Asimptotinė statistika

Vytautas Kazakevičius

December 14, 2007

Turinys

1 Asimptotiniai skirstiniai **1**
1.1 Atsitiktiniai dydžiai ir vektoriai 1
1.1.1 Diskretieji atsitiktiniai dydžiai 1
1.1.2 Diskretieji atsitiktiniai vektoriai 2
1.1.3 Tolydieji atsitiktiniai dydžiai 3
1.1.4 Tolydieji atsitiktiniai vektoriai 5
1.2 Asimptotiniai skirstiniai 9
1.2.1 Uždavinio formulavimas 9
1.2.2 Empiriniai vidurkiai 9
1.2.3 Normalusis skirstinys 10
1.2.4 Pagrindiniai asimptotinių skirstinių ieškojimo metodai . . 10
1.2.5 Pavyzdžiai 10

2 Stochastinis konvergavimas **17**
2.1 Apibrėžimas ir pavyzdžiai 17
2.1.1 Apibrėžimas 17
2.1.2 Pavyzdžiai 18
2.1.3 Kaip įrodyti, kad ribos nėra 20
2.2 Portmanto lema 21
2.2.1 Formulavimas ir įrodymas 21
2.2.2 Įrodymo paaiškinimai 22
2.2.3 Komentarai 26
2.3 Konvergavimas į išsigimusį dydį 27
2.3.1 Teoremos 27
2.3.2 Įrodymų paaiškinimai 28
2.3.3 Komentarai 29
2.4 Tolydaus atvaizdžio principas 29
2.4.1 Teorema 29
2.4.2 Įrodymo paaiškinimai 31
2.4.3 Komentarai 32
2.5 O-simbolika 33

3 Asimptotiniai skirstiniai (tęsinys) **35**
3.1 Vektorinis delta metodas 35
3.1.1 Matricinė notacija 35
3.1.2 Atsitiktinių matricų vidurkiai 36
3.1.3 Normalusis skirstinys 37
3.1.4 Empiriniai vidurkiai 37
3.1.5 Asimptotikų ieškojimo metodai 37

3.1.6 Pavyzdžiai . . . 38
3.2 Ypatingi atvejai . . . 41
3.2.1 Atvejis, kai $g'(a) = 0$ . . . 41
3.2.2 Atvejis, kai g priklauso nuo n . . . 42
3.3 Pavyzdžiai . . . 44

4 Taikymai **49**
4.1 Pasikliautinieji intervalai . . . 49
4.1.1 Uždavinio formulavimas . . . 49
4.1.2 Standartinė procedūra . . . 49
4.1.3 Modifikuota procedūra . . . 50
4.1.4 Dispersijos stabilizavimo metodas . . . 52
4.2 Hipotezių tikrinimas . . . 53
4.2.1 Bendra teorija . . . 53
4.2.2 Nepriklausomumo tikrinimas, remiamtis empiriniu koreliacijos koeficientu . . . 54
4.3 Klaidingos modelio specifikacijos efektas . . . 57
4.3.1 Hipotezės apie dispersiją tikrinimas . . . 57

5 M-įvertiniai **61**
5.1 Maksimalaus tikėtinumo įvertiniai . . . 61
5.2 M-įvertinių pagrįstumas. Klasikinė teorema . . . 65
5.2.1 M-įvertiniai ir Z-įvertiniai . . . 65
5.2.2 Klasikinė teorema . . . 66
5.2.3 Įrodymų paaiškinimai . . . 67
5.2.4 Komentarai . . . 69
5.3 Valdo teorema . . . 69
5.3.1 Valdo teorema . . . 69
5.3.2 Įrodymo paaiškinimai . . . 70
5.4 Kitos teoremos . . . 72
5.4.1 Teoremos . . . 72
5.4.2 Įrodymų paaiškinimai . . . 73

1 skyrius

Asimptotiniai skirstiniai

1.1 Atsitiktiniai dydžiai ir vektoriai

1.1.1 Diskretieji atsitiktiniai dydžiai

Diskretaus atsitiktinio dydžio skirstinys paprastai užrašomas lentele

X	x_1	$\dots$	x_r
	p_1	$\dots$	p_r

Toks užrašas reiškia, kad atsitiktinis dydis įgyja reikšmes $x_1, \dots, x_r$ ir

$$\mathrm{P}\{X = x_j\} = p_j \quad \text{su } j = 1, \dots, r.$$

Visos tikimybės p_j yra neneigiami skaičiai ir $\sum_{j=1}^{r} p_j = 1$.

Pagrindinės tokio atsitiktinio dydžio charakteristikos skaičiuojamos taip:

$$\mathrm{P}\{X \in A\} = \sum_{x_i \in A} p_i, \quad \mathrm{E}\, f(X) = \sum_{i=1}^{r} f(x_i) p_i, \quad \mathrm{D}\, f(X) = \mathrm{E}\, f^2(X) - \big[\mathrm{E}\, f(X)\big]^2.$$

Be to, reikia mokėti rasti bet kokios transformacijos $Y = \varphi(X)$ skirstinį. Galimos Y reikšmės yra $\varphi(x_1), \dots, \varphi(x_r)$ (kai kurie iš tų skaičių gali sutapti) ir su bet kokia galima reikšme y

$$\mathrm{P}\{Y = y\} = \sum_{\varphi(x_i) = y} p_i.$$

1.1 PAVYZDYS. Atsitiktinio dydžio X skirstinys užrašytas lentele

X	0	1	2	3
	0.3	0.4	0.1	0.2

Apskaičiuokite:

a) $\mathrm{P}\{(X-1)(X-2) \neq 0\}$;

b) $\mathrm{E} \cos \dfrac{\pi X}{4}$;

c) atsitiktinio dydžio $Y = |X - 1|$ skirstinį.

Sprendimas.

$$\mathrm{P}\{(X-1)(X-2) \neq 0\} = 0.3 + 0.2 = 0.5;$$
$$\mathrm{E}\cos\frac{\pi X}{4} = 1 \cdot 0.3 + \frac{\sqrt{2}}{2} \cdot 0.4 + 0 - \frac{\sqrt{2}}{2} \cdot 0.2 = 0.3 + 0.1\sqrt{2}.$$

Y	0	1	2
	0.4	0.4	0.2

1.1.2 Diskretieji atsitiktiniai vektoriai

Jei X ir Y yra du diskretūs atsitiktiniai dydžiai, tai jų *bendras skirstinys*, arba *atsitiktinio vektoriaus* (X, Y) *skirstinys*, aprašomas lentele

$X\ Y$	y_1	$\ldots$	y_s
x_1	p_{11}	$\ldots$	p_{1s}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_r	p_{r1}	$\ldots$	p_{rs}

Toks užrašas reiškia, kad X įgyja reikšmes $x_1, \ldots, x_r$, Y įgyja reikšmes $y_1, \ldots, y_s$ ir

$$\mathrm{P}\{X = x_j, Y = y_k\} = p_{jk} \quad \text{su } j = 1, \ldots, r \text{ ir } k = 1, \ldots, s.$$

Visos tikimybės p_{jk} yra neneigiami skaičiai ir $\sum_{j,k} p_{jk} = 1$.

Pagrindinės (X, Y) vektoriaus charakteristikos skaičiuojamos taip:

$$\mathrm{P}\{(X,Y) \in A\} = \sum_{(x_j, y_k) \in A} p_{jk}, \quad \mathrm{E}\, f(X,Y) = \sum_{j,k} f(x_j, y_k) p_{jk}.$$

Be to, reikia mokėti rasti bet kokios transformacijos $Z = \varphi(X, Y)$ skirstinį. Galimos Z reikšmės yra $\varphi(x_j, y_k)$ (kai kurie iš tų skaičių gali sutapti) ir su bet kokia galima reikšme z

$$\mathrm{P}\{Z = z\} = \sum_{\varphi(x_j, y_k) = z} p_{jk}.$$

Jei X ir Y nepriklausomi, užtenka nurodyti tik X ir Y skirstinius:

X	x_1	$\ldots$	x_r
	p_1	$\ldots$	p_r

Y	y_1	$\ldots$	y_s
	q_1	$\ldots$	q_s

Bendro skirstinio tikimybės tada skaičiuojamos pagal formulę

$$p_{jk} = p_j q_k, \quad j = 1, \ldots, r, \ k = 1, \ldots, s.$$

1.2 PAVYZDYS. X ir Y yra nepriklausomi atsitiktiniai dydžiai, pasiskirstę taip, kaip parodyta lentelėse:

X	0	1
	0.3	0.7

Y	0	1	2
	0.2	0.3	0.5

Apskaičiuokite:

a) $\mathrm{P}\{Y = X\}$;

b) $\mathrm{D}(XY)$;

c) atsitiktinio dydžio $Z = XY$ skirstinį.

Sprendimas. Bendras X ir Y skirstinys aprašomas lentele:

X Y	0	1	2	
0	0.06	0.09	0.15	0.3
1	0.14	0.21	0.35	0.7
	0.2	0.3	0.5	

Tada
$$\mathrm{P}\{Y = X\} = 0.06 + 0.21 = 0.27.$$

Be to,
$$\mathrm{D}(XY) = \mathrm{E}(X^2Y^2) - \big[\mathrm{E}(XY)\big]^2.$$

Kadangi
$$\mathrm{E}(XY) = 0 + 0 + 0 + 0 + 0.21 + 0.7 = 0.91,$$
$$\mathrm{E}(X^2Y^2) = 0 + 0 + 0 + 0 + 0.21 + 1.4 = 1.61,$$

gaunu
$$\mathrm{D}(XY) = 1.61 - 0.91^2 = 0.7819.$$

Z skirstinys aprašomas lentele

Z	0	1	2
	0.44	0.21	0.35

1.1.3 Tolydieji atsitiktiniai dydžiai

(Absoliučiai) tolydaus atsitiktinio dydžio X skirstinys aprašomas, nurodant jo tankį $p(x)$, kuris yra neneigiama funkcija, apibrėžta kokioje nors aibėje $D \subset \mathbb{R}$ ir tokia, kad
$$\int_D p(x)\mathrm{d}x = 1.$$

Pavyzdžiui, *tolygiojo* $(a; b)$ *intervale* skirstinio tankis
$$p(x) = \frac{1}{b-a}, \quad \text{kai } a < x < b.$$

Pagrindinės atsitiktinio dydžio X charakteristikos skaičiuojamos taip:
$$\mathrm{P}\{X \in A\} = \int_{A\cap D} p(x)\mathrm{d}x, \qquad \mathrm{E}\, f(X) = \int_D f(x)p(x)\mathrm{d}x.$$

Be to, reikia mokėti aprašyti bet kokios transformacijos $Y = \varphi(X)$ skirstinį. Jei φ yra
$$\varphi(x) = \begin{cases} y_1, & \text{kai } x \in D_1; \\ \vdots & \\ y_s, & \text{kai } y \in D_s; \end{cases}$$

pavidalo (čia $(D_1, \ldots, D_s)$ yra D aibės skaidinys), tai Y yra diskretus atsitiktinis dydis, įgyjantis reikšmes $y_1, \ldots, y_s$ su tikimybėmis

$$\mathrm{P}\{Y = y_k\} = \int_{D_k} p(x)\mathrm{d}x, \quad k = 1, \ldots, s.$$

Jei φ yra glodi funkcija, tai Y turi tankį $q(y)$, kuris skaičiuojamas taip:

$$q(y) = \big(G(y)\big)';$$
$$G(y) = \mathrm{P}\{Y \leqslant y\} = \mathrm{P}\{\varphi(X) \leqslant y\} = \int_{\substack{x \in D \\ \varphi(x) \leqslant y}} p(x)\mathrm{d}x.$$

1.3 PAVYZDYS. Atsitiktinis dydis X turi tankį

$$p(x) = x\mathrm{e}^{-x}, \quad \text{kai } x > 0.$$

Apskaičiuokite:

a) $\mathrm{P}\{1 < X < 2\}$;

b) $\mathrm{E}\,X^3$;

c) atsitiktinių dydžių

$$Y = \begin{cases} 0, & \text{kai } X < 1; \\ 1, & \text{kai } X \geqslant 1; \end{cases}$$

ir $Z = \ln X$ skirstinius.

Sprendimas.

$$\begin{aligned}
\mathrm{P}\{1 < X < 2\} &= \int_1^2 x\mathrm{e}^{-x}\mathrm{d}x \\
&= -\int_1^2 x\mathrm{d}\mathrm{e}^{-x} \\
&= -x\mathrm{e}^{-x}\Big|_1^2 + \int_1^2 \mathrm{e}^{-x}\mathrm{d}x \\
&= -2\mathrm{e}^{-2} + \mathrm{e}^{-1} - \mathrm{e}^{-x}\Big|_1^2 \\
&= -2\mathrm{e}^{-2} + \mathrm{e}^{-1} - \mathrm{e}^{-2} + \mathrm{e}^{-1} \\
&= -3\mathrm{e}^{-2} + 2\mathrm{e}^{-1} \\
&= \frac{2\mathrm{e} - 3}{\mathrm{e}^2}.
\end{aligned}$$

$$\mathrm{E}(X^3) = \int_0^\infty x^4\mathrm{e}^{-x}\mathrm{d}x = \Gamma(5) = 4! = 24.$$

Kadangi

$$\mathrm{P}\{Y = 1\} = \mathrm{P}\{X \geqslant 1\} = \int_1^\infty x\mathrm{e}^{-x}\mathrm{d}x = -\int_1^\infty x\mathrm{d}\mathrm{e}^{-x} = -x\mathrm{e}^{-x}\Big|_1^\infty = \mathrm{e}^{-1},$$

Y skirstinys aprašomas lentele:

Y	0	1
	$1-\mathrm{e}^{-1}$	e^{-1}

Kadangi

$$\begin{aligned}
\mathrm{P}\{Z \leqslant z\} &= \mathrm{P}\{\ln X \leqslant z\} \\
&= \mathrm{P}\{X \leqslant \mathrm{e}^z\} \\
&= \int_0^{\mathrm{e}^z} x\mathrm{e}^{-x}\mathrm{d}x \\
&= -\int_0^{\mathrm{e}^z} x\mathrm{d}\mathrm{e}^{-x} \\
&= -x\mathrm{e}^{-x}\Big|_0^{\mathrm{e}^z} + \int_0^{\mathrm{e}^z} \mathrm{e}^{-x}\mathrm{d}x \\
&= -\mathrm{e}^z\mathrm{e}^{-\mathrm{e}^z} - \mathrm{e}^{-x}\Big|_0^{\mathrm{e}^z} \\
&= 1-\mathrm{e}^{-\mathrm{e}^z} - \mathrm{e}^z\mathrm{e}^{-\mathrm{e}^z}
\end{aligned}$$

ir

$$(1-\mathrm{e}^{-\mathrm{e}^z} - \mathrm{e}^z\mathrm{e}^{-\mathrm{e}^z})' = \mathrm{e}^{-\mathrm{e}^z}\mathrm{e}^z - \mathrm{e}^z\mathrm{e}^{-\mathrm{e}^z} + \mathrm{e}^{2z}\mathrm{e}^{-\mathrm{e}^z} = \mathrm{e}^{2z}\mathrm{e}^{-\mathrm{e}^z},$$

atsitiktinis dydis Z turi tankį

$$q(z) = \mathrm{e}^{2z}\mathrm{e}^{-\mathrm{e}^z}, \quad -\infty < z < \infty.$$

1.1.4 Tolydieji atsitiktiniai vektoriai

Tolydžiojo atsitiktinio vektoriaus (X,Y) skirstinys aprašomas, nurodant tankį $p(x,y)$, kuris yra neneigiama funkcija, apibrėžta tam tikroje srityje $D \subset \mathbb{R}^2$ ir tenkina sąlygą

$$\int_D p(x,y)\mathrm{d}x\mathrm{d}y = 1.$$

Pagrindinės (X,Y) vektoriaus charakteristikos skaičiuojamos taip:

$$\mathrm{P}\{(X,Y)\in A\} = \int_{D\cap A} p(x,y)\mathrm{d}x\mathrm{d}y, \quad \mathrm{E}\, f(X,Y) = \int_D f(x,y)p(x,y)\mathrm{d}x\mathrm{d}y.$$

Be to, reikia mokėti rasti bet kokios transformacijos $Z=\varphi(X,Y)$ skirstinį. Jei φ yra

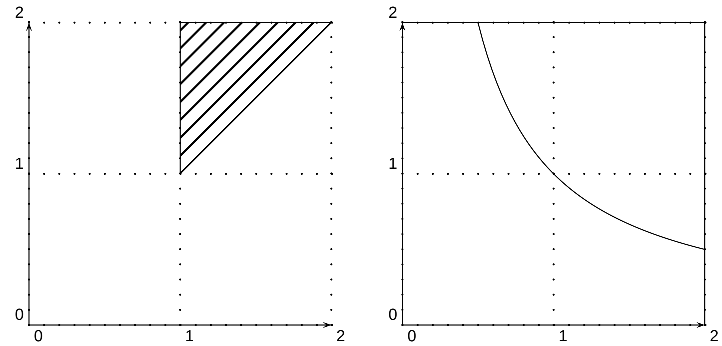
$$\varphi(x,y) = \begin{cases} z_1, & \text{kai } (x,y)\in D_1; \\ \vdots \\ z_s, & \text{kai } (x,y)\in D_s; \end{cases}$$

pavidalo (čia $(D_1,\dots,D_s)$ yra D aibės skaidinys), tai Z yra diskretus atsitiktinis dydis, įgyjantis reikšmes $z_1,\dots,z_s$ su tikimybėmis

$$\mathrm{P}\{Z=z_k\} = \int_{D_k} p(x,y)\mathrm{d}x\mathrm{d}y, \quad k=1,\dots,s.$$

Jei φ yra glodi funkcija, tai Z turi tankį $q(z)$, kuris skaičiuojamas taip:

$$q(z) = \big(G(z)\big)';$$



(a) (b)

1.1 pav.

$$G(z) = \mathrm{P}\{Z \leqslant z\} = \mathrm{P}\{\varphi(X,Y) \leqslant z\} = \int_{\substack{(x,y)\in D\\ \varphi(x,y)\leqslant z}} p(x,y)\mathrm{d}x\mathrm{d}y.$$

Jei X ir Y nepriklausomi, užtenka nurodyti tik X ir Y tankius $p(x)$ ir $q(y)$. Bendro skirstinio tankis tada yra marginalių tankių sandauga:

$$p(x,y) = p(x)q(y).$$

1.4 PAVYZDYS. Atsitiktiniai dydžiai X ir Y nepriklausomi ir tolygiai pasiskirstę $(0;2)$ intervale. Apskaičiuokite:

a) $\mathrm{P}\{Y \geqslant X \geqslant 1\}$;

b) $\mathrm{E}|XY-1|$;

c) atsitiktinio dydžio $Z = X+Y$ skirstinį.

