skirstiniu su f laisvės laipsniais, gauname patikslintą tikėtinumų santykio kriterijų (sutapatinami vidurkiai): hipotezė atmetama patikslintu asimptotiniu tikėtinumų santykio reikšmingumo lygmens α kriterijumi, kai

$$pv_a^{(1)} = \mathbf{P}\{\chi^2_f > \delta z^*\} < \alpha, \tag{5.1.28}$$

čia z^* yra statistikos Z^* realizacija.

Analogiškai skyreliui 3.5.5 galima sudaryti patikslinimus sutapatinant 2 ar 3 momentus.

Kaip ir skyrelyje 3.5.5.4, aproksimacijos (4.5.20) tikslumą galima padidinti skleidžiant funkciją $\ln\varphi(\delta t)$ ne (it), o $(1-2it)$ laipsniais:

$$\ln\varphi(\delta t) = -\frac{\nu}{2}\ln(1-2it) + \sum_{s=1}^l \frac{\omega_s}{(n\delta)^s}\left[\frac{1}{(1-2it)^s} - 1\right] + O(1/n^{l+1}). \tag{5.1.29}$$

Kaip ir 3.5.5.4 skyrelyje tikslinga parinkti δ taip, kad ω_1 skleidinyje (5.1.29) virstų 0. Gauname, kad δ reikia parinkti taip:

$$\delta = 1 - \left[\sum_j \left(\frac{1}{\nu_j}\right) - \frac{1}{\nu}\right] \frac{2k^2+3k-1}{6(k+1)(m-1)}. \tag{5.1.30}$$

Apsiribojęs nariu $s = 1$ (dalinis vidurkio sutapatinimas), gausime patikslintą kriterijų P reikšmių terminais: hipotezė atmetama apytiksliu reikšmingumo lygmens α kriterijumi, kai

$$pv_a^{(2)} = \mathbf{P}\{\chi^2_f > \delta\tilde{z}\} < \alpha, \tag{5.1.31}$$

čia $\tilde{z}$ yra statistikos $\tilde{Z}$ realizacija.

Imdami ir narį $s = 2$, gauname tokį P reikšmės patikslinimą (analogišką teoremai 3.5.6):

$$pv_a^{(3)} = \mathbf{P}\{\chi^2_\nu > \delta\tilde{z}\} - \omega_2[\mathbf{P}\{\chi^2_{f+4} > \delta\tilde{z}\} - \mathbf{P}\{\chi^2_f > \delta\tilde{z}\}] + O(1/n^3), \quad (5.1.32)$$

$$\omega_2 = \frac{k(k+1)[(k-1)(k+2)(\sum_j (1/\nu_j^2) - 1/\nu^2) - 6(m-1)(1-\delta)^2]}{48\delta^2}.$$

5.1.5 pavyzdys. (**5.1.3** pavyzdžio tęsinys). Patikslinkime asimptotinį tikėtinumų santykio kriterijų remdamiesi pateiktais aproksimacijų patikslinimais.

Randame pirmojo semiinvarianto (vidurkio) reikšmę $\kappa_1 = 6,37418$, $\delta\tilde{z} = 10,11754$, ir asimptotinę patikslintą (sutapatinant vidurkius) P reikšmę

$$pv_a^{(1)} = \mathbf{P}\{\chi^2_6 > 10,11754\} = 0,1198.$$

Apsiribojęs pirmuoju asimptotinio skleidinio nariu pagal (5.1.29), gauname $\delta = 0,9364$, $\delta\tilde{z} = 10,0649$ ir asimptotinę patikslintą P reikšmę

$$pv_a^{(2)} = \mathbf{P}\{\chi^2_6 > 10,0649\} = 0,1219.$$

Pridėję antrąjį asimptotinio skleidinio narį pagal (5.1.32) gausime $\omega_2 = 0,00081$ ir

$$pv_a^{(3)} = pv_a^{(2)} - 0,0003 = 0,1216.$$

Matome, kad pataisa lyginant su $pv_a^{(2)}$ yra maža. Tačiau skirtumas nuo pv_a, kai patikslinimas neatliekamas, gana didelis.

Gauname tokią išvadą. Kai imtys nedidelės, asimptotinį tikėtinumų santykio kriterijų reikia koreguoti atliekant, pavyzdžiui, pateiktus patikslinimus. Reikšmingiausią indėlį mažinant paklaidą sudaro tikslus ar apytikslis vidurkio sutapatinimas. Tolesnių pataisų įtaka kur kas mažesnė.

5.1.6 pavyzdys. (**5.1.4** pavyzdžio tęsinys). Patikslinkime asimptotinį tikėtinumų santykio kriterijų remdamiesi pateiktais aproksimacijų patikslinimais.

Randame pirmojo semiinvarianto (vidurkio) reikšmę $\kappa_1 = 12,8648$, $\delta\tilde{z} = 31,06665$, ir asimptotinę patikslintą (sutapatinant vidurkius) P reikšmę

$$pv_a^{(1)} = \mathbf{P}\{\chi^2_{12} > 31,06665\} = 0,00192.$$

Apsiribojęs pirmuoju asimptotinio skleidinio nariu pagal (5.1.29), gauname $\delta = 0,93343$, $\delta\tilde{z} = 31,0884$ ir asimptotinę patikslintą P reikšmę

$$pv_a^{(2)} = \mathbf{P}\{\chi^2_{12} > 31,0884\} = 0,00191.$$

Prijungę antrąjį asimptotinio skleidinio narį pagal (5.1.32), gausime $\omega_2 = 0,00295$ ir

$$pv_a^{(3)} = pv_a^{(2)} - 0,00005 = 0,00186.$$

Gauname analogiškas išvadas kaip ir **5.1.5** pavyzdyje.

5.2. Proporcingumo hipotezės tikrinimas

5.2.1. Proporcingumo (sferiškumo) hipotezė

Tarkime, kad $\mathbf{X}_1, ..., \mathbf{X}_n$ yra paprastoji imtis a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Reikia patikrinti hipotezę

$$H : \boldsymbol{\Sigma} = \sigma^2 \boldsymbol{\Sigma}_0, \quad (5.2.1)$$

kad kovariacijų matrica $\boldsymbol{\Sigma}$ proporcinga žinomai matricai $\boldsymbol{\Sigma}_0$. Matrica $|\boldsymbol{\Sigma}_0| > 0$ teigiamai apibrėžta, o σ^2 – nežinomas proporcingumo koeficientas.

5.2.2. Tikėtinumų santykio kriterijus

Rasime tikėtinumų santykio statistiką hipotezei H tikrinti.

5.2.1 teorema. *Tarkim, kad* $|\mathbf{\Sigma}| > 0, n > k$ *ir žinoma matrica* $|\mathbf{\Sigma}_0| > 0$ *taip pat teigiamai apibrėžta. Tada tikėtinumų santykio statistika* Λ *hipotezei* H *tikrinti turi tokį pavidalą*

$$\Lambda = \frac{\max_{\boldsymbol{\mu},\mathbf{\Sigma}=\mathbf{\Sigma}_0} L(\boldsymbol{\mu},\mathbf{\Sigma})}{\max_{\boldsymbol{\mu},\mathbf{\Sigma}} L(\boldsymbol{\mu},\mathbf{\Sigma})} = \frac{|\boldsymbol{S}\mathbf{\Sigma}_0^{-1}|^{n/2}}{(Tr(\boldsymbol{S}\mathbf{\Sigma}_0^{-1})/k)^{nk/2}}, \tag{5.2.2}$$

čia

$$\bar{\mathbf{X}} = \sum_{j=1}^{n} \mathbf{X}_j, \quad \boldsymbol{S} = \sum_{j=1}^{n} (\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})^T.$$

Įrodymas. Kadangi žinoma matrica $\mathbf{\Sigma}_0$ simetrinė ir teigiamai apibrėžta, tai egzistuoja tokia neišsigimusi kvadratinė matrica $\boldsymbol{C}$ (1 priedas, (8.2.10)), kad

$$\boldsymbol{C}\mathbf{\Sigma}_0\boldsymbol{C}^T = \boldsymbol{I}. \tag{5.2.3}$$

Atlikime a. v.. $\mathbf{X}_j$ tiesines transformacijas naudodami matricą $\boldsymbol{C}$:

$$\boldsymbol{Y}_j = \boldsymbol{C}\mathbf{X}_j, \quad j = 1, ..., n. \tag{5.2.4}$$

Tada vektoriai $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_n$ taip pat yra paprastoji imtis, gauta stebint normalųjį vektorių $\boldsymbol{Y} \sim N_k(\boldsymbol{\nu}, \mathbf{\Psi})$:

$$\mathbf{E}(\boldsymbol{Y}_j) = \boldsymbol{C}\boldsymbol{\mu} = \boldsymbol{\nu}, \quad \boldsymbol{V}(\boldsymbol{Y}_j) = \sigma^2\boldsymbol{C}\mathbf{\Sigma}\boldsymbol{C}^T = \mathbf{\Psi}.$$

Atsitiktinių vektorių $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_n$ terminais hipotezė H ekvivalenti hipotezei $H^* : \mathbf{\Psi} = \sigma^2\boldsymbol{I}$, čia $\boldsymbol{I}$ vienetinė matrica.

Randame tikėtinumų santykio statistiką hipotezei H^* tikrinti:

$$\Lambda^* = \frac{\max_{\boldsymbol{\nu},\mathbf{\Psi}=\sigma^2\boldsymbol{I}} L(\boldsymbol{\nu},\mathbf{\Psi})}{\max_{\boldsymbol{\nu},\mathbf{\Psi}} L(\boldsymbol{\nu},\mathbf{\Psi})} = \frac{L(\tilde{\boldsymbol{\nu}}, \tilde{\sigma}^2\boldsymbol{I})}{L(\hat{\boldsymbol{\nu}}, \hat{\mathbf{\Psi}})}$$

čia $\hat{\boldsymbol{\nu}}$ ir $\hat{\mathbf{\Psi}}$ yra parametrų $\boldsymbol{\nu}$ ir $\mathbf{\Psi}$ DT įvertiniai, o $\tilde{\boldsymbol{\nu}}$ ir $\tilde{\sigma}^2$ yra parametrų $\boldsymbol{\nu}$ ir σ^2 DT įvertiniai surasti tarus, kad hipotezė H^* teisinga.

Remiantis 1.1 skyreliu parametrų $\boldsymbol{\nu}$ ir $\mathbf{\Psi}$ didžiausiojo tikėtinumo įvertiniai yra

$$\hat{\boldsymbol{\nu}} = \bar{\boldsymbol{Y}} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{Y}_i, \quad \hat{\mathbf{\Psi}} = \frac{1}{n}\sum_{i=1}^{n} (\boldsymbol{Y}_i - \bar{\boldsymbol{Y}})(\boldsymbol{Y}_i - \bar{\boldsymbol{Y}})^T = \frac{1}{n}\boldsymbol{S}^* = \frac{1}{n}[S_{ij}^*]_{k\times k},$$

o tikėtinumo funkcijos maksimumas

$$L(\hat{\boldsymbol{\nu}}, \hat{\mathbf{\Psi}}) = (2\pi)^{-nk/2} n^{nk/2} |\boldsymbol{S}^*|^{-n/2} \exp\{-\frac{nk}{2}\}.$$

Kai hipotezė H^* teisinga, skaitiklyje turime tikėtinumo funkciją, gautą iš k nepriklausomų paprastųjų imčių, gautų stebint vienmačius normaliuosius dydžius su vienodomis dispersijomis. Taigi $\tilde{\boldsymbol{\nu}} = \hat{\boldsymbol{\nu}}$, o parametro σ^2 DT įvertinys yra $\tilde{\sigma}^2 = (S_{11}^* + ... + S_{kk}^*)/(nk) = Tr(\boldsymbol{S}^*)/(nk)$. Tikėtinumo funkcijos sąlyginis maksimumas

$$L(\tilde{\boldsymbol{\nu}}, \tilde{\sigma}^2\boldsymbol{I}) = (2\pi)^{-nk/2}(nk)^{nk/2}(S_{11}^* + ... + S_{kk}^*)^{-nk/2}\exp\{-\frac{nk}{2}\}.$$

Padalinę gauname, kad tikėtinumų santykis hipotezei H^* tikrinti yra

$$\Lambda^* = \frac{|\boldsymbol{S}^*|^{n/2}}{((S_{11}^* + ... + S_{kk}^*)/k)^{nk/2}}. \tag{5.2.5}$$

Remdamiesi (5.2.3), (5.2.4) gauname

$$\boldsymbol{S}^* = \boldsymbol{C}\boldsymbol{S}\boldsymbol{C}^T, \quad \boldsymbol{\Sigma}_0 = (\boldsymbol{C}^T\boldsymbol{C})^{-1}, \quad \boldsymbol{\Sigma}_0^{-1} = \boldsymbol{C}\boldsymbol{C}^T,$$

todėl

$$|\boldsymbol{S}^*| = |\boldsymbol{C}\boldsymbol{S}\boldsymbol{C}^T| = |\boldsymbol{S}||\boldsymbol{C}\boldsymbol{C}^T| = |\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}|,$$

$$Tr(\boldsymbol{S}^*) = Tr(\boldsymbol{C}\boldsymbol{S}\boldsymbol{C}^T) = Tr(\boldsymbol{S}\boldsymbol{C}^T\boldsymbol{C}) = Tr(\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}). \tag{5.2.6}$$

Įrašę šias išraiškas į (5.2.5), gauname (5.2.2). Statistika Λ^* nepriklauso nuo matricos $\boldsymbol{C}$ parinkimo ir sutampa su Λ. ▲

Reikia pažymėti, kad jei $\theta_1, ..., \theta_k$ yra lygties $|\boldsymbol{S}^* - \theta\boldsymbol{I}| = 0$ šaknys, tai statistiką Λ^* galima išreikšti šaknų terminais:

$$\Lambda^* = \left(\frac{\prod_i \theta_i^{1/k}}{\sum_i \theta_i/k}\right)^{nk/2}. \tag{5.2.7}$$

Statistika Λ^* yra šaknų geometrinio ir aritmetinio vidurkių santykio tam tikras laipsnis. Jeigu į (5.2.7) įrašytume lygties $|\boldsymbol{\Psi} - \theta\boldsymbol{I}| = 0$ šaknis, kai H^* teisinga, tai gautume 1. Taigi, kai hipotezė teisinga, tikėtinumų santykio statistikos reikšmės bus sukoncentruotos 1 aplinkoje.

Tikrinama hipotezė H atmetama reikšmingumo lygmens α tikėtinumų santykio kriterijumi, kai teisinga nelygybė

$$\Lambda < \Lambda_{1-\alpha},$$

čia $\Lambda_{1-\alpha}$ yra statistikos Λ eilės $1-\alpha$ kritinė reikšmė.

Kartais rekomenduojama sudarant statistiką naudoti nepaslinktuosius dispersijų įvertinius. Modifikuotoji tikėtinumų santykio statistika $\tilde{\Lambda}$ gaunama pakeičiant n į $\nu = n - 1$:

$$\tilde{\Lambda} = \frac{|\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}|^{\nu/2}}{(Tr(\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1})/k)^{\nu k/2}}$$

Hipotezė H atmetama modifikuotuoju reikšmingumo lygmens α tikėtinumų santykio kriterijumi, kai teisinga nelygybė

$$\tilde{\Lambda} < \tilde{\Lambda}_{1-\alpha}.$$

Norint rasti kritines reikšmes, reikia ištirti tikėtinumų santykio statistikos savybes esant teisingai hipotezei.

5.2.3. Tikėtinumų santykio statistikos momentai

Ieškosime statistikos

$$U = \Lambda^{2/n} k^{-k} = \tilde{\Lambda}^{2/\nu} k^{-k} \tag{5.2.8}$$

momento $\mathbf{E}(U^h)$, kai hipotezė H yra teisinga.

5.2.2 teorema. *Jeigu hipotezė H teisinga ir $n > k$, tai momentas $\mathbf{E}(U^h)$ turi tokį pavidalą:*

$$\mathbf{E}(U^h) = \frac{\Gamma(k(n-1)/2)}{\Gamma(k(n-1)/2 + kh)} \prod_{i=1}^{k} \frac{\Gamma((n-i)/2 + h)}{\Gamma((n-i)/2)}. \tag{5.2.9}$$

Įrodymas. Remiantis 5.2.1 teoremos įrodymu statistikos Λ skirstinys sutampa su statistikos Λ^*, apibrėžtos (5.2.5) formule, skirstiniu. Todėl

$$U \overset{d}{\sim} \frac{|\boldsymbol{S}^*|}{(S_{11}^* + ... + S_{kk}^*)^k}, \tag{5.2.10}$$

čia matrica $\boldsymbol{S}^*$, kai hipotezė teisinga, turi Višarto skirstinį $\boldsymbol{S}^* \sim W_k(\nu, \sigma^2 \boldsymbol{I})$. Remdamiesi (1.6.13) tankio išraiška gauname

$$\mathbf{E}(U^h) = \int \cdots \int \frac{|\boldsymbol{S}^*|^h}{(Tr(\boldsymbol{S}^*))^{kh}} f(\boldsymbol{S}^* | k, \nu, \sigma^2 \boldsymbol{I}) d\boldsymbol{S}^*$$

$$\int \cdots \int \frac{|\boldsymbol{S}^*|^h}{(Tr(\boldsymbol{S}^*))^{kh}} \frac{|\boldsymbol{S}^*|^{\frac{\nu-k-1}{2}} \exp\{-\frac{1}{2\sigma^2} Tr(\boldsymbol{S}^*)\}}{K(k, \nu, \sigma^2 \boldsymbol{I})} d\boldsymbol{S}^*$$

$$= \frac{K(k, \nu + 2h, \sigma^2 \boldsymbol{I})}{K(k, \nu, \sigma^2 \boldsymbol{I})} \int \cdots \int (Tr(\boldsymbol{S}^*))^{-kh} \frac{|\boldsymbol{S}^*|^{\frac{\nu+2h-k-1}{2}} \exp\{-\frac{1}{2\sigma^2} Tr(\boldsymbol{S}^*)\}}{K(k, \nu + 2h, \sigma^2 \boldsymbol{I})} d\boldsymbol{S}^*.$$

Integralas reiškia funkcijos $(S_{11}^* + ... + S_{kk}^*)$ momentą $(S_{11}^* + ... + S_{kk}^*)^{-kh}$, kai vidurkinama pagal Višarto skirstinio $W_k(\nu + 2h, \sigma^2 \boldsymbol{I})$ tankį. Kadangi funkcija priklauso tik nuo matricos $\boldsymbol{S}^*$ diagonalinių elementų, tai iš pradžių galima suintegruoti pagal visus argumentus $S_{ij}^*, i \neq j$. Tada po integralo ženklu gausime a. v. $(S_{11}^*, ..., S_{kk}^*)^T$ tankį, kuris yra lygus Višarto skirstinių $W_1(\nu + 2h, \sigma^2)$ tankių sandaugai. Remiantis trečia Višarto skirstinio savybe $S_{11}^* + ... + S_{kk}^* \sim W_1(k\nu + 2kh, \sigma^2)$. Taigi

$$\mathbf{E}(U^h) = \frac{K(k, \nu + 2h, \sigma^2 \boldsymbol{I})}{K(k, \nu, \sigma^2 \boldsymbol{I})} \int \cdots \int \frac{(Tr(\boldsymbol{S}^*))^{\frac{k\nu-2}{2}} \exp\{-\frac{1}{2\sigma^2} Tr(\boldsymbol{S}^*)\}}{K(1, k\nu + 2kh, \sigma^2)} \times$$

$$\times dS_{11}^* \cdots dS_{kk}^* = \frac{K(k, \nu + 2h, \sigma^2 \boldsymbol{I})}{K(k, \nu, \sigma^2 \boldsymbol{I})} \frac{K(1, k\nu, \sigma^2)}{K(1, k\nu + 2kh, \sigma^2)},$$

nes likęs integralas lygus 1.

Remdamiesi normuojančių konstantų iš (1.6.13) išraiškomis gauname (5.2.9). ▲

5.2.4. Tikėtinumų santykio skirstinys

Momentus (5.2.9) galima traktuoti kaip a. d. $\ln U$ momentus generuojančią funkciją, kuri, kai $n > k$, egzistuoja h kitimo intervale $|h| < 1/2$, apimančiame tašką $h = 0$. Todėl momentai (5.2.9) visiškai nusako a. d. $\ln U$, o kartu ir a. d. U skirstinį.

5.2.3 teorema. *Jeigu hipotezė H teisinga, $n > k$, tai a. d. U skirstinys sutampa su skirstiniu polinomo nuo nepriklausomų a. d., turinčių beta skirstinius:*

$$U \overset{d}{\sim} \prod_{i=2}^{k} \xi_{i1} \prod_{j=1}^{k-1} [\xi_{1j}^j (1 - \xi_{1j})]; \tag{5.2.11}$$

čia $\xi_{21}, ..., \xi_{k1}, \xi_{11}, ..., \xi_{1,k-1}$ yra nepriklausomi a. d., turintys beta skirstinius:

$$\xi_{i1} \sim Be((n-i)/2, (i-1/2)), \quad \xi_{1j} \sim Be(j(n-1)/2, (n-1)/2).$$

Įrodymas. Tikėtinumų santykį Λ^* galima išreikšti dviejų tikėtinumų santykių sandauga

$$\Lambda^* = \Lambda_1^* \cdot \Lambda_2^* = \left(\frac{\max_{\boldsymbol{\nu}, \boldsymbol{\Psi} = \Delta} L(\boldsymbol{\nu}, \boldsymbol{\Psi})}{\max_{\boldsymbol{\nu}, \boldsymbol{\Psi}} L(\boldsymbol{\nu}, \boldsymbol{\Psi})} \right) \cdot \left(\frac{\max_{\boldsymbol{\nu}, \boldsymbol{\Psi} = \sigma^2 \boldsymbol{I}} L(\boldsymbol{\nu}, \boldsymbol{\Psi})}{\max_{\boldsymbol{\nu}, \boldsymbol{\Psi} = \Delta} L(\boldsymbol{\nu}, \boldsymbol{\Psi})} \right), \tag{5.2.12}$$

čia Δ yra diagonali matrica. Taigi hipotezę H^* galima tikrinti kaip sudėtinę hipotezę. Visų pirma tikriname hipotezę, kad vektoriaus $\boldsymbol{Y}$ koordinatės yra nepriklausomos. Šiai hipotezei tikėtinumų santykis yra Λ_1^*. Paskui tikriname hipotezę, kad vektoriaus $\boldsymbol{Y}$ koordinačių dispersijos yra lygios, jei žinoma, kad jos yra nepriklausomos. Tokiai hipotezei tikrinti tikėtinumų santykis yra Λ_2^*.

Pirma hipotezė yra hipotezės, nagrinėtos skyrelyje 4.5, atskiras atvejis. Formulėje (4.5.3) reikia imti $m = k, k_1 = k_2 = ... = k_m = 1$. Antra hipotezė yra hipotezės, nagrinėtos 5.1 skyrelyje, atskiras atvejis. Formulėje (5.1.5) vietoje parametrų $k; m; n_1, ..., n_m$, reikia imti $1; k; n_1 = ... = n_k = n$. Gauname

$$\Lambda^* = \Lambda_1^* \cdot \Lambda_2^* = \left(\frac{|\boldsymbol{S}^*|}{S_{11}^* \cdot ... \cdot S_{kk}^*} \right)^{n/2} \cdot \left(\frac{k^{nk/2} (S_{11}^* \cdot ... \cdot S_{kk}^*)^{n/2}}{(S_{11}^* + ... + S_{kk}^*)^{nk/2}} \right). \tag{5.2.13}$$

Pirmas daugiklis išreiškiamas empiriniais koreliacijos koeficientais

$$r_{ij}^* = S_{ij}^* \sqrt{S_{ii}^* S_{jj}^*},$$

o antrasis – tik diagonaliniais elementaiss S^*_{ii}. Kai stebimo vektoriaus koordinatės nepriklausomos, tai vektorius, sudarytas iš stebėjimus atitinkančios Višarto matricos diagonalinių elementų, ir vektorius, sudarytas iš koreliacijos koeficientų, yra nepriklausomi (žr. 4.4 pratimą). Todėl tikėtinumų santykiai Λ_1^* ir Λ_2^* yra nepriklausomi.

4.5 skyrelio (4.5.8) formulėje buvo gauta

$$(\Lambda_1^*)^{2/n} \overset{d}{\sim} \prod_{i=2}^{k} \xi_{i1}, \quad \xi_{i1} \sim Be\left(\frac{n-i}{2}, \frac{i-1}{2}\right), \quad i = 2, ..., k. \tag{5.2.14}$$

5.1 skyrelio (5.1.14) formulėje buvo gauta

$$(\Lambda_2^*)^{2/n} k^{-k} \overset{d}{\sim} \left\{ \prod_{j=1}^{k-1} \left[\xi_{1j}^{j\nu/2} (1-\xi_{1j})^{\nu/2} \right] \right\}^{2/\nu} = \prod_{j=1}^{k-1} \left[\xi_{1j}^{j} (1-\xi_{1j}) \right], \tag{5.2.15}$$

$$\xi_{1j} \sim Be\left(\frac{j(n-1)}{2}, \frac{n-1}{2}\right), \quad j = 1, ..., k-1.$$

Sujungę (5.2.14) ir (5.2.15), gauname (5.2.11). ▲

Formulę (5.2.11) galima naudoti ieškant a. d. U kritinių reikšmių skaitinio modeliavimo metodais.

Atvejis $k = 2$. Kai skirstinys dvimatis, statistikos U skirstinys įgauna labai paprastą pavidalą. Pritaikę gama funkcijos nuo dvigubo argumento savybę (3.5.8) gauname, kad momentas (5.2.9) įgyja tokią išraišką:

$$\mathbf{E}(U^h) = \frac{\Gamma(n-1)}{\Gamma(n-1+2h)} \frac{\Gamma((n-1+2h)/2)\Gamma((n-2+2h)/2)}{\Gamma((n-1)/2)\Gamma((n-2)/2)} =$$

$$= \frac{1}{4^h} \frac{\Gamma(n-1)\Gamma(n-2+2h)}{\Gamma(n-1+2h)\Gamma(n-2)} = \frac{1}{4^h} \frac{n-2}{n-2+2h}.$$

Nesunku patikrinti, kad tai a. d. $Z^2/4$ momentas $\mathbf{E}(Z^2/4)^h$, kai a. d. $Z \sim Be(n-2, 1)$.

Gauname, kad hipotezė H atmetama reikšmingumo lygmens α tikėtinumų santykio kriterijumi, kai

$$pv = \mathbf{P}\{Z^2 < 4u\} = (2\sqrt{u})^{n-2} < \alpha, \tag{5.2.16}$$

čia u yra statistikos U realizacija.

5.2.1 pavyzdys. (**4.5.1** pavyzdžio tęsinys). Lentelėje **4.5.1** yra pateikta $n = 30$ dvimačio a. v. realizacijų. Tarę, kad šie stebėjimai yra dvimačio normaliojo vektoriaus $\mathbf{X} \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ paprastosios imties realizacija, patikrinsime hipotezę H, kad kovariacinė matrica $\boldsymbol{\Sigma}$ proporcinga žinomai matricai $\boldsymbol{\Sigma}_0$:

$$H : \boldsymbol{\Sigma} = \sigma^2 \boldsymbol{\Sigma}_0, \quad \boldsymbol{\Sigma}_0 = \begin{pmatrix} 1 & 0,5 \\ 0,5 & 1,5 \end{pmatrix}.$$

Matricos $\boldsymbol{S}$ realizacija surasta **4.5.1** pavyzdyje. Randame

$$\boldsymbol{\Sigma}_0^{-1} = \begin{pmatrix} 1,2 & -0,4 \\ -0,4 & 0,8 \end{pmatrix}, \quad \boldsymbol{S} = \begin{pmatrix} 0,2137 & 0,2023 \\ 0,2023 & 0,3897 \end{pmatrix}, \quad \boldsymbol{S}\boldsymbol{\Sigma}_0^{-1} = \begin{pmatrix} 0,1755 & 0,0764 \\ 0,0764 & 0,2308 \end{pmatrix}.$$

Apskaičiuojame

$$|\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}| = 0,03469, \quad Tr(\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}) = 0,40636, \quad U = 0,2101$$

ir randame P reikšmę

$$pv = (2\sqrt{u})^{n-2} = 0,0877.$$

Hipotezė atmetama, jai kriterijaus reikšmingumo lygmuo viršija 0,0877.

5.2.5. Tikėtinumų santykio statistikos asimptotinis skirstinys

Jeigu imties dydis $n \to \infty$, tai remiantis tikėtinumų santykio asimptotinėmis savybėmis (1 dalis, 4.5.4 skyrelis)

$$-2\ln\Lambda \xrightarrow{d} \chi_f^2, \tag{5.2.17}$$

čia laisvės laipsnių skaičius

$$f = k(k+1)/2 - 1, \tag{5.2.18}$$

nes iš pradžių buvo $k + k(k+1)/2$ nežinomas parametras (vidurkių vektoriaus ir kovariacijų matricos elementai), o esant teisingai hipotezei lieka $k+1$ nežinomas parametras (vidurkių vektoriaus elementai ir vienas kovariacinės matricos elementas), t. y. hipotezė H uždeda parametrams f apribojimų. Pagal asimptotinį tikėtinumų santykio reikšmingumo lygmens α kriterijų hipotezė atmetama, kai

$$Z = -2\ln\Lambda > \chi_\alpha^2(f), \tag{5.2.19}$$

arba P reikšmių terminais, kai

$$pv_a = \mathbf{P}\{\chi_f^2 > z\} < \alpha, \tag{5.2.20}$$

čia z yra statistikos Z realizacija.

Naudojant modifikuotą tikėtinumų santykio statistiką $\tilde{\Lambda}$, kuri gaunama imant nepaslinktuosius kovariacinių matricų įvertinius, asimptotinis skirstinys išliks nepakitęs:

$$-2\ln\tilde{\Lambda} \xrightarrow{d} \chi_f^2. \tag{5.2.21}$$

Pagal asimptotinį modifikuotą tikėtinumų santykio reikšmingumo lygmens α kriterijų hipotezė H atmetama, kai

$$\tilde{Z} = -2\ln\tilde{\Lambda} > \chi_\alpha^2(f), \tag{5.2.22}$$

arba P reikšmių terminais, kai

$$\tilde{pv}_a = \mathbf{P}\{\chi_f^2 > \tilde{z}\} < \alpha, \tag{5.2.23}$$

čia $\tilde{z}$ yra statistikos $\tilde{Z}$ realizacija.

5.2.2 pavyzdys. (**5.2.1** pavyzdžio tęsinys). Palyginti išspręskime tą patį uždavinį kaip ir **5.2.1** pavyzdyje, taikydami asimptotinį tikėtinumų santykio kriterijų.

Randame statistikos $Z = -2\ln\Lambda$ realizaciją $z = 5,2163$ ir ją atitinkančią P reikšmę

$$pv_a = \mathbf{P}\{\chi_f^2 > z\} = \mathbf{P}\{\chi_2^2 > 5,2163\} = 0,0734.$$

Taikydami modifikuotąjį asimptotinį tikėtinumų santykio kriterijų gauname statistikos $\tilde{Z} = -2\ln\tilde{\Lambda}$ realizaciją $\tilde{z} = 5,0424$ ir ją atitinkančią P reikšmę

$$\tilde{pv}_a = \mathbf{P}\{\chi_f^2 > \tilde{z}\} = \mathbf{P}\{\chi_2^2 > 5,0424\} = 0,0804.$$

Matome, kad modifikuotasis kriterijus šiame pavyzdyje yra šiek tiek tikslesnis.

5.2.6. Asimptotinio skirstinio patikslinimai

Lygybėje (5.2.9) imdami $h = -it\nu$ ir padauginę iš $k^{-itk\nu}$ gauname a. d. $-2\ln\tilde{\Lambda}$ charakteristinę funkciją

$$\psi(t) = k^{-itk\nu}\frac{\Gamma(k\nu/2)}{\Gamma(k\nu/2 - it\nu k)}\prod_{j=1}^{k}\frac{\Gamma((\nu+1-j)/2 - it\nu)}{\Gamma((\nu+1-j)/2)}. \tag{5.2.24}$$

Skleisdami funkciją $\ln\psi(t)$ Teiloro eilute (it) laipsniais, gauname

$$\ln\psi(t) = \sum_s \kappa_s\frac{(it)^s}{s!}, \tag{5.2.25}$$

čia $\kappa_1, \kappa_2, ...,$ yra a. d. $\tilde{Z} = -2\ln\Lambda$ semiinvariantai.

Pirmasis semiinvariantas

$$\kappa_1 = -\nu k\ln k + \nu k\varphi_1\left(\frac{k\nu}{2}\right) - \sum_{j=1}^{k}\nu\varphi_1\left(\frac{\nu+1-j}{2}\right),$$

čia $\varphi_1(x)$ yra funkcijos $\ln\Gamma(x)$ pirmoji išvestinė (digama funkcija).

Analogiškai gauname kitus semiinvariantus

$$\kappa_s = (-1)^{s+1}(k\nu)^s\varphi_s\left(\frac{k\nu}{2}\right) + \sum_{j=1}^{k}(-1)^s\nu^s\varphi_s\left(\frac{\nu+1-j}{2}\right),$$

čia $\varphi_s(x)$ yra funkcijos $\ln\Gamma(x)$ s-oji išvestinė.

Jeigu vietoje statistikos $\tilde{Z} = -2\ln\tilde{\Lambda}$ imsime statistiką $Z^* = \delta\tilde{Z}$, kai $\delta = f/\kappa_1$, tai a. d. Z^* vidurkis $\mathbf{E}Z^* = f$ tiksliai sutampa su asimptotinio skirstinio (5.2.21) vidurkiu f.

Aproksimuodami statistikos Z^* skirstinį χ^2 skirstiniu su f laisvės laipsniais, gauname patikslintą tikėtinumų santykio kriterijų (sutapatinami vidurkiai): hipotezė atmetama patikslintu asimptotiniu tikėtinumų santykio reikšmingumo lygmens α kriterijumi, kai

$$pv_a^{(1)} = \mathbf{P}\{\chi_f^2 > \delta z^*\} < \alpha, \tag{5.2.26}$$

čia z^* yra statistikos Z^* realizacija.

Analogiškai skyreliui 3.5.5 galima sudaryti patikslinimus sutapatinant 2 ar 3 momentus.

Kaip ir skyrelyje 3.5.5.4, aproksimacijos (5.2.21) tikslumą galima padidinti skleidžiant funkciją $\ln\varphi(\delta t)$ ne (it), o $(1-2it)$ laipsniais:

$$\ln\varphi(\delta t) = -\frac{\nu}{2}\ln(1-2it) + \sum_{s=1}^{l}\frac{\omega_s}{(n\delta)^s}\left[\frac{1}{(1-2it)^s}-1\right] + O(1/n^{l+1}). \quad (5.2.27)$$

Kaip ir 3.5.5.4 skyrelyje tikslinga parinkti δ taip, kad ω_1 skleidinyje (5.2.27) virstų 0. Gauname, kad δ reikia parinkti taip:

$$\delta = 1 - (2k^2+k+2)/(6k\nu). \quad (5.2.28)$$

Apsiriboję tik nariu $s=1$ (dalinis vidurkio sutapatinimas), gausime patikslintą kriterijų P reikšmių terminais: hipotezė atmetama apytiksliu reikšmingumo lygmens α kriterijumi, kai

$$pv_a^{(2)} = \mathbf{P}\{\chi_f^2 > \delta\tilde{z}\} < \alpha, \quad (5.2.29)$$

čia $\tilde{z}$ yra statistikos $\tilde{Z}$ realizacija.

Paėmę ir narį $s=2$, gauname tokį P reikšmės patikslinimą (analogišką teoremai 3.5.6):

$$pv_a^{(3)} = \mathbf{P}\{\chi_\nu^2 > \delta\tilde{z}\} - \omega_2[\mathbf{P}\{\chi_{f+4}^2 > \delta\tilde{z}\} - \mathbf{P}\{\chi_f^2 > \delta\tilde{z}\}] + O(1/n^3), \quad (5.2.30)$$

$$\omega_2 = (k+2)(k-1)(k-2)(2k^3+6k^2+3k+2)/(288k^2\nu^2\delta^2).$$

5.2.3 pavyzdys. (**5.2.1** pavyzdžio tęsinys). Spręsime uždavinį iš pavyzdžio **5.2.1** naudodami pateiktus aproksimacijos patikslinimus. Remiantis 5.2.4 skyreliu atveju $k=2$ statistikos $-2\ln\tilde{\Lambda}$ charakteristinė funkcija

$$\psi(t) = (1-2it\nu/(\nu-1))^{-1}.$$

Taigi vidurkis $\kappa_1 = 2\nu/(\nu-1)$ ir parinkus $\delta = 2/\kappa_1 = (\nu-1)/(\nu)$ (sutapatinus vidurkius), statistikos $\delta\tilde{Z}$ skirstinys yra χ^2 skirstinys su 2 laisvės laipsniais. Todėl asimptotinė P reikšmė $pv_a^{(1)}$ sutampa su tikslia P reikšme pv, surasta **5.2.1** pavyzdyje, ir lygi $0,0877$.

Analogiškai taikydami skleidimą (5.2.27) ir parinkę $\omega_1 = 0$, gausime $\delta = (\nu-1)/\nu$ ir koeficietai $\omega_2 = \omega_3 = ... = 0$ lygūs 0. Gauname, kad statistikos $\delta\tilde{Z}$ skirstinys yra χ^2 skirstinys su 2 laisvės laipsniais ir $pv_a^{(2)} = pv = 0,0877$.

Palyginę su **5.2.2** pavyzdžiu galime daryti išvadą, kad asimptotinį tikėtinumų santykio kriterijų tikslinga koreguoti, naudojant, pavyzdžiui, pateiktus patikslinimus.

5.3. Kovariacinės matricos lygybės žinomai matricai hipotezės tikrinimas

5.3.1. Kovariacinės matricos lygybės žinomai matricai hipotezė

Tarkime, kad $\mathbf{X}_1, ..., \mathbf{X}_n$ yra paprastoji imtis a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Reikia patikrinti hipotezę

$$H : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0, \tag{5.3.1}$$

kad kovariacinė matrica $\boldsymbol{\Sigma}$ yra lygi žinomai matricai $\boldsymbol{\Sigma}_0$. Tarsime, kad matrica $|\boldsymbol{\Sigma}_0| > 0$ teigiamai apibrėžta.

5.3.2. Tikėtinumų santykio kriterijus

Rasime tikėtinumų santykio statistiką hipotezei H tikrinti.

5.3.1 teorema. *Tarkim, kad $|\boldsymbol{\Sigma}| > 0, n > k$, ir žinoma matrica $|\boldsymbol{\Sigma}_0| > 0$ teigiamai apibrėžta. Tada tikėtinumų santykio statistika Λ hipotezei H tikrinti turi tokį pavidalą*

$$\Lambda = \left(\frac{e}{n}\right)^{nk/2} |\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}|^{n/2} \exp\{-\frac{1}{2}Tr(\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1})\}, \tag{5.3.2}$$

čia

$$\bar{\mathbf{X}} = \sum_{j=1}^{n} \mathbf{X}_j, \quad \boldsymbol{S} = \sum_{j=1}^{n} (\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})^T.$$

Įrodymas. Kadangi žinomoji matrica $\boldsymbol{\Sigma}_0$ simetrinė ir teigiamai apibrėžta, tai egzistuoja tokia neišsigimusi kvadratinė matrica $\boldsymbol{C}$ (1 priedas, (8.2.10)), kad

$$\boldsymbol{C}\boldsymbol{\Sigma}_0\boldsymbol{C}^T = \boldsymbol{I}. \tag{5.3.3}$$

Atlikime a. v. $\mathbf{X}_j$ tiesines transformacijas naudodami matricą $\boldsymbol{C}$:

$$\boldsymbol{Y}_j = \boldsymbol{C}\mathbf{X}_j, \quad j = 1, ..., n. \tag{5.3.4}$$

Tada vektoriai $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_n$ taip pat yra paprastoji imtis, gauta stebint normalųjį vektorių $\boldsymbol{Y} \sim N_k(\boldsymbol{\nu}, \boldsymbol{\Psi})$:

$$\mathbf{E}(\boldsymbol{Y}_j) = \boldsymbol{C}\boldsymbol{\mu} = \boldsymbol{\nu}, \quad \boldsymbol{V}(\boldsymbol{Y}_j) = \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}^T = \boldsymbol{\Psi}.$$

Atsitiktinių vektorių $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_n$ terminais hipotezė H ekvivalenti hipotezei $H^* : \boldsymbol{\Psi} = \boldsymbol{I}$, čia $\boldsymbol{I}$ vienetinė matrica.

Randame tikėtinumų santykio statistiką hipotezei H^* tikrinti:

$$\Lambda^* = \frac{\max_{\boldsymbol{\nu}, \boldsymbol{\Psi}=\boldsymbol{I}} L(\boldsymbol{\nu}, \boldsymbol{\Psi})}{\max_{\boldsymbol{\nu}, \boldsymbol{\Psi}} L(\boldsymbol{\nu}, \boldsymbol{\Psi})} = \frac{L(\tilde{\boldsymbol{\nu}}, \boldsymbol{I})}{L(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\Psi}})},$$

čia $\hat{\boldsymbol{\nu}}$ ir $\hat{\boldsymbol{\Psi}}$ yra parametrų $\boldsymbol{\nu}$ ir $\boldsymbol{\Psi}$ DT įvertiniai, o $\tilde{\boldsymbol{\nu}}$ ir yra parametro $\boldsymbol{\nu}$ DT įvertinys surastas tarus, kad hipotezė H^* teisinga.

Remiantis 1.1 skyreliu parametrų $\boldsymbol{\nu}$ ir $\boldsymbol{\Psi}$ didžiausiojo tikėtinumo įvertiniai yra

$$\hat{\boldsymbol{\nu}} = \bar{\boldsymbol{Y}} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{Y}_i, \quad \hat{\boldsymbol{\Psi}} = \frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{Y}_i - \bar{\boldsymbol{Y}})(\boldsymbol{Y}_i - \bar{\boldsymbol{Y}})^T = \frac{1}{n}\boldsymbol{S}^* = \frac{1}{n}[S^*_{ij}]_{k\times k},$$

o tikėtinumo funkcijos maksimumas

$$L(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\Psi}}) = (2\pi)^{-nk/2} n^{nk/2} |\boldsymbol{S}^*|^{-n/2} \exp\{-\frac{nk}{2}\}.$$

Kai hipotezė H^* teisinga, skaitiklyje turime tikėtinumo funkciją, gautą iš k nepriklausomų paprastųjų imčių, gautų stebint vienmačius normaliuosius dydžius su dispersijomis lygiomis 1. Taigi $\tilde{\boldsymbol{\nu}} = \hat{\boldsymbol{\nu}}$, o tikėtinumo funkcijos sąlyginis maksimumas

$$L(\tilde{\boldsymbol{\nu}}, \boldsymbol{I}) = (2\pi)^{-nk/2} \exp\{-\frac{1}{2}Tr(\boldsymbol{S}^*)\}.$$

Padalinę gauname, kad tikėtinumų santykis hipotezei H^* tikrinti yra

$$\Lambda^* = \left(\frac{e}{n}\right)^{nk/2} |\boldsymbol{S}^*|^{n/2} \exp\{-\frac{1}{2}Tr(\boldsymbol{S}^*)\}. \qquad (5.3.5)$$

Remdamiesi (5.3.3), (5.3.4) gauname

$$\boldsymbol{S}^* = \boldsymbol{C}\boldsymbol{S}\boldsymbol{C}^T, \quad \boldsymbol{\Sigma}_0 = (\boldsymbol{C}^T\boldsymbol{C})^{-1}, \quad \boldsymbol{\Sigma}_0^{-1} = \boldsymbol{C}\boldsymbol{C}^T,$$

todėl

$$|\boldsymbol{S}^*| = |\boldsymbol{C}\boldsymbol{S}\boldsymbol{C}^T| = |\boldsymbol{S}||\boldsymbol{C}\boldsymbol{C}^T| = |\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}|,$$
$$Tr(\boldsymbol{S}^*) = Tr(\boldsymbol{C}\boldsymbol{S}\boldsymbol{C}^T) = Tr(\boldsymbol{S}\boldsymbol{C}^T\boldsymbol{C}) = Tr(\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}). \qquad (5.3.6)$$

Įrašę šias išraiškas į (5.3.5), gauname (5.3.2). Statistika Λ^* nepriklauso nuo matricos $\boldsymbol{C}$ parinkimo ir sutampa su Λ. ▲

Tikrinamoji hipotezė H atmetama reikšmingumo lygmens α tikėtinumų santykio kriterijumi, kai teisinga nelygybė

$$\Lambda < \Lambda_{1-\alpha},$$

čia $\Lambda_{1-\alpha}$ yra statistikos Λ eilės $1-\alpha$ kritinė reikšmė. Kartais rekomenduojama sudarant statistiką naudoti nepaslinktuosius dispersijų įvertinius. Modifikuotoji tikėtinumų santykio statistika $\tilde{\Lambda}$ gaunama pakeičiant n į $\nu = n-1$:

$$\tilde{\Lambda} = \left(\frac{e}{\nu}\right)^{\nu k/2} |\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}|^{\nu/2} \exp\{-\frac{1}{2}Tr(\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1})\}, \qquad (5.3.7)$$

Hipotezė H atmetama modifikuotuoju reikšmingumo lygmens α tikėtinumų santykio kriterijumi, kai teisinga nelygybė

$$\tilde{\Lambda} < \tilde{\Lambda}_{1-\alpha}.$$

Norint rasti kritines reikšmes, reikia ištirti tikėtinumų santykio statistikos savybes, kai hipotezė teisinga.

5.3.3. Tikėtinumų santykio statistikos momentai

Ieškosime statistikos Λ momento $\mathbf{E}(\Lambda^h)$, kai hipotezė H yra teisinga.

5.3.2 teorema. *Jeigu hipotezė H teisinga ir $n > k$, tai momentas $\mathbf{E}(\Lambda^h)$ turi tokį pavidalą:*

$$\mathbf{E}(\Lambda^h) = \left(\frac{2e}{n}\right)^{\frac{nkh}{2}} \left(\frac{1}{1+h}\right)^{\frac{(n+nh-1)k}{2}} \prod_{j=1}^{k} \frac{\Gamma((n+nh-j)/2)}{\Gamma((n-j)/2)}. \tag{5.3.8}$$

Įrodymas. Remiantis 5.3.1 teoremos įrodymu statistikos Λ skirstinys sutampa su statistikos Λ^*, apibrėžtos (5.3.5) formule, skirstiniu. Todėl

$$\Lambda \overset{d}{\sim} \left(\frac{e}{n}\right)^{nk/2} |\boldsymbol{S}^*|^{n/2} \exp\{-\frac{1}{2}Tr(\boldsymbol{S}^*)\}, \tag{5.3.9}$$

čia matrica $\boldsymbol{S}^*$, kai hipotezė teisinga, turi Višarto skirstinį $\boldsymbol{S}^* \sim W_k(\nu, \boldsymbol{I})$. Remdamiesi (1.6.13) tankio išraiška gauname

$$\mathbf{E}(\Lambda^h) = \left(\frac{e}{n}\right)^{\frac{nkh}{2}} \int |\boldsymbol{S}^*|^{\frac{nh}{2}} \exp\{-\frac{1}{2}hTr(\boldsymbol{S}^*)\} \frac{|\boldsymbol{S}^*|^{\frac{n-k-2}{2}} \exp\{-\frac{1}{2}Tr(\boldsymbol{S}^*)\}}{K(k, n-1, \boldsymbol{I})} d\boldsymbol{S}^*$$

$$= \left(\frac{e}{n}\right)^{\frac{nkh}{2}} \frac{K(k, n-1+nh, [(1+h)\boldsymbol{I}]^{-1})}{K(k, n-1, \boldsymbol{I})} \int f(\boldsymbol{S}^*|k, n+nh-1, [(1+h)\boldsymbol{I}]^{-1}) d\boldsymbol{S}^*.$$

Kadangi po integralu yra Višarto skirstinio tankis, tai integralas lygus 1 ir, remdamiesi normuojančių konstantų iš (1.6.13) išraiškomis, gauname (5.3.8). ▲

5.3.4. Tikėtinumų santykio skirstinys

Momentus (5.3.8) galima traktuoti kaip a. d. $\ln\Lambda$ momentus generuojančiąją funkciją, kuri, kai $n > k$, egzistuoja h kitimo intervale $|h| < 1/2$, apimančiame tašką $h = 0$. Todėl momentai (5.3.8) visiškai nusako a. d. $\ln\Lambda$, o kartu ir a. d. Λ skirstinį.

5.3.3 teorema. *Jeigu hipotezė H teisinga, $n > k$, tai a. d. Λ skirstinys sutampa su skirstiniu polinomo nuo nepriklausomų atsitiktinių dydžių:*

$$\Lambda^{2/n} \overset{d}{\sim} \prod_{i=2}^{k} \xi_{i1} \prod_{j=1}^{k} [(\eta_j/n)\exp\{1-\eta_j/n\}]; \tag{5.3.10}$$

čia $\xi_{21}, ..., \xi_{k1}, \eta_1, ..., \eta_k$ yra nepriklausomi atsitiktiniai dydžiai. Atsitiktiniai dydžiai $\xi_{21}, ..., \xi_{k1}$ turi beta skirstinius:

$$\xi_{i1} \sim Be((n-i)/2, (i-1/2)), \quad i = 2, ..., k.$$

A. d. $\eta_1, ..., \eta_k$ yra nepriklausomi ir vienodai pasiskirstę

$$\eta_j \sim \chi^2(n-1), \quad j = 1, ..., k.$$

Įrodymas. Tikėtinumų santykio Λ skirstinys sutampa su Λ^* skirstiniu. Tikėtinumų santykį Λ^* galima išreikšti dviejų tikėtinumų santykių sandauga

$$\Lambda^* = \Lambda_1^* \cdot \Lambda_2^* = \left(\frac{\max_{\boldsymbol{\nu},\boldsymbol{\Psi}=\Delta} L(\boldsymbol{\nu},\boldsymbol{\Psi})}{\max_{\boldsymbol{\nu},\boldsymbol{\Psi}} L(\boldsymbol{\nu},\boldsymbol{\Psi})}\right) \cdot \left(\frac{\max_{\boldsymbol{\nu},\boldsymbol{\Psi}=\boldsymbol{I}} L(\boldsymbol{\nu},\boldsymbol{\Psi})}{\max_{\boldsymbol{\nu},\boldsymbol{\Psi}=\Delta} L(\boldsymbol{\nu},\boldsymbol{\Psi})}\right), \qquad (5.3.11)$$

čia Δ yra diagonali, o $\boldsymbol{I}$ – vienetinė matricos. Taigi hipotezę H^* galima tikrinti kaip sudėtinę hipotezę. Visų pirma tikriname hipotezę, kad vektoriaus $\boldsymbol{Y}$ koordinatės yra nepriklausomos. Šiai hipotezei tikėtinumų santykis yra Λ_1^*. Paskui tikriname hipotezę, kad vektoriaus $\boldsymbol{Y}$ koordinačių dispersijos yra lygios 1, jei žinoma, kad jos yra nepriklausomos. Tokiai hipotezei tikrinti tikėtinumų santykis yra Λ_2^*.

Pirmoji hipotezė yra atskiras hipotezės, nagrinėtos skyrelyje 4.5, atvejis. Formulėje (4.5.3) reikia imti $m = k, k_1 = k_2 = ... = k_m = 1$. Suradę antrosios hipotezės tikėtinumų santykį gauname

$$\Lambda^* = \Lambda_1^* \cdot \Lambda_2^* = \left(\frac{|\boldsymbol{S}^*|}{S_{11}^* \cdot ... \cdot S_{kk}^*}\right)^{n/2} \cdot \left(\left(\frac{e}{n}\right)^{\frac{nk}{2}} \frac{(S_{11}^* \cdot ... \cdot S_{kk}^*)^{n/2}}{\exp\{\frac{1}{2}Tr(\boldsymbol{S}^*)\}}\right). \qquad (5.3.12)$$

Pirmasis daugiklis išreiškiamas empiriniais koreliacijos koeficientais $r_{ij}^* = S_{ij}^* \sqrt{S_{ii}^* S_{jj}^*}$, o antrasis – diagonaliniais elementais S_{ii}^*. Kai stebimo vektoriaus koordinatės nepriklausomos, tai vektorius, sudarytas iš stebėjimus atitinkančios Višarto matricos diagonalinių elementų, ir vektorius, sudarytas iš koreliacijos koeficientų, yra nepriklausomi (žr. 4.4 pratimą). Todėl tikėtinumų santykiai Λ_1^* ir Λ_2^* yra nepriklausomi.

4.5 skyrelio (4.5.8) formulėje buvo gauta

$$(\Lambda_1^*)^{2/n} \overset{d}{\sim} \prod_{i=2}^{k} \xi_{i1}, \quad \xi_{i1} \sim Be\left(\frac{n-i}{2}, \frac{i-1}{2}\right), \quad i = 2, ..., k. \qquad (5.3.13)$$

Kadangi esant teisingai nepriklausomumo hipotezei ir vienetinėms dispersijoms, a. d. $S_{11}^*, ..., S_{kk}^*$ yra nepriklausomi ir vienodai pasiskirstę $S_{jj}^* \sim \chi^2(n-1)$, tai

$$(\Lambda^*)^{2/n} = \prod_{j=1}^{k} \left(\frac{eS_{jj}^*}{n \exp\{S_{jj}^*/n\}}\right) = \prod_{j=1}^{k} [(\eta_j/n) \exp\{1 - \eta_j/n\}]. \qquad (5.3.14)$$

Sujungę (5.3.13) ir (5.3.14), gauname (5.3.10). ▲

Formulę (5.3.10) galima naudoti ieškant a. d. Λ kritinių reikšmių skaitinio modeliavimo metodais.

5.3.5. Tikėtinumų santykio statistikos asimptotinis skirstinys

Jeigu imties dydis $n \to \infty$, tai remiantis tikėtinumų santykio asimptotinėmis savybėmis (1 dalis, 4.5.4 skyrelis)

$$-2 \ln \Lambda \xrightarrow{d} \chi_f^2, \qquad (5.3.15)$$

čia laisvės laipsnių skaičius

$$f = k(k+1)/2, \tag{5.3.16}$$

nes iš pradžių buvo $k + k(k+1)/2$ nežinomas parametras (vidurkių vektoriaus ir kovariacijų matricos elementai), o esant teisingai hipotezei lieka k nežinomų parametrų (vidurkių vektoriaus elementai), t. y. hipotezė H uždeda parametrams f apribojimų. Pagal asimptotinį tikėtinumų santykio reikšmingumo lygmens α kriterijų hipotezė atmetama, kai

$$Z = -2\ln\Lambda > \chi^2_\alpha(f), \tag{5.3.17}$$

arba P reikšmių terminais, kai

$$pv_a = \mathbf{P}\{\chi^2_f > z\} < \alpha, \tag{5.3.18}$$

čia z yra statistikos Z realizacija.

Naudojant modifikuotąją tikėtinumų santykio statistiką $\tilde{\Lambda}$, kuri gaunama imant nepaslinktuosius kovariacinių matricų įvertinius, asimptotinis skirstinys išliks nepakitęs:

$$-2\ln\tilde{\Lambda} \xrightarrow{d} \chi^2_f. \tag{5.3.19}$$

Pagal asimptotinį modifikuotą tikėtinumų santykio reikšmingumo lygmens α kriterijų hipotezė H atmetama, kai

$$\tilde{Z} = -2\ln\tilde{\Lambda} > \chi^2_\alpha(f), \tag{5.3.20}$$

arba P reikšmių terminais, kai

$$\tilde{pv}_a = \mathbf{P}\{\chi^2_f > \tilde{z}\} < \alpha, \tag{5.3.21}$$

čia $\tilde{z}$ yra statistikos $\tilde{Z}$ realizacija.

5.3.1 pavyzdys. (**2.2.1** pavyzdžio tęsinys). Tarkime, kad **2.2.1** lentelėje pateiktus duomenis galima interpretuoti kaip didumo $n = 24$ paprastosios imties, gautos stebint trimatį normalųjį vektorių $\mathbf{X} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, realizaciją. Reikia patikrinti hipotezę

$$H : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0 = 10^{-2}\begin{pmatrix} 0,8 & 0,4 & 0,8 \\ 0,4 & 1,4 & 1,2 \\ 0,8 & 1,2 & 1,8 \end{pmatrix},$$

kad kovariacinė matrica $\boldsymbol{\Sigma}$ lygi žinomai matricai $\boldsymbol{\Sigma}_0$.

Matricos $\boldsymbol{S}$ realizacija rasta **2.2.1** pavyzdyje. Randame

$$\boldsymbol{\Sigma}_0^{-1} = 100\begin{pmatrix} 2,4107 & 0,5357 & -1,4286 \\ 0,5357 & 1,7857 & -1,4286 \\ -1,4286 & -1,4286 & 2,1429 \end{pmatrix},$$

$$\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1} = 100\begin{pmatrix} 0,71990,3955 - 0,5264 \\ 0,47960,5671 - 0,5007 \\ -0,0929 - 0,08230,2629 \end{pmatrix}.$$

Apskaičiuojame statistikų $|\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}|$, $Tr(\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1})$ ir $Z = -2\ln\Lambda$ realizacijas

$$|\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}| = 100^2 3,9224, \quad Tr(\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}) = 154,991, \quad z = 57,9619,$$

ir randame asimptotinio tikėtinumų santykio kriterijaus P reikšmę

$$pv_a = \mathbf{P}\{\chi_f^2 > z\} = \mathbf{P}\{\chi_6^2 > 57,9619\} = 1,2 \cdot 10^{-10}.$$

Analogiškai pagal modifikuotąjį asimptotinį tikėtinumų santykio kriterijų (naudojami nepasliktieji kovariacinės matricos įvertiniai, t. y. n keičiama į $\nu = n-1$) randame

$$\tilde{z} = 59,0682, \quad \tilde{pv}_a = \mathbf{P}\{\chi_6^2 > 59,0682\} = 7 \cdot 10^{-11}.$$

Matome, kad modifikuotas kriterijus šiame pavyzdyje šiek tiek skiriasi.

5.3.6. Asimptotinio skirstinio patikslinimai

Nagrinėsime statistikos $\tilde{Z} = -2\ln\tilde{\Lambda}$ asimptotinio skirstinio patikslinimus. Turime

$$\tilde{\Lambda} = \left(\frac{e}{n}\right)^{nk/2} |\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}|^{n/2} \exp\{-\frac{1}{2}Tr(\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1})\} \overset{d}{\sim}$$

$$\tilde{\Lambda}_1^*\tilde{\Lambda}_2^* = \left(\frac{|\boldsymbol{S}^*|}{S_{11}^* \cdot ... \cdot S_{kk}^*}\right)^{\nu/2} \cdot \left(\left(\frac{e}{\nu}\right)^{\frac{\nu k}{2}} \frac{(S_{11}^* \cdot ... \cdot S_{kk}^*)^{\nu/2}}{\exp\{\frac{1}{2}Tr(\boldsymbol{S}^*)\}}\right).$$

Remdamiesi 4.5.3 skyrelio (4.5.6) formule (imdami $m = k, k_1 = ... = k_m = 1$) gauname

$$\mathbf{E}\left(\tilde{\Lambda}_1^*\right)^{\frac{2h}{\nu}} = \mathbf{E}\left(\frac{|\boldsymbol{S}^*|}{S_{11}^* \cdot ... \cdot S_{kk}^*}\right)^{\nu/2} = \prod_{j=2}^{k}\left(\frac{\Gamma\left(\frac{\nu+1-j}{2}+h\right)\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu+1-j}{2}\right)\Gamma\left(\frac{\nu}{2}+h\right)}\right). \tag{5.3.22}$$

Antrasis daugiklis

$$\left(\tilde{\Lambda}_2^*\right)^{\frac{2}{\nu}} = \prod_{j=1}^{k}\left(\frac{S_{jj}^*}{\nu}\exp\left\{1 - \frac{S_{jj}^*}{\nu}\right\}\right). \tag{5.3.23}$$

Kadangi a. d. $S_{11}^*, ..., S_{kk}^*$ yra nepriklausomi ir vienodai pasiskirstę, tai

$$\mathbf{E}\left(\tilde{\Lambda}_2^*\right)^{\frac{2h}{\nu}} = \prod_{j=1}^{k}\left[\mathbf{E}\left(\frac{S_{jj}^*}{\nu}\exp\left\{1 - \frac{S_{jj}^*}{\nu}\right\}\right)^h\right]^k =$$

$$\left[\left(\frac{2e}{\nu}\right)^h \frac{1}{(1+2h/\nu)^{\nu/2+h}} \frac{\Gamma\left(\frac{\nu}{2}+h\right)}{\Gamma\left(\frac{\nu}{2}\right)}\right]^k. \tag{5.3.24}$$

Lygybėse (5.3.22) ir (5.3.24) vietoje h imdami $(-it\nu)$ ir sudauginę gausime a. d. $\tilde{Z} = -2\ln\tilde{\Lambda}$ charakteristinę funkciją

$$\psi(t) = \psi_1(t)[\tilde{\psi}(t)]^k, \tag{5.3.25}$$

čia

$$\psi_1(t) = \prod_{j=2}^{k}\left(\frac{\Gamma\left(\frac{\nu+1-j}{2}-it\nu\right)\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu+1-j}{2}\right)\Gamma\left(\frac{\nu}{2}-it\nu\right)}\right),$$

$$\tilde{\psi}(t) = \left[\left(\frac{2e}{\nu}\right)^{-it\nu} \frac{1}{(1-2it)^{\nu/2-it\nu}} \frac{\Gamma\left(\frac{\nu}{2}-it\nu\right)}{\Gamma\left(\frac{\nu}{2}\right)}\right].$$

Remdamiesi 4.5.3 skyreliu galime tvirtinti, kad $\nu \to \infty$ charakteristinė funkcija $\psi_1(t)$ artėja į $(1-2it)^{-f_1/2}$, $f_1 = k(k-1)/2$, todėl charakteristinė funkcija $\tilde{\psi}(t)$ turėtų artėti į $(1-2it)^{-1/2}$, t. y. į χ^2 skirstinio su 1 laisvės laipsniu charakteristinę funkciją. Tuo galime įsitikinti ir tiesiogiai. Remiantis (5.3.23)

$$-2\ln\tilde{\Lambda}_2^* = \sum_{j=1}^{k}\left[(\nu - S_{jj}^*) - \nu\ln\frac{S_{jj}^*}{\nu}\right] = \sum_{j=1}^{k}\left[(\nu - S_{jj}^*) - \nu\ln\left(1+\frac{S_{jj}^*-\nu}{\nu}\right)\right] =$$

$$\sum_{j=1}^{k}\left[\frac{(S_{jj}^*-\nu)^2}{2\nu} + O_P\left(\frac{1}{\sqrt{\nu}}\right)\right]. \qquad (5.3.26)$$

Kadangi $(S_{jj}^*-\nu)/\sqrt{2\nu} \xrightarrow{d} Y \sim N(0,\ 1)$, tai $-2\ln\tilde{\Lambda}_2^* \xrightarrow{d} \chi_k^2$, kai $\nu \to \infty$.