Sprendimas. Bendras X ir Y tankis yra

$$p(x,y) = \frac{1}{4}, \quad \text{kai } 0 < x,y < 2.$$

Todėl (žr. 1.1 pav.)

$$\begin{aligned}
\mathrm{P}\{Y \geqslant X \geqslant 1\} &= \int\limits_{\substack{0<x<2\\ 0<y<2\\ y\geqslant x\geqslant 1}} \frac{1}{4}\mathrm{d}x\mathrm{d}y\\
&= \frac{1}{4}\int_1^2 \mathrm{d}x \int_x^2 \mathrm{d}y\\
&= \frac{1}{4}\int_1^2 (2-x)\mathrm{d}x\\
&= \frac{1}{4}\Big(2x - \frac{x^2}{2}\Big)\Big|_1^2\\
&= \frac{1}{4}\Big(4-2-2+\frac{1}{2}\Big)
\end{aligned}$$

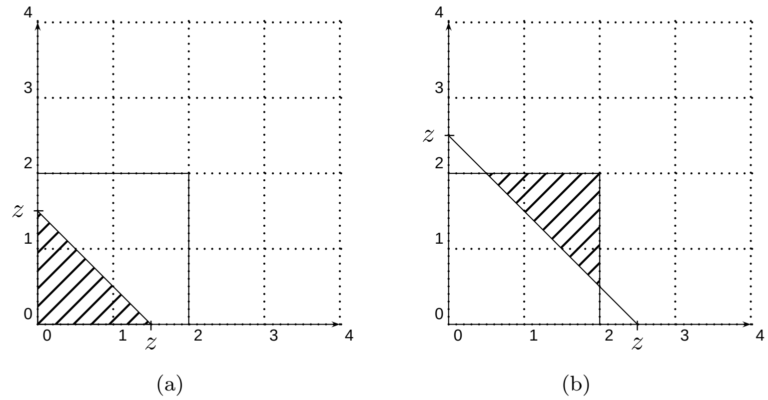
$$= \frac{1}{8},$$

o

$$\begin{aligned}
\mathrm{E}|XY-1| &= \int\limits_{\substack{0<x<2\\0<y<2}} |xy-1|\frac{1}{4}\mathrm{d}x\mathrm{d}y\\
&= \frac{1}{4}\int\limits_{\substack{0<x<2\\0<y<2\\xy>1}} (xy-1)\mathrm{d}x\mathrm{d}y + \frac{1}{4}\int\limits_{\substack{0<x<2\\0<y<2\\xy<1}} (1-xy)\mathrm{d}x\mathrm{d}y\\
&= \frac{1}{4}\int_{1/2}^{2}\mathrm{d}x\int_{1/x}^{2}(xy-1)\mathrm{d}y + \frac{1}{4}\int_{0}^{1/2}\mathrm{d}x\int_{0}^{2}(1-xy)\mathrm{d}y\\
&\qquad + \frac{1}{4}\int_{1/2}^{2}\mathrm{d}x\int_{0}^{1/x}(1-xy)\mathrm{d}y\\
&= \frac{1}{4}\int_{1/2}^{2}\Big(x\frac{y^2}{2}-y\Big)\Big|_{1/x}^{2}\mathrm{d}x + \frac{1}{4}\int_{0}^{1/2}\Big(y-x\frac{y^2}{2}\Big)\Big|_{0}^{2}\mathrm{d}x\\
&\qquad + \frac{1}{4}\int_{1/2}^{2}\Big(y-x\frac{y^2}{2}\Big)\Big|_{0}^{1/x}\mathrm{d}x\\
&= \frac{1}{4}\int_{1/2}^{2}\Big(2x-2-\frac{1}{2x}+\frac{1}{x}\Big)\mathrm{d}x + \frac{1}{4}\int_{0}^{1/2}(2-2x)\mathrm{d}x\\
&\qquad + \frac{1}{4}\int_{1/2}^{2}\Big(\frac{1}{x}-\frac{1}{2x}\Big)\mathrm{d}x\\
&= \frac{1}{4}\int_{1/2}^{2}\Big(2x-2+\frac{1}{2x}\Big)\mathrm{d}x + \frac{1}{4}(2x-x^2)\Big|_{0}^{1/2} + \frac{1}{4}\int_{1/2}^{2}\frac{\mathrm{d}x}{2x}\\
&= \frac{1}{4}\int_{1/2}^{2}\Big(2x-2+\frac{1}{x}\Big)\mathrm{d}x + \frac{1}{4}\Big(1-\frac{1}{4}\Big)\\
&= \frac{1}{4}(x^2-2x+\ln x)\Big|_{1/2}^{2} + \frac{3}{16}\\
&= \frac{1}{4}\Big(4-4+\ln 2-\frac{1}{4}+1-\ln\frac{1}{2}\Big) + \frac{3}{16}\\
&= \frac{3}{8}+\frac{\ln 2}{2}.
\end{aligned}$$

Kadangi $0 < X, Y < 2$, galimos sumos $Z = X + Y$ reikšmės užpildo $(0; 4)$ intervalą. Jei $0 < z < 2$, tai (žr. 1.2 pav.)

$$\begin{aligned}
\mathrm{P}\{Z \leqslant z\} &= \mathrm{P}\{X+Y \leqslant z\}\\
&= \int\limits_{\substack{0<x<2\\0<y<2\\x+y\leqslant z}} \frac{1}{4}\mathrm{d}x\mathrm{d}y\\
&= \frac{1}{4}\int_{0}^{z}\mathrm{d}x\int_{0}^{z-x}\mathrm{d}y\\
&= \frac{1}{4}\int_{0}^{z}(z-x)\mathrm{d}x
\end{aligned}$$



1.2 pav.

$$
\begin{aligned}
&= \frac{1}{4}\Big(zx - \frac{x^2}{2}\Big)\Big|_0^z \\
&= \frac{z^2}{8}.
\end{aligned}
$$

Jei $2 < z < 4$, tai

$$
\begin{aligned}
\mathrm{P}\{Z > z\} &= \mathrm{P}\{X + Y > z\} \\
&= \int\limits_{\substack{0<x<2\\0<y<2\\x+y>z}} \frac{1}{4}\mathrm{d}x\mathrm{d}y \\
&= \frac{1}{4}\int_{z-2}^{2}\mathrm{d}x\int_{z-x}^{2}\mathrm{d}y \\
&= \frac{1}{4}\int_{z-2}^{2}(2 - z + x)\mathrm{d}x \\
&= \frac{1}{4}\Big((2-z)x + \frac{x^2}{2}\Big)\Big|_{z-2}^{2} \\
&= \frac{1}{4}\Big(4 - 2z + 2 + (z-2)^2 - \frac{(z-2)^2}{2}\Big) \\
&= 2 - z + \frac{z^2}{8};
\end{aligned}
$$

todėl

$$
\mathrm{P}\{Z \leqslant z\} = -1 + z - \frac{z^2}{8}.
$$

Išdiferencijavęs gautas lygybes, gaunu, kad Z turi tankį

$$
p(z) = \begin{cases} z/4, & \text{kai } 0 < z < 2; \\ 1 - z/4, & \text{kai } 2 < z < 4. \end{cases}
$$

1.2 Asimptotiniai skirstiniai

1.2.1 Uždavinio formulavimas

Jei atsitiktinių dydžių seka T_n konverguoja pagal pasiskirstymą į atsitiktinį dydį T, rašau $T_n \to T$. Tikslų šio termino apibrėžimą duosiu kitame skyriuje. Ten taip pat įrodysiu ir daugumą tokio konvergavimo savybių. Dabar gi reikia išmokti dirbti su šia sąvoka.

Pagrindinis uždavinys, kurį reikia mokėti spręsti — rasti sekos T_n *asimptotinį skirstinį*. Šie žodžiai reiškia, kad reikia rasti tokias skaičių sekas a_n ir b_n, kad

$$b_n(T_n - a_n) \to T;$$

čia T — koks nors neišsigimęs atsitiktinis dydis (t.y. dydis, nelygus konstantai su tikimybe 1). Tokiu atveju galima sakyti, kad jei n didelis, tai T_n skirstinys maždaug sutampa su $a_n + b_n^{-1}T$ dydžio skirstiniu (kuris ir vadinamas asimptotiniu T_n skirstiniu).

Mano duodamuose uždaviniuose dažniausiai bus galima imti $a_n = a$ ir $b_n = \sqrt{n}$; čia a — tam tikra konstanta. Bet apskritai gali atsitikti visko. Pavyzdžiui, jei

$$\sqrt{n}(T_n - a) \to T,$$

o $U_n = nT_n$, tai

$$n^{-1/2}(U_n - na) \to T,$$

t.y. šiuo atveju $a_n = na$, o $b_n = n^{-1/2}$.

1.2.2 Empiriniai vidurkiai

Mano duodamuose uždaviniuose T_n bus

$$T_n = \varphi_n(X_1, \dots, X_n)$$

pavidalo; čia $X_1, \dots, X_n$ yra nepriklausomi vienodai pasiskirstę atsitiktiniai dydžiai, o φ_n — kokios nors gražios funkcijos. Jei X yra atsitiktinis dydis, pasiskirstęs taip pat, kaip kiekvienas X_i, dažnai sakysiu, kad $X_1, \dots, X_n$ yra *nepriklausomos X kopijos*, arba *imtis iš X skirstinio*. Atsitiktinius dydžius T_n paprastai vadinsiu *statistikomis*.

Labai svarbios yra

$$\overline{f(X)} = \frac{f(X_1) + \cdots + f(X_n)}{n}$$

pavidalo statistikos (empiriniai vidurkiai). Pavyzdžiui,

$$\overline{X} = \frac{X_1 + \cdots + X_n}{n}, \quad \overline{X^2} = \frac{X_1^2 + \cdots + X_n^2}{n}$$

ir pan.

Reikia žinoti tokias naudingas empirinių vidurkių savybes:

$$\overline{c} = c, \quad \overline{cX} = c\overline{X}, \quad \overline{X+Y} = \overline{X} + \overline{Y}.$$

Bet apskritai

$$\overline{XY} \neq \overline{X}\,\overline{Y}.$$

Pavyzdžiui, visada $\overline{X^2} \geqslant \left[\overline{X}\right]^2$ ir lygybė galima tik tada, kai visi X_i sutampa.

Jei f yra A aibės indikatorius 1_A, tai vietoje $1_A(X)$ aš mėgstu rašyti $1_{\{X\in A\}}$. Tokiu atveju

$$\overline{1_{\{X\in A\}}} = \frac{N}{n};$$

čia N yra A aibės elementų skaičius $X_1, \dots, X_n$ imtyje.

1.2.3 Normalusis skirstinys

Simboliu $N(\mu, \sigma^2)$ aš žymiu atsitiktinį dydį, turintį normalųjį skirstinį su vidurkiu μ ir dispersija σ^2. Gerai žinoma, kad jei X pasiskirstęs normaliai, tai $a+bX$ taip pat pasiskirstęs normaliai, tik su kitais parametrais: vidurkiu $a+b\mu$ ir dispersija $b^2\sigma^2$. Todėl aš dažnai rašau

$$a + bN(\mu, \sigma^2) = N(a + b\mu, b^2\sigma^2).$$

Pavyzdžiui,

$$3N(0, 0.4) = N(0, 3.6).$$

1.2.4 Pagrindiniai asimptotinių skirstinių ieškojimo metodai

Ieškodami asimptotinių statistikų skirstinių, naudojamės tokiais teiginiais:

- (didžiųjų skaičių dėsnis) jei $\mathrm{E}\,X = a$, tai
$$\overline{X} \to a;$$
- (centrinė ribinė teorema) jei $\mathrm{E}\,X = a$, $\mathrm{D}\,X = \sigma^2$, tai
$$\sqrt{n}(\overline{X} - a) \to N(0, \sigma^2);$$
- (tolydaus atvaizdžio principas) jei $T_n \to T$, $U_n^{(1)} \to u_1, \dots, U_n^{(k)} \to u_k$, tai
$$\varphi(T_n, U_n^{(1)}, \dots, U_n^{(k)}) \to \varphi(T, u_1, \dots, u_k);$$
- (delta metodas) jei $b_n \to \infty$ ir $b_n(T_n - a) \to T$, tai
$$b_n\big(g(T_n) - g(a)\big) \to g'(a)T.$$

Tikslūs paskutinių dviejų teiginių formulavimai ir įrodymai bus duoti kitame skyriuje.

1.2.5 Pavyzdžiai

1.5 PAVYZDYS. Atsitiktinio dydžio X skirstinys užrašytas lentele

X	-1	0	1	2
	0.2	0.2	0.2	0.4

o $X_1, \dots, X_n$ yra nepriklausomos X kopijos. Raskite nurodytų statistikų asimptotinį skirstinį:

a) $\dfrac{N^0}{N^+ - N^-}$;

b) $\dfrac{\sqrt{(N^+)^2 + n^2}}{N^+ + n}$;

c) $\dfrac{n\overline{X} - N^0}{N^+}$;

čia N^+, N^- ir N^0 yra, atitinkamai, teigiamų, neigiamų ir lygių nuliui narių skaičius imtyje $X_1, \ldots, X_n$.

Sprendimas. a. Pažymiu

$$T_n = \frac{N^0}{N^+ - N^-} = \frac{N^0/n}{N^+/n - N^-/n} = \frac{\overline{1_{\{X=0\}}}}{\overline{1_{\{X>0\}}} - \overline{1_{\{X<0\}}}}.$$

Pagal didžiųjų skaičių dėsnį

$$\begin{aligned}
\overline{1_{\{X=0\}}} &\to \mathrm{E}\,1_{\{X=0\}} = 0.2;\\
\overline{1_{\{X>0\}}} &\to \mathrm{E}\,1_{\{X>0\}} = 0.2 + 0.4 = 0.6;\\
\overline{1_{\{X<0\}}} &\to \mathrm{E}\,1_{\{X<0\}} = 0.2;
\end{aligned}$$

todėl pagal tolydaus atvaizdžio principą

$$T_n \to \frac{0.2}{0.6 - 0.2} = \frac{1}{2}.$$

Turiu:

$$T_n - \frac{1}{2} = \frac{\overline{1_{\{X=0\}}} - \frac{1}{2}\overline{1_{\{X>0\}}} + \frac{1}{2}\overline{1_{\{X<0\}}}}{\overline{1_{\{X>0\}}} - \overline{1_{\{X<0\}}}} = \frac{\overline{1_{\{X=0\}} - \frac{1}{2}1_{\{X>0\}} + \frac{1}{2}1_{\{X<0\}}}}{\overline{1_{\{X>0\}}} - \overline{1_{\{X<0\}}}}.$$

Kadangi

$$\begin{aligned}
\mathrm{E}\Big[1_{\{X=0\}} - \frac{1}{2}1_{\{X>0\}} + \frac{1}{2}1_{\{X<0\}}\Big] &= \mathrm{E}\,1_{\{X=0\}} - \frac{1}{2}\,\mathrm{E}\,1_{\{X>0\}} + \frac{1}{2}\,\mathrm{E}\,1_{\{X<0\}}\\
&= 0.2 - \frac{1}{2}0.6 + \frac{1}{2}0.2\\
&= 0,\\
\mathrm{E}\Big[1_{\{X=0\}} - \frac{1}{2}1_{\{X>0\}} + \frac{1}{2}1_{\{X<0\}}\Big]^2 &= \frac{1}{4}0.2 + 0.2 + \frac{1}{4}(0.2 + 0.4)\\
&= 0.4,\\
\mathrm{D}\Big[1_{\{X=0\}} - \frac{1}{2}1_{\{X>0\}} + \frac{1}{2}1_{\{X<0\}}\Big] &= 0.4,
\end{aligned}$$

iš centrinės ribinės teoremos gaunu

$$\sqrt{n}\Big[\overline{1_{\{X=0\}}} - \frac{1}{2}\overline{1_{\{X>0\}}} + \frac{1}{2}\overline{1_{\{X<0\}}}\Big] \to N(0, 0.4).$$

Reiškia, vėl pagal tolydaus atvaizdžio principą,

$$\sqrt{n}\Big[T_n - \frac{1}{2}\Big] \to \frac{N(0, 0.4)}{0.6 - 0.2} = \frac{1}{0.4}N(0, 0.4) = N\Big(0, \frac{1}{0.4}\Big) = N\Big(0, \frac{5}{2}\Big).$$

b. Pažymiu

$$T_n = \frac{\sqrt{(N^+)^2 + n^2}}{N^+ + n} = \frac{\sqrt{(N^+/n)^2 + 1}}{N^+/n + 1} = g(N^+/n) = g(\overline{1_{\{X>0\}}});$$

čia $g(t) = \sqrt{t^2 + 1}/(t + 1)$.

Kadangi

$$\mathrm{E}\, 1^2_{\{X>0\}} = \mathrm{E}\, 1_{\{X>0\}} = 0.6,$$

to dydžio dispersija lygi $0.6 - 0.36 = 0.24$; todėl pagal centrinę ribinę teoremą

$$\sqrt{n}(\overline{1_{\{X>0\}}} - 0.6) \to N(0, 0.24).$$

Kadangi $g(0.6) = \sqrt{1.36}/1.6$,

$$g'(t) = \frac{\frac{1}{2\sqrt{t^2+1}} \cdot 2t(t+1) - \sqrt{t^2+1}}{(t+1)^2} = \frac{t-1}{(t+1)^2\sqrt{t^2+1}}$$

ir $g'(0.6) = -0.4/(2.56\sqrt{1.36})$, pritaikęs delta metodą gaunu

$$\sqrt{n}\Big(T_n - \frac{\sqrt{1.36}}{1.6}\Big) \to -\frac{1}{6.4\sqrt{1.36}} N(0, 0.24) = N\Big(0, \frac{0.24}{6.4^2 1.36}\Big) \approx N(0, 0.0043).$$

c. Pažymiu

$$T_n = \frac{n\overline{X} - N^0}{N^+} = \frac{\overline{X} - N^0/n}{N^+/n} = \frac{\overline{X} - \overline{1_{\{X=0\}}}}{\overline{1_{\{X>0\}}}}.$$

Iš didžiųjų skaičių dėsnio

$$\begin{aligned} \overline{X} &\to \mathrm{E}\, X = -0.2 + 0 + 0.2 + 0.8 = 0.8; \\ \overline{1_{\{X=0\}}} &\to \mathrm{E}\, 1_{\{X=0\}} = 0.2; \\ \overline{1_{\{X>0\}}} &\to \mathrm{E}\, 1_{\{X>0\}} = 0.6; \end{aligned}$$

todėl

$$T_n \to \frac{0.8 - 0.2}{0.6} = 1.$$

Toliau:

$$T_n - 1 = \frac{\overline{X} - \overline{1_{\{X=0\}}} - \overline{1_{\{X>0\}}}}{\overline{1_{\{X>0\}}}} = \frac{\overline{X - 1_{\{X=0\}} - 1_{\{X>0\}}}}{\overline{1_{\{X>0\}}}}.$$

Kadangi

$$\begin{gathered} \mathrm{E}\, X - 1_{\{X=0\}} - 1_{\{X>0\}} = \mathrm{E}\, X - \mathrm{E}\, 1_{\{X=0\}} - \mathrm{E}\, 1_{\{X>0\}} = 0.8 - 0.2 - 0.6 = 0, \\ \mathrm{E}[X - 1_{\{X=0\}} - 1_{\{X>0\}}]^2 = 0.2 + 0.2 + 0 + 0.4 = 0.8, \end{gathered}$$

šio dydžio dispersija taip pat yra 0.8. Todėl iš centrinės ribinės teoremos

$$\sqrt{n}\,\overline{X - 1_{\{X=0\}} - 1_{\{X>0\}}} \to N(0, 0.8).$$

Tada iš tolydaus atvaizdžio principo

$$\sqrt{n}(T_n - 1) \to \frac{N(0, 0.8)}{0.6} = N\Big(0, \frac{2}{9}\Big).$$

1.6 PAVYZDYS. Atsitiktinio dydžio X tankis

$$p(x) = \begin{cases} 1/2, & \text{kai } -1 < x < 0; \\ x, & \text{kai } 0 < x < 1; \end{cases}$$

o $X_1, \ldots, X_n$ yra nepriklausomos X kopijos. Raskite nurodytų statistikų asimptotinį skirstinį:

a) $\dfrac{N^+}{2n + N^-}$;

b) $\ln(\overline{X} + 2)$;

c) $\dfrac{\overline{X} - \overline{X^3}}{\overline{X^2}}$;

čia N^+ ir N^- yra, atitinkamai, teigiamų ir neigiamų narių skaičius imtyje $X_1, \ldots, X_n$.

Sprendimas. a. Pažymiu

$$T_n = \frac{N^+}{2n + N^-} = \frac{N^+/n}{2 + N^-/n} = \frac{\overline{1_{\{X>0\}}}}{2 + \overline{1_{\{X<0\}}}} = \frac{\overline{1_{\{X>0\}}}}{3 - \overline{1_{\{X>0\}}}} = g(\overline{1_{\{X>0\}}});$$

čia

$$g(t) = \frac{t}{3 - t}.$$

Kadangi

$$\mathrm{E}\, 1^2_{\{X>0\}} = \mathrm{E}\, 1_{\{X>0\}} = \int_0^1 x \, \mathrm{d}x = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2},$$

to dydžio dispersija lygi $\frac{1}{2} - \frac{1}{4} = \frac{1}{4}$. Todėl iš centrinės ribinės teoremos

$$\sqrt{n}(\overline{1_{\{X>0\}}} - 0.5) \to N(0, 0.25).$$

Kadangi

$$g'(t) = \frac{3 - t - t \cdot (-1)}{(3 - t)^2} = \frac{3}{(3 - t)^2},$$

pritaikęs delta metodą gaunu

$$\sqrt{n}(T_n - g(0.5)) \to g'(0.5) N(0, 0.25);$$
$$\sqrt{n}\Big(T_n - \frac{0.5}{2.5}\Big) \to \frac{3}{2.5^2} N(0, 0.25);$$
$$\sqrt{n}\Big(T_n - \frac{1}{5}\Big) \to \frac{12}{25} N(0, 0.25) = N\Big(0, \frac{1.44}{25}\Big) = N(0, 0.0576).$$

b. Pažymiu

$$T_n = \ln(\overline{X} + 2) = g(\overline{X});$$

čia $g(x) = \ln(x + 2)$.

Kadangi

$$\mathrm{E}\, X = \int_{-1}^1 x p(x) \mathrm{d}x$$

$$
\begin{aligned}
&= \int_{-1}^{0} x \cdot \frac{1}{2}\mathrm{d}x + \int_{0}^{1} x^2\mathrm{d}x \\
&= \frac{1}{2}\frac{x^2}{2}\Big|_{-1}^{0} + \frac{x^3}{3}\Big|_{0}^{1} \\
&= -\frac{1}{4} + \frac{1}{3} \\
&= \frac{1}{12}
\end{aligned}
$$

ir analogiškai

$$
\begin{aligned}
\mathrm{E}\, X^2 &= \frac{1}{2}\int_{-1}^{0} x^2\mathrm{d}x + \int_{0}^{1} x^3\mathrm{d}x \\
&= \frac{x^3}{6}\Big|_{-1}^{0} + \frac{x^4}{4}\Big|_{0}^{1} \\
&= \frac{1}{6} + \frac{1}{4} \\
&= \frac{5}{12},
\end{aligned}
$$

atsitiktinio dydžio X dispersija yra $\frac{5}{12} - \frac{1}{144} = \frac{59}{144}$. Todėl iš centrinės ribinės teoremos

$$\sqrt{n}\Big(\overline{X} - \frac{1}{12}\Big) \to N\Big(0, \frac{59}{144}\Big).$$

Kadangi

$$g'(x) = \frac{1}{x+2},$$

pritaikęs delta metodą gaunu

$$\sqrt{n}\Big(T_n - \ln\frac{25}{12}\Big) \to \frac{12}{25}N\Big(0, \frac{59}{144}\Big) = N\Big(0, \frac{59}{625}\Big).$$

c. Pažymiu

$$T_n = \frac{\overline{X} - \overline{X^3}}{\overline{X^2}}.$$

Jau žinau, kad $\mathrm{E}\, X = \frac{1}{12}$ ir $\mathrm{E}\, X^2 = \frac{5}{12}$. Be to,

$$
\begin{aligned}
\mathrm{E}\, X^3 &= \frac{1}{2}\int_{-1}^{0} x^3\mathrm{d}x + \int_{0}^{1} x^4\mathrm{d}x \\
&= \frac{x^4}{8}\Big|_{-1}^{0} + \frac{x^5}{5}\Big|_{0}^{1} \\
&= -\frac{1}{8} + \frac{1}{5} \\
&= \frac{3}{40}.
\end{aligned}
$$

Todėl iš didžiųjų skaičių dėsnio

$$\overline{X} \to \frac{1}{12},$$

$$\overline{X^2} \to \frac{5}{12},$$
$$\overline{X^3} \to \frac{3}{40}$$

ir pagal tolydaus atvaizdžio principą

$$T_n \to \frac{\frac{1}{12} - \frac{3}{40}}{\frac{5}{12}} = \frac{\frac{1}{120}}{\frac{5}{12}} = \frac{1}{50}.$$

Toliau

$$T_n - \frac{1}{50} = \frac{\overline{X} - \overline{X^3} - \frac{1}{50}\overline{X^2}}{\overline{X^2}} = \frac{\overline{X - X^3 - \frac{1}{50}X^2}}{\overline{X^2}}.$$

Kadangi

$$\mathrm{E}\Big[X - X^3 - \frac{1}{50}X^2\Big] = \frac{1}{12} - \frac{3}{40} - \frac{1}{50} \cdot \frac{5}{12} = 0$$

ir

$$\begin{aligned}
&\mathrm{E}\Big[X - X^3 - \frac{1}{50}X^2\Big]^2 \\
&= \frac{1}{2}\int_{-1}^{0}\Big(x - x^3 - \frac{1}{50}x^2\Big)^2 \mathrm{d}x + \int_{0}^{1}\Big(x - x^3 - \frac{1}{50}x^2\Big)^2 x \mathrm{d}x \\
&= \frac{1}{2}\int_{-1}^{0}\Big(x^2 - \frac{1}{25}x^3 - \frac{4999}{2500}x^4 + \frac{1}{25}x^5 + x^6\Big)\mathrm{d}x \\
&\quad + \int_{0}^{1}\Big(x^3 - \frac{1}{25}x^4 - \frac{4999}{2500}x^5 + \frac{1}{25}x^6 + x^7\Big)\mathrm{d}x \\
&= \frac{1}{6} + \frac{1}{200} - \frac{4999}{25000} - \frac{1}{300} + \frac{1}{14} + \frac{1}{4} - \frac{1}{125} - \frac{4999}{15000} + \frac{1}{175} + \frac{1}{8},
\end{aligned}$$

to dydžio dispersija taip pat lygi

$$\sigma^2 = \frac{1}{6} + \frac{1}{200} - \frac{4999}{25000} - \frac{1}{300} + \frac{1}{14} + \frac{1}{4} - \frac{1}{125} - \frac{4999}{15000} + \frac{1}{175} + \frac{1}{8}.$$

Todėl iš centrinės ribinės teoremos

$$\sqrt{n}\,\overline{X - X^3 - \frac{1}{50}X^2} \to N(0, \sigma^2),$$

o tada pagal tolydaus atvaizdžio principą

$$\sqrt{n}\Big(T_n - \frac{1}{50}\Big) \to \frac{N(0, \sigma^2)}{\frac{5}{12}} = N\Big(0, \frac{144\sigma^2}{25}\Big).$$

2 skyrius

Stochastinis konvergavimas

2.1 Apibrėžimas ir pavyzdžiai

2.1.1 Apibrėžimas

Tegu X_n ir X yra atsitiktiniai vektoriai su reikšmėmis $\mathbb{R}^k$ erdvėje. Sakome, kad X_n *konverguoja pagal pasiskirstymą* į X, ir rašome $X_n \to X$, jei

$$\mathrm{E}\, f(X_n) \to \mathrm{E}\, f(X)$$

su bet kokia aprėžta tolydžia funkcija $f\colon \mathbb{R}^k \to \mathbb{R}$.

Aišku, kad jei $X_n \to X$, tai ir bet koks (X_n) sekos posekis konverguoja į X.

Jei $X_n \to X$ ir $X_n \to Y$, tai X ir Y skirstiniai sutampa. Tai įrodoma taip. Jei O yra atviras $\mathbb{R}^k$ poaibis, tai egzistuoja tokia tolydžių neneigiamų funkcija seka (f_m), kad $f_m \uparrow 1_O$. Aišku, kad

$$\mathrm{E}\, f_m(X) = \lim_{n\to\infty} \mathrm{E}\, f_m(X_n) = \mathrm{E}\, f_m(Y).$$

Todėl

$$\mathrm{P}\{X \in O\} = \lim_{m\to\infty} \mathrm{E}\, f_m(X) = \lim_{m\to\infty} \mathrm{E}\, f_m(Y) = \mathrm{P}\{Y \in O\}.$$

Tegu dabar $\mathcal{M}$ yra aibė visų Borelio aibių A, su kuriomis $\mathrm{P}\{X \in A\} = \mathrm{P}\{Y \in A\}$. Iš tikimybių savybių išplaukia, kad

$\mathbb{R}^k \in \mathcal{M}$;

$A, B \in \mathcal{M},\ A \subset B \Rightarrow B \setminus A \in \mathcal{M}$;

$A_n \in \mathcal{M}$, $A_n \uparrow A \Rightarrow A \in \mathcal{M}$.

Reiškia, $\mathcal{M}$ yra monotoniška poaibių klasė. Kadangi jai priklauso visos atviros aibės, o atvirų aibių sankirta vėl atvira, iš teoremos apie monotonišką poaibių klasę išplaukia, kad $\mathcal{M}$ sutampa su visų Borelio aibių sistema.

Jei visi X_n vektoriai išsigimę, t.y. beveik tikrai įgyja vienintelę reikšmę x_n, tai jie turi ribą tada ir tik tada, kai vektorių seka (x_n) konverguoja. Be to, jei $x_n \to x$, tai $X_n \to X$; čia X yra išsigimęs atsitiktinis vektorius, beveik tikrai įgyjantis reikšmę x.

Įrodysiu tai. Tegu $x_n \to x$. Jei f yra bet kokia tolydžioji funkcija, tai

$$\mathrm{E}\, f(X_n) = f(x_n) \to f(x) = \mathrm{E}\, f(X).$$

Reiškia, $X_n \to X$.

Dabar įrodysiu, kad jei (x_n) diverguoja, tai (X_n) taip pat neturi ribos. Čia galimi keli atvejai. Tegu iš pradžių $\|x_n\| \to \infty$. Tegu $f_m(x) = g(\|x\|/m)$; čia

$$g(t) = \begin{cases} 1, & \text{kai } t \leqslant 1; \\ 2-t, & \text{kai } 1 < t \leqslant 2; \\ 0, & \text{kai } t > 2. \end{cases}$$

Aišku, kad f_m yra tolydi funkcija, $0 \leqslant f_m \leqslant 1$; be to $f(x) = 1$, kai $\|x\| \leqslant m$, ir $f(x) = 0$, kai $\|x\| > 2m$. Jei $X_n \to X$, iš čia išplauktų

$$\mathrm{P}\{\|X\| \leqslant m\} \leqslant \mathrm{E}\, f_m(X) = \lim_{n\to\infty} \mathrm{E}\, f_m(X_n) = \lim_{n\to\infty} f_m(x_n) = 0.$$

Tada $\mathrm{P}\{\|X\| < \infty\} = 0$, o to negali būti. Reiškia, $X_n \not\to X$ su jokiu X.

Bendruoju atveju, jei (x_n) nekonverguoja, tai arba $\|x_{n_k}\| \to \infty$ su kokiu nors posekiu (n_k), arba kokie nors du jos posekiai turi skirtingas ribas. Pirmuoju atveju tam tikras (X_{n_k}) posekis neturi ribos, antruoju — du (X_n) sekos posekiai turi skirtingas ribas. Abiem atvejais (X_n) seka ribos neturi.

2.1.2 Pavyzdžiai

2.1 PAVYZDYS. Tegu X_n skirstinys aprašomas lentele

X_n	$\frac{1}{n}$	1	$\frac{n}{n+1}$
	0.4	0.3	0.3

Tada

$$\mathrm{E}\, f(X_n) = 0.4 f\Big(\frac{1}{n}\Big) + 0.3 f(1) + 0.3 f\Big(\frac{n}{n+1}\Big).$$

Aišku, kad $\frac{1}{n} \to 0$, $\frac{n}{n+1} \to 1$. Jei f tolydi, tai

$$f\Big(\frac{1}{n}\Big) \to f(0), \quad f\Big(\frac{n}{n+1}\Big) \to f(1)$$

ir todėl

$$\mathrm{E}\, f(X_n) \to 0.4 f(0) + 0.3 f(1) + 0.3 f(1) = 0.4 f(0) + 0.6 f(1).$$

Reiškinys dešinėje pusėje sutampa su $\mathrm{E}\, f(X)$, jei X yra atsitiktinis dydis, kurio skirstinys aprašomas lentele

X	0	1
	0.4	0.6

Reiškia, $X_n \to X$.

2.2 PAVYZDYS. Tegu X_n skirstinys aprašomas lentele

X_n	1	2	3
	$\frac{1}{n}$	$\frac{2}{n}$	$1 - \frac{3}{n}$

Tada

$$\mathrm{E}\,f(X_n) = \frac{1}{n}f(1) + \frac{2}{n}f(2) + \Big(1 - \frac{3}{n}\Big)f(3) \to f(3) = \mathrm{E}\,f(X);$$

čia X yra išsigimęs atsitiktinis dydis, įgyjantis reikšmę 3 su tikimybe 1. Tokį dydį aš žymiu tiesiog 3. Taigi $X_n \to 3$.

2.3 PAVYZDYS. Tegu X_n skirstinys aprašomas lentele

X_n	0	1	n
	$\frac{1}{2}$	$\frac{1}{2} - \frac{1}{n}$	$\frac{1}{n}$

Tada

$$\mathrm{E}\,f(X_n) = \frac{1}{2}f(0) + \Big(\frac{1}{2} - \frac{1}{n}\Big)f(1) + \frac{1}{n}f(n).$$

Jei f aprėžta, tai $f(n)/n \to 0$; todėl

$$\mathrm{E}\,f(X_n) \to \frac{1}{2}f(0) + \frac{1}{2}f(1) = \mathrm{E}\,f(X);$$

čia X yra atsitiktinis dydis, kurio skirstinys aprašomas lentele

X	0	1
	0.5	0.5

Taigi $X_n \to X$.

2.4 PAVYZDYS. Tegu X_n skirstinys aprašomas lentele

X_n	$\frac{1}{n}$	$\frac{2}{n}$	$\cdots$	1
	$\frac{1}{n}$	$\frac{1}{n}$	$\cdots$	$\frac{1}{n}$

Tada

$$\mathrm{E}\,f(X_n) = \frac{1}{n}f(1/n) + \frac{1}{n}f(2/n) + \cdots + \frac{1}{n}f(1)$$

Reiškinys dešinėje pusėje yra funkcijos f integralinė suma $(0;1)$ intervale, atitinkanti intervalo skaidinį į n lygių dalių. Todėl

$$\mathrm{E}\,f(X_n) \to \int_0^1 f(x)\mathrm{d}x = \mathrm{E}\,f(X);$$

čia X yra atsitiktinis dydis, pasiskirstęs tolygiai $(0;1)$ intervale. Reiškia, $X_n \to X$.