Skleisdami $\ln\psi(t)$ Teiloro eilute (it) laipsniais, gausime

$$\psi(t) = \sum_s (\kappa_s + k\tilde{\kappa}_s)\frac{(it)^s}{s!},$$

čia κ_s yra semiinvariantai, atitinkantys funkciją $\ln\psi_1(t)$ (jų išraiškos pateiktos 4.5.3 skyrelyje); $\tilde{\kappa}_s$ – semiinvariantai atitinkantys funkciją $\ln\tilde{\psi}(t)$. Pirmieji du semiinvariantai yra

$$\kappa_1 = \nu\{k\ln(\nu/2) - \varphi_1(\nu/2) - \sum_{j=2}^{k}\varphi_1((\nu+1-j)/2)\},$$

$$\kappa_2 = \nu^2\{k\varphi_2(\nu/2) + \sum_{j=2}^{k}\varphi_2((\nu+1-j)/2)\} - 2k\nu.$$

Jeigu vietoje statistikos $\tilde{Z} = -2\ln\tilde{\Lambda}$ imsime statistiką $Z^* = \delta\tilde{Z}$, kai $\delta = f/\kappa_1$, tai a. d. Z^* vidurkis $\mathbf{E}Z^* = f$ tiksliai sutampa su asimptotinio skirstinio (5.3.19) vidurkiu f. Aproksimuodami statistikos Z^* skirstinį χ^2 skirstiniu su f laisvės laipsniais, gauname patikslintą tikėtinumų santykio kriterijų (sutapatinami vidurkiai): hipotezė atmetama patikslintu asimptotiniu tikėtinumų santykio reikšmingumo lygmens α kriterijumi, kai

$$pv_a^{(1)} = \mathbf{P}\{\chi_f^2 > \delta z^*\} < \alpha, \qquad (5.3.27)$$

čia z^* yra statistikos Z^* realizacija.

Parinkdami du parametrus δ ir f iš lygčių

$$\mathbf{E}(\delta\tilde{Z}) = \delta\kappa_1 = f,$$

$$\boldsymbol{V}(\delta\tilde{Z}) = \delta^2\kappa_2 = 2f,$$

ir aproksimuodami statistikos Z^* skirstinį χ^2 skirstiniu su f laisvės laipsniais, gauname patikslintą tikėtinumų santykio kriterijų (sutapatinami du momentai): hipotezė atmetama patikslintu asimptotiniu tikėtinumų santykio reikšmingumo lygmens α kriterijumi, kai

$$pv_a^{(2)} = \mathbf{P}\{\chi_f^2 > \delta z^*\} < \alpha, \tag{5.3.28}$$

čia z^* yra statistikos Z^* realizacija. Reikia pažymėti, kad šioje aproksimacijoje f nebūtinai sveikasis skaičius. Todėl randant $pv_a^{(2)}$ reikia prisiminti χ^2 skirstinio ir gama skirstinio sąryšį.

Kitas patikslinimas, kaip ir pirmesniuose skyreliuose, gali būti gautas skleidžiant charakteristinę funkciją $(1-2it)$ laipsniais. Skyrelyje 4.5.7 gautas toks charakteristinės funkcijos $\psi_1(\delta t)$ skleidimas (reikia imti $m = k, k_1 = k_2 = ... = k_m = 1$):

$$\psi_1(\delta t) = \left(\frac{1}{1-2it}\right)^{f_1/2}\left\{1 + \frac{\omega_2}{(n\delta)^2}\left[\frac{1}{(1-2it)^2} - 1\right]\right\} + O\left(\frac{1}{n^3}\right), \tag{5.3.29}$$

čia

$$\delta = \frac{6n-2k-11}{6n}, \quad \omega_2 = \frac{2f_1(4f_1-13)}{216}, \quad f_1 = \frac{k(k-1)}{2}.$$

Vietoje statistikos $-2\ln\Lambda$ imdami statistiką

$$U = -2\delta\ln\Lambda_1^* + T, \quad T = \sum_{j=1}^{k}\left[\frac{(S_{jj}^* - \nu)^2}{2\nu}\right], \tag{5.3.30}$$

gausime, kad jos charakteristinė funkcija apytiksliai lygi sandaugai (5.3.28) ir $(1-2it)^{k/2}$. Naudodami pirmąjį skleidinio narį gauname kriterijų: hipotezė H atmetama asimptotiniu reikšmingumo lygmens α kriterijumi, kai

$$pv_a^{(3)} = \mathbf{P}\{\chi^2_{f_1+k\nu} > u\} < \alpha, \tag{5.3.31}$$

čia u yra statistikos U realizacija.

Imdami ir narį su ω_2 gauname kriterijų: hipotezė H atmetama asimptotiniu reikšmingumo lygmens α kriterijumi, kai

$$pv_a^{(4)} = \mathbf{P}\{\chi^2_{f_1+k\nu} > u\} - \frac{\omega_2}{(n\delta)^2}\{\mathbf{P}\{\chi^2_{f_1+k\nu+4} > u\} - \mathbf{P}\{\chi^2_{f_1+k\nu} > u\}\} < \alpha. \tag{5.3.32}$$

5.3.2 pavyzdys. (**5.3.1** pavyzdžio tęsinys). Spręsime uždavinį iš pavyzdžio **5.3.1** naudodami pateiktus aproksimacijos patikslinimus.

Statistikos $\tilde{Z}$ realizacija $\tilde{z}$ gauta **5.3.1** pavyzdyje $\tilde{z} = 12,3868$. Naudodami SAS paketą randame

$$\kappa_1 = 6,23662, \quad \delta = f/\kappa_1 = 0,96206, \quad \delta\tilde{z} = 11,91684,$$

ir asimptotinę patikslintą P reikšmę

$$pv_a^{(1)} = \mathbf{P}\{\chi_f^2 > \delta\tilde{z}\} = \mathbf{P}\{\chi_6^2 > 11,91684\} = 0,06385.$$

Sutapatindami du momentus papildomai apskaičiuojame

$$\kappa_2 = 12,97364, \quad \delta = 2\kappa_1/\kappa_2 = 0,96143, \quad f = 2\kappa_1^2/\kappa_2 = 5,99607,$$

ir randame patikslintą asimptotinę P reikšmę

$$pv_a^{(2)} = \mathbf{P}\{U > \delta\tilde{z}\} = \mathbf{P}\{U > 11,90904\} = 0,06390,$$

čia $U \sim G(1/2, f/2)$ turi gama skirstinį su parametrais $1/2$ ir $f/2$.

Randame $\delta = 0,9056$ ir $\omega_2 = -0,0278$ ir naudodamiesi **5.3.1** pavyzdžiu apskaičiuojame statistikos U realizaciją

$$u = -2\ln\frac{|\boldsymbol{S\Sigma}_0^{-1}|}{S_{11}^* \cdot ... \cdot S_{kk}^*} + T = 11,6295.$$

Asimptotinė P reikšmė

$$pv_a^{(3)} = \mathbf{P}\{\chi_6^2 > u\} = 0,0708.$$

Asimptotinė P reikšmė $pv_a^{(4)}$ keturių ženklų po kablelio tikslumu sutampa su $pv_a^{(3)}$.

Palyginę **5.3.1** ir **5.3.2** pavyzdžius matome, kad, kai n nedideli, tikėtinumų santykio kriterijus gali būti netikslus. Rekomenduojama jį patikslinti bent jau sutapatinant vidurkius. Tolesni patikslinimai yra mažiau reikšmingi.

5.4. Pratimai

5.1. Tarkime, $X_{i1}, ..., X_{in_i}, i = 1, ..., m$, yra m paprastųjų imčių, gautų stebint nepriklausomus a. d. $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, ..., m$. a) Raskite tikėtinumų santykio statistiką atsižvelgdami į Bartleto pataisą hipotezei $H : \sigma_1^2 = ... = \sigma_m^2$ tikrinti. b) Sudarykite asimptotinį hipotezės H tikrinimo kriterijų.

5.2. Turime tris vienodo didumo n imtis, gautas stebint nepriklausomus a. v. $\mathbf{X}_i \sim N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$. Sudarykite tikėtinumų santykio statistiką Λ_1 hipotezei $H_1 : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ ir tikėtinumų santykio statistiką Λ_2 hipotezei $H_1 : \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_3$ tikrinti, kai žinoma, kad $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$. Įrodykite, kad a. d. Λ_1 ir Λ_2 yra nepriklausomi, kai $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_3$.

5.3. Tegu $\boldsymbol{Y}^{(j)}, j = 1, ..., m$, yra nepriklausomi k-mačiai vektoriai su nuliniais vidurkiais ir kovariacijų matricomis $\boldsymbol{\Sigma}_j$. Tarkim, $\boldsymbol{C} = [c_{ij}]_{m\times m}$ ortogonali matrica, kurios paskutinioji eilutė yra $(1/\sqrt{m}, ..., 1/\sqrt{m})$. Atlikime transformaciją

$$\boldsymbol{Z}^{(j)} = \sum_{i=1}^{m} c_{ji}\boldsymbol{Y}^{(i)}, \quad j = 1, ..., m.$$

Įrodykite, kad

$$\mathbf{Cov}\,(\boldsymbol{Z}^{(m)}, \boldsymbol{Z}^{(j)}) = \mathbf{0}, \quad j = 1, ..., m-1,$$

tada ir tik tada, kai $\boldsymbol{\Sigma}_1 = ... = \boldsymbol{\Sigma}_m$.

5.4. (**5.3** pratimo tęsinys.) Remiantis **5.3** pratimu raskite hipotezės $H : \boldsymbol{\Sigma}_1 = ... = \boldsymbol{\Sigma}_m$ tikrinimo kriterijų, kaip kriterijų dėl a. v. $\boldsymbol{Z}^{(m)}$ nepriklausomumo nuo a. v. $\boldsymbol{Z}^{(1)}, ..., \boldsymbol{Z}^{(m-1)}$ remiantis jungtinio vektoriaus $\boldsymbol{U} = ((\boldsymbol{Z}^{(1)})^T, ..., (\boldsymbol{Z}^{(m-1)})^T)^T$ paprastąja imtimi $\boldsymbol{U}_1, ..., \boldsymbol{U}_n$. Raskite tikėtinumų statistikos skirstinį, kai $k = 2$.

5.5. Tegu $\mathbf{X}_1, ..., \mathbf{X}_n$ yra paprastoji imtis a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Nagrinėkime hipotezes $H : \boldsymbol{\mu} = \boldsymbol{\mu}_0, \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$, $H_1 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ ir $H_2 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$, kai žinoma, kad $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$; čia $\boldsymbol{\Sigma}_0$ žinoma, $|\boldsymbol{\Sigma}_0| > 0$. Tegu Λ, Λ_1 ir Λ_2 yra šias hipotezes atitinkančios tikėtinumų santykio statistikos.

a) Raskite statistikos Λ išraišką ir asimptotinį skirstinį, kai H teisinga.

b) Įrodykite, kad $\Lambda = \Lambda_1\Lambda_2$.

c) Įrodykite, kad Λ_1 ir Λ_2 nepriklausomi, kai hipotezė H teisinga.

5.6. Tegu $\mathbf{X}_1, ..., \mathbf{X}_n$ yra paprastoji imtis a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Nagrinėkime hipotezes $H : \boldsymbol{\mu} = \mathbf{0}, \boldsymbol{\Sigma} = \sigma^2\boldsymbol{\Sigma}_0$, $H_1 : \boldsymbol{\Sigma} = \sigma^2\boldsymbol{\Sigma}_0$ ir $H_2 : \boldsymbol{\mu} = \mathbf{0}$, kai žinoma, kad $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{\Sigma}_0$; čia

$\mathbf{\Sigma}_0$ žinoma, o σ – nežinomas. Tegu Λ, Λ_1 ir Λ_2 yra šias hipotezes atitinkančios tikėtinumų santykio statistikos.

a) Raskite statistikos Λ_2 skirstinį, kai H_2 teisinga.

b) Raskite statistikos Λ išraišką ir asimptotinį skirstinį, kai H teisinga.

5.7. Pagal 8 vienodo didumo $n = 20$ paprastąsias imtis, gautas stebint nepriklausomus a. d. $X_i \sim N(\mu_i, \sigma_i^2)$ gauti nepaslinktieji dispersijų įverčiai: 12,4; 6,8; 17,1; 10,4; 12,0; 15,3; 5,1; 8,9. Patikrinti hipotezę $H : \sigma_1^2 = ... = \sigma_8^2$.

5.8. Buvo matuojamas aliuminio lydinių tempiamasis atsparumas (X_1) ir jų tvirtumas (X_2). Gauta po 12 matavimų 5 nepriklausomose imtyse. Pagal matavimo duomenis apskaičiuotos matricos [2]

$$\boldsymbol{S}_1 = \begin{pmatrix} 78,948 & 214,18 \\ 214,18 & 1247,18 \end{pmatrix}, \quad \boldsymbol{S}_2 = \begin{pmatrix} 223,695 & 657,62 \\ 657,62 & 2519,31 \end{pmatrix},$$

$$\boldsymbol{S}_3 = \begin{pmatrix} 57,448 & 190,63 \\ 190,63 & 1241,78 \end{pmatrix}, \quad \boldsymbol{S}_4 = \begin{pmatrix} 187,618 & 375,91 \\ 375,91 & 1473,44 \end{pmatrix},$$

$$\boldsymbol{S}_5 = \begin{pmatrix} 88,456 & 259,18 \\ 259,18 & 1171,73 \end{pmatrix}.$$

Patikrinkite hipotezę apie penkių kovariacinių matricų lygybę.

5.9. Tegu X_1 ir X_2 yra taurėlapio ilgis ir plotis, o X_3 ir X_4 – vainiklapio ilgis ir plotis. Atlikta po 50 matavimų *Iris versicolor*, *Iris setosa* ir *Iris virginica*. Pagal gautus duomenis apskaičiuotos matricos [2]:

$$\boldsymbol{S}_1 = \begin{pmatrix} 13,0552 & 4,1740 & 8,9620 & 2,7332 \\ 4,1740 & 4,8250 & 4,0500 & 2,0190 \\ 8,9620 & 4,0500 & 10,8200 & 3,5820 \\ 2,7332 & 2,0190 & 3,5820 & 1,9162 \end{pmatrix},$$

$$\boldsymbol{S}_2 = \begin{pmatrix} 6,0882 & 4,8616 & 0,8014 & 0,5062 \\ 4,8616 & 7,0408 & 0,5732 & 0,4556 \\ 0,8014 & 0,5732 & 1,4778 & 0,2974 \\ 0,5062 & 0,4556 & 0,2974 & 0,5442 \end{pmatrix},$$

$$\boldsymbol{S}_3 = \begin{pmatrix} 19,8128 & 4,5944 & 14,8612 & 2,4056 \\ 4,5944 & 5,0962 & 3,4976 & 2,3338 \\ 14,8612 & 3,4976 & 14,9248 & 2,3924 \\ 2,4056 & 2,3338 & 2,3924 & 3,6962 \end{pmatrix}.$$

a) Patikrinkite hipotezę $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$.

b) Patikrinkite hipotezę $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}_3$.

5.10. Pagal **1.9** pratimo duomenis patikrinkite hipotezę, kad atveju a) ir atveju b) kovariacinės matricos yra lygios.

5.11. (**2.7** pratimo tęsinys). Pagal pratimo **2.7** duomenis patikrinkite hipotezę

$$H : \mathbf{\Sigma} = \sigma^2 \mathbf{\Sigma}_0 = \sigma^2 \begin{pmatrix} 6 & 8 \\ 8 & 15 \end{pmatrix}.$$

5.12. Pagal pratimo **2.9** duomenis patikrinkite hipotezę, kad a. v. $(X_1, X_2, X_3)^T$ kovariacinė matrica proporcinga matricai

$$\mathbf{\Sigma}_0 = \begin{pmatrix} 10 & 5 & 2 \\ 5 & 5 & 2 \\ 2 & 2 & 5 \end{pmatrix}.$$

5.13. Pagal pratimo **1.13** duomenis patikrinkite hipotezę, kad a. v. $(X_1, X_2, X_3, X_4)^T$ kovariacinė matrica $\mathbf{\Sigma}$ lygi žinomai matricai

$$\mathbf{\Sigma}_0 = \begin{pmatrix} 0,1 & -0,5 & 0 & 0 \\ -0,5 & 25 & 0,5 & 0,5 \\ 0 & 0,5 & 1 & 0 \\ 0 & 0,5 & 0 & 1 \end{pmatrix}.$$

5.14. Pagal pratimo **1.13** duomenis patikrinkite hipotezę, kad a. v. $(X_1, X_2, X_3)^T$ kovariacinė matrica $\mathbf{\Sigma}$ lygi žinomai matricai

$$\mathbf{\Sigma}_0 = \begin{pmatrix} 3 & 10 & 0 \\ 10 & 200 & -6 \\ 0 & -6 & 3 \end{pmatrix}.$$

5.15. Pagal pratimo **2.18** duomenis patikrinkite hipotezę $H : \mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$.

5.16. Pagal pratimo **3.14** duomenis patikrinkite hipotezę $H : \mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}_3$.

5.17. Pagal pratimo **4.24** duomenis patikrinkite hipotezę, kad a. v. $(X_1, X_2, X_3)^T$ sąlyginio skirstinio, kai a. v. $(X_4, X_5, X_6)^T$ fiksuotas, kovariacinė matrica $\mathbf{\Sigma}_{11\cdot 2}$ proporcinga matricai

$$\mathbf{\Sigma}_0 = \begin{pmatrix} 34 & 0 & 10 \\ 0 & 6 & 0 \\ 10 & 0 & 1 \end{pmatrix}.$$

Atsakymai ir nurodymai.

5.1. a) Tikėtinumų santykio statistika

$$\Lambda = \frac{\nu^{\nu/2}}{\nu_1^{\nu_1/2} \cdot \ldots \cdot \nu_m^{\nu_m/2}} \frac{S_1^{\nu/2} \ldots \cdot S_m^{\nu_m/2}}{S^{\nu/2}},$$

čia $S_i = \sum_j (X_{ij} - X_{i.})^2, \nu_i = n_1 - 1, i = 1, ..., m;\ S = S_1 + ... + S_m, \nu = \nu_1 + ... + \nu_m - m$.
b) Hipotezė H atmetama asimptotiniu reikšmingumo lygmens α kriterijumi, kai

$$Z = -2\ln\Lambda = \nu \ln S - \nu_1 \ln S_1 - ... - \nu_m \ln S_m - \nu \ln \nu + \nu_1 \ln \nu_1 + ... + \nu_m \ln \nu_m < \alpha,$$

arba, P reikšmių terminais, kai $pv_a = \mathbf{P}\{\chi^2_{m-1} > Z\} < \alpha$, čia Z yra statistikos realizacija.
5.2. Tikėtinumų santykio statistikos

$$\Lambda_1 = \frac{|\boldsymbol{S}_1|^{nk/2} |\boldsymbol{S}_2|^{nk/2} 2^{nk}}{|\boldsymbol{S}_1 + \boldsymbol{S}_2|^{nk}}, \quad \Lambda_2 = \frac{|\boldsymbol{S}_1 + \boldsymbol{S}_2|^{nk} |\boldsymbol{S}_3|^{nk/2} 3^{3nk/2}}{|\boldsymbol{S}_1 + \boldsymbol{S}_2 + \boldsymbol{S}_3|^{3nk/2}},$$

čia $\boldsymbol{S}_1, \boldsymbol{S}_2, \boldsymbol{S}_3$ yra Višarto matricos, sudarytos iš nepriklausomų imčių. Siekiant įsitikinti Λ_1 ir Λ_2 nepriklausomumu pakanka patikrinti, kad $\mathbf{E}(\Lambda_1\Lambda_2)^h = \mathbf{E}(\Lambda_1^h)\mathbf{E}(\Lambda_2^h)$.
5.4. Tegu $\boldsymbol{S} = \sum_j \boldsymbol{U}_j \boldsymbol{U}_j^T = [S_{ij}]_{mk\times mk}$, o $\boldsymbol{S}_{11} = [S_{ij}]_{k(m-1)\times k(m-1)}$, kai $i, j = 1, ..., k(m-1)$; $\boldsymbol{S}_{22} = [S_{ij}]_{k\times k}$, kai $i, j = km - k + 1, ..., km$. Tikėtinumų santykio statistika $\Lambda = (|\boldsymbol{S}|/(\boldsymbol{S}_{11}\boldsymbol{S}_{22}))^{n/2}$. Kai hipotezė teisinga, statistika $Z = -2\ln\Lambda$ asimptotiškai turi χ^2 skirstinį su $\nu = k^2 m$ laisvės laipsnių. Jeigu $k = 2$, tai remdamiesi skyreliu 4.5.4 gauname, kad statistika $U = \Lambda^{2/n} \overset{d}{\sim} \eta^2_{21}, \eta_{21} \sim Be(n-2m-4, 2(m-1))$ ir $F = (1-\sqrt{U})(n-2m-4)/(2(m-1)\sqrt{U}) \sim F(4(m-1), 2(n-2m-4))$. **5.5.** a) Tikėtinumų santykio statistika

$$\Lambda = \left(\frac{e}{n}\right)^{nk/2} |\boldsymbol{S}\mathbf{\Sigma}_0^{-1}|^{n/2} \exp\{-\frac{1}{2}[Tr(\boldsymbol{S}\mathbf{\Sigma}_0^{-1}) + n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T \mathbf{\Sigma}_0^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)]\},$$

čia $\boldsymbol{S} = \sum_i (\bar{\mathbf{X}} - \boldsymbol{\mu}_0)(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T$. b)

$$\Lambda_1 = \left(\frac{e}{n}\right)^{nk/2} |\boldsymbol{S}\mathbf{\Sigma}_0^{-1}|^{n/2} \exp\{-\frac{1}{2}Tr(\boldsymbol{S}\mathbf{\Sigma}_0^{-1})\}, \Lambda_2 = \exp\{-\frac{1}{2}n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T \mathbf{\Sigma}_0^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)\}.$$

c)Λ_1 priklauso tik nuo $\boldsymbol{S}$, Λ_2 – tik nuo $\bar{\mathbf{X}}$, o $\bar{\mathbf{X}}$ ir $\boldsymbol{S}$ yra nepriklausomi. **5.6.** a)

$$\Lambda_1 = (Tr(\boldsymbol{S}\mathbf{\Sigma}_0^{-1})/[Tr(\boldsymbol{S}\mathbf{\Sigma}_0^{-1} + n\bar{\mathbf{X}}^T \mathbf{\Sigma}_0^{-1}\bar{\mathbf{X}})^{nk/2}.$$

Jeigu $U = \Lambda_2^{2/(nk)}$, tai statistika $F = (1 - U)k(n-1)/(kU) \sim F(k, k(n-1))$. b)

$$\Lambda = \Lambda_1\Lambda_2 = |\boldsymbol{S}\mathbf{\Sigma}_0^{-1}|^{n/2} k^{nk/2} / [Tr(\boldsymbol{S}\mathbf{\Sigma}_0^{-1}) + n\bar{\mathbf{X}}^T \mathbf{\Sigma}_0^{-1}\bar{\mathbf{X}}]^{nk/2}$$

ir $-2\ln\Lambda \xrightarrow{d} \chi^2_\nu$, $\nu = (k+1)(k+2)-2$. **5.7.** Statistika $Z = -2\ln\Lambda$ įgijo reikšmę 10,2525. Kadangi $pv_a = \mathbf{P}\{\chi^2_7 > 10,2525\} = 01747$, tai atmesti hipotezę nėra pagrindo. **5.8.** Statistika $Z = -2\ln\Lambda$ įgijo reikšmę 10,7985; $pv_a = \mathbf{P}\{\chi^2_{12} > 10,7985\} = 0,5463$. Atmesti hipotezę nėra pagrindo. **5.9.** a) Statistika $Z = -2\ln\Lambda$ įgijo reikšmę 71,3025; $pv_a = \mathbf{P}\{\chi^2_{10} > 71,3025\} = 2,5\cdot 10^{-11}$. Hipotezė atmestina. b) Statistika $Z = -2\ln\Lambda$ įgijo reikšmę 149,66. Hipotezė atmestina. **5.10.** Statistika $-2\ln\Lambda$ įgijo reikšmę 25,0316; asimptotinė P reikšmė $pv = \mathbf{P}\{\chi^2_3 > 25,0316\} = 1,5\cdot 10^{-5}$; hipotezė atmetama. **5.11.** Statistika U iš (5.2.8) įgijo reikšmę $u = 0,2475$. Kadangi $k=2$, tai turime išnagrinėtą atskirą atvejį ir $pv = (2\sqrt{u})^8$=0,9605; atmesti hipotezę nėra pagrindo. **5.12.** Statistika $Z = -2\ln\Lambda$ įgijo reikšmę 2,3708; $pv_a = \mathbf{P}\{\chi^2_5 > 2,3708\} = 0,7958$. Atmesti hipotezę nėra pagrindo. **5.13.** Statistika $Z = -2\ln\tilde\Lambda$ įgijo reikšmę 9,3996; $pv_a = \mathbf{P}\{\chi^2_{10} > 9,3996\} = 0,4946$; hipotezė neatmetama. **5.14.** Statistika $Z = -2\ln\tilde\Lambda$ įgijo reikšmę 9,5763; $pv_a = \mathbf{P}\{\chi^2_6 > 9,5763\} = 0,1437$. Atmesti hipotezę nėra pagrindo. **5.15.** Statistika $Z = -2\ln\tilde\Lambda$ įgijo reikšmę 26,0674; $pv_a = \mathbf{P}\{\chi^2_3 > 26.0674\} = 9\cdot 10^{-6}$; hipotezė atmetama. **5.16.** Statistika $Z = -2\ln\tilde\Lambda$ įgijo reikšmę 251,6067; $pv_a = \mathbf{P}\{\chi^2_{10} > 111,7974\} = 5,4\cdot 10^{-42}$; hipotezė atmetama. **5.17.** Statistika U iš (5.2.8) įgijo reikšmę $u = 1,2643$. Asimptotinė P reikšmė $pv_a = \mathbf{P}\{\chi^2_2 > 1,2643\} = 0,5394$; atmesti hipotezę nėra pagrindo.

6 skyrius

Diskriminantinė analizė

Tarkime, kad objektą (individą) reikia priskirti vienai iš galimų m klasių remiantis tam tikro atsitiktinio vektoriaus $\mathbf{X} = (X_1, ..., X_k)^T$ matavimais.

Pavyzdžiui, gydytojas, turėdamas paciento tam tikrų simptomų matavimus, bando nustatyti, kuria iš galimų m ligų serga pacientas. Išleidžiamosios kontrolės metu, remiantis gaminio parametrų matavimais, juos bandoma suskirstyti į dvi grupes: gerus gaminius, kurie bus sėkmingai eksploatuojami garantinio laikotarpio metu, ir defektinius gaminius, kurie vartotojo bus reklamuoti. Literatūroje tokių uždavinių sprendimas vadinamas *diskriminantine analize, klasifikavimu, identifikavimu.*

Skyreliuose 6.1, 6.2 aptarsime atvejį, kai žinomi vektoriaus $\mathbf{X}$ skirstiniai kiekvienai iš m klasių. Skyrelyje 6.3 aptarsime atvejį, kai tikimybiniai skirstiniai nėra žinomi ir juos tenka vertinti remiantis statistiniais duomenimis.

6.1. Dviejų klasių atvejis

6.1.1. Klasifikavimo tikslumo tikimybės

Tarkime, objektai, dėl kurių turime priimti sprendimus, gali priklausyti vienai iš dviejų klasių. Dėl trumpumo pažymėkime ξ a. d., kuris įgyja reikšmę 1, jeigu objektas priklauso pirmai klasei, ir įgyja reikšmę 2, jeigu objektas priklauso antrai klasei. Tegu η yra a. d., kuris įgyja reikšmę 1 arba 2, jeigu objektas priskiriamas pirmai arba antrai klasei. Tada klasifikavimo tikslumą apibūdina tikimybės

$$\alpha_{ij} = \mathbf{P}\{\eta = i | \xi = j\}, \quad i, j = 1,\ 2, \tag{6.1.1}$$

kurias surašykime į lentelę

6.1.1 lentelė. Klasifikavimo tikslumo tikimybės

$\xi \backslash \eta$	1	2	Σ
1	α_{11}	α_{21}	1
2	α_{12}	α_{22}	1

Šioje lentelėje α_{11} ir α_{22} yra teisingų sprendimų tikimybės, o α_{12} ir α_{21} – klaidingų. Klasifikavimo taisyklė tuo geresnė, kuo didesnės tikimybės α_{11} ir α_{22} ir kuo mažesnės tikimybės α_{12} ir α_{21}.

Tarkime, a. v. $\mathbf{X}$ tankis σ baigtinio mato $\boldsymbol{\mu}$ atžvilgiu, kai $\xi = i$, yra $f_i(\boldsymbol{x})$, t. y.

$$(\mathbf{X}|\xi = i) \sim f_i(\boldsymbol{x}), \quad i = 1,\ 2. \tag{6.1.2}$$

Randomizuota sprendimo priėmimo taisyklė yra mati funkcija $\boldsymbol{\varphi}(\boldsymbol{x}) = (\varphi_1(\boldsymbol{x}), \varphi_2(\boldsymbol{x}))^T$, $\boldsymbol{x} \in \mathbf{R}^\mathbf{k}$; čia $\varphi_i(\boldsymbol{x})$ reiškia tikimybę priskirti objektą i-ajai klasei, kai a. v. $\mathbf{X}$ įgijo reikšmę $\boldsymbol{x}$:

$$\varphi_i(\boldsymbol{x}) = \mathbf{P}\{\eta = i|\mathbf{X} = \boldsymbol{x}\}, \quad i = 1,\ 2; \quad \varphi_1(\boldsymbol{x}) + \varphi_2(\boldsymbol{x}) \equiv 1. \tag{6.1.3}$$

Tokių funkcijų klasę žymėsime $\boldsymbol{\Phi}$. Tikimybės α_{ij} yra $\boldsymbol{\varphi}(\boldsymbol{x})$ funkcijos

$$\alpha_{ij} = \alpha_{ij}(\boldsymbol{\varphi}) = \int \varphi_i(\boldsymbol{x}) f_j(\boldsymbol{x}) d\boldsymbol{\mu}. \tag{6.1.4}$$

Reikia rasti optimalią sprendimų priėmimo taisyklę $\boldsymbol{\varphi}(\boldsymbol{x})$ iš aibės $\boldsymbol{\Phi}$.

Apibrėžkime nuostolių funkciją. Tarkime, nuostoliai lygūs 0, jeigu priimtas teisingas sprendimas; lygūs $c_{ij} > 0$, kai priimtas sprendimas $\eta = i$, o $\xi = j$, $i \neq j = 1,\ 2$. Tada rizikos funkcija $R = (R_1, R_2)^T$ susideda iš dviejų komponenčių

$$R_1(\boldsymbol{\varphi}) = c_{21}\alpha_{21}(\boldsymbol{\varphi}) = c_{21}\mathbf{P}\{\eta = 2|\xi = 1\} = c_{21} \int \varphi_2(\boldsymbol{x}) f_1(\boldsymbol{x}) d\boldsymbol{\mu},$$

$$R_2(\boldsymbol{\varphi}) = c_{12}\alpha_{12}(\boldsymbol{\varphi}) = c_{12}\mathbf{P}\{\eta = 1|\xi = 2\} = c_{12} \int \varphi_1(\boldsymbol{x}) f_2(\boldsymbol{x}) d\boldsymbol{\mu}. \tag{6.1.5}$$

6.1.2. Sprendimų priėmimo taisyklės

Jeigu žinoma, kaip dažnai kurios klasės objektai pasirodo, t. y. žinomos apriorinės klasių tikimybės

$$\omega_1 = \mathbf{P}\{\xi = 1\}, \quad \omega_2 = \mathbf{P}\{\xi = 2\}, \quad \omega_1 + \omega_2 = 1, \tag{6.1.6}$$

galime apibrėžti vidutinę rizikos funkciją

$$R(\boldsymbol{\varphi}) = R_1(\boldsymbol{\varphi})\omega_1 + R_2(\boldsymbol{\varphi})\omega_2 = c_{12}\alpha_{12}(\boldsymbol{\varphi})\omega_2 + c_{21}\alpha_{21}(\boldsymbol{\varphi})\omega_1$$

ir suformuluoti optimalios sprendimų taisyklės $\boldsymbol{\varphi}^*$ radimo uždavinį:

$$\inf_{\boldsymbol{\varphi} \in \boldsymbol{\Phi}} R(\boldsymbol{\varphi}) = R(\boldsymbol{\varphi}^*). \tag{6.1.7}$$

6.1.1 teorema. *Funkcionalas (6.1.7) įgyja minimalią reikšmę, kai sprendimų funkcija (6.1.3) turi tokį pavidalą*

$$\varphi_1^*(\boldsymbol{x}) = \begin{cases} 1, & \textit{kai} \quad f_1(\boldsymbol{x}) \geq cf_2(\boldsymbol{x}), \\ 0, & \textit{kai} \quad f_1(\boldsymbol{x}) < cf_2(\boldsymbol{x}), \end{cases} \tag{6.1.8}$$

$$c = c_{12}\omega_2/(c_{21}\omega_1), \quad \varphi_2^*(\boldsymbol{x}) = 1 - \varphi_1^*(\boldsymbol{x}).$$

Jeigu su visais c tikimybė $\mathbf{P}\{f_1(\mathbf{X}) = cf_2(\mathbf{X})|\xi = i\} = 0$, $i = 1, 2$, tai sprendinys (6.1.8) yra vienintelis beveik visur mato $\boldsymbol{\mu}$ atžvilgiu.

Įrodymas. Remiantis (6.1.5) funkcionalas

$$R(\boldsymbol{\varphi}) = \int [c_{12}\varphi_1(\boldsymbol{x})f_2(\boldsymbol{x})\omega_2 + c_{21}\varphi_2(\boldsymbol{x})f_1(\boldsymbol{x})\omega_1]d\boldsymbol{\mu}$$

$$= c_{21}\omega_1 + \int \varphi_1(\boldsymbol{x})[c_{12}f_2(\boldsymbol{x})\omega_2 - c_{21}f_1(\boldsymbol{x})\omega_1]d\boldsymbol{\mu}.$$

Integralas įgyja minimalią reikšmę, kai pointegralinis reiškinys kiekviename taške $\boldsymbol{x}$ yra minimalus. Taigi reikia imti $\varphi_1^*(\boldsymbol{x}) = 1$, kai $c_{12}f_2(\boldsymbol{x})\omega_2 < c_{21}f_1(\boldsymbol{x})\omega_1$, ir reikia imti $\varphi_1^*(\boldsymbol{x}) = 0$, kai teisinga priešinga nelygybė. Kai $c_{12}f_2(\boldsymbol{x})\omega_2 = c_{21}f_1(\boldsymbol{x})\omega_1$, tai pointegralinis reiškinys lygus 0, ir galima parinkti bet kokią reikšmę $0 \leq \varphi_1^*(\boldsymbol{x}) \leq 1$, $\varphi_2^*(\boldsymbol{x}) = 1 - \varphi_1^*(\boldsymbol{x})$, nes nuo to funkcionalo R reikšmė nepakinta. Jei aibės, kurioje galioja nurodyta lygybė, matas lygus 0, tai $\boldsymbol{\varphi}$ apibrėžta vienareikšmiškai, išskyrus aibę, kurios $\boldsymbol{\mu}$ matas lygus 0. ▲

Jeigu $c_{12} = c_{21}$, tai sprendimų priėmimo taisyklę galima suformuluoti taip:

$$\varphi_1^*(\boldsymbol{x}) = \begin{cases} 1, & \text{kai} \quad f_1(\boldsymbol{x}) > cf_2(\boldsymbol{x}), \\ 0, & \text{kai} \quad f_1(\boldsymbol{x}) < cf_2(\boldsymbol{x}), \end{cases} \tag{6.1.9}$$

$$c = \omega_2/\omega_1, \quad \varphi_2^*(\boldsymbol{x}) = 1 - \varphi_1^*(\boldsymbol{x}).$$

Kai $f_1(\boldsymbol{x}) = cf_2(\boldsymbol{x})$, tai $0 \leq \varphi_1^*(\boldsymbol{x}) \leq 1$ galima parinkti bet kokiu būdu, nes nuo to funkcionalo reikšmė nepakinta. Pavyzdžiui, galime imti $\varphi_1^*(\boldsymbol{x}) = \gamma$, $0 \leq \gamma \leq 1$ visuose taškuose $\boldsymbol{x}$, kuriuose $f_1(\boldsymbol{x}) = cf_2(\boldsymbol{x})$. Tokią sprendimų priėmimo taisyklę vadinsime *Bejeso taisykle* apriorinio skirstinio $(\omega_1, \omega_2)^T$ atžvilgiu. Ji minimizuoja suminę klasifikavimo klaidą $\alpha_{12}\omega_2 + \alpha_{21}\omega_1$.

Bejeso klasifikavimo taisyklę galima gauti ir kitokiu būdu. Jeigu neturime jokios papildomos informacijos, tai objektą natūralu priskirti i-ajai klasei, jeigu apriorinė tikimybė $\omega_i > \omega_j$. Jei gauta a. v. $\mathbf{X}$ reikšmė $\mathbf{X} = \boldsymbol{x}$, tai remdamiesi Bejeso formule apskaičiuojame aposteriorines klasių tikimybes

$$\omega_i(\boldsymbol{x}) = \mathbf{P}\{\xi = i|\mathbf{X} = \boldsymbol{x}\} = \frac{f_i(\boldsymbol{x})\omega_i}{f_1(\boldsymbol{x})\omega_1 + f_2(\boldsymbol{x})\omega_2}, \quad i = 1,\ 2; \tag{6.1.10}$$

$$\omega_1(\boldsymbol{x}) + \omega_2(\boldsymbol{x}) = 1.$$

Objektą, kurio gauta reikšmė $\mathbf{X} = \boldsymbol{x}$, priskirsime i-ajai klasei, jeigu $\omega_i(\boldsymbol{x}) > \omega_j(\boldsymbol{x})$, arba $\omega_i(\boldsymbol{x}) > 1/2$. Akivaizdu, kad ši taisyklė sutampa su (6.1.9).

Sprendinį (6.1.8) galima interpretuoti taip. Apibrėžiame *diskriminantines funkcijas* $g_1(\boldsymbol{x}) = c_{12}f_2(\boldsymbol{x})\omega_2$ ir $g_2(\boldsymbol{x}) = c_{21}f_1(\boldsymbol{x})\omega_1$, kurių palyginimas sudalija erdvę $\mathbf{R^k}$ į tris nesikertančias dalis $\Omega_1 \cup \Omega_2 \cup \Omega_0 = \mathbf{R^k}$

$$\Omega_i = \{\boldsymbol{x} : g_i(\boldsymbol{x}) < g_j(\boldsymbol{x})\}, \quad i \neq j; \quad \Omega_0 = \{\boldsymbol{x} : g_1(\boldsymbol{x}) = g_2(\boldsymbol{x})\}. \tag{6.1.11}$$

Jeigu stebėtoji reikšmė $\boldsymbol{x}$ patenka į Ω_1, tai objektas priskiriamas pirmai klasei; jeigu patenka į Ω_2 – antrai klasei; jei patenka į pasienio taškų aibę Ω_0, tai objektas gali būti priskirtas bet kuriai klasei.

Suprantama, kai yra dvi klasės, pakanka apsiriboti viena diskriminantine funkcija $g(\boldsymbol{x}) = g_1(\boldsymbol{x}) - g_2(\boldsymbol{x})$. Sprendimas priimamas palyginant $g(\boldsymbol{x})$ su 0.

Kai $\boldsymbol{\mu}(\Omega_0) = 0$, klaidingo klasifikavimo tikimybė yra

$$P\{e\} = \int_{\Omega_1} f_2(\boldsymbol{x})\omega_2 d\boldsymbol{\mu} + \int_{\Omega_2} f_1(\boldsymbol{x})\omega_1 d\boldsymbol{\mu}. \tag{6.1.12}$$

Bejeso klasifikavimo taisyklės diskriminantinę funkciją galima apibrėžti naudojant aposteriorines klasių tikimybes

$$g(\boldsymbol{x}) = \omega_2(\boldsymbol{x}) - \omega_1(\boldsymbol{x}). \tag{6.1.13}$$

6.1.1 pavyzdys. Panagrinėkime paprastą iliustracinį pavyzdį. Tegu $(X|\xi = 1) \sim N(0,\ 1)$, $(X|\xi = 2) \sim N(3,\ 1)$; apriorinės tikimybės $\omega_1 = 1/4, \omega_2 = 3/4$; nuostoliai $c_{12} = 1, c_{21} = 4$. Tada sprendimų priėmimo taisyklė, minimizuojanti rizikos funkciją R, yra

$$\varphi_1^*(x) = \begin{cases} 1, & \text{kai} \quad f_1(x) \geq cf_2(x), \\ 0, & \text{kai} \quad f_1(x) < cf_2(x). \end{cases} = \begin{cases} 1, & \text{kai} \quad x \leq 1,596, \\ 0, & \text{kai} \quad x > 1,596. \end{cases}$$

Bejeso klasifikavimo taisyklėje $\varphi_1^*(x) = 1$, kai $x \leq 1,134$ ir $\varphi_1^*(x) = 0$, kai $x > 1,134$. Suminė mažiausioji klasifikavimo klaida

$$\mathbf{P}\{e\} = \Phi(1,134 - 3)\omega_2 + (1 - \Phi(1,134))\omega_1 = 0,055.$$

Minimakso principas

Jeigu apriorinės klasių tikimybės ω_1 ir ω_2 nežinomos, tai rizikos funkcija $R(\boldsymbol{\varphi})$ yra dvimatis vektorius (6.1.5) ir sprendimų taisykles reikia parinkti atsižvelgiant į abi komponentes.

Sakysime, kad sprendimų priėmimo taisyklė $\boldsymbol{\varphi}$ yra *ne blogesnė* už $\boldsymbol{\varphi}^*$, jeigu $R_1(\boldsymbol{\varphi}) \leq R_1(\boldsymbol{\varphi}^*)$ ir $R_2(\boldsymbol{\varphi}) \leq R_2(\boldsymbol{\varphi}^*)$ ir *geresnė* už $\boldsymbol{\varphi}^*$, jei bent viena nelygybė griežta.

6.1.1 apibrėžimas. Sprendimų priėmimo taisyklę $\boldsymbol{\varphi}$ vadinsime *leistina*, jeigu nėra taisyklės, geresnės už $\boldsymbol{\varphi}$.

Natūralu ieškant optimalios taisyklės apsiriboti leistinų taisyklių aibėmis. Tačiau tokių taisyklių aibė gali būti gana plati. Todėl reikalingi principai, kurie leistų iš šios aibės išskirti tam tikra prasme priimtiniausią elementą. Vienas iš dažniausiai naudojamų tokio tipo principų yra vadinamasis *minimakso* principas.

6.1.3 apibrėžimas. Sakysime, kad sprendimų priėmimo taisyklė $\boldsymbol{\varphi}^*$ tenkina minimakso principą, jeigu ji minimizuoja rizikos funkcijos

$$R(\boldsymbol{\varphi}) = (R_1(\boldsymbol{\varphi}), R_2(\boldsymbol{\varphi}))^T$$

maksimalią komponentę:

$$\boldsymbol{\varphi}^* = \{\boldsymbol{\varphi} : \min_{\boldsymbol{\varphi}\in\Phi^*} (\max(R_1(\boldsymbol{\varphi}), R_2(\boldsymbol{\varphi})))\}. \tag{6.1.14}$$

Įrodyta, kad šio uždavinio sprendinys yra Bejeso taisyklė, surasta turint tokį (mažiausiai palankų) apriorinių tikimybių rinkinį ω_1^*, ω_2^*, kai rizikos funkcijos (6.1.5) komponentės $R_1(\boldsymbol{\varphi}^*)$ ir $R_2(\boldsymbol{\varphi}^*)$ yra lygios (žr. [14]).

6.1.2 pavyzdys. (**6.1.1** pavyzdžio tęsinys). Rasime sprendimų priėmimo taisyklę, tenkinančią minimakso principą **6.1.1** pavyzdžio sąlygomis, kai apriorinės klasių tikimybės nežinomos.

Kai klasių tikimybės nežinomos, tai pagal Bejeso taisyklę (6.1.9) funkcija $\varphi_1(x) = 1$, kai $x \le c(\omega_1) = 1,5 - (1/3)\ln((1-\omega_1/(\omega_1)))$. Palyginę rizikos funkcijos komponentes, gauname lygtį

$$c_{12}\Phi(c(\omega_1) - 3) = c_{21}(1 - \Phi((c(\omega_1)))),$$

iš kurios randame, kad mažiausiai palankus apriorinių tikimybių rinkinys yra $(\omega_1, \omega_2) = (0,745,\ 0,255)$. Šio rinkinio abi rizikos funkcijos komponentės vienodos ir lygios $0,1266$. Suminė klasifikavimo klaida $\mathbf{P}\{e\} = 0,056$, suprantama, yra didesnė už tą, kuri buvo gauta **6.1.1** pavyzdyje.

6.1.3. Klasifikavimas, kai yra apribojimų

Ne visada galima apsiriboti nuostolių funkcijos minimizavimu, o reikia, kad būtų patenkinti tam tikri apribojimai. Pavyzdžiui, vartotojas gali pageidauti, jog produkcijos kontrolė būtų organizuota taip, kad jam patenkančios produkcijos defektinių gaminių dalis būtų maža. Chirurgas gali reikalauti, kad tarp siunčiamų operuoti ligonių būtų maža dalis tokių, kuriems operacija nereikalinga, ir pan.

Hipotezių tikrinimo uždavinys

Jeigu apribosime vieną rizikos funkcijos komponentę ir stengsimės minimizuoti kitą komponentę, tai ekvivalentu tokiam uždaviniui:

$$\begin{cases} \alpha_{21}(\boldsymbol{\varphi}) \to \min, \\ \alpha_{12}(\boldsymbol{\varphi}) \le \alpha, \end{cases} \quad \text{arba} \quad \begin{cases} \alpha_{11}(\boldsymbol{\varphi}) \to \max, \\ \alpha_{12}(\boldsymbol{\varphi}) \le \alpha. \end{cases} \tag{6.1.15}$$

Toks uždavinys vadinamas hipotezių tikrinimo uždaviniu. Tikrinama paprastoji hipotezė $H : \xi = 1$, kai alternatyva $\bar{H} : \xi = 2$ taip pat paprastoji. Skaičius $0 < \alpha < 1$ vadinamas kriterijaus *reikšmingumo lygmeniu*; tikimybė $\alpha_{11}(\boldsymbol{\varphi})$ vadinama *kriterijaus galia*. Uždavinio (6.1.15) sprendinys $\boldsymbol{\varphi}^* \in \boldsymbol{\Phi}$ vadinamas *galingiausiuoju kriterijumi*. Uždavinio sprendinys $\boldsymbol{\varphi}^*$ duotas fundamentaliojoje Neimano ir Pirsono lemoje (žr. I dalį, 5.2.1 skyrelį):

$$\boldsymbol{\varphi}_1^*(\boldsymbol{x}) = \begin{cases} 1, & f_1(\boldsymbol{x}) > cf_2(\boldsymbol{x}), \\ \delta, & f_1(\boldsymbol{x}) = cf_2(\boldsymbol{x}), \\ 0, & f_1(\boldsymbol{x}) < cf_2(\boldsymbol{x}) \end{cases} \tag{6.1.16}$$

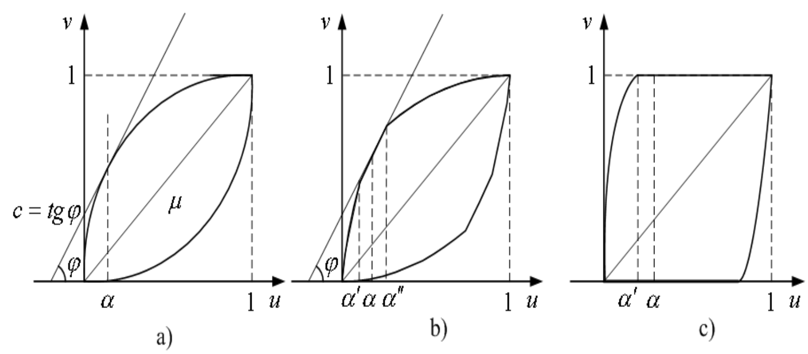
Konstantos c, δ randamos iš sąlygos

$$\alpha_{12}(\boldsymbol{\varphi}^*) = \alpha. \tag{6.1.17}$$

Pateiksime geometrinę Neimano ir Pirsono lemos interpretaciją. Apibrėžkime vienetinio kvadrato sritį

$$\boldsymbol{M} = \{(u,v) = (\alpha_{12}(\boldsymbol{\varphi}), \alpha_{11}(\boldsymbol{\varphi})) : \boldsymbol{\varphi} \in \boldsymbol{\Phi}\}. \tag{6.1.18}$$

Remdamiesi aibės $\boldsymbol{\Phi}$ iškilumu gauname, kad aibė $\boldsymbol{M}$ taip pat iškila. Įrodyta, kad aibė $\boldsymbol{M}$ yra uždara (žr. [12]). Aibei $\boldsymbol{M}$ priklauso taškas (0, 0), kai $\boldsymbol{\varphi}_1(\boldsymbol{x}) \equiv 0$; priklauso taškas (1, 1), kai $\boldsymbol{\varphi}_1(\boldsymbol{x}) \equiv 1$, taigi ir kvadrato diagonalė, jungianti šiuos taškus. Aibė $\boldsymbol{M}$ simetriška taško (1/2, 1/2) atžvilgiu. Norint tuo įsitikinti pakanka imti funkciją $(\boldsymbol{\varphi}_1(\boldsymbol{x}), \boldsymbol{\varphi}_2(\boldsymbol{x}))$ ir funkciją $(1 - \boldsymbol{\varphi}_1(\boldsymbol{x}), 1 - \boldsymbol{\varphi}_2(\boldsymbol{x}))$. Jeigu tankiai sutampa, tai $\boldsymbol{M}$ yra diagonalė; jeigu tankiai singuliarūs vienas kito atžvilgiu, tai $\boldsymbol{M}$ sutampa su vienetiniu kvadratu (žr. 6.1.1 pav.).



6.1.1 pav. Geometrinė interpretacija

Naujais terminais uždavinys (6.1.15) įgauna tokį pavidalą:

$$\begin{cases} v \to \max, \\ u \le \alpha, \end{cases} \qquad (u,v) \in \boldsymbol{M}. \tag{6.1.19}$$

Pagal Lagranžo neapibrėžtinių daugiklių metodą reikia maksimizuoti

$$g(u,v) = v - \lambda u, \quad (u,v) \in \boldsymbol{M}.$$

Taigi atsakymą atitinka toks taškas ant viršutinės gaubiamosios srities $\boldsymbol{M}$ linijos, kad lietėja (atraminė tiesė) $g(u,v)$ įgyja maksimalią reikšmę. Kai λ prabėga reikšmes nuo 0 iki ∞, tai gauname visus srities $\boldsymbol{M}$ viršutinės gaubiamosios linijos taškus; parametras λ lygus kampo φ, kuri sudaro lietėja su abscisių ašimi, tangentui, $\lambda = \operatorname{tg} \varphi$.

Matome, kad atsakymą atitiks liestinė, pravesta taške $u = \alpha$. Jeigu šiame taške liestinė turi tik vieną bendrą tašką su sritimi $\boldsymbol{M}$, tai randomizacijos nereikia (6.1.1 pav. a)). Jeigu liestinės su sritimi $\boldsymbol{M}$ bendri taškai sudaro intervalą, kai $u \in [\alpha', \alpha''], \alpha' < \alpha < \alpha''$ (6.1.1 pav, b)), tai reikalinga randomizacija ir $\delta = (\alpha - \alpha')/(\alpha'' - \alpha')$. Sprendinys vienintelis, išskyrus atvejį, kai taškuose $u \in [\alpha', \alpha]$, $\alpha' < \alpha$, kriterijaus galia $v = 1$ (6.1.1 pav. c)).

6.1.3 pavyzdys. Tegu, kaip ir **6.1.1** pavyzdyje $(X|\xi = 1) \sim N(0,\ 1)$ ir $(X|\xi = 2) \sim N(3,\ 1)$. Reikia rasti klasifikavimo taisyklę, kad $\alpha_{12} \leq \alpha = 0,05$, o $\alpha_{21} \to \min$. Randame, kad $\eta = 1$, kai $x < c$. Konstanta c randama iš sąlygos

$$\alpha_{12} = \mathbf{P}\{X < c|\xi = 2\} = \Phi(c - 3) = 0,05.$$

Randame $c = 3 + z_{0,05} = 1,355$.

Netiesiniai apribojimai.

Tiesinių funkcionalų apribojimai (6.1.15) ne visada atitinka realią situaciją. Norėdami tuo įsitikinti panagrinėkime paprastą pavyzdį.

Atliekama išleidžiamoji produkcijos kontrolė, po kurios gerais pripažinti gaminiai siunčiami vartotojui. Tegu $\xi = 1$, jei gaminys geras, ir $\xi = 2$, jei gaminys defektinis. Vartotoją pirmiausia domina, kokią jo gautos produkcijos dalį sudaro defektiniai gaminiai (t. y. tikimybė $\mathbf{P}\{\xi = 2|\eta = 1\}$), o ne kaip dažnai defektinis gaminys pripažįstamas geru (tikimybė $\alpha_{12} = \mathbf{P}\{\eta = 1|\xi = 2\}$). Pavyzdžiui, jeigu $\omega_2 = 1$, tai kad ir kokia maža būtų tikimybė α_{12}, vartotojui pateks vien defektiniai gaminiai (kitokių ir nebuvo). Taigi tikimybės $\mathbf{P}\{\xi = 2|\eta = 1\}$ mažumą lemia ne tik tikimybės α_{12} mažumas, bet ir apriorinės tikimybės ω_1, ω_2.

Apibrėžkime tikimybes

$$\beta_{ji} = \mathbf{P}\{\xi = j|\eta = i\}, \quad i, j = 1,\ 2 \tag{6.1.20}$$

ir surašykime jas į lentelę.

6.1.2 lentelė. Aposteriorinės klasifikavimo tikslumo tikimybės

ξ \ η	1	2
1	β_{11}	β_{12}
2	β_{21}	β_{22}
$\sum$	1	1

Klasifikavimo tikimybės α_{ij}, pateiktos 6.1.1 lentelėje, apibūdina klaidas, kurias atliksime taikydami klasifikatorių, t. y. prieš atlikdami klasifikavimą. Tikimybės β_{ji} apibūdina klaidas, kurias gavome, kai klasifikavimas jau atliktas, t. y. apibūdina po klasifikavimo gautų aibių užterštumą kitų aibių elementais. Tiek aptartame pavyzdyje, tiek ir kituose klasifikavimo uždaviniuose, matyt, svarbiau yra žinoti, kokius rezultatus gavome atlikę klasifikavimą. Klasifikavimo tikimybes β_{ji} natūralu vadinti aposteriorinėmis klasifikavimo klaidų tikimybėmis, o tikimybes α_{ij} – apriorinėmis.

Nagrinėkime uždavinį

$$\begin{cases} \alpha_{11} \to \max, \\ \beta_{21} \leq \beta. \end{cases} \tag{6.1.21}$$

6.1.2 teorema. *Tegu $\beta < \omega_2$. Tada uždavinio (6.1.21) sprendinys turi (6.1.16)* *struktūrą. Parametrai c, δ randami iš sąlygos*

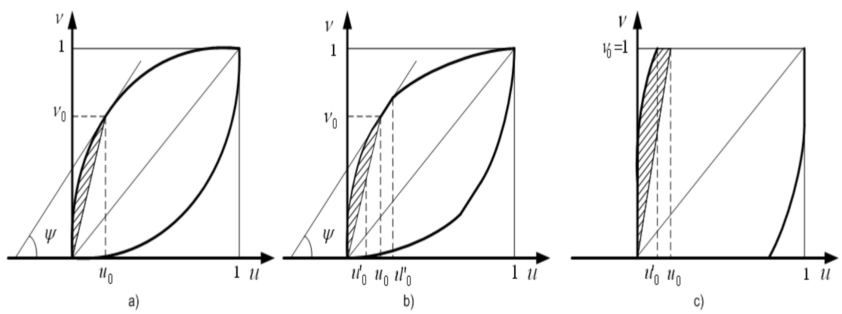
$$\beta_{21}(\boldsymbol{\varphi}^*) = \beta. \tag{6.1.22}$$

Sprendinys apibrėžiamas vienareikšmiškai, išskyrus tą atvejį, kai egzistuoja sprendinys $\boldsymbol{\varphi}^*$, kuriam $\beta_{21}(\boldsymbol{\varphi}^*) < \beta$, o $\alpha_{11}(\boldsymbol{\varphi}^*) = 1$.

Įrodymas. Naudosime geometrinę interpretaciją. Perėjus prie srities $\boldsymbol{M}$ kintamųjų (u, v) terminais uždavinį galima pertvarkyti į tokį:

$$\begin{cases} v \to \max, \\ v \geq \frac{\omega_2(1-\beta)}{\omega_1\beta}u, \end{cases} \qquad (u, v) \in \boldsymbol{M}.$$

Srities $\boldsymbol{M}$ poaibis, kuriame ieškomas sprendinys, pavaizduotas 6.1.2 pav. (užštrichuotas).



6.1.2 pav. Geometrinė interpretacija

Matome, kad ir šiame uždavinyje sprendinys yra taškas (u_0, v_0) ant viršutinės gaubiamosios srities $\boldsymbol{M}$ linijos, kuris gaunamas susikertant šiai linijai su tiese $v = u(\omega_2(1-\beta))/(\omega_1\beta)$.

Jeigu $\beta \geq \omega_2$, tai apribojimas patenkintas visiems viršutinę gaubiamąją liniją atitinkantiems sprendiniams (neesminis apribojimas) ir galima imti sprendinį $\varphi_1(\boldsymbol{x}) \equiv 1$, $\alpha_{11} = 1$.

Jeigu $\beta < \omega_2$, tai sprendinys turės (6.1.16) pavidalą, kuriame parametrus c, δ randame iš sąlygos

$$\beta_{21}(\boldsymbol{\varphi}^*) = \beta, \quad \text{arba} \quad \alpha_{12}(\boldsymbol{\varphi}^*) = u_0. \tag{6.1.23}$$

Jeigu taške $u = u_0$ liestinė turi tik vieną bendrą tašką su sritimi $\boldsymbol{M}$, tai randomizacijos nereikia (6.1.2 pav. a)). Jeigu liestinės su sritimi $\boldsymbol{M}$ bendri taškai sudaro intervalą, kai $u \in [u_0', u_0''], u_0' < u_0 < u_0''$ (6.1.2 pav, b)), tai reikalinga randomizacija ir $\delta = (u_0 - u_0')/(u_0'' - u_0')$. Sprendinys vienintelis, išskyrus atvejį, kai taškuose $u \in [u_0', u_0]$, $u_0' < u_0$, kriterijaus galia $v = 1$ (6.1.2 pav. c)). ▲

6.1.4 pavyzdys. (**6.1.1** pavyzdžio tęsinys). **6.1.1** pavyzdžio sąlygomis rasime sprendimų priėmimo taisyklę, kad tikimybė α_{11} būtų maksimali, o aposteriorinė tikimybė $\beta_{21} \leq 0,05$.

Šiame pavyzdyje randomizacija nereikalinga. Objektas priskiriamas pirmai klasei, kai $x < d$. Remiantis (6.1.22) konstantai d rasti gauname lygtį

$$\Phi(d) = 57\Phi(d-3).$$

Iš šios lygties randame $d = 0,7975$. Matome, kad šio uždavinio sprendimas ir atsakymas iš esmės skiriasi nuo **6.1.3** pavyzdžio, kuriame buvo apribota apriorinė klaidos tikimybė.

6.1.1 pastaba. Dažnai spendžiant klasifikavimo uždavinius ne tik apribojimai, bet ir funkcionalai, kuriuos maksimizuojame ar minimizuojami, yra netiesinės φ funkcijos. Tokių uždavinių sprendimas aptariamas [11].

Hierarchinė klasifikacija

Sprendimų priėmimo taisyklė (6.1.16) faktiškai priklauso tik nuo vieno parametro. Jo parinkimas leidžia valdyti vieną iš klasifikavimo klaidų tikimybių. Kai yra fiksuotas informacijos kiekis valdyti abi klasifikavimo klaidų tikimybes negalime. Galimybė valdyti abiejų rūšių klaidas atsiranda, jeigu įtartinais atvejais galutinis sprendimas nepriimamas. Suprantama, pageidautina, kad atsisakymas priimti galutinį sprendimą būtų kuo retesnis, tačiau abiejų rūšių klaidų tikimybės neviršytų leistinų ribų.

Su objektais, dėl kurių galutinis sprendimas nepriimtas, gali būti elgiamasi įvairiai. Pavyzdžiui, atliekant produkcijos kokybės kontrolę „įtartiniems" gaminiams gali būti mažinama kaina, žeminamas rūšingumas ir pan. Kita galimybė – likusius objektus klasifikuoti remiantis papildomai surinkta tikslesne informacija, kurios panaudoti visiems objektams klasifikuoti negalima dėl per didelių laiko ar lėšų sąnaudų.

Taigi natūraliai gauname *hierarchinės klasifikacijos* procedūrą. Pirmame etape, remdamiesi minimalia informacija, suklasifikuojame objektus taip, kad abiejų rūšių klaidų tikimybės būtų apribotos, o likusių nesuklasifikuotų objektų dalis būtų minimali. Antrame etape, remdamiesi tikslesne informacija, klasifikuojame objektus, dėl kurių sprendimas pirmame etape nebuvo priimtas. Jeigu antrojo etapo turima informacija nepakankama, kad būtų galima išklasifikuoti visus likusius objektus esant apribotoms klaidų tikimybėms, tai vėl galime dėl dalies objektų galutinio sprendimo nepriimti. Tada klasifikavimas gali būti užbaigtas trečiame etape, remiantis nauja papildoma informacija, ir t. t.

Nagrinėkime pirmą etapą.

Taip, kaip ir pirmiau, tegu a. d. ξ įgyja reikšmę 1 arba 2, kai objektas priklauso pirmai arba antrai klasei, o a. d. η įgyja reikšmę 1 arba 2, kai objektas priskiriamas pimai arba antrai klasei ir įgyja reikšmę 0, jeigu galutinis sprendimas nepriimtas. Klasifikavimo tikslumą apibūdina apriorinės tikimybės

$$\alpha_{ij} = \mathbf{P}\{\eta = i | \xi = j\}, \quad i = 1,\ 2,\ 0; \quad j = 1,\ 2, \tag{6.1.24}$$

arba aposteriorinės (sprendimą priėmus) tikimybės

$$\beta_{ji} = \mathbf{P}\{\xi = j | \eta = i\} = \frac{\alpha_{ij}\omega_j}{\alpha_{i1}\omega_1 + \alpha_{i2}\omega_2}, \quad j = 1,\ 2, \quad i = 1,\ 2,\ 0, \tag{6.1.25}$$

čia $\omega_i = \mathbf{P}\{\xi = i\}, \omega_1 + \omega_2 = 1$, apriorinės klasių tikimybės.

Tarkim, sprendimas priimamas remiantis a. v. $\mathbf{X}$, kurio tankiai σ-baigtinio mato $\boldsymbol{\mu}$ atžvilgiu yra

$$(\mathbf{X}|\xi = i) \sim f_i(\boldsymbol{x}), \quad i = 1,\ 2.$$

Randomizuotoji sprendimų priėmimo taisyklė šiuo atveju yra trimatė funkcija $\boldsymbol{\varphi} = \boldsymbol{\varphi}(\boldsymbol{x}) = (\varphi_1(\boldsymbol{x}), \varphi_2(\boldsymbol{x}), \varphi_0(\boldsymbol{x}))$; čia

$$\varphi_i(\boldsymbol{x}) = \mathbf{P}\{\eta = i|\mathbf{X} = \boldsymbol{x}\}, \quad \varphi_1(\boldsymbol{x}) + \varphi_2(\boldsymbol{x}) + \varphi_0(\boldsymbol{x}) \equiv 1.$$

Tokių sprendimų priėmimo taisyklių aibę vėl žymėsime $\boldsymbol{\Phi}$. Klasifikavimo tikimybės α_{ij} yra tiesiniai $\boldsymbol{\varphi}$ funkcionalai:

$$\alpha_{ij} = \alpha_{ij}(\boldsymbol{\varphi}) = \int \varphi_i(\boldsymbol{x}) f_j(\boldsymbol{x}) d\boldsymbol{\mu}, \quad i = 1,\ 2,\ 0; \quad j = 1,\ 2. \tag{6.1.26}$$

Nagrinėsime sąlyginio ekstremumo radimo uždavinį, kai apribotos abiejų rūšių klaidų tikimybės α_{12} ir α_{21} ir minimizuojamos atsisakymų nuo priimti sprendimą tikimybės:

$$\begin{cases} \alpha_{01} \to \min, \\ \alpha_{02} \to \min, \\ \alpha_{12} \le \alpha_1, \\ \alpha_{21} \le \alpha_2, \end{cases} \quad \text{arba} \quad \begin{cases} \alpha_{11} \to \max, \\ \alpha_{22} \to \max, \\ \alpha_{12} \le \alpha_1, \\ \alpha_{21} \le \alpha_2, \end{cases} \quad \boldsymbol{\varphi} \in \boldsymbol{\Phi}. \tag{6.1.27}$$

6.1.3 teorema. *Tegu $\varphi_1^*(\boldsymbol{x})$ ir $\varphi_2^*(\boldsymbol{x})$ yra sąlyginio ekstremumo radimo uždavinių*

$$\begin{cases} \alpha_{11} \to \max, \\ \alpha_{12} \le \alpha_1, \end{cases} \quad \textit{ir} \quad \begin{cases} \alpha_{22} \to \max, \\ \alpha_{21} \le \alpha_2, \end{cases} \tag{6.1.28}$$

sprendiniai. Jeigu su visais $\boldsymbol{x}$ suma $\varphi_1^*(\boldsymbol{x}) + \varphi_2^*(\boldsymbol{x}) \le 1$, tai funkcija $\boldsymbol{\varphi} = (\varphi_1^*(\boldsymbol{x}), \varphi_2^*(\boldsymbol{x}), 1 - \varphi_1^*(\boldsymbol{x}) - \varphi_2^*(\boldsymbol{x}))$ yra daugiaekstremaliojo uždavinio (6.1.27) sprendinys.

Įrodymas. Pirmasis uždavinys priklauso tik nuo sprendimų funkcijos pirmosios komponentės φ_1, o antrasis tik nuo antrosios komponentės φ_2, todėl juos galime nagrinėti atskirai. Tokių uždavinių sprendiniai pateikti (6.1.16). Gauname

$$\boldsymbol{\varphi}_1^*(\boldsymbol{x}) = \begin{cases} 1, & f_1(\boldsymbol{x}) > c_1 f_2(\boldsymbol{x}), \\ \delta_1, & f_1(\boldsymbol{x}) = c_1 f_2(\boldsymbol{x}), \\ 0, & f_1(\boldsymbol{x}) < c_1 f_2(\boldsymbol{x}) \end{cases} \tag{6.1.29}$$

Konstantos c_1, δ_1 randamos iš sąlygos $\alpha_{12}(\boldsymbol{\varphi}_1^*) = \alpha_1$.

Analogiškai, spręsdami kitą uždavinį gauname

$$\boldsymbol{\varphi}_2^*(\boldsymbol{x}) = \begin{cases} 1, & f_2(\boldsymbol{x}) > c_2 f_1(\boldsymbol{x}), \\ \delta_2, & f_2(\boldsymbol{x}) = c_2 f_1(\boldsymbol{x}), \\ 0, & f_2(\boldsymbol{x}) < c_2 f_1(\boldsymbol{x}) \end{cases} \tag{6.1.30}$$

Konstantos c_2, δ_2 randamos iš sąlygos $\alpha_{21}(\boldsymbol{\varphi}_2^*) = \alpha_2$.

Šie sprendiniai maksimizuoja tikimybes α_{11} ir α_{22}. Jeigu jie nepersidengia, t. y. jeigu $c_1c_2 < 1$, arba $c_1c_2 = 1$, tačiau $\delta_1 + \delta_2 \le 1$, tada galime imti $\varphi_0^*(\boldsymbol{x}) = 1 - \varphi_1^*(\boldsymbol{x}) - \varphi_2^*(\boldsymbol{x})$. Funkcija $\boldsymbol{\varphi}^* = (\varphi_1^*, \varphi_2^*, \varphi_0^*) \in \boldsymbol{\Phi}$ ir maksimizuoja abi tikimybes α_{11} ir α_{22}, t. y. ji yra daugiaekstremaliojo uždavinio (6.1.27) sprendinys. ▲

6.1.2 pastaba. Jeigu apriorinės tikimybės ω_1, ω_2 žinomos, tai visiškai analogiškai galime spręsti uždavinį (6.1.27), kuriame apribotos abi aposteriorinės klaidų tikimybės β_{12} ir β_{21}.

6.1.3 pastaba. Jeigu teoremoje pateiktų atskirų uždavinių sprendiniai persidengia, tai reiškia, kad kai yra (6.1.27) uždavinio apribojimai sprendimo atsisakyti nereikia, t. y. $\alpha_{01} = \alpha_{02} = 0$. Sprendinys, kuris maksimizuotų abi tikimybes α_{11} ir α_{22}, neegzistuoja. Tokiu atveju galime imti sprendimų priėmimo taisyklę, kuri maksimizuotų teisingo klasifikavimo tikimybę $\alpha_{11} + \alpha_{22}$. Abu (6.1.27) apribojimai automatiškai tenkinami.

6.1.5 pavyzdys (**6.1.1** pavyzdžio tęsinys). **6.1.1** pavyzdžio sąlygomis rasime sprendimų priėmimo taisyklę, kad sprendimo atsisakymo tikimybės α_{01} ir α_{02} būtų minimalios ir a) abi klasifikavimo klaidos α_{12} ir α_{21} neviršytų 0,05; b) abi aposteriorinės klasifikavinmo klaidos β_{12} ir β_{21} neviršytų 0,05 (apriorinės klasių tikimybės vienodos).

a) Remiantis (6.1.29) $\varphi_1^*(x) = 1$, kai $x \le c$ o c randamas iš sąlygos $\Phi(c-3) = 0,05$; gaunama $\varphi_1^*(x) = 1$, kai $x \le 1,3551$; analogiškai $\varphi_2^*(x) = 1$, kai $x \ge 1,6449$, $\varphi_0^*(x) = 1$, kai $1,3551 < x < 1,6449$.

b) Remiantis (6.1.23) $\varphi_1^*(x) = 1$, kai $x \le c$ o c randamas iš sąlygos $\Phi(c-3)/(\Phi(c-3) + \Phi(c)) = 0,05$; gaunama $\varphi_1^*(x) = 1$, kai $x \le 1,3338$; analogiškai $\varphi_2^*(x) = 1$, kai $x \ge 1,6662$, $\varphi_0^*(x) = 1$, kai $1,3338 < x < 1,6662$.

6.1.4. Normaliojo skirstinio atvejis

Tarkime, apibūdinančio objektus a. v. $\mathbf{X}$ skirstinys kiekvienai iš klasių yra k-matis normalusis

$$(\mathbf{X}|\xi = i) \sim N_k(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), \quad i = 1,\ 2. \tag{6.1.31}$$

Kai skirstinys yra eksponentinio tipo, tai diskriminantine funkcija patogu imti tankių santykio logaritmą

$$g(\boldsymbol{x}) = \ln f_2(\boldsymbol{x}) - \ln f_1(\boldsymbol{x}) + d, \quad d = \ln \frac{c_{21}\omega_1}{c_{12}\omega_2}.$$

Gauname

$$g(\boldsymbol{x}) = \frac{1}{2}\{(\boldsymbol{x} - \boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}_2^{-1}(\boldsymbol{x} - \boldsymbol{\mu}_1) - (\boldsymbol{x} - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}_1^{-1}(\boldsymbol{x} - \boldsymbol{\mu}_2)\} - \tag{6.1.32}$$

$$-\frac{1}{2}\ln|\boldsymbol{\Sigma}_2| + \frac{1}{2}\ln|\boldsymbol{\Sigma}_1| + d.$$

Bendru atveju, kai $\boldsymbol{\Sigma}_1 \ne \boldsymbol{\Sigma}_2$, skiriamasis paviršius $g(\boldsymbol{x}) = 0$, gali būti bet koks antrojo laipsnio k-matės erdvės paviršius. Atveju $k = 2$ tai gali būti tiesė, apskritimas, elipsė, parabolė, hiperbolė, hiperbolė, išsigimstanti į dvi tieses. Panagrinėsime atskirus atvejus.

1. Tegu kovariacinės matricos $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \sigma^2 \boldsymbol{I}$. Tada

$$g(\boldsymbol{x}) = \frac{1}{2\sigma^2}\{(\boldsymbol{x}-\boldsymbol{\mu}_1)^T(\boldsymbol{x}-\boldsymbol{\mu}_1) - (\boldsymbol{x}-\boldsymbol{\mu}_2)^T(\boldsymbol{x}-\boldsymbol{\mu}_2)\} + d.$$

Jeigu $d = 0$, tai gauname tokią taisyklę: objektas, kuriam gauta a. v. $\mathbf{X}$ realizacija $\boldsymbol{x}$, priskiriamas pirmai klasei, jeigu atstumas tarp $\boldsymbol{x}$ ir $\boldsymbol{\mu}_1$ yra mažesnis už atstumą tarp $\boldsymbol{x}$ ir $\boldsymbol{\mu}_2$, t. y. diskriminantinė funkcija yra hiperplokštuma, statmena atkarpai, jungiančiai taškus $\boldsymbol{\mu}_1$ ir $\boldsymbol{\mu}_2$ ir einanti per šios atkarpos vidurį. Jeigu $c_{21} > c_{12}$, tai ši hiperplokštuma bus pastumta link taško $\boldsymbol{\mu}_2$.

Suprastinę kvadratinius narius gausime minėtos hiperplokštumos lygtį

$$g(\boldsymbol{x}) = \frac{1}{\sigma^2}(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)^T\boldsymbol{x} + \frac{1}{2\sigma^2}(\boldsymbol{\mu}_1^T\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2^T\boldsymbol{\mu}_2) + d = 0; \tag{6.1.33}$$

objektas priskiriamas pirmai klasei, kai $g(\boldsymbol{x}) < 0$.

2. Tegu bendresniu atveju kovariacinės matricos yra vienodos, tačiau nebūtinai diagonaliosios $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}$. Tada

$$g(\boldsymbol{x}) = \frac{1}{2}\{(\boldsymbol{x}-\boldsymbol{\mu}_1)^T\mathbf{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_1) - (\boldsymbol{x}-\boldsymbol{\mu}_2)^T\mathbf{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_2)\} + d.$$

Reiškinys $(\boldsymbol{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$ vadinamas *Machalanobio atstumu* tarp taškų $\boldsymbol{x}$ ir $\boldsymbol{\mu}$. Jeigu $d = 0$, tai gauname tokią taisyklę: objektas, kuriam $\mathbf{X} = \boldsymbol{x}$, priskiriamas tai klasei, kurios Machalanobio atstumas tarp $\boldsymbol{x}$ ir klasės vidurkio yra mažesnis. Suprastinę kvadratinius narius įsitikiname, kad ir šiuo atveju diskriminantinė funkcija yra hiperplokštuma

$$g(\boldsymbol{x}) = w^T\boldsymbol{x} + w_1 - w_2 + d, \tag{6.1.34}$$

$$w^T = (\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)^T\mathbf{\Sigma}^{-1}, \quad w_i = \frac{1}{2}\boldsymbol{\mu}_i^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_i, \quad i = 1,\ 2.$$

Jei $d = 0$, tai ši hiperplokštuma taip pat eina per atkarpos, jungiančios $\boldsymbol{\mu}_1$ ir $\boldsymbol{\mu}_2$, vidurį, tačiau nebūtinai jai statmena.

Rasime klasifikavimo klaidos tikimybę.

Atsitiktinio dydžio $g(\mathbf{X})$ vidurkis, kai $\xi = 1$ ir $\xi = 2$, yra

$$\mathbf{E}(g(\mathbf{X})|\xi = 1) = -\frac{1}{2}a + d, \quad \mathbf{E}(g(\mathbf{X})|\xi = 2) = \frac{1}{2}a + d,$$

$$d = \ln(c_{21}/c_{12}), \quad a = (\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^T\mathbf{\Sigma}^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2),$$

o dispersija vienoda ir lygi

$$\boldsymbol{V}(g(\mathbf{X})|\xi = i) = a, \quad i = 1,\ 2.$$

Todėl klaidingo klasifikavimo tikimybė

$$P\{e\} = \mathbf{P}\{g(\mathbf{X}) < 0|\xi = 2\}\omega_2 + \mathbf{P}\{g(\mathbf{X}) > 0|\xi = 1\}\omega_1 =$$

$$= \omega_1(1 - \Phi(\frac{(a/2-d)}{\sqrt{a}})) + \omega_2(1 - \Phi(\frac{(-a/2-d)}{\sqrt{a}})). \tag{6.1.35}$$

Jeigu apriorinės tikimybės vienodos ir $d = 0$, tai $P\{e\} = 1 - \Phi(\sqrt{a}/2)$.

6.2. Klasifikavimas, kai klasių daugiau negu dvi

6.2.1. Klasifikavimo tikslumo charakteristikos

Tarkime, objektai, dėl kurių turime priimti sprendimus, gali priklausyti vienai iš $m > 2$ klasių. Pažymėkime ξ a. d., kuris įgyja reikšmę i, jeigu objektas priklauso i-ajai klasei, $i = 1, 2, ..., m$. Tegu η yra a. d., kuris įgyja reikšmę i, jeigu objektas priskiriamas i-ajai klasei, i = 1,2,...,m. Tada klasifikavimo tikslumą galima apibūdinti tikimybėmis

$$\alpha_{ij} = \mathbf{P}\{\eta = i | \xi = j\}, \quad i, j = 1,\ 2, ..., m, \tag{6.2.1}$$

kurios surašytos į lentelę

6.2.1 lentelė. Klasifikavimo tikslumo tikimybės

$\xi \backslash \eta$	1	2	$\cdots$	m	Σ
1	α_{11}	α_{21}	$\cdots$	α_{m1}	1
2	α_{12}	α_{22}	$\cdots$	α_{m2}	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
m	α_{1m}	α_{2m}	$\cdots$	α_{mm}	1

Šioje lentelėje tik ant diagonalės yra teisingų sprendimų tikimybės; α_{ii} yra tikimybė, kad i-osios klasės objektas priskiriamas i-ajai klasei. Visos kitos tikimybės apibūdina klasifikavimo klaidas; α_{ij} yra tikimybė, kad j-osios klasės objektas priskiriamas i-ajai klasei, $i \neq j = 1, ..., m$. Taigi m klasių atveju gauname $m(m-1)$ skirtingo tipo klasifikavimo klaidų. Klasifikavimo taisyklė tuo geresnė, kuo didesnės tikimybės α_{ii} ir kuo mažesnės tikimybės $\alpha_{ij}, i \neq j = 1, ..., m$. Praktiniu požiūriu, matyt, svarbesnės yra aposteriorinės (klasifikavimą atlikus) klaidų tikimybės

$$\beta_{ji} = \mathbf{P}\{\xi = j | \eta = i\} = \frac{\alpha_{ij}\omega_j}{\alpha_{i1}\omega_1 + ... + \alpha_{im}\omega_m}, \quad j, i = 1, ..., m, \tag{6.2.2}$$

čia $\omega_i = \mathbf{P}\{\xi = i\}$ yra apriorinė i-osios klasės tikimybė, $\omega_1 + ... + \omega_m = 1$. Tikimybė $\beta_{ji}, j \neq i$, reiškia objektų aibės, kurie klasifikuojant rezultate buvo priskirti j-ajai klasei, užterštumą i-osios klasės objektais.

Tarkime, kad klasifikavimą atliksime remdamiesi objektą apibūdinančių požymių vektoriaus $\mathbf{X} = (X_1, ..., X_k)^T$ matavimais. Tegu a. v. tankis σ baigtinio mato $\boldsymbol{\mu}$ atžvilgiu, kai $\xi = i$, yra $f_i(\boldsymbol{x})$, t. y.

$$(\mathbf{X} | \xi = i) \sim f_i(\boldsymbol{x}), \quad i = 1, ..., m. \tag{6.2.3}$$

Randomizuota sprendimo priėmimo taisyklė yra mati funkcija $\boldsymbol{\varphi}(\boldsymbol{x}) = (\varphi_1(\boldsymbol{x}), ..., \varphi_m(\boldsymbol{x}))^T$, $\boldsymbol{x} \in \mathbf{R}^{\mathbf{k}}$; čia $\varphi_i(\boldsymbol{x})$ reiškia tikimybę priskirti objektą i-ajai klasei, kai a. v. $\mathbf{X}$ įgijo reikšmę $\boldsymbol{x}$:

$$\varphi_i(\boldsymbol{x}) = \mathbf{P}\{\eta = i | \mathbf{X} = \boldsymbol{x}\}, \quad i = 1, ..., m; \quad \varphi_1(\boldsymbol{x}) + ... + \varphi_m(\boldsymbol{x}) \equiv 1. \tag{6.2.4}$$

Tokių funkcijų klasę žymėsime $\mathbf{\Phi}$. Tikimybės α_{ij} yra tiesiniai $\boldsymbol{\varphi}(\boldsymbol{x})$ funkcionalai

$$\alpha_{ij} = \alpha_{ij}(\boldsymbol{\varphi}) = \int \varphi_i(\boldsymbol{x}) f_j(\boldsymbol{x}) d\boldsymbol{\mu}, \quad i, j = 1, ..., m. \tag{6.2.5}$$

Reikia rasti optimalią sprendimų priėmimo taisyklę $\boldsymbol{\varphi}(\boldsymbol{x})$ iš aibės $\mathbf{\Phi}$.

6.2.2. Sprendimų priėmimo taisyklės

Apibrėžkime nuostolių funkciją. Tarkime, nuostoliai lygūs 0, jeigu priimtas teisingas sprendimas; lygūs $c_{ij} > 0$, kai priimtas sprendimas $\eta = i$, o $\xi = j$, $i \neq j = 1, ..., m$. Tada rizikos funkcija $R = (R_1, ..., R_m)^T$ susideda iš m komponenčių

$$R_i(\boldsymbol{\varphi}) = \sum_{j \neq i} c_{ij} \alpha_{ij}(\boldsymbol{\varphi}) = \int \varphi_i(\boldsymbol{x}) \sum_{j \neq i} c_{ij} f_j(\boldsymbol{x}) d\boldsymbol{\mu}, \quad i = 1, ..., m. \tag{6.2.6}$$

Apriorinės klasių tikimybės žinomos

Jeigu apriorinės klasių tikimybės $\omega_i, i = 1, ..., m$, yra žinomos, tai yra galimybė apibrėžti vidutinę nuostolių funkciją

$$R(\boldsymbol{\varphi}) = \sum_{i=1}^{m} R_i(\boldsymbol{\varphi}) \omega_i = \sum_{i=1}^{m} \int \varphi_i(\boldsymbol{x}) g_i(\boldsymbol{x}) d\boldsymbol{\mu},$$

$$g_i(\boldsymbol{x}) = \sum_{j \neq i} c_{ij} f_j(\boldsymbol{x}) \omega_j, \quad i = 1, ..., m,$$

ir suformuluoti optimalios sprendimų priėmimo taisyklės $\boldsymbol{\varphi}^*$ radimo uždavinį:

$$\sup_{\boldsymbol{\varphi} \in \mathbf{\Phi}} R(\boldsymbol{\varphi}) = R(\boldsymbol{\varphi}^*). \tag{6.2.7}$$

Apibrėžkime sprendimų priėmimo funkciją $\boldsymbol{\varphi}^* = (\varphi_1^*, ..., \varphi_m^*)$ taip

$$\varphi_i^*(\boldsymbol{x}) = 1, \quad \text{kai} \quad g_i(\boldsymbol{x}) < g_j(\boldsymbol{x}), \quad \forall \ j \neq i, \quad i = 1, ..., m. \tag{6.2.8}$$

Jeigu kokiame nors taške $\boldsymbol{x}$ yra keletas sutampančių minimalių funkcijų $g_i(\boldsymbol{x})$ reikšmių, pavyzdžiui,

$$g_{i_1}(\boldsymbol{x}) = ... = g_{i_r}(\boldsymbol{x}) < g_{i_{r+1}}(\boldsymbol{x}) \leq g_{i_{r+2}}(\boldsymbol{x}) \leq ... \leq g_{i_m}(\boldsymbol{x}), \quad 2 \leq r \leq m,$$

o $(i_1, ..., i_m)$ yra skaičių $(1, ..., m)$ kėlinys, tai $\boldsymbol{\varphi}^*$ apibrėžkime taip:

$$0 \leq \varphi_{i_j}^*(\boldsymbol{x}) \leq 1, \quad \varphi_{i_1}^*(\boldsymbol{x}) + ... + \varphi_{i_r}^*(\boldsymbol{x}) = 1, \quad \varphi_{i_{r+1}}^*(\boldsymbol{x}) = ... = \varphi_{i_m}^*(\boldsymbol{x}) = 0. \tag{6.2.9}$$

6.2.1 teorema. *Funkcija $\boldsymbol{\varphi}^*$, apibrėžta (6.2.8), (6.2.9) formulėmis, minimizuoja funkcionalą $R(\boldsymbol{\varphi})$. Jeigu $\mathbf{P}\{g_j(\mathbf{X}) = g_r(\mathbf{X}) | \xi = i\} = 0, j \neq r; i, j, r = 1, ..., m$, tai $\boldsymbol{\varphi}^*$ nusakyta vienareikšmiškai, išskyrus, galbūt nulinio $\boldsymbol{\mu}$ mato aibę.*

Įrodymas. Integralas (6.2.7) minimalus, jei pointegralinis reiškinys minimalus kiekviename taške $\boldsymbol{x}$. Pagal funkcijos $\boldsymbol{\varphi}^*$ parinkimą akivaizdu, kad ji minimizuoja (6.2.7) pointegralinį reiškinį kiekviename taške $\boldsymbol{x}$. ▲

Jeigu kainos $c_{ij} = 1,\ i \neq j = 1, ..., m$, tai sprendimų priėmimo funkciją (6.2.8), (6.2.9) galima perrašyti taip:

$$\varphi_i^*(\boldsymbol{x}) = 1, \quad \text{kai} \quad f_i(\boldsymbol{x})\omega_i > f_j(\boldsymbol{x})\omega_j, \quad \forall\ j \neq i, \quad i = 1, ..., m. \tag{6.2.10}$$

Jei kokiame nors taške $\boldsymbol{x}$ yra keletas sutampančių maksimalių funkcijų $f_i(\boldsymbol{x})\omega_i$ reikšmių, pavyzdžiui,

$$f_{i_1}(\boldsymbol{x})\omega_{i_1} = ... = f_{i_r}(\boldsymbol{x})\omega_{i_r} > f_{i_{r+1}}(\boldsymbol{x})\omega_{i_{r+1}} \geq f_{i_{r+2}}(\boldsymbol{x})\omega_{i_{r+2}} \geq ... \geq f_{i_m}(\boldsymbol{x})\omega_{i_m},$$

o $(i_1, ..., i_m)$ yra skaičių $(1, ..., m)$ kėlinys, tai $\boldsymbol{\varphi}^*$ apibrėžkime taip:

$$0 \leq \varphi_{i_j}^*(\boldsymbol{x}) \leq 1, \quad \varphi_{i_1}^*(\boldsymbol{x}) + ... + \varphi_{i_r}^*(\boldsymbol{x}) = 1, \quad \varphi_{i_{r+1}}^*(\boldsymbol{x}) = ... = \varphi_{i_m}^*(\boldsymbol{x}) = 0. \tag{6.2.11}$$

Tokia sprendimų priėmimo taisyklė vadinama *Bejeso taisykle* apriorinio skirstinio $(\omega_1, ..., \omega_m)$ atžvilgiu. Ji minimizuoja suminę svertinę klasifikavimo klaidą

$$\sum_{i \neq j} \alpha_{ij}\omega_j.$$

Bejeso klasifikavimo taisyklę galima gauti ir kitokiu būdu. Neturint jokios papildomos informacijos, objektą natūralu priskirti i-ajai klasei, jeigu apriorinė tikimybė $\omega_i > \omega_j,\ \ \forall\ j \neq i$. Jeigu gauta a. v. $\mathbf{X}$ reikšmė $\mathbf{X} = \boldsymbol{x}$, tai remdamiesi Bejeso formule apskaičiuojame aposteriorines klasių tikimybes

$$\omega_i(\boldsymbol{x}) = \mathbf{P}\{\xi = i | \mathbf{X} = \boldsymbol{x}\} = \frac{f_i(\boldsymbol{x})\omega_i}{f_1(\boldsymbol{x})\omega_1 + ... + f_m(\boldsymbol{x})\omega_m}, \quad i = 1, ..., m; \tag{6.2.12}$$

$$\omega_1(\boldsymbol{x}) + ... + \omega_m(\boldsymbol{x}) = 1.$$

Objektą, kurio gauta reikšmė $\mathbf{X} = \boldsymbol{x}$, priskirsime i-ajai klasei, jeigu

$$\omega_i(\boldsymbol{x}) > \omega_j(\boldsymbol{x}),\ \ \forall\ j \neq i. \tag{6.2.13}$$

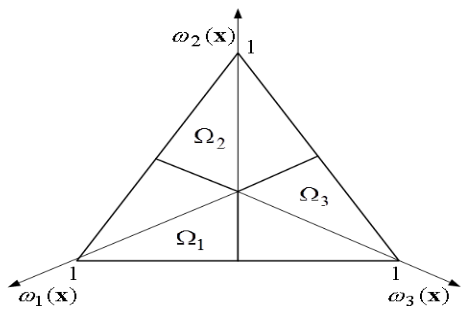
Akivaizdu, kad ši taisyklė sutampa su (6.2.10), (6.2.11).

Apriorinės tikimybės $\omega_1(\boldsymbol{x}), ..., \omega_m(\boldsymbol{x})$ įgyja reikšmes simplekse

$$\{(\omega_1, ..., \omega_m) : 0 \leq \omega_i \leq 1, \omega_1 + ... + \omega_m = 1\},$$

kurį hiperplokštumos (6.2.13) sudalija į m nesikertančių sričių $\Omega_1, ..., \Omega_m$. Objektas priskiriamas i-ajai klasei, jeigu vektorius $(\omega_1(\boldsymbol{x}), ..., \omega_m(\boldsymbol{x}))^T \in \Omega_i,\ i = 1, ..., m$. Taigi kintamųjų $\omega_i(\boldsymbol{x}), i = 1, ..., m$, terminais sprendimų priėmimo sritys turi gana paprastą pavidalą, o kintamųjų $\boldsymbol{x}$ terminais jos gali būti kur kas sudėtingesnės (žr. skyrelį 6.1.4).

Pavyzdžiui, trijų klasių aposteriorinės tikimybės $(\omega_1(\boldsymbol{x}), \omega_2(\boldsymbol{x}), \omega_3(\boldsymbol{x}))$ įgyja reikšmes taisyklingame trikampyje, kuris gaunamas susikertant plokštumai, einančiai per taškus (1, 0, 0), (0, 1, 0) ir (0, 0, 1), su koordinačių plokštumomis. Dėl simetrijos suprojektavę koordinačių ašis į šį trikampį, gausime vadinamąją centrinę koordinačių sistemą. Ši koordinačių sistema ir Bejeso sprendimų priėmimo taisyklių sritys $\Omega_1, \Omega_2, \Omega_3$ pavaizduotos **6.2.1** paveiksle.



6.2.1 pav. Sprendimų priėmimo sritys

6.2.1 pavyzdys. Panagrinėsime paprastą iliustracinį pavyzdį. Tarkime, $m = 3$. Klasifikavimą atliekame remdamiesi a. d. X matavimu; $(X|\xi = 1) \sim N(0,\ 1)$, $(X|\xi = 2) \sim N(3,\ 1)$, o kai $\xi = 3$ a. d. X skirstinys yra mišinys normaliųjų skirstinių $N(-2,\ 1)$ ir $N(5,\ 1)$ su svoriais 1/2. Tegu apriorinės tikimybės $\omega_1 = \omega_2 = \omega_3 = 1/3$ vienodos. Nuostoliai $c_{12} = c_{21} = 2$, $c_{13} = c_{23} = 4$, $c_{31} = c_{32} = 3$.

Rasime sprendimų priėmimo taisyklę $\boldsymbol{\varphi}^*$, minimizuojančią vidutinius nuostolius $R(\boldsymbol{\varphi})$.

Randame

$$g_1(x) = 2(\phi(x+2) + \phi(x-3) + \phi(x-5))/3, \quad g_2(x) = 2(\phi(x) + \phi(x+2) + \phi(x-5))/3,$$
$$g_3(x) = \phi(x) + \phi(x-3).$$

Sulyginę šias diskriminantines funkcijas, gauname sprendinį

$$\varphi_1^*(x) = 1, \quad \text{kai} \quad -1,203 < x < 1,5, \quad \varphi_2^*(x) = 1, \quad \text{kai} \quad 1,5 < x < 4,203,$$
$$\varphi_3^*(x) = 1 - \varphi_1^*(x) - \varphi_2^*(x).$$

Minimali rizikos funkcijos reikšmė yra $R(\boldsymbol{\varphi}^*) = 0,6399$. Šios taisyklės klasifikavimo klaida yra 0,1918.

Rasime Bejeso taisyklę, minimizuojančią suminę klasifikavimo klaidą. Remdamiesi (6.2.10) gauname tokią taisyklę:

$$\varphi_1^*(x) = 1, \quad \text{kai} \quad -1,3466 < x < 1,5, \quad \varphi_2^*(x) = 1, \quad \text{kai} \quad 1,5 < x < 4,3466,$$
$$\varphi_3^*(x) = 1 - \varphi_1^*(x) - \varphi_2^*(x).$$

Minimali klasifikavimo klaida 0,1894.

Minimakso principas

Jeigu apriorinės klasių tikimybės $\omega_i,\ i = 1, ..., m$, nežinomos, tai parenkant optimalią sprendimų priėmimo taisyklę reikia atsižvelgti į visas rizikos funkcijos komponentes $R_1(\boldsymbol{\varphi}), ..., R_m(\boldsymbol{\varphi})$. Kaip ir atveju, kai yra dvi klasės, tuo tikslu galime remtis minimakso principu, t. y. ieškoti tokios taisyklės, kuri minimizuotų maksimaliąją rizikos funkcijos komponentę

$$\min_{\boldsymbol{\varphi} \in \boldsymbol{\Phi}} (\max(R_1(\boldsymbol{\varphi}), ..., R_m(\boldsymbol{\varphi}))).$$

Pasirodo, kad tokio uždavinio sprendinys yra Bejeso taisyklė (6.2.10), (6.2.11), atitinkanti tam tikrą (mažiausiai palankų) apriorinių tikimybių rinkinį $\omega_1^*, ..., \omega_m^*$.

6.2.2 teorema. *Tarkime, kad egzistuoja toks apriorinių tikimybių vektorius* $(\omega_1^*, ..., \omega_m^*)^T$ *su teigiamomis koordinatėmis, kad jį atitinkančio Bejeso sprendinio* $\boldsymbol{\varphi}^*$ *rizikos funkcijos komponentės* $R_1(\boldsymbol{\varphi}^*) = ... = R_m(\boldsymbol{\varphi}^*)$ *yra vienodos. Tada sprendinys* $\boldsymbol{\varphi}^*$ *tenkina minimakso principą.*

Įrodymas. Įrodymą galima rasti knygoje [14]. ▲

6.2.2 pavyzdys. Tarkime, analogiškai **6.2.1** pavyzdžiui, $(X|\xi = 1) \sim N(0,\ 1)$, $(X|\xi = 2) \sim N(3,\ 1)$, o kai $\xi = 3$ a.d. X skirstinys yra normaliųjų skirstinių $N(-2,\ 1)$ ir $N(5,\ 1)$ mišinys su svoriais 1/2. Tegu klaidingų sprendimų pasekmės įvertinamos taip: $c_{12} = c_{21} = 2$, $c_{13} = c_{31} = c_{32} = c_{23} = 3$, o apriorinės klasių tikimybės nežinomos. Rasime sprendinį $\boldsymbol{\varphi}$, tenkinantį minimakso primcipą.

Bejeso taisyklė, atitinkanti apriorinį skirstinį $\omega_1, \omega_2, 1 - \omega_1 - \omega_2$, nusakoma sprendimų priėmimo sritimis

$$\Omega_1 = \{x : a_1 < x < a_2\}, \quad \Omega_2 = \{x : a_2 < x < a_3\}, \quad \Omega_3 = \{x : x < a_1, \text{arba } x > a_3\},$$

čia

$$a_1 = \frac{1}{2}\ln\frac{1 - \omega_1 - \omega_2}{2\omega_1}, \quad a_2 = \frac{3}{2} - \ln\frac{\omega_2}{\omega_1}, \quad a_3 = 3 - a_1.$$

Prilyginę rizikos funkcijos komponentes vieną kitai, gausime dviejų lygčių sistemą tikimybių ω_1^*, ω_2^* atžvilgiu. Išsprendę gauname $\omega_1^* = \omega_2^* = 0,27585$, $\omega_3^* = 0,4483$. Su šiuo aprioriniu skirstiniu rizikos funkcijos komponentės yra vienodos ir lygios 0,5318. Taisyklės $\boldsymbol{\varphi}^*$ suminė svertinė klasifikavimo klaida yra 0,1916.

6.2.1 pastaba. Klasifikavimo į keletą klasių uždaviniai, kai yra apribojimų ir kai yra galimybė atsisakyti priimti galutinį sprendimą, detaliau nagrinėjami knygoje [11].

6.2.3. Normaliojo skirstinio atvejis

Tarkime, kad a.v. $\mathbf{X} = (X_1, ..., X_k)^T$ skirstinys, kai objektas priklauso i-ajai klasei yra normalusis $N_k(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), i = 1, ..., m$. Tankio funkcija

$$f_i(\boldsymbol{x}) = f_i(\boldsymbol{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) = (2\pi)^{-k/2}|\boldsymbol{\Sigma}_i|^{-1/2}\exp\{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu}_i)^T\boldsymbol{\Sigma}_i^{-1}(\boldsymbol{x} - \boldsymbol{\mu}_i)\}.$$

Tegu nuostolių funkcijos konstantos $c_{ii} = 0, i = 1, ..., m$, $c_{ij} = 1, i \neq j = 1, ...m$. Tada (6.2.7) uždavinys reiškia, kad minimizuojama suminė svertinė klasifikavimo klaida. Remiantis 6.2.1 teorema priimamas sprendimas $\eta = i$, kai

$$f_i(\boldsymbol{x})\omega_i > f_j(\boldsymbol{x})\omega_j, \quad \forall j \neq i, \quad j = 1, ..., m.$$

Logaritmuodami gauname, kad sprendimas $\eta = i$ priimamas, kai

$$g_i(\boldsymbol{x}) > g_j(\boldsymbol{x}), \quad \forall j \neq i, \quad j = 1, ..., m, \tag{6.2.14}$$

$$g_i(\boldsymbol{x}) = -\frac{1}{2}\ln|\boldsymbol{\Sigma}_i| - \frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu}_i)^T\boldsymbol{\Sigma}_i^{-1}(\boldsymbol{x} - \boldsymbol{\mu}_i) + \ln\omega_i, \quad i = 1, ..., m.$$

Sprendimų priėmimo sritys (6.2.14) gali turėti sudėtingą pavidalą. Todėl patogiau atsakymą užrašyti aposteriorinių klasių tikimybių terminais. Sprendimas $\eta = i$ priimamas, kai

$$\omega_i(\boldsymbol{x}) > \omega_j(\boldsymbol{x}), \quad \forall j \neq i, \quad j = 1, ..., m, \tag{6.2.15}$$

$$\omega_i(\boldsymbol{x}) = \frac{f_i(\boldsymbol{x})\omega_i}{f_1(\boldsymbol{x})\omega_1 + ... + f_m(\boldsymbol{x})\omega_m}, \quad i = 1, ..., m.$$

Tikimybių $\omega_i(\boldsymbol{x})$ terminais klasės atskiriamos tiesinėmis funkcijomis.

Panagrinėkime atvejį, kai kovariacinės matricos $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = ... = \boldsymbol{\Sigma}_m = \boldsymbol{\Sigma}$ yra vienodos. Tada nelygybės (6.2.14) reiškia, kad reikia palyginti tiesines diskriminantines funkcijas

$$g_i(\boldsymbol{x}) = \boldsymbol{\mu}_i^T \boldsymbol{\Sigma}^{-1} \boldsymbol{x} - \frac{1}{2} \boldsymbol{\mu}_i^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_i + \ln \omega_i, \quad i = 1, ..., m.$$

Sprendimas $\eta = i$ priimamas, kai

$$g_i(\boldsymbol{x}) > g_j(\boldsymbol{x}), \quad \forall j \neq i, \quad j = 1, ..., m. \tag{6.2.16}$$

Šios nelygybės sudalija erdvę $\mathbf{R}^k$ į m sričių $\Omega_1, \Omega_2, ..., \Omega_m$. Objektas priskiriamas i-ajai klasei, jei stebėtoji reikšmė $\mathbf{X} = \boldsymbol{x}$ patenka į sritį Ω_i. Sprendinys nusakytas vienareikšmiškai beveik visur Lebego mato prasme, nes $\mathbf{P}\{g_i(\mathbf{X}) - g_j(\mathbf{X}) = 0 | \xi = l\} = 0$, $i \neq j,\ i, j, l = 1, ..., m$. Sritį Ω_i apriboja $m-1$ hiperplokštuma

$$g_i(\boldsymbol{x}) - g_j(\boldsymbol{x}) = (\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^T \boldsymbol{\Sigma}^{-1} \boldsymbol{x} - \frac{1}{2}(\boldsymbol{\mu}_i^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_i - \boldsymbol{\mu}_j^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_j) + \ln \frac{\omega_i}{\omega_j}$$

$$= (\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^T \boldsymbol{\Sigma}^{-1} \boldsymbol{x} - \frac{1}{2}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_i + \boldsymbol{\mu}_j) + \ln \frac{\omega_i}{\omega_j} = 0, j \neq i = 1, ..., m. \tag{6.2.17}$$

6.3. Klasifikavimas neturint visos informacijos

Pirmesniame skyrelyje aptarėme klasifikavimo uždavinius, kai turima visa informacija, t. y. žinomi kiekvienos klasės a. v. $\mathbf{X}$ tankiai $(\mathbf{X}|\xi = i) \sim f_i(\boldsymbol{x}), i = 1, ..., m$, ir apriorinės klasių tikimybės $\omega_i = \mathbf{P}\{\xi = i\}, i = 1, ..., m$.

Dažniausiai nei tankiai, nei apriorinės tikimybės nežinomos. Tipišku atveju turimos vadinamosios *apmokančiosios imtys* $\mathbf{X}_1^{(i)}, ..., \mathbf{X}_{n_i}^{(i)}, \quad i = 1, ..., m$. Yra žinoma, kad imtis $\mathbf{X}_1^{(i)}, ..., \mathbf{X}_{n_i}^{(i)}$ gauta stebint a. v. $\mathbf{X}$, kai objektai priklausė i-ajai klasei, t. y. kai $\xi = i$. Naudojant apmokančiąsias imtis stengiamasi įvertinti nežinomus tankius $f_i(\boldsymbol{x}), i = 1, ..., m$, ir apriorines tikimybes $\omega_1, ..., \omega_m$, o paskui gautus įvertinius naudoti įrašant į klasifikavimo taisyklę vietoje nežinomų tankių ir apriorinių tikimybių.

Rasti apriorinių tikimybių įvertinius dažniausiai nebūna sudėtinga. Pavyzdžiui, jei komplektuojant apmokančiąsias imtis objektas iš i-osios klasės pasirodo su tikimybe ω_i nepriklausomai nuo kitų objektų, tai tikimybę galima įvertinti santykiniu dažniu

$$\hat{\omega}_i = \frac{n_i}{n_1 + ... + n_m}, \quad i = 1, ..., m. \tag{6.3.1}$$

Jeigu stebėjimų skaičius $n = n_1 + ... + n_m$ fiksuotas, tai a. d. n_i turi binominius skirstinius $n_i \sim B(n, \omega_i)$. Taigi

$$\mathbf{E}\hat{\omega}_i = \omega_i, \quad \boldsymbol{V}\hat{\omega}_i = \frac{\omega_i(1-\omega_i)}{n}, \quad \hat{\omega}_i \xrightarrow{b.t.} \omega_i, \quad n \to \infty. \tag{6.3.2}$$

Kur kas sudėtingiau gauti gana tikslius tankių $f_i(\boldsymbol{x}), i = 1, ..., m$ įvertinius, nes tam dažniausiai nepakanka turimų apmokančiųjų imčių didumų. Uždavinys šiek tiek supaprastėja, kai yra pakankamos prielaidos tarti, kad tankių $f_i(\boldsymbol{x}), i = 1, ..., m$ funkcinė išraiška žinoma, nežinomi tik parametrai $\boldsymbol{\theta}_i$, nuo kurių priklauso tankiai $f_i(\boldsymbol{x}) = f_i(\boldsymbol{x}|\boldsymbol{\theta}_i), i = 1, ..., m$. Gavę parametro $\boldsymbol{\theta}_i$ įvertinį $\hat{\boldsymbol{\theta}}_i$, kaip tankio įvertinį galime imti $\hat{f}_i(\boldsymbol{x}) = f_i(\boldsymbol{x}|\hat{\boldsymbol{\theta}}_i)$; jį ir naudoti klasifikatoriuose vietoje nežinomo tankio $f_i(\boldsymbol{x})$. Suprantama, tada klasifikavimo tikslumas priklausys ne tik nuo sprendimų priėmimo taisyklės, bet ir nuo įvertinių $\hat{\boldsymbol{\theta}}_i$ tikslumo.

Jeigu imtyje $\mathbf{X}_i^{(i)}, ..., \mathbf{X}_{n_i}^{(i)}$ nėra jokios informacijos apie parametrą $\boldsymbol{\theta}_j, j \neq i$, tai vertindami parametrą $\boldsymbol{\theta}_i$ naudojame tik šią imtį, t. y. turime klasikinį parametrų vertinimo uždavinį, kurį sprendžiame atskirai kiekvienai klasei. Tiksliau, turime imtį $\mathbf{X}_1, ..., \mathbf{X}_n$, gautą stebint a. v. $\mathbf{X}$, kurio tankis $f(\boldsymbol{x}|\boldsymbol{\theta})$ priklauso nuo nežinomo parametro $\boldsymbol{\theta}$ (klasės indeksą praleidome). Remiantis turima imtimi reikia rasti parametro $\boldsymbol{\theta}$ įvertinį.

Parametro $\boldsymbol{\theta}$ vertinimo metodus (momentų ir didžiausiojo tikėtinumo) ir gautųjų įvertinių savybes nagrinėjome 1 dalies 4 skyriuje. Klasifikavimo uždaviniuose vertinant parametrą $\boldsymbol{\theta}$ dažnai naudojamas ir Bejeso metodas.

Jeigu turimos informacijos nepakanka, kad būtų galima patikimai parinkti konkretų tankio pavidalą, tada galima naudoti neparametrinius tankių įvertinius histogramos ar branduolinių įvertinių pavidalo (žr. 1 dalies 3 skyrių).

Diskriminantinėje analizėje dažnai naudojami metodai, kurie nereikalauja tankių įvertinių radimo. Jų esmė yra tokia: tariama, kad klasifikatoriaus diskriminantinės funkcijos turi konkretų (dažniausiai tiesinį) pavidalą, priklausantį nuo nežinomų parametrų. Parametrai vertinami pagal apmokančiąsias imtis naudojant tam tikrus kriterijus.

6.3.1. Parametriniai tankių įvertiniai

6.3.1.1. Didžiausiojo tikėtinumo įvertiniai

Pailiustruosime DT metodo taikymą vertinant tankius. Tarkime, kad a. v. $\mathbf{X}$ skirstinys kiekvienoje klasėje yra normalusis $(\mathbf{X}|\xi = i) \sim N_k(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, $i = 1, ..., m$. Tada nežinomo parametro $\boldsymbol{\theta}_i = (\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ nepaslinktasis įvertinys (žr.

6.1 skyrelį) yra

$$\hat{\boldsymbol{\mu}}_i = \bar{\mathbf{X}}^{(i)} = \frac{1}{n_i}\sum_{j=1}^{n_i}\mathbf{X}_j^{(i)}, \quad \hat{\mathbf{\Sigma}}_i = \boldsymbol{S}_i = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}(\mathbf{X}_j^{(i)} - \bar{\mathbf{X}}^{(i)})(\mathbf{X}_j^{(i)} - \bar{\mathbf{X}}^{(i)})^T. \tag{6.3.3}$$

Įstatę šiuos įvertinius į (6.1.39), gausime klasifikavimo taisyklės įvertinį.

Jeigu yra pagrindo tvirtinti, kad kovariacijų matricos vienodos $\mathbf{\Sigma}_1 = ... = \mathbf{\Sigma}_m = \mathbf{\Sigma}$, tai $\mathbf{\Sigma}$ įvertiniui sudaryti natūralu panaudoti visas imtis

$$\hat{\mathbf{\Sigma}} = \boldsymbol{S} = \frac{1}{n-m}\sum_{i=1}^{m}(n_i - 1)\boldsymbol{S}_i, \quad n = n_1 + ... + n_m.$$

Kai yra dvi klasės, vietoje diskriminantinės funkcijos (6.1.41) gauname jos įvertinį

$$\hat{g}(\boldsymbol{x}) = (\bar{\mathbf{X}}^{(1)} - \bar{\mathbf{X}}^{(2)})^T\hat{\mathbf{\Sigma}}^{-1}\boldsymbol{x} - \frac{1}{2}((\bar{\mathbf{X}}^{(1)})^T\hat{\mathbf{\Sigma}}^{-1}\bar{\mathbf{X}}^{(1)} - (\bar{\mathbf{X}}^{(2)})^T\hat{\mathbf{\Sigma}}^{-1}\bar{\mathbf{X}}^{(2)}) + d$$

$$= (\bar{\mathbf{X}}^{(1)} - \bar{\mathbf{X}}^{(2)})^T\hat{\mathbf{\Sigma}}^{-1}\boldsymbol{x} - \frac{1}{2}(\bar{\mathbf{X}}^{(1)} - \bar{\mathbf{X}}^{(2)})^T\hat{\mathbf{\Sigma}}^{-1}(\bar{\mathbf{X}}^{(1)} + \bar{\mathbf{X}}^{(2)}) + d. \tag{6.3.4}$$

6.3.1 teorema. *Jeigu apmokančiųjų imčių didumai* $n_1 \to \infty$, $n_2 \to \infty$ *ir* $|\mathbf{\Sigma}| > 0$, *tai*

$$\hat{g}(\boldsymbol{x}) \xrightarrow{b.t.} g(\boldsymbol{x}), \quad (\hat{g}(\mathbf{X})|\xi = 1) \xrightarrow{d} Z_1 \sim N(-a/2 + d, a), \tag{6.3.5}$$

$$(\hat{g}(\mathbf{X})|\xi = 2) \xrightarrow{d} Z_2 \sim N(a/2 + d, a), \quad a = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T\mathbf{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2).$$

Įrodymas. Remiantis empirinių momentų savybėmis (1 dalis, 3 skyrelis), galima tvirtinti, kad

$$\bar{\mathbf{X}}^{(1)} \xrightarrow{b.t.} \boldsymbol{\mu}_1, \quad \bar{\mathbf{X}}^{(2)} \xrightarrow{b.t.} \boldsymbol{\mu}_2, \quad \hat{\mathbf{\Sigma}}^{-1} \xrightarrow{b.t.} \mathbf{\Sigma}^{-1}.$$

Remiantis teoremomis apie sekų konvergavimą (1 dalis, 4.5.4 skyrelis), galima tvirtinti, kad

$$\hat{g}(\boldsymbol{x}) \xrightarrow{b.t.} g(\boldsymbol{x}), \quad (\hat{g}(\mathbf{X})|\xi = i) \xrightarrow{d} (g(\mathbf{X})|\xi = i), \quad i = 1,\ 2. \qquad \blacktriangle$$

Tokiu būdu, kai yra pakankamai dideli n_1 ir n_2, klasifikavimo taisyklė ir klasifikavimo klaidos yra apytiksliai tokios pačios kaip ir tuo atveju, kai parametrai $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \mathbf{\Sigma}$ yra žinomi.

Kai yra dvi klasės, o a. v. $\mathbf{X}$ skirstiniai yra normalieji su vienoda kovariacijų matrica ir apriorinės tikimybės lygios, suminė klasifikavimo klaida (6.1.42) yra

$$P\{e\} = 1 - \Phi(\sqrt{a}/2), \quad a = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T\mathbf{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2). \tag{6.3.6}$$

Taigi $P\{e\} \to 0$, kai $a \to \infty$, Jeigu a. v. $\mathbf{X}$ koordinatės nepriklausomos, tai

$$a = \sum_{i=1}^{k} (\mu_{i1} - \mu_{i2})^2 / \sigma_i^2; \tag{6.3.7}$$

čia μ_{i1} ir μ_{i2} yra i-osios koordinatės vidurkis, kai objektas atitinkamai priklauso pirmai ir antrai klasei, o σ_i^2 yra i-osios koordinatės dispersija.

Iš čia matome, kad klaidai (6.3.7) mažėti didžiausią įtaką turi požymiai, kurių skirtumas $|\mu_{i1} - \mu_{i2}|$, palyginti su σ_i, yra didžiausias. Kita vertus, bet koks požymis, kurio $|\mu_{i1} - \mu_{i2}| \neq 0$, naudingas klasifikavimo klaidai mažinti. Atrodytų, kad, norint pagerinti klasifikavimo tikslumą, reikia į klasifikatorių įtraukti kuo daugiau požymių, t. y. padidinti a. v. $\mathbf{X}$ dimensiją.

Tačiau padėtis iš esmės pasikeičia, kai tankiai nežinomi ir juos tenka vertinti iš apmokančiųjų imčių. Naujų požymių įtraukimas didina vertinamų parametrų skaičių, tada, užuot patikslinę klasifikavimą, galime gauti priešingą rezultatą. Pavyzdžiui, k-mačio normaliojo skirstinio atveju, kai jį visiškai apibūdina pirmieji ir antrieji momentai, reikia įvertinti k vidurkių ir $k(k+1)/2$ kovariacinės matricos elementų. Vektoriaus $\mathbf{X}$ dimensiją padidinus vienetu reikalingų įvertinti parametrų skaičių padidėja $k+2$. Taigi, labai padidinus a. v. $\mathbf{X}$ dimensiją ir norint pakankamai tiksliai įvertinti visus parametrus, apmokančiųjų imčių didumai turėtų padidėti tiek, kad jas gauti būtų beveik nerealu.

Sumažinti stebimo vektoriaus dimensiją yra viena iš svarbiausiųjų daugiamatės matematinės statistikos problemų, sprendžiant įvairius uždavinius (klasifikavimas, regresinė analizė ir kt.). Šios problemos nagrinėjamos daugelyje knygų (žr., pvz.,[13], [14], [15]).

Sprendžiant klasifikavimo uždavinį, taikomi įvairūs vertinamų parametrų skaičiaus mažinimo metodai: atmetami tie požymiai, kuriems $|\hat{\mu}_{i1} - \hat{\mu}_{i2}|$, palyginti su $\hat{\sigma}_i$, yra mažas; nuliui prilyginami tie koreliacijos koeficientai,kurių įverčiai įgijo mažas reikšmes; tariama, kad kovariacinė matrica yra diagonalioji, t. y. vertinami tik vidurkiai ir dispersijos; naudojamos ne a. v. $\mathbf{X}$ koordinatės, o jų funkcijos (dažniausiai tiesinės), paliekant mažesnį jų skaičių negu a. v. dimensija (pvz., imama keletas pirmųjų pagrindinių komponenčių; žr. 6.7 skyrelį), ir kt.

6.3.1.2. Bejeso metodas tankiams vertinti

Kartais dar prieš gaunant apmokančiąsias imtis turima tam tikra išankstinė (apriorinė) informacija apie nežinomus parametrus $\boldsymbol{\theta}_i$. Sudarant klasifikatorių natūralu panaudoti ne tik apmokančiąsias imtis, bet ir šią išankstinę informaciją. Pavyzdžiui, tarkime, kad ilgą laiką atliekant išleidžiamąją produkcijos kontrolę nustatyta, kad defektinių gaminių dalis p vidutiniškai lygi $p_0 = 0,02$ ir niekada neviršija $0,1$. Todėl vietoje parametro p kitimo srities intervalo $[0,\, 1]$ galima apsiriboti siauresne kitimo sritimi $[0,\, 0,1]$.

Tarsime, kad vertinant i-osios klasės parametrą $\boldsymbol{\theta}_i$, kitų klasių $j \neq i$ apmokančiosios imtys jokios informacijos nesuteikia, t. y. klasės indeksą praleisime.

Paprastai išankstinė turima informacija formalizuojama taip. Tariama, kad parametras $\boldsymbol{\theta}$ turi apriorinį skirstinį, nusakytą tokiu tankiu $g(\boldsymbol{\theta})$ (σ-baigtinio mato $\boldsymbol{\nu}$ atžvilgiu), kad jo tikimybinė masė sukoncentruota taško $\boldsymbol{\theta}_0 = \mathbf{E}\boldsymbol{\theta}$ aplinkoje. Sklaida apie tašką $\boldsymbol{\theta}_0$ priklauso nuo informacijos tikslumo ir patikimumo: kuo tikslesnė išankstinė informacija, tuo mažesnė sklaida apie tašką $\boldsymbol{\theta}_0$ parenkama. Tegu apmokančioji imtis $\mathcal{X} = (\mathbf{X}_1, ..., \mathbf{X}_n)$ yra paprastoji imtis a. v. $\mathbf{X}$, kurio sąlyginis tankis (σ-baigtinio mato $\boldsymbol{\mu}$ atžvilgiu), kai $\boldsymbol{\theta}$ fiksuotas, yra $(\mathbf{X}|\boldsymbol{\theta}) \sim f(\boldsymbol{x}|\boldsymbol{\theta})$. Tada imties sąlyginis tankis

$$L(\boldsymbol{x}_1, ..., \boldsymbol{x}_n|\boldsymbol{\theta}) = \prod_{j=1}^{n} f(\boldsymbol{x}_j|\boldsymbol{\theta}). \tag{6.3.8}$$

Pagal Bejeso formulę gauname aposteriorinį parametro $\boldsymbol{\theta}$ tankį

$$g(\boldsymbol{\theta}|\mathcal{X}) = \frac{1}{c} L(\mathbf{X}_1, ..., \mathbf{X}_n|\boldsymbol{\theta}) g(\boldsymbol{\theta}), \tag{6.3.9}$$

čia c nepriklausanti nuo $\boldsymbol{\theta}$ normuojanti konstanta

$$c = \int L(\mathbf{X}_1, ..., \mathbf{X}_n|\boldsymbol{\theta}) g(\boldsymbol{\theta}) d\boldsymbol{\nu}.$$

Parametro $\boldsymbol{\theta}$ Bejeso įvertiniu imame vidurkį

$$\hat{\boldsymbol{\theta}}_B = \mathbf{E}(\boldsymbol{\theta}|\mathcal{X}) = \int \boldsymbol{\theta} g(\boldsymbol{\theta}|\mathcal{X}) d\boldsymbol{\nu}. \tag{6.3.10}$$

Naudodami tankį (6.3.9) galime gauti a. v. $\mathbf{X}$ sąlyginį tankį apmokančiosios imties $\mathcal{X}$ atžvilgiu:

$$f(\boldsymbol{x}|\mathcal{X}) = \int f(\boldsymbol{x}|\boldsymbol{\theta}) g(\boldsymbol{\theta}|\mathcal{X}) d\boldsymbol{\nu}, \tag{6.3.11}$$

jį ir galime naudoti sudarydami klasifikatorių.