2.5 PAVYZDYS. Jei X_n yra tolydūs atsitiktiniai dydžiai su tankiais $p_n(x)$, $p(x)$ yra dar vienas tankis ir $p_n(x) \to p(x)$ su visais x, tai X_n konverguoja į atsitiktinį dydį X, kurio tankis yra $p(x)$. Tai įrodoma taip.

Jei f yra tolydi aprėžta funkcija, tai

$$\begin{aligned}|\mathrm{E}\,f(X_n) - \mathrm{E}\,f(X)| &= \Big|\int f(x)p_n(x)\mathrm{d}x - \int f(x)p(x)\mathrm{d}x\Big| \\ &\leqslant c\int \big|p_n(x) - p(x)\big|\mathrm{d}x\end{aligned}$$

$$= c\int_{p_n \geqslant p} \big(p_n(x) - p(x)\big)\mathrm{d}x + c\int_{p > p_n} \big(p(x) - p_n(x)\big)\mathrm{d}x.$$

Kadangi p_n yra tankis,

$$\int_{p_n \geqslant p} p_n(x)\mathrm{d}x = 1 - \int_{p_n < p} p_n(x)\mathrm{d}x.$$

Analogiškai gaunu

$$\int_{p_n \geqslant p} p(x)\mathrm{d}x = 1 - \int_{p_n < p} p(x)\mathrm{d}x.$$

Reiškia,

$$|\mathrm{E}\, f(X_n) - \mathrm{E}\, f(X)| \leqslant 2c\int_{p > p_n} \big(p(x) - p_n(x)\big)\mathrm{d}x \leqslant 2c\int_{p > p_n} p(x)\mathrm{d}x \to 0,$$

nes $1_{\{p > p_n\}}(x) \to 0$ su visais x.

2.1.3 Kaip įrodyti, kad ribos nėra

Yra du pagrindiniai atvejai, kada X_n seka neturi ribos pagal pasiskirstymą:

- kai du sekos posekiai turi skirtingas ribas;
- kai dalis skirstinio masės dingsta begalybėje.

Pirmąjį atvejį iliustruoja toks pavyzdys.

2.6 PAVYZDYS. Tegu X_n skirstinys aprašomas lentele

X_n	1	2
	$\frac{1}{2} + \frac{(-1)^n n}{2n+1}$	$\frac{1}{2} - \frac{(-1)^n n}{2n+1}$

Jei $n = 2k$, tai lentelė supaprastėja:

X_{2k}	1	2
	$\frac{1}{2} + \frac{2k}{4k+1}$	$\frac{1}{2} - \frac{2k}{4k+1}$

Šiuo atveju

$$\mathrm{E}\, f(X_{2k}) = \Big(\frac{1}{2} + \frac{2k}{4k+1}\Big) f(1) + \Big(\frac{1}{2} - \frac{2k}{4k+1}\Big) f(2) \to f(1);$$

todėl $X_{2k} \xrightarrow[k\to\infty]{} 1$.

Jei $n = 2k+1$, X_n skirstinys aprašomas tokia lentele:

X_{2k+1}	1	2
	$\frac{1}{2} - \frac{2k}{4k+1}$	$\frac{1}{2} + \frac{2k}{4k+1}$

Šiuo atveju

$$\mathrm{E}\, f(X_{2k+1}) = \Big(\frac{1}{2} - \frac{2k}{4k+1}\Big) f(1) + \Big(\frac{1}{2} + \frac{2k}{4k+1}\Big) f(2) \to f(2);$$

todėl $X_{2k+1} \xrightarrow[k\to\infty]{} 2$.

Taigi X_n seka ribos neturi.

Antrą atvejį iliustruoja kitas pavyzdys.

2.7 PAVYZDYS. Tegu X_n skirstinys aprašomas lentele

X_n	1	n
	0.4	0.6

Tada

$$\mathrm{E}\, f(X_n) = 0.4 f(1) + 0.6 f(n).$$

Galima sugalvoti tokią aprėžtą tolydžią funkciją f, kad seka $f(n)$ neturėtų ribos (pavyzdžiui, jei $f(x) = \cos \pi x$, tai $f(n) = \cos \pi n = (-1)^n$). Tada $\mathrm{E}\, f(X_n)$ neturi ribos ir, reiškia, $X_n \not\to X$ su jokiu X. Tačiau paprastai šis faktas įrodomas, pasirėmus vienu teiginiu, kurį ruošiuosi įrodyti vėliau šiame skyriuje.

Teiginys, kurį ruošiuosi įrodyti vėliau, skamba taip: jei $X_n \to X$, tai $X_n = O_P(1)$, t.y.

$$\sup_n \mathrm{P}\{|X_n| \geqslant k\} \xrightarrow[k\to\infty]{} 0.$$

Iš čia išplaukia, kad su bet kokiu posekiu n_k

$$\mathrm{P}\{|X_{n_k}| \geqslant k\} \xrightarrow[k\to\infty]{} 0.$$

Nagrinėjamame pavyzdyje paimu $n_k = k$. Tada

$$\mathrm{P}\{|X_k| \geqslant k\} = 0.6 \not\to 0;$$

todėl $X_n \not\to X$ su jokiu X.

2.2 Portmanto lema

2.2.1 Formulavimas ir įrodymas

2.1 teorema. *Tokie penki teiginiai yra ekvivalentūs:*

1) $X_n \to X$*;*

2) $\mathrm{E}\, f(X_n) \to \mathrm{E}\, f(X)$ *su bet kokia aprėžta Lipšico funkcija* f*;*

3) $\underline{\underline{\lim}}\, \mathrm{P}\{X_n \in O\} \geqslant \mathrm{P}\{X \in O\}$ *su bet kokia atvira aibe* O*;*

4) $\overline{\lim}\, \mathrm{P}\{X_n \in C\} \leqslant \mathrm{P}\{X \in C\}$ *su bet kokia uždara aibe* C*;*

5) $\mathrm{P}\{X_n \in A\} \to \mathrm{P}\{X \in A\}$ *su bet kokia Borelio aibe* A*, tenkinančia sąlygą* $\mathrm{P}\{X \in \partial A\} = 0$.

Įrodymas. ($1 \Rightarrow 2$) Akivaizdu, nes kiekviena Lipšico funkcija tolydi.

($2 \Rightarrow 3$) Raskime tokią aprėžtų Lipšico funkcijų seką (f_m), kad $f_m \uparrow 1_O$ (tinka, pavyzdžiui, $f_m(x) = \big[m d(x, O^c)\big] \wedge 1$). Tada su visais m

$$\underline{\lim}\, \mathrm{P}\{X_n \in O\} \geqslant \lim_{n\to\infty} \mathrm{E}\, f_m(X_n) = \mathrm{E}\, f_m(X); \tag{2.1}$$

todėl

$$\underline{\lim}\, \mathrm{P}\{X_n \in O\} \geqslant \lim_{m\to\infty} \mathrm{E}\, f_m(X) = \mathrm{P}\{X \in O\}. \tag{2.2}$$

($3 \Rightarrow 4$) Tegu C uždara ir $O = C^c$. Tada O atvira ir todėl

$$\overline{\lim}\, \mathrm{P}\{X_n \in C\} = 1 - \underline{\lim}\, \mathrm{P}\{X_n \in O\} \leqslant 1 - \mathrm{P}\{X \in O\} = \mathrm{P}\{X \in C\}. \tag{2.3}$$

Analogiškai įrodoma $4 \Rightarrow 3$ implikacija.

($4 \Rightarrow 5$) Tegu A yra Borelio poaibis, $C = \overline{A}$ ir $O = \operatorname{Int} A$. Tada

$$\begin{aligned} \mathrm{P}\{X \in O\} &\leqslant \varliminf \mathrm{P}\{X_n \in O\} \leqslant \varliminf \mathrm{P}\{X_n \in A\} \\ &\leqslant \varlimsup \mathrm{P}\{X_n \in A\} \leqslant \varlimsup \mathrm{P}\{X_n \in C\} \leqslant \mathrm{P}\{X \in C\}. \end{aligned} \tag{2.4}$$

Iš $\mathrm{P}\{X \in \partial A\} = 0$ išplaukia, kad $\mathrm{P}\{X \in O\} = \mathrm{P}\{X \in C\} = \mathrm{P}\{X \in A\}$. Reiškia, $\mathrm{P}\{X_n \in A\} \to \mathrm{P}\{X \in A\}$.

($5 \Rightarrow 1$) Tegu f yra bet kokia aprėžta tolydi funkcija. Tarkime $f(x) \in (-c; c)$ su visais x.

Randu tokią $(-c; c)$ intervalo skaidinių seką $(c_{n0}, \ldots, c_{ns_n})$, kad

$$c_{nj} - c_{n,j-1} \leqslant \frac{1}{n} \quad \text{ir} \quad \mathrm{P}\{f(X) = c_{nj}\} = 0$$

su visais j ir n. Pažymiu

$$A_{nj} = \{x \mid c_{n,j-1} \leqslant f(x) < c_{nj}\}, \quad g_n = \sum_{j=1}^{s_n} c_{n,j-1} 1_{A_{nj}} \quad \text{ir} \quad h_n = \sum_{j=1}^{s_n} c_{nj} 1_{A_{nj}}.$$

Aišku, kad g_n, h_n yra Borelio funkcijos $|g_n| \leqslant c$, $|h_n| \leqslant c$, $g_n \leqslant f \leqslant h_n$, $g_n \to f$ ir $h_n \to f$.

Dėl f tolydumo aibės $O_{nj} = \{x \mid c_{n,j-1} < f(x) < c_{nj}\}$ yra atviros, o aibės $C_{nj} = \{x \mid c_{n,j-1} \leqslant f(x) \leqslant c_{nj}\}$ — uždaros. Reiškia,

$$\partial A_{nj} \subset C_{nj} \setminus O_{nj} = \{x \mid f(x) = c_{n,j-1}\} \cup \{x \mid f(x) = c_{nj}\} \tag{2.5}$$

ir todėl $\mathrm{P}(X \in \partial A_{nj}) = 0$. Su kiekvienu m

$$\mathrm{E}\, f(X_n) \geqslant \mathrm{E}\, g_m(X_n) \to \mathrm{E}\, g_m(X); \tag{2.6}$$

todėl

$$\varliminf \mathrm{E}\, f(X_n) \geqslant \lim_{m\to\infty} \mathrm{E}\, g_m(X) = \mathrm{E}\, f(X). \tag{2.7}$$

Analogiškai gaunu, kad su visais m

$$\mathrm{E}\, f(X_n) \leqslant \mathrm{E}\, h_m(X_n) \to \mathrm{E}\, h_m(X);$$

todėl

$$\varlimsup \mathrm{E}\, f(X_n) \geqslant \lim_{m\to\infty} \mathrm{E}\, h_m(X) = \mathrm{E}\, f(X).$$

Reiškia, $\mathrm{E}\, f(X_n) \to \mathrm{E}\, f(X)$. $\square$

2.2.2 Įrodymo paaiškinimai

1. Kas yra Lipšico funkcija?

Funkcija $f\colon \mathbb{R}^k \to \mathbb{R}$ vadinama Lipšico, jei egzistuoja tokia konstanta q, kad su visais $x, y \in \mathbb{R}^k$

$$\bigl|f(x) - f(y)\bigr| \leqslant q\|x - y\|.$$

2. Kodėl kiekviena Lipšico funkcija tolydi?

Jei $x_n \to x$, tai

$$\bigl|f(x_n) - f(x)\bigr| \leqslant q\|x_n - x\| \to 0,$$

t.y. $f(x_n) \to f(x)$.

3. Kodėl f_m funkcija aprėžta?
Nes $0 \leqslant f_m(x) \leqslant 1$ su visais x.

4. Kodėl f_m yra Lipšico funkcija?
Nes

$$\bigl|f_m(x) - f_m(y)\bigr| \leqslant \bigl|md(x, O^c) - md(y, O^c)\bigr| \leqslant md(x, y) = m\|x - y\|.$$

Antras sąryšis išplaukia iš nelygybės

$$\bigl|d(x, A) - d(y, A)\bigr| \leqslant d(x, y),$$

teisingos bet kokioje metrinėje erdvėje.

Pirmas sąryšis išplaukia iš nelygybės

$$|a \wedge 1 - b \wedge 1| \leqslant |a - b|,$$

teisingos su visais $a, b \geqslant 0$. Nelygybė įrodoma, perrenkant visus galimus variantus:

jei $a, b \leqslant 1$, ji virsta $|a - b| \leqslant |a - b|$ nelygybe;
jei $a \leqslant 1 < b$, ji virsta $1 - a \leqslant b - a$ nelygybe, kuri teisinga, nes $b > 1$;
jei $a, b > 1$, ji virsta $0 \leqslant |a - b|$ nelygybe.

5. Kodėl $f_m \uparrow 1_O$?
Šis klausimas išsiskaido į du: kodėl (f_m) yra nemažėjanti seka ir kodėl $f_m \to 1_O$?

Atsakymas į pirmą klausimą: jei $m < n$, tai $md(X, O^c) \leqslant nd(x, O^c)$ (nes $d(x, O^c) \geqslant 0$) ir todėl $md(X, O^c) \wedge 1 \leqslant nd(x, O^c) \wedge 1$ (tai galima įrodyti, perrenkant variantus).

Atsakymas į antrą klausimą: jei $x \in O^c$, tai

$$f_m(x) = (m \cdot 0) \wedge 1 = 0 \to 0 = 1_O(x);$$

jei $x \in O$, tai $d(x, O^c) > 0$ (nes O^c uždara); todėl $md(x, O^c) \to \infty$ ir

$$f_m(x) \to 1 = 1_O(x).$$

6. Kodėl teisingi sąryšiai (2.1) grandinėje?
Su visais n

$$\mathrm{P}\{X_n \in O\} = \mathrm{E}\,1_O(X_n) \geqslant \mathrm{E}\,f_m(X_n)$$

(nes $1_O \geqslant f_m$); todėl

$$\varliminf_{n\to\infty} \mathrm{P}\{X_n \in O\} \geqslant \varliminf_{n\to\infty} \mathrm{E}\,f_m(X_n).$$

Tačiau f_m yra Lipšico funkcija; todėl iš 2 prielaidos gauname $\mathrm{E}\,f_m(X_n) \to \mathrm{E}\,f_m(X)$, t.y.

$$\varliminf_{n\to\infty} \mathrm{E}\,f_m(X_n) = \lim_{n\to\infty} \mathrm{E}\,f_m(X_n) = \mathrm{E}\,f_m(X).$$

7. Kodėl teisingi sąryšiai (2.2) grandinėje?

Iš (2.1) išplaukia, kad su visais m

$$\varliminf_{n\to\infty} \mathrm{P}\{X_n \in O\} \geqslant \mathrm{E}\, f_m(X).$$

Iš čia

$$\varliminf_{n\to\infty} \mathrm{P}\{X_n \in O\} \geqslant \varlimsup_{m\to\infty} \mathrm{E}\, f_m(X).$$

Tačiau iš $0 \leqslant f_m \uparrow 1_O$ ir Lebego teoremos apie monotonišką konvergavimą išplaukia

$$\mathrm{E}\, f_m(X) \uparrow \mathrm{E}\, 1_O(X) = \mathrm{P}\{X \in O\}.$$

Reiškia,

$$\varlimsup_{m\to\infty} \mathrm{E}\, f_m(X) = \lim_{m\to\infty} \mathrm{E}\, f_m(X) = \mathrm{P}\{X \in O\}.$$

8. Kodėl teisingi sąryšiai (2.3) grandinėje?

Pirma lygybė gaunama pasirėmus tuo, kad

$$\mathrm{P}\{X_n \in C\} = \mathrm{P}\{X_n \notin O\} = 1 - \mathrm{P}\{X_n \in O\},$$

ir tuo, kad

$$\varlimsup(1 - a_n) = 1 - \varliminf a_n$$

su bet kokia seka (a_n).

Antras sąryšis išplaukia iš 3 teoremos prielaidos, o paskutinė lygybė — vėl iš to, kad $C = O^c$.

9. Kodėl teisingi sąryšiai (2.4) grandinėje?

Pirma nelygybė teisinga, nes iš 4 prielaidos išplaukia 3 prielaida, o O aibė atvira. Antra nelygybė teisinga, nes $O \subset A$. Trečia nelygybė teisinga, nes apatinė sekos riba visada neviršija viršutinės. Ketvirta nelygybė teisinga, nes $A \subset C$. Paskutinė nelygybė išplaukia iš 4 prielaidos.

10. Kaip iš $\mathrm{P}\{X \in \partial A\} = 0$ išplaukia $\mathrm{P}\{X \in O\} = \mathrm{P}\{X \in C\} = \mathrm{P}\{X \in A\}$?

Kadangi $O \subset A \subset C$ ir $\partial A = C \setminus O$,

$$0 = \mathrm{P}\{X \in \partial A\} = \mathrm{P}\{X \in C\} - \mathrm{P}\{X \in O\},$$

t.y. $\mathrm{P}\{X \in O\} = \mathrm{P}\{X \in C\}$. Kita vertus, iš $O \subset A \subset C$ gauname

$$\mathrm{P}\{X \in O\} \leqslant \mathrm{P}\{X \in A\} \leqslant \mathrm{P}\{X \in C\}.$$

Kraštiniai šios nelygybių grandinės nariai vienodi; todėl visos nelygybės virsta lygybėmis.

11. Kodėl daroma išvada, kad $\mathrm{P}\{X_n \to A\} \to \mathrm{P}\{X \in A\}$?

Kraštiniai (2.4) grandinės nariai lygūs $\mathrm{P}\{X \in A\}$. Reiškia, tam skaičiui lygūs ir visi kiti nariai; pavyzdžiui,

$$\varliminf_{n\to\infty} \mathrm{P}\{X_n \in A\} = \varlimsup_{n\to\infty} \mathrm{P}\{X_n \in A\} = \mathrm{P}\{X \in A\}.$$

Jei skaičių sekos apatinė ir viršutinė riba sutampa, tai ta sutampanti reikšmė ir yra sekos riba.