6.3.1 pavyzdys. Tarkime, $\mathcal{X} = (X_1, ..., X_n)$ yra paprastoji imtis, kurios elementai, kai parametras p fiksuotas, turi Bernulio skirstinius $B(1,\ p)$, o p savo ruožtu yra a. d., turintis beta skirstinį $p \sim Be(\gamma, \eta)$, $p_0 = \mathbf{E}p = \gamma/(\gamma+\eta)$. Tada, pažymėję $S_n = X_1 + ... + X_n$, gauname

$$L(X_1, ..., X_n|p) = p^{S_n}(1-p)^{n-S_n}, \quad g(p) = \frac{\Gamma(\gamma+\eta)}{\Gamma(\gamma)\Gamma(\eta)} p^{\gamma-1}(1-p)^{\eta-1}$$

ir

$$g(p|\mathcal{X}) = \frac{\Gamma(\gamma+\eta+n)}{\Gamma(\gamma+S_n)\Gamma(\eta+n-S_n)} p^{\gamma+S_n-1}(1-p)^{\eta+n-S_n-1}$$

yra beta skirstinys su parametrais $\gamma + S_n$ ir $\eta + n - S_n$.

Parametro p Bejeso įvertinys

$$\hat{p}_B = \frac{\gamma+S_n}{\gamma+\eta+n} = \frac{\gamma}{\gamma+\eta}\frac{\gamma+\eta}{\gamma+\eta+n} + \bar{X}\frac{n}{\gamma+\eta+n}. \tag{6.3.12}$$

Matome, kad $\hat{p}_B$ susideda iš dviejų komponenčių. Pirmoji rodo apriorinės informacijos, o antroji apmokančiosios imties indėlį. Jeigu γ ir η fiksuoti, o $n \to \infty$, tai $\hat{p}_B$ artėja į DT įvertinį $\bar{X}$, t. y. apriorinės informacijos indėlis artėja į 0. Atvirkščiai, jeigu n ir $p_0 = \gamma/(\gamma+\eta)$ fiksuoti, o $\eta \to \infty$ (sklaida apie p_0 mažėja), tai $\hat{p}_B$ artėja į p_0, t. y. į įvertinį, kurį gautume, jei apmokančiosios imties nebūtų. Kitais atvejais gauname tarpinius variantus, kai panaudojama ir išankstinė informacija, ir apmokančioji imtis.

Pagal (6.3.11) sąlyginis X skirstinys imties $\mathcal{X}$ atžvilgiu yra Bernulio $B(1,\hat{p}_B)$.

6.3.1 pastaba. Šiame pavyzdyje parametro p aposteriorinis skirstinys $g(p|\mathcal{X})$ turi tą patį pavidalą kaip ir apriorinis skirstinys $g(p)$. Tokiu atveju sakoma, kad apriorinis beta skirstinys yra *sujungtinis* su Bernulio skirstiniu. Iš šio pavyzdžio matome, kad sujungtinį apriorinį skirstinį galima parinkti, jeigu išpildytos tokios sąlygos.

1. Egzistuoja parametro $\boldsymbol{\theta}$ dimensijos pakankamoji statistika $T = T(\mathcal{X})$, t. y.

$$L(\mathbf{X}_1, ..., \mathbf{X}_n|\boldsymbol{\theta}) = L_1(T(\mathcal{X})|\boldsymbol{\theta})h(\mathbf{X}_1, ..., \mathbf{X}_n),$$

čia h nepriklauso nuo parametro $\boldsymbol{\theta}$.

2. Apriorinio skirstinio tankis $g(\boldsymbol{\theta})$ gali būti parinktas proporcingas funkcijai L_1 (proporcingumo koeficientas nepriklauso nuo $\boldsymbol{\theta}$), o sandauga gL_1 yra tokio pat pavidalo kaip ir g.

Šias sąlygas tenkina, pavyzdžiui, eksponentinio tipo skirstiniai (detaliau žr. [5]).

Apskritai kalbant, aprioriniai skirstiniai dažniausiai parenkami iš sujungtinių skirstinių aibės. Tai gerokai palengvina aposteriorinių skirstinių radimą.

6.3.2 pavyzdys. Panagrinėkime šiek tiek sudėtingesnį pavyzdį. Tarkime, paprastoji imtis $\mathcal{X} = (\mathbf{X}_1, ..., \mathbf{X}_n)$ gauta stebint a. v. $\mathbf{X} = (X_1, ..., X_k)^T$, kurio skirstinys yra k-matis normalusis su žinoma neišsigimusia kovariacine matrica $\boldsymbol{\Sigma}$ ir nežinomu vidurkių vektoriumi $\boldsymbol{\mu}$. Vektorius $\boldsymbol{\mu}$ savo ruožtu yra k-matis normalusis vektorius $\boldsymbol{\mu} \sim N_k(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ su žinomu vidurkių vektoriumi $\boldsymbol{\mu}_0$ ir žinoma neišsigimusia kovariacine matrica $\boldsymbol{\Sigma}_0$.

Tikėtinumo funkcija

$$L(\mathbf{X}_1, ..., \mathbf{X}_n|\boldsymbol{\mu}) = (2\pi)^{-nk/2}|\boldsymbol{\Sigma}|^{-n/2} \exp\{-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{X}_i - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{X}_i - \boldsymbol{\mu})\}$$

$$= h(\mathcal{X}) \exp\{-\frac{n}{2}(\bar{\mathbf{X}} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu})\},$$

o apriorinis tankis

$$g(\boldsymbol{\mu}) = (2\pi)^{-nk/2}|\boldsymbol{\Sigma}|_0^{-n/2} \exp\{-\frac{1}{2}(\boldsymbol{\mu} - \boldsymbol{\mu}_0)^T\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\mu} - \boldsymbol{\mu}_0)\}.$$

Gauname aposteriorinį tankį

$$g(\boldsymbol{\mu}|\mathcal{X}) = c \exp\{-\frac{1}{2}[\boldsymbol{\mu}^T(n\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Sigma}_0^{-1})\boldsymbol{\mu} - 2\boldsymbol{\mu}^T(n\boldsymbol{\Sigma}^{-1}\bar{\mathbf{X}} + \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_0)]\},$$

kurį galime suvesti į tokį pavidalą

$$g(\boldsymbol{\mu}|\mathcal{X}) = c' \exp\{-\frac{1}{2}(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}_B)^T\boldsymbol{\Sigma}_n^{-1}(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}_B)\},$$

jeigu parinksime $\hat{\boldsymbol{\mu}}_B$ ir $\boldsymbol{\Sigma}_n$ iš lygčių sistemos

$$\boldsymbol{\Sigma}_n^{-1} = n\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Sigma}_0^{-1},$$

$$\boldsymbol{\Sigma}_n^{-1}\hat{\boldsymbol{\mu}}_B = n\boldsymbol{\Sigma}^{-1}\bar{\mathbf{X}} + \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_0.$$

Pasinaudoję lygybe,

$$(\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1})^{-1} = \boldsymbol{A}(\boldsymbol{A} + \boldsymbol{B})^{-1}\boldsymbol{B} = \boldsymbol{B}(\boldsymbol{A} + \boldsymbol{B})^{-1}\boldsymbol{A},$$

kuri teisinga visoms vienodos dimensijos neišsigimusioms matricoms, gauname

$$\mathbf{\Sigma}_n = \mathbf{\Sigma}_0(\mathbf{\Sigma}_0 + \frac{1}{n}\mathbf{\Sigma})^{-1}\frac{1}{n}\mathbf{\Sigma},$$

$$\hat{\boldsymbol{\mu}}_B = \frac{1}{n}\mathbf{\Sigma}(\mathbf{\Sigma}_0 + \frac{1}{n}\mathbf{\Sigma})^{-1}\boldsymbol{\mu}_0 + \mathbf{\Sigma}_0(\mathbf{\Sigma}_0 + \frac{1}{n}\mathbf{\Sigma})^{-1}\bar{\mathbf{X}}.$$

Taigi aposteriorinis $\boldsymbol{\mu}$ skirstinys normalusis $N_k(\hat{\boldsymbol{\mu}}_B, \mathbf{\Sigma}_n)$.

Kaip ir **6.3.1** pavyzdyje Bejeso įvertinys $\hat{\boldsymbol{\mu}}_B$ susideda iš dviejų komponenčių. Pirmoji rodo apriorinės informacijos, o antroji apmokančiosios imties indėlį. Jeigu $\boldsymbol{\mu}_0$, $\mathbf{\Sigma}_0$ ir $\mathbf{\Sigma}$ fiksuoti, o $n \to \infty$, tai $\hat{\boldsymbol{\mu}}_B$ artėja į DT įvertinį $\bar{\mathbf{X}}$, t. y. apriorinės informacijos indėlis artėja į 0. Atvirkščiai, jeigu n, $\boldsymbol{\mu}_0$ ir $\mathbf{\Sigma}$ fiksuoti, o $\mathbf{\Sigma}_0 \to \mathbf{0}$ (sklaida apie $\boldsymbol{\mu}_0$ mažėja), tai $\hat{\boldsymbol{\mu}}_B$ artėja į $\boldsymbol{\mu}_0$, t. y. į įvertinį, kurį gautume, jei apmokančiosios imties nebūtų.

Ieškant a. v. $\mathbf{X}$ sąlyginio tankio $\mathcal{X}$ atžvilgiu pakanka pažymėti, kad a. v. $\mathbf{X}$ yra dviejų nepriklausomų normaliųjų vektorių $N_k(\hat{\boldsymbol{\mu}}_B, \mathbf{\Sigma}_n)$ ir $N_k(\mathbf{0}, \mathbf{\Sigma})$ suma. Taigi

$$(\mathbf{X}|\mathcal{X}) \sim N_k(\hat{\boldsymbol{\mu}}_B, \mathbf{\Sigma} + \mathbf{\Sigma}_n).$$

Keletas kitų sujungtinių apriorinių skirstinių pavyzdžių pateikiama 6.65–6.74 pratimuose.

6.3.2. Neparametriniai tankių įvertiniai

Parametrinių tankių įvertinių alternatyva yra neparametriniai tankių įvertiniai, kai jokios prielaidos apie tankio funkcinį pavidalą nepriimamos. Vienmačiu atveju neparametriniai tankių įvertiniai nagrinėti 1 dalies 3.6 skyrelyje. Analogiški tankių įvertiniai sudaromi ir daugiamačiu atveju.

6.3.2.1 Histograma

Tarkime, $\mathbf{X}_1, ..., \mathbf{X}_n$ yra paprastoji imtis a. v. $\mathbf{X}$, turinčio absoliučiai tolydųjį k-matį skirstinį, kurio su tankio funkcija $f(\boldsymbol{x})$ nežinoma, o $\boldsymbol{x}_1, ..., \boldsymbol{x}_n$ yra imties realizacija. Sudalinkime k-matę erdvę $\mathbf{R}^k$ į vienodo didumo k-mačius intervalus

$$\boldsymbol{I} = \{(x_1, ..., x_k) : a_i' < x_i \le a_i'', \ \ i = 1, ..., k\},$$

kurių tūriai

$$V = V(\boldsymbol{I}) = \prod_{i=1}^{k}(a_i'' - a_i') = \prod_{i=1}^{k} h_i.$$

Jeigu su visais $i = 1, ..., k$ briaunų ilgiai $h_i \to 0$, tai su bet kuriuo $\boldsymbol{x} \in \boldsymbol{I}$ tikimybė

$$\mathbf{P}\{\mathbf{X} \in \boldsymbol{I}\} = \int_{\boldsymbol{I}} f(\boldsymbol{y})d\boldsymbol{y} \approx f(\boldsymbol{x}) \cdot V,$$

taigi tankio reikšmė nedaug skiriasi nuo

$$f_n(\boldsymbol{x}) = \frac{1}{nV}\sum_{j=1}^{n} \boldsymbol{I}_{\{\boldsymbol{I}\}}(\boldsymbol{x}_j) = \frac{m}{nV}, \qquad (6.3.13)$$

čia m yra skaičius tų $\boldsymbol{x}_j$, kurie pateko į intervalą $\boldsymbol{I}$.

Sudarę $f_n(\boldsymbol{x})$ visiems intervalams, gausime funkciją, vadinamą *histograma.*

Funkcija f_n yra realizacija atsitiktinės funkcijos $\hat{f}_n$, kuri su visais $\boldsymbol{x} \in \boldsymbol{I}$ apibrėžiama lygybe

$$\hat{f}_n(\boldsymbol{x}) = \frac{1}{nV}\sum_{j=1}^{n} \boldsymbol{I}_{\{\boldsymbol{I}\}}(\mathbf{X}_j) = \frac{M}{nV}, \tag{6.3.14}$$

čia M yra atsitiktinės imties elementų $\mathbf{X}_j$, patekusių į intervalą $\boldsymbol{I}$, skaičius.

6.3.2 teorema. *Jeigu funkcija $f(\boldsymbol{x})$ tolydi taško $\boldsymbol{x}$ aplinkoje, tai*

$$\hat{f}_n(\boldsymbol{x}) \overset{b.t.}{\to} f(\boldsymbol{x}), \quad kai \quad nV \to \infty, \quad h_i \to 0, \quad i = 1, ..., k. \tag{6.3.15}$$

Įrodymas. Bet kuriame f tolydumo taške $\boldsymbol{x} \in \boldsymbol{I}$

$$|\mathbf{E}(\hat{f}_n(\boldsymbol{x})) - f(\boldsymbol{x})| = |\frac{1}{V}\mathbf{P}\{\mathbf{X} \in \boldsymbol{I}\} - f(\boldsymbol{x})| =$$

$$|\frac{1}{V}\int_{\boldsymbol{I}} f(\boldsymbol{y})d\boldsymbol{y} - f(\boldsymbol{x})| \leq \sup_{\boldsymbol{y}\in\boldsymbol{I}} |f(\boldsymbol{y}) - f(\boldsymbol{x})| \to 0, \tag{6.3.16}$$

kai $h_i \to 0,\ i = 1, ..., k$.

Kadangi

$$|\hat{f}_n(\boldsymbol{x}) - f(\boldsymbol{x})| \leq |\hat{f}_n(\boldsymbol{x}) - \mathbf{E}(\hat{f}_n(\boldsymbol{x}))| + |\mathbf{E}(\hat{f}_n(\boldsymbol{x})) - f(\boldsymbol{x})|,$$

tereikia įvertinti dešinės nelygybės pusės pirmąjį dėmenį. Gauname

$$\hat{f}_n(\boldsymbol{x}) - \mathbf{E}(\hat{f}_n(\boldsymbol{x})) = \frac{1}{n}\sum_{j=1}^{n}(Y_j - \mathbf{E}Y_j), \quad Y_j = \frac{1}{V}\boldsymbol{I}_{\boldsymbol{I}}(\mathbf{X}_j),$$

A. d. Y_j nepriklausomi ir vienodai pasiskirstę

$$\mathbf{E}Y_j = \frac{1}{V}\mathbf{P}\{\mathbf{X} \in \boldsymbol{I}\}, \quad \boldsymbol{V}Y_j = \frac{1}{V^2}\mathbf{P}\{\mathbf{X} \in \boldsymbol{I}\}(1 - \mathbf{P}\{\mathbf{X} \in \boldsymbol{I}\}).$$

Kadangi

$$\frac{\boldsymbol{V}Y_j}{n} \leq \frac{1}{nV^2}\mathbf{P}\{\mathbf{X} \in \boldsymbol{I}\} = \frac{1}{nV^2}\int_{\boldsymbol{I}} f(\boldsymbol{y})d\boldsymbol{y} \leq \frac{1}{nV} sup_{\boldsymbol{y}\in\boldsymbol{I}} f(\boldsymbol{y}) \to 0,$$

kai $nV \to \infty$, tai, remdamiesi stipriuoju DSD ir (6.3.16), gauname, kad

$$\hat{f}_n(\boldsymbol{x}) \overset{b.t.}{\to} f(\boldsymbol{x}). \qquad \blacktriangle$$

6.3.2 pastaba. Kai n fiksuotas, ypač svarbu tinkamai parinkti intervalus $\boldsymbol{I}$. Jei šie intervalai labai maži, tai kiekviename iš jų bus 0 arba 1 iš realizacijų $\boldsymbol{x}_j$. Gausime aukštus stulpelius, atitinkančius 1, ir intervalus be stulpelių, atitinkančius 0. Iš tokio grafiko mažai ką galima pasakyti apie tankio pavidalą. Kita vertus, jeigu $\boldsymbol{I}$ dideli, tai turėsime keletą intervalų su stebėjimais ir visus kitus be jų. Tai taip pat nėra vaizdu.

6.3.2.2 Branduoliniai tankių įvertiniai

Jeigu imtis paprastoji, tai kiekvienam imties realizacijos elementui $\boldsymbol{x}_j$ priskiriama vienodo didumo $1/n$ tikimybinė masė. Kitaip negu sudarant histogramą, tankį galima bandyti priartinti išsklaidant tikimybinę masę $1/n$ taško $\boldsymbol{x}_j$ aplinkoje.

Tarkime, $K(\boldsymbol{x}) = K(x_1, ..., x_k)$ yra unimodalus tankis, įgyjantis maksimumą taške $\mathbf{0} = (0, ..., 0)$. Jis vadinamas *branduoliu*. Tada

$$f_{in}(x_1, ...x_k) = \frac{1}{nV} K\left(\frac{x_1 - x_{1i}}{h_1}, ..., \frac{x_k - x_{ki}}{h_k}\right),$$
$$V = h_1 \cdot h_2 \cdot ... \cdot h_k, \quad \boldsymbol{x}_i = (x_{1i}, ..., x_{ki}),$$

yra unimodali funkcija, kurios maksimumas taške $\boldsymbol{x}_i$. Taigi didumo $1/n$ tikimybinę masę, priskirtą elementui $\boldsymbol{x}_i$, išsklaidome taško $\boldsymbol{x}_i$ aplinkoje remdamiesi branduoliu $K(\boldsymbol{x})$.

Tokį masės išsklaidymą atlikus su visais imties realizacijos elementais $\boldsymbol{x}_1, ...,$ $\boldsymbol{x}_n$, taške $\boldsymbol{x} = (x_1, ..., x_k)^T \in \mathbf{R}^k$ tikimybinės masės tankis bus

$$f_n(x_1, ..., x_k) = \frac{1}{nV} \sum_{i=1}^{n} K\left(\frac{x_1 - x_{1i}}{h_1}, ..., \frac{x_k - x_{ki}}{h_k}\right).$$

Gautoji tankio funkcija f_n yra realizacija atsitiktinės funkcijos

$$\hat{f}_n(x_1, ..., x_k) = \frac{1}{nV} \sum_{i=1}^{n} K\left(\frac{x_1 - X_{1i}}{h_1}, ..., \frac{x_k - X_{ki}}{h_k}\right),$$

čia $\mathbf{X}_i = (X_{1i}, ..., X_{ki})^T$ yra i-asis atsitiktinės imties elementas.

Atsitiktinė funkcija $\hat{f}_n$ vadinama tankio f *branduoliniu įvertiniu*.

Branduoliu $K(\boldsymbol{x})$ dažnai imama k-mačio normaliojo skirstinio $N_k(\mathbf{0}, \boldsymbol{I})$ tankio funkcija, arba, bendriau, skirstinio $N_k(\mathbf{0}, \boldsymbol{\Sigma})$ su neišsigimusia kovariacine matrica $\boldsymbol{\Sigma}$ tankio funkcija.

Atsitiktinės funkcijos $\hat{f}_n$ vidurkis

$$\mathbf{E}(\hat{f}_n(x_1, ...x_k)) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{V} \mathbf{E}(K\left(\frac{x_1 - X_{1i}}{h_1}, ..., \frac{x_k - X_{ki}}{h_k}\right))$$
$$= \frac{1}{V} \int \cdots \int K\left(\frac{x_1 - y_1}{h_1}, ..., \frac{x_k - y_k}{h_k}\right) f(y_1, ..., y_k) dy_1 ... dy_k$$
$$= \int \cdots \int K(z_1, ..., z_k) f(x_1 - h_1 y_1, ..., x_k - h_k y_k) dz_1 ... dz_k$$
$$= \int_{\mathbf{R}^k} K(\boldsymbol{z}) f(\boldsymbol{x} - \boldsymbol{h}^T \boldsymbol{z}) d\boldsymbol{z}, \quad \boldsymbol{h} = (h_1, ..., h_k)^T, \tag{6.3.17}$$

o dispersija

$$\boldsymbol{V}(\hat{f}_n(x_1, ...x_k)) \leq \mathbf{E}(\hat{f}_n^2(x_1, ...x_k))$$
$$= \frac{1}{nV^2} \int_{\mathbf{R}^k} K^2(\boldsymbol{z}) f(\boldsymbol{x} - \boldsymbol{h}^T \boldsymbol{z}) d\boldsymbol{z} \leq \frac{\sup_{\boldsymbol{z}} K(\boldsymbol{z}) \mathbf{E}(\hat{f}_n(\boldsymbol{x}))}{nV}. \tag{6.3.18}$$

6.3.3 teorema. *Jeigu tankio funkcija $f(\boldsymbol{x})$ tolydi ir aprėžta taško $\boldsymbol{x}$ aplinkoje ir išpildyta sąlyga*

$$\int_{\mathbf{R}^k} K(\boldsymbol{z})f(\boldsymbol{x}-\boldsymbol{h}^T\boldsymbol{z})d\boldsymbol{z} \to f(\boldsymbol{x}), \quad kai \quad \boldsymbol{h} \to \mathbf{0},$$

tai

$$\hat{f}_n(\boldsymbol{x}) \overset{kv.v.}{\to} f(\boldsymbol{x}), \quad kai \quad nV \to \infty.$$

Įrodymas. Išplaukia iš vidurkio ir dispersijos išraiškų (6.3.17) ir (6.3.18). ▲

6.3.3. Diskriminantinių funkcijų vertinimas

Minėjome, kad sudarant klasifikatorius kartais tankių vertinimo klausimas apeinamas. Tariama, kad klases atskiriančios diskriminantinės funkcijos turi tam tikrą žinomą pavidalą, priklausantį nuo nežinomų parametrų. Remdamiesi kokiu tai kriterijumi įvertinę nežinomus parametrus, gauname ir diskriminantinių funkcijų įvertinius.

6.3.3.1. Fišerio tiesinė diskriminantinė funkcija, kai yra dvi klasės

Tarkime objektai klasifikuojami į dvi klases. Objektus apibūdina k-matis a. v. $\mathbf{X} = (X_1, ..., X_k)^T$. Jeigu kiekvienos klasės a. v. $\mathbf{X}$ skirstinys yra normalusis, tai diskriminantinė funkcija yra tiesinė (6.1.40). Pagal analogiją Fišeris pasiūlė įvesti tokią parametrizaciją: apibrėžiame tiesinę funkciją

$$g(\boldsymbol{x}) = z = \boldsymbol{\beta}^T\boldsymbol{x} \tag{6.3.19}$$

ir, lygindami jos reikšmę su pasirinktu slenksčiu, objektą priskiriame pirmai arba antrai klasei. Nežinomų parametrų vektorius $\boldsymbol{\beta} = (\beta_1, ..., \beta_k)^T$ įvertinamas naudojant apmokančiąsias imtis $\mathbf{X}_1^{(i)}, ..., \mathbf{X}_{n_i}^{(i)}, i = 1,\ 2$. Tiesinę transformaciją (6.3.19) galime interpretuoti kaip k-mačių vektorių projektavimą į tiesę, kurios kryptis sutampa su vektoriaus $\boldsymbol{\beta}$ kryptimi, t. y. kaip perėjimą nuo k-mačio a. v. $\mathbf{X}$ prie vienmačio a. d. Z. Atlikus projektavimą apmokančiųjų imčių elementai bus suprojektuoti į minėtą tiesę ir gausime vienmačių a. d. rinkinius

$$Z_j^{(1)} = \boldsymbol{\beta}^T\mathbf{X}_j^{(1)}, \quad j = 1, ..., n_1; \quad Z_j^{(2)} = \boldsymbol{\beta}^T\mathbf{X}_j^{(2)}, \quad j = 1, ..., n_2. \tag{6.3.20}$$

Aritmetiniai vidurkiai

$$\bar{\mathbf{X}}^{(i)} = \frac{1}{n_i}\sum_{j=1}^{n_i}\mathbf{X}_j^{(i)}, \quad i = 1,\ 2; \quad \bar{\mathbf{X}} = \frac{1}{n}(n_1\bar{\mathbf{X}}^{(1)} + n_2\bar{\mathbf{X}}^{(2)}), \quad n = n_1 + n_2,$$

bus suprojektuoti į atitinkamus a. d. $Z_j^{(i)}$ aritmetinius vidurkius

$$\bar{Z}^{(i)} = \frac{1}{n_i}\sum_{j=1}^{n_i}Z_j^{(i)}, \quad i = 1,\ 2; \quad \bar{Z} = \frac{1}{n}(n_1\bar{Z}^{(1)} + n_2\bar{Z}^{(2)}), \quad n = n_1 + n_2. \tag{6.3.21}$$

Bendra jungtinės imties $Z_1^{(1)},...,Z_{n_1}^{(1)},Z_1^{(2)},...,Z_{n_2}^{(2)}$ sklaida

$$\tilde{S}_T=\sum_{i=1}^{2}\sum_{j=1}^{n_i}(Z_j^{(i)}-\bar{Z})^2=\sum_{i=1}^{2}\sum_{j=1}^{n_i}(Z_j^{(i)}-\bar{Z}^{(i)})^2+\sum_{i=1}^{2}n_i(\bar{Z}^{(i)}-\bar{Z})^2=$$

$$=\sum_{i=1}^{2}\sum_{j=1}^{n_i}(Z_j^{(i)}-\bar{Z}^{(i)})^2+\frac{n_1n_2}{n_1+n_2}(\bar{Z}^{(1)}-\bar{Z}^{(2)})^2=\tilde{S}_V+\tilde{S}_I \tag{6.3.22}$$

yra suma sklaidos apmokančiųjų imčių viduje $\tilde{S}_V$ ir likusios sklaidos $\tilde{S}_I$, kuri apibūdina atstumą tarp imčių. Norint, kad $Z_j^{(1)}$ ir $Z_j^{(2)}$ būtų geriau atskiriami, vektorių $\boldsymbol{\beta}$ reikia parinkti taip, kad $\tilde{S}_I$ būtų kuo didesnė.

Fišerio diskriminantinė funkcija parenkama taip, kad būtų maksimizuojama kriterijaus funkcija

$$J(\boldsymbol{\beta})=\frac{\tilde{S}_I}{\tilde{S}_V}\to\max_{\boldsymbol{\beta}}. \tag{6.3.23}$$

Keletą kitokių kriterijaus funkcijų, naudojamų parenkant tiesinę diskriminantinę funkciją, galima rasti knygoje [4].

Įrašę į J($\boldsymbol{\beta}$) išraiškas pradines apmokančiąsias imtis, gauname

$$\tilde{S}_V=\boldsymbol{\beta}^T\sum_{i=1}^{2}\sum_{j=1}^{n_i}(\mathbf{X}_j^{(i)}-\bar{\mathbf{X}}^{(i)})(\mathbf{X}_j^{(i)}-\bar{\mathbf{X}}^{(i)})^T\boldsymbol{\beta}=\boldsymbol{\beta}^T\boldsymbol{S}_V\boldsymbol{\beta},$$

$$\tilde{S}_I=\boldsymbol{\beta}^T(\bar{\mathbf{X}}^{(1)}-\bar{\mathbf{X}}^{(2)})(\bar{\mathbf{X}}^{(1)}-\bar{\mathbf{X}}^{(2)})^T\boldsymbol{\beta}=\boldsymbol{\beta}^T\boldsymbol{S}_I\boldsymbol{\beta}$$

Taigi reikia maksimizuoti

$$J(\boldsymbol{\beta})=\frac{\boldsymbol{\beta}^T\boldsymbol{S}_I\boldsymbol{\beta}}{\boldsymbol{\beta}^T\boldsymbol{S}_V\boldsymbol{\beta}}\to\max_{\boldsymbol{\beta}}. \tag{6.3.24}$$

6.3.4 teorema. *Jeigu $n_1>k$ ir $n_2>k$, o matrica $\boldsymbol{S}_V$ teigiamai apibrėžta su tikimybe 1, tai Fišerio diskriminantinė funkcija turi tokį pavidalą*

$$\hat{g}(\boldsymbol{x})=\boldsymbol{x}^T\hat{\boldsymbol{\beta}}+w_0,\quad \hat{\boldsymbol{\beta}}=\boldsymbol{S}_V^{-1}(\bar{\mathbf{X}}^{(1)}-\bar{\mathbf{X}}^{(2)}). \tag{6.3.25}$$

Objektas priskiriamas pirmai klasei, kai $\hat{g}(\boldsymbol{x})>0$.

Įrodymas. Funkcijos $J(\boldsymbol{\beta})$ maksimizavimas yra ekvivalentus tam, kad

$$\boldsymbol{\beta}^T\boldsymbol{S}_I\boldsymbol{\beta}\to\max_{\boldsymbol{\beta}},\quad \boldsymbol{\beta}^T\boldsymbol{S}_V\boldsymbol{\beta}=1.$$

Sprendinys $\hat{\boldsymbol{\beta}}$ turi tenkinti lygtį

$$\boldsymbol{S}_I\hat{\boldsymbol{\beta}}=\lambda\boldsymbol{S}_V\hat{\boldsymbol{\beta}},\quad \text{arba}\quad \boldsymbol{S}_V^{-1}\boldsymbol{S}_I\hat{\boldsymbol{\beta}}=\lambda\hat{\boldsymbol{\beta}},$$

čia λ yra charakteringosios lygties

$$|\boldsymbol{S}_I - \lambda \boldsymbol{S}_V| = 0, \quad \text{arba} \quad |\boldsymbol{S}_V^{-1}\boldsymbol{S}_I - \lambda \boldsymbol{I}| = 0$$

šaknis. Kadangi $Rang(\boldsymbol{S}_I) = 1$, tai yra tik viena nenulinė lygties šaknis. Vektoriaus $\boldsymbol{S}_I\boldsymbol{\beta}$ kryptis sutampa su vektoriaus $\bar{\mathbf{X}}^{(1)} - \bar{\mathbf{X}}^{(2)}$ kryptimi, o daugiklis neturi reikšmės, tai galime imti $\hat{\boldsymbol{\beta}} = \boldsymbol{S}_V^{-1}(\bar{\mathbf{X}}^{(1)} - \bar{\mathbf{X}}^{(2)})$.

Reikia pastebėti, jei a. v. skirstinys abiejų klasių atveju yra k-matis normalusis su vienoda kovariacine matrica, tai (6.3.25) pakeitę empirines charakteristikas teorinėmis ir atitinkamai parinkę slenkstį w_0, gausime diskriminantinę funkciją (6.1.41).

Slenkstis w_0 gali būti parinktas empiriniu būdu. Tuo tikslu su įvairiomis w_0 reikšmėmis įvertinamos klasifikavimo klaidos ir parenkama priimtina w_0 reikšmė. Norint korektiškai įvertinti klasifikavimo klaidas, iš kiekvienos klasės reikia turėti pakankamo dydžio aibes objektų, kurie nebuvo naudojami sudarant klasifikatorių (testinės aibės). Jeigu apmokančiųjų imčių didumai pakankami, tai jas galima suskaidyti į dvi dalis: pirmosios dalys naudojamos klasifikatoriui sudaryti, o antrosios dalys – įvertinti jo tikslumą.

Jeigu apmokančiųjų aibių didumai nepakankami, rekomenduojama taikyti tokią procedūrą. Iš pirmos apmokančiosios imties pašaliname elementą $\mathbf{X}_1^{(1)}$; sudarome klasifikatorių naudodami likusias apmokančiąsias didumo $n_1 - 1$ ir n_2 imtis; gautąjį klasifikatorių išbandome klasifikuodami stebėjimą $\mathbf{X}_1^{(1)}$. Analogišką procedūrą pakartojame su kiekvienu pirmos ir antros imties elementu. Aprašytoji procedūra leidžia klasifikavimo tikslumui vertinti naudoti visus $n_1 + n_2$ stebėjimus iš esmės nesumažinant kriterijaus tikslumo (jis sudaromas naudojant vienu stebėjimu mažiau). Tiesa, skaičiavimų apimtis gerokai padidėja, tačiau tai nesudaro ypatingų sunkumų, nes daugelyje specializuotų matematinės statistikos TPP tokia procedūra numatyta.

6.3.3 pavyzdys. Lentelėje pateikti duomenys apie lašišų charakteristikų matavimus; čia $(X_{1i}^{(1)}, X_{2i}^{(1)})^T$ yra Aliaskos lašišų matavimai, o $(X_{1i}^{(2)}, X_{2i}^{(2)})^T$ – Kanados lašišų matavimai [9]. Sudarysime klasifikatorių Aliaskos ir Kanados lašišoms atskirti remiantis turimomis apmokančiosiomis imtimis.

$X_{1i}^{(1)}$	$X_{2i}^{(1)}$	$X_{1i}^{(1)}$	$X_{2i}^{(1)}$	$X_{1i}^{(1)}$	$X_{2i}^{(1)}$	$X_{1i}^{(1)}$	$X_{2i}^{(1)}$	$X_{1i}^{(1)}$	$X_{2i}^{(1)}$
108	368	114	428	114	396	105	388	84	511
131	355	123	372	100	470	121	403	91	469
105	469	123	372	84	399	85	451	74	451
86	506	109	420	102	429	83	453	101	474
99	402	112	394	101	469	53	427	80	398
87	423	104	407	85	444	95	411	95	433
94	440	111	422	109	397	76	442	92	404
117	489	126	423	106	442	95	426	99	481
79	432	105	434	82	431	87	402	94	491
99	403	119	474	118	381	70	397	87	480

$X_{1i}^{(2)}$	$X_{2i}^{(2)}$	$X_{1i}^{(2)}$	$X_{2i}^{(2)}$	$X_{1i}^{(2)}$	$X_{2i}^{(2)}$	$X_{1i}^{(2)}$	$X_{2i}^{(2)}$	$X_{1i}^{(2)}$	$X_{2i}^{(2)}$
129	420	149	393	90	385	153	403	133	375
148	371	108	330	145	337	150	354	128	383
179	407	135	355	123	364	154	390	123	349
152	381	170	386	145	376	155	349	144	373
166	377	152	301	115	354	109	325	140	388
124	389	153	397	134	383	117	344	150	339
156	419	152	301	117	355	128	400	124	341
131	345	136	438	126	345	144	403	125	346
140	362	122	306	118	379	163	370	153	352
144	345	148	383	120	369	145	355	108	339

Tardami, kad buvo stebėti normalieji a. v. $N_2(\boldsymbol{\mu}^{(1)}, \boldsymbol{\Sigma})$ ir $N_2(\boldsymbol{\mu}^{(2)}, \boldsymbol{\Sigma})$ su vienodomis kovariacinėmis matricomis, randame pirmos ir antros imties vidurkių vektorių ir kovariacinių matricų įverčius

$$\hat{\boldsymbol{\mu}}^{(1)} = \begin{pmatrix} \bar{X}_{1.}^{(1)} \\ \bar{X}_{2.}^{(1)} \end{pmatrix} = \begin{pmatrix} 98,380 \\ 429,660 \end{pmatrix}, \quad \hat{\boldsymbol{\mu}}^{(2)} = \begin{pmatrix} \bar{X}_{1.}^{(2)} \\ \bar{X}_{2.}^{(2)} \end{pmatrix} = \begin{pmatrix} 137,460 \\ 366,620 \end{pmatrix},$$

$$\hat{\boldsymbol{\Sigma}}^{(1)} = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(X_{1i}^{(1)} - \bar{X}_{1.}^{(1)})(X_{1i}^{(1)} - \bar{X}_{1.}^{(1)})^T = \begin{pmatrix} 260,608 & -188,093 \\ -188,093 & 1399,086 \end{pmatrix},$$

$$\hat{\boldsymbol{\Sigma}}^{(2)} = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(X_{1i}^{(2)} - \bar{X}_{1.}^{(2)})(X_{1i}^{(2)} - \bar{X}_{1.}^{(2)})^T = \begin{pmatrix} 326,090 & 133,505 \\ 133,505 & 893,261 \end{pmatrix},$$

ir bendrą kovariacinės matricos įvertį

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{2}(\hat{\boldsymbol{\Sigma}}^{(1)} + \hat{\boldsymbol{\Sigma}}^{(2)}) = \begin{pmatrix} 293,349 & -27,294 \\ -27,294 & 1146,174 \end{pmatrix}, \hat{\boldsymbol{\Sigma}}^{-1} = \begin{pmatrix} 0,003416 & 0,000081 \\ 0,000081 & 0,000874 \end{pmatrix}.$$

Parinkę slenkstį $d = 0$, gauname klasifikavimo taisyklės įvertį: objektas, kuriam vektorius $\mathbf{X}$ lygus $\mathbf{X}^0 = (X_1^0, X_2^0)^T$, priskiriamas pirmai klasei, kai

$$\hat{g}(\mathbf{X}^0) = (\hat{\boldsymbol{\mu}}^{(1)} - \hat{\boldsymbol{\mu}}^{(2)})^T\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{X}_0 - \frac{1}{2}(\hat{\boldsymbol{\mu}}^{(1)} - \hat{\boldsymbol{\mu}}^{(2)})^T\hat{\boldsymbol{\Sigma}}^{-1}(\hat{\boldsymbol{\mu}}^{(1)} + \hat{\boldsymbol{\mu}}^{(2)}) =$$

$$-0,12839X_1^0 + 0,05194X_2^0 - 5,54121 \geq 0,$$

ir priskiriamas antrai klasei priešingu atveju.

Pritaikykime šią klasifikavimo taisyklę apmokančiųjų imčių objektams klasifikuoti. Gauname, kad $V_{11} = 44$ pirmos imties objektai, teisingai priskiriami pirmai grupei, o $V_{21} = 6$ objektai, priskiriami antrai grupei. Antros imties $V_{22} = 49$ objektai teisingai priskiriami antrai grupei ir tik vienas objektas $V_{12} = 1$ klaidingai priskiriamas pirmai grupei. Klaidingo klasifikavimo tikimybių įverčiai $\hat{\alpha}_{21} = V_{21}/50 = 0,12$, $\hat{\alpha}_{12} = V_{12}/50 = 0,02$.

6.3.3.2. Fišerio diskriminantinės funkcijos, kai klasių yra daugiau negu dvi

Klasifikuojant objektus į $m > 2$ klasių reikia apibrėžti $m - 1$ diskriminantinę funkciją

$$g_j(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{\beta}_j + w_j^0, \quad j = 1, ..., m - 1, \tag{6.3.26}$$

kurias lygindami suklasifikuojame objektus į m klasių. Nežinomų parametrų $\boldsymbol{\beta}_j = (\beta_{j1}, ..., \beta_{jk})^T, j = 1, ..., m-1$, įvertiniai gaunami iš nepriklausomų apmokančiųjų imčių $(\mathbf{X}_1^{(i)}, ..., \mathbf{X}_{n_i}^{(i)}), i = 1, ..., m$. Transfomacijas $g_j(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{\beta}_j, j = 1, ..., m-1$, galime interpretuoti kaip k-matės erdvės taško $\boldsymbol{x} \in \mathbf{R}^k$ projektavimą į $\mathbf{R}^{m-1}$ erdvę. Imčių elementai $\mathbf{X}_j^{(i)}$ bus suprojektuoti į taškus $\boldsymbol{Z}_j^{(i)} = \boldsymbol{B}^T\mathbf{X}_j^{(i)}$; čia $\boldsymbol{B}$ yra matrica $k \times (m-1)$, kurios stulpeliai yra vektoriai $\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_{m-1}$. Aritmetiniai vidurkiai

$$\bar{\mathbf{X}}^{(i)} = \frac{1}{n_i}\sum_{j=1}^{n_i}\mathbf{X}_j^{(i)}, \quad \bar{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{m} n_i\bar{\mathbf{X}}^{(i)}, \quad n = n_1 + ... + n_m,$$

bus suprojektuoti į vidurkius

$$\bar{\boldsymbol{Z}}^{(i)} = \boldsymbol{B}^T\bar{\mathbf{X}}^{(i)}, \quad i = 1, ..., m, \quad \bar{\boldsymbol{Z}} = \boldsymbol{B}^T\bar{\mathbf{X}}.$$

Bendra jungtinės imties $\{\boldsymbol{Z}_1^{(i)}, ..., \boldsymbol{Z}_{n_i}^{(i)}, i = 1, ..., m\}$ sklaida apibūdinama matrica

$$\tilde{\boldsymbol{S}}_T = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(\boldsymbol{Z}_j^{(i)} - \bar{\boldsymbol{Z}})(\boldsymbol{Z}_j^{(i)} - \bar{\boldsymbol{Z}})^T = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(\boldsymbol{Z}_j^{(i)} - \bar{\boldsymbol{Z}}^{(i)})(\boldsymbol{Z}_j^{(i)} - \bar{\boldsymbol{Z}}^{(i)})^T +$$

$$+\sum_{i=1}^{m} n_i(\bar{\boldsymbol{Z}}^{(i)} - \bar{\boldsymbol{Z}})(\boldsymbol{Z}^{(i)} - \bar{\boldsymbol{Z}})^T = \tilde{\boldsymbol{S}}_V + \tilde{\boldsymbol{S}}_I \tag{6.3.27}$$

yra suma sklaidos apmokančiųjų imčių viduje $\tilde{\boldsymbol{S}}_V$ ir likusios sklaidos $\tilde{\boldsymbol{S}}_I$, kuri apibūdina atstumą tarp imčių. Gauname

$$\tilde{\boldsymbol{S}}_V = \boldsymbol{B}^T\sum_{i=1}^{m}\sum_{j=1}^{n_i}(\mathbf{X}_j^{(i)} - \bar{\mathbf{X}}^{(i)})(\mathbf{X}_j^{(i)} - \bar{\mathbf{X}}^{(i)})^T\boldsymbol{B} = \boldsymbol{B}^T\boldsymbol{S}_V\boldsymbol{B},$$

$$\tilde{\boldsymbol{S}}_I = \sum_{i=1}^{m} n_i(\bar{\mathbf{X}}^{(i)} - \bar{\mathbf{X}})(\mathbf{X}^{(i)} - \bar{\mathbf{X}})^T\boldsymbol{B} = \boldsymbol{B}^T\boldsymbol{S}_I\boldsymbol{B}. \tag{6.3.28}$$

Matricą $\boldsymbol{B}$ reikia parinkti taip, kad sklaida tarp imčių būtų kuo didesnė, lyginant su sklaida imčių viduje. Sklaidos matu imdami determinantus, gauname daugiamatį (6.3.24) kriterijaus analogą

$$J(\boldsymbol{B}) = \frac{|\tilde{\boldsymbol{S}}_I|}{|\tilde{\boldsymbol{S}}_V|} = \frac{|\boldsymbol{B}^T\boldsymbol{S}_I\boldsymbol{B}|}{|\boldsymbol{B}^T\boldsymbol{S}_V\boldsymbol{B}|} \to \max_{\boldsymbol{B}}. \tag{6.3.29}$$

6.3.5 teorema. *Tegu $n_1 + ... + n_m - m \geq k$, $k \geq m$; matrica $\boldsymbol{S}_V$ teigiamai apibrėžta su tikimybe 1, o* Rang$(\boldsymbol{S}_I) = m - 1$ *su tikimybe 1. Tada (6.3.29) sprendinys yra matrica $\hat{\boldsymbol{B}}$, kurios stulpelius sudaro tikriniai vektoriai $\hat{\boldsymbol{\beta}}_1, ..., \hat{\boldsymbol{\beta}}_{m-1}$, atitinkantys charakteringosios lygties*

$$|\boldsymbol{S}_I - \lambda\boldsymbol{S}_V| = 0 \tag{6.3.30}$$

šaknis $\lambda_1 > \lambda_2 > ... > \lambda_{m-1}$.

Įrodymas. Kadangi $|\boldsymbol{S}_V| > 0$, tai egzistuoja neišsigimusi kvadratinė matrica $\boldsymbol{C}$, kad $\boldsymbol{S}_V = \boldsymbol{C}\boldsymbol{C}^T$. Tada lygtis

$$|\boldsymbol{S}_I - \lambda \boldsymbol{S}_V| = 0 \Leftrightarrow |\boldsymbol{S}_I - \lambda \boldsymbol{C}\boldsymbol{C}^T| = 0 \Leftrightarrow |(\boldsymbol{C}^{-1})^T \boldsymbol{S}_I \boldsymbol{C}^{-1} - \lambda \boldsymbol{I}| = 0. \quad (6.3.31)$$

Matrica $(\boldsymbol{C}^{-1})^T \boldsymbol{S}_I \boldsymbol{C}^{-1}$ simetrinė, neneigiamai apibrėžta ir turi rangą lygų $m-1$. Todėl egzistuoja $m-1$ nenulinių skirtingų lygties (6.3.30) sprendinių $\lambda_1, ..., \lambda_{m-1}$ ir juos atitinkančių $m-1$ tikrinių vektorių $\boldsymbol{L}_1, ..., \boldsymbol{L}_{m-1}$, kad matrica $\boldsymbol{L}$, kurios stulpelius sudaro vektoriai $\boldsymbol{L}_1, ..., \boldsymbol{L}_{m-1}$, tenkina sąlygas

$$\boldsymbol{L}^T (\boldsymbol{C}^{-1})^T \boldsymbol{S}_I \boldsymbol{C}^{-1} \boldsymbol{L} = \boldsymbol{\Lambda}, \quad \boldsymbol{L}\boldsymbol{L}^T = \boldsymbol{I};$$

čia $\boldsymbol{\Lambda}$ – diagonalioji matrica, kurios diagonaliniai elementai yra $\lambda_1, ..., \lambda_{m-1}$. Parinkę matricą $\boldsymbol{B} = \boldsymbol{C}^{-1}\boldsymbol{L}$, gausime

$$J(\boldsymbol{B}) = \frac{|\boldsymbol{B}^T \boldsymbol{S}_I \boldsymbol{B}|}{|\boldsymbol{B}^T \boldsymbol{S}_V \boldsymbol{B}|} = \frac{|\boldsymbol{L}^T (\boldsymbol{C}^{-1})^T \boldsymbol{S}_I \boldsymbol{C}^{-1} \boldsymbol{L}|}{\boldsymbol{L}^T (\boldsymbol{C}^{-1})^T \boldsymbol{S}_V \boldsymbol{C}^{-1} \boldsymbol{L}} = \frac{|\boldsymbol{\Lambda}|}{|\boldsymbol{L}^T \boldsymbol{L}|} = \lambda_1 \lambda_2 \cdots \lambda_{m-1}.$$

Taigi matricos $\hat{\boldsymbol{B}}$ stulpelius sudaro tikriniai vektoriai, tenkinantys sąlygą

$$\boldsymbol{S}_V^{-1} \boldsymbol{S}_I \hat{\boldsymbol{\beta}}_i = \lambda_i \hat{\boldsymbol{\beta}}_i, \quad i = 1, ..., m-1. \qquad \blacktriangle$$

Gauname $m-1$ įvertintą diskriminantinę funkciją

$$\hat{g}(\boldsymbol{x}) = \hat{\beta}_i^T \boldsymbol{x}, \quad i = 1, 2, ..., m-1. \quad (6.3.32)$$

Remiantis šiomis dikriminantinėmis funkcijomis klasifikavimas atliekamas tokiu būdu. Tarkime, kad objektą, kuriam vektorius $\mathbf{X} = \boldsymbol{x}$, reikia priskirti vienai iš m klasių. Remdamiesi (6.3.32) suprojektuojame tašką $\boldsymbol{x}$ ir vidurkių įverčius $\bar{\mathbf{X}}^{(i)}$, $i = 1, 2, ..., m$ į $\mathbf{R}^{m-1}$ erdvę

$$\hat{\boldsymbol{x}} = \hat{\boldsymbol{B}}^T \boldsymbol{x}, \quad \hat{\bar{\boldsymbol{Z}}}^{(i)} = \hat{\boldsymbol{B}}^T \bar{\mathbf{X}}^{(i)}, \quad i = 1, ..., m.$$

Randame atstumų tarp projekcijų kvadratus

$$D_i = ||\hat{\boldsymbol{x}} - \hat{\bar{\boldsymbol{Z}}}^{(i)}||^2 = (\boldsymbol{x} - \bar{\mathbf{X}}^{(i)})^T \hat{\boldsymbol{B}} \hat{\boldsymbol{B}}^T (\boldsymbol{x} - \bar{\mathbf{X}}^{(i)}), \quad i = 1, ..., m.$$

Objektas, kuriam $\mathbf{X} = \boldsymbol{x}$, priskiriamas i-jai klasei, jeigu

$$D_i < D_j, \quad \forall \; j \neq i, \quad j = 1, ..., m. \quad (6.3.33)$$

6.3.4 pavyzdys. Lentelėje pateikti duomenys apie stojimo į aukštesniąją studijų pakopą rezultatus (I grupė – priimtas; II grupė – nepriimtas; III grupė – galutinis sprendimas nepriimtas) priklausomai nuo mokymosi rezultatų žemesnėje grandyje X_1 ir vidurinėje grandyje X_2 [9].

I gr				II gr				III gr			
$X_1^{(1)}$	$X_2^{(1)}$	$X_1^{(1)}$	$X_2^{(1)}$	$X_1^{(2)}$	$X_2^{(2)}$	$X_1^{(2)}$	$X_2^{(2)}$	$X_1^{(3)}$	$X_2^{(3)}$	$X_1^{(3)}$	$X_2^{(3)}$
2,96	596	3,47	552	2,54	446	2,13	408	2,86	494	3,12	463
3,14	473	3,35	520	2,43	425	2,41	469	2,85	496	3,08	440
3,22	482	3,39	543	2,20	474	2,55	538	3,14	419	3,03	419
3,29	527	3,28	523	2,36	531	2,31	505	3,28	371	3,00	509
3,69	505	3,21	530	2,57	542	2,41	489	2,89	447	3,03	438
3,46	693	3,58	564	2,35	406	2,19	411	3,15	313	3,05	399
3,03	626	3,33	565	2,51	412	2,35	321	3,50	402	2,85	483
3,19	663	3,40	431	2,51	458	2,60	394	2,89	485	3,01	453
3,63	447	3,38	605	2,36	399	2,55	528	2,80	444	3,03	414
3,59	588	3,26	664	2,36	482	2,72	399	3,13	416	3,04	446
3,30	563	3,60	609	2,66	420	2,85	381	3,01	471		
3,40	553	3,37	559	2,68	414	2,90	384	2,79	490		
3,50	572	3,80	521	2,48	533			2,89	431		
3,78	591	3,76	646	2,46	509			2,91	446		
3,44	692	3,24	467	2,63	504			2,75	546		
3,48	528			2,44	336			2,73	467		

Remdamiesi šiais duomenimis rasime klasifikavimo taisyklę trijų grupių objektams atskirti.

Tarkime, kad a. v. $\mathbf{X}^{(i)} = (X_1^{(i)}, X_2^{(i)})^T \sim N_2(\boldsymbol{\mu}^{(i)}, \boldsymbol{\Sigma})$, $i = 1,\ 2,\ 3$ turi dvimačius normaliuosius skirstinius su vienodomis kovariacinėmis matricomis. Randame parametrų įverčius (žr. 6.3.1.1 skyrelį)

$$\bar{\mathbf{X}}^{(1)} = \begin{pmatrix} 3,4039 \\ 561,2258 \end{pmatrix}, \quad \bar{\mathbf{X}}^{(2)} = \begin{pmatrix} 2,4825 \\ 447,0714 \end{pmatrix}, \quad \bar{\mathbf{X}}^{(3)} = \begin{pmatrix} 2,9927 \\ 446,2308 \end{pmatrix},$$

$$\hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 0,0360678 & -2,018759 \\ -2,018759 & 3655,90112 \end{pmatrix}, \quad \hat{\boldsymbol{\Sigma}}^{-1} = \begin{pmatrix} 28,609766 & 0,01579808 \\ 0,01579808 & 0,00028225 \end{pmatrix}.$$

Tardami, kad apriorinės klasių tikimybės $\omega_1 = \omega_2 = \omega_3 = 1/3$ yra vienodos, randame įvertintas diskriminantines funkcijas

$$\begin{aligned}\hat{g}_1(\boldsymbol{x}) &= (\bar{\mathbf{X}}^{(1)})^T \hat{\boldsymbol{\Sigma}}^{-1} \boldsymbol{x} - \frac{1}{2} (\bar{\mathbf{X}}^{(1)})^T \hat{\boldsymbol{\Sigma}}^{-1} \bar{\mathbf{X}}^{(1)} + \ln(1/3) \\ &= 106,25022 x_1 + 0,21218 x_2 - 241,47081.\end{aligned}$$

Analogiškai

$$\hat{g}_2(\boldsymbol{x}) = 78,08662 x_1 + 0,16541 x_2 - 134,99781, \quad \hat{g}_3(\boldsymbol{x}) = 92,66975 x_1 + 0,17323 x_2 - 178,41463.$$

Objektas, kurio $\mathbf{X} = \boldsymbol{x} = (x_1, x_2)^T$, priskiriamas i-ajai klasei, kai $\hat{g}_i(\boldsymbol{x}) > \hat{g}_j(\boldsymbol{x})$, $\forall\ j \neq i$, $j = 1,\ 2,\ 3$.

Pritaikysime šią taisyklę lentelėje pateiktoms apmokančiosioms imtims suklasifikuoti. Gauname tokią klasifikavimo lentelę

	$\eta = 1$	$\eta = 2$	$\eta = 3$	$\sum$
$\xi = 1$	27	0	4	31
$\xi = 2$	0	26	2	28
$\xi = 3$	1	0	25	26

Matome, kad suklasifikuota gana tiksliai. Suminės klasifikavimo klaidos tikimybės įvertis $7/85 = 0,082$.

Naudodami pateiktos lentelės duomenis rasime Fišerio klasifikavimo taisyklę (6.3.33) trims klasėms atskirti. Randame

$$\boldsymbol{S}_V = 82\hat{\boldsymbol{\Sigma}}, \quad \boldsymbol{S}_I = \sum_{i=1}^{3} n_i (\bar{\mathbf{X}}^{(i)} - \bar{\mathbf{X}})(\bar{\mathbf{X}}^{(i)} - \bar{\mathbf{X}}), \quad \bar{\mathbf{X}} = \sum_{i=1}^{3} n_i \bar{\mathbf{X}}^{(i)} / (n_1 + n_2 + n_3)$$

ir sudarome charakteringąją lygtį

$$|\boldsymbol{S}_V^{-1} \boldsymbol{S}_I - \lambda \boldsymbol{I}| = 0, \quad \lambda^2 - 5,83666\lambda + 1,07619 = 0.$$

Išsprendę kvadratinę lygtį gauname tikrines reikšmes $\lambda_1 = 0,19061$, $\lambda_2 = 5,64605$. Tikriniai vektoriai $\hat{\beta}_1, \hat{\beta}_2$, atitinkantys šias tikrines reikšmes ir tenkinantys sąlygas $\hat{\beta}_i^T \hat{\mathbf{\Sigma}}^{-1} \hat{\beta}_i = 1$, yra

$$\hat{\beta}_1 = (-1,87669,\ 0,01445)^T, \quad \hat{\beta}_2 = (5,00877,\ 0,00857)^T, \quad \hat{\boldsymbol{B}} = (\hat{\beta}_1 \vdots \hat{\beta}_2).$$

Suradę atstumus D_i, gausime (6.3.33) klasifikavimo taisyklę.

Pritaikysime šią taisyklę lentelėje pateiktoms apmokančiosioms imtims suklasifikuoti. Gauname tokią pačią klasifikavimo lentelę. Be to, lygiai tie patys objektai klaidingai priskiriami kitoms klasėms, kaip ir pateiktoje klasifikavimo lentelėje.

6.3.3.3. Logistinė regresija

Kai yra dvi klasės, logistinės regresijos modelyje (žr. 3 dalies 5.3 skyrelį) tariama, kad įvykio $\{\xi = 1\}$, t. y. objekto, priklausymo pirmai klasei aposteriorinė tikimybė turi tokį pavidalą

$$\omega_1(\boldsymbol{x}) = \mathbf{P}\{\xi = 1 | \mathbf{X} = \boldsymbol{x}\} = \frac{e^{\beta_0 + \boldsymbol{\beta}^T \boldsymbol{x}}}{1 + e^{\beta_0 + \boldsymbol{\beta}^T \boldsymbol{x}}}.$$

Remiantis šiomis tikimybėmis sudaromas šansų santykis

$$\frac{\omega_1(\boldsymbol{x})}{\omega_2(\boldsymbol{x})} = e^{\beta_0 + \boldsymbol{\beta}^T \boldsymbol{x}}.$$

Tada objektas priskirtinas pirmai klasei, kai

$$\frac{\omega_1(\boldsymbol{x})}{\omega_2(\boldsymbol{x})} > 1 \Leftrightarrow \beta_0 + \boldsymbol{\beta}^T \boldsymbol{x} > 0. \tag{6.3.34}$$

Bendresniu atveju šansų santykis gali būti lyginamas su slenksčiu $c \neq 1$.

Logistinės regresijos modelį galime interpretuoti kaip objektų klasifikavimo taisyklę su tiesine diskriminantine funkcija

$$g(\boldsymbol{x}) = \beta_0 + \boldsymbol{\beta}^T \boldsymbol{x}. \tag{6.3.35}$$

Kaip matome, parametrizacija čia įvedama tokiu būdu: tariama, kad žinomo pavidalo diskriminantinė funkcija priklauso nuo nežinomų parametrų, kuriuos tenka vertinti pagal apmokančiąsias imtis.

Pagal apmokančiąsias imtis gavę parametrų $\beta_0, \boldsymbol{\beta}$ įvertinius $\hat{\beta}_0, \hat{\boldsymbol{\beta}}$, gauname diskriminantinės funkcijos įvertinį

$$\hat{g}(\boldsymbol{x}) = \hat{\beta}_0 + \hat{\boldsymbol{\beta}}^T \boldsymbol{x}$$

ir klasifikavimo taisyklę: objektas priskiriamas pirmai klasei, kai $\hat{g}(\boldsymbol{x}) > c$.

Skirtumas nuo tiesinio Fišerio klasifikatoriaus yra tas, kad taikomi skirtingi metodai parametrams vertinti: logistinėje regresijoje taikomas DT metodas, o ieškant Fišerio klasifikatoriaus parametrų įvertiniai randami iš (6.3.24) sąlygos.

6.3.5 pavyzdys (**6.3.3** pavyzdžio tęsinys.) Rasime klasifikatorių dviems lašišų rūšims atskirti pagal **6.3.3** pavyzdžio duomenis naudodami logistinę regresiją. Naudodami SAS programų paketą gauname diskriminantinės funkcijos įvertį

$$\hat{g}(\boldsymbol{x}) = -3,9252 - 0,1260x_1 + 0,0485x_2.$$

Matome, kad gautas įvertis mažai skiriasi nuo Fišerio diskriminantinės funkcijos įverčio, gauto **6.3.3** pratime. Pritaikę klasifikavimo taisyklę **6.3.3** pavyzdžio duomenims klasifikuoti, gauname klasifikavimo lentelę

	$\eta=1$	$\eta=2$	$\sum$
$\xi=1$	46	4	50
$\xi=2$	3	47	50

6.3.3.4. Daugianarė logistinė regresija

Tarkime, kad objektus, kuriuos apibūdina a. v. $\mathbf{X}=(X_1,...,X_k)^T$, reikia suklasifikuoti į $m>2$ klasių. Tegu $\boldsymbol{e}_1=(1,\ 0,...,0)^T$, $\boldsymbol{e}_2=(0,\ 1,...,0)^T$,...,$\boldsymbol{e}_{m-1}=(0,\ 0,...,1)^T$ yra dimensijos $m-1$ vienetiniai vektoriai, o $\boldsymbol{e}_m=(0,\ 0,...,0)^T$ – nulinis vektorius. Tarsime, kad a. v. $\boldsymbol{Y}$ įgijo reikšmę $\boldsymbol{e}_j$, jeigu objektas priklauso j-ai klasei, $j=1,...,m$.

Daugianarės logistinės regresijos modelis gaunamas analogiškai kaip dvinarės logistinės regresijos modelis palyginant sąlygines tikimybes $\mathbf{P}\{\boldsymbol{Y}=\boldsymbol{e}_j|\mathbf{X}=\boldsymbol{x}\}$ ir $\mathbf{P}\{\boldsymbol{Y}=\boldsymbol{e}_m|\mathbf{X}=\boldsymbol{x}\}$, $j=1,...m-1$. Tariame, kad

$$\frac{\mathbf{P}\{\boldsymbol{Y}=\boldsymbol{e}_j|\mathbf{X}=\boldsymbol{x}\}}{\mathbf{P}\{\boldsymbol{Y}=\boldsymbol{e}_m|\mathbf{X}=\boldsymbol{x}\}}=\exp\{\alpha_j+\boldsymbol{\beta}_j^T\boldsymbol{x}\}=\exp\{\mu_j(\boldsymbol{x})\},\quad j=1,...,m-1, \tag{6.3.36}$$

čia α_j ir $\boldsymbol{\beta}_j=(\beta_{j1},...,\beta_{jk})^T$, $j=1,...,m-1$ yra nežinomi parametrai.

Iš (6.3.36) gauname, kad sąlyginės tikimybės

$$\pi_j(\boldsymbol{x})=\mathbf{P}\{\boldsymbol{Y}=\boldsymbol{e}_j|\mathbf{X}=\boldsymbol{x}\}=\frac{e^{\mu_j(\boldsymbol{x})}}{1+e^{\mu_1(\boldsymbol{x})}+...+e^{\mu_{m-1}(\boldsymbol{x})}},\quad j=1,...,m-1,$$

$$\pi_m(\boldsymbol{x})=\mathbf{P}\{\boldsymbol{Y}=\boldsymbol{e}_m|\mathbf{X}=\boldsymbol{x}\}=\frac{1}{1+e^{\mu_1(\boldsymbol{x})}+...+e^{\mu_{m-1}(\boldsymbol{x})}}. \tag{6.3.37}$$

Modelį apibūdina $(k+1)(m-1)=km+m-k-1$ nežinomas parametras.

Objektas, kurio $\mathbf{X}=\boldsymbol{x}$, priskirtinas j-ai klasei, jeigu sąlyginė tikimybė $\pi_j(\boldsymbol{x})$ yra didesnė už kitas tikimybes $\pi_l(\boldsymbol{x})$, $l\neq j$, t. y.

$$\frac{\pi_j(\boldsymbol{x})}{\pi_l(\boldsymbol{x})}\geq 1,\quad \forall\ l\neq j,\quad l=1,...,m.$$

Logaritmuodami gauname, kad ši taisyklė ekvivalenti tokiai: objektas priskiriamas j-ai klasei, jeigu

$$\mu_j(\boldsymbol{x})\geq 0,\quad \mu_j(\boldsymbol{x})\geq\mu_l(\boldsymbol{x}),\quad l\neq j,\quad j,l=1,...,m-1;$$

objektas priskiriamas m-ai klasei, jeigu

$$\mu_j(\boldsymbol{x})<0,\quad \forall\ j=1,...,m-1.$$

Klases atskiriame palygindami tiesines diskriminantines funkcijas $\mu_1(\boldsymbol{x}),...,\mu_{m-1}(\boldsymbol{x})$.

Tarkime, kad stebint a. v. $\mathbf{X} = (X_1, ..., X_k)^T$ gauta imtis $(\mathbf{X}_1^{(j)}, ..., \mathbf{X}_{n_j}^{(j)})$, kai objektas priklausė j-ai klasei, t. y. kai $\boldsymbol{Y} = \boldsymbol{e}_j$, $j = 1, ..., m$. DT metodu radę parametrų $\alpha_j, \boldsymbol{\beta}_j^T$ įvertinius $\hat{\alpha}, \hat{\boldsymbol{\beta}}_j^T$, galime įvertinti tikimybes

$$\hat{\pi}_j(\boldsymbol{x}) = \frac{e^{\hat{\mu}_j(\boldsymbol{x})}}{1 + e^{\hat{\mu}_1(\boldsymbol{x})} + ... + e^{\hat{\mu}_{m-1}(\boldsymbol{x})}}, \quad j = 1, ..., m-1,$$

$$\hat{\pi}_m(\boldsymbol{x}) = \frac{1}{1 + e^{\hat{\mu}_1(\boldsymbol{x})} + ... + e^{\hat{\mu}_{m-1}(\boldsymbol{x})}}. \tag{6.3.38}$$

Gauname įvertintą klasifikavimo taisyklę: objektas, kurio $\mathbf{X} = \boldsymbol{x}$, priskiriamas j-ai klasei,jeigu

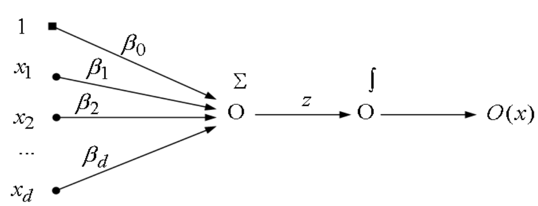
$$\hat{\mu}_j(\boldsymbol{x}) \geq 0, \quad \hat{\mu}_j(\boldsymbol{x}) \geq \hat{\mu}_l(\boldsymbol{x}), \quad l \neq j, \quad j, l = 1, ..., m-1;$$

objektas priskiriamas m-ai klasei, jeigu

$$\hat{\mu}_j(\boldsymbol{x}) < 0, \forall \ j = 1, ..., m-1.$$

6.3.3.5. Neuroniniai tinklai

Pastaruoju metu labai intensyviai vystomi klasifikavimo algoritmai, grindžiami vadinamaisiais *neuroniniais tinklais*. Neuroniniu tinklu vadinamas matematinis modelis, gautas imituojant nervų sistemą, sudarytą iš neuronų ir jų jungčių. Paprasčiausias neuroninis tinklas yra perceptronas, turintis tik vieną neuroną, transformuojantį įėjimo vektorių $\boldsymbol{x} = (x_1, ..., x_k)^T$ ir imituojantį išėjimą $g(\boldsymbol{x})$.



6.6.1 pav. Neuroninis tinklas su vienu neuronu

Pažymėkime

$$z = z(\boldsymbol{x}) = \beta_0 + \boldsymbol{\beta}^T \boldsymbol{x};$$

tada kiekvienas įėjimas x_j sustiprinamas ($\beta_j > 1$) arba susilpninamas ($\beta_j < 1$) ir jie sumuojami. Gautas signalas z transformuojamas pasitelkus perdavimo funkciją (6.6.1 pav. simbolis $\int$) ir gaunama išėjimo funkcija $g(\boldsymbol{x})$. Jeigu perdavimo funkcija yra sigmoidė

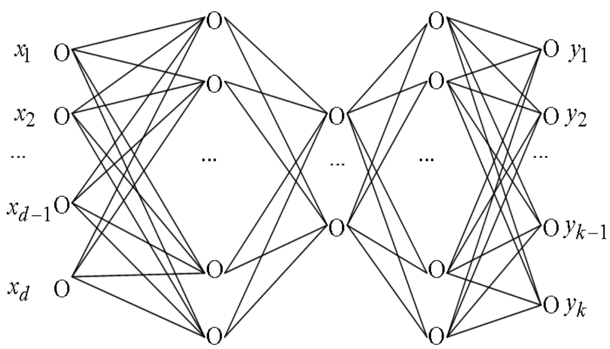
$$\varphi(z) = \frac{e^z}{1 + e^z},$$

tai perceptrono matematinis modelis sutampa su logistine regresija $g(\boldsymbol{x}) = \varphi(z(\boldsymbol{x}))$.

Bendriau, tarkime, kad turime keletą neuronų ir perdavimo funkcijas $\varphi(z_j(\boldsymbol{x}))$; čia $z_j = \beta_{0j} + \beta_{1j}x_1 + ... + \beta_{kj}x_k, j = 1, ..., m$. Imame jų sumą. Tada išėjimą galima išreikšti taip

$$g(\boldsymbol{x}) = \omega_0 + \sum_{j=1}^{m} \omega_j \varphi(z_j(\boldsymbol{x})). \tag{6.3.39}$$

Ši išraiška apibūdina vieno paslėpto sluoksnio neuroninį tinklą, kuriame yra m neuronų (6.6.2 pav).



6.6.2 pav. Vieno paslėpto sluoksnio neuroninis tinklas

Bendru atveju neuroniniame tinkle gali būti keletas paslėptų neuronų sluoksnių, turinčių skirtingą neuronų skaičių. Kiekvieno sluoksnio neuronų įėjimai yra ankstesnio sluoksnio išėjimai.

Vertinant nežinomus parametrus dažniausiai naudojami kriterijai, susiję su klasifikavimo klaidų ar jų funkcijų minimizavimu. Todėl ir perceptrono atveju atsakymai gali skirtis nuo logistinės regresijos, kurioje nežinomi parametrai vertinami DT metodu.

Daugelyje matematinės statistikos TPP yra sukurtos ir įtrauktos paprogramės, kurios remiantis apmokančiosiomis imtimis leidžia parinkti tam tikra prasme geriausią neuroninio tinklo architektūrą, įvertinti parametrus ir atlikti gautojo klasifikatoriaus tikslumo patikrinimą su testine aibe. Plačiau apie neuroninius tinklus žr.[15].

6.4. Pratimai

6.1 skyrelis

6.1. Klasifikuojant į dvi klases remiamasi a.d. X ir $(X|\xi = 1) \sim K(\mu_1, \sigma)$, $(X|\xi = 2) \sim K(\mu_2, \sigma)$, $\mu_1 < \mu_2$. Tegu $\omega_1 = \omega_2 = 1/2$. a) Raskite Bejeso klasifikavimo taisyklę, atitinkančią šias apriorines klasių tikimybes, ir apskaičiuokite klasifikavimo klaidą. b) Tarkim, nuostoliai apibrėžiami kainomis $c_{11} = c_{22} = 0, c_{12} = 4, c_{21} = 1$. Raskite taisyklę, minimizuojančią vidutinius nuostolius.

6.2. (**6.1** pratimo tęsinys). Tegu $\sigma = 1, \mu_1 = 0, \mu_2 = 3$. a) Tarkime, apriorinės klasių tikimybės nežinomos, o nuostoliai apibrėžti **6.1** pratimo p. b). Raskite taisyklę, tenkinančią minimakso principą. b) Tarkime, apriorinės klasių tikimybės $\omega_1 = \omega_2 = 1/2$ ir yra galimybė atsisakyti priimti galutinį sprendimą. Raskite taisyklę, minimizuojančią atsisakymų priimti sprendimą tikimybes, kai apribotos tikimybės $\beta_{12} \leq 0,1$, $\beta_{21} \leq 0,1$.

6.3. Įrodykite, kad dviejų klasių atveju Bejeso klasifikatoriaus suminė klaidų tikimybė tenkina sąlygą

$$\mathbf{P}\{e\} \leq \sqrt{\omega_1\omega_2}\rho \leq \rho/2, \quad \rho = \int [f_1(\boldsymbol{x})f_2(\boldsymbol{x})]^{1/2} d\boldsymbol{\mu}.$$

6.4. Klasifikuojant į dvi klases remiamasi a.v. $\mathbf{X} = (X_1, ..., X_k)^T$ stebėjimu. Vektoriaus $\mathbf{X}$ koordinatės nepriklausomos ir turi Bernulio skirstinius $(X_i|\xi = 1) \sim B(1, p_i), (X_i|\xi = 2) \sim B(1, q_i), i = 1, ..., k$, o apriorinės klasių tikimybės yra $\omega_1, \omega_2 = 1 - \omega_1$. Tegu rizikos funkcija nusakoma kainomis $c_{21} > c_{11} \geq 0$, $c_{12} > c_{22} \geq 0$. Raskite diskriminantinę funkciją, minimizuojančią vidutinius nuostolius.

6.5. (**6.4** pratimo tęsinys). Tarkime, tikimybės $p_i = p > 1/2$, $q_i = 1 - p$, $i = 1, ..., k$, $c_{12} - c_{22} = c_{21} - c_{11}$, o apriorinės klasių tikimybės vienodos $\omega_1 = \omega_2 = 1/2$. Raskite Bejeso klasifikavimo taisyklę ir klasifikavimo klaidą. Aptarkite klasifikavimo klaidos ribą, kai $p \to 1/2$; $k \to \infty$.

6.6. Klasifikuojant į dvi klases remiamasi a.d. X ir $(X|\xi = 1) \sim \mathcal{E}(\lambda_1)$, $(X|\xi = 2) \sim \mathcal{E}(\lambda_2)$, $\lambda_1 < \lambda_2$. Tegu $\omega_1 = \omega_2 = 1/2$, nuostoliai apibrėžiami kainomis $c_{11} = c_{22} = 0, c_{12} = 4, c_{21} = 2$. Raskite taisyklę, minimizuojančią vidutinius nuostolius.

6.7. (**6.6** pratimo tęsinys). Tegu $\lambda_1 = 1, \lambda_2 = 5$ ir apriorinės klasių tikimybės nežinomos. Rakite taisyklę, tenkinančią minimakso principą.

6.8. (**6.6** pratimo tęsinys). Tarkime, yra galimybė atsisakyti priimti galutinį sprendimą. Nuostoliai dėl atsisakymų priimti sprendimą apsprendžiami kainomis $c_{01} = c_{02} = 1$. Raskite taisyklę, minimizuojančią vidutinius nuostolius.

6.9. Klasifikuojant į dvi klases remiamasi a.d. X ir $(X|\xi = 1) \sim N(0,\ 1)$, $(X|\xi = 2) \sim K(0,\ 1)$. Tegu $\omega_1 = \omega_2 = 1/2$. a) Raskite klasifikavimo taisyklę, minimizuojančią vidutinę klaidos tikimybę, ir apskaičiuokite tą tikimybę. b) Tegu nuostolius apsprendžia kainos $c_{11} = c_{22} = 0$, $c_{12} = 1$, $c_{21} = 3$. Raskite taisyklę, minimizuojančią vidutinius nuostolius, ir šios taisyklės klasifikavimo klaidos tikimybę.

6.10. (**6.9** pratimo tęsinys). Tarkime, apriorinės klasių tikimybės nežinomos. Raskite taisyklę $\boldsymbol{\varphi}^*$, minimizuojančią vidutinius nuostolius $R(\boldsymbol{\varphi}, \omega_1)$, ir grafiškai pavaizduokite funkciją $R(\boldsymbol{\varphi}^*, \omega_1)$ argumentu imdami apriorinę tikimybę $\omega_1 = \mathbf{P}\{\xi = 1\}, 0 \leq \omega_1 \leq 1$.

6.11. (**6.9** pratimo tęsinys). Raskite taisyklę, maksimizuojančią tikimybę α_{11}, kai yra apribojimas $\beta_{21} \leq 0,1$.

6.12. Tarkime, kad a.d. X skirstiniai dviejų klasių atveju yra $(X|\xi = 1) \sim B(3, p_1), (X|\xi = 2) \sim B(3, p_2)$; $p_1 = 0,01, p_2 = 0,1$. Nuostolius apsprendžia kainos $c_{11} = c_{22} = 0$, $c_{12} = 1$, $c_{21} = 2$. a) Nagrinėjame tris nerandomizuotas sprendimų priėmimo taisykles: $\varphi_1^{(i)}(x) = 1$, kai $x = 0, ..., i$, $\varphi_1^{(i)}(x) = 0$, kai $x = i+1, ..., 3$; $\varphi_2^{(i)} = 1 - \varphi_1^{(i)}$, $i = 0; 1; 2$. Kuri iš šių trijų taisyklių tenkina minimakso principą ? b) Raskite sprendimų priėmimo taisyklę, minimizuojančią vidutinius nuostolius, kai apriorinės klasių tikimybės yra (ω_1, ω_2).

6.13. Tarkime, kad a.d. X skirstiniai dviejų klasių atveju yra geometriniai ir jų parametrai $p_1 = 0,2, p_2 = 0,1$. Nuostoliai apsprendžiami kainomis $c_{11} = c_{22} = 0$, $c_{12} = 2$, $c_{21} = 1$. a) Nagrinėjame penkias nerandomizuotas sprendimų priėmimo taisykles: $\varphi_1^{(i)}(x) = 1$, kai $x = 0, ..., i$, $\varphi_1^{(i)}(x) = 0$, kai $x = i+1, i+2, ...$; $\varphi_2^{(i)} = 1 - \varphi_1^{(i)}$, $i = 0; 1; 2; 3; 4$. Kuri iš šių trijų taisyklių tenkina minimakso principą ? b) Raskite sprendimų priėmimo taisyklę, minimizuojančią vidutinius nuostolius, kai apriorinės klasių tikimybės yra (ω_1, ω_2).

6.14. Tarkime, kad a. d. X skirstiniai dviejų klasių atveju yra Puasono ir jų parametrai $\lambda_1 = 1, \lambda_2 = 2$. Nuostolius apsprendžia kainos $c_{11} = c_{22} = 0$, $c_{12} = 100$, $c_{21} = 50$. a) Nagrinėjame septynias nerandomizuotas sprendimų priėmimo taisykles: $\varphi_1^{(1)}(x) \equiv 0$; $\varphi_1^{(i)}(x) = 1$, kai $x = 0, ..., i-1$, $\varphi_1^{(i)}(x) =$, kai $x = i, i+1, ...$; $\varphi_2^{(i)}(x) = 1 - \varphi_1^{(i)}$, $i = 2; 3; 4; 5; 6$; $\varphi_1^{(7)}(x) \equiv 1$. a) Pavaizduokite grafiškai šių sprendimų rizikos funkcijos komponentes $R_1(\boldsymbol{\varphi}^{(i)}), R_2(\boldsymbol{\varphi}^{(i)})$. b) Įvedę randomizaciją, raskite minimakso sprendinį ir pavaizduokite jį grafiškai.

6.15. Tarkime, kad a. d. X skirstiniai dviejų klasių atveju yra Bernulio $(X|\xi = 1) \sim B(1,\ 3/4), (X|\xi = 2) \sim B(1,\ 1/3)$. Nuostolius apsprendžia kainos $c_{11} = c_{22} = 0$, $c_{12} = 5$, $c_{21} = 10$. Tegu apriorinės klasių tikimybės yra $\omega_1, \omega_2 = 1 - \omega_1$. Raskite sprendimų priėmimo taisyklę, minimizuojančią vidutinius nuostolius, ir pavaizduokite juos grafiškai argumentu imdami ω_1. Su kokiais ω_1 atlikti a, d. X stebėjimą neapsimoka, jeigu stebėjimo kaina $c = 0,1$?

6.16. (**6.15** pratimo tęsinys). Tarkime, kad **6.15** pratimo sąlygomis prieš priimant sprendimą galima stebėti a. v. $\mathbf{X} = (X_1, ..., X_n)^T$, kurio koordinatės yra nepriklausomi a. d., turintys Bernulio skirstinius $B(1,\ 3/4)$ pirmos klasės atveju ir $B(1,\ 1/3)$ antros klasės atveju. a) Raskite sprendimų priėmimo taisyklę, minimizuojančią vidutinius nuostolius, atsižvelgiant į apriorines klasių tikimybes (ω_1, ω_2). b) Raskite imties didumą n, kuriam esant gaunami mažiausieji vidutiniai nuostoliai, jeigu apriorinės klasių tikimybės vienodos, o vieno matavimo kaina $c = 0,01$.

6.17. Sprendimai $\eta = 1; 2; 0$ priimami remiantis a. v. $\mathbf{X}$, kurio tankiai $(\mathbf{X}|\xi = 1) \sim f_1(\boldsymbol{x})$ ir $(\mathbf{X}|\xi = 2) \sim f_2(\boldsymbol{x})$. Nuostolius apsprendžia kainos $c_{11} = c_{22} = 0$, $c_{12} = 10$, $c_{21} = 8, c_{01} = 4, c_{02} = 3$. Su kokiomis apriorinėmis dviejų klasių tikimybėmis ω_1 ir $\omega_2 = 1 - \omega_1$ sprendimas $\eta = 0$ minimizuoja vidutinius nuostolius, kai stebėta reikšmė $\mathbf{X} = \boldsymbol{x}$?

6.18. Tarkime, kad a. d. X skirstiniai dviejų klasių atveju yra Bernulio $(X|\xi = 1) \sim B(1,\ 3/4), (X|\xi = 2) \sim B(1,\ 1/4)$. Priimami trys sprendimai $\eta = 1,\ 2,\ 0$. Nuostolius apsprendžia kainos $c_{11} = c_{22} = 0$, $c_{12} = c_{21} = 10$, $c_{01} = c_{02} = 3$. Tegu apriorinės klasių tikimybės yra $\omega_1, \omega_2 = 1 - \omega_1$. Raskite sprendimų priėmimo taisyklę, minimizuojančią vidutinius nuostolius, ir pavaizduokite ją grafiškai argumentu imdami ω_1.

6.19. (**6.18** pratimo tęsinys). Tarkime, kad **6.18** pratimo sąlygomis prieš priimant sprendimą galima stebėti a. v. $\mathbf{X} = (X_1, ..., X_n)^T$, kurio koordinatės yra nepriklausomi a. d., turintys Bernulio skirstinius $B(1,\ 3/4)$ pirmos klasės atveju ir $B(1,\ 1/4)$ antros klasės atveju. Raskite imties didumą n, kuriam esant gaunami mažiausieji vidutiniai nuostoliai, jeigu apriorinės klasių tikimybės vienodos, o vieno matavimo kaina $c = 0,1$.

6.20. Tarkime, kad klasifikuojant objektus į dvi klases galima atlikti n nepriklausomų a. d. stebėjimų, kurių skirstiniai pirmos klasės atveju yra normalieji $N(-1,\ 9)$, o antros klasės atveju – $N(1,\ 9)$. Nuostolius apsprendžia kainos $c_{11} = c_{22} = 0$, $c_{12} = c_{21} = 1000$, o vieno stebėjimo kaina lygi $c = 1$. Raskite optimalų imties didumą ir nuostolių funkcijos minimalią reikšmę, jeigu apriorinės klasių tikimybės yra vienodos.

6.21. Tegu $X_1, ..., X_n$, yra n. a. d., vieno iš kurių (jo numeris nežinomas) tankis yra $g(x)$, o visų likusių koordinačių tankiai yra vienodi ir lygūs $f(x)$. Tarkime, $\omega_i > 0$ reiškia, kad a. d. X_i tankis yra $g(x)$, $\omega_1 + ... + \omega_n = 1$. Raskite aposteriorinę tikimybę to, kad a. d. X_1 tankis yra $g(x)$, jei buvo stebėta reikšmė $\boldsymbol{x} = (x_1, ..., x_n)$.

6.22. Tarkim, dviejų klasių aposteriorinės tikimybės yra ω_1 ir ω_2, o a. d. X tankiai yra $(X|\xi = i) \sim f_i(x)$, $i = 1,\ 2$. Pažymėkime $\omega_1(x) = \mathbf{P}\{\xi = 1|X = x\}$ aposteriorinę pirmos klasės tikimybę. a) Įrodykite, kad $\mathbf{E}(\omega_1(X)) = \omega_1$. b) Įrodykite, kad $\mathbf{E}(\omega_1(X)|\xi = 1) \geq \omega_1$.

6.23. Tarkime, kad dviejų klasių atveju stebimo a. v. skirstiniai yra normalieji $N_k(\boldsymbol{\mu}_1, \mathbf{\Sigma})$ ir $N_k(\boldsymbol{\mu}_2, \mathbf{\Sigma})$ su vienodomis kovariacinėmis matricomis. Turimos paprastosios nepriklausomos apmokančiosios imtys $\mathbf{X}_i^{(1)}, i = 1, ..., n_1$ ir $\mathbf{X}_i^{(2)}, i = 1, ..., n_2$. Reikia naują nepriklausomą stebėjimą $\mathbf{X}$ priskirti vienai iš dviejų klasių. Tikrinkime hipotezę $H_1 : \mathbf{X}, \mathbf{X}_1^{(1)}, ..., \mathbf{X}_{n_1}^{(1)} \sim$

$N_k(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}), \mathbf{X}_1^{(2)}, ..., \mathbf{X}_{n_2}^{(2)} \sim N_k(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$ su alternatyviąja hipoteze $H_2 : \mathbf{X}_1^{(1)}, ..., \mathbf{X}_{n_1}^{(1)} \sim N_k(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$, $\mathbf{X}, \mathbf{X}_1^{(2)}, ..., \mathbf{X}_{n_2}^{(2)} \sim N_k(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$. Įrodykite, kad tikėtinumų santykio statistika šioms hipotezėms tikrinti yra

$$\Lambda = \frac{1 + n_2(\mathbf{X} - \bar{\mathbf{X}}^{(2)})^T \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \bar{\mathbf{X}}^{(2)})/(n_2 + 1)}{1 + n_1(\mathbf{X} - \bar{\mathbf{X}}^{(1)})^T \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \bar{\mathbf{X}}^{(1)})/(n_1 + 1)}.$$

6.2 skyrelis

6.24. Tarkime, objektas priskiriamas i-ajai klasei, kai aposteriorinė tikimybė $\omega_i(\boldsymbol{x}) \geq \omega_j(\boldsymbol{x})$, $j \neq i$, $i, j = 1, ..., m$. Įrodykite, kad $\omega_i(\boldsymbol{x}) \geq 1/m$, o suminė klaidos tikimybė neviršija $(m-1)/m$. Pateikite pavyzdį, kai suminės klaidos tikimybė lygi $(m-1)/m$.

6.25. Tarkime, objektai klasifikuojami į m klasių esant galimybei atsisakyti priimti sprendimą, o nuostolius apsprendžia kainos $c_{ii} = 0, i = 1, ..., m$, $c_{ij} = c > 0, i \neq j, i, j = 1, ..., m$, $c_{0i} = c_0 > 0, i = 1, ..., m$. Tegu apriorinės klasių tikimybės yra $\omega_i, \omega_1 + ... + \omega_m = 1$. Raskite klasifikatoriaus, minimizuojančio vidutinius nuostolius, diskriminantines funkcijas. Aptarkite gautą taisyklę, kai $c_0 \to 0$; kai $c_0 > c$.

6.26. Tegu vektoriaus $\mathbf{X} = (X_1, ..., X_k)^T$ koordinatės yra nepriklausomos ir turi Bernulio skirstinius $(X_i|\xi = j) \sim B(1, p_{ij}), i = 1, ..., k, j = 1, ..., m$. Raskite klasifikatoriaus, minimizuojančio vidutinę klaidos tikimybę, diskriminantines funkcijas.

6.27. Tegu a.d. X skirstiniai trijų klasių atveju yra: $(X|\xi = 1) \sim N(0,\ 1)$; $(X|\xi = 2) \sim K(0,\ 1)$; $(X|\xi = 3) \sim L(1)$; čia $L(1)$ žymi Laplaso skirstinį, kurio tankis $f_3(x) = \exp\{-|x|\}/2, -\infty < x < \infty$. Tegu apriorinės klasių tikimybės yra vienodos, o nuostolius apsprendžia kainos $c_{11} = c_{22} = c_{33} = 0$; $c_{ij} = 5, i \neq j = 1, 2, 3$. Raskite klasifikavimo taisyklę, minimizuojančią vidutinį nuostolį, ir apskaičiuokite jo reikšmę.

6.28. (**6.27** pratimo tęsinys). Tarkime, yra galimybė atsisakyti priimti sprendimą, o atsisakymo priimti sprendimą nuostolius apsprendžia kainos $c_{0i} = 3, i = 1, 2, 3$. Raskite klasifikavimo taisyklę, minimizuojančią vidutinį nuostolį, ir apskaičiuokite jo reikšmę.

6.29. Tegu a.v. $\mathbf{X}$ tankiai trijų klasių atveju yra: $(\mathbf{X}|\xi = i) \sim f_i(\boldsymbol{x}), i = 1, 2, 3$. Tegu apriorinės klasių tikimybės yra vienodos, o nuostolius apsprendžia kainos $c_{11} = c_{22} = c_{33} = 0$; $c_{ij} = 4, i \neq j = 1, 2, 3$. Raskite klasifikavimo taisyklę, minimizuojančią vidutinį nuostolį, aposteriorinių klasių tikimybių $\omega_i(\boldsymbol{x}) = \mathbf{P}\{\xi = i|\mathbf{X} = \boldsymbol{x}\}, i = 1, 2, 3$, terminais ir pavaizduokite ją grafiškai.

6.30. (**6.29** pratimo tęsinys). Tarkime, yra galimybė atsisakyti priimti sprendimą, o atsisakymo priimti sprendimą nuostolius apsprendžia kainos $c_{0i} = 2; i = 1, 2, 3$. Raskite klasifikavimo taisyklę, minimizuojančią vidutinį nuostolį, aposteriorinių klasių tikimybių $\omega_i(\boldsymbol{x})$ terminais ir pavaizduokite ją grafiškai.

6.31. Tegu a.v. $\mathbf{X} = (X, Y)^T$ skirstiniai trijų klasių atveju yra: $(\mathbf{X}|\xi = i) \sim N_2(\boldsymbol{\mu}_i, \boldsymbol{I})$, kai vidurkių vektoriai išsidėstę taisyklingojo trikampio viršūnėse: $\boldsymbol{\mu}_1 = (-1,5; 0)$, $\boldsymbol{\mu}_2 = (1,5; 0)$; $\boldsymbol{\mu}_3 = (0; 3\sqrt{3}/2)$. Tegu apriorinės klasių tikimybės yra vienodos ir priimami trys sprendimai $\eta = i, i = 1, 2, 3$. Raskite Bejeso klasifikavimo taisyklę, minimizuojančią vidutinę klasifikavimo klaidą, ir suraskite jos reikšmę. Pavaizduokite sprendinį geometriškai a.v. $\mathbf{X}$ ir aposteriorinių klasių tikimybių terminais.

6.32. (**6.31** pratimo tęsinys). Tarkime yra galimybė atsisakyti nuo sprendimo priėmimo, o atsisakymo priimti sprendimą nuostoliai apsprendžiami kainomis $c_{ii} = 0; c_{ij} = 50, i \neq j = 1, 2, 3; c_{0i} = 30, i = 1, 2, 3$. Raskite klasifikavimo taisyklę minimizuojančią vidutinį nuostolį ir pavaizduokite ją grafiškai a.v. $\mathbf{X}$ ir aposteriorinių klasių tikimybių terminais.

6.33. (**6.32** pratimo tęsinys). Apskaičiuokite **6.32** pratime rastos klasifikavimo taisyklės vidutinį nuostolį.

6.3 skyrelis

6.34. Pagal paprastąją imtį $\mathbf{X}_1, ..., \mathbf{X}_n$, gautą stebint normalųjį a.v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, gauti įvertiniai

$$\hat{\boldsymbol{\mu}}_n = \bar{\mathbf{X}}_n = \frac{1}{n}\sum_{j=1}^{n}\mathbf{X}_j, \quad \hat{\boldsymbol{\Sigma}}_n = \frac{1}{n-1}\boldsymbol{S} = \frac{1}{n-1}\sum_{j=1}^{n}(\mathbf{X}_j - \bar{\mathbf{X}}_n)(\mathbf{X}_j - \bar{\mathbf{X}}_n)^T.$$

Tarkime, papildomai gautas dar vienas nepriklausantis nuo ankstesnių stebėjimas $\mathbf{X}_{n+1}$. Įrodykite, kad

$$\bar{\mathbf{X}}_{n+1} = \bar{\mathbf{X}}_n + \frac{1}{n+1}(\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n),$$

$$\hat{\boldsymbol{\Sigma}}_{n+1} = \frac{n-1}{n}\hat{\boldsymbol{\Sigma}}_n + \frac{1}{n+1}(\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n)(\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n)^T.$$

6.35.(6.34 tęsinys). Įrodykite, kad $\hat{\boldsymbol{\Sigma}}_{n+1}^{-1}$ galima surasti tokiu būdu

$$\hat{\boldsymbol{\Sigma}}_{n+1}^{-1} = \frac{n}{n-1}\left[\hat{\boldsymbol{\Sigma}}_n^{-1} - \frac{\hat{\boldsymbol{\Sigma}}_n^{-1}(\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n)(\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n)^T\hat{\boldsymbol{\Sigma}}_n^{-1}}{(n^2-1)/n + (\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n)^T\hat{\boldsymbol{\Sigma}}_n^{-1}(\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n)}\right].$$

6.36. Tarkime, kad paprastoji imtis $\mathcal{X} = (X_1, ..., X_n)$ gauta stebint a.d. X, kuris esant fiksuotam vidurkiui $\mu = \mathbf{E}X$ turi normalųjį skirstinį $(X|\mu) \sim N(\mu, \sigma^2)$, parametras σ žinomas. Vidurkis μ savo ruožtu yra a.d., turintis normalųjį skirstinį $\mu \sim N(\mu_0, \beta^2)$ su žinomais parametrais μ_0, β^2. Raskite parametro μ Bejeso įvertinį. Aptarkite gautojo įvertinio elgesį, kai $n \to \infty$, o β fiksuotas ir kai $\beta \to 0$, o n fiksuotas. Raskite a.d. X sąlyginį skirstinį imties $\mathcal{X}$ atžvilgiu.

6.37. Tarkime, kad paprastoji imtis $\mathcal{X} = (X_1, ..., X_n)$ gauta stebint a.d. X, kuris esant fiksuotam vidurkiui $\lambda = \mathbf{E}X$ turi Puasono skirstinį $(X|\lambda) \sim \mathcal{P}(\lambda)$. Vidurkis λ savo ruožtu yra a.d., turintis gama skirstinį $\lambda \sim G(\gamma, \eta)$ su žinomais parametrais γ, η. Raskite parametro λ Bejeso įvertinį ir a.d. X sąlyginį skirstinį imties $\mathcal{X}$ atžvilgiu.

6.38. Tarkime, kad paprastoji imtis $\mathcal{X} = (X_1, ..., X_n)$ gauta stebint a.d. X, kuris esant fiksuotam parametrui p turi binominį skirstinį $(X|p) \sim B(m, p)$. Tikimybė p savo ruožtu yra a.d., turintis beta skirstinį $p \sim Be(\gamma, \eta)$ su žinomais parametrais γ, η. Raskite parametro p Bejeso įvertinį ir a.d. X sąlyginį skirstinį imties $\mathcal{X}$ atžvilgiu.

6.39. Tarkime, kad paprastoji imtis $\mathcal{X} = (X_1, ..., X_n)$ gauta stebint a.d. X, kuris esant fiksuotam parametrui p turi neigiamąjį binominį skirstinį $(X|p) \sim B^-(m, p)$, parametras m žinomas. Tikimybė p savo ruožtu yra a.d., turintis beta skirstinį $p \sim Be(\gamma, \eta)$ su žinomais parametrais γ, η. Raskite parametro p Bejeso įvertinį ir. a.d. X sąlyginį skirstinį imties $\mathcal{X}$ atžvilgiu.

6.40. Tarkime, kad paprastoji imtis $\mathcal{X} = (X_1, ..., X_n)$ gauta stebint a.d. X, kuris esant fiksuotam parametrui λ turi eksponentinį $(X|\lambda) \sim \mathcal{E}(\lambda)$ skirstinį. Parametras λ savo ruožtu yra a.d., turintis gama skirstinį $\lambda \sim G(\gamma, \eta)$ su žinomais parametrais γ, η. Raskite parametro λ Bejeso įvertinį ir a.d. X sąlyginį skirstinį imties $\mathcal{X}$ atžvilgiu.

6.41. Tarkime, kad paprastoji imtis $\mathcal{X} = (X_1, ..., X_n)$ gauta stebint a.d. X, kuris esant fiksuotam parametrui σ turi normalųjį skirstinį $(X|\sigma) \sim N(\mu, \sigma^2)$, vidurkis μ žinomas. Parametras $\theta = 1/\sigma^2$ savo ruožtu yra a.d., turintis gama skirstinį $\theta \sim G(\lambda, \eta)$ su žinomais parametrais λ, η. Raskite parametro θ aposteriorinį skirstinį ir a.d. X sąlyginį skirstinį imties $\mathcal{X}$ atžvilgiu.

6.42. Tarkime, kad paprastoji imtis $\mathcal{X} = (\mathbf{X}_1, ..., \mathbf{X}_n)$ gauta stebint a.v. $\mathbf{X}$, kuris turi polinominį skirstinį $\mathcal{P}_k(1, \boldsymbol{p})$, $\boldsymbol{p} = (p_1, ..., p_k)^T$, $0 \le p_i \le 1$, $p_1 + ... + p_k = 1$. Parametro $\boldsymbol{p}$ apriorinis skirstinys yra Dirichlė skirstinys su žinomu parametrų vektoriumi $(\alpha_1, ..., \alpha_k)^T$, $\alpha_i > 0$. Raskite parametro $\boldsymbol{p}$ aposteriorinį skirstinį ir a.v. $\mathbf{X}$ sąlyginį skirstinį imties $\mathcal{X}$ atžvilgiu.

6.43. Tarkime, a. v. $\mathbf{X} = (X_1, ..., X_k)^T$ skirstinys yra daugiamatis Bernulio:

$$\mathbf{P}\{\mathbf{X} = \boldsymbol{x}|\boldsymbol{\theta}\} = \prod_{i=1}^{k} \theta_i^{x_i}(1-\theta_i)^{1-x_i}, \quad \boldsymbol{x} = (x_1, ..., x_k)^T,$$

o parametras $\boldsymbol{\theta} = (\theta_1, ..., \theta_k)^T$ turi tolygųjį apriorinį skirstinį vienetiniame k-mačiame kube. Gauta paprastoji a. v. $\mathbf{X}$ imtis $\mathcal{X} = (\mathbf{X}_1, ..., \mathbf{X}_n)$. Raskite parametro $\boldsymbol{\theta}$ aposteriorinį skirstinį ir a. v. $\mathbf{X}$ sąlyginį skirstinį imties $\mathcal{X}$ atžvilgiu.

6.44. Defektinio gaminio pagaminimo tikimybė p turi apriorinį skirstinį $U(0;1)$. Pagal paprastąją didumo n imtį surastas p aposteriorinis skirstinys yra $Be(7;95)$. Kokio didumo imtis buvo tikrinta ir koks defektinių gaminių skaičius joje buvo aptiktas?

6.45. Tarkim, kad vidutinis defektų skaičius λ pagamintoje tam tikro ilgio magnetinėje juostoje nežinomas, tačiau turi apriorinį gama skirstinį, kurio vidurkis 2 ir dispersija 1. Tegu defektų skaičius juostoje esant fiksuotam λ turi Puasono skirstinį $\mathcal{P}(\lambda)$. Patikrinus n juostų pasirodė, kad aposteriorinio λ skirstinio vidurkis lygus 1,6, o dispersija 0,16. Kokio didumo imtis buvo tikrinta ir kiek defektų surasta?

6.46. Defektinio gaminio pagaminimo tikimybė p turi apriorinį skirstinį $Be(1;99)$. Gaminiai atsitiktinai atrenkami ir tikrinami, kol bus rasti 5 defektiniai. Aposteriorinio skirstinio vidurkis 0,02. Kiek gerų gaminių buvo surasta tikrinimo metu?

6.47. Pritariančių liberalių partijų pažiūroms rinkėjų dalies p apriorinis skirstinys yra $Be(1;10)$. a) Tegu tarp 1000 atsitiktinai apklaustų rinkėjų 123 pritaria liberalių partijų pažiūroms. Koks tokių rinkėjų dalies aposteriorinis skirstinys? b) Tarkime, rinkėjai atsitiktinai apklausiami tol, kol atsiras 123 rinkėjai, pritariantys liberalių partijų pažiūroms. Koks tokių rinkėjų dalies aposteriorinis skirstinys?

6.48. Gaminamų elektros lempučių darbo laikas turi eksponentinį skirstinį su parametru λ. Savo ruožtu parametras λ turi apriorinį gama skirstinį su variacijos koeficientu 0,5. Kokio didumo turi būti paprastoji imtis, kad aposteriorinio skirstinio variacijos koeficientas sumažėtų iki 0,1?

6.49. A. d. X turi normalųjį skirstinį $N(\mu;4)$. Apriorinio μ skirstinio dispersija $\beta^2 = 9$. Kokią mažiausią paprastąją imtį reikia turėti, kad būtų galima sukonstruoti parametro μ ilgio 1 pasikliovimo intervalą, kai pasikliovimo lygmuo $Q = 0,95$?

6.50. Tarkime, $\mathbf{X} = (X_1, X_2)^T \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$; vidurkių vektorius $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$ nežinomas, o $\boldsymbol{\Sigma}^{-1} = [\sigma^{ij}]_{2\times 2}$ žinoma; $\sigma^{11} = 1/3$, $\sigma^{22} = 4/3$, $\sigma^{12} + \sigma^{21} = -1/3$. Vektoriaus $\boldsymbol{\mu}$ apriorinis skirstinys yra dvimatis normalusis su kovariacine matrica $\boldsymbol{\Sigma}_0$; $\boldsymbol{\Sigma}_0^{-1} = [\beta^{ij}]_{2\times 2}$ žinoma; $\beta^{11} = 1, \beta^{22} = 6$, $\beta^{12} = \beta^{21} = -1$. Koks turi būti imties didumas, kad aposteriorinio skirstinio a. d. $\mu_1 - \mu_2$ dispersija neviršytų 0,01?

6.51. Gaminiai apibūdinami parametru X, kurio skirstinys normalusis $X \sim N(\mu, \sigma^2)$. Gaminys yra geras, kai $a < X < b$, ir defektinis priešingu atveju. Kontrolės metu stebime a. d. $Y = X + e$; čia matavimo paklaida e nepriklauso nuo X ir turi normalųjį skirstinį $e \sim N(0, \beta^2)$. Raskite Bejeso klasifikavimo taisyklę ir suminę klaidą, kai $\mu = 0, \sigma^2 = 1, \beta^2 = 1/4, -a = b = 1,6$.

6.52. (**6.51** pratimo tęsinys). Tarkime, kad parametrai μ, σ^2, β^2 nežinomi. Modeliuodami a. d. su nurodytomis **6.51** pratime parametrų reikšmėmis, gaukite apmokančiąsias imtis. Įvertinkite skirtingų klasių a. d. Y tankius dviem būdais: a) didžiausiojo tikėtinumo metodu įvertinę parametrus μ, σ^2, β^2; b) naudodami neparametrinius tankio įverčius branduoliu imdami normaliojo skirstinio tankį. Raskite klasifikavimo taisyklės įverčius ir palyginkite juos su **6.51** pratime gauta taisykle.

6.53. Tarkime, kad trijų klasių atveju stebimo vektoriaus $\mathbf{X}$ skirstinys yra normalusis $(\mathbf{X}|\xi = i) \sim N_k(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}), i = 1,2,3$, su vienoda kovariacine matrica. Raskite klasifikatorių minimizuojantį suminę klasifikavimo klaidą $\sum_{i\neq j} \alpha_{ij}$ ir raskite klasifikavimo klaidų tikimybes $\alpha_{ij}, i \neq j = 1,2,3$.

6.54 (**6.53** pratimo tęsinys). Klasikiniu tapusiame Fišerio eksperimente [8] buvo matuota $n = 50$ vilkdalgio *Iris setosa* taurėlapio ilgis ir plotis, bei vainiklapio ilgis ir plotis (žymėsime (X_1, X_2, X_3, X_4)). Analogiški matavimai buvo atlikti kitos vilkdalgio rūšies *Iris versicolor* (žymėsime (Y_1, Y_2, Y_3, Y_4)) ir trečios vilkdalgio rūšies *Iris virginica* (žymėsime (Z_1, Z_2, Z_3, Z_4)). Matavimo rezultatai pateikiami lentelėse.

X_1	X_2	X_3	X_4	Y_1	Y_2	Y_3	Y_4	X_1	X_2	X_3	X_4	Y_1	Y_2	Y_3	Y_4
5,1	3,5	1,4	0,2	7,0	3,2	4,7	1,4	5,0	3,0	1,6	0,2	6,6	3,0	4,4	1,4
4,9	3,0	1,4	0,2	6,4	3,2	4,5	1,5	5,0	3,4	1,6	0,4	6,8	2,8	4,8	1,4
4,7	3,2	1,3	0,2	6,9	3,1	4,9	1,5	5,2	3,5	1,5	0,2	6,7	3,0	5,0	1,7
4,6	3,1	1,5	0,2	5,5	2,3	4,0	1,3	5,2	3,4	1,4	0,2	6,0	2,9	4,5	1,5
5,0	3,6	1,4	0,2	6,5	2,8	4,6	1,5	4,7	3,2	1,6	0,2	5,7	2,6	3,5	1,0
5,4	3,9	1,7	0,4	5,7	2,8	4,5	1,3	4,8	3,1	1,6	0,2	5,5	2,4	3,8	1,1
4,6	3,4	1,4	0,3	6,3	3,3	4,7	1,6	5,4	3,4	1,5	0,4	5,5	2,4	3,7	1,0
5,0	3,4	1,5	0,2	4,9	2,4	3,3	1,0	5,2	4,1	1,5	0,1	5,8	2,7	3,9	1,2
4,4	2,9	1,4	0,2	6,6	2,9	4,6	1,3	5,5	4,2	1,4	0,2	6,0	2,7	5,1	1,6
4,9	3,1	1,5	0,1	5,2	2,7	3,9	1,4	4,9	3,1	1,5	0,2	5,4	3,0	4,5	1,5
5,4	3,7	1,5	0,2	5,0	2,0	3,5	1,0	5,0	3,2	1,2	0,2	6,0	3,4	4,5	1,6
4,8	3,4	1,6	0,2	5,9	3,0	4,2	1,5	5,5	3,5	1,3	0,2	6,7	3,1	4,7	1,5
4,8	3,0	1,4	0,1	6,0	2,2	4,0	1,0	4,9	3,6	1,4	0,1	6,3	2,3	4,4	1,3
4,3	3,0	1,1	0,1	6,1	2,9	4,7	1,4	4,4	3,0	1,3	0,2	5,6	3,0	4,1	1,3
5,8	4,0	1,2	0,2	5,6	2,9	3,6	1,3	5,1	3,4	1,5	0,2	5,5	2,5	4,0	1,3
5,7	4,4	1,5	0,4	6,7	3,1	4,4	1,4	5,0	3,5	1,3	0,3	5,5	2,6	4,4	1,2
5,4	3,9	1,3	0,4	5,6	3,0	4,5	1,5	4,5	2,3	1,3	0,3	6,1	3,0	4,6	1,4
5,1	3,5	1,4	0,3	5,8	2,7	4,1	1,0	4,4	3,2	1,3	0,2	5,8	2,6	4,0	1,2
5,7	3,8	1,7	0,3	6,2	2,2	4,5	1,5	5,0	3,5	1,6	0,6	5,0	2,3	3,3	1,0
5,1	3,8	1,5	0,3	5,6	2,5	3,9	1,1	5,1	3,8	1,9	0,4	5,6	2,7	4,2	1,3
5,4	3,4	1,7	0,2	5,9	3,2	4,8	1,8	4,8	3,0	1,4	0,3	5,7	3,0	4,2	1,2
5,1	3,7	1,5	0,4	6,1	2,8	4,0	1,3	5,1	3,8	1,6	0,2	5,7	2,9	4,2	1,3
4,6	3,6	1,0	0,2	6,3	2,5	4,9	1,5	4,6	3,2	1,4	0,2	6,2	2,9	4,3	1,3
5,1	3,3	1,7	0,5	6,1	2,8	4,7	1,2	5,3	3,7	1,5	0,2	5,1	2,5	3,0	1,1
4,8	3,4	1,9	0,2	6,4	2,9	4,3	1,3	5,0	3,3	1,4	0,2	5,7	2,8	4,1	1,3

Z_1	Z_2	Z_3	Z_4	Z_1	Z_2	Z_3	Z_4	Z_1	Z_2	Z_3	Z_4	Z_1	Z_2	Z_3	Z_4
6,3	3,3	6,0	2,5	5,7	2,5	5,0	2,0	6,2	2,8	4,8	1,8	6,0	3,0	4,8	1,8
5,8	2,7	5,1	1,9	5,8	2,8	5,1	2,4	6,1	3,0	4,9	1,8	6,9	3,1	5,4	2,1
7,1	3,0	5,9	2,1	6,4	3,2	5,3	2,3	6,4	2,8	5,6	2,1	6,7	3,1	5,6	2,4
6,3	2,9	5,6	1,8	6,5	3,0	5,5	1,8	7,2	3,0	5,8	1,6	6,9	3,1	5,1	2,3
6,5	3,0	5,8	2,2	7,7	3,8	6,7	2,2	7,4	2,8	6,1	1,9	5,8	2,7	5,1	1,9
7,6	3,0	6,6	2,1	7,7	2,6	6,9	2,3	7,9	3,8	6,4	2,0	6,8	3,2	5,9	2,3
4,9	2,5	4,5	1,7	6,0	2,2	5,0	1,5	6,4	2,8	5,6	2,2	6,7	3,3	5,7	2,5
7,3	2,9	6,3	1,8	6,9	3,2	5,7	2,3	6,3	2,8	5,1	1,5	6,7	3,0	5,2	2,3
6,7	2,5	5,8	1,8	5,6	2,8	4,9	2,0	6,1	2,6	5,6	1,4	6,3	2,5	5,0	1,9
7,2	3,6	6,1	2,5	7,7	2,8	6,7	2,0	7,7	3,0	6,1	2,3	6,5	3,0	5,2	2,0
6,5	3,2	5,1	2,0	6,3	2,7	4,9	1,8	6,3	3,4	5,6	2,4	6,2	3,4	5,4	2,3
6,4	2,7	5,3	1,9	6,7	3,3	5,7	2,1	6,4	3,1	5,5	1,8	5,9	3,0	5,1	1,8
6,8	3,0	5,5	2,1	7,2	3,2	6,0	1,8								

Tardami, kad buvo stebimi normalieji a. v. $N_4(\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$, $i = 1,\ 2,\ 3$, sudarykite klasifikavimo taisyklę minimizuojančią suminę klasifikavimo klaidą trims vilkdalgio rūšims atskirti. Įvertinkite klasifikavimo klaidų tikimybes.

6.55. (**6.54** pratimo tęsinys). Raskite Fišerio klsifikavimo taisyklę trims vilkdalgio rūšims atskirti ir palyginkite ją su **6.54** pratime gauta taisykle.

Atsakymai ir nurodymai.

6.1 skyrelis

6.1. a) $\varphi_1(x) = 1$, kai $x \le (\mu_1+\mu_2)/2$, $\varphi_1(x) = 0$, kai $x > (\mu_1+\mu_2)/2$; $\varphi_2(x) = 1-\varphi_1(x)$; $\mathbf{P}\{e\} = 1/2 - (1/\pi)\operatorname{arctg}((\mu_2-\mu_1)/2\sigma)$. b) Jeigu $9\sigma^2 \le 4(\mu_2-\mu_1)^2$ tai $\varphi_1(x) = 1$, kai $x_1 < x < x_2$ ir $\varphi_1(x) = 0$, priešingu atveju; $\varphi_2(x) = 1-\varphi_1(x)$; čia $x_1 = (4\mu_1 - \mu_2 - \sqrt{4(\mu_2-\mu_1)^2 - 9\sigma^2})/3$, $x_2 = (4\mu_1 - \mu_2 + \sqrt{4(\mu_2-\mu_1)^2 - 9\sigma^2})/3$. Jeigu $9\sigma^2 > 4(\mu_2-\mu_1)^2$, tai $\varphi_1(x) \equiv 1$. **6.2.** a) Priimamas sprendimas $\eta = 1$, kai $-0,7250 \le x \le 0,0393$; priimamas sprendimas $\eta = 2$, kai $x < -0,7250$ arba $x > 0,0393$. Mažiausiai palankus apriorinių klasių tikimybių rinkinys yra $\omega_1 = 0,093, \omega_2 = 0,907$. b) Priimamas sprendimas $\eta = 1$, kai $-1,590 \le x \le 0,448$; priimamas sprendimas $\eta = 2$, kai $2,552 \le x \le 4,590$; priimamas sprendimas $\eta = 0$, kai $x < -1,590$, $0,448 < x < 2,552$, arba $x > 4,590$. **6.3. Nurodymas.** $\mathbf{P}\{e\} \le \min(\int \omega_1 f_1(x) d\boldsymbol{\mu}, \int \omega_2 f_2(x) d\boldsymbol{\mu}) \le \int(\omega_1 f_1(x) + \omega_2 f_2(x)/2 d\boldsymbol{\mu})$. Lieka pasinaudoti nelygybe tarp aritmetinio ir geometrinio vidurkio. **6.4.** Priimamas sprendimas $\eta = 1$, kai

$$\prod_{i=1}^{k}\left[\frac{p_i(1-q_i)}{(1-p_i)q_i}\right]^{X_i} \ge c\prod_{i=1}^{k}\left[\frac{1-q_i}{1-p_i}\right], \quad c = \frac{(c_{12}-c_{22})\omega_2}{(c_{21}-c_{11})\omega_1}.$$

Priešingu atveju priimamas sprendimas $\eta = 2$. **6.5.** Priimamas sprendimas $\eta = 1$, kai $S_k = X_1 + ... + X_k \ge k/2$; priimamas sprendimas $\eta = 2$, kai $S_k < k/2$. Aproksimuodami normaliuoju skirstiniu gauname $\mathbf{P}\{e\} = 1 - \Phi(\sqrt{k}(1/2-q)/\sqrt{pq})$; $\mathbf{P}\{e\} \to 0$, kai $k \to \infty$; $\mathbf{P}\{e\} \to 1/2$, kai $p \to 1/2$. **6.6.** Priimamas sprendimas $\eta = 1$, kai $x \ge [\ln(2\lambda_2/\lambda_1)]/(\lambda_2-\lambda_1)$; priimamas sprendimas $\eta = 2$ priešingu atveju. **6.7.** Priimamas sprendimas $\eta = 1$, kai $x \ge 0,3365$; priimamas sprendimas $\eta = 2$, kai $x < 0,3365$. Mažiausiai palankus apriorinių klasių tikimybių rinkinys yra $\omega_1 = 0,5655, \omega_2 = 0,4345$. **6.8.** Priimamas sprendimas $\eta = 1$, kai $x \ge [\ln(3\lambda_2/\lambda_1)]/(\lambda_2-\lambda_1)$; priimamas sprendimas $\eta = 2$, kai $x \le [\ln(\lambda_2/\lambda_1)]/(\lambda_2-\lambda_1)$; priimamas sprendimas $\eta = 0$, kai $[\ln(\lambda_2/\lambda_1)]/(\lambda_2-\lambda_1) < x < [\ln(3\lambda_2/\lambda_1)]/(\lambda_2-\lambda_1)$. **6.9.** a) Priimamas sprendimas $\eta = 1$, kai $-1,8512 \le x \le 1,8512$; priimamas sprendimas $\eta = 2$ kai $|x| > 1,8512$; $\mathbf{P}\{e\} = 0,3744$. b) Priimamas sprendimas $\eta = 1$, kai $|x| \le 2,5965$; priimamas sprendimas $\eta = 2$ kai $|x| > 2,5965$; $\mathbf{P}\{e\} = 0,3877$. **6.10.** $\varphi_1^*(x) = 1$, kai $|x| \le h(\omega_1)$ ir $\varphi_1^*(x) = 0$, kai $|x| > h(\omega_1)$; $\varphi_2^*(x) = 1-\varphi_1^*(x)$; $R(\boldsymbol{\varphi}^*, \omega_1) = ((1-\omega_1)/\pi)\operatorname{arctg}(h(\omega_1)) + 3\omega_1(1-\Phi(h(\omega_1)))$; čia funkcija $h(\omega_1)$ apibrėžiama lygybe $\sqrt{\pi/2}(1+h^2(\omega_1))\exp\{-h^2(\omega_1)/2\} = (1-\omega_1)/\omega_1$. **6.11.** $\varphi_1(x) = 1$, kai $|x| \le 0,1543$; $\varphi_2(x) = 1-\varphi_1(x)$. **6.12.** a) $(\varphi_1^{(0)}(x), \varphi_2^{(0)}(x))$; b) Jeigu $0 \le \omega_1 \le 0,273$, tai $\varphi_1(x) \equiv 1$; jeigu $0,273 < \omega_1 \le 0,805$, tai $\varphi_1(0) = 1$, $\varphi_1(1) = \varphi_1(2) = \varphi_1(3) = 0$, $\varphi_2(x) = 1-\varphi_1(x)$; jeigu $0,805 < \omega_1 \le 0,978$, tai $\varphi_1(0) = \varphi_1(1) = 1$, $\varphi_1(2) = \varphi_1(3) = 0$, $\varphi_2(x) = 1-\varphi_1(x)$; jeigu $0,978 < \omega_1 \le 0,998$, tai $\varphi_1(0) = \varphi_1(1) = \varphi_1(2) = 1$, $\varphi_1(3) = 0$, $\varphi_2(x) = 1-\varphi_1(x)$; jeigu $0,998 \le \omega_1 \le 1$, tai $\varphi_1(x) \equiv 1$. **6.13.** a) $(\varphi_1^{(3)}(x), \varphi_2^{(3)}(x))$; b) $\varphi_1(x) \equiv 0$, kai $\omega_1 \le 0,5$; $\varphi_1(x) = 1$, $x = 1,...,k$, $\varphi_2(x) = 1$, $x = k+1, k+2, ...$, kai $1/(+(8/9)^{k-1}) \le \omega_1 \le 1/(1+(8/9)^k)$, $k = 1,2,...$ **6.14.** a) Randame

i	1	2	3	4	5	6	7
$R_1(\boldsymbol{\varphi}^{(i)})$	0	13,533	40,601	67,668	83,712	94,736	100
$R_2(\boldsymbol{\varphi}^{(i)})$	50	31,606	13,216	4,015	0,949	0,183	0

b) $\varphi_1(x) = 1$, kai $x = 0;1$, $\varphi_1(x) = 0,6933$, kai $x = 2$, $\varphi_1(x) = 0$, kai $x = 3,4,...$; $\varphi_2(x) = 1-\varphi_1(x)$. **6.15.** Jeigu $0 \le \omega_1 \le 1/3$, tai $\varphi_1(x) \equiv 0$; jeigu $1/3 < \omega_1 < 1$, tai $\varphi_1(x) \equiv 1$; $R(\boldsymbol{\varphi}) = 10\omega_1$, kai $0 \le \omega_1 \le 1/3$; $R(\boldsymbol{\varphi}) = 5(1-\omega_1)$, kai $1/3 < \omega_1 \le 1$. Neapsimoka, jei $\omega_1 < 0,01$. **6.16.** Tegu $S_n = X_1 + ... + X_n$. Tada $\varphi_1(S_n) \equiv 1$, kai $S_n = m, m \in \{m : m \ge (n\ln(8/3) + \ln c)/\ln 6; m = 0,1,...,n\}$. b) $n = 30$. **6.17.** $3/(3+4\Lambda) \le \omega_1 \le 7/(7+4\Lambda)$, $\Lambda = f_1(\boldsymbol{x})/f_2(\boldsymbol{x})$. **6.18.** Jeigu $0 \le \omega_1 \le 2/16$, tai priimamas sprendimas $\eta = 2$; jeigu $2/16 < \omega_1 \le 7/16$, tai taške $x = 0$ priimamas sprendimas $\eta = 2$, o taške $x = 1$ – sprendimas $\eta = 0$; jeigu $7/16 < \omega_1 \le 9/16$, tai taške $x = 0$ priimamas sprendimas $\eta = 2$, o taške $x = 1$ – sprendimas $\eta = 1$; jeigu $9/16 < \omega_1 \le 7/8$, tai taške $x = 0$ priimamas sprendimas $\eta = 0$, o taške $x = 1$ – sprendimas $\eta = 1$; jeigu $7/8 \le \omega_1 \le 1$, tai visur priimamas sprendimas

$\eta = 1$. Nuostolių funkcija

$$R(\boldsymbol{\varphi}, \omega_1) = \begin{cases} 10\omega_1, \quad \text{kai} \quad 0 \le \omega_1 \le 2/16, \\ 4\omega_1 + 3/4, \quad \text{kai} \quad 2/16 < \omega_1 \le 7/16, \\ 5/2, \quad \text{kai} \quad 7/16 < \omega_1 \le 9/16, \\ 19/4 - 4\omega_1, \quad \text{kai} \quad 9/16 < \omega_1 \le 14/16, \\ 10(1-\omega_1), \quad \text{kai} \quad 14/16 < \omega_1 \le 1. \end{cases}$$

6.19. $n = 10$. **6.20.** $n = 32, R = 46,8366$. **6.21.** $\omega_1(x) = g(x_1)f(x_2)\cdot\ldots\cdot f(x_n)\omega_1/(\sum_i (g(x_i)\omega_i \prod_{j\ne i} f(x_j))$.