12. Kodėl egzistuoja tokia skaidinių seka (c_{nj})?

Yra tik ne daugiau kaip skaiti aibė skaičių a, su kuriais $\mathrm{P}\{f(X)=a\}>0$ (nes su kiekvienu k yra ne daugiau kaip k skirtingų a, su kuriais $\mathrm{P}\{f(X)=a\}\geqslant 1/k$). Todėl šalia kiekvieno a kiek norima arti yra toks a', kad $\mathrm{P}\{f(X)=a'\}=0$.

Skaidinys (c_{nj}) gali būti rastas taip: padaliname $(-c;c)$ intervalą bet kaip į $<\frac{1}{3n}$ ilgio gabalus ir šalia kiekvieno dalijimo taško c'_{nj} randame tokį c_{nj}, kad $|c_{nj}-c'_{nj}|<\frac{1}{3n}$ ir $\mathrm{P}\{f(X)=c_{nj}\}=0$.

13. Kodėl $|g_n|\leqslant c$?

Funkcija g_n įgyja tik reikšmes $c_{n,j-1}$, $j=1,\dots,s_n$. Taigi $g_n(x)\geqslant c_{n0}=-c$ ir $g_n(x)\leqslant c_{ns_n}=c$ su visais x.

14. Kodėl $g_n\leqslant f\leqslant h_n$?

Koks bebūtų x, atsiras toks j, kad $c_{n,j-1}\leqslant f(x)<c_{nj}$. Tuomet $g_n(x)=c_{n,j-1}\leqslant f(x)$, o $h_n(x)=c_{nj}\geqslant f(x)$.

15. Kodėl $g_n\to f$?

Jei $c_{n,j-1}\leqslant f(x)<c_{nj}$, tai

$$\bigl|g_n(x)-f(x)\bigr|=f(x)-c_{n,j-1}\leqslant c_{nj}-c_{n,j-1}\leqslant\frac{1}{n}.$$

16. Paaiškinkite smulkiau, kodėl O_{nj} atvira.

O_{nj} yra atviros aibės $(c_{n,j-1};c_{nj})$ pirmavaizdis tolydžiosios funkcijos f atžvilgiu; todėl atvira.

17. Kodėl teisingi sąryšiai (2.5) grandinėje?

Kadangi C_{nj} uždara ir $\supset A_{nj}$, ji didesnė už $\overline{A_{nj}}$. Aibė O_{nj} atvira ir $\subset A_{nj}$; todėl ji mažesnė už $\operatorname{Int}A_{nj}$. Reiškia,

$$C_{nj}\setminus O_{nj}\supset\overline{A_{nj}}\setminus\operatorname{Int}A_{nj}=\partial A_{nj}.$$

Antras sąryšis įrodomas taip. Tegu $x\in C_{nj}\setminus O_{nj}$. Tada $x\in C_{nj}$, t.y.

$$c_{n,j-1}\leqslant f(x)\leqslant c_{nj}.$$

Be to, $x\notin O_{nj}$, t.y. arba $c_{n,j-1}<f(x)$, arba $f(x)<c_{nj}$ sąryšis neteisingas. Pirmu atveju gaunu $f(x)=c_{n,j-1}$, antruoju — $f(x)=c_{nj}$.

18. Kodėl $\mathrm{P}\{X\in\partial A\}=0$?

Iš (2.5)

$$\mathrm{P}\{X\in\partial A\}\leqslant\mathrm{P}\{f(X)=c_{n,j-1}\}+\mathrm{P}\{f(X)=c_{nj}\},$$

o pagal c_{nj} parinkimą $\mathrm{P}\{f(X)=c_{n,j-1}\}=\mathrm{P}\{f(X)=c_{nj}\}=0$.

19. Kodėl teisingi sąryšiai (2.6) grandinėje?

Pirma nelygybė teisinga dėl to, kad $g_n\leqslant f$. Antras sąryšis — dėl to, kad

$$\mathrm{E}\,g_m(X_n)=\sum_{j=1}^{s_m}c_{m,j-1}\,\mathrm{P}\{X_n\in A_{mj}\}$$

ir pagal 5 prielaidą $\mathrm{P}\{X_n\in A_{mj}\}\to\mathrm{P}\{X\in A_{mj}\}$ (nes $\mathrm{P}\{X\in\partial A_{mj}\}=0$).

20. Kodėl teisingi sąryšiai (2.7) grandinėje?

Pirmas sąryšis išplaukia iš (2.6), o antras — iš Lebego teoremos apie aprėžtą konvergavimą (nes $g_m\to f$ ir $|g_m|\leqslant c$).

2.2.3 Komentarai

21. Jei $X_n \to X$, tai su bet kokia atvira O

$$\varliminf_{n\to\infty} \mathrm{P}\{X_n \in O\} \geqslant \mathrm{P}\{X \in O\}.$$

Nelygybės pakeisti lygybe apskritai negalima. Tegu, pavyzdžiui, $X_n = 1/n$; tada $X = 0$. Aibė $O = (0;\infty)$ atvira, bet su visais n

$$\mathrm{P}\{X_n \in O\} = 1, \quad \text{o} \quad \mathrm{P}\{X \in O\} = 0.$$

Taip pat negalima $\varliminf$ pakeisti į lim, nes seka $\mathrm{P}\{X_n \in O\}$ gali neturėti ribos. Pavyzdžiui, jei $X_n = (-1)^n/n$, $X = 0$ ir $O = (0;\infty)$, tai

$$\mathrm{P}\{X_n \in O\} = \begin{cases} 0, & \text{kai } n \text{ nelyginis;} \\ 1, & \text{kai } n \text{ lyginis;} \end{cases}$$

ir ribos neturi.

22. Jei $X_n \to X$, tai su bet kokia uždara C

$$\varlimsup_{n\to\infty} \mathrm{P}\{X_n \in C\} \leqslant \mathrm{P}\{X \in C\}.$$

Nelygybės pakeisti lygybe apskritai negalima. Tegu, pavyzdžiui, $X_n = -1/n$; tada $X = 0$. Aibė $C = [0;\infty)$ uždara, bet su visais n

$$\mathrm{P}\{X_n \in C\} = 0, \quad \text{o} \quad \mathrm{P}\{X \in C\} = 1.$$

Taip pat negalima $\varlimsup$ pakeisti į lim, nes seka $\mathrm{P}\{X_n \in C\}$ gali neturėti ribos. Pavyzdžiui, jei $X_n = (-1)^n/n$, $X = 0$ ir $C = [0;\infty)$, tai

$$\mathrm{P}\{X_n \in C\} = \begin{cases} 0, & \text{kai } n \text{ nelyginis;} \\ 1, & \text{kai } n \text{ lyginis;} \end{cases}$$

ir ribos neturi.

23. Jei $X_n \to X$, tai

$$\mathrm{P}\{X_n \in A\} \to \mathrm{P}\{X \in A\}$$

su bet kokia aibe A, tenkinančia sąlygą $\mathrm{P}\{X \in \partial A\} = 0$. Pastarosios sąlygos praleisti apskritai negalima. Tegu, pavyzdžiui, $X_n = 1/n$, $X = 0$ ir $A = (0;1)$. Tada su visais n

$$\mathrm{P}\{X_n \in A\} = 1, \quad \text{bet} \quad \mathrm{P}\{X \in A\} = 0.$$

Priežastis, dėl kurios $\mathrm{P}\{X_n \in A\}$ nekonverguoja į $\mathrm{P}\{X \in A\}$, paprasta: šiuo atveju $\partial A = \{0, 1\}$ ir

$$\mathrm{P}\{X \in \partial A\} = 0.$$

24. Jei (X_n) yra atsitiktinių dydžių seka ir $X_n \to X$, tai Portmanto lemos 5 teiginyje paėmęs $A = (-\infty; x]$ gaunu

$$\mathrm{P}\{X_n \leqslant x\} \xrightarrow[n\to\infty]{} \mathrm{P}\{X \leqslant x\},$$

jei tik $\mathrm{P}\{X = x\} = 0$.

Funkcija $F(x) = \mathrm{P}\{X \leqslant x\}$ vadinama atsitiktinio dydžio X pasiskirstymo funkcija, o sąlyga $\mathrm{P}\{X = x\} = 0$ reiškia, kad x yra tos funkcijos tolydumo taškas. Taigi jei $X_n \to X$, F_n yra X_n dydžio, o F — X dydžio pasiskirstymo funkcija, tai $F_n(x) \to F(x)$ kiekviename F funkcijos tolydumo taške x.

Teisingas ir atvirkščias teiginys: jei $F_n(x) \to F(x)$ su kiekvienu x, kuriame F tolydi, tai $X_n \to X$. Jis įrodomas panašiai kaip $5 \Rightarrow 1$ implikacija Portmanto lemoje. Aš tą faktą praleidau, nes toliau tikiuosi apseiti be pasiskirstymo funkcijų.

2.3 Konvergavimas į išsigimusį dydį

2.3.1 Teoremos

2.2 teorema. *Tokie du teiginiai ekvivalentūs:*

1) $X_n \to a$*;*

2) $\mathrm{P}\{\|X_n - a\| \geqslant \varepsilon\} \to 0$ *su visais* ε*.*

Įrodymas. ($1 \Rightarrow 2$) Tegu f yra tolydi neneigiama funkcija su tokiomis savybėmis: $f(a) = 0$, $f(x) = 1$, kai $\|x - a\| \geqslant \varepsilon$. Pavyzdžiui, galiu paimti

$$f(x) = \left[\varepsilon^{-1}\|x - a\|\right] \wedge 1.$$

Tada

$$\mathrm{P}\{\|X_n - a\| \geqslant \varepsilon\} \leqslant \mathrm{E}\, f(X_n) \to f(a) = 0. \tag{2.8}$$

($2 \Rightarrow 1$) Tegu f yra bet kokia tolydi aprėžta funkcija. Randu tokį c, kad $|f(x)| \leqslant c$ su visais x, ir, fiksavęs ε, tokį δ, kad $\bigl|f(x) - f(a)\bigr| < \varepsilon$, kai $\|x - a\| < \delta$. Tada

$$\begin{aligned}
&\bigl|\mathrm{E}\, f(X_n) - f(a)\bigr| \\
&\leqslant \mathrm{E}\bigl|f(X_n) - f(a)\bigr| \\
&= \mathrm{E}\bigl|f(X_n) - f(a)\bigr| 1_{\{\|X_n - a\| < \delta\}} + \mathrm{E}\bigl|f(X_n) - f(a)\bigr| 1_{\{\|X_n - a\| \geqslant \delta\}} \\
&\leqslant \varepsilon + 2c\, \mathrm{P}\{\|X_n - a\| \geqslant \delta\} \\
&= \varepsilon + o(1).
\end{aligned} \tag{2.9}$$

Reiškia,

$$\overline{\lim}\bigl|\mathrm{E}\, f(X_n) - f(a)\bigr| \leqslant \varepsilon. \tag{2.10}$$

Gauta nelygybė teisinga su visais ε; todėl $\mathrm{E}\, f(X_n) \to f(a)$. $\square$

Iš įrodytos teoremos išplaukia ir tokie paprasti teiginiai:

1) $X_n \to a \iff \|X_n - a\| \to 0$;

2) jei $\|X_n - a\| \leqslant Y_n \to 0$, tai $X_n \to a$.

Įrodysiu dar du teiginius.

2.3 teorema. *1. Jei* $X_n \to X$ *ir* $\|X_n - Y_n\| \to 0$*, tai* $Y_n \to X$*.*

2. Jei $X_n \to X$*,* $Y_n \to a$*, tai* $(X_n, Y_n) \to (X, a)$*.*

Įrodymas. 1. Tegu f yra aprėžta Lipšico funkcija: $\left|f(x)\right| \leqslant c$ su visais x ir $\left|f(x) - f(y)\right| \leqslant q\|x - y\|$ su visais x, y. Kadangi

$$\mathrm{E}\, f(Y_n) - \mathrm{E}\, f(X) = \mathrm{E}\, f(Y_n) - \mathrm{E}\, f(X_n) + \mathrm{E}\, f(X_n) - \mathrm{E}\, f(X), \tag{2.11}$$

pakanka parodyti, kad $\mathrm{E}\, f(Y_n) - \mathrm{E}\, f(X_n) \to 0$. Tai išplaukia iš sąryšių

$$\left|\mathrm{E}\, f(Y_n) - \mathrm{E}\, f(X_n)\right| \leqslant \mathrm{E}\left|f(Y_n) - f(X_n)\right| \leqslant q\|Y_n - X_n\| \to 0. \tag{2.12}$$

2. Kadangi

$$\|(X_n, Y_n) - (X_n, a)\| = \|Y_n - a\| \to 0, \tag{2.13}$$

pakanka įrodyti, kad $(X_n, a) \to (X, a)$. Tai išplaukia iš tokio samprotavimo. Jei $f(x, y)$ yra aprėžta tolydi dviejų kintamųjų funkcija, tai $f(x, a)$ yra aprėžta tolydi vieno kintamojo funkcija ir, reiškia, $\mathrm{E}\, f(X_n, a) \to \mathrm{E}\, f(X, a)$. $\square$

2.3.2 Įrodymų paaiškinimai

Iš pradžių aiškinu 2.2 teoremos įrodymą.

25. Kodėl teisingi sąryšiai (2.8) grandinėje?

Pirma nelygybė teisinga dėl to, kad $1_{\{\|x-a\| \geqslant \varepsilon\}} \leqslant f(x)$, antras sąryšis — dėl to, kad f tolydi ir aprėžta, o $X_n \to a$. Paskutinė lygybė teisinga dėl f funkcijos parinkimo.

26. Kodėl egzistuoja toks δ?

Nes f tolydi a taške.

27. Kodėl teisingi sąryšiai (2.9) grandinėje?

Pirmas sąryšis išplaukia iš vidurkio savybių:

$$f(a) = \mathrm{E}\, f(a) \quad \text{ir} \quad \left|\mathrm{E}\, X - \mathrm{E}\, Y\right| \leqslant \mathrm{E}|X - Y|.$$

Antras — dėl to, kad

$$1_{\{\|X_n - a\| < \delta\}} + 1_{\{\|X_n - a\| \geqslant \delta\}} = 1$$

ir $\mathrm{E}(X + Y) = \mathrm{E}\, X + \mathrm{E}\, Y$.

Trečia nelygybė gauta taip: pirmame integrale įvertinau

$$\left|f(X_n) - f(a)\right| \leqslant \varepsilon$$

(taip galima daryti, nes $\left|f(x) - f(a)\right| < \varepsilon$, kai $\|x - a\| < \delta$, o pirmame integrale $\|X_n - a\| < \delta$); antrame integrale —

$$\left|f(X_n) - f(a)\right| \leqslant \left|f(X_n)\right| + \left|f(a)\right| \leqslant 2c$$

(tai teisėta, nes $\left|f(x)\right| \leqslant c$ su visais x).

Paskutinis sąryšis išplaukia iš teoremos 2 prielaidos.

28. Kodėl teisingas (2.10) sąryšis?

Suskaičiavau viršutines ribas (2.9) nelygybėje.

Dabar aiškinu 2.3 teoremos įrodymą.

29. Kodėl pakanka įrodyti, kad $\mathrm{E}\, f(Y_n) - \mathrm{E}\, f(X_n) \to 0$?

Jei tai įrodysiu, iš (2.11) išplauks, kad su bet kokia aprėžta Lipšico funkcija

$$\mathrm{E}\, f(Y_n) \to \mathrm{E}\, f(X)$$

(nes iš $X_n \to X$ išplaukia $\mathrm{E}\, f(X_n) \to f(X)$). Tada $Y_n \to X$ gausiu, pasirėmęs Portmanto lema.

30. Kodėl teisingi sąryšiai (2.12) grandinėje?

Pirmas sąryšis: įkėliau modulį į vdurkį; antras sąryšis: nes $\big|f(x) - f(y)\big| \leqslant q\|x - y\|$ su visais x ir y; trečias sąryšis: nes duota $\|X_n - Y_n\| \to 0$.

31. Kodėl teisingi sąryšiai (2.13) grandinėje?

Pirmas sąryšis: jei $x = (x_1, \ldots, x_k)$, $y = (y_1, \ldots, y_l)$ ir $a = (a_1, \ldots, a_l)$, tai

$$\begin{aligned}
&\|(x, y) - (x, a)\| \\
&= \sqrt{(x_1 - x_1)^2 + \cdots + (x_k - x_k)^2 + (y_1 - a_1)^2 + \cdots + (y_l - a_l)^2} \\
&= \sqrt{(y_1 - a_1)^2 + \cdots + (y_l - a_l)^2} \\
&= \|y - a\|.
\end{aligned}$$

Antras sąryšis išplaukia iš teoremos sąlygos $Y_n \to a$.

32. Kodėl pakanka įrodyti, kad $(X_n, a) \to (X, a)$?

Tą įrodžius, bus galima pasiremti 1 teoremos teiginiu su (X_n, a) vietoje X_n, (X, a) vietoje X ir (X_n, Y_n) vietoje Y_n.

2.3.3 Komentarai

33. Jei X neišsigimęs, tai sąryšis $X_n \to X$ apskritai nėra ekvivalentus sąryšiui

$$\forall \varepsilon \ \ \mathrm{P}\{\|X_n - X\| \geqslant \varepsilon\} \to 0, \tag{2.14}$$

jau vien dėl to, kad pirmojo sąryšio apibrėžime figūruoja tik atskirų dydžių skirstiniai, o antrasis išreiškiamas per (X_n, X) vektoriaus bendrą skirstinį.

Pavyzdžiui, jei visi X_n yra nepriklausomi ir vienodai pasiskirstę, tai $X_n \to X_1$, nes $\mathrm{E}\, f(X_n) = \mathrm{E}\, f(X_1)$ su visais n. Kita vertus,

$$\mathrm{P}\{\|X_n - X_1\| \geqslant \varepsilon\} = \mathrm{P}\{\|X_2 - X_1\| \geqslant \varepsilon\}$$

su visais $n \geqslant 2$ ir jei ta tikimybė lygi 0 su visais ε, tai $X_2 = X_1$, o tai neįmanoma, jei X_1 ir X_2 nepriklausomi ir neišsigimę.

34. Jei teisinga (2.14) formulė, sakoma, kad X_n konverguoja prie X *pagal tikimybę*, ir rašoma $X_n \xrightarrow{P} X$. Taigi $X_n \to a \iff X_n \xrightarrow{P} a$, bet apskritai $X_n \to X$ ir $X_n \xrightarrow{P} X$ sąryšiai nėra ekvivalentūs. Aš ruošiuosi išsiversti be konvergavimo pagal tikimybę sąvokos.