6.2 skyrelis

6.26. $\varphi_j(\mathbf{X}) = 1$, jeigu su visais $j' \ne j = 1,...,m$ yra teisingos nelygybės

$$\prod_{i=1}^{k} \left[\frac{p_{ij}(1-p_{ij'})}{(1-p_{ij})p_{ij'}}\right]^{X_i} \ge \prod_{i=1}^{k} \left[\frac{1-p_{ij'}}{1-p_{ij}}\right] \frac{\omega'_j}{\omega_j}.$$

6.27. Priimamas sprendimas $\eta = 1$, kai $0,26 \le |x| \le 1,74$; priimamas sprendimas $\eta = 2$, kai $2,265 \le |x|$; priimamas sprendimas $\eta = 3$, kai $|x| < 0,26$ arba $1,74 < |x| < 2,265$; $R(\boldsymbol{\varphi}) = 4,536$. **6.28.** Priimamas sprendimas $\eta = 1$, kai $0,653 \le |x| \le 1,329$; priimamas sprendimas $\eta = 2$, kai $2,345 \le |x|$; priimamas sprendimas $\eta = 3$, kai $|x| < 0,0464$; priimamas sprendimas $\eta = 0$, kai $0,0464 < |x| < 0,653$ arba $1,329 < |x| < 2,345$ **6.29.** Priimamas sprendimas $\eta = i$, kai $\omega_i(\boldsymbol{x}) > \omega_j(\boldsymbol{x})$, $\omega_i(\boldsymbol{x}) > \omega_l(\boldsymbol{x})$, $i \ne j \ne l = 1;2;3$. **6.30.** Priimamas sprendimas $\eta = i$, kai $\omega_i(\boldsymbol{x}) > \omega_j(\boldsymbol{x})$, $\omega_i(\boldsymbol{x}) > \omega_l(\boldsymbol{x})$, $\omega_i(\boldsymbol{x}) > \omega_j(\boldsymbol{x}) + \omega_l(\boldsymbol{x})$, $i \ne j \ne l = 1;2;3$. **6.31.** Sprendinys sudalija plokštumą $\mathbf{R}^2$ į tris nesikertančias sritis: sprendimas $\eta = 1$ priimamas, kai $X \le 0, \quad 2X + 2\sqrt{3}Y - 3 \le 0$; sprendimas $\eta = 2$ priimamas kai $X > 0, \quad 2X - 2\sqrt{3}Y + 3 \ge 0$; sprendimas $\eta = 3$ priimamas kai $2X + 2\sqrt{3}Y - 3 > 0, \quad 2X - 2\sqrt{3}Y + 3 < 0$. Aposteriorinių klasių tikimybių terminais: $\eta = i$, kai $\omega_i(\boldsymbol{x}) > \omega_j(\boldsymbol{x})$, $\omega_i(\boldsymbol{x}) > \omega_l(\boldsymbol{x})$, $i \ne j \ne l = 1;2;3$. $\mathbf{P}\{e\} = 0,1153$. **Nurodymas.** Dėl simetrijos visos klaidų tikimybės yra vienodos, todėl $\mathbf{P}\{e\} = 6\alpha_{12}\omega_2 = 2\alpha_{12}$; $\alpha_{12} = \mathbf{P}\{X \le 0, \quad 2X + 2\sqrt{3}Y - 3 \le 0|\xi = 2\} = \mathbf{P}\{Z_1 \le -3/2, Z_2 \le 0\}$ čia a.v. $(Z_1, Z_2)^T = (X - 3/2, (X - 3/2 + \sqrt{3}Y)/2)$ turi dvimatį normalųjį skirstinį su nuliniu vidurkių vektoriumi, vienetinėmis dispersijomis ir koreliacijos koeficientu $\rho = 1/2$. **6.32.** Sprendimas $\eta = 1$ priimamas, kai $X \le 0, \quad 2X + 2\sqrt{3}Y - 3 \le 0, \quad \exp\{3X\} + \exp\{3(X + \sqrt{3}Y - 3/2)/2\} \le 3/2$; sprendimas $\eta = 2$ priimamas, kai $X > 0, \quad 2X - 2\sqrt{3}Y + 3 \ge 0, \quad \exp\{-3X\} + \exp\{3(-X + \sqrt{3}Y - 3/2)/2\} \le 3/2$; sprendimas $\eta = 3$ priimamas, kai $2X + 2\sqrt{3}Y - 3 > 0, \quad 2X - 2\sqrt{3}Y + 3 < 0, \quad \exp\{-3(-X + \sqrt{3}Y - 3)/2\} + \exp\{-3(X + \sqrt{3}Y - 3/2)/2\} \le 3/2$: likusiuose taškuose $(X,Y)^T$ priimamas sprendimas $\eta = 0$. Aposteriorinių klasių tikimybių terminais: $\eta = i$, kai $\omega_i(\boldsymbol{x}) > \omega_j(\boldsymbol{x})$, $\omega_i(\boldsymbol{x}) > \omega_l(\boldsymbol{x})$, $3\omega_i > 2(\omega_j + \omega_l)$, $i \ne j \ne l = 1;2;3$, kituose simplekso taškuose priimamas sprendimas $\eta = 0$.

6.3 skyrelis

6.36. $\hat{\mu}_B = \bar{X}n\beta^2/(n\beta^2 + \sigma^2) + \mu_0\sigma^2/(n\beta^2 + \sigma^2)$; $\hat{\mu}_B \xrightarrow{P} \bar{X}$, kai $n \to \infty$, o β fiksuotas; $\hat{\mu}_B \xrightarrow{P} \mu_0$, kai $\beta \to 0$, o n fiksuotas; $X|\mathcal{X} \sim N(\hat{\mu}_B, \sigma^2\beta^2/(n\beta^2 + \sigma^2))$. **6.37.** $\hat{\lambda}_B = (\eta + S_n)/(\gamma + n)$, $S_n = X_1 + ... + X_n$; $(X|\mathcal{X}) \sim B^-(\eta + S_n, p)$. **6.38.** $\hat{p}_B = (\gamma + S_n)/(\gamma + \eta + mn)$, $S_n = X_1 + ... + X_n$; a.d. X sąlyginis skirstinys $\mathcal{X}$ atžvilgiu yra apibendrintas binominis:

$$\mathbf{P}\{X = k|\mathcal{X}\} = C_m^k \frac{\Gamma(\gamma + \eta + mn)}{\Gamma(\gamma + S_n)\Gamma(\eta + mn - S_n)} \frac{\Gamma(\gamma + S_n + k)\Gamma(\eta + mn + m - k - S_n)}{\Gamma(\gamma + \eta + mn + m)},$$

$k = 0,1,...,m$. **6.39.** $\hat{p}_B = (\gamma + mn)/(\gamma + \eta + mn + S_n)$, $S_n = X_1 + ... + X_n$;

$$\mathbf{P}\{X = k|\mathcal{X}\} = \frac{\Gamma(m+k)}{k!\Gamma(m)} \frac{\Gamma(\gamma + \eta + mn + S_n)}{\Gamma(\gamma + mn)\Gamma(\eta + S_n)} \frac{\Gamma(\gamma + mn + m)\Gamma(\eta + S_n + k)}{\Gamma(\gamma + \eta + mn + S_n + m + k)},$$

$k = 0,1,...,m$. **6.40.** $\hat{\lambda}_B = (\eta + n)/(\gamma + S_n)$, $S_n = X_1 + ... + X_n$; $f(x|\mathcal{X}) = (\gamma + \eta + 1)(\gamma + S_n)^{\gamma+\eta}/(\gamma + S_n + x^{\gamma+\eta+1})$. **6.41.** $(\theta|\mathcal{X}) \sim G(\lambda + n(\bar{X} - \mu)^2/2, \eta + n/2)$; a.d. $(X - \mu)\sqrt{2\eta + n}/\sqrt{2\lambda + n(\bar{X} - \mu)^2}$ sąlyginis skirstinys $\mathcal{X}$ atžvilgiu yra Stjudento skirstinys su $2\eta + n$ laisvės laipsnių. 6.42. Parametro $\boldsymbol{p}$ aposteriorinis skirstinys yra Dirichlė skirstinys su parametrų vektoriumi $(\alpha_1 + S_1, ..., \alpha_k + S_k)^T$; $\boldsymbol{S} = (S_1, ..., S_k)^T = \sum_j \mathbf{X}_j$; $(\mathbf{X}|\mathcal{X}) \sim$

$\mathcal{P}_k(1, \boldsymbol{p}^*)$, $\boldsymbol{p}^* = (p_1^*, ..., p_k^*)^T$, $p_i^* = (\alpha_i + S_i)/\sum_i(\alpha_i + S_i)$, $i = 1, ..., k$. **6.43.** A. v. $(\theta_1, ..., \theta_k)^T$ aposteriorinio skirstinio koordinatės yra nepriklausomi a. d. ir $\theta_j \sim Be(S_j + 1, n - S_j + 1)$, $j = 1, ..., k$, čia $\boldsymbol{S} = (S_1, ..., S_k^T) = \sum_j \mathbf{X}_j$; a. v. $\mathbf{X}$ sąlyginis skirstinys atžvilgiu $\mathcal{X}$ yra daugiamatis Bernulio su parametru $(\theta_1^*, ..., \theta_k^*)^T$, $\theta_j^* = (S_j + 1)/(n + 2)$, $j = 1, ..., k$. **6.44.** Tikrinta didumo $n = 100$ imtis, kurioje buvo surasti 6 defektiniai gaminiai. **6.45.** Patikrintos 8 juostos; surasti 8 defektai. **6.46.** 195. **6.47.** a) $Be(124; 887)$; b) $Be(124; 1011)$. **6.48.** $n \geq 96$. **6.49.** $n \geq 62$. **6.50.** $n \geq 297$. **6.51.** Priimamas sprendimas $\eta = 1$, kai $|Y| \leq 1,7185$; priimamas sprendimas $\eta = 2$, kai $|Y| > 1,7185$; $\mathbf{P}\{e\} = 0,0852$. **Nurodymas.** Tegu sąlyginiai tankiai $(Y||X| < 1,6) \sim f_1(y)$ ir $(Y||X| > 1,6) \sim f_2(y)$. Tada $\varphi(y) = 1$, kai $f_1(y)/f_2(y) > \omega_2/\omega_1$; $\omega_1 = 2\Phi(1,6) - 1$, $\omega_2 = 1 - \omega_1$. $(Y|a < X < b) \sim f(y|a,b) = \varphi(y)[\Phi(2(b-y)) - \Phi(2(a-y))]/[\Phi(2b/\sqrt{5}) - \Phi(2a/\sqrt{5})]$. **6.53.** Pažymėkime $a_{ij} = \boldsymbol{\mu}_i^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_j, i, j = 1, 2, 3$. Objektas, kuriam $\mathbf{X} = \boldsymbol{x}$ priskiriamas pirmai klasei, jeigu patenkintos nelygybės

$$g_{12}(\boldsymbol{x}) = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}^{-1} \boldsymbol{x} - a_{11}/2 + a_{22}/2 > 0,$$

$$g_{13}(\boldsymbol{x}) = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_3)^T \boldsymbol{\Sigma}^{-1} \boldsymbol{x} - a_{11}/2 + a_{33}/2 > 0.$$

Taigi, pavyzdžiui, $\alpha_{12} = \mathbf{P}\{g_{12}(\mathbf{X}) > 0, g_{12}(\mathbf{X}) > 0 | \xi = 2\}$. Atsitiktinio vektoriaus $(g_{12}(\mathbf{X}), g_{12}(\mathbf{X}))^T$ sąlyginis skirstinys, kai $\xi = 2$, yra dvimatis normalusis su vidurkių vektoriumi $\boldsymbol{\nu} = (a_{12} - (a_{11} + a_{22})/2, a_{12} - a_{23} - (a_{11} - a_{33})/2)$ ir kovariacijų matrica $\Gamma = [\gamma_{ij}]_{2\times 2}$, $\gamma_{11} = a_{11} + a_{22} - 2a_{12}$, $\gamma_{22} = a_{11} + a_{33} - 2a_{13}$, $\gamma_{12} = a_{11} + a_{23} - a_{12} - a_{13}$. Analogiškai gaunamos kitos klasifikavimo klaidų tikimybės. **6.54.** Randame vidurkių vektorių įverčius

$$\hat{\boldsymbol{\mu}}_1 = \bar{\mathbf{X}} = \begin{pmatrix} 5,006 \\ 3,428 \\ 1,462 \\ 0,246 \end{pmatrix}, \quad \hat{\boldsymbol{\mu}}_2 = \bar{\boldsymbol{Y}} = \begin{pmatrix} 5,936 \\ 2,770 \\ 4,260 \\ 1,326 \end{pmatrix}, \quad \hat{\boldsymbol{\mu}}_3 = \bar{\boldsymbol{Z}} = \begin{pmatrix} 6,588 \\ 2,974 \\ 5,552 \\ 2,026 \end{pmatrix}$$

ir jungtinės imties kovariacinės matricos atvirkštinės įvertį

$$\hat{\boldsymbol{\Sigma}}^{-1} = \begin{pmatrix} 17,45127 & -9,82939 & -11,99752 & 5,82635 \\ -9,82939 & 21,35522 & 4,91129 & -16,28068 \\ -11,99752 & 4,91129 & 32,24613 & -39,98703 \\ 5,82635 & -16,28068 & -39,98703 & 131,49894 \end{pmatrix}.$$

Diskriminantinių funkcijų įverčiai

$$\hat{g}_1(\boldsymbol{x}) = \hat{\boldsymbol{\mu}}_1^T \hat{\boldsymbol{\Sigma}}^{-1} \boldsymbol{x} - \frac{1}{2} \hat{\boldsymbol{\mu}}_1^T \hat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\mu}}_1$$

$$= 37,5588x_1 + 27,1750x_2 - 5,9167x_3 - 52,7558x_4 - 129,7736,$$

$$\hat{g}_2(\boldsymbol{x}) = 32,9796x_1 + 0,1406x_2 + 26,7327x_3 - 6,4824x_4 - 150,7163,$$

$$\hat{g}_3(\boldsymbol{x}) = 30,9303x_1 - 6,9628x_2 + 33,5833x_3 + 34,3741x_4 - 219,5788.$$

Objektas $\boldsymbol{x} = (x_1, x_2, x_3, x_4)^T$ priskiriamas i-ai klasei, jeigu

$$\hat{g}_i(\boldsymbol{x}) > \hat{g}_j(\boldsymbol{x}), \quad \forall\ j \neq i, \quad j = 1,\ 2,\ 3.$$

Randame $a_{ij} = \boldsymbol{\mu}_i^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_j$, $i, j = 1,\ 2,\ 3$ įverčius

$$\hat{a}_{11} = 259,5472, \quad \hat{a}_{12} = 203,0646, \quad \hat{a}_{13} = 188,5232,$$

$$\hat{a}_{22} = 301,4325, \quad \hat{a}_{23} = 352,9599, \quad \hat{a}_{33} = 439,1575.$$

Remdamiesi 6.53 pratime pateiktomis momentų išraiškomis gauname klasifikavimo klaidų įverčius $\hat{\alpha}_{12} = F(-6,2220, -3,3505 \,|\, 0,9899) = 2 \times 10^{-10}$, $\hat{\alpha}_{13} = F(-11,5313, -8,9674 \,|\, 0,9899) = 5 \times 10^{-31}$, $\hat{\alpha}_{21} = F(-6,2220, 14,1647 \,|\, -09017) = 2 \times 10^{-10}$, $\hat{\alpha}_{23} = F(11,5313, -2,9441 \,|\, -0,9017) = 0,0016195$, $\hat{\alpha}_{31} = F(-8,9674, -14,1647 \,|\, 0,9539) = 8 \times 10^{-46}$, $\hat{\alpha}_{32} = F(3,3505, -2,9441 \,|\, 0,9539) = 0,0016195$; čia $F(u_1, u_2|\rho) = \mathbf{P}\{U_1 < u_1, U_2 < u_2 \,|\, \rho\}$, kai a. v. $(U_1, U_2)^T$ turi dvimatį normalųjį skirstinį, kurio koordinatės turi nulinius vidurkius, vienetines dispersijas ir koreliacijos koeficientą ρ. **Nurodymas.** Dvimačio normaliojo skirstinio pasiskirstymo funkcijos skaičiavimas yra numatytas, pavyzdžiui, SAS programų pakete.

7 skyrius

Kanoniniai kintamieji

Spręsdami daugiamatės matematinės statistikos uždavinius susiduriame su sunkumais, kurie susiję su didele vektoriaus dimensija, taip pat su galima sudėtinga jo koordinačių priklausomybe. Todėl nagrinėjamos tokios a. v. transformacijos (dažniausiai tiesinės), kad naujai gautų vektorių koordinatės būtų nekoreliuotos, o vektoriaus dimensijos sumažinimas būtų atliekamas taip, kad nebūtų prarandama daug informacijos.

7.1. Pagrindinės komponentės

Parinkime a. v. tiesinę transformaciją taip, kad naujai gauto vektoriaus koordinatės būtų nekoreliuotos ir išrikiuotos jų dispersijų mažėjimo tvarka. Tada gauto vektoriaus koordinatės vadinamos *pagrindinėmis komponentėmis.*

7.1.1. Pagrindinių komponenčių savybės

Tegu a. v. $\mathbf{X} = (X_1, ..., X_k)^T$ kovariacijų matrica $\boldsymbol{V}(\mathbf{X}) = \boldsymbol{\Sigma} = [\sigma_{ij}]_{k\times k}$. Kadangi pagrindinės komponentės priklauso tik nuo $\boldsymbol{\Sigma}$, tai tarsime, kad $\boldsymbol{\mu} = \mathbf{E}(\mathbf{X}) = \mathbf{0}$. Priešingu atveju galima nagrinėti a. v. $\mathbf{X} - \boldsymbol{\mu}$. Parinkime vektorių $\boldsymbol{L} \in \mathbf{R^k}$. Tada a. d. $Y = \boldsymbol{L}^T\mathbf{X}$ dispersija yra

$$\boldsymbol{V}Y = \mathbf{E}(\boldsymbol{L}^T\mathbf{X}) = \boldsymbol{L}^T\boldsymbol{\Sigma}\boldsymbol{L}, \tag{7.1.1}$$

kurią galima padaryti kiek norint didelę didinant vektoriaus $\boldsymbol{L}$ ilgį. Todėl natūralu nagrinėti tik tokias transformacijas, kai vektorius $\boldsymbol{L}$ yra normuotas, t. y. tenkina sąlygą

$$\boldsymbol{L}^T\boldsymbol{L} = 1. \tag{7.1.2}$$

7.1.1 teorema. *Pagrindinės komponentės $Y_1 = \boldsymbol{L}_1^T\mathbf{X}, ..., Y_k = \boldsymbol{L}_k^T\mathbf{X}$ gaunamos tada, kai $\boldsymbol{L}_1, ..., \boldsymbol{L}_k$ yra tikriniai vektoriai*

$$\boldsymbol{\Sigma}\boldsymbol{L}_i = \lambda_i\boldsymbol{L}_i, \quad i = 1, ..., k, \tag{7.1.3}$$

atitinkantys charakteringosios lygties

$$|\mathbf{\Sigma} - \lambda \boldsymbol{I}| = 0 \tag{7.1.4}$$

šaknis $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0$. *Pagrindinės komponentės nekoreliuotos ir išrikiuotos taip, kad jų dispersijos*

$$\boldsymbol{V}Y_i = \boldsymbol{L}_i^T \mathbf{\Sigma} \boldsymbol{L}_i = \lambda_i, \quad i = 1, ..., k,$$

tenkina sąlygą

$$\boldsymbol{V}Y_1 \geq \cdots \geq \boldsymbol{V}Y_k.$$

Perėjimas nuo a. v. $\mathbf{X} = (X_1, ..., X_k)^T$ *prie a. v.* $\boldsymbol{Y} = (Y_1, ..., Y_k)^T$ *nepakeičia apibendrintosios dispersijos ir koordinačių dispersijų sumos:*

$$|\boldsymbol{V}(\boldsymbol{Y})| = |\boldsymbol{V}(\mathbf{X})| = |\mathbf{\Sigma}|, \quad \sum_{i=1}^{k} \boldsymbol{V}(Y_i) = \sum_{i=1}^{k} \boldsymbol{V}(X_i). \tag{7.1.5}$$

Įrodymas. Iš algebros kurso žinome (1 priedas (8.2.23)), kad kiekvienai simetrinei neneigiamai apibrėžtai matricai $\mathbf{\Sigma}$ galioja jos spektrinis išdėstymas

$$\mathbf{\Sigma} = \lambda_1 \boldsymbol{L}_1 \boldsymbol{L}_1^T + ... + \lambda_k \boldsymbol{L}_k \boldsymbol{L}_k^T, \quad \boldsymbol{I} = \boldsymbol{L}_1 \boldsymbol{L}_1^T + ... + \boldsymbol{L}_k \boldsymbol{L}_k^T,$$

$$\boldsymbol{L}_i^T \mathbf{\Sigma} \boldsymbol{L}_i = \lambda_i, \quad \boldsymbol{L}_i^T \mathbf{\Sigma} \boldsymbol{L}_j = 0, \quad i \neq j, \quad i, j = 1, ..., k; \tag{7.1.6}$$

čia $\lambda_1, ..., \lambda_k$ yra charakteringosios lygties (7.1.4) šaknys, o $\boldsymbol{L}_1, ..., \boldsymbol{L}_k$ šias šaknis atitinkantys tikriniai vektoriai, t. y. vektoriai randami iš (7.1.3) sąlygų.

Ieškokime vektoriaus $\boldsymbol{L}$, kuris tenkintų sąlygą (7.1.2) ir maksimizuotų (7.1.1). Remiantis Lagranžo neapibrėžtinių daugiklių metodu, reikia maksimizuoti

$$\varphi = \varphi(\boldsymbol{L}) = \boldsymbol{L}^T \mathbf{\Sigma} \boldsymbol{L} - \lambda \boldsymbol{L}^T \boldsymbol{L}$$

Imdami išvestinę

$$\frac{\partial \varphi}{\partial \boldsymbol{L}} = 2\mathbf{\Sigma}\boldsymbol{L} - 2\lambda \boldsymbol{L}$$

ir prilyginę ją 0, gauname lygtį

$$(\mathbf{\Sigma} - \lambda \boldsymbol{I})\boldsymbol{L} = \mathbf{0}.$$

Kadangi $\boldsymbol{L}^T \boldsymbol{L} = 1$, tai matrica $\mathbf{\Sigma} - \lambda \boldsymbol{I}$ turi būti išsigimusi, t. y. λ turi tenkinti (7.1.4) lygtį. Taigi, vektorius $\boldsymbol{L}$ yra tikrinis vektorius atitinkantis charakteringosios lygties (7.1.4) šaknį λ. Norint, kad a. d. $Y = \boldsymbol{L}^T \mathbf{X}$ dispersija būtų maksimali, iš (7.1.6) matome, kad λ turėtų būti didžiausioji lygties (7.1.4) šaknis. Taigi pirmoji pagrindinė komponentė yra $Y_1 = \boldsymbol{L}_1 \mathbf{X}$, kai $\boldsymbol{L}_1$ yra tikrinis vektorius, atitinkantis didžiausiąją lygties (7.1.4) šaknį λ_1. Analogiškai įsitikiname, kad i-oji pagrindinė komponentė yra $Y_i = \boldsymbol{L}_i \mathbf{X}$, kai $\boldsymbol{L}_i$ yra tikrinis vektorius, atitinkantis i-ąją pagal didumą lygties (7.1.4) šaknį λ_i. Jeigu yra sutampančių šaknų, tai jas atitinkančios pagrindinės komponentės turės vienodas dispersijas.

Pagrindinių komponenčių vektorius $\boldsymbol{Y} = (Y_1, ..., Y_k)^T$ gaunamas atliekant ortonormuotą transformaciją $\boldsymbol{Y} = \boldsymbol{L}^T\mathbf{X}$; čia $\boldsymbol{L}$ – matrica, kurios stulpelius sudaro tikriniai vektoriai $\boldsymbol{L}_1, ..., \boldsymbol{L}_k$. Remiantis (7.1.6) vektoriaus $\boldsymbol{Y}$ kovariacinė matrica yra diagonali su diagonaliaisiais elementais $\lambda_1, ..., \lambda_k$. Gauname

$$|\boldsymbol{V}(\boldsymbol{Y})| = |\boldsymbol{L}^T\boldsymbol{\Sigma}\boldsymbol{L}| = |\boldsymbol{L}^T||\boldsymbol{\Sigma}||\boldsymbol{L}| = |\boldsymbol{\Sigma}||\boldsymbol{L}^T\boldsymbol{L}| = |\boldsymbol{\Sigma}|,$$

$$\sum_{i=1}^{k}\boldsymbol{V}(Y_i) = Tr(\boldsymbol{L}^T\boldsymbol{\Sigma}\boldsymbol{L}) = Tr(\boldsymbol{\Sigma}\boldsymbol{L}^T\boldsymbol{L}) = Tr(\boldsymbol{\Sigma}\boldsymbol{I}) = \sum_{i=1}^{k}\boldsymbol{V}(X_i). \quad \blacktriangle$$

Nagrinėkime normalųjį a. v. $\mathbf{X} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, kai kovariacine matrica $\boldsymbol{\Sigma}$ teigiamai apibrėžta. Tankio funkcija

$$\varphi(\boldsymbol{x}|\boldsymbol{\Sigma}) = (2\pi)^{-k/2}|\boldsymbol{\Sigma}|^{-1/2}\exp\{-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{x}\}$$

viršija parinktą konstantą, kai $\boldsymbol{x}$ patenka į elipsoido

$$\boldsymbol{x}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{x} = C \tag{7.1.7}$$

vidų. Pagrindinė šio elipsoido ašis yra atkarpa $(-y, +y)$, kai y yra toks elipsoido taškas, kad atstumas $\boldsymbol{x}^T\boldsymbol{x}$ įgyja maksimalią reikšmę. Remdamiesi Lagranžo neapibrėžtiniais daugikliais nagrinėkime funkciją

$$\phi(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{x} - \lambda\boldsymbol{x}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{x}.$$

Diferencijuodami $\boldsymbol{x}$ atžvilgiu gauname lygtį

$$2\boldsymbol{x} - 2\lambda\boldsymbol{\Sigma}^{-1}\boldsymbol{x} = 0, \quad \text{arba} \quad \boldsymbol{\Sigma}\boldsymbol{x} - \lambda\boldsymbol{x} = \mathbf{0}.$$

Ši lygtis sutampa su (7.1.3). Taigi pagrindinių komponenčių vektorių kryptys sutampa su elipsoido (7.1.7) pagrindinių ašių kryptimis. Transformacija $\boldsymbol{Y} = \boldsymbol{L}^T\mathbf{X}$ yra toks koordinačių ašių pasukimas, kad naujos koordinačių sistemos ašys sutampa su elipsoido ašimis. Naujoje koordinačių sistemoje elipsoido (7.1.7) lygtis yra

$$\boldsymbol{y}^T\boldsymbol{\Lambda}^{-1}\boldsymbol{y} = \sum_{i=1}^{k}\frac{y_i^2}{\lambda_i} = C,$$

o i-osios pagrindinės ašies ilgis lygus $2\sqrt{\lambda_i C}$.

Gali atsitikti taip, kad keleto pirmųjų pagrindinių komponenčių $Y_1, ..., Y_r$, $r < k$, suma $\lambda_1 + ... + \lambda_r$ sudaro beveik visą bendros dispersijos $\lambda_1 + ... + \lambda_k$ dalį, o likusioms komponentėms $Y_{r+1}, ..., Y_k$ tenka tik nežymi suminės dispersijos dalis. Kai skirstinys normalusis, tai reikštų, kad elipsoidas (7.1.7) yra stipriai ištemptas keleto pagrindinių ašių kryptimi. Tada dažnai naudojamas toks vektoriaus $\mathbf{X}$ dimensijos sumažinimo metodas: vietoje vektoriaus $\mathbf{X}$ naudojame r-matį vektorių $(Y_1, ..., Y_r)^T$, sudarytą iš pirmųjų pagrindinių komponenčių, o likusias komponentes $Y_{r+1}, ..., Y_k$ tiesiog atmetame. Kyla klausimas, ar tai geriausias būdas sumažinti vektoriaus dimensiją nuo k iki r.

7.1.2 teorema. *Tegu* $Y_1 = \boldsymbol{B}_1^T \mathbf{X}, ..., Y_r = \boldsymbol{B}_r^T \mathbf{X}$ *yra* r *tiesinių a. v.* $\mathbf{X}$ *funkcijų, tenkinančių sąlygas*

$$\boldsymbol{V}(Y_i) = \boldsymbol{B}_i^T \boldsymbol{\Sigma} \boldsymbol{B}_i = 1, \quad \mathbf{Cov}\,(Y_i, Y_j) = \boldsymbol{B}_i^T \boldsymbol{\Sigma} \boldsymbol{B}_j = 0, \quad i \neq j, \quad i, j = 1, ..., r. \tag{7.1.8}$$

Pažymėkime σ_i^2 *liekamąją dispersiją tiesiškai prognozuodami koordinatę* X_i *remdamiesi a. d.* $Y_1, ..., Y_r$. *Tada* $\sum_{i=1}^k \sigma_i^2$ *įgyja minimalią reikšmę, kai* $Y_1, ..., Y_r$ *sutampa su pirmosiomis* r *pagrindinėmis komponentėmis.*

Įrodymas. Liekamoji dispersija

$$\sigma_i^2 = \min_{\beta_1, ..., \beta_r} \mathbf{E}(X_i - \beta_1 Y_1 - ... - \beta_r Y_r)^2 = \sigma_{ii} - [\mathbf{Cov}\,(X_i, \boldsymbol{B}_1^T \mathbf{X})]^2 - ...$$

$$-[\mathbf{Cov}\,(X_i, \boldsymbol{B}_r^T \mathbf{X})]^2 = \sigma_{ii} - \boldsymbol{B}_1^T \boldsymbol{\Sigma}_i \boldsymbol{\Sigma}_i^T \boldsymbol{B}_1 - ... - \boldsymbol{B}_r^T \boldsymbol{\Sigma}_i \boldsymbol{\Sigma}_i^T \boldsymbol{B}_r,$$

čia $\boldsymbol{\Sigma}_i$ yra matricos $\boldsymbol{\Sigma}$ i-asis stulpelis. Tada

$$\sum_{i=1}^k \sigma_i^2 = \sum_{i=1}^k \sigma_{ii} - \boldsymbol{B}_1^T \sum_{i=1}^k (\boldsymbol{\Sigma}_i \boldsymbol{\Sigma}_i^T) \boldsymbol{B}_1 - ... - \boldsymbol{B}_r^T \sum_{i=1}^k (\boldsymbol{\Sigma}_i \boldsymbol{\Sigma}_i^T) \boldsymbol{B}_r$$

$$= Tr(\boldsymbol{\Sigma}) - \boldsymbol{B}_1^T \boldsymbol{\Sigma\Sigma} \boldsymbol{B}_1 - ... - \boldsymbol{B}_r^T \boldsymbol{\Sigma\Sigma} \boldsymbol{B}_r.$$

Norint minimizuoti $\sum_{i=1}^k \sigma_i^2$, reikia maksimizuoti sumą

$$\boldsymbol{B}_1^T \boldsymbol{\Sigma\Sigma} \boldsymbol{B}_1 + ... + \boldsymbol{B}_r^T \boldsymbol{\Sigma\Sigma} \boldsymbol{B}_r,$$

kai $\boldsymbol{B}_1, ..., \boldsymbol{B}_r$ tenkina sąlygas (7.1.8). Analogiškai 7.1.1 teoremai įsitikiname, kad vektorius $\boldsymbol{B}_i$ yra tikrinis vektorius, atitinkantis i-ąją pagal didumą charakteringosios lygties

$$|\boldsymbol{\Sigma\Sigma} - \lambda \boldsymbol{\Sigma}| = 0 \tag{7.1.9}$$

šaknį λ_i. Tačiau lygties (7.1.9) tikrinės reikšmės ir tikriniai vektoriai sutampa su lygties (7.1.4) tikrinėmis reikšmėmis ir tikriniais vektoriais. ▲

Vadinasi, a. v. $\mathbf{X}$ dimensijos sumažinimas nuo k iki r naudojant pirmąsias pagrindines komponentes leidžia prarasti minimalų informacijos kiekį tuo požiūriu, kad pagal $Y_1, ..., Y_r$ galima tiksliausia atkurti vektorių $\mathbf{X}$, bent jau naudojant tiesines prognozes.

7.1.1 pastaba. Interpretuojant pagrindinę komponentę Y_i naudinga žinoti, kokį indėlį ją sudarant suvaidino pradinio vektoriaus koordinatės $X_1, ..., X_k$. Tuo tikslu galima palyginti komponentės Y_i koreliacijos koeficientus su kiekvienu a. d. $X_1, ..., X_k$. Turime

$$\rho(X_j, \boldsymbol{L}_i^T \mathbf{X}) = \frac{\mathbf{Cov}\,(X_j, \boldsymbol{L}_i^T \mathbf{X})}{\sqrt{\sigma_{jj}}\sqrt{\lambda_i}} = \frac{\lambda_i \boldsymbol{L}_i^T \boldsymbol{I}_j}{\sqrt{\sigma_{jj}}\sqrt{\lambda_i}} = \frac{L_{ij}}{\sqrt{\sigma_{jj}}} \sqrt{\lambda_i}, \tag{7.1.10}$$

čia $\boldsymbol{I}_j$ yra vektorius, kurio j-oji koordinatė lygi 1, o visos kitos – 0; L_{ij} yra vektoriaus $\boldsymbol{L}_i = (L_{i1}, ..., L_{ik})^T$ j-oji koordinatė. Taigi lygindami a. d. $X_1, ..., X_k$ indėlį į i-ąją pagrindinę komponentę palyginame santykius $L_{i1}/\sqrt{\sigma_{11}}, ..., L_{ik}/\sqrt{\sigma_{kk}}$.

7.1.2 pastaba. Apibrėžiant pagrindines komponentes vietoje kovariacijų matricos galima nagrinėti koreliacijos koeficientų matricą, t. y. vietoje a. v. $\mathbf{X} = (X_1, ..., X_k)^T$ nagrinėti a. v. $\boldsymbol{Z} = (X_1/\sqrt{\sigma_{11}}, ..., X_k/\sqrt{\sigma_{kk}})$. Jeigu a. v. $\mathbf{X}$ koordinatės matuojamos skirtingais vienetais, neturinčiais nieko bendro, tai pagrindines komponentes dažnai lengviau interpretuoti nagrinėjant a. v. $\boldsymbol{Z}$ su bedimensinėmis koordinatėmis. Tačiau reikia pažymėti, kad pagrindinės komponentės, gautos iš koreliacijos koeficientų matricos, nesutampa su tomis, kurios gautos naudojant kovariacijų matricą.

7.1.1 pavyzdys. Tarkime, trimačio a. v. $\mathbf{X} = (X_1, X_2, X_3)^T$ kovariacijų matrica yra

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix}.$$

Rasime šią matricą atitinkančias pagrindines komponentes. Charakteringoji lygtis $|\boldsymbol{\Sigma} - \lambda \boldsymbol{I}| = 0$ yra ekvivalenti lygčiai $(1-\lambda)^3 - 3(1-\lambda)\rho^2 + 2\rho^3 = 0$. Didžiausioji šios lygties šaknis $\lambda_1 = 1 + 2\rho$, o kitos dvi šaknys vienodos $\lambda_2 = \lambda_3 = 1 - \rho$. Suradę tikrinius vektorius $\boldsymbol{L}_1, \boldsymbol{L}_2, \boldsymbol{L}_2$, randame pagrindines komponentes $Y_1 = \boldsymbol{L}_1^T \mathbf{X} = (X_1 + X_2 + X_3)/\sqrt{3}$, $Y_2 = \boldsymbol{L}_2^T \mathbf{X} = (X_1 - X_2)/\sqrt{2}$, $Y_3 = \boldsymbol{L}^T \mathbf{X} = (X_1 + X_2 - 2X_3)/\sqrt{6}$. Atsitiktinio vektoriaus $\boldsymbol{Y} = (Y_1, Y_2, Y_3)^T$ kovariacijų matrica yra diagonali

$$\boldsymbol{V}(\boldsymbol{Y}) = \begin{pmatrix} 1+2\rho & 0 & 0 \\ 0 & 1-\rho & 0 \\ 0 & 0 & 1-\rho \end{pmatrix}.$$

Pirmosios komponentės Y_1 dispersija sudaro $(1+2\rho)/3$ dalį bendros dispersijų sumos. Jeigu, pavyzdžiui, $\rho = 1/2$, tai šis santykis lygus $2/3$.

7.1.2 pavyzdys. Tegu a. v. $\mathbf{X} = (X_1, X_2, X_3)^T$ kovariacijų matrica yra

$$\boldsymbol{\Sigma} = \begin{pmatrix} 4 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Rasime pagrindines komponentes. Charakteringoji lygtis $|\boldsymbol{\Sigma} - \lambda \boldsymbol{I}| = 0$ yra ekvivalenti lygčiai $((4-\lambda)^2 - 4)(1-\lambda) = 0$. Šios lygties šaknys yra $\lambda_1 = 6$, $\lambda_2 = 2$, $\lambda_3 = 1$. Suradę tikrinius vektorius $\boldsymbol{L}_1, \boldsymbol{L}_2, \boldsymbol{L}_2$, randame pagrindines komponentes $Y_1 = \boldsymbol{L}_1^T \mathbf{X} = (X_1 + X_2)/\sqrt{2}$, $Y_2 = \boldsymbol{L}_2^T \mathbf{X} = (X_1 - X_2)/\sqrt{2}$, $Y_3 = \boldsymbol{L}_3^T \mathbf{X} = X_3$. Atsitiktinio vektoriaus $\boldsymbol{Y} = (Y_1, Y_2, Y_3)^T$ kovariacijų matrica yra diagonali

$$\boldsymbol{V}(\boldsymbol{Y}) = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Pirmosios komponentės Y_1 dispersija sudaro $2/3$ dalį, o pirmųjų dviejų komponenčių dispersijų suma – $8/9$ dalį bendros dispersijų sumos.

Jeigu šiame pavyzdyje būtume perėję prie normuoto vektoriaus $\boldsymbol{Z} = (Z_1, Z_2, Z_3)^T = (X_1/2, X_2/2, X_3)^T$, t. y. vietoje kovariacijų matricos naudotume koreliacijos koeficientų matricą, tai būtume gavę kitokias pagrindinių komponenčių išraiškas. Tikriniai vektoriai $\lambda_1 = 3/2, \lambda_2 = 1, \lambda_3 = 1/2$ ir juos atitiktų pagrindinės komponentės $Y_1 = (Z_1 + Z_2)/\sqrt{2} = (X_1 + X_2)/\sqrt{8}$, $Y_2 = Z_3 = X_3$, $Y_3 = (Z_1 - Z_2)/\sqrt{2} = (X_1 - X_2)/\sqrt{8}$.

Pirmosios komponentės Y_1 dispersija sudaro $1/2$ dalį, o pirmųjų dviejų komponenčių dispersijų suma – $5/6$ dalį bendros dispersijų sumos.

7.1.2. Pagrindinių komponenčių ir jų dispersijų DT įvertiniai

Tegu $\mathbf{X}_1, ..., \mathbf{X}_n$ yra paprastoji imtis a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Ieškosime parametrų $\lambda_1, ..., \lambda_k$ ir vektorių $\boldsymbol{L}_1, ..., \boldsymbol{L}_k$ didžiausiojo tikėtinumo įvertinių.

7.1.3 teorema. *Tarkime, kad kovariacijų matrica $\mathbf{\Sigma}$ teigiamai apibrėžta $|\mathbf{\Sigma}| > 0$, $n > k$, o charakteringoji lygtis $|\hat{\mathbf{\Sigma}} - \lambda \boldsymbol{I}| = 0$ turi k skirtingų šaknų $\lambda_1 > ... > \lambda_k > 0$. Tada parametrų $\lambda_1, ..., \lambda_k$ didžiausiojo tikėtinumo įvertiniai yra lygties*

$$|\hat{\mathbf{\Sigma}} - \lambda \boldsymbol{I}| = 0, \quad \hat{\mathbf{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^T, \quad \bar{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i, \qquad (7.1.11)$$

šaknys $\hat{\lambda}_1 > \cdots > \hat{\lambda}_k > 0$, o vektorių $\boldsymbol{L}_1, ..., \boldsymbol{L}_k$ įvertiniai $\hat{\boldsymbol{L}}_1, ..., \hat{\boldsymbol{L}}_k$ tenkina sąlygas

$$\hat{\mathbf{\Sigma}}\hat{\boldsymbol{L}}_i = \hat{\lambda}_i \hat{\boldsymbol{L}}_i, \quad i = 1, ..., k, \quad \hat{\boldsymbol{L}}_i^T \hat{\boldsymbol{L}}_i = 1. \qquad (7.1.12)$$

Įrodymas. Jeigu charakteringosios lygties (7.1.4) šaknys visos skirtingos ir teigiamos $\lambda_1 > \lambda_2 > ... > \lambda_k > 0$, tai matricos $\mathbf{\Sigma}$ spektrinis išdėstymas (7.1.6) apibrėžiamas vienareikšmiškai, išskyrus tai, kad galima pakeisti vektorių $\boldsymbol{L}_i$ ženklus priešingais. Šį neapibrėžtumą galima panaikinti, pavyzdžiui, tariant, kad vektoriaus $\boldsymbol{L}_i$ pirmoji nenulinė koordinatė yra teigiama. Tada egzistuoja abipus vienareikšmė $\boldsymbol{\mu}, \lambda_1, ..., \lambda_k, \boldsymbol{L}_1, ..., \boldsymbol{L}_k$ ir $\boldsymbol{\mu}, \mathbf{\Sigma}$ priklausomybė. Todėl remiantis DT įvertinių invariantiškumo principu parametrų $\lambda_1, ..., \lambda_k$, $\boldsymbol{L}_1, ..., \boldsymbol{L}_k$ DT įvertiniai gaunami imant (7.1.3), (7.1.4) vietoje kovariacinės matricos $\mathbf{\Sigma}$ jos įvertinį $\hat{\mathbf{\Sigma}}$, t. y. iš (7.1.11), (7.1.12) sąlygų.

Jeigu $|\mathbf{\Sigma}| > 0$ ir $n > k$, tai su tikimybe 1 determinantas $|\hat{\mathbf{\Sigma}}| > 0$. Todėl su tikimybe 1 įvertiniai $\hat{\lambda}_1 > ... > \hat{\lambda}_k > 0$ ir įvertiniai $\hat{\boldsymbol{L}}_1, ..., \hat{\boldsymbol{L}}_k$ apibrėžiami vienareikšmiškai (neskaitant ženklo). Kadangi remiantis (7.1.6)

$$\hat{\mathbf{\Sigma}} = \hat{\lambda}_1 \hat{\boldsymbol{L}}_1^T \hat{\boldsymbol{L}}_1 + ... + \hat{\lambda}_k \hat{\boldsymbol{L}}_k^T \hat{\boldsymbol{L}}_k,$$

tai $\hat{\boldsymbol{L}}_i$ ženklo keitimas nekeičia $\hat{\mathbf{\Sigma}}$. Tikėtinumo funkcijos maksimumas priklauso tik nuo $\hat{\mathbf{\Sigma}}$, todėl jis nepakinta imant bet kurį (7.1.11), (7.1.12) sprendinį. ▲

7.1.3 pavyzdys. (**6.3.3** pavyzdžio tęsinys). Įvertinsime pagrindines komponentes naudodami $n = 50$ matavimų $\boldsymbol{Y}_i = (Y_{1i}, Y_{2i}, Y_{3i}, Y_{4i})^T$, $i = 1, ..., 50$, vilkdalgio rūšiai *Iris versicolor*.

Randame koreliacinės matricos $\mathbf{\Sigma}$ įvertį

$$\hat{\mathbf{\Sigma}} = \frac{1}{49}\sum_{i=1}^{50}(\boldsymbol{Y}_i - \bar{\boldsymbol{Y}})(\boldsymbol{Y}_i - \bar{\boldsymbol{Y}})^T = \frac{1}{49}\begin{pmatrix} 13,0552 & 4,1740 & 8,9620 & 2,7332 \\ 4,1740 & 4,8250 & 4,0500 & 2,0190 \\ 8,9620 & 4,0500 & 10,8200 & 3,5820 \\ 2,7332 & 2,0190 & 3,5820 & 1,9162 \end{pmatrix}$$

Naudodami SAS programų paketą, randame tikrinių reikšmių įverčius

$$\hat{\lambda}_1 = 0,4879, \quad \hat{\lambda}_2 = 0,0724, \quad \hat{\lambda}_3 = 0,0548, \quad \hat{\lambda}_4 = 0,0098,$$

ir tikrinių vektorių $\boldsymbol{L}_1, \boldsymbol{L}_2, \boldsymbol{L}_3, \boldsymbol{L}_4$ įverčius

$\hat{\boldsymbol{L}}_1$	$\hat{\boldsymbol{L}}_2$	$\hat{\boldsymbol{L}}_3$	$\hat{\boldsymbol{L}}_4$
0,6867	0,6690	0,2651	0,1023
0,3053	-0,5675	0,7296	-0,2289
0,6237	-0,3433	-0,6272	-0,3160
0,2150	-0,3353	-0,0637	0,9150

Jau pirmoji pagrindinė komponentė

$$\hat{Z} = \hat{\boldsymbol{L}}_1^T \boldsymbol{Y} = 0,6867Y_1 + 0,3053Y_2 + 0,6237Y_3 + 0,2150Y_4$$

paaiškina didesnę dalį sklaidos: jos dispersija sudaro apytiksliai 78 procentus vektoriaus koordinačių dispersijų sumos. Matyt, daugelyje praktinių situacijų vietoje keturmačio vektoriaus $\boldsymbol{Y}$ pakaks naudoti vienmatį a. d. Z.

7.1.3. Hipotezės dėl tikrinių reikšmių

Daugelis svarbių daugiamatės matematinės statistikos uždavinių gali būti suformuluota kovariacijų matricos tikrinių reikšmių terminais. Pavyzdžiui, tiesinės hipotezės daugiamačiuose tiesiniuose modeliuose gali būti suformuluotos kaip hipotezės dėl minimalios tikrinės reikšmės ar hipotezės dėl tam tikrų simetrinių tikrinių reikšmių funkcijų (žr. 6.3.4 skyrelį). Hipotezė, kad kovariacinė matrica proporcinga žinomai matricai (žr. 6.5.2 skyrelį), galima suformuluoti taip: visos kovariacinės matricos tikrinės reikšmės yra vienodos. Kalbant apie pagrindines komponentes, pirmiausia, matyt, tikslinga formuluoti tokius klausimus.

1. Ar atliekama a. v. $\mathbf{X}$ transformacija, kurios metu gauname pagrindinių komponenčių vektorių $\boldsymbol{Y}$, kurio koordinatės nepriklausomi a. d., yra vertinga? Suprantama, kad tokia transformacija neturi prasmės, jeigu a. v. $\mathbf{X}$ koordinatės yra nekoreliuotos (kai skirstinys normalusis, nepriklausomos), nes tada pirmosios komponentės sutaps su $\mathbf{X}$ koordinatėmis, turinčiomis didžiausias dispersijas. Vektorių nepriklausomumo hipotezių tikrinimo kriterijai buvo nagrinėjami 6.4 skyrelyje.

2. Tarkime, kad pirmųjų r pagrindinių komponenčių dispersijų suma sudaro didesnę dalį visų dispersijų sumos. Kyla klausimas, ar tarp likusių pagrindinių komponenčių nėra išsiskiriančių, kurias galbūt reikėtų pridėti prie r parinktųjų. Norint atsakyti į šį klausimą galima patikrinti sferiškumo hipotezę $H : \mathbf{\Lambda}_r = \lambda \boldsymbol{I}$, čia $\mathbf{\Lambda}_r$ yra a. v. $(Y_{r+1}, ..., Y_k)^T$ kovariacijų matrica. Tokių hipotezių tikrinimo uždavinį aptarėme 6.5 skyrelyje.

7.2. Kanoninės koreliacijos

Dauginis koreliacijos koeficientas (4.4 skyrelis) yra tiesinės a. d. X_1 ir a. v. $(X_2, ..., X_k)^T$ priklausomybės matas. Jis gaunamas kaip įprastinis a. d. X_1 ir geriausios jo tiesinės prognozės $\hat{X}_1$, sukonstruotos remiantis a. v. $(X_2, ..., X_k)^T$, koreliacijos koeficientas. Apibendrinsime šią sąvoką, kad ji apibūdintų dviejų atsitiktinių vektorių sąryšį.

7.2.1. Kanoninių koreliacijų apibrėžimas ir jų savybės

Suskirstykime a. v. $\mathbf{X} = (X_1, ..., X_k)^T$ į dvi dalis $\mathbf{X}_1 = (X_1, ..., X_r)^T$ ir $\mathbf{X}_2 = (X_{r+1}, ..., X_k)^T$. Kovariacijų matrica $\mathbf{\Sigma}$ bus sudalyta į blokus

$$\boldsymbol{V}(\mathbf{X}) = \mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix}, \tag{7.2.1}$$

čia $\mathbf{\Sigma}_{11} = \boldsymbol{V}(\mathbf{X}_1)$, $\mathbf{\Sigma}_{22} = \boldsymbol{V}(\mathbf{X}_2)$, $\mathbf{\Sigma}_{12} = \mathbf{Cov}\,(\mathbf{X}_1, \mathbf{X}_2)$. Tarsime, kad $|\mathbf{\Sigma}_{11}| > 0, |\mathbf{\Sigma}_{22}| > 0$.

Ieškosime tokių normuotų tiesinių formų $Y_1 = \boldsymbol{L}^T\mathbf{X}_1$ ir $Y_2 = \boldsymbol{M}^T\mathbf{X}_2$, $\boldsymbol{L} \in \mathbf{R}^r$, $\boldsymbol{M} \in \mathbf{R}^s$, $s = k - r$, kad jų koreliacijos koeficientas būtų kuo didesnis. Tiksliau nagrinėsime uždavinį

$$\mathbf{Cov}\,(Y_1, Y_2) = \boldsymbol{L}^T\boldsymbol{\Sigma}_{12}\boldsymbol{M} \to \max_{\boldsymbol{L},\boldsymbol{M}}, \tag{7.2.2}$$

kai išpildytos sąlygos

$$\boldsymbol{V}(Y_1) = \boldsymbol{L}^T\boldsymbol{\Sigma}_{11}\boldsymbol{L} = 1, \quad \boldsymbol{V}(Y_2) = \boldsymbol{M}^T\boldsymbol{\Sigma}_{22}\boldsymbol{M} = 1.$$

Surastųjų tiesinių formų $\boldsymbol{L}^T\mathbf{X}_1$ ir $\boldsymbol{M}^T\mathbf{X}_2$ koreliacijos koeficientai vadinami *kanoninėmis koreliacijomis*, o a. d. $\boldsymbol{L}^T\mathbf{X}_1$ ir $\boldsymbol{M}^T\mathbf{X}_2$ – *kanoniniais kintamaisiais.*

7.2.1 teorema. *Kanoniniai kintamieji $\boldsymbol{L}_1^T\mathbf{X}_1, ..., \boldsymbol{L}_r^T\mathbf{X}_1$ gaunami, kai $\boldsymbol{L}_1, ..., \boldsymbol{L}_r$ yra tikriniai vektoriai, atitinkantys charakteringosios lygties*

$$|\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} - \rho^2\boldsymbol{\Sigma}_{11}| = 0 \tag{7.2.3}$$

šaknis $\rho_1^2 \geq \rho_2^2 \geq ... \geq \rho_r^2 \geq 0$. Kanoniniai kintamieji $\boldsymbol{M}_1^T\mathbf{X}_2, ..., \boldsymbol{M}_s^T\mathbf{X}_2$, $s = k-r$, gaunami, kai vektoriai $\boldsymbol{M}_1, ..., \boldsymbol{M}_s$ yra tikriniai vektoriai, atitinkantys charakteringosios lygties

$$|\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} - \rho^2\boldsymbol{\Sigma}_{22}| = 0 \tag{7.2.4}$$

šaknis $\rho_1^2 \geq \rho_2^T \geq ... \geq \rho_s^2 \geq 0$. Nenulinės charakteringųjų lygčių (7.2.3) ir (7.2.4) šaknys sutampa, o jų skaičius $m = Rang(\boldsymbol{\Sigma}_{12})$.

Pažymėkime $\boldsymbol{L}$ ir $\boldsymbol{M}$ matricas, kurių stulpeliai yra atitinkamai vektoriai $\boldsymbol{L}_1, ..., \boldsymbol{L}_r$ ir $\boldsymbol{M}_1, ..., \boldsymbol{M}_s$. Tada kanoninių kintamųjų vektoriaus $\boldsymbol{Y} = (Y_1, ..., Y_k)^T$ $= ((\boldsymbol{L}^T\mathbf{X}_1)^T, (\boldsymbol{M}^T\mathbf{X}_2)^T)^T$ kovariacijų matrica yra tokio pavidalo

$$\boldsymbol{V}(\boldsymbol{Y}) = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{R} \\ \boldsymbol{R}^T & \boldsymbol{I}_s \end{pmatrix}, \tag{7.2.5}$$

čia $\boldsymbol{I}_r$ ir $\boldsymbol{I}_s$ yra atitinkamai $r \times r$ ir $s \times s$ vienetinės matricos, o $\boldsymbol{R}$ – diagonalioji matrica, kurios pirmieji diagonalieji elementai yra $\rho_1 \geq \rho_2 \geq ... \geq \rho_m$, o likusieji lygūs 0. Taigi kanoninių kintamųjų vektoriaus $\boldsymbol{Y}$ koordinatės yra nekoreliuotos, išskyrus poras $\boldsymbol{L}_1^T\mathbf{X}_1$ ir $\boldsymbol{M}_1^T\mathbf{X}_2$, ..., $\boldsymbol{L}_m^T\mathbf{X}_1$ ir $\boldsymbol{M}_m^T\mathbf{X}_2$. Kanoninės koreliacijos išrikiuotos nemažėjančia tvarka $\rho_1 \geq \rho_2 \geq ... \geq \rho_m$.

Įrodymas. Naudodami Lagranžo neapibrėžtinius daugiklius (7.2.2) spręsti, gauname

$$F = \boldsymbol{L}^T\boldsymbol{\Sigma}_{12}\boldsymbol{M} - \frac{\lambda_1}{2}\boldsymbol{L}^T\boldsymbol{\Sigma}_{11}\boldsymbol{L} - \frac{\lambda_2}{2}\boldsymbol{M}^T\boldsymbol{\Sigma}_{22}\boldsymbol{M} \to \max_{\boldsymbol{L},\boldsymbol{M}}.$$

Diferencijuodami F pagal $\boldsymbol{L}$ ir pagal $\boldsymbol{M}$ gauname lygtis

$$\boldsymbol{\Sigma}_{12}\boldsymbol{M} - \lambda_1\boldsymbol{\Sigma}_{11}\boldsymbol{L} = 0, \quad -\lambda_2\boldsymbol{\Sigma}_{22}\boldsymbol{M} + \boldsymbol{\Sigma}_{21}\boldsymbol{L} = 0. \tag{7.2.6}$$

Iš pirmos lygties gauname $\lambda_1 = \boldsymbol{L}^T \boldsymbol{\Sigma}_{12} \boldsymbol{M}$, o iš antrosios $\lambda_2 = \boldsymbol{M}^T \boldsymbol{\Sigma}_{21} \boldsymbol{L}$, todėl $\lambda_1 = \lambda_2 = \rho$. Padauginę pirmą lygtį iš $\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}$ ir sudėję su antrąja, padauginta iš ρ, gauname

$$(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} - \rho^2 \boldsymbol{\Sigma}_{22})\boldsymbol{M} = 0.$$

Taigi vektorius $\boldsymbol{M}$ yra tikrinis vektorius, atitinkantis tikrinę charakteringosios lygties (7.2.4) šaknį. Tegu $\rho_1^2 \geq ... \geq \rho_s^2$ yra charakteristinės lygties (7.2.4) šaknys, o $\boldsymbol{M}$ – matrica, kurios stulpeliai yra šias šaknis atitinkantys tikriniai vektoriai. Iš algebros kurso (1 priedas (8.2.22)) žinome, kad matrica $\boldsymbol{M}$ tenkina tokias sąlygas

$$\boldsymbol{M}^T \boldsymbol{\Sigma}_{22} \boldsymbol{M} = \boldsymbol{I}, \quad \boldsymbol{M}^T \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} \boldsymbol{M} = \boldsymbol{R}_2, \tag{7.2.7}$$

čia $\boldsymbol{R}_2$ yra diagonalioji matrica su diagonaliniais elementais $\rho_1^2, ..., \rho_s^2$.

Analogiškai gauname, kad vektoriai $\boldsymbol{L}_i$ turi būti tikriniai vektoriai, atitinkantys (7.2.3) lygties šaknis. Tegu $\rho_1^2 \geq ... \geq \rho_r^2$ yra charakteringosios lygties (7.2.3) šaknys, o $\boldsymbol{L}$ – matrica, kurios stulpeliai yra šias šaknis atitinkantys tikriniai vektoriai. Tada

$$\boldsymbol{L}^T \boldsymbol{\Sigma}_{11} \boldsymbol{L} = \boldsymbol{I}, \quad \boldsymbol{L}^T \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} \boldsymbol{L} = \boldsymbol{R}_1. \tag{7.2.8}$$

čia $\boldsymbol{R}_1$ yra diagonalioji matrica su diagonaliniais elementais $\rho_1^2, ..., \rho_r^2$.

Lygčių (7.2.3) ir (7.2.4) nenulinės šaknys yra vienodos, todėl jos žymimos vienodais simboliais $\rho_1^2, \rho_2^2, \cdots$. Nenulinių šaknų skaičius yra lygus matricos $\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$ arba $\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$ rangui, kuris sutampa su matricos $\boldsymbol{\Sigma}_{12}$ rangu $m \leq \min(r, s)$. Lygtys (7.2.6) yra tenkinamos imant $\pm\rho_i$, kai ρ_i yra teigiama ρ_i^2 šaknis. Jeigu reikia, pakeičiant $\boldsymbol{L}_i$ arba $\boldsymbol{M}_i$ ženklą, galima pasiekti, kad koeficientai ρ_i būtų teigiami, t. y. $\rho_1 \geq \rho_2 \geq ... \geq \rho_m > 0$.

Iš (7.2.7) sąlygos $\boldsymbol{M}^T \boldsymbol{\Sigma}_{22} \boldsymbol{M} = \boldsymbol{I}$ išplaukia, kad kanoniniai kintamieji $\boldsymbol{M}_1^T \mathbf{X}_2$, ..., $\boldsymbol{M}_s^T \mathbf{X}_2$ yra nekoreliuoti ir turi vienetines dispersijas. Analogiškas tvirtinimas teisingas dėl kanoninių kintamųjų $\boldsymbol{L}_1^T \mathbf{X}_1, ..., \boldsymbol{L}_r^T \mathbf{X}_1$.

Imkime du kanoninius kintamuosius $\boldsymbol{L}_i^T \mathbf{X}_1$ ir $\boldsymbol{M}_i^T \mathbf{X}_2$ iš skirtingų grupių su vienodais indeksais. Jų kovariacija

$$\mathbf{Cov}\,(\boldsymbol{L}_i^T \mathbf{X}_1, \boldsymbol{M}_i^T \mathbf{X}_2) = \boldsymbol{L}_i^T \boldsymbol{\Sigma}_{12} \boldsymbol{M}_i.$$

Padauginę (7.2.6) pirmąją lygtį iš $\boldsymbol{L}_i^T$ gausime

$$\boldsymbol{L}_i^T \boldsymbol{\Sigma}_{12} \boldsymbol{M}_i = \rho_i \boldsymbol{L}_i^T \boldsymbol{\Sigma}_{11} \boldsymbol{L}_i = \rho_i.$$

Taigi skirtingų grupių kanoninių kintamųjų su vienodu indeksu koreliacijos koeficientas (kanoninė koreliacija) lygus ρ_i; jis teigiamas, kai $i = 1, ..., m$, ir lygus 0, kai $i > m$.

Imkime du kanoninius kintamuosius $\boldsymbol{L}_j^T \mathbf{X}_1$ ir $\boldsymbol{M}_i^T \mathbf{X}_2$ iš skirtingų grupių su skirtingais indeksais $i \neq j$. Jų kovariacija

$$\mathbf{Cov}\,(\boldsymbol{L}_j^T \mathbf{X}_1, \boldsymbol{M}_i^T \mathbf{X}_2) = \boldsymbol{L}_j^T \boldsymbol{\Sigma}_{12} \boldsymbol{M}_i.$$

Padauginkime (7.2.6) pirmąją lygtį iš $\boldsymbol{L}_j^T$. Gausime

$$\boldsymbol{L}_j^T \boldsymbol{\Sigma}_{12} \boldsymbol{M}_i = \rho_i \boldsymbol{L}_j^T \boldsymbol{\Sigma}_{11} \boldsymbol{L}_i = 0, \quad i \neq j. \qquad \blacktriangle$$

Aprašant dviejų vektorių priklausomumą svarbiausi yra tie kanoniniai kintamieji, kurių kanoninės koreliacijos yra didžiausios. Jeigu, pavyzdžiui,pirmosios kelios kanoninės koreliacijos $\rho_1 \geq ... \geq \rho_l$ yra dominuojančios, palyginti su koeficientais $\rho_{l+1}, ..., \rho_m$, tai apibūdinant vektorių priklausomybę dažnai pakanka apsiriboti pirmaisiais l kanoniniais kintamaisiais iš abiejų grupių.

7.2.1 pavyzdys. Panagrinėkime paprastą iliustracinį pavyzdį. Tarkime, a. v. $\mathbf{X} = (X_1, X_2, X_3, X_4)^T$ kovariacinė matrica $\boldsymbol{\Sigma}$ turi tokį pavidalą:

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 & \delta & \delta \\ 0 & 1 & \delta & \delta \\ \delta & \delta & 1 & 0 \\ \delta & \delta & 0 & 1 \end{pmatrix}.$$

Suskaidykime vektorių $\mathbf{X}$ į du vektorius $\mathbf{X}_1 = (X_1, X_2)^T$ ir $\mathbf{X}_2 = (X_3, X_4)^T$. Rasime kanoninius kintamuosius ir kanonines a. v $\mathbf{X}_1$ ir $\mathbf{X}_2$ koreliacijas.

Charakteringoji lygtis $|\boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12} - \rho^2 \boldsymbol{\Sigma}_{22}| = 0$ turi vieną nenulinę šaknį $\rho_1^2 = 4\delta^2$ (matricos $\boldsymbol{\Sigma}_{12}$ rangas lygus 1), o kita šaknis $\rho_2^2 = 0$. Šias šaknis atitinka tikriniai vektoriai $\boldsymbol{L}_1 = (1;\ 1)^T/\sqrt{2}$ ir $\boldsymbol{L}_2 = (1;\ -1)^T/\sqrt{2}$. Gauname kanoninius kintamuosius

$$Y_1 = \boldsymbol{L}_1^T \mathbf{X}_1 = (X_1 + X_2)/\sqrt{2}, \quad Y_2 = \boldsymbol{L}_2^T \mathbf{X}_1 = (X_1 - X_2)/\sqrt{2}.$$

Analogiškai iš charakteringosio lygties $|\boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{\Sigma}_{21} - \rho^2 \boldsymbol{\Sigma}_{11}| = 0$ gauname šaknis $\rho_1^2 = 4\delta^2, \rho_2^2 = 0$, kurias atitinka tikriniai vektoriai $\boldsymbol{M}_1 = (1;\ 1)^T/\sqrt{2}$ ir $\boldsymbol{M}_2 = (1;\ -1)^T/\sqrt{2}$ ir kanoniniai kintamieji

$$Y_3 = \boldsymbol{M}_1^T \mathbf{X}_2 = (X_3 + X_4)/\sqrt{2}, \quad Y_4 = \boldsymbol{M}_2^T \mathbf{X}_2 = (X_3 - X_4)/\sqrt{2}.$$

Atsitiktinio vektoriaus $\boldsymbol{Y} = (Y_1, Y_2, Y_3, Y_4)^T$ kovariacijų matrica yra

$$\boldsymbol{V}(\boldsymbol{Y}) = \begin{pmatrix} 1 & 0 & 2\delta & 0 \\ 0 & 1 & 0 & 0 \\ 2\delta & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Nenulinė kanoninė koreliacija lygi 2δ, t. y. a. d. Y_1 ir Y_3 koreliacijos koeficientui.

7.2.2. Kanoninių koreliacijų DT įvertiniai.

Tegu $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n$ yra paprastoji imtis a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Suskaidę vektorių $\mathbf{X}$ į r-matį ir $s = (k - r)$-matį vektorius, gausime šių dalinių vektorių paprastąsias imtis $\mathbf{X}_1^{(1)}, ..., \mathbf{X}_1^{(n)}$ ir $\mathbf{X}_2^{(1)}, ..., \mathbf{X}_2^{(n)}$, gautas stebint normaliuosius a. v. $N_r(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$ ir $N_s(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22})$. Parametrų DT įvertiniai

$$\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}} = \frac{1}{n} \sum_{j=1}^{n} \mathbf{X}_j, \quad \hat{\boldsymbol{\mu}}_i = \bar{\mathbf{X}}_i = \frac{1}{n} \sum_{j=1}^{n} \mathbf{X}_i^{(j)}, \quad i = 1,\ 2;$$

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n} \sum_{j=1}^{n} (\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})^T = \frac{1}{n} \begin{pmatrix} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} \\ \boldsymbol{S}_{21} & \boldsymbol{S}_{22} \end{pmatrix}, \qquad (7.2.9)$$

$$\boldsymbol{S}_{ii}=\sum_{j=1}^{n}(\mathbf{X}_i^{(j)}-\bar{\mathbf{X}}_i)(\mathbf{X}_i^{(j)}-\bar{\mathbf{X}}_i)^T,\quad i=1,\ 2;\quad \boldsymbol{S}_{12}=\sum_{j=1}^{n}(\mathbf{X}_1^{(j)}-\bar{\mathbf{X}}_1)(\mathbf{X}_2^{(j)}-\bar{\mathbf{X}}_2)^T.$$

7.2.2 teorema. *Tegu $n>k$, matricos $\boldsymbol{\Sigma}_{11}$ ir $\boldsymbol{\Sigma}_{22}$ teigiamai apibrėžtos, matricos $\boldsymbol{\Sigma}_{12}$ rangas $m=\min(r,s)$, o lygčių (7.2.10), (7.2.11) nenulinės šaknys yra skirtingos. Tada parametrų $\rho_1^2>...>\rho_m^2>0$ įvertiniai $\hat\rho_1^2>...>\hat\rho_m^2>0$ yra lygčių sistemos*

$$|\boldsymbol{S}_{12}\boldsymbol{S}_{22}^{-1}\boldsymbol{S}_{21}-\hat\rho^2\boldsymbol{S}_{11}|=0, \tag{7.2.10}$$

arba lygčių sistemos

$$|\boldsymbol{S}_{21}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12}-\hat\rho^2\boldsymbol{S}_{22}|=0 \tag{7.2.11}$$

šaknys. Vektorių $\boldsymbol{L}_1,...,\boldsymbol{L}_r$ įvertiniai $\hat{\boldsymbol{L}}_1,...,\hat{\boldsymbol{L}}_r$ yra tikriniai vektoriai, atitinkantys lygties (7.2.10) šaknis, o vektorių $\boldsymbol{M}_1,...,\boldsymbol{M}_s$ įvertiniai $\hat{\boldsymbol{M}}_1,...,\hat{\boldsymbol{M}}_s$ yra tikriniai vektoriai, atitinkantys lygties (7.2.11) šaknis. Matricos $\hat{\boldsymbol{L}}$ ir $\hat{\boldsymbol{M}}$ sudarytos atitinkamai iš vektorių $\hat{\boldsymbol{L}}_1,...,\hat{\boldsymbol{L}}_r$ ir $\hat{\boldsymbol{M}}_1,...,\hat{\boldsymbol{M}}_s$ tenkina sąlygas

$$\hat{\boldsymbol{L}}^T\boldsymbol{S}_{11}\hat{\boldsymbol{L}}=\boldsymbol{I}_r,\quad \hat{\boldsymbol{M}}^T\boldsymbol{S}_{22}\hat{\boldsymbol{M}}=\boldsymbol{I}_s.$$

Įrodymas. Suformuluotomis sąlygomis lygčių (7.2.3) ir (7.2.4) pirmosios m šaknys yra teigiamos ir skirtingos. Todėl jas atitinkantys tikriniai vektoriai $\boldsymbol{L}_1,...,\boldsymbol{L}_m$ ir $\boldsymbol{M}_1,...,\boldsymbol{M}_m$ nusakomi vienareikšmiškai (išskyrus ženklą). Šį neapibrėžtumą galima panaikinti parenkant, pavyzdžiui, kad pirmasis nenulinis šių vektorių elementas būtų teigiamas. Jeigu $r>m=s$, tai vektoriai $\boldsymbol{L}_{m+1},...,\boldsymbol{L}_r$, atitinkantys nulines lygties (7.2.3) šaknis $\rho_{m+1}^2,...,\rho_r^2$ apibrėžiami nevienareikšmiškai. Šį neapibrėžtumą galima panaikinti pasirenkant bet kokius apribojimus su sąlyga, kad šie vektoriai tenkina sąlygą (7.2.8). Tada tarp parametrų $\boldsymbol{\mu},\boldsymbol{\Sigma}$ ir parametrų $\boldsymbol{\mu},\rho_1^2,...,\rho_m^2,\boldsymbol{M},\boldsymbol{L}$ yra abipusė vienareikšmė priklausomybė. Šių parametrų DT įvertiniai gaunami pakeičiant (7.2.3), (7.2.4) lygtyse matricas $\boldsymbol{\Sigma}_{11},\boldsymbol{\Sigma}_{12},\boldsymbol{\Sigma}_{22}$ jų įvertiniais $\boldsymbol{S}_{11}/n,\boldsymbol{S}_{12}/n,\boldsymbol{S}_{22}/n$. ▲

7.2.2 pavyzdys (**1.8** pratimo tęsinys.) Rasime kanoninius kintamuosius ir kanonines koreliacijas, apibūdinančias a. v. $\mathbf{X}^{(1)}=(X_1,X_2)^T$ ir a. v. $\mathbf{X}^{(2)}=(X_3,X_4)^T$ priklausomybę, **1.8** pratimo sąlygomis.

Kovariacinės matricos $\boldsymbol{\Sigma}$ įvertis $\hat{\boldsymbol{\Sigma}}$ surastas 1.8 pratime. Jos blokai

$$\frac{\boldsymbol{S}_{11}}{n-1}=\begin{pmatrix}95,293 & 52,868\\ 52,868 & 54,360\end{pmatrix},\quad \frac{\boldsymbol{S}_{12}}{n-1}=\begin{pmatrix}69,662 & 46,112\\ 51,312 & 35,053\end{pmatrix},$$

$$\frac{\boldsymbol{S}_{22}}{n-1}=\begin{pmatrix}100,807 & 56,540\\ 56,540 & 45,023\end{pmatrix},\quad \boldsymbol{S}_{21}=\boldsymbol{S}_{12}^T.$$

Spręsdami charakteringąją lygtį (7.2.10) arba (7.2.11) randame tikrines reikšmes $\hat\rho_1^2=0,621751$, $\hat\rho_2^2=0,002885$ ir jas atitinkančius tikrinius vektorius $\hat{\boldsymbol{L}}_1=(0,06986,\quad 0,05230)^T$, $\hat{\boldsymbol{L}}_2=(-0,14825,\quad 0,11856)^T$, $\hat{\boldsymbol{M}}_1=(0,07122,\quad 0,04787)^T$, $\hat{\boldsymbol{M}}_2=(-0,14487,\quad 0,90734)^T$. Kanoniniai kintamieji

$$Y_1=0,06986X_1+0,05230X_2,\quad Y_2=-0,14825X_1+0,11856X_2,$$

$$Y_3=0,07122X_3+0,04787X_4,\quad Y_4=-0,14487X_3+0,90734Y_4.$$

Atsitiktinio vektoriaus $\boldsymbol{Y} = (Y_1, Y_2, Y_3, Y_4)^T$ kovariacinė matrica turi tokią struktūrą:

$$\hat{\boldsymbol{V}}(\boldsymbol{Y}) = \begin{pmatrix} 1 & 0 & \hat{\rho}_1 & 0 \\ 0 & 1 & 0 & \hat{\rho}_2 \\ \hat{\rho}_1 & 0 & 1 & 0 \\ 0 & \hat{\rho}_2 & 0 & 1 \end{pmatrix}.$$

Pirmoji kanoninė koreliacija $\hat{\rho}_1 = \sqrt{0,621751} = 0,7885$, o antroji yra daug mažesnė $\hat{\rho}_2 = \sqrt{0,002885} = 0,0537$. Apibūdinant a. v. $\mathbf{X}^{(1)} = (X_1, X_2)^T$ ir $\mathbf{X}^{(2)} = (X_3, X_4)^T$ priklausomybę matyt pakanka apsiriboti kanoniniais kintamaisiais Y_1 ir Y_3, kurių koreliacijos koeficiento įvertis yra 0,7885.

7.2.3. Hipotezės dėl kanoninių koreliacijų

Pirmasis uždavinys, kuris kyla nagrinėjant kanonines koreliacijas, matyt, yra atsakymas į klausimą, ar apskritai yra prasminga apibrėžti kanonines koreliacijas. Aišku, kad kanoninės koreliacijos neturi prasmės, jeigu vektoriai $\mathbf{X}_1$ ir $\mathbf{X}_2$ yra nekoreliuoti, t. y. reikia patikrinti hipotezę $H : \boldsymbol{\Sigma}_{12} = \mathbf{0}$. Hipotezę apie dviejų normaliųjų vektorių nepriklausomumą nagrinėjome 6.4.5.4 skyrelyje. Buvo įrodyta, kad tikėtinumų santykio statistikos laipsnis $\Lambda^{2/n}$ yra pasiskirstęs taip pat kaip nepriklausomų a. d. sandauga

$$\prod_{i=1}^{r} \xi_{2j}, \quad \xi_{2j} \sim Be(\frac{n-s-j}{2}, \frac{s}{2}), \quad j = 1, ..., r.$$

Asimptotinis $n \to \infty$ statistikos $-2\ln\Lambda$ skirstinys yra χ^2 skirstinys su $\nu = rs$ laisvės laipsnių. Remdamiesi šiais faktais sudarome tikslų ar apytikslį tikėtinumų santykio kriterijų hipotezei H tikrinti.

Minėjome, kad vertingiausi yra tie kanoniniai kintamieji $\boldsymbol{L}_1^T\mathbf{X}_1, ..., \boldsymbol{L}_l^T\mathbf{X}_1$ ir $\boldsymbol{M}_1^T\mathbf{X}_2, ..., \boldsymbol{M}_l^T\mathbf{X}_2$, kurių kanoninės koreliacijos $\rho_1 \geq \rho_2 \geq ... \geq \rho_l$ yra dominuojančios, palyginti su kanoninėmis koreliacijomis $\rho_{l+1}, ..., \rho_m$. Apsiribojant pirmosiomis l kanoninėmis koreliacijomis kyla klausimas, ar tarp atmestųjų koreliacijų $\rho_{l+1}, ..., \rho_m$ nėra išsiskiriančiųjų, kurias galbūt vertėtų pridėti prie atrinktųjų. Tuo tikslu galima patikrinti hipotezę, kad $\rho_{l+1} = ... = \rho_m$, t. y. kad atsitiktinių vektorių $(\boldsymbol{L}_{l+1}^T\mathbf{X}_1, ..., \boldsymbol{L}_m^T\mathbf{X}_1)^T$ ir $(\boldsymbol{M}_{l+1}^T\mathbf{X}_2, ..., \boldsymbol{M}_m^T\mathbf{X}_2)^T$ kovariacijų matrica proporcinga vienetinei matricai. Tokių hipotezių tikrinimo kriterijai sukonstruoti 6.5 skyrelyje.

7.3. Faktorinė analizė

Tarkime, kad stebime a. v. $\mathbf{X} = (X_1, ..., X_k)^T$ su kovariacijų matrica $\boldsymbol{\Sigma} = [\sigma_{ij}]_{k\times k}$. Aprašant a. v. $\mathbf{X}$ kovariacijų struktūrą, galima bandyti suskaidyti jo koordinates į grupes taip, kad vienos grupės kintamuosius vienytų koks nors tiesiogiai nestebimas faktorius. Pavyzdžiui, testais tikrinant studentų žinias, galima tarti, kad visi atsakymai priklauso nuo tam tikrų tiesiogiai nematuojamų faktorių: kūrybiškumas, matematiniai gabumai, vaizduotė, atmintis ir kt.

7.3.1. Matematinis modelis

Įvesdami pagrindines komponentes 7.1.1 skyrelyje atlikome a. v. $\mathbf{X} = (X_1, ..., X_k)^T$ ortonormuotą transformaciją

$$\boldsymbol{Y} = \boldsymbol{L}^T\mathbf{X}, \quad \boldsymbol{V}(\boldsymbol{Y}) = \boldsymbol{L}^T\boldsymbol{\Sigma}\boldsymbol{L} = \boldsymbol{\Delta};$$

$$Y_i = \boldsymbol{L}_i^T\mathbf{X} = \sum_{j=1}^{k} L_{ji}X_j, \quad \boldsymbol{V}(Y_i) = \sum_{r,s=1}^{k} L_{ir}\sigma_{rs}L_{js}. \tag{7.3.1}$$

Kadangi $\boldsymbol{L}^T\boldsymbol{L} = \boldsymbol{I}$, tai atvirkštinė transformacija, kuria a. v. $\mathbf{X}$ išreiškiamas pagrindinių komponenčių vektoriumi $\boldsymbol{Y} = (Y_1, ..., Y_k)^T$, yra

$$\mathbf{X} = \boldsymbol{L}\boldsymbol{Y}, \quad \boldsymbol{V}(\mathbf{X}) = \boldsymbol{\Sigma} = \boldsymbol{L}\boldsymbol{V}(\boldsymbol{Y})\boldsymbol{L}^T = \boldsymbol{L}\boldsymbol{\Delta}\boldsymbol{L}^T;$$

$$X_i = \sum_{j=1}^{k} L_{ij}Y_j, \quad \sigma_{ii} = \sum_{j=1}^{k} L_{ij}^2\boldsymbol{V}(Y_j), \quad \sigma_{ij} = \sum_{r=1}^{k} L_{ir}\boldsymbol{V}(Y_r)L_{jr}. \tag{7.3.2}$$

Šios formulės nusako naują kovariacinės matricos $\boldsymbol{\Sigma}$ struktūrą, arba faktorizaciją. Dispersijos σ_{ii} ir kovariacijos $\sigma_{ij}, i \neq j$, yra pagrindinių komponenčių dispersijų funkcijos. Faktorizacija (7.3.2) atlikta, kai faktorių (pagrindinių komponenčių) skaičius k sutampa su vektoriaus $\mathbf{X}$ dimensija.

Faktorinėje analizėje nagrinėjamas bendresnis modelis, kai $\boldsymbol{\Sigma}$ bandoma faktorizuoti mažesniu faktorių $F_1, ..., F_m$ skaičiumi $m < k$. Išreiškiant a. v. $\mathbf{X}$ faktorių $F_1, ..., F_m$ tiesine daugdara tenka pridėti paklaidos komponentę $\boldsymbol{e}$, kuri apibūdina tą sklaidos dalį, kuri negali būti paaiškinta faktoriais $F_1, ..., F_m$. Tiksliau, nagrinėjame modelį

$$\mathbf{X} = \boldsymbol{A}\boldsymbol{F} + \boldsymbol{e};$$

$$X_i = a_{i1}F_1 + ... + a_{im}F_m + e_i, \quad i = 1, ..., k; \tag{7.3.3}$$

čia $\boldsymbol{F} = (F_1, ..., F_m)^T$, matrica $\boldsymbol{A} = [a_{ij}]_{k\times m}$ ir vektorius $\boldsymbol{e} = (e_1, ..., e_k)^T$. Kintamieji $F_1, ..., F_m$ vadinami *bendraisiais*, arba *latentiniais*, faktoriais, nes jais išreiškiami visi kintamieji $X_1, ..., X_k$. Kintamieji $e_1, ..., e_k$ vadinami *specifiniais* faktoriais, nes kiekvienas e_i susijęs tik su vienu kintamuoju X_i. Priimamos prielaidos, kad a. d. $F_1, ..., F_m, e_1, ..., e_k$ yra nekoreliuoti. A. d. $F_1, ..., F_m$ turi vienetines dispersijas, o a. d. $e_1, ..., e_k$ dispersijos $\boldsymbol{V}(e_i) = \beta_i^2, i = 1, ..., k$, vadinamos *specifinėmis* dispersijomis, arba i-osios koordinatės X_i *specifiškumu*. Koeficientai a_{ij} vadinami *faktorių svoriais*.

Atsižvelgę į priimtas prielaidas gauname kovariacinės matricos $\boldsymbol{\Sigma}$ faktorizaciją, analogišką (7.3.2)

$$\boldsymbol{V}(\mathbf{X}) = \boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^T + \boldsymbol{B};$$

$$\sigma_{ii} = \sum_{j=1}^{k} a_{ij}^2 + \beta_i^2 = h_i^2 + \beta_i^2, \tag{7.3.4}$$

$$\sigma_{ij} = \sum_{r=1}^{k} a_{ir}a_{jr}, \quad i \neq j, \quad i,j = 1,...,k.$$

Dydis $h_i^2 = \sigma_{ii} - \beta_i^2$ vadinamas i-osios komponentės X_i *bendrumu*. Pagrindinių komponenčių schemoje dispersija σ_{ii} išskaidoma į k dalių, tenkančių pagrindinėms komponentėms. Faktorinės analizės schemoje dispersija σ_{ii} suskaidoma į dvi dalis: dispersija, tenkanti bendriesiems faktoriams h_i^2 (bendrumas), ir likusioji dispersijos dalis β_i^2, kuri faktoriais $F_1,...,F_m$ nepaaiškinama (specifiškumas).

Išoriškai (7.3.3) lygybė primena tiesinės regresijos modelį – žinodami F_i ir a_{ij} reikšmes galėtume prognozuoti a. d. X_i reikšmes. Tačiau faktorinės analizės uždavinys yra atvirkštinis – žinome tik X_i reikšmes, o norime išsiaiškinti, ką galima pasakyti apie bendruosius faktorius $F_1,...,F_m$.

Matrica $\boldsymbol{A}^T\boldsymbol{A}$, kurios elementai a_{ij} yra lygūs σ_{ij}, kai $i \neq j$, ir diagonaliniai elementai $a_{ii} = h_i^2$ yra bendrumai, vadinama *redukuotąja* kovariacine matrica. Redukuotoji matrica rodo, ar bendrieji faktoriai gerai paaiškina pradinių kintamųjų priklausomybės struktūrą. Kuo h_i^2 artimesnis σ_{ii}, tuo daugiau informacijos išsaugome pereidami nuo pradinių kintamųjų $X_1,...,X_k$ prie bendrųjų faktorių $F_1,...F_m$. Jeigu specialūs kintamieji $e_i = 0$, tai $\boldsymbol{A}^T\boldsymbol{A}$ sutampa su $\boldsymbol{\Sigma}$ ir faktoriai $F_1,...,F_m$ išsaugo visą informaciją apie kintamuosius $X_1,...,X_k$.

Faktorinės analizės tikslas yra įvertinti faktorių svorius a_{ij} ir specifines dispersijas $\beta_i^2, i = 1,...,k,\ j = 1,...,m$. Taip pat reikia padaryti sprendimą dėl priimtino bendrųjų faktorių skaičiaus m. Suradus parametrų įvertinius, reikia gauti kiekvienam kintamajam X_i faktorių reikšmių įvertinius.

Kai minėtieji įvertiniai gauti, iškyla uždavinys dėl priimtinos faktorių interpretacijos. Kad būtų lengviau juos interpretuoti, dažnai atliekamos bendrųjų faktorių transformacijos (pasukimai). Dėl šio etapo subjektyvumo, ši faktorinės analizės dalis sukelia daugiausia abejonių.

Faktorinė analizė jau išsiskyrė į atskirą statistikos mokslo sritį, todėl mes čia neturime galimybės detaliau aptarti faktorinės analizės problemų. Besidomintiems rekomenduotume literatūrą [6], [7], [9].

7.3.2. Parametrų įvertiniai

7.3.2.1 DT metodas

Tarkime, kad $\mathbf{X}_1,...,\mathbf{X}_n$ yra paprastoji imtis a. v. $\mathbf{X} = (X_1,...,X_k)^T$ su nuliniu vidurkiu ir teigiamai apibrėžta kovariacine matrica $\boldsymbol{\Sigma} = [\sigma_{ij}]_{k\times k}$. Įvertinę kovariacinę matricą $\boldsymbol{\Sigma}$ gausime $k(k+1)/2$ skirtingų įvertinių $\hat{\sigma}_{ij}$ (matrica $\boldsymbol{\Sigma}$ simetrinė). O dešinėje (7.3.4) lygybių pusėje yra km parametrų a_{ij} ir k parametrų β_i^2, iš viso $k(m+1)$ nežinomas parametras. Jeigu $(k+1)/2 < m+1$, tai sąryšių (7.3.4) nepakanka, kad būtų galima įvertinti visus parametrus. Parametrų skaičių galima sumažinti pareikalavus, pavyzdžiui, kad matrica $\boldsymbol{A}$ tenkintų sąlygą

$$\boldsymbol{A}^T\boldsymbol{B}^{-1}\boldsymbol{A} = \boldsymbol{\Delta}, \tag{7.3.5}$$

čia $\mathbf{\Delta}$ – diagonalioji matrica. Tada nežinomi parametrai susieti dar $m(m-1)/2$ papildomais sąryšiais. Kad parametrai būtų įvertinami, turėtų galioti nelygybė

$$\frac{k(k+1)}{2} \geq k(m+1) - \frac{m(m-1)}{2},$$

arba

$$(k-m)^2 \geq k+m. \tag{7.3.6}$$

Pavyzdžiui, jei $k=5$, tai nelygybę (7.3.6) tenkina $m \leq 2$; jeigu $k=10$, tai m turi neviršyti 6.

Jeigu $(k-m)^2 > k+m$, tai į (7.3.4) ir (7.3.5) įrašę kovariacinės matricos įvertinį gausime lygčių sistemą, kuri turi be galo daug sprendinių. Tada įvertį galima gauti DT metodu.

Tarkime, kad $\boldsymbol{F} \sim N_m(\mathbf{0}, \boldsymbol{I})$ ir $\boldsymbol{e} \sim N_k(\mathbf{0}, \boldsymbol{B})$. Tada a. v. $\mathbf{X}$ taip pat normalusis $\mathbf{X} \sim N(\mathbf{0}, \boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{B})$. Imties $\mathbf{X}_1, ..., \mathbf{X}_n$ tikėtinumo funkcija

$$L(\mathbf{\Sigma}) = (2\pi)^{-nk/2}|\mathbf{\Sigma}|^{-n/2}\exp\{-\frac{1}{2}\sum_{j=1}^{n}\mathbf{X}_j^T\mathbf{\Sigma}^{-1}\mathbf{X}_j\}, \quad \mathbf{\Sigma} = \boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{B}. \tag{7.3.7}$$

Maksimizuodami $L(\mathbf{\Sigma})$, kai yra papildomų apribojimų (7.3.5), gausime parametrų DT įvertinius $\tilde{\mathbf{\Sigma}} = \hat{\boldsymbol{A}}^T\hat{\boldsymbol{A}} + \hat{\boldsymbol{B}}$, $\hat{\boldsymbol{A}}^T\hat{\boldsymbol{B}}^{-1}\hat{\boldsymbol{A}} = \hat{\mathbf{\Delta}}$. DT lygčių sistema gana sudėtinga ir sprendinio tenka ieškoti taikant artutinius metodus. Kai kuriuose matematinės statistikos TPP gautos lygčių sistemos sprendimas yra realizuotas (žr. [16]).

Gautą įvertinį galima panaudoti hipotezei, kad kovariacinė matrica gali būti faktorizuota (7.3.4) ir (7.3.5) lygybėmis, tikrinti. Tuo tikslu sudarome tikėtinumo santykio statistiką

$$\Lambda = \frac{\max_{\mathbf{\Sigma}=\boldsymbol{A}^T\boldsymbol{A}+\boldsymbol{B}} L(\mathbf{\Sigma})}{\max_{\mathbf{\Sigma}} L(\mathbf{\Sigma})} = \frac{L(\tilde{\mathbf{\Sigma}})}{L(\hat{\mathbf{\Sigma}})}, \tag{7.3.8}$$

čia $\hat{\mathbf{\Sigma}}$ yra besąlyginis kovariacinės matricos $\mathbf{\Sigma}$ DT įvertinys

$$\hat{\mathbf{\Sigma}} = \frac{1}{n}\sum_{j=1}^{n}\mathbf{X}_j\mathbf{X}_j^T. \tag{7.3.9}$$

Jeigu suformuluotoji hipotezė teisinga, tai a. d. $-2\ln\Lambda$ asimptotiškai $n \to \infty$ turi χ^2 skirstinį su $\nu = [(k-m)^2-(k+m)]/2$ laisvės laipsnių. Hipotezė atmetama apytiksliu tikėtinumų santykio kriterijumi, kai

$$-2\ln\Lambda > \chi^2_\alpha(\nu). \tag{7.3.10}$$

Hipotezės atmetimas reiškia, kad faktorių $F_1, ..., F_m$ nepakanka kovariacinei matricai $\mathbf{\Sigma}$ faktorizuoti pagal (7.3.4) ir (7.3.5), t. y. faktorių skaičių reikėtų padidinti.

7.3.1 pastaba. Jeigu stebimo a. v. $\mathbf{X}$ vidurkis $\mathbf{E}(\mathbf{X}) = \boldsymbol{\mu} \neq \mathbf{0}$, tai (7.3.4) ir (7.3.5) reikia imti centruotą vektorių $\mathbf{X} - \boldsymbol{\mu}$, o (7.3.8) skaitiklį ir vardiklį reikia maksimizuoti ir parametro $\boldsymbol{\mu}$ atžvilgiu.

7.3.2 pastaba. Prieš realizuojant faktorinę analizę vertėtų atlikti pradinį kovariacinės matricos $\boldsymbol{\Sigma}$ tyrimą remiantis jos DT įvertiniu $\hat{\boldsymbol{\Sigma}}$. Pirmiausia reikia atsakyti į klausimą, ar apskritai faktorinė analizė turi prasmę. Jeigu a. v. $\mathbf{X}$ koordinatės nepriklausomos, tai, aišku, jokių bendrų faktorių nėra ir dispersija $\sigma_{ii} = \beta_i^2$ sutampa su specifiškumu. Jeigu nepriklausomumo hipotezė atmetama, tai reikia išskirti tokias koordinates, kurios nepriklauso nuo visų likusiųjų. Pagaliau, likusį vektorių reikia bandyti sudalyti į keletą mažesnės dimensijos vektorių taip, kad skirtingų vektorių koordinatės būtų nepriklausomos. Tada faktorinę analizę būtų tikslinga taikyti kiekvienam gautajam mažesnės dimensijos vektoriui atskirai. Visų čia išvardytų hipotezių tikrinimo kriterijai pateikiami 6.4 skyrelyje.

7.3.3.2 Pagrindinių komponenčių metodas

Kitas metodas, parenkant bendruosius faktorius, įvertinant jų svorius ir specifines dispersijas, yra grindžiamas pagrindinėmis komponentėmis. Šis metodas yra paprastesnis ir lengviau paaiškinamas, tačiau jį naudojant yra daugiau subjektyvumo ir, griežtai kalbant, gauti įvertiniai netenkina sąlygų (7.3.4).

Pagrindiniais faktoriais imkime atitinkamai normuotas pirmąsias m pagrindines komponentes, apibrėžtas (7.3.1)

$$F_j = \frac{Y_j}{\boldsymbol{V}(Y_j)^{1/2}}, \quad j = 1, ..., m. \tag{7.3.11}$$

Tada, palyginę (7.3.3) ir (7.3.1), matome, kad faktorių svoriai yra

$$a_{ij} = L_{ij}[\boldsymbol{V}(Y_j)]^{1/2} \quad i = 1, ..., k, \quad j = 1, ..., m, \tag{7.3.12}$$

o specifiniai faktoriai e_i nusakomi lygybėmis

$$\mathrm{e}_i = \sum_{j=m+1}^{k} L_{ij} Y_j, \quad i = 1, ..., k. \tag{7.3.13}$$

Gauname faktorinį modelį (7.3.3)

$$X_i = \sum_{j=1}^{m} a_{ij} F_j + e_i = \sum_{j=1}^{m} L_{ij}[\boldsymbol{V}(Y_j)]^{1/2} F_j + e_i, \quad i = 1, ..., k. \tag{7.3.14}$$

Šiame modelyje $F_1, ..., F_m$ turi vienetines dispersijas ir yra nekoreliuoti. Kiekvienas faktorius F_j nekoreliuotas su bet kuriuo e_i. Tačiau

$$\mathbf{Cov}\,(e_i, e_j) = \sum_{l=m+1}^{k} a_{il} a_{jl} \boldsymbol{V}(Y_l), \quad i \neq j \quad i, j = 1, ..., k,$$

nebūtinai lygi nuliui, t. y. viena iš modelio (7.3.3) sąlygų gali būti neįvykdyta.

Kintamojo $X_i,\ i=1,...,k$, bendrumas ir specifiškumas yra

$$h_i^2=\sum_{j=1}^m a_{ij}^2=\sum_{j=1}^m L_{ij}^2\boldsymbol{V}(y_j),\quad \beta_i^2=\sum_{j=m+1}^k L_{ij}^2\boldsymbol{V}(Y_j). \tag{7.3.15}$$

Kai pagrindinės komponentės randamos iš empirinės kovariacijų matricos, tai formulėse (7.3.11) – (7.3.15) parametrus reikia pakeisti jų įvertiniais $\hat{L}_{ij},\hat{\boldsymbol{V}}(Y_j)=\hat{\lambda}_j,\ i,j=1,...,k$, iš 6.7.3 teoremos.

7.3.3 pastaba. Faktorinėje analizėje vietoje kovariacijų matricos dažniau naudojama koreliacijos koeficientų matrica, nes naudojant bedimensinius dydžius $(X_i-\mu_i)/\sqrt{\sigma_{ii}}$ gautus faktorius paprasčiau interpretuoti.

7.3.4 pastaba. Bendrųjų faktorių skaičius m dažniausiai parenkamas a) atsižvelgiant į stebimų kovariančių $X_1,...,X_k$ prasmę; b) imamos tos pagrindinės komponentės, kurių tikrinės reikšmės viršija parinktą konstantą (pavyzdžiui, naudojant koreliacijos koeficientų matricą, jei viršija 1); c) parenkama tiek pagrindinių komponenčių, kad jų dispersijų suma sudarytų didesniąją visų dispersijų σ_{ii} sumos dalį.

7.3.3.3 Faktorių interpretavimas

Tai, matyt, sunkiausias ir subjektyviausias faktorinės analizės etapas. Kadangi kintamojo X_i ir faktoriaus F_j kovariacija yra a_{ij}, tai interpretuojant faktorių F_j natūralu atsižvelgti į tuos kintamuosius X_i, su kuriais faktorius F_j turi didžiausias koreliacijas. Dažnai rekomenduojama tokia taisyklė: faktorius F_j laikytinas susijusiu su tais kintamaisiais X_i, kuriems įvertinys $\hat{a}_{ij}=\hat{L}_{ij}\sqrt{\hat{\lambda}_j}$ absoliučiuoju didumu viršija 0,4.

Naudojant gautąją pradinę faktorių svorių matrica, „prasminius“ faktorius gana sunku identifikuoti. Dažnai $\hat{a}_{ij}$ viršija 0,4 daugeliui faktorių F_j, t.y. kintamasis X_i susijęs su daugeliu faktorių. Interpretacijai palengvinti dažnai naudojamas vadinamasis *faktorių pasukimas*.

Vietoje (7.3.4) nagrinėkime modelį

$$\mathbf{X}=\tilde{\boldsymbol{A}}\tilde{\boldsymbol{F}}+\boldsymbol{e},\quad X_i=\sum_{j=1}^m\tilde{a}_{ij}\tilde{F}_j+e_i,\quad i=1,...,k;$$

čia $\tilde{\boldsymbol{F}}=\boldsymbol{C}\boldsymbol{F}$, o $\boldsymbol{C}=[C_{ij}]_{m\times m}$ ortogonali matrica $\boldsymbol{C}\boldsymbol{C}^T=\boldsymbol{I}$. Tada naujo bendrųjų faktorių vektoriaus $\tilde{\boldsymbol{F}}$ koordinatės taip pat nekoreliuotos ir turi vienetines dispersijas. Vietoje (7.3.4) gauname kitą kovariacijų (ar koreliacijų) matricos faktorizaciją

$$\boldsymbol{V}(\mathbf{X})=\tilde{\boldsymbol{A}}\tilde{\boldsymbol{A}}^T+\boldsymbol{B}, \tag{7.3.16}$$

kuri išsaugo nepakitusius bendrumus h_i^2 ir specifiškumus β_i^2, tačiau svoriai $\tilde{a}_{ij}$ gali nesutapti su svoriais a_{ij}. Ieškoma tokios transformacijos, kad matrica $\tilde{\boldsymbol{A}} = [\tilde{a}_{ij}]_{m\times m}$ turėtų kuo paprastesnę struktūrą, t. y. kad dauguma svorių $\tilde{a}_{ij}$ būtų artimi nuliui, o tik nedidelė jų dalis turėtų palyginti didesnes reikšmes. Tada kiekvienas kintamasis X_i laikytinas susijusiu tik su keletu (ar vienu) faktorių.

Praktiškai labiausiai paplitęs sukimo metodas *varimax*, kai transformacija parenkama taip, kad būtų maksimizuojamas reiškinys

$$\sum_{j=1}^{m}\sum_{i=1}^{k}(\tilde{a}_{ij}^2 - (\bar{a_j^2}))^2, \quad \bar{a_j^2} = \frac{1}{k}\sum_{i=1}^{k}\tilde{a}_{ij}^2. \tag{7.3.17}$$

Taigi šis metodas reiškia, kad maksimizuojama suma faktorių svorių kvadratų sklaidų, apibrėžtų kiekvienam faktoriui $\tilde{F}_j$, t. y. didesnieji svoriai turės tendenciją padidėti, o mažesnieji – sumažėti.

Faktorinėje analizėje yra gauta daug įvairių grafinių ir analizinių metodų, siekiant gauti kuo paprastesnę ir lengviau interpretuojamą faktorių svorių matricos struktūrą. Interpretavimui palengvinti kartais naudojamos ir neortogonalios transformacijos, t. y. atsisakoma sąlygos, kad faktoriai $\tilde{F}_1, ..., \tilde{F}_m$ būtų nekoreliuoti.

Kartais reikia rasti faktoriaus F_j prognozę konkrečiam vektoriui $\mathbf{X}_0 = (X_{10}, ..., X_{k0})^T$. Pavyzdžiui, atlikus studentų žinių tikrinimo testų duomenų faktorinę analizę, buvo išskirti faktoriai: kūrybiškumas, matematiniai gabumai ir kt. Kokias šių faktorių reikšmes galime priskirti studentui, kurio testo duomenys yra $\mathbf{X}_0$?

Jeigu pagrindiniai faktoriai išskirti pagrindinių komponenčių metodu, tai galime tiesiog imti

$$\hat{F}_j = \hat{Y}_j/\sqrt{\hat{\lambda}_j},$$

čia $\hat{Y}_j$ yra j-osios pagrindinės komponentės įvertinys, o $\hat{\lambda}_j$ jos tikrinės reikšmės įvertinys.

Kitas būdas – regresinės analizės metodika. Tariama, kad faktoriai yra priklausomi kintamieji, o X_i – nepriklausomi kintamieji. Įvertinę regresijos koeficientus, gauname faktoriaus reikšmės prognozę

$$\hat{F}_j = \sum_{i=1}^{k}\hat{b}_{ij}Z_{i0}, \quad j = 1, ..., m; \tag{7.3.18}$$

čia $Z_{i0} = (X_{i0} - \bar{X}_i)/s_i$ yra standartizuotoji kintamojo X_{i0} reikšmė, o $\hat{b}_{ij}$ – regresijos koeficientų įvertiniai. Pažymėjus $\hat{\boldsymbol{b}}_j = (\hat{b}_{1j}, ..., \hat{b}_{kj})^T$, $\hat{\boldsymbol{L}}_j = (\hat{L}_{1j}, ..., \hat{L}_{kj})^T$, o $\boldsymbol{R}$ – empirinę koreliacijų matricą, tai $\hat{\boldsymbol{b}} = \boldsymbol{R}^{-1}\hat{\boldsymbol{L}}_j$.

7.3.1 pavyzdys. Pateiksime pavyzdį iš knygos [9]. Tiriama penkių skirtingo tipo akcijų grąža, žymėsime $\mathbf{X} = (X_1, X_2, X_3, X_4, X_5)^T$. Už 1975 metų sausio – 1976 metų gruodžio laikotarpį gauta $n = 100$ a. v. $\mathbf{X}$ stebėjimų savaitiniais intervalais, iš kurių įvertinta koreliacijos koeficientų matrica:

$$\hat{\boldsymbol{R}} = \begin{pmatrix} 1,000 & 0,577 & 0,509 & 0,387 & 0,462 \\ 0,577 & 1,000 & 0,599 & 0,389 & 0,322 \\ 0,509 & 0,599 & 1,000 & 0,436 & 0,426 \\ 0,387 & 0,389 & 0,436 & 1,000 & 0,523 \\ 0,462 & 0,322 & 0,426 & 0,523 & 1,000 \end{pmatrix}$$

Naudodami pagrindinių komponenčių ir DT metodus bandysime a. v. $\mathbf{X}$ apibūdinti dviem faktoriais F_1 ir F_2.

Randame tikrines reikšmes $\hat{\lambda}_1 = 2,857$, $\hat{\lambda}_2 = 0,809$, $\hat{\lambda}_3 = 0,540$, $\hat{\lambda}_4 = 0,452$, $\hat{\lambda}_5 = 0,343$. Pirmųjų dviejų tikrinių reikšmių suma sudaro 0,73 vektoriaus $\mathbf{X}$ koordinačių dispersijų įverčių sumos. Joms atitinka tikrinių vektorių įverčiai

$$\hat{\boldsymbol{L}}_1 = (0,464 \;\; 0,457 \;\; 0,470 \;\; 0,421 \;\; 0,421)^T, \quad \hat{\boldsymbol{L}}_2 = (0,240 \;\; 0,509 \;\; 0,260 \;\; -0,526 \;\; -0,582)^T.$$

Pagrindinių komponenčių metodu remiantis (7.3.12) svorių matricos elementai $\hat{a}_{1j} = \hat{L}_{1j}\sqrt{\hat{\lambda}_1}$, $\hat{a}_{2j} = \hat{L}_{2j}\sqrt{\hat{\lambda}_2}$, $j = 1, ..., 5$; čia $\hat{L}_{1j}, \hat{L}_{2j}$ yra vektorių $\hat{\boldsymbol{L}}_1, \hat{\boldsymbol{L}}_2$ koordinatės. Svoriai $\hat{a}_{1j}, \hat{a}_{2j}$ pateikiami lentelės antrame ir trečiame stulpeliuose. Ketvirtame stulpelyje pateikiami specifinių dispersijų įverčiai $\hat{\beta}_j^2$.

Naudojant SAS programų paketą DT metodu gaunami svorių matricos įverčiai $\hat{b}_{1j}$ ir $\hat{b_{2j}}$.

	$\hat{a}_{1j}$	$\hat{a}_{2j}$	$\hat{\beta}_j^2$	$\hat{b}_{1j}$	$\hat{b}_{1j}$	$\hat{b}_{1j}^*$	$\hat{b}_{2j}^*$	$\hat{\beta}_j^2$
X_1	0,784	0,216	0,338	0,684	0,189	$\mathbf{0,601}$	0,377	0,496
X_2	0,772	0,458	0,194	0,694	0,517	$\mathbf{0,850}$	0,164	0,251
X_3	0,794	0,223	0,315	0,681	0,248	$\mathbf{0,643}$	0,335	0,475
X_4	0,712	-0,473	0,269	0,621	-0,073	0,365	$\mathbf{0,507}$	0,609
X_5	0,712	-0,523	0,220	0,792	-0,442	0,208	$\mathbf{0,887}$	0,177

Matome, kad abu metodai duoda panašius rezultatus. Faktorius sunku interpretuoti. Sudarant faktorių F_1, visi kintamieji X_i dalyvauja su beveik vienodais svoriais.

Atlikę DT metodu gautų faktorių ortogonalią transformaciją naudojant metodą *varimax*, gauname naujus svorius (žymėta $\hat{b}_{1j}^*$ ir $\hat{b}_{2j}^*$). Situacija pasikeičia: sudarant faktorių $\boldsymbol{F}_1^*$ didžiausią įtaką turi kintamieji X_1, X_2, X_3, o sudarant faktorių $\boldsymbol{F}_2^*$ – kintamieji X_4, X_5.

7.4. Pratimai

7.1. Raskite kovariacinės matricos $\mathbf{\Sigma} = [\sigma_{ij}]_{2\times 2}$, kai $\sigma_{11} = \sigma_{22} = 1$, $\sigma_{12} = \sigma_{21} = \rho$ tikrines reikšmes ir tikrinius vektorius.

7.2. Įrodykite, kad kovariacinės matricos $\mathbf{\Sigma} = [\sigma_{ij}]_{k\times k}$, kai $\sigma_{11} = ... = \sigma_{kk} = 1$, $\sigma_{ij} = \rho$, $i \neq j = 1, ..., k$, didžiausioji tikrinė reikšmė yra $\lambda_1 = 1 + (k-1)\rho$, o visos kitos tikrinės reikšmės vienodos ir lygios $\lambda_j = 1 - \rho, j = 2, ..., k$. Pirmasis tikrinis vektorius yra proporcingas vektoriui $(1, ..., 1)^T$.

7.3. Tegu $\mathbf{\Sigma} = \mathbf{\Phi} + \sigma^2\boldsymbol{I}$; čia $\mathbf{\Phi}$ – neneigiamai apibrėžta simetrinė matrica. Įrodykite, kad kiekvienas matricos $\mathbf{\Phi}$ tikrinis vektorius yra ir matricos $\mathbf{\Sigma}$ tikrinis vektorius, o kiekviena matricos $\mathbf{\Sigma}$ tikrinė reikšmė yra matricos $\mathbf{\Phi}$ tikrinės reikšmės ir σ^2 suma.

7.4. Tegu $\boldsymbol{Y} = (Y_1, ..., Y_k)^T$ ir $\mathbf{X} = (X_1, ..., X_m)^T$ du atsitiktiniai vektoriai, o $\mathbf{\Sigma}$ bendra vektoriaus $(\boldsymbol{Y}^T, \mathbf{X}^T)^T$ kovariacijų matrica

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix};$$

čia $\mathbf{\Sigma}_{11} = \boldsymbol{V}(\boldsymbol{Y}), \mathbf{\Sigma}_{22} = \boldsymbol{V}(\boldsymbol{Y}), \mathbf{\Sigma}_{12} = \mathbf{Cov}\,(\boldsymbol{Y}, \mathbf{X})$. Pažymėkime σ_i^2 liekamąją kvadratų sumą prognozuojant Y_i geriausiu būdu parinkta tiesine prognoze $\hat{Y}_i = b_1X_1 + ... + b_mX_m$.

a) Įrodykite, kad vektorius $(b_1, ..., b_m)^T$, minimizuojantis $\sigma_1^2 + ... + \sigma_k^2$, yra tikrinis vektorius, atitinkantis didžiausiąją lygties

$$|\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{12} - \lambda\mathbf{\Sigma}_{22}| = 0.$$

šaknį. b) Įrodykite, kad r tiesinių a. v. $\mathbf{X}$ funkcijų, minimizuojančių $\sigma_1^2 + ... + \sigma_k^2$, yra r tikrinių vektorių, atitinkančių didžiausias pateiktos lygties šaknis.

7.5. Tegu $\boldsymbol{Y} = (Y_1, ..., Y_k)^T$ yra k-matis vektorius. Nagrinėkime r tiesinių $\boldsymbol{Y}$ funkcijų $\boldsymbol{Z}_j = \boldsymbol{L}_j^T\boldsymbol{Y}, j = 1, ..., r$, ir tegu σ_i^2 liekamoji kvadratų suma prognozuojant Y_i tiesine a. v. $\boldsymbol{Z} = (Z_1, ..., Z_r)^T$ funkcija. Parinkite vektorius $\boldsymbol{L}_1, ..., \boldsymbol{L}_r$ taip, kad būtų minimizuojama suma $\sum_i \omega_i^2\sigma_i^2$, čia ω_i^2 žinomi svoriai.

7.6. Raskite tikrinių reikšmių ir pagrindinių komponenčių įverčius remiantis kovariacijų matricos įverčiu, gautu pagal **1.8.** pratimo duomenis. Palyginkite gautus atsakymus su tikrinių reikšmių ir pagrindinių komponenčių įverčiais, gautais naudojant koreliacijų matricos įvertį.

7.7. Raskite tikrinių reikšmių ir pagrindinių komponenčių įverčius pagal **2.9** pratimo duomenis. Kurią dalį bendros sklaidos nusako pirmoji pagrindinė komponentė?

7.8. Raskite tikrinių reikšmių ir pagrindinių komponenčių įverčius pagal **4.19** pratimo duomenis. Kiek reikia paimti pagrindinių komponenčių, kad jos paaiškintų ne mažiau kaip 0,9 bendros sklaidos?

7.9. Raskite tikrinių reikšmių ir pagrindinių komponenčių įverčius pagal **5.9** pratimo duomenis trimis atvejais. Palyginkite pagrindines komponentes, gautas trims skirtingoms iriso rūšims.

7.10. Raskite tikrinių reikšmių ir pagrindinių komponenčių įverčius pagal **2.17** pratimo duomenis. Kurie kintamieji daro didžiausią įtaką apibrėžiant pirmąją pagrindinę komponentę?

7.11. Raskite kanoninių koreliacijų ir kanoninių kintamųjų įverčius tarp a. v. $(X_1, X_2)^T$ ir $(Z_1, ..., Z_6)^T$ pagal **3.6** pratimo duomenis. Raskite tikrinių reikšmių ir pagrindinių komponenčių įverčius naudodami a. v. $(Z_1, ..., Z_6)^T$ kovariacijų matricos įvertį.

7.12. Pagal egzaminų rezultatus vertinant $n = 220$ studentų šešių dalykų : škotų kalba, anglų kalba, istorija, aritmetika, algebra, geometrija (žymėsime $X_1, X_2, ..., X_6$), gautas koreliacijos koeficientų matricos įvertis [9]

$$\hat{\boldsymbol{R}} = \begin{pmatrix} 1,000 & 0,439 & 0,410 & 0,288 & 0,329 & 0,248 \\ 0,439 & 1,000 & 0,351 & 0,354 & 0,320 & 0,329 \\ 0,410 & 0,351 & 1,000 & 0,164 & 0,190 & 0,181 \\ 0,288 & 0,354 & 0,164 & 1,000 & 0,595 & 0,470 \\ 0,329 & 0,320 & 0,190 & 0,595 & 1,000 & 0,464 \\ 0,248 & 0,329 & 0,181 & 0,470 & 0,464 & 1,000 \end{pmatrix}.$$

Parinkite du faktorius ir pateikite jų interpretaciją.

7.13. Atsitiktinio vektoriaus $(X_1, X_2, X_3)^T$ kovariacijų matrica yra

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & 0,6 & 0,5 \\ 0,6 & 1 & 0,4 \\ 0,5 & 0,4 & 1 \end{pmatrix}.$$

Sudarykite faktorinės analizės modelį, kai $m = 1$. Apskaičiuokite faktoriaus svorių matricą. Kurie kintamieji daro didžiausią įtaką apibrėžiant bendrąjį faktorių? Kokią dalį bendrosios dispersijos paaiškina bendrasis faktorius?

7.14. Naudojant koreliacijos koeficientų matricą gauta dviejų faktorių svorių matrica

F_1	0,7	0,8	0,7	0,8	0,6	0,5	0,6	0,7
F_2	0,3	0,0	0,0	0,6	0,5	0,0	0,4	0,6

Koks yra kiekvieno faktoriaus indėlis į bendrąją dispersiją?

7.15. Atlikite faktorinę analizę pagal **4.20** pratimo duomenis. Išskirkite du pagrindinius faktorius ir pasiūlykite jų interpretaciją.

7.16. Pradinių klasių mokiniai sprendė uždavinius (kintamasis ARIT); demonstravo erudiciją (kintamasis INF); pagal spalvas komponavo kubelius (kintamasis KK); iš dalių surinkinėjo objektus (OS); vertino objektų panašumą (PAN); išdėstydavo paveiksliukus nuoseklia tvarka (PI); užbaigdavo paveiksliuką (PU); bandė įsiminti ir pakartoti skaičių seką (SE); ieškojo piešinyje simbolių (SP); demonstravo žodyno turtingumą (ŽOD). Šių kintamųjų koreliacijų matricos įvertis pateikiamas lentelėje [3].

	INF	ARIT	PAN	SE	ŽOD	SP	PU	OS	PI	KK
INF	1,00	0,34	0,40	0,27	0,59	0,09	0,25	0,27	0,22	0,26
ARIT	0,34	1,00	0,36	0,28	0,33	0,18	0,32	0,38	0,29	0,30
PAN	0,40	0,36	1,00	0,22	0,35	0,08	0,31	0,26	0,25	0,20
SE	0,27	0,28	0,22	1,00	0,29	0,16	0,14	0,18	0,15	0,22
ŽOD	0,59	0,33	0,35	0,29	1,00	0,08	0,27	0,24	0,28	0,26
SP	0,09	0,18	0,08	0,16	0,08	1,00	0,19	0,13	0,22	0,17
PU	0,25	0,32	0,31	0,14	0,27	0,19	1,00	0,36	0,36	0,40
OS	0,27	0,38	0,26	0,18	0,24	0,13	0,36	1,00	0,30	0,60
PI	0,22	0,29	0,25	0,15	0,28	0,22	0,36	0,30	1,00	0,25
KK	0,26	0,30	0,20	0,22	0,26	0,17	0,40	0,60	0,25	1,00

Išskirkite keturis faktorius ir pateikite jų interpretaciją.

Atsakymai ir nurodymai

7.1. $\lambda_1 = 1+\rho, \lambda_2 = 1-\rho$; $\boldsymbol{L}_1 = (1/\sqrt{2}; 1/\sqrt{2})^T$, $\boldsymbol{L}_2 = (1/\sqrt{2}; -1/\sqrt{2})^T$. **7.12.** DT metodu gauname svorių matricos elementų įverčius $\hat{a}_{1i}$ ir $\hat{a}_{2i}, i = 1, ..., 6$

	$\hat{a}_{1i}$	$\hat{a}_{2i}$	$\hat{a}^*_{1i}$	$\hat{a}^*_{2i}$	$\hat{h}^2_i$
X_1	0,553	0,429	0,369	0,594	0,490
X_2	0,568	0,288	0,433	0,467	0,406
X_3	0,392	0,450	0,211	0,558	0,356
X_4	0,740	-0,273	0,789	0,001	0,623
X_5	0,724	-0,211	0,752	0,054	0,568
X_6	0,595	-0,132	0,604	0,083	0,372

Atlikę ortogonalią transformaciją naudodami *varimax* metodą, gauname vaizdesnę svorių matricą su elementais $\hat{a}^*_{1i}$ ir $\hat{a}^*_{2i}$. Faktorius $\boldsymbol{F}^*_1$ apibūdina gabumus tiksliesiems mokslams, o faktorius $\boldsymbol{F}^*_2$ – gabumus humanitariniams mokslams. **7.13.** Randame tikrines reikšmes $\lambda_1 = 2,00446$, $\lambda_2 = 0,61309$, $\lambda_3 = 0,38245$ ir jas atitinkančius tikrinius vektorius $\boldsymbol{L}_1 = (0,61332;\ 0,57990;\ 0,53623)^T$, $\boldsymbol{L}_2 = (-0,17918;\ -0,55906;\ 0,80953)^T$,
$\boldsymbol{L}_3 = (-0,76924;\ 0,59257;\ 0,23902)^T$. Svorių matricos elementai $a_{11} = 0,86833$, $a_{12} = 0,82102$, $a_{13} = 0,75919$. Faktorius F_1 paaiškina 0,668 dalį sklaidos. **7.14.** 0,465; 0,1525.

8 skyrius

1 Priedas. Tiesinės algebros elementai

Daugiamatėje statistikoje yra patogu naudoti vektorinius ir matricinius žymenis. Tai labai supaprastina formules ir padeda geriau suvokti dėstomą medžiagą. Reikalingos matematinėje statistikoje tiesinės algebros žinios yra susistemintos knygoje [14], kurioje duota ir kompaktiški pateikiamų faktų įrodymai. Šiame priede pateikiami tiesinės algebros faktai, kuriais remiamasi šioje knygoje.

8.1. Vektoriai

1P.1 apibrėžimas. Dimensijos k vektoriumi (vektoriumi – stulpeliu) $\boldsymbol{x}$ vadinsime stulpelį, kurio elementai $x_1, ..., x_k$ yra realūs skaičiai, t. y. vektorių traktuosime kaip k-matės Euklido erdvės $\mathbf{R}^k$ elementą. Transponuotą vektorių (vektorių – eilutę) žymėsime $\boldsymbol{x}^T = (x_1, ..., x_k)$.

1. *Vektorių sudėtis.* Vienodos dimensijos vektorių $\boldsymbol{x}^T = (x_1, ..., x_k)$ ir $\boldsymbol{y}^T = (y_1, ..., y_k)$ suma yra vektorius $(\boldsymbol{x}+\boldsymbol{y})^T = (x_1+y_1, ..., x_k+y_k)$, kurio elementai gaunami sudedant atitinkamus dėmenų elementus. Sumos operacija tenkina komutatyvumo

$$\boldsymbol{x}+\boldsymbol{y} = \boldsymbol{y}+\boldsymbol{x} \tag{8.1.1}$$

ir distributyvumo

$$(\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z} = \boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z}) \tag{8.1.2}$$

dėsnius. Egzistuoja nulinis $\mathbf{0}^T = (0, ..., 0)$ ir atvirkštinis $(-\boldsymbol{x})^T = (-x_1, ..., -x_k)$ vektoriai, kad

$$\boldsymbol{x}+\mathbf{0} = \boldsymbol{x}, \quad \boldsymbol{x}+(-\boldsymbol{x}) = \mathbf{0}. \tag{8.1.3}$$

2. *Daugyba iš skaliaro.* Vektoriaus $\boldsymbol{x}^T = (x_1, ..., x_k)$ sandauga iš skaliaro $c \in \mathbf{R}$ suprantamas vektorius $(c\boldsymbol{x})^T = (cx_1, ..., cx_k)$, kuris gaunamas padauginant iš c visas vektoriaus $\boldsymbol{x}$ koordinates. Šis veiksmas tenkina distributyvumo dėsnius

$$c(\boldsymbol{x}+\boldsymbol{y}) = c\boldsymbol{x}+c\boldsymbol{y}, \quad (c_1+c_2)\boldsymbol{x} = c_1\boldsymbol{x}+c_2\boldsymbol{x}, \tag{8.1.4}$$

ir asociatyvumo dėsnį
$$c_1(c_2\boldsymbol{x}) = (c_1c_2)\boldsymbol{x}. \tag{8.1.5}$$

1P.2 apibrėžimas. Dimensijos k vektorių, kuriems apibrėžtos sudėties ir daugybos iš skaliaro operacijos, visumą vadiname *tiesine k-mate Euklido erdve* $\mathbf{R}^k$.

1P.3 apibrėžimas. Erdvės $\mathbf{R}^k$ poaibis $\mathcal{M}$ uždaras sudėties ir daugybos iš skaliaro operacijų atžvilgiu, t. y. jei $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{M}$, tai $c\boldsymbol{x} + d\boldsymbol{y} \in \mathcal{M}$ su visais $c, d \in \mathbf{R}$, vadinamas tiesiniu erdvės $\mathbf{R}^k$ *poerdviu*. Suprantama, kad bet kuris tiesinis poerdvis yra tiesinė vektorinė erdvė.

Pavyzdžiui, aibė, susidedanti iš nulinio vektoriaus $\mathbf{0}$, arba visų vektorių aibė yra tiesiniai poerdviai. Sujungę visus galimus aibės $S = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_m\}$ vektorių tiesinius darinius
$$c_1\boldsymbol{x}_1 + ... + c_m\boldsymbol{x}_m, \quad c_1, ..., c_m \in \mathbf{R},$$
gausime tiesinį poerdvį $\mathcal{M}(S)$, generuotą vektorių aibės S.

1P.4 apibrėžimas. Sakome, kad vektoriai $\boldsymbol{x}_1, ..., \boldsymbol{x}_m$ yra *tiesiškai priklausomi*, jei egzistuoja skaliarai $c_1, ..., c_m \in \mathbf{R}$ ne visi lygūs 0, kad patenkinta sąlyga
$$c_1\boldsymbol{x}_1 + ... + c_m\boldsymbol{x}_m = \mathbf{0}. \tag{8.1.6}$$

Jeigu tokių skaliarų neegzistuoja, sakome, kad vektoriai $\boldsymbol{x}_1, ..., \boldsymbol{x}_m$ yra *tiesiškai nepriklausomi*. Iš šio apibrėžimo išplaukia tokios išvados:

1) Nulinis vektorius sudaro aibę tiesiškai priklausomų vektorių.

2) Bet kuri vektorių aibė, kuriai priklauso nulinis vektorius $\mathbf{0}$, yra tiesiškai priklausomų vektorių aibė.

3) Aibė nenulinių vektorių $\boldsymbol{x}_1, ..., \boldsymbol{x}_m$ yra tiesiškai priklausoma tada ir tik tada, kai kuris nors vektorius yra tiesinis kitų vektorių darinys.

1P.5 apibrėžimas. Tiesinės vektorinės erdvės $\mathcal{M}$ poaibis, kuris generuoja tiesinę erdvę $\mathcal{M}$, vadinamas tiesinės vektorinės erdvės $\mathcal{M}$ *baze*.

1) Kiekviena vektorinė erdvė turi bazę.

2) Jeigu $\boldsymbol{a}_1, ..., \boldsymbol{a}_m$ ir $\boldsymbol{b}_1, ..., \boldsymbol{b}_r$ yra dvi tos pačios erdvės bazės, tai $m = r$.

3) Kiekvienas erdvės $\mathcal{M}$ vektorius vieninteliu būdu išreiškiamas per jos bazę.