2.4 Tolydaus atvaizdžio principas

2.4.1 Teorema

2.4 teorema. *Tegu $X_n \to X$, o φ_n ir φ yra funkcijos iš $\mathbb{R}^k$ į $\mathbb{R}^l$. Jei egzistuoja tokia Borelio aibė $A \subset \mathbb{R}^k$, kad*

$$\mathrm{P}\{X \in A\} = 1,$$

$$x_n \to x \in A \Rightarrow \varphi_n(x_n) \to \varphi(x),$$

tai $\varphi_n(X_n) \to \varphi(X)$.

Įrodymas. Iš pradžių pastebiu, kad jei $x \in A$, tai

$$\forall \varepsilon\ \exists \delta\ \exists k\ \forall n \geqslant k\ \forall y \in U(x,\delta)\ \big|\big|\varphi_n(y) - \varphi(x)\big|\big| < \varepsilon. \tag{2.15}$$

Tikrai, priešingu atveju atsirastų toks ε, kad

$$\forall \delta\ \forall k\ \exists n \geqslant k\ \exists y \in U(x,\delta)\ \big|\big|\varphi_n(y) - \varphi(x)\big|\big| \geqslant \varepsilon.$$

Imdamas $\delta = 1/k$, rasčiau tokius $n_k \geqslant k$ ir $y_k \in U(x, 1/k)$, kad

$$\big|\big|\varphi_{n_k}(y_k) - \varphi(x)\big|\big| \geqslant \varepsilon$$

su visais k. Tegu $x_n = y_k$, kai $n = n_k$, ir $x_n = x$ su likusiais n. Tada $x_n \to x$, bet $\varphi_{n_k}(x_{n_k}) \not\to \varphi(x)$, o tai prieštarauja A aibės apibrėžimui.

Tegu dabar O yra bet koks atviras $\mathbb{R}^l$ poaibis. Jei $x \in A$ ir $\varphi(x) \in O$, tai iš (2.15) išplaukia, jog egzistuoja toks δ ir toks k, kad su visais $y \in U(x,\delta)$

$$\forall n \geqslant k\ \varphi_n(y) \in O;$$
$$\forall n \geqslant k\ y \in \varphi_n^{-1}(O);$$
$$y \in \bigcap_{n\geqslant k} \varphi_n^{-1}(O).$$

Reiškia, egzistuoja toks k, kad x yra vidinis aibės $\bigcap_{n\geqslant k} \varphi_n^{-1}(O)$ taškas, t.y

$$x \in \bigcup_{k\geqslant 1} \operatorname{Int} \bigcap_{n\geqslant k} \varphi_n^{-1}(O). \tag{2.16}$$

Iš įrodyto teiginio išplaukia, kad

$$\mathrm{P}\{\varphi(X) \in O\} \leqslant \mathrm{P}\Big\{X \in \bigcup_{k\geqslant 1} \operatorname{Int} \bigcap_{n\geqslant k} \varphi_n^{-1}(O)\Big\}. \tag{2.17}$$

Kai k didėja, aibės $\operatorname{Int}\bigcap_{n\geqslant k} \varphi_n^{-1}(O)$ didėja; todėl fiksavęs ε galiu rasti tokį k, kad

$$\mathrm{P}\{\varphi(X) \in O\} \leqslant \mathrm{P}\{X \in \operatorname{Int} \bigcap_{n\geqslant k} \varphi_n^{-1}(O)\} + \varepsilon. \tag{2.18}$$

Iš $X_n \to X$ dabar išplaukia

$$\begin{aligned}
\mathrm{P}\{\varphi(X) \in O\} &\leqslant \varliminf_{n\to\infty} \mathrm{P}\{X_n \in \operatorname{Int} \bigcap_{n\geqslant k} \varphi_n^{-1}(O)\} + \varepsilon \\
&\leqslant \varliminf_{n\to\infty} \mathrm{P}\{X_n \in \bigcap_{n\geqslant k} \varphi_n^{-1}(O)\} + \varepsilon \\
&\leqslant \varliminf_{n\to\infty} \mathrm{P}\{X_n \in \varphi_n^{-1}(O)\} + \varepsilon \\
&= \varliminf_{n\to\infty} \mathrm{P}\{\varphi_n(X_n) \in O\} + \varepsilon
\end{aligned} \tag{2.19}$$

Kadangi ε pasirinkau laisvai, iš čia gaunu

$$\varliminf_{n\to\infty} \mathrm{P}\{\varphi_n(X_n) \in O\} \geqslant \mathrm{P}\{\varphi(X) \in O\}.$$

Nelygybė teisinga su bet kokia atvira O; todėl $\varphi_n(X_n) \to \varphi(X)$. $\square$

2.4.2 Įrodymo paaiškinimai

35. Kodėl $x_n \to x$ ir $\varphi_{n_k}(x_{n_k}) \not\to \varphi(x)$ ir kur čia prieštara?

$x_n \to x$, nes $x_n \in U(x, 1/n)$, t.y. $\|x_n - x\| < 1/n$ su visais n.

$\varphi_{n_k}(x_{n_k}) \not\to \varphi(x)$, nes $\left|\varphi_{n_k}(x_{n_k}) - \varphi(x)\right| \geqslant \varepsilon$ su visais k.

Iš pastarojo sąryšio išplaukia $\varphi_n(x_n) \not\to \varphi(x)$. Reiškia, x negali priklausyti A aibei, nes jei priklausytų, iš $x_n \to x$ išplauktų $\varphi_n(x_n) \to \varphi(x)$. Gavau prieštarą, nes x paėmiau iš A.

36. Kaip iš (2.15) išplaukia, jog egzistuoja tokie δ ir k, kad $\forall n \geqslant k\ \varphi_n(y) \in O$ su visais $y \in U(x, \delta)$?

Kadangi O atvira, o $\varphi(x) \in O$, egzistuoja toks ε, kad $U(\varphi(x), \varepsilon) \subset O$. Pasirėmęs (2.15), randu tokius δ ir k, kad $\varphi_n(y) \in U(\varphi(x), \varepsilon)$ su visais $n \geqslant k$ ir $y \in U(x, \delta)$. Tada su visais $y \in U(x, \delta)$

$$\forall n \geqslant k\ \varphi_n(y) \in U(\varphi(x), \varepsilon) \subset O.$$

37. Kodėl teisingas (2.16) sąryšis?

Jei $y \in \bigcap_{n \geqslant k} \varphi_n^{-1}(O)$ su visais $y \in U(x, \delta)$, tai

$$U(x, \delta) \subset \bigcap_{n \geqslant k} \varphi_n^{-1}(O).$$

Reiškia, x yra vidinis $\bigcap_{n \geqslant k} \varphi_n^{-1}(O)$ taškas, t.y. priklauso tos aibės vidui. Reiškia, egzistuoja toks k, kad

$$x \in \operatorname{Int}\Big(\bigcap_{n \geqslant k} \varphi_n^{-1}(O)\Big).$$

Reiškia, x priklauso visų tokių aibių junginiui.

38. Kodėl teisingas (2.17) sąryšis?

Įrodžiau, kad iš $x \in A$ ir $\varphi(x) \in O$ išplaukia (2.16). Reiškia, iš $\varphi(x) \in O$ išplaukia (2.16) arba $x \notin A$. Taigi

$$\mathrm{P}\{\varphi(X) \in O\} \leqslant \mathrm{P}\Big\{X \in \bigcup_{k \geqslant 1} \operatorname{Int} \bigcap_{n \geqslant k} \varphi_n^{-1}(O)\Big\} + \mathrm{P}\{X \notin A\}.$$

Tačiau pagal vieną iš teoremos sąlygų $\mathrm{P}\{X \notin A\} = 0$.

39. Kodėl aibės $\operatorname{Int} \bigcap_{n \geqslant k} \varphi_n^{-1}(O)$ didėja?

Kuo daugiau aibių sukertu, tuo mažesnę aibę gaunu. Reiškia,

$$\operatorname{Int} \bigcap_{n \geqslant k} \varphi_n^{-1}(O) \subset \operatorname{Int} \bigcap_{n \geqslant k+1} \varphi_n^{-1}(O).$$

Kita vertus, kuo didesnė aibė, tuo didesnis jos vidus.

40. Kodėl su tam tikru k teisinga (2.18) formulė? Lebego teorema apie monotonišką konvergavimą sako, kad jei $A_k \uparrow A$ (t.y. jei A_k aibės didėja ir $A = \bigcup_k A_k$), tai $\mathrm{P}(A_k) \uparrow \mathrm{P}(A)$. Reiškia, su tam tikru k

$$\mathrm{P}(A_k) \geqslant \mathrm{P}\Big(\bigcup_k A_k\Big) - \varepsilon.$$

Aš panaudoju šį faktą su $\{X \in \operatorname{Int} \bigcap_{n \geqslant k} \varphi_n^{-1}(O)\}$ vietoje A_k.

41. Kodėl teisingi sąryšiai (2.19) grandinėje?

Pirma nelygybė: suskaičiavau apatines ribas (2.17) nelygybėje ir pasirėmiau Portmanto lemos 1 $\Rightarrow$ 3 implikacija su $\operatorname{Int} \bigcap_{n \geqslant k} \varphi_n^{-1}(O)$ vietoje O (taip daryti turiu teisę, nes bet kokios aibės vidus yra atvira aibė).

Antra nelygybė: pasirėmiau tuo, kad $\operatorname{Int} A \subset A$ su bet kokia A.

Trečia nelygybė: pasirėmiau tuo, kad $\bigcap_{n \geqslant k} A_n \subset A_n$, kai $n \geqslant k$ (o nuo pirmų sekos narių apatinė riba nepriklauso).

Paskutinis sąryšis: pagal pirmavaizdžio apibrėžimą

$$x \in \varphi_n^{-1}(O) \iff \varphi_n(x) \in O.$$

42. Kodėl įrodymo gale daroma išvada $\varphi_n(X_n) \to \varphi(X)$?

Pasirėmiau Portmanto lema.

2.4.3 Komentarai

43. Kai $\varphi_n = \varphi$ su visais n, sąlyga $x_n \to x \in A \Rightarrow \varphi_n(x_n) \to \varphi(x)$ reiškia, kad kiekviename taške $x \in A$ funkcija φ tolydi. Todėl iš įrodytos teoremos išplaukia tokia išvada: jei $X_n \to X$, $\mathrm{P}\{X \in A\} = 1$ ir φ tolydi A aibėje, tai $\varphi(X_n) \to \varphi(X)$.

Jei $T_n \to T$, $U_n^{(1)} \to u_1, \ldots, U_n^{(k)} \to u_k$, tai iš praeito skyrelio rezultatų išplaukia, kad

$$(T_n, U_n^{(1)}, \ldots, U_n^{(k)}) \to (T, u_1, \ldots, u_k).$$

Todėl

$$\varphi(T_n, U_n^{(1)}, \ldots, U_n^{(k)}) \to \varphi(T, u_1, \ldots, u_k),$$

jei φ tolydi kiekviename taške $(t, u_1, \ldots, u_k)$ su $t \in A$; čia A — tokia aibė, kad $\mathrm{P}\{T \in A\} = 1$. Tai ir yra tolydaus atvaizdžio principas, kurį naudojome pirmame skyriuje.

44. Iš įrodytos teoremos galima gauti ir delta metodo pagrindimą. Tegu $b_n \to \infty$, $b_n(T_n - a) \to T$, o g — diferencijuojama a taške funkcija. Įrodysiu, kad tada $b_n(g(T_n) - g(a)) \to g'(a)T$.

Kairioji šio sąryšio pusė lygi $\varphi_n(U_n)$ su $U_n = b_n(T_n - a)$ ir tam tikra funkcija φ_n. Kad rasčiau φ_n, išsireiškiu

$$T_n = a + b_n^{-1} U_n;$$

tada

$$b_n\big(g(T_n) - g(a)\big) = b_n\big(g(a + b_n^{-1} U_n) - g(a)\big).$$

Taigi

$$\varphi_n(u) = b_n\big(g(a + b_n^{-1} u) - g(a)\big).$$

Jei $u_n \to u$, tai

$$\varphi_n(u_n) = \frac{g(a + u_n/b_n) - g(a)}{u_n/b_n} u_n \to g'(a)u.$$

Iš tolydaus atvaizdžio principo tada išplaukia $\varphi_n(U_n) \to g'(a)T$.

2.5 O-simbolika

Matematinėje analizėje plačiai naudojami o ir O simboliai. Primenu, kad $o(1)$ simboliu žymimos artėjančios į 0 sekos, o $O(1)$ simboliu — aprėžtos sekos. Toje pat formulėje gali būti keli o ir O ženklai ir tada jie žymi nebūtinai tą pačią seką. Jei rašome lygybę, kurios abiejose pusėse yra o ir O simboliai, suprantame, kad bet kokia seką, kurią galima užrašyti kairėje pusėje parašytu pavidalu, galima užrašyta ir dešinės pusės pavidalu. Pavyzdžiui,

$$O(1)o(1) = o(1)$$

lygybė reiškia, kad aprėžtos ir artėjančios į 0 sekos sandauga artėja į 0.

Stochastinėje analizėje naudojami panašūs o_P ir O_P simboliai. Simboliu $o_P(1)$ žymima bet kokia atsitiktinių dydžių (arba vektorių) seka, artėjanti pagal tikimybę (= pagal pasiskirstymą) į 0. Simboliu $O_P(1)$ žymima bet kokia *aprėžta pagal tikimybę* atsitiktinių dydžių (vektorių) seka, t.y. bet kokia seka (X_n), kuriai

$$\sup_n \mathrm{P}\{\|X_n\| \geqslant k\} \xrightarrow[k\to\infty]{} 0.$$

Įrodysiu keletą teiginių su naujais simboliais.

1. $o_P(1)+o_P(1) = o_P(1)$. Kitaip tariant, jei $X_n \to 0$, $Y_n \to 0$, tai $X_n+Y_n \to 0$. Tai išplaukia iš tolydaus atvaizdžio principo.

2. Jei $X_n \to X$, tai $X_n = O_P(1)$. Fiksuoju ε ir randu tokį c, kad

$$\mathrm{P}\{\|X\| \geqslant c\} < \varepsilon \quad \text{ir} \quad \mathrm{P}\{\|X\| = c\} = 0.$$

Tada $\mathrm{P}\{\|X_n\| \geqslant c\} \to \mathrm{P}\{\|X\| \geqslant c\}$; todėl atsiras toks n_1, kad su visais $n \geqslant n_1$

$$\mathrm{P}\{\|X_n\| \geqslant c\} < \varepsilon.$$

Dabar su kiekvienu $i = 1, \ldots, n_1 - 1$ randu tokį c_i, kad

$$\mathrm{P}\{\|X_n\| \geqslant c_i\} < \varepsilon.$$

Tegu k yra bet koks skaičius didesnis už c ir visus c_i. Tada

$$\mathrm{P}\{\|X_n\| \geqslant k\} < \varepsilon$$

su visais n, t.y.

$$\sup_n \mathrm{P}\{\|X_n\| \geqslant k\} \leqslant \varepsilon.$$

Šį teiginį naudojau pirmame šio skyriaus skyrelyje, norėdamas įrodyti, kad kokia nors atsitiktinių dydžių seka neturi ribos pagal pasiskirstymą. Iš jo taip pat išplaukia $o_P(1) = O_P(1)$ lygybė, t.y. kad kiekviena artėjanti į 0 atsitiktinių vektorių seka aprėžta pagal tikimybę.

3. $O_P(1) + O_P(1) = O_P(1)$. Kitaip tariant, dviejų aprėžtų pagal tikimybę sekų suma vėl yra aprėžta pagal tikimybę. Tegu $X_n = O_P(1)$ ir $Y_n = O_P(1)$. Fiksuoju ε ir randu tokius k_1 ir k_2, kad

$$\sup_n \mathrm{P}\{\|X_n\| \geqslant k_1\} < \varepsilon \quad \text{ir} \quad \sup_n \mathrm{P}\{\|X_n\| \geqslant k_2\} < \varepsilon.$$

Tada su visais $k \geqslant k_1 + k_2$ ir visais n

$$\mathrm{P}\{\|X_n + Y_n\| \geqslant k\} \leqslant \mathrm{P}\{\|X_n\| \geqslant k_1\} + \mathrm{P}\{\|Y_n\| \geqslant k_2\} < 2\varepsilon.$$

Jei bent viena iš sekų skaliarinė, analogiškai įrodoma ir $O_P(1)O_P(1) = O_P(1)$ lygybė.

4. Jei viena iš sekų skaliarinė, tai $O_P(1)o_P(1) = o_P(1)$. Tegu, pavyzdžiui, $X_n = O_P(1)$, $Y_n = o_P(1)$ ir X_n seka skaliarinė. Fiksuoju ε ir randu tokį k, kad

$$\sup_n \mathrm{P}\{|X_n| \geqslant k\} < \varepsilon.$$

Tada su bet kokiu δ

$$\mathrm{P}\{\|X_n Y_n\| \geqslant \delta\} \leqslant \mathrm{P}\{|X_n| \geqslant k\} + \mathrm{P}\{\|Y_n\| \geqslant \delta/k\} \leqslant \varepsilon + o(1).$$

Taigi

$$\varlimsup_{n\to\infty} \mathrm{P}\{\|X_n Y_n\| \geqslant \delta\} \leqslant \varepsilon.$$

Kadangi ε pasirinkau laisvai, iš čia išplaukia $\mathrm{P}\{\|X_n Y_n\| \geqslant \delta\} \to 0$.

3 skyrius

Asimptotiniai skirstiniai (tęsinys)

3.1 Vektorinis delta metodas

3.1.1 Matricinė notacija

Vektorių $x = (x_1, \dots, x_k)$ aš mėgstu užrašyti kaip vieno stulpelio $k \times 1$ matricą:

$$x = (x_1, \dots, x_k) = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}.$$

Vienos eilutės matrica yra visai kitas objektas:

$$(x_1, \dots, x_k) \neq \begin{pmatrix} x_1 & \cdots & x_k \end{pmatrix}.$$

Skirtumas tarp simbolių kairėje ir dešinėje tas, kad vektoriaus komponentes aš vieną nuo kitos atskiriu kableliais, o $1 \times k$ matricos komponentes — tarpais.