Pavyzdžiui, k-matės erdvės $\mathbf{R}^k$ elementai $\boldsymbol{e}_1^T = (1,\ 0, ..., 0)$, $\boldsymbol{e}_2^T = (0,\ 1, ..., 0)$, ..., $\boldsymbol{e}_k^T = (0,\ 0, ..., 1)$, sudaro bazę, nes jie yra tiesiškai nepriklausomi ir bet kuris vektorius $\boldsymbol{x}^T = (x_1, ..., x_k) \in \mathbf{R}^k$ išreiškiamas šiais vektoriais:
$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + ... + x_k\boldsymbol{e}_k.$$

Vektorių skaliarinė sandauga. Dviejų vienodos dimensijos vektorių $\boldsymbol{x}^T = (x_1, ..., x_k)$ ir $\boldsymbol{y}^T = (y_1, ..., y_k)$ *skaliarine sandauga* vadiname sumą
$$\boldsymbol{x}^T\boldsymbol{y} = \boldsymbol{y}^T\boldsymbol{x} = \sum_{j=1}^{k} x_j y_j. \tag{8.1.7}$$

Skaliarinė sandauga tenkina tokias sąlygas.

1) $\boldsymbol{x}^T\boldsymbol{x} = 0$ tada ir tik tada, kai $\boldsymbol{x} = \mathbf{0}$.
2) Patenkintos sąlygos

$$c(\boldsymbol{x}^T\boldsymbol{y}) = (c\boldsymbol{x}^T)\boldsymbol{y}, \quad (\boldsymbol{x}+\boldsymbol{y})^T\boldsymbol{z} = \boldsymbol{x}^T\boldsymbol{z} + \boldsymbol{y}^T\boldsymbol{z}. \tag{8.1.8}$$

3) Tenkinama Koši ir Švarco nelygybė

$$(\boldsymbol{x}^T\boldsymbol{y})^2 \leq (\boldsymbol{x}^T\boldsymbol{x})(\boldsymbol{y}^T\boldsymbol{y}). \tag{8.1.9}$$

Teigiama kvadratinė šaknis iš sandaugos $\boldsymbol{x}^T\boldsymbol{x}$ vadinama vektoriaus $\boldsymbol{x}$ *norma arba ilgiu*

$$||\boldsymbol{x}|| = +\sqrt{\boldsymbol{x}^T\boldsymbol{x}}. \tag{8.1.10}$$

Ji tenkina trikampio nelygybę

$$||\boldsymbol{x}-\boldsymbol{y}|| + ||\boldsymbol{y}-\boldsymbol{z}|| \geq ||\boldsymbol{x}-\boldsymbol{z}||. \tag{8.1.11}$$

Sakome, kad vektoriai $\boldsymbol{x}$ ir $\boldsymbol{y}$ yra *ortogonalūs*, jeigu skaliarinė sandauga

$$\boldsymbol{x}^T\boldsymbol{y} = \boldsymbol{y}^T\boldsymbol{x} = 0 \tag{8.1.12}$$

lygi nuliui. Bendriau, kampas θ tarp dviejų nenulinių vektorių apibrėžiamas lygybe

$$\cos\theta = \frac{\boldsymbol{x}^T\boldsymbol{y}}{||\boldsymbol{x}|| \cdot ||\boldsymbol{y}||}.$$

Nenuliniai poromis ortogonalūs vektoriai $\boldsymbol{x}_1, ..., \boldsymbol{x}_m$ yra tiesiškai nepriklausomi. Ortogonalūs vektoriai, generuojantys tiesinę erdvę $\mathcal{M}$, vadinami *ortogonalia* erdvės $\mathcal{M}$ *baze*, o jei jie turi vienetinius ilgius, tai – *ortonormuota baze*. Tiesinės erdvės ortonormuota bazė visada egzistuoja. Pavyzdžiui, vektoriai $\boldsymbol{e}_1^T = (1,\ 0, ..., 0)$, $\boldsymbol{e}_2^T = (0,\ 1, ..., 0)$, ..., $\boldsymbol{e}_k^T = (0,\ 0, ..., 1)$ sudaro erdvės $\mathbf{R}^k$ ortonormuotą bazę.

8.2. Matricos ir determinantai

Matrica $\boldsymbol{A}$ yra lentelė, užpildyta realiais skaičiais. Kiekvieną matricos elementą a_{ij} numeruosime dviem indeksais: indeksas i nurodo eilutės, o indeksas j stulpelio, kurių sankirtoje yra elementas a_{ij}, numerius. Žymėsime $\boldsymbol{A} = [a_{ij}]_{m\times n}$, čia m yra eilučių skaičius, o n – stulpelių skaičius. Matricos $\boldsymbol{A}$ stulpeliai yra dimensijos m vektoriai, o eilutės – dimensijos n transponuoti vektoriai.

Matricų sudėtis. Dviejų vienodos dimensijos $(m \times n)$ matricų $\boldsymbol{A} = [a_{ij}]_{m\times n}$ ir $\boldsymbol{B} = [b_{ij}]_{m\times n}$ suma yra matrica $\boldsymbol{A}+\boldsymbol{B} = [a_{ij}+b_{ij}]_{m\times n}$, kurios elementai gaunami sudedant atitinkamus matricų $\boldsymbol{A}$ ir $\boldsymbol{B}$ elementus. Šis veiksmas tenkina komutatyvumo ir distributyvumo savybes

$$\boldsymbol{A}+\boldsymbol{B} = \boldsymbol{B}+\boldsymbol{A}, \quad \boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C}) = (\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}.$$

Daugyba iš skaliaro. Matricos $\boldsymbol{A} = [a_{ij}]_{m\times n}$ sandauga iš skaliaro $c \in \mathbf{R}$ suprantama kaip matrica, kurios kiekvienas elementas padaugintas iš skaliaro c:

$$c\boldsymbol{A} = [ca_{ij}]_{m\times n}.$$

Šis veiksmas, akivaizdu, tenkina sąlygas

$$(c+d)\boldsymbol{A} = c\boldsymbol{A} + d\boldsymbol{A}, \quad c(\boldsymbol{A}+\boldsymbol{B}) = c\boldsymbol{A} + c\boldsymbol{B}.$$

Matricų sandauga. Matricų $\boldsymbol{A} = [a_{ij}]_{m\times n}$ ir $\boldsymbol{B} = [b_{ij}]_{n\times r}$ sandauga $\boldsymbol{AB}$ apibrėžta tik tada, kai matricos $\boldsymbol{A}$ stulpelių skaičius sutampa su matricos $\boldsymbol{B}$ eilučių skaičiumi. Daugybos rezultatas yra matrica $\boldsymbol{AB} = [c_{ij}]_{m\times r}$, kurios elementas c_{ij} gaunamas imant matricos $\boldsymbol{A}$ i-osios eilutės ir matricos $\boldsymbol{B}$ j-ojo stulpelio skaliarinę sandaugą:

$$c_{ij} = \sum_{s=1}^{n} a_{is}b_{sj}, \quad i = 1,...,m, \quad j = 1,...,r. \tag{8.2.1}$$

Matricos $\boldsymbol{AB}$ eilučių skaičius sutampa su matricos $\boldsymbol{A}$ eilučių skaičiumi, o stulpelių skaičius lygus matricos $\boldsymbol{B}$ stulpelių skaičiui.

Matricų daugyba netenkina komutatyvumo dėsnio, nes, pavyzdžiui, sandauga $\boldsymbol{AB}$ gali būti apibrėžta, o sandauga $\boldsymbol{BA}$ – neapibrėžta.

Asociatyvumo ir distributyvumo dėsniai yra tenkinami

$$\boldsymbol{ABC} = (\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC}), \quad \boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{BC},$$

$$(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{C}+\boldsymbol{D}) = \boldsymbol{A}(\boldsymbol{C}+\boldsymbol{D}) + \boldsymbol{B}(\boldsymbol{C}+\boldsymbol{D}),$$

jeigu matricų dimensijos yra tokios, kad visi daugybos veiksmai yra apibrėžti.

Nulinė ir vienetinė matricos. Nulinė matrica $\mathbf{0}$ yra tokia, kurios visi elementai lygūs 0. Dimensijos $(m \times m)$ kvadratinė matrica vadinama *vienetine*, jeigu jos visi diagonaliniai elementai lygūs 1, o visi kiti elementai lygūs 0. Žymėsime $\boldsymbol{I} = \boldsymbol{I}_m$. Akivaizdu, kad

$$\boldsymbol{A} + \mathbf{0} = \boldsymbol{A}, \quad \boldsymbol{AI} = \boldsymbol{IA} = \boldsymbol{A}. \tag{8.2.2}$$

Transponuota matrica. Matricą $\boldsymbol{A}^T$, kurią gauname iš matricos $\boldsymbol{A}$ pakeitę jos eilutes stulpeliais ir atvirkščiai, vadiname transponuota matrica, t. y.

$$\boldsymbol{A} = [a_{ij}]_{m\times n}, \quad \boldsymbol{A}^T = [a_{ji}]_{n\times m}.$$

Transponavimo operacija tenkina sąlygas

$$(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T, \quad (\boldsymbol{ABC}) = \boldsymbol{C}^T\boldsymbol{B}^T\boldsymbol{A}^T, ... \tag{8.2.3}$$

Matricos pėdsakas. Kvadratinės matricos $\boldsymbol{A} = [a_{ij}]_{m\times m}$ diagonalinių elementų sumą vadiname matricos *pėdsaku*. Žymėsime

$$Tr(\boldsymbol{A}) = \sum_{j=1}^{m} a_{jj}.$$

Pėdsakas tenkina sąlygas

$$Tr(\boldsymbol{A}+\boldsymbol{B}) = Tr(\boldsymbol{A}) + Tr(\boldsymbol{B}), \quad Tr(\boldsymbol{AB}) = Tr(\boldsymbol{BA}). \tag{8.2.4}$$

Blokinės matricos. Kartais matricą $\boldsymbol{A}$ yra patogu suskaidyti į blokus

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{B} & \boldsymbol{C} \\ \boldsymbol{D} & \mathbf{E} \end{pmatrix}.$$

Transponavimo operaciją galima atlikti taip:

$$\boldsymbol{A}^T = \begin{pmatrix} \boldsymbol{B}^T & \boldsymbol{D}^T \\ \boldsymbol{C}^T & \mathbf{E}^T \end{pmatrix}.$$

Blokinių matricų daugyba atliekama pagal įprastines matricų daugybos taisykles

$$\begin{pmatrix} \boldsymbol{B} & \boldsymbol{C} \\ \boldsymbol{D} & \mathbf{E} \end{pmatrix}\begin{pmatrix} \boldsymbol{G} \\ \boldsymbol{H} \end{pmatrix} = \begin{pmatrix} \boldsymbol{BG}+\boldsymbol{CH} \\ \boldsymbol{DG}+\mathbf{E}\boldsymbol{H} \end{pmatrix},$$

$$\begin{pmatrix} \boldsymbol{B} & \boldsymbol{C} \\ \boldsymbol{D} & \mathbf{E} \end{pmatrix}\begin{pmatrix} \boldsymbol{G} & \boldsymbol{J} \\ \boldsymbol{H} & \boldsymbol{L} \end{pmatrix} = \begin{pmatrix} \boldsymbol{BG}+\boldsymbol{CH} & \boldsymbol{BJ}+\boldsymbol{CL} \\ \boldsymbol{DG}+\mathbf{E}\boldsymbol{H} & \boldsymbol{DJ}+\mathbf{E}\boldsymbol{L} \end{pmatrix},$$

jeigu tik visi daugybos veiksmai yra apibrėžti.

Matricos rangas. Matricą $\boldsymbol{A} = [a_{ij}]_{m\times n}$ sudaro n dimensijos m vektorių (matricos stulpeliai) arba m dimensijos n transponuotų vektorių (matricos eilutės). Tiesiškai nepriklausomų stulpelių (arba eilučių) skaičius vadinamas *matricos rangu.* Žymėsime $Rang(\boldsymbol{A})$. Akivaizdu, kad

$$Rang(\boldsymbol{A}) = Rang(\boldsymbol{A}^T\boldsymbol{A}), \quad Rang(\boldsymbol{A}) \le \min(m,\ n) \tag{8.2.5}$$

$$Rang(\boldsymbol{AB}) \le \min(Rang(\boldsymbol{A}), Rang(\boldsymbol{B})), \quad Rang(\boldsymbol{A}+\boldsymbol{B}) \le Rang(\boldsymbol{A})+Rang(\boldsymbol{B}).$$

Jeigu matrica $\boldsymbol{A} = [a_{ij}]_{n\times n}$ yra *idempotentinė* , t. y. $\boldsymbol{AA} = \boldsymbol{A}$, tai

$$Rang(\boldsymbol{A}) + Rang(\boldsymbol{I}-\boldsymbol{A}) = n. \tag{8.2.6}$$

Atvirkštinė matrica. Kvadratinė matrica $\boldsymbol{A} = [a_{ij}]_{n\times n}$ vadinama *neišsigimusia*, jeigu jos rangas lygus n. Tokiu atveju egzistuoja vienintelė matrica $\boldsymbol{A}^{-1} = [a^{ij}]_{n\times n}$, vadinama *atvirkštine* matricai $\boldsymbol{A}$, kad patenkintos sąlygos

$$\boldsymbol{AA}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}. \tag{8.2.7}$$

Jeigu $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ vienodos dimensijos neišsigimusios matricos, tai

$$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}, \quad (\boldsymbol{ABC})^{-1} = \boldsymbol{C}^{-1}\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}, \quad (\boldsymbol{A}^T)^{-1} = (\boldsymbol{A}^{-1})^T. \tag{8.2.8}$$

Kvadratinė matrica $\boldsymbol{A}$ vadinama *ortogonalia*, jeigu $\boldsymbol{AA}^T = \boldsymbol{I}$. Tokiu atveju atvirkštinė matrica sutampa su transponuota $\boldsymbol{A}^T = \boldsymbol{A}^{-1}$ ir

$$\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{AA}^T = \boldsymbol{I}. \tag{8.2.9}$$

Matricos pertvarkymas. 1) Tarkime, $\boldsymbol{A}$ yra simetriška neišsigimusi dimensijos m matrica. Tada

a) egzistuoja neišsigimusi dimensijos m matrica $\boldsymbol{B}$, kad

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{B}^T; \tag{8.2.10}$$

b) egzistuoja neišsigimusi trikampė dimensijos m matrica $\boldsymbol{C}$, kad

$$\boldsymbol{A} = \boldsymbol{C}\boldsymbol{C}^T; \tag{8.2.11}$$

c) egzistuoja neišsigimusi dimensijos m matrica $\boldsymbol{B}$, kad

$$\boldsymbol{B}\boldsymbol{A}\boldsymbol{B}^T = \boldsymbol{I}; \tag{8.2.12}$$

d) egzistuoja ortogonali dimensijos m matrica $\boldsymbol{B}$, kad

$$\boldsymbol{B}\boldsymbol{A}\boldsymbol{B}^T = \boldsymbol{\Delta}, \quad \boldsymbol{A} = \boldsymbol{C}\boldsymbol{\Delta}\boldsymbol{C}^T, \quad \boldsymbol{C} = \boldsymbol{B}^{-1}, \tag{8.2.13}$$

čia $\boldsymbol{\Delta} = [\delta_{ij}]_{m\times m}$ turi diagonalinį pavidalą, t. y. $\delta_{ij} = 0$, kai $i \neq j$.

2) Tegu $\boldsymbol{A} = [a_{ij}]_{m\times m}$ idempotentinė rango $r \leq m$ matrica. Tada egzistuoja ortonormuoti vektoriai $\boldsymbol{B}_1, ..., \boldsymbol{B}_r$, kad

$$\boldsymbol{A} = \boldsymbol{B}_1\boldsymbol{B}_1^T + ... + \boldsymbol{B}_r\boldsymbol{B}_r^T. \tag{8.2.14}$$

3) Tegu $\boldsymbol{A}$ simetriška neišsigimusi dimensijos m matrica. Tada egzistuoja ortogonalūs vektoriai $\boldsymbol{B}_1, ..., \boldsymbol{B}_m$, kad su visais $\boldsymbol{t} \in \mathbf{R}^m$ kvadratinė forma $\boldsymbol{t}^T\boldsymbol{A}\boldsymbol{t}$ gali būti pertvarkyta į kvadratų sumą

$$\boldsymbol{t}^T\boldsymbol{A}\boldsymbol{t} = (\boldsymbol{B}_1^T\boldsymbol{t})^2 + ... + (\boldsymbol{B}_m^T\boldsymbol{t})^2. \tag{8.2.15}$$

Determinantas. Kvadratinės matricos $\boldsymbol{A} = [a_{ij}]_{m\times m}$ *determinantas* yra jos elementų a_{ij} skaliarinė funkcija:

$$|\boldsymbol{A}| = \sum \pm(a_{1i_1}a_{2i_2}...a_{mi_m});$$

sumavimas atliekamas pagal visus galimus skaičių $(1, 2, ..., m)$ perstatinius $(i_1, i_2, ..., i_m)$. Sandauga imama su teigiamu ženklu, jei perstatinis $(i_1, i_2, ..., i_m)$ gaunamas perstatant aibės $(1, 2, ..., m)$ elementus lyginį skaičių kartų ir su neigiamu ženklu priešingu atveju.

Pažymėkime $\boldsymbol{A}_{ij}$ elemento a_{ij} algebrinį papildinį. Jis lygus sandaugai $(-1)^{i+j}$ iš determinanto matricos, gaunamos išbraukus i-ąją eilutę ir j-ąjį stulpelį. Tada

$$|\boldsymbol{A}| = \sum_{i=1}^{m} a_{ri}\boldsymbol{A}_{ri}, \quad \text{su visais} \quad r = 1, ..., m; \tag{8.2.16}$$

$$|\boldsymbol{A}| = \sum_{r=1}^{m} a_{ri}\boldsymbol{A}_{ri}, \quad \text{su visais} \quad i = 1, ..., m. \tag{8.2.17}$$

Pateikiame dar keletą savybių.

1) $|\boldsymbol{A}| = 0$ tada ir tik tada, kai $Rang(\boldsymbol{A}) \neq m$.
2) Jeigu $|\boldsymbol{A}| \neq 0$, tai $\boldsymbol{A}^{-1} = (\boldsymbol{A}_{ij}/|\boldsymbol{A}|)^T$.
3) Jei $\boldsymbol{A}$ diagonalinė ar trikampė matrica, tai $|\boldsymbol{A}|$ lygus diagonalinių elementų sandaugai.
4) Jei $\boldsymbol{A}$ ir $\boldsymbol{B}$ vienodos dimensijos kvadratinės matricos, tai

$$|\boldsymbol{AB}| = |\boldsymbol{A}||\boldsymbol{B}|. \tag{8.2.18}$$

5) Jeigu blokinėje matricoje

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{B} & \boldsymbol{C} \\ \boldsymbol{D} & \mathbf{E} \end{pmatrix}.$$

determinantas $|\boldsymbol{B}| \neq 0$, tai

$$|\boldsymbol{A}| = |\boldsymbol{B}||\mathbf{E} - \boldsymbol{D}\boldsymbol{B}^{-1}\boldsymbol{C}|. \tag{8.2.19}$$

Tikrinės reikšmės ir tikriniai vektoriai. Tegu $\boldsymbol{A} = [a_{ij}]_{m\times m}$ kvadratinė simetriška matrica, t. y. $a_{ij} = a_{ji}$. Charakteringosios lygties

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = 0 \tag{8.2.20}$$

šaknys $\lambda_1, ..., \lambda_m$ vadinamos matricos $\boldsymbol{A}$ *tikrinėmis reikšmėmis*, o vektoriai $\boldsymbol{L}_1, ...,$ $\boldsymbol{L}_m$, tenkinantys sąlygas

$$\boldsymbol{A}\boldsymbol{L}_i = \lambda_i \boldsymbol{L}_i, \quad i = 1, ..., m, \tag{8.2.21}$$

vadinami matricos $\boldsymbol{A}$ *tikriniais vektoriais.*
1) Visos charakteringosios lygties šaknys yra realios neneigiamos, o tikriniai vektoriai gali būti parinkti realūs.
2) Jeigu matricos $\boldsymbol{A}$ rangas lygus r ir $r < m$, tai nulis yra charakteringosios lygties kartotinumo $m - r$ šaknis.
3) Skirtingas $\lambda_i \neq \lambda_j$ atitinkantys tikriniai vektoriai $\boldsymbol{L}_i$ ir $\boldsymbol{L}_j$ yra ortogonalūs $\boldsymbol{L}_i^T \boldsymbol{L}_j = 0$.
4) Jeigu matrica $\boldsymbol{A}$ teigiamai apibrėžta, tai visos tikrinės reikšmės teigiamos, o jei neneigiamai apibrėžta, tai – neneigiamos.
5) Egzistuoja ortogonali matrica $\boldsymbol{L}$, kad

$$\boldsymbol{L}^T \boldsymbol{A} \boldsymbol{L} = \boldsymbol{\Delta}, \quad \boldsymbol{A} = \boldsymbol{L}\boldsymbol{\Delta}\boldsymbol{L}^T; \tag{8.2.22}$$

čia $\boldsymbol{\Delta} = [\delta_{ij}]_{m\times m}$ diagonalinė matrica; $\delta_{ii} = \lambda_i, i = 1, ..., m$, $\delta_{ij} = 0$, kai $i \neq j$; $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_m \geq 0$ yra matricos $\boldsymbol{A}$ tikrinės reikšmės, o matricos $\boldsymbol{L}$ i-asis stulpelis yra tikrinis vektorius, atitinkantis tikrinę reikšmę λ_i. Detaliau išskleidę gauname

$$\begin{gathered} \boldsymbol{A} = \lambda_1 \boldsymbol{L}_1 \boldsymbol{L}_1^T + ... + \lambda_m \boldsymbol{L}_m \boldsymbol{L}_m^T, \\ \boldsymbol{I} = \boldsymbol{L}_1 \boldsymbol{L}_1^T + ... + \boldsymbol{L}_m \boldsymbol{L}_m^T, \end{gathered} \tag{8.2.23}$$

vadinamąjį *spektrinį* matricos $\boldsymbol{A}$ skleidinį.

6) Jeigu matrica $\boldsymbol{A}$ teigiamai apibrėžta, tai

$$|\boldsymbol{A}| = \prod_{i=1}^{m} \lambda_i, \quad Tr(\boldsymbol{A}) = \sum_{i=1}^{m} \lambda_i. \tag{8.2.24}$$

Tiesinės lygčių sistemos. Tiesinę m lygčių sistemą kintamųjų $x_1, ..., x_m$ atžvilgiu matricine forma galima užrašyti taip:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{8.2.25}$$

čia $\boldsymbol{A} = [a_{ij}]_{m \times m}$ žinomų koeficientų matrica, $\boldsymbol{b}^T = (b_1, ..., b_m)$ laisvasis narys, o $\boldsymbol{x}^T = (x_1, ..., x_m)$ nežinomas vektorius. Arba ekvivalenčia forma

$$x_1 \boldsymbol{a}_1 + ... + x_m \boldsymbol{a}_m = \boldsymbol{b}, \tag{8.2.26}$$

čia $\boldsymbol{a}_1, ..., \boldsymbol{a}_m$ yra matricos $\boldsymbol{A}$ stulpeliai.

1) Homogeninė lygčių sistema ($\boldsymbol{b} = \boldsymbol{0}$) turi nenulinį sprendinį tada ir tik tada, kai vektoriai $\boldsymbol{a}_1, ...\boldsymbol{a}_m$ yra tiesiškai priklausomi.

2) Nehomogeninė lygčių sistema ($\boldsymbol{b} \neq \boldsymbol{0}$) turi sprendinį tada ir tik tada, kai vektorius $\boldsymbol{b}$ gali būti išreikštas tiesiniu vektorių $\boldsymbol{a}_1, ..., \boldsymbol{a}_m$ dariniu. Bendras šios sistemos sprendinys yra suma kurio nors jos sprendinio ir bendrojo homogeninės sistemos sprendinio.

3) Nehomogeninė lygčių sistema turi vienintelį sprendinį, kai vektoriai $\boldsymbol{a}_1, ...,$ $\boldsymbol{a}_m$ yra tiesiškai nepriklausomi, t. y. kai $Rang(\boldsymbol{A}) = m$. Sprendinys turi tokį pavidalą

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}. \tag{8.2.27}$$

Šiuo atveju homogeninė lygčių sistema turi tik trivialų sprendinį $\boldsymbol{x} = \boldsymbol{0}$.

4) Tegu $\boldsymbol{A} = [a_{ij}]_{m \times n}$ dimensijos $(m \times n)$ matrica. *Apibedrintąja atvirkštine* matricai $\boldsymbol{A}$ vadiname tokią matricą $\boldsymbol{A}^- = [a_{ij}^-]_{n \times m}$, kad vektorius $\boldsymbol{x} = \boldsymbol{A}^-\boldsymbol{b}$ yra suderintos lygčių sistemos $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ sprendinys.

a) apibendrintoji atvirkštinė matrica egzistuoja ir tenkina sąlygą $\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{A} = \boldsymbol{A}$;

b) suderintos lygčių sistemos $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ sprendinys turi tokį pavidalą

$$\boldsymbol{x} = \boldsymbol{A}^-\boldsymbol{b} + (\boldsymbol{A}^-\boldsymbol{A} - \boldsymbol{I})\boldsymbol{z},$$

čia $\boldsymbol{z} \in \mathbf{R}^n$ bet koks vektorius;

c) apibendrintąją atvirkštinę matricą galima surasti taip. Jeigu matricos $\boldsymbol{A}$ rangas lygus r, $r < \min(m,\ n)$, tai egzistuoja dimensijos $(r \times r)$ neišsigimęs matricos $\boldsymbol{A}$ blokas $\boldsymbol{B}$. Perstatant eilutes ir stulpelius matricą $\boldsymbol{A}$ galima užrašyti tokiu pavidalu:

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{B} & \boldsymbol{C} \\ \boldsymbol{D} & \mathbf{E} \end{pmatrix}.$$

Tada apibedrintąja atvirkštine galima imti matricą

$$\boldsymbol{A}^- = \begin{pmatrix} \boldsymbol{B}^{-1} & \boldsymbol{K} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix}, \tag{8.2.28}$$

kai matrica $\boldsymbol{K}$ tenkina sąlygas

$$\boldsymbol{B}\boldsymbol{K} = \boldsymbol{C}, \quad \boldsymbol{D}\boldsymbol{K} = \mathbf{E}.$$

9 skyrius

2 priedas. Atsitiktiniai vektoriai

9.1. Atsitiktinio vektoriaus skirstinys

Tarkime, turime tikimybinę erdvę $(\Omega, \mathcal{F}, P)$.

2P.1 apibrėžimas. Realią vienareikšmę k-matę funkciją $\mathbf{X} = \mathbf{X}(\omega) = (X_1(\omega), ..., X_k(\omega))^T$, apibrėžtą aibėje Ω ir tokią, kad su visais $\boldsymbol{x} = (x_1, ..., x_k)^T \in \mathbf{R}^k$

$$\{\omega : X_1(\omega) \leq x_1, ..., X_k(\omega) \leq x_k\} \in \mathcal{F},$$

vadiname k-mačiu *atsitiktiniu vektoriumi* (a. v.).

1) Bet koks a. v. vienareikšmiškai nusakomas jo pasiskirstymo funkcijos

$$F(\boldsymbol{x}) = F(x_1, ..., x_k) = \mathbf{P}\{X_1 \leq x_1, ..., X_k \leq x_k\}, \tag{9.1.1}$$

apibrėžtos su visais $\boldsymbol{x} = (x_1, ..., x_k)^T \in \mathbf{R}^k$.

2) Jeigu a. v. $\mathbf{X}$ įgyjamų reikšmių aibė yra baigtinė arba skaiti, tai tokio a. v. skirstinys vadinamas *diskrečiuoju*. Jo skirstinys visiškaiai nusakomas išvardijant galimas reikšmes $\boldsymbol{x}_1, \boldsymbol{x}_2, ... \in \mathbf{R}^k$ ir jų įgijimo tikimybes

$$p_i = \{\mathbf{X} = \boldsymbol{x}_i\}, \quad i = 1, 2, ...; \quad \sum_i p_i = 1. \tag{9.1.2}$$

3) Absoliučiai tolydžiojo a. v. $\mathbf{X}$ skirstinys visiškai nusakomas jo *tankio funkcija*

$$f(\boldsymbol{x}) = f(x_1, ..., x_k) = \frac{\partial^k}{\partial x_1 ... \partial x_k} F(x_1, ..., x_k). \tag{9.1.3}$$

4) Bet kokio a. v. $\mathbf{X}$ skirstinį visiškai nusako jo *charakteristinė funkcija*

$$\varphi_{\boldsymbol{X}}(\boldsymbol{t}) = \varphi_{\boldsymbol{X}}(t_1, ..., t_k) = \mathbf{E}e^{i\boldsymbol{t}^T\boldsymbol{X}} = \mathbf{E}e^{i(t_1X_1 + ... + t_kX_k)}, \tag{9.1.4}$$

nusakyta su visais $\boldsymbol{t} = (t_1, ..., t_k)^T \in \mathbf{R}^k$. Pateiksime keletą charakteristinės funkcijos savybių.

a) A. v. $\boldsymbol{Y} = \boldsymbol{a} + \boldsymbol{B}\mathbf{X}$ ($\boldsymbol{a}$ fiksuotas dimensijos k vektorius, $\boldsymbol{B}$ fiksuota dimensijos k kvadratinė matrica) charakteristinė funkcija

$$\varphi_{\boldsymbol{Y}}(\boldsymbol{t}) = e^{i\boldsymbol{t}^T\boldsymbol{a}}\varphi_{\boldsymbol{X}}(\boldsymbol{t}^T\boldsymbol{B}_1, ..., \boldsymbol{t}^T\boldsymbol{B}_k), \tag{9.1.5}$$

čia $\boldsymbol{B}_1, ..., \boldsymbol{B}_k$ yra matricos $\boldsymbol{B}$ stulpeliai.

b) Nepriklausomų a. v. $\mathbf{X}_1, ..., \mathbf{X}_n$ sumos $\boldsymbol{S}_n = \mathbf{X}_1 + ... + \mathbf{X}_n$ charakteristinė funkcija lygi dėmenų charakteristinių funkcijų sandaugai

$$\varphi_{\boldsymbol{S}_n}(\boldsymbol{t}) = \prod_{j=1}^{n} \varphi_{\boldsymbol{X}_j}(\boldsymbol{t}). \tag{9.1.6}$$

c) Remiantis charakteristinės funkcijos apibrėžimu įrodoma Kramero ir Voldo teorema. A. v. $\mathbf{X}$ tikimybinis skirstinys nusakytas tada ir tik tada, kai nusakyti skirstiniai vienmačių a. d.

$$Y_{\boldsymbol{L}} = \boldsymbol{L}^T\mathbf{X} = L_1X_1 + ... + L_kX_k$$

su visais vektoriais $\boldsymbol{L} = (L_1, ..., L_k)^T \in \mathbf{R}^k$.

9.2. Marginalieji ir sąlyginiai skirstiniai

Nagrinėsime $(k+s)$-matį absoliučiai tolydųjį a. v. $\mathbf{X} = (\boldsymbol{Y}^T, \boldsymbol{Z}^T)^T = (Y_1, ..., Y_k, Z_1, ..., Z_s)^T$. Vektorių $\mathbf{X}, \boldsymbol{Y}$ ir $\boldsymbol{Z}$ pasiskirstymo funkcijas žymėsime $F(\boldsymbol{y}, \boldsymbol{z}) = F(y_1, ..., y_k, z_1, ..., z_s)$, $G(\boldsymbol{y}) = G(y_1, ..., y_k)$ ir $H(\boldsymbol{z}) = H(z_1, ..., z_s)$, o tankių funkcijas $f(\boldsymbol{y}, \boldsymbol{z}), g(\boldsymbol{y})$ ir $h(\boldsymbol{z})$.

Marginalieji skirstiniai. Vektoriaus, sudaryto iš bet kurių pradinio vektoriaus koordinačių, skirstinys vadinamas *marginaliuoju* skirstiniu pradinio vektoriaus atžvilgiu. Pavyzdžiui, a. v. $\boldsymbol{Y}$ skirstinys vadinamas marginaliuoju jungtinio vektoriaus $\mathbf{X} = (\boldsymbol{Y}^T, \boldsymbol{Z}^T)^T$ atžvilgiu.

1) Vektoriaus $\boldsymbol{Y}$ pasiskirstymo funkcija gaunama iš a. v. $\mathbf{X}$ pasiskirstymo funkcijos įrašant $+\infty$ vietoje argumentų, atitinkančių likusias pradinio vektoriaus koordinates

$$G(\boldsymbol{y}) = G(y_1, ..., y_k) = F(y_1, ..., y_k, +\infty, ..., +\infty). \tag{9.2.1}$$

2) Marginaliojo skirstinio tankis gaunamas integruojant:

$$g(\boldsymbol{y}) = \int_{\mathbf{R}^s} f(\boldsymbol{y}, \boldsymbol{z})d\boldsymbol{z}. \tag{9.2.2}$$

3) A. v. $\boldsymbol{Y}$ charakteristinė funkcija

$$\varphi_{\boldsymbol{Y}}(\boldsymbol{t}) = \varphi_{\boldsymbol{Y}}(t_1, ..., t_k) = \mathbf{E}e^{i\boldsymbol{t}^T\boldsymbol{Y}} = \mathbf{E}e^{i(t_1Y_1 + ... + t_kY_k)}$$

gaunama iš jungtinio vektoriaus $\mathbf{X} = (\boldsymbol{Y}^T, \boldsymbol{Z}^T)^T$ charakteristinės funkcijos

$$\varphi_{\boldsymbol{X}}(\boldsymbol{t}, \boldsymbol{\theta}) = \varphi_{\boldsymbol{X}}(t_1, ..., t_k, \theta_1, ..., \theta_s) = \mathbf{E}e^{i(\boldsymbol{t}^T\boldsymbol{Y} + \boldsymbol{\theta}^T\boldsymbol{Z})}$$

įrašant vietoje $\boldsymbol{\theta}$ nulinį vektorių $\mathbf{0}$:

$$\varphi_{\boldsymbol{Y}}(\boldsymbol{t}) = \varphi_{\boldsymbol{X}}(\boldsymbol{t}, \mathbf{0}) = \varphi_{\boldsymbol{X}}(t_1, ...t_k, 0, ..., 0). \tag{9.2.3}$$

Sąlyginiai skirstiniai. Tarkime, kad a. v. $\boldsymbol{Z} = \boldsymbol{z}$ yra fiksuotas. Jeigu $h(\boldsymbol{z}) \neq 0$, tai a. v. $\boldsymbol{Y}$ sąlyginio skirstinio tankis

$$g(\boldsymbol{y}|\boldsymbol{z}) = \frac{f(\boldsymbol{y}, \boldsymbol{z})}{h(\boldsymbol{z})}. \tag{9.2.4}$$

Turėdami a. v. $\boldsymbol{Y}$ sąlyginį tankį, kai $\boldsymbol{Z} = \boldsymbol{z}$ fiksuotas, ir besąlyginį a. v. $\boldsymbol{Z}$ tankį, galima atkurti a. v. $\boldsymbol{Y}$ besąlyginį tankį:

$$g(\boldsymbol{y}) = \int_{\mathbf{R}^s} g(\boldsymbol{y}|\boldsymbol{z})h(\boldsymbol{z})d\boldsymbol{z}. \tag{9.2.5}$$

Ši formulė apibendrina pilnosios tikimybės formulę. Analogiškai apibendrinamos ir Bejeso formulės

$$h(\boldsymbol{z}|\boldsymbol{y}) = \frac{g(\boldsymbol{y}|\boldsymbol{z})h(\boldsymbol{z})}{\int_{\mathbf{R}^s} g(\boldsymbol{y}|\boldsymbol{z})h(\boldsymbol{z})d\boldsymbol{z}}. \tag{9.2.6}$$

Nepriklausomi vektoriai. Jeigu kiekvieno iš dviejų vektorių sąlyginiai skirstiniai, kai kito vektoriaus reikšmės fiksuotos, nepriklauso nuo fiksuotųjų reikšmių ir sutampa su besąlyginiais skirstiniais, tokie vektoriai vadinami nepriklausomais.

Suformuluosime keletą nepriklausomumo kriterijų. Atsitiktiniai vektoriai $\boldsymbol{Y}$ ir $\boldsymbol{Z}$ nepriklausomi tada ir tik tada, kai

1) pasiskirstymo funkcija $F(\boldsymbol{y}, \boldsymbol{z})$ yra lygi marginaliųjų pasiskirstymo funkcijų sandaugai

$$F(\boldsymbol{y}, \boldsymbol{z}) = G(\boldsymbol{y})H(\boldsymbol{z}), \quad \boldsymbol{y} \in \mathbf{R}^k, \quad \boldsymbol{z} \in \mathbf{R}^s; \tag{9.2.7}$$

2) tankio funkcija $f(\boldsymbol{y}, \boldsymbol{z})$ yra lygi marginaliųjų tankio funkcijų sandaugai

$$f(\boldsymbol{y}, \boldsymbol{z}) = g(\boldsymbol{y})h(\boldsymbol{z}), \quad \boldsymbol{y} \in \mathbf{R}^k, \quad \boldsymbol{z} \in \mathbf{R}^s; \tag{9.2.8}$$

3) charakteristinė funkcija $\varphi_{\boldsymbol{X}}(\boldsymbol{t}, \boldsymbol{\theta})$ yra lygi marginaliųjų charakteristinių funkcijų sandaugai

$$\varphi_{\boldsymbol{X}}(\boldsymbol{t}, \boldsymbol{\theta}) = \varphi_{\boldsymbol{Y}}(\boldsymbol{t})\varphi_{\boldsymbol{Z}}(\boldsymbol{\theta}), \quad \boldsymbol{t} \in \mathbf{R}^k, \quad \boldsymbol{\theta} \in \mathbf{R}^s. \tag{9.2.9}$$

Analogiškai formuluojami ir didesnio skaičiaus atsitiktinių vektorių nepriklausomumo kriterijai.

9.3. Atsitiktinių vektorių funkcijos

Tarkime, $\mathbf{X} = (X_1, ..., X_k)^T$ yra absoliučiai tolydusis a. v., kurio pasiskirstymo funkcija $F(\boldsymbol{x}) = F(x_1, ..., x_k)$ ir tankis $f(\boldsymbol{x}) = f(x_1, ..., x_k)$. Nagrinėsime a. v. $\boldsymbol{Y} = (Y_1, ..., Y_k)^T$, $Y_i = h_i(X_1, ..., X_k)$, $i = 1, ..., k$, tikimybinį skirstinį.

1) Tarkime, kad atliekama transformacija yra abipus vienareikšmė ir egzistuoja atvirkštinė transformacija $x_i = h_i^{-1}(y_1, ..., y_k), i = 1, ..., k$. Jeigu funkcijų $h_i^{-1}, i = 1, ..., k$, dalinės išvestinės pagal visus argumentus yra tolydžios ir jakobianas

$$|\boldsymbol{J}| = \left|\frac{\mathcal{D}(x_1, ..., x_k)}{\mathcal{D}(y_1, ..., y_k)}\right|$$

nelygus nuliui, tai a. v. $\boldsymbol{Y}$ yra absoliučiai tolydusis ir jo tankio funkcija

$$g(\boldsymbol{y}) = g(y_1, ..., y_k) = f(h_1^{-1}(y_1, ..., y_k), ..., h_k^{-1}(y_1, ..., y_k))|\boldsymbol{J}|. \tag{9.3.1}$$

2) Tegu atliekama transformacija nėra abipus vienareikšmė, tačiau a. v. $\mathbf{X}$ įgyjamų reikšmių sritį galima taip suskirstyti į n nesikertančių sričių $D_1, ..., D_n$, kad srities D_i transformacija į ją atitinkančią vektoriaus $\boldsymbol{Y}$ įgyjamų reikšmių sritį D_i^* būtų abipus vienareikšmė. Jeigu egzistuoja atvirkštinės transformacijos $x_r = h_{r_i}^{-1}(y_1, ..., y_k)$, $r = 1, ..., k$, iš srities D_i^* į sritį D_i, funkcijų $h_{r_i}^{-1}$ dalinės išvestinės pagal kiekvieną argumentą yra tolydžios, o atitinkami jakobianai $\boldsymbol{J}_i$ srityse D_i^*, $i = 1, ..., n$, nelygūs nuliui, tai a. v. $\boldsymbol{Y}$ yra absoliučiai tolydusis ir jo tankio funkcija

$$g(\boldsymbol{y}) = g(y_1, ..., y_k) = \sum_{i=1}^{n} \boldsymbol{I}_{D_i^*}(y_1, ..., y_k) f_i(x_1, ..., x_k)|\boldsymbol{J}_i|; \tag{9.3.2}$$

čia $\boldsymbol{I}_{D_i^*}(y_1, ..., y_k)$ yra aibės D_i^* indikatorius, o indeksas prie tankio funkcijos $f(x_1, ..., x_k)$ nurodo, kad jos argumentus x_r reikia pakeisti į $h_{r_i}^{-1}(y_1, ..., y_k)$.

Atsitiktinio vektoriaus momentai

1) Atsitiktinio vektoriaus $\mathbf{X} = (X_1, ..., X_k)^T$ *vidurkis* yra vektorius, sudarytas iš jo koordinačių vidurkių

$$\mathbf{E}(\mathbf{X}) = (\mathbf{E}X_1, ..., \mathbf{E}X_k)^T = (\mu_1, ..., \mu_k)^T = \boldsymbol{\mu}. \tag{9.3.3}$$

2) Atsitiktinio vektoriaus $\mathbf{X} = (X_1, ..., X_k)^T$ antrųjų mišriųjų momentų matricą vadiname *kovariacijų matrica*

$$\boldsymbol{V}(\mathbf{X}) = \boldsymbol{\Sigma} = [\sigma_{ij}]_{k\times k}, \tag{9.3.4}$$

čia

$$\sigma_{ij} = \mathbf{Cov}\,(X_i, X_j) = \mathbf{E}[(X_i - \mathbf{E}X_i)(X_j - \mathbf{E}X_j)] = \mathbf{E}(X_iX_j) - \mathbf{E}X_i\mathbf{E}X_j,$$

$$\sigma_{ii} = \mathbf{Cov}\,(X_i, X_i) = \boldsymbol{V}X_i, \quad i, j = 1, ..., k,$$

arba trumpiau

$$\boldsymbol{V}(\mathbf{X}) = \mathbf{E}[(\mathbf{X} - \mathbf{E}(\mathbf{X}))(\mathbf{X} - \mathbf{E}(\mathbf{X}))^T] = \mathbf{E}(\mathbf{X}\mathbf{X}^T) - \mathbf{E}(\mathbf{X})(\mathbf{E}(\mathbf{X}))^T.$$

Naudojant kovariacijas sudaroma *koreliacijos koeficientų* matrica

$$\boldsymbol{R} = [\rho_{ij}]_{k\times k}, \quad \rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}, \quad i,j = 1,...,k. \tag{9.3.5}$$

3) Tarkime, $\boldsymbol{A} = [a_{ij}]_{m\times k}$ yra matrica su pastoviais koeficientais, o $\mathbf{X} = (X_1,...,X_k)^T$ k-matis vektorius. Tada $\boldsymbol{Y} = \boldsymbol{A}\mathbf{X}$ yra m-matis atsitiktinis vektorius. Vektoriaus $\boldsymbol{Y}$ pirmieji momentai

$$\mathbf{E}(\boldsymbol{Y}) = \boldsymbol{A}\boldsymbol{\mu}, \quad \boldsymbol{V}(\boldsymbol{Y}) = \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^T = [\gamma_{ij}]_{m\times m}. \tag{9.3.6}$$

Kvadratinės formos

$$Q = \mathbf{X}^T\boldsymbol{A}\mathbf{X}$$

vidurkis

$$\mathbf{E}Q = Tr(\boldsymbol{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}^T\boldsymbol{A}\boldsymbol{\mu}. \tag{9.3.7}$$

4) Tarkime, $\mathbf{X}_1,...,\mathbf{X}_n$ yra vienodai pasiskirstę nepriklausomi atsitiktiniai vektoriai, kurių $\mathbf{E}(\mathbf{X}_j) = \boldsymbol{\mu}$ ir $\boldsymbol{V}(\mathbf{X}_j) = \boldsymbol{\Sigma}$. Tada centruotos ir normuotos sumos

$$\bar{\boldsymbol{S}}_n = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}(\mathbf{X}_j - \boldsymbol{\mu})$$

pirmieji momentai

$$\mathbf{E}(\bar{\boldsymbol{S}}_n) = \mathbf{0}, \quad \boldsymbol{V}(\bar{\boldsymbol{S}}_n) = \boldsymbol{\Sigma}. \tag{9.3.8}$$

Remdamiesi charakteristinių funkcijų savybe (9.1.6), Kramero ir Voldo teorema ir vienmate centrine ribine teorema (CRT) gauname, pavyzdžiui, tokį paprasčiausią daugiamatės CRT variantą. Jeigu $|\boldsymbol{\Sigma}| > 0$ ir $n \to \infty$, tai

$$\bar{\boldsymbol{S}}_n \xrightarrow{d} \boldsymbol{Y} \sim N_k(\mathbf{0}, \boldsymbol{\Sigma}), \quad \bar{\boldsymbol{S}}_n\boldsymbol{\Sigma}^{-1}\bar{\boldsymbol{S}}_n \xrightarrow{d} \chi^2_k. \tag{9.3.9}$$

Jeigu $\boldsymbol{\Sigma}$ yra rango $r \le k$ idempotentinė matrica, tai

$$\bar{\boldsymbol{S}}_n\boldsymbol{\Sigma}^{-}\bar{\boldsymbol{S}}_n \xrightarrow{d} \chi^2_r. \tag{9.3.10}$$

10 skyrius

Daugiamačio normaliojo skirstinio savybės

Pateiksime keletą daugiamačio normaliojo skirstinio savybių. Jų įrodymus galima rasti, pavyzdžiui, knygoje [14].

Yra keletas daugiamačio normaliojo skirstinio apibrėžimų. Vienas iš jų grindžiamas Kramero ir Voldo teorema.

3P.1 apibrėžimas. Atsitiktinio vektoriaus $\mathbf{X} = (X_1, ..., X_k)^T$ skirstinys yra k-matis normalusis su vidurkių vektoriumi $\boldsymbol{\mu} = \mathbf{E}(\mathbf{X})$ ir kovariacijų matrica $\boldsymbol{V}(\mathbf{X}) = \boldsymbol{\Sigma}$, jeigu bet kurios tiesinės formos $\boldsymbol{L}^T\mathbf{X} = L_1X_1 + ... + L_kX_k$ skirstinys yra vienmatis normalusis su visais $\boldsymbol{L} = (L_1, ..., L_k)^T \in \mathbf{R}^k$. Remiantis (9.3.6) šios tiesinės formos vidurkis ir dispersija yra $\mathbf{E}(\boldsymbol{L}^T\mathbf{X}) = \boldsymbol{L}^T\boldsymbol{\mu}$, $\boldsymbol{V}(\boldsymbol{L}^T\mathbf{X}) = \boldsymbol{L}^T\boldsymbol{\Sigma}\boldsymbol{L}$.

1 savybė. Atsitiktinio vektoriaus $\mathbf{X}$ charakteristinė funkcija

$$\varphi(\boldsymbol{t}) = \mathbf{E}(\exp(i\boldsymbol{t}^T\mathbf{X})) = \exp\left(i\boldsymbol{t}^T\boldsymbol{\mu} - \frac{\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}}{2}\right). \tag{10.0.1}$$

Taigi a. v. $\mathbf{X}$ skirstinį vienareikšmiškai nusako vidurkių vektorius $\boldsymbol{\mu}$ ir kovariacijų matrica $\boldsymbol{\Sigma}$. Sutrumpintai žymėsime $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

2 savybė. Tegu $\mathbf{X}^{(1)}$ ir $\mathbf{X}^{(2)}$ yra vektoriai, sudaryti iš skirtingų vektoriaus $\mathbf{X}$ koordinačių. Tada a. v. $\mathbf{X}^{(1)}$ ir $\mathbf{X}^{(2)}$ yra nepriklausomi tada ir tik tada, kai jų kovariacijų matrica $\mathbf{Cov}\,(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}) = \mathbf{0}$. Šis tvirtinimas yra teisingas ir dėl didesnio skaičiaus vektorių, sudarytų iš skirtingų vektoriaus $\mathbf{X}$ koordinačių. Pavyzdžiui, a. d. $X_1, ..., X_k$ yra nepriklausomi tada ir tik tada, kai kovariacijų matrica $\boldsymbol{\Sigma}$ yra diagonalinė.

3 savybė. Jeigu $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, tai bet kurio vektoriaus, sudaryto iš $s < k$ skirtingų vektoriaus $\mathbf{X}$ koordinačių, skirstinys yra s-matis normalusis.

4 savybė. Jeigu vektorius $\boldsymbol{Y} = (Y_1, ..., Y_r)^T$ gautas atlikus tiesinę vektoriaus $\mathbf{X}$ transformaciją $Y = \boldsymbol{C}\mathbf{X}$; čia $\boldsymbol{C} = [C_{ij}]_{r\times k}$ yra matrica, turinti r eilučių ir k stulpelių, tai

$$\boldsymbol{Y} \sim N_r(\boldsymbol{C\mu}, \boldsymbol{C}^T\boldsymbol{\Sigma}\boldsymbol{C}). \tag{10.0.2}$$

5 savybė. Tegu $\mathbf{X}_1, ..., \mathbf{X}_n$ yra nepriklausomi a. v. ir $\mathbf{X}_i \sim N_k(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), i = 1, ..., n$, o $c_1, ..., c_n$ – konstantos. Tada

$$\boldsymbol{Y} = c_1\mathbf{X}_1 + ... + c_n\mathbf{X}_n \sim N_k\left(\sum_{i=1}^{n} c_i\boldsymbol{\mu}_i, \sum_{i=1}^{n} c_i^2\boldsymbol{\Sigma}_i\right). \tag{10.0.3}$$

Kai $c_1 = ... = c_n = 1/n$ ir a. v. $\mathbf{X}_1, ..., \mathbf{X}_n$ vienodai pasiskirstę $\mathbf{X}_i \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, gauname

$$\bar{\mathbf{X}} = \frac{1}{n}\sum_{i-1}^{n} \mathbf{X}_i \sim N_k(\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{\Sigma}). \tag{10.0.4}$$

6 savybė. A. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tada ir tik tada, kai teisingas dėstinys

$$\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{B}\boldsymbol{Z}, \quad \boldsymbol{B}\boldsymbol{B}^T = \boldsymbol{\Sigma}; \tag{10.0.5}$$

čia $\boldsymbol{B}$ matrica, turinti k eilučių ir m stulpelių, $\boldsymbol{Z} = (Z_1, ..., Z_m)^T$ yra vektorius, kurio koordinatės nepriklausomi a. d. ir $Z_i \sim N(0,\ 1), i = 1, ..., m$. Taigi normalųjį vektorių galime gauti atlikę a. v. $\boldsymbol{Z} = (Z_1, ..., Z_m)^T \sim N_m(\mathbf{0}, \boldsymbol{I})$ tiesinę transformaciją. Šią savybę galima panaudoti kaip kitą normaliojo a. v. apibrėžimą.

3P.2 apibrėžimas. Atsitiktinis vektorius $\mathbf{X}$ vadinamas k-mačiu normaliuoju, jeigu jis tenkina (10.0.5) lygybę.

7 savybė. Jeigu 6 savybėje $m = k$ ir $Rang(\boldsymbol{B}) = Rang(\boldsymbol{\Sigma}) = k$, tai atlikę atvirkštinę transformaciją

$$\boldsymbol{Z} = \boldsymbol{B}^{-1}(\mathbf{X} - \boldsymbol{\mu})$$

remdamiesi (9.3.1) iš a. v. $\boldsymbol{Z} \sim N_k(\mathbf{0}, \boldsymbol{I})$ tankio

$$\varphi(\boldsymbol{z}|\mathbf{0}, \boldsymbol{I}) = (2\pi)^{-k/2}\exp\{-\frac{1}{2}\boldsymbol{z}^T\boldsymbol{z}\}$$

gauname a. v. $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tankio funkciją (pakeitimo jakobianas yra $1/|\boldsymbol{B}| = 1/|\boldsymbol{\Sigma}|^{1/2}$)

$$\varphi(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = (2\pi)^{-k/2}|\boldsymbol{\Sigma}|^{-1/2}\exp\{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\}. \tag{10.0.6}$$

Kadangi $\boldsymbol{Z}^T\boldsymbol{Z} \sim \chi^2(k)$, tai

$$\boldsymbol{Z}^T\boldsymbol{Z} = (\mathbf{X} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu}) \sim \chi^2(k). \tag{10.0.7}$$

Šioje lygybėje vietoje $\boldsymbol{\mu}$ įrašę kitą vektorių $\boldsymbol{\nu}$, gautume

$$(\mathbf{X}-\boldsymbol{\nu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\nu}) \sim \chi^2(k;\delta) \tag{10.0.8}$$

necentrinį χ^2 skirstinį su k laisvės laipsnių ir necentriškumo parametru

$$\delta = (\boldsymbol{\nu}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{\nu}-\boldsymbol{\mu}).$$

8 savybė. Tegu $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Kvadratinė forma

$$Q = (\mathbf{X}-\boldsymbol{\mu})^T\boldsymbol{A}(\mathbf{X}-\boldsymbol{\mu}) \sim \chi^2(m) \tag{10.0.9}$$

tada ir tik tada, kai

$$\boldsymbol{\Sigma}(\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}-\boldsymbol{A})\boldsymbol{\Sigma} = \mathbf{0};$$

čia $m = Tr(\boldsymbol{A}\boldsymbol{\Sigma})$. Jeigu $Rang(\boldsymbol{\Sigma}) = k$, tai būtina ir pakankama sąlyga yra paprastesnė: $\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A} = \boldsymbol{A}$.

9 savybė. Tegu $\boldsymbol{Z} \sim N_k(\mathbf{0}, \boldsymbol{I})$. Apibrėžkime

$$Y = \boldsymbol{b}^T\boldsymbol{Z}, \quad Q_i = \boldsymbol{Z}^T\boldsymbol{A}_i\boldsymbol{Z}, \quad i = 1, ..., m, \tag{10.0.10}$$

tiesinę ir kvadratines formas; čia $\boldsymbol{A}_i$ — kvadratinės simetriškos matricos ir $Rang(\boldsymbol{A}_i) = k_i$. Tada:

a) jeigu $\boldsymbol{b}^T\boldsymbol{A}_i = \mathbf{0}$, tai Y ir Q_i nepriklausomi;

b) jeigu $\boldsymbol{A}_j^T\boldsymbol{A}_i = \mathbf{0}$, tai Q_j ir Q_i nepriklausomos;

c) jeigu $\boldsymbol{Z}^T\boldsymbol{Z} = Q_1 + ... + Q_m$, tai sąlyga $k_1 + ... + k_m = k$ yra būtina ir pakankama, kad kvadratinės formos $Q_1, ..., Q_m$ būtų nepriklausomos ir turėtų χ^2 skirstinius $Q_j \sim \chi^2(k_j)$, $j = 1, ..., m$.

10 savybė. Tegu $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, o $\mathbf{X}^{(1)}$ ir $\mathbf{X}^{(2)}$ yra r-matis ir $(k-r)$-matis vektoriai, sudaryti iš skirtingų a. v. $\mathbf{X}$ koordinačių. Pažymėkime $\boldsymbol{\mu}^{(1)} = \mathbf{E}(\mathbf{X}^{(1)})$, $\boldsymbol{\mu}^{(2)} = \mathbf{E}(\mathbf{X}^{(2)})$, $\boldsymbol{\Sigma}_{11} = \boldsymbol{V}(\mathbf{X}^{(1)})$, $\boldsymbol{\Sigma}_{22} = \boldsymbol{V}(\mathbf{X}^{(2)})$, $\boldsymbol{\Sigma}_{12} = \mathbf{Cov}\,(\mathbf{X}^{(1)}, \mathbf{X}^{(2)})$, $\boldsymbol{\Sigma}_{21} = \boldsymbol{\Sigma}_{12}^T$. Tarkime, $|\boldsymbol{\Sigma}_{11}| > 0$ Tada sąlyginis a. v. $\mathbf{X}^{(2)}$ skirstinys, kai a. v. $\mathbf{X}^{(1)} = x^{(1)}$ yra fiksuotas yra $(k-r)$-matis normalusis su parametrais

$$\mathbf{E}(\mathbf{X}^{(2)}|\mathbf{X}^{(1)} = x^{(1)}) = \boldsymbol{\mu}^{(2)} + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{x}^{(1)} - \boldsymbol{\mu}^{(1)}),$$
$$\boldsymbol{V}(\mathbf{X}^{(2)}|\mathbf{X}^{(1)} = x^{(1)}) = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}. \tag{10.0.11}$$

Kai $\mathbf{X}^{(2)} = X_k$ yra vienmatis, gauname

$$\begin{aligned}\mathbf{E}(X_k|\mathbf{X}^{(1)} = \boldsymbol{x}^{(1)}) &= \mu_k + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{x}^{(1)} - \boldsymbol{\mu}^{(1)})\\ &= \alpha + \beta_1 x_1 + ... + \beta_{k-1}x_{k-1},\\ \boldsymbol{V}(X_k|\mathbf{X}^{(1)} = \boldsymbol{x}^{(1)}) &= \sigma_{kk} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} = \frac{1}{\sigma^{kk}};\end{aligned} \tag{10.0.12}$$

čia σ^{kk} — matricos $\boldsymbol{\Sigma}^{-1}$ paskutinis kampinis elementas. Taigi bet kurios a. v. $\mathbf{X}$ koordinatės *regresija* kitų koordinačių atžvilgiu yra tiesinė.

Literatūra

1. Abramowitz M., Stegun I. Handbook of mathematical functions. Vertimas į rusų kalbą.– Maskva: Nauka, 1979.

2. Anderson T. W. An introduction to Multivariate Statistical Analysis. New York – John Wiley. Vertimas į rusų kalbą.– Maskva: Fizmatgiz, 1963.

3. Čekanavičius V., Murauskas G. Statistika ir jos taikymai. Vilnius: TEV, II dalis – 2002.

4. Duda R. O., Hart P. E. Pattern Classsification and Scene Analysis. John Wiley and Sons, 1973. Vertimas į rusų kalbą.– Maskva: Mir, 1976.

5. DeGroot M. H. Optimal Statistical Decisions. McGraw-Hill Company, 1970. Vertimas į rusų kalbą.– Maskva: Mir, 1974

6. Gorsuch R. L. Factor Analysis. 2nd Edition. Lawrence Erlbaum Assoc. Inc., 1983.

7. Harman H. H. Modern Factor Analysis. USA: 1976.

8. Fisher R. A. The Use of Multiple Measurements in Taxonomic Problems. Annals of Eugenics, 7(1936).

9. Jonson R. A., Wichern D. W. Applied Multivariate Statistical Analysis. 5th edition. Prentice Hall Inc., 2002.

10. Kendall M. G., Stuart A. The advanced Theory of Statistics. Volume 3. Design and Analysis and Time-Series. second edition. London: Charles Griffin Company Limited, ...Vertimas į rusų kalbą.– Maskva: Fizmatgiz,1976.

11. Kruopis J., Vaišvila A., Kalnius R. Mechatronikos gaminių kokybė. Atrankinė kontrolė. Vilnius: VU leidykla, 2005.

12. Lehmann E.L. Testing statistical hypotheses. Vertimas į rusų kalbą. – Maskva: Nauka, 1979.

13. Mardia K. V., Kent J. T., Bibby J. M. Multivariate Analysis. 2nd Edition. New York: Academic Press, 1979.

14. Rao C. R. Linear statistical inference and its applications. Vertimas į rusų kalbą. – Maskva: Nauka, 1968.

15. Raudys Š. Statistical and Neural Classifiers. An Integrated Approach to Desiggn. Springer, 2001.

16. SAS. Help and Documentation. Kompaktinė plokštelė (platinama su SAS sistema).

Dalykinė rodyklė

Vilijandas Bagdonavičius, Julius Jonas Kruopis
Matematinė statistika: vadovėlis
Ketvirta dalis. Daugiamatė statistika. – Vilnius: Vilniaus universitetas, 2015. – 230 p.
ISBN 978–609–459–518–9

Šoje vadovėlio dalyje matematinės statistikos uždaviniai sprendžiami tariant, kad imties elementai yra atsitiktiniai vektoriai. Apsiribojama daugiamačio normaliojo skirstinio atveju, kai rezultatai ir metodai yra įgavę labiausiai užbaigtą pavidalą. Pateikiami vidurkių vektoriaus ir kovariacijų matricos nepaslinktieji įvertiniai ir jų savybės. Sudaryti kriterijai hipotezėms apie vidurkių vektoriaus ir kovariacijų matricos reikšmes ir daugiamačiai dispersinės ir regresinės analizės analogai. Nagrinėjami klasifikavimo ir atsitiktinių vektorių nepriklausomumo hipotezių tikrinimo uždaviniai. Aptariama stebimo vektoriaus dimensijos sumažinimo problema įvedant kanoninius kintamuosius.

519.2(075.8)

Vlijandas Bagdonavičius, Julius Jonas Kruopis
Matematinė statistika. IV dalis. Daugiamatė statistika
Vadovėlis

Lietuvių kalbos redaktorė *Danutė Petrauskienė*
Maketuotoja *Rūta Levulienė*

Išleido *Vilniaus universiteto leidykla*