Jei $g(x_1, \dots, x_k)$ yra k kintamųjų realioji funkcija, tai jos išvestine vadinama $1 \times k$ matrica, sudaryta iš tos funkcijos dalinių išvestinių:

$$g'(x) = \begin{pmatrix} \frac{\partial g}{\partial x_1}(x) & \dots & \frac{\partial g}{\partial x_k}(x) \end{pmatrix}.$$

Analogiška $k \times 1$ matrica (t.y. k-matis vektorius) vadinamas funkcijos gradientu ir žymimas ∇g:

$$\nabla g(x) = \left(\frac{\partial g}{\partial x_1}(x), \dots, \frac{\partial g}{\partial x_k}(x) \right) = \begin{pmatrix} \frac{\partial g}{\partial x_1}(x) \\ \vdots \\ \frac{\partial g}{\partial x_k}(x) \end{pmatrix}.$$

Pavyzdžiui, jei

$$g(x, y, z) = \frac{x}{y - z^2},$$

tai

$$g' = \begin{pmatrix} \frac{1}{y-x^2} & -\frac{x}{(y-z^2)^2} & \frac{2xz}{(y-z^2)^2} \end{pmatrix}, \quad \nabla g \begin{pmatrix} \frac{1}{y-x^2} \\ -\frac{x}{(y-z^2)^2} \\ \frac{2xz}{(y-z^2)^2} \end{pmatrix}.$$

Jei $g(x) = g(x_1, \dots, x_k)$ yra k kintamųjų vektorinė funkcija (su reikšmėmis $\mathbb{R}^l$ erdvėje), tai ji yra

$$\big(g_1(x), \dots, g_l(x)\big) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_l(x) \end{pmatrix}$$

pavidalo; čia $g_j(x)$ yra skaliarinės funkcijos, vadinamos g funkcijos komponentėmis. Tokios funkcijos išvestinė yra $l \times k$ matrica

$$g'(x) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x) & \cdots & \frac{\partial g_1}{\partial x_k}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_l}{\partial x_1}(x) & \cdots & \frac{\partial g_l}{\partial x_k}(x) \end{pmatrix}.$$

3.1.2 Atsitiktinių matricų vidurkiai

Jei X yra atsitiktinė matrica (atskiru atveju — $k \times 1$ matrica, t.y. vektorius), jos vidurkiu vadinama skaičių matrica, sudaryta iš atitinkamų X matricos komponenčių vidurkių. Taigi, pavyzdžiui,

$$\mathrm{E}\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} \mathrm{E}\,X_1 \\ \mathrm{E}\,X_2 \\ \mathrm{E}\,X_3 \end{pmatrix}, \quad \mathrm{E}\begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} = \begin{pmatrix} \mathrm{E}\,X_{11} & \mathrm{E}\,X_{12} \\ \mathrm{E}\,X_{21} & \mathrm{E}\,X_{22} \end{pmatrix}$$

ir pan.

Pagrindinės vektorinių vidurkių savybės tokios pat, kaip skaliarinių vidurkių: konstantos vidurkis lygus jai pačiai, sumos vidurkis lygus vidurkių sumai, pastovų daugiklį galima iškelti už vidurkio ženklo (kadangi matricų daugyba nekomutatyvi, pastaroji savybė užrašoma dviem lygybėmis):

$$\mathrm{E}\,C = C, \quad \mathrm{E}\,X + Y = \mathrm{E}\,X + \mathrm{E}\,Y, \quad \mathrm{E}\,AX = A\,\mathrm{E}\,X, \quad \mathrm{E}\,XB = (\mathrm{E}\,X)B.$$

Transponuotos matricos vidurkis gaunamas transponuojant vidurkių matricą:

$$\mathrm{E}\,X^\top = (\mathrm{E}\,X)^\top.$$

Dispersijos vaidmenį daugiamatėje teorijoje atlieka *kovariacijų matrica*, kurią aš žymėsiu V simboliu: jei $X = (X_1, \dots, X_k)$ yra atsitiktinis vektorius, tai

$$\mathrm{V}(X) = \mathrm{E}(X - \mathrm{E}\,X)(X - \mathrm{E}\,X)^\top = \mathrm{E}(XX^\top) - (\mathrm{E}\,X)(\mathrm{E}\,X)^\top.$$

Kovariacijų matricos elementai yra kovariacijos tarp X vektoriaus komponenčių:

$$\mathrm{V}(X) = \begin{pmatrix} \mathrm{cov}(X_1, X_1) & \cdots & \mathrm{cov}(X_1, X_k) \\ \vdots & \ddots & \vdots \\ \mathrm{cov}(X_k, X_1) & \cdots & \mathrm{cov}(X_k, X_k) \end{pmatrix};$$

čia

$$\mathrm{cov}(X_i, X_j) = \mathrm{E}(X_i - \mathrm{E}\,X_i)(X_j - \mathrm{E}\,X_j) = \mathrm{E}\,X_iX_j - \mathrm{E}\,X_i\,\mathrm{E}\,X_j.$$

Iš vidurkio savybių nesunkiai gaunamos tokios kovariacijų matricos savybės:

$$\mathrm{V}(X + C) = \mathrm{V}\,X, \quad \mathrm{V}(AX) = A\,\mathrm{V}(X)A^\top.$$

3.1.3 Normalusis skirstinys

Simboliu $N(\mu, \Sigma)$ aš žymiu atsitiktinį vektorių, pasiskirsčiusį normaliai su vidurkiu μ ir kovariacijų matrica Σ. Gerai žinoma, kad jei X turi tokį skirstinį, tai $A + BX$ taip pat pasiskirstęs normaliai, tik su kitais parametrais: vidurkiu $A + B\mu$ ir kovariacijų matrica $B\Sigma B^\top$. Todėl aš dažnai rašau

$$A + BN(\mu, \Sigma) = N(A + B\mu, B\Sigma B^\top).$$

Dažniausiai $\mu = 0$, $A = 0$, o B yra vienos eilutės matrica; tada $B\Sigma B^\top$ yra skaičius: jei $\Sigma = (\sigma_{ij} \mid 1 \leqslant i, j \leqslant k)$, $B = (b_1 \ \cdots \ b_k)$, tai

$$BN(0, \Sigma) = N(0, \sigma^2);$$

čia

$$\sigma^2 = B\Sigma B^\top = \sum_{i,j=1}^{k} \sigma_{ij} b_i b_j$$

3.1.4 Empiriniai vidurkiai

Tegu $X_1, \ldots, X_n$ yra nepriklausomos atsitiktinio k-mačio vektoriaus X kopijos, f — funkcija iš $\mathbb{R}^k$ į $\mathbb{R}^l$ ir f_j — tos funkcijos komponentės. Tada simboliu $\overline{f(X)}$ žymėsiu empirinį vidurkį:

$$\overline{f(X)} = \frac{1}{n} \sum_{i=1}^{n} f(X_i).$$

Aišku, kad $\overline{f(X)}$ yra l-matis vektorius, kurio komponentės yra empiriniai vidurkiai, atitinkantys f funkcijos komponentes:

$$\overline{f(X)} = \big(\overline{f_1(X)}, \ldots, \overline{f_l(X)}\big) = \begin{pmatrix} \overline{f_1(X)} \\ \vdots \\ \overline{f_l(X)} \end{pmatrix}.$$

Dažniausiai susidursime su situacija, kai X yra atsitiktinis dydis (skaliaras), o f — vektorinė funkcija iš $\mathbb{R}$ į $\mathbb{R}^l$. Pavyzdžiui, jei $f(x) = (x, x^2)$, tai

$$\overline{f(X)} = \overline{(X, X^2)} = \big(\overline{X}, \overline{X^2}\big) = \begin{pmatrix} \overline{X} \\ \overline{X^2} \end{pmatrix} = \begin{pmatrix} \frac{1}{n} \sum_{i=1}^n X_i \\ \frac{1}{n} \sum_{i=1}^n X_i^2 \end{pmatrix} = \frac{1}{n} \sum_{i=1}^{n} \begin{pmatrix} X_i \\ X_i^2 \end{pmatrix}.$$

Pagrindinės empirinio vidurkio savybės panašios į paprasto vidurkio: konstantos vidurkis lygus tai konstantai, sumos vidurkis lygus vidurkių sumai, pastovų daugiklį galima iškelti iš vidurkio. Kitaip tariant,

$$\overline{c} = c, \quad \overline{f(X) + g(X)} = \overline{f(X)} + \overline{g(X)}, \quad \overline{Af(X)} = A\overline{f(X)}, \quad \overline{f(X)B} = \overline{f(X)}B.$$

3.1.5 Asimptotikų ieškojimo metodai

Pagrindinių teiginių, kuriais remiamės ieškome asimptotinių skirstinių, analogai vektoriniu atveju skamba taip:

- (didžiųjų skaičių dėsnis) jei $\mathrm{E}\, X = a$, tai
$$\overline{X} \to a;$$
- (centrinė ribinė teorema) jei $\mathrm{E}\, X = a$, $\mathrm{V}(X) = \Sigma$, tai
$$\sqrt{n}(\overline{X} - a) \to N(0, \Sigma);$$
- (tolydaus atvaizdžio principas) jei $T_n \to T$, $U_n^{(1)} \to u_1, \ldots, U_n^{(k)} \to u_k$, tai
$$\varphi(T_n, U_n^{(1)}, \ldots, U_n^{(k)}) \to \varphi(T, u_1, \ldots, u_k);$$
- (delta metodas) jei $b_n \to \infty$ ir $b_n(T_n - a) \to T$, tai
$$b_n\big(g(T_n) - g(a)\big) \to g'(a)T.$$

Pakomentuosiu paskutinį teiginį. Vektorinio delta metodo aprašymas nesiskiria nuo skaliarinio, bet tie patys simboliai dabar žymi nebe skaliarinius, o vektorinius dydžius (arba matricas). Tegu g yra skaliarinė k kintamųjų funkcija $g(t_1, \ldots, t_k)$ (šis atvejis bus dažniausiai sutinkamas). Tada T_n turi būti k-matis atsitiktinis vektorius:
$$T_n = \begin{pmatrix} T_{n1} \\ \vdots \\ T_{nk} \end{pmatrix}.$$
Sąryšis $b_n(T_n - a)$ paprastai gaunamas iš vektorinės centrinės ribinės teoremos; tada T yra normaliai pasiskirstęs atsitiktinis vektorius — koks nors $N(0, \Sigma)$. Simbolis $g'(a)$ žymi g funkcijos dalinių išvestinių matricą; nagrinėjamu atveju ji sudaryta iš vienos eilutės. Tada
$$g'(a)T = g'(a)N(0, \Sigma) = N(0, \sigma^2)$$
su
$$\sigma^2 = g'(a)\Sigma g'(a)^\top = \sum_{i,j=1}^{k} \sigma_{ij}\partial_i\partial_j;$$
čia σ_{ij} yra Σ matricos komponentės, o
$$\partial_i = \frac{\partial g}{\partial t_i}(a).$$

3.1.6 Pavyzdžiai

3.1 PAVYZDYS. Tegu $X_1, \ldots, X_n$ yra imtis iš skirstinio

X	1	2	3
	0.2	0.3	0.5

Raskite statistikos
$$T_n = \frac{N_3 - N_1}{\sqrt{N_1 N_3}}$$
asimptotinį skirstinį; čia N_j žymi j reikšmės pasikartojimų skaičių imtyje.

Sprendimas. Turime:

$$T_n = \frac{N_3/n - N_1/n}{\sqrt{N_3/n \cdot N_1/n}} = g\big(\overline{1_{\{X=1\}}}, \overline{1_{\{X=3\}}}\big);$$

čia

$$g(u,v) = \frac{v-u}{\sqrt{uv}}.$$

Taigi iš pradžių reikia rasti vektorinės sekos

$$\big(\overline{1_{\{X=3\}}}, \overline{1_{\{X=1\}}}\big) = \overline{(1_{\{X=3\}}, 1_{\{X=1\}})}$$

asimptotiką, o po to pritaikyti vektorinį delta metodą.

Kadangi

$$\begin{aligned} \mathrm{E}\, 1^2_{\{X=1\}} &= \mathrm{E}\, 1_{\{X=1\}} = 0.2, \\ \mathrm{E}\, 1^2_{\{X=3\}} &= \mathrm{E}\, 1_{\{X=3\}} = 0.5, \\ \mathrm{E}\, 1_{\{X=1\}} 1_{\{X=3\}} &= 0, \end{aligned}$$

gaunu

$$\begin{aligned} \mathrm{cov}(1_{\{X=1\}}, 1_{\{X=1\}}) &= 0.2 - 0.04 = 0.16, \\ \mathrm{cov}(1_{\{X=3\}}, 1_{\{X=3\}}) &= 0.5 - 0.25 = 0.25, \\ \mathrm{cov}(1_{\{X=1\}}, 1_{\{X=3\}}) &= 0 - 0.1 = -0.1. \end{aligned}$$

Reiškia,

$$\sqrt{n}\left[\begin{pmatrix} \overline{1_{\{X=1\}}} \\ \overline{1_{\{X=3\}}} \end{pmatrix} - \begin{pmatrix} 0.2 \\ 0.5 \end{pmatrix}\right] \to N(0,\Sigma);$$

čia

$$\Sigma = \begin{pmatrix} 0.16 & -0.1 \\ -0.1 & 0.25 \end{pmatrix}.$$

Dabar taikau vektorinį delta metodą. Iš pradžių randu dalines g funkcijos išvestines:

$$\begin{aligned} \frac{\partial g}{\partial u} &= \frac{-\sqrt{uv} - (v-u)\frac{v}{2\sqrt{uv}}}{uv} = -\frac{uv+v^2}{2(uv)^{3/2}}, \\ \frac{\partial g}{\partial v} &= \frac{\sqrt{uv} - (v-u)\frac{u}{2\sqrt{uv}}}{uv} = \frac{uv+u^2}{2(uv)^{3/2}}. \end{aligned}$$

Taigi

$$g(0.2, 0.5) = \frac{0.3}{\sqrt{0.1}} = 0.3\sqrt{10},$$

$$g'(0.2,0.5) = \frac{1}{2\cdot 0.1^{3/2}}\big(-0.1 - 0.25 \;\; 0.1 + 0.04\big) = \frac{1}{2\cdot 0.1^{3/2}}\big(-0.35 \;\; 0.14\big)$$

ir pagal delta metodą

$$\sqrt{n}(T_n - 0.3\sqrt{10}) \to N(0,\sigma^2);$$

čia

$$\begin{aligned}\sigma^2 &= \frac{1}{4\cdot 0.1^3}\bigl(-0.35\ 0.14\bigr)\begin{pmatrix}0.16 & -0.1\\ -0.1 & 0.25\end{pmatrix}\begin{pmatrix}-0.35\\ 0.14\end{pmatrix}\\ &= 250\bigl(-0.35\ 0.14\bigr)\begin{pmatrix}-0.07\\ 0.07\end{pmatrix}\\ &= 250\cdot 0.0343\\ &= 8.575.\end{aligned}$$

3.2 PAVYZDYS. Tegu $X_1,\dots,X_n$ yra imtis iš tolygaus $(0;1)$ intervale skirstinio ir

$$T_n = \frac{\overline{X}}{\sqrt{\overline{X^2}-\overline{X}^2}}.$$

Raskite T_n statistikos asimptotinį skirstinį.

Sprendimas. Aišku, kad $T_n = g(\overline{X},\overline{X^2})$; čia

$$g(u,v) = \frac{u}{\sqrt{v-u^2}}.$$

Kadangi

$$\begin{aligned}\mathrm{E}\,X &= \int_0^1 x\mathrm{d}x = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2},\\ \mathrm{E}\,X^2 &= \int_0^1 x^2\mathrm{d}x = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3},\\ \mathrm{E}\,X^3 &= \int_0^1 x^3\mathrm{d}x = \frac{x^4}{4}\Big|_0^1 = \frac{1}{4},\\ \mathrm{E}\,X^4 &= \int_0^1 x^4\mathrm{d}x = \frac{x^5}{5}\Big|_0^1 = \frac{1}{5},\end{aligned}$$

gaunu

$$\begin{aligned}\mathrm{cov}(X,X) &= \frac{1}{3}-\frac{1}{4} = \frac{1}{12},\\ \mathrm{cov}(X,X^2) &= \frac{1}{4}-\frac{1}{6} = \frac{1}{12},\\ \mathrm{cov}(X^2,X^2) &= \frac{1}{5}-\frac{1}{9} = \frac{4}{45}.\end{aligned}$$

Todėl iš centrinės ribinės teoremos

$$\sqrt{n}\left[\begin{pmatrix}\overline{X}\\ \overline{X^2}\end{pmatrix} - \begin{pmatrix}1/2\\ 1/3\end{pmatrix}\right] \to N(0,\Sigma);$$

čia

$$\Sigma = \begin{pmatrix}1/12 & 1/12\\ 1/12 & 4/45\end{pmatrix}.$$

Randu g funkcijos išvestines:

$$\frac{\partial g}{\partial u} = \frac{\sqrt{v-u^2} - u\frac{-2u}{2\sqrt{v-u^2}}}{v-u^2} = \frac{v}{(v-u^2)^{3/2}},$$
$$\frac{\partial g}{\partial v} = -\frac{u}{2(v-u^2)^{3/2}}.$$

Taigi

$$g(1/2, 1/3) = \frac{1/2}{\sqrt{1/3-1/4}} = \sqrt{3},$$
$$g'(1/2, 1/3) = 12^{3/2}\big(1/3 \quad -1/4\big) = 2\sqrt{3}(4 \quad -3)$$

ir pagal delta metodą

$$\sqrt{n}(T_n - \sqrt{3}) \to N(0, \sigma^2);$$

čia

$$\begin{aligned} \sigma^2 &= 12(4 \quad -3)\begin{pmatrix} 1/12 & 1/12 \\ 1/12 & 4/45 \end{pmatrix}\begin{pmatrix} 4 \\ -3 \end{pmatrix} \\ &= 12(4 \quad -3)\begin{pmatrix} 1/12 \\ 1/15 \end{pmatrix} \\ &= 12 \cdot \frac{2}{15} \\ &= \frac{8}{5}. \end{aligned}$$

3.2 Ypatingi atvejai

3.2.1 Atvejis, kai $g'(a) = 0$

Primenu, kad užduotis rasti asimptotinį T_n sekos skirstinį reiškia, jog reikia rasti tokias skaičių sekas a_n ir b_n, kad $b_n(T_n - a_n)$ konverguotų į neišsigimusį atsitiktinį dydį T. Jei $T_n = g(U_n)$ ir $b_n(U_n - a) \to U$, tai taikydamas delta metodą gaunu

$$b_n\big(T_n - g(a)\big) \to g'(a)U.$$

Deja, dydis $g'(a)U$ išsigimęs, jei $g'(a) = 0$. Ką daryti tokiu atveju?

Tarkime, $g''(a) \neq 0$. Pažymėjęs

$$V_n = b_n(U_n - a),$$

gaunu

$$T_n = g(a + V_n/b_n) = g(a) + \frac{\varphi_n(V_n)}{b_n^2};$$

čia

$$\varphi_n(x) = b_n^2\big[g(a + x/b_n) - g(a)\big].$$

Iš Teiloro formulės

$$g(a+h) = g(a) + g'(a)h + \frac{1}{2}g''(a)h^2 + h^2 r(h) = g(a) + \frac{1}{2}g''(a)h^2 + h^2 r(h)$$

su tam tikra funkcija $r(h)$, artėjančia į 0, kai $h \to 0$. Taigi jei $x_n \to x$, tai

$$\varphi_n(x_n) = b_n^2\Big[\frac{1}{2}\big(g''(a) + r(x_n/b_n)\big)\frac{x_n^2}{b_n^2}\Big] = \frac{x_n^2}{2}\big(g''(a) + r(x_n/b_n)\big) \to \frac{g''(a)x^2}{2}.$$

Taigi iš tolydaus atvaizdžio principo

$$b_n^2\big(T_n - g(a)\big) = \varphi_n(V_n) \to \frac{g''(a)U^2}{2}.$$

Tai ir yra ieškomoji T_n asimptotika, jei $g''(a) \neq 0$.

Jei ir $g''(a) = 0$ tektų taikyti Teiloro formulę su didesniu skaičiumi narių. Bet praktikoje tokie uždaviniai pasitaiko retai.

Jei ne tik $g'(a) = 0$, bet ir $g(a) = 0$, tai asimptotinį skirstinį galima rasti žymiai greičiau, nenaudojant Teiloro formulės. Štai pavyzdys. Tegu $X_1, \dots, X_n$ yra imtis iš tolygaus $(-1; 1)$ intervale skirstinio ir $T_n = \overline{X}^2$. Reikia rasti (T_n) sekos asimptotiką.

Kadangi

$$\mathrm{E}\,X = \int_{-1}^{1}\frac{x}{2}\mathrm{d}x = 0,\ \mathrm{E}\,X^2 = \int_{-1}^{1}\frac{x^2}{2}\mathrm{d}x = \frac{1}{3},$$

dispersija $\mathrm{D}\,X = 1/3$ ir iš centrinės ribinės teoremos gaunu

$$\sqrt{n}\overline{X} \to N(0, 1/3).$$

Tada iš tolydaus atvaizdžio principo iškart

$$n\overline{X}^2 \to N^2(0, 1/3) = \Big[\frac{1}{\sqrt{3}}N(0,1)\Big]^2 = \frac{1}{3}\chi^2(1).$$

Kitaip tariant,

$$3nT_n \to \chi^2(1).$$

Jei būčiau taikęs delta metodą su funkcija $g(x) = x^2$, būčiau gavęs

$$\sqrt{n}\big(T_n - 0\big) \to 0 \cdot N(0,1),$$

t.y. $\sqrt{n}T_n \to 0$ (nes šiuo atveju $a = 0$, $g(0) = 0$ ir $g'(0) = 2 \cdot 0 = 0$). Tada „antros eilės delta metodas“ duotų tą patį atsakymą, kurį gavau iškart: $nT_n \to N^2(0, 1/3)$ (nes $g''(0)/2 = 1$).

3.2.2 Atvejis, kai g priklauso nuo n

Jei $T_n = g_n(U_n)$ ir $b_n(U_n - a) \to U$, kažkokios bendros teorijos nėra. Tenka tiesiog kartoti delta metodo pagrindimą ir uždavinį suvesti į tokį pavidalą, kad būtų galima taikyti tolydaus atvaizdžio principą. Todėl iškart panagrinėsiu pavyzdį.

Tegu $X_1, \dots, X_n$ yra imtis iš skirstinio

X	-1	0	1	2
	0.2	0.2	0.2	0.4

Reikia rasti statistikos

$$T_n = \frac{\sqrt{N^2+1}}{N+1}$$

asimptotinį skirstinį; čia N yra teigiamų narių skaičius imtyje.

Aišku, kad

$$T_n = \frac{\sqrt{(N/n)^2+1/n^2}}{N/n+1/n} = g_n(N/n);$$

čia

$$g_n(t) = \frac{\sqrt{t^2+1/n^2}}{t+1/n}.$$

Kadangi $N/n = \overline{1_{\{X>0\}}}$ ir

$$\mathrm{E}\,1_{\{X>0\}} = \mathrm{E}\,1^2_{\{X>0\}} = 0.6,$$

iš centrinės ribinės teoremos gaunu

$$U_n = \sqrt{n}(\overline{1_{\{X>0\}}} - 0.6) \to N(0, 0.24).$$

Be to, $\overline{1_{\{X>0\}}} = 0.6 + U_n/\sqrt{n}$ ir

$$T_n = g_n(0.6 + U_n/\sqrt{n}).$$

Randu reiškinio $g_n(0.6+u_n/\sqrt{n})$ asimptotiką, kai $u_n \to u$:

$$\begin{aligned}
&g_n(0.6+u_n/\sqrt{n})\\
&= \frac{\sqrt{0.36+1.2u_n/\sqrt{n}+u_n^2/n+1/n^2}}{0.6+u_n/\sqrt{n}+1/n}\\
&= \Big(1+\frac{10u_n}{3\sqrt{n}}+\frac{25u_n^2}{9n}+\frac{25}{9n^2}\Big)^{1/2}\Big(1+\frac{5u_n}{3\sqrt{n}}+\frac{5}{3n}\Big)^{-1}\\
&= \Big[1+\frac{1}{2}\Big(\frac{10u_n}{3\sqrt{n}}+\frac{25u_n^2}{9n}+\frac{25}{9n^2}\Big)-\frac{1}{8}\Big(\frac{10u_n}{3\sqrt{n}}+\frac{25u_n^2}{9n}+\frac{25}{9n^2}\Big)^2\\
&\qquad+\frac{1}{16}\Big(\frac{10u_n}{3\sqrt{n}}+\frac{25u_n^2}{9n}+\frac{25}{9n^2}\Big)^3+o(n^{-3/2})\Big]\\
&\quad\times\Big[1-\Big(\frac{5u_n}{3\sqrt{n}}+\frac{5}{3n}\Big)+\Big(\frac{5u_n}{3\sqrt{n}}+\frac{5}{3n}\Big)^2-\Big(\frac{5u_n}{3\sqrt{n}}+\frac{5}{3n}\Big)^3+o(n^{-3/2})\Big]\\
&= \Big[1+\frac{5u_n}{3\sqrt{n}}+o(n^{-3/2})\Big]\\
&\qquad\times\Big[1-\frac{5u_n}{3\sqrt{n}}-\frac{5}{3n}+\frac{25u_n^2}{9n}+\frac{50u_n}{9n^{3/2}}-\frac{125u_n^3}{27n^{3/2}}+o(n^{-3/2})\Big]\\
&= 1-\frac{5}{3n}+\frac{25u_n}{9n^{3/2}}+o(n^{-3/2}).
\end{aligned}$$

Kitaip tariant, jei $u_n \to u$, tai

$$n^{3/2}\Big(g_n(0.6+u_n/\sqrt{n})-1+\frac{5}{3n}\Big) \to \frac{25u}{9}.$$

Tada iš tolydaus atvaizdžio principo:

$$n^{3/2}\Big(T_n-1+\frac{5}{3n}\Big) \to \frac{25}{9}N(0,0.24) = N(0, 50/27).$$

3.3 Pavyzdžiai

3.3 PAVYZDYS. X ir Y yra nepriklausomi vienodai pasiskirstę atsitiktiniai dydžiai, kurių skirstinys aprašomas lentele

X, Y	1	2
	0.4	0.6

Tegu $(X_1, Y_1), \ldots, (X_n, Y_n)$ yra nepriklausomos (X, Y) kopijos, o N_{jk} žymi poros (j, k) pasikartojimų skaičių imtyje. Raskite nurodytų statistikų asimptotinius skirstinius:

a) $T_n = \dfrac{N_{12} - N_{21}}{N_{11} + N_{22}}$;

b) $T_n = \dfrac{N_{11}}{N_{11}^2 + N_{22}^2}$.

Sprendimas. a. Kadangi $N_{11} + N_{22} = n - N_{11} - N_{21}$, T_n statistiką galiu užrašyti taip:

$$T_n = \frac{N_{12} - N_{21}}{n - N_{12} - N_{21}} = \frac{N_{12}/n - N_{21}/n}{1 - N_{12}/n - N_{21}/n} = g(\overline{1_{\{X=1,Y=2\}}}, \overline{1_{\{X=2,Y=1\}}});$$

čia

$$g(u, v) = \frac{u - v}{1 - u - v}.$$

Bendras (X, Y) skirstinys aprašomas lentele

$X \quad Y$	1	2	
1	0.16	0.24	0.4
2	0.24	0.36	0.6
	0.4	0.6	

Iš jos matyti, kad

$$\mathrm{E}\,1_{\{X=1,Y=2\}} = \mathrm{E}\,1^2_{\{X=1,Y=2\}} = \mathrm{E}\,1_{\{X=2,Y=1\}} = \mathrm{E}\,1^2_{\{X=2,Y=1\}} = 0.24,$$
$$\mathrm{E}\,1_{X=1,Y=2} 1_{X=2,Y=1} = \mathrm{E}\,0 = 0.$$

Reiškia,

$$\mathrm{D}\,1_{\{X=1,Y=2\}} = \mathrm{D}\,1_{\{X=2,Y=1\}} = \frac{114}{625}, \quad \mathrm{cov}(1_{\{X=1,Y=2\}}, 1_{\{X=2,Y=1\}}) = -\frac{36}{625}.$$

Tada iš centrinės ribinės teoremos

$$\sqrt{n}\Big[\begin{pmatrix} \overline{1_{\{X=1,Y=2\}}} \\ \overline{1_{\{X=2,Y=1\}}} \end{pmatrix} - \begin{pmatrix} 6/25 \\ 6/25 \end{pmatrix}\Big] \to N(0, \Sigma);$$

čia

$$\Sigma = \begin{pmatrix} 114/625 & -36/625 \\ -36/625 & 114/625 \end{pmatrix}.$$

Taikau vektorinį delta metodą. Iš pradžių randu g funkcijos išvestines:

$$\frac{\partial g}{\partial u} = \frac{1 - u - v + (u - v)}{(1 - u - v)^2} = \frac{1 - 2v}{(1 - u - v)^2},$$

$$\frac{\partial g}{\partial v} = \frac{-(1-u-v)+(u-v)}{(1-u-v)^2} = \frac{-1+2u}{(1-u-v)^2}.$$

Įstatęs $u = v = 0.24$, gaunu $g(0.24, 0.24) = 0$ ir

$$g'(0.24, 0.24) = \begin{pmatrix} \frac{25}{13} & -\frac{25}{13} \end{pmatrix}.$$

Taigi

$$\sqrt{n}T_n \to N(0, \sigma^2);$$

čia

$$\sigma^2 = \begin{pmatrix} \frac{25}{13} & -\frac{25}{13} \end{pmatrix} \begin{pmatrix} 114/625 & -36/625 \\ -36/625 & 114/625 \end{pmatrix} \begin{pmatrix} \frac{25}{13} \\ -\frac{25}{13} \end{pmatrix}$$
$$= \frac{1}{169} \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} 150 \\ -150 \end{pmatrix} = \frac{300}{169}.$$

b. Turiu

$$nT_n = \frac{nN_{11}}{N_{11}^2 + N_{22}^2} = \frac{N_{11}/n}{(N_{11}/n)^2 + (N_{22}/n)^2} = g(\overline{1_{\{X=Y=1\}}}, \overline{1_{\{X=Y=2\}}});$$

čia

$$g(u, v) = \frac{u}{u^2+v^2}.$$

Iš (X, Y) skirstinio lentelės matyti, kad

$$\mathrm{E}\,1_{\{X=Y=1\}} = \frac{4}{25}, \quad \mathrm{D}\,1_{\{X=Y=1\}} = \frac{84}{625},$$
$$\mathrm{E}\,1_{\{X=Y=2\}} = \frac{9}{25}, \quad \mathrm{D}\,1_{\{X=Y=2\}} = \frac{144}{625},$$
$$\mathrm{cov}(1_{\{X=Y=1\}}, 1_{\{X=Y=2\}}) = -\frac{36}{625}.$$

Tada iš centrinės ribinės teoremos

$$\sqrt{n}\left[\begin{pmatrix} \overline{1_{\{X=Y=1\}}} \\ \overline{1_{\{X=Y=2\}}} \end{pmatrix} - \begin{pmatrix} 4/25 \\ 9/25 \end{pmatrix}\right] \to N(0, \Sigma);$$

čia

$$\Sigma = \begin{pmatrix} 84/625 & -36/625 \\ -36/625 & 144/625 \end{pmatrix}.$$

Taikau vektorinį delta metodą. Iš pradžių randu g funkcijos išvestines:

$$\frac{\partial g}{\partial u} = \frac{u^2+v^2-2u^2}{(u^2+v^2)^2} = \frac{v^2-u^2}{(u^2+v^2)^2},$$
$$\frac{\partial g}{\partial v} = -\frac{2v}{(u^2+v^2)^2}.$$

Įstatęs $u = 0.16$ ir $v = 0.36$, gaunu $g(0.16, 0.36) = \frac{100}{97}$ ir

$$g'(0.16, 0.36) = \frac{625}{97^2} \begin{pmatrix} 65 & -450 \end{pmatrix}.$$

Taigi

$$\sqrt{n}\Big(nT_n - \frac{100}{97}\Big) \to N(0,\sigma^2);$$

čia

$$\sigma^2 = \frac{625}{97^4}\begin{pmatrix}65 & -450\end{pmatrix}\begin{pmatrix}84 & -36\\ -36 & 144\end{pmatrix}\begin{pmatrix}65\\ -450\end{pmatrix}$$
$$= \frac{625}{97^4}\begin{pmatrix}65 & -450\end{pmatrix}\begin{pmatrix}21660\\ -67140\end{pmatrix} = \frac{625}{97^4}31620900 \approx 223.$$

3.4 PAVYZDYS. Tegu $X_1,\dots,X_n$ yra imtis iš skirstinio, aprašomo tankio funkcija

$$p(x) = \begin{cases} 2/3, & \text{kai } 0 < x < 1;\\ 1/x^4, & \text{kai } x > 1.\end{cases}$$

Raskite nurodytų statistikų asimptotinius skirstinius:

a) $T_n = \dfrac{N_{>1}}{N_{<2}}$;

b) $T_n = \dfrac{N_{>1}}{N_{>1} + \overline{X}}$;

čia $N_{>1}$ (atitinkamai, $N_{<2}$) yra skaičius imties narių, didesnių už 1 (atitinkamai, mažesnių už 2).

Sprendimas. a. Aišku, kad

$$T_n = g(\overline{1_{\{X>1\}}}, \overline{1_{\{X<2\}}});$$

čia $g(u,v) = u/v$.

Kadangi

$$\mathrm{E}\,1^2_{\{X>1\}} = \mathrm{E}\,1_{\{X>1\}} = \int_1^\infty \frac{\mathrm{d}x}{x^4} = -\frac{x^{-3}}{3}\Big|_1^\infty = \frac{1}{3},$$
$$\mathrm{E}\,1^2_{\{X<2\}} = \mathrm{E}\,1_{\{X<2\}} = \int_0^1 \frac{2}{3}\mathrm{d}x + \int_1^2 \frac{\mathrm{d}x}{x^4} = \frac{2}{3} - \frac{x^{-3}}{3}\Big|_1^2 = \frac{2}{3} - \frac{1}{24} + \frac{1}{3} = \frac{23}{24},$$
$$\mathrm{E}\,1_{\{X>1\}}1_{\{X<2\}} = \int_1^2 \frac{\mathrm{d}x}{x^4} = -\frac{x^{-3}}{3}\Big|_1^2 = -\frac{1}{24} + \frac{1}{3} = \frac{7}{24},$$

gaunu

$$\mathrm{D}\,1_{\{X>1\}} = \frac{2}{9}, \quad \mathrm{D}\,1_{\{X<2\}} = \frac{23}{24^2}, \quad \mathrm{cov}(1_{\{X>1\}}, 1_{\{X<2\}}) = -\frac{1}{36}.$$

Todėl iš centrinės ribinės teoremos

$$\sqrt{n}\Big[\begin{pmatrix}\overline{1_{\{X>1\}}}\\ \overline{1_{\{X<2\}}}\end{pmatrix} - \begin{pmatrix}1/3\\ 23/24\end{pmatrix}\Big] \to N(0,\Sigma);$$

čia

$$\Sigma = \begin{pmatrix}2/9 & -1/36\\ -1/36 & 23/24^2\end{pmatrix}.$$

Taikau vektorinį delta metodą. Iš pradžių randu g funkcijos išvestines:

$$\frac{\partial g}{\partial u} = \frac{1}{v}, \quad \frac{\partial g}{\partial v} = -\frac{u}{v^2}.$$

Įstatęs $u = 1/3$ ir $v = 23/24$, gaunu $g(1/3, 23/24) = 8/23$ ir

$$g'(1/3, 23/24) = \begin{pmatrix} 24/23 & -8 \cdot 24/23^2 \end{pmatrix} = \frac{24}{23^2} \begin{pmatrix} 23 & -8 \end{pmatrix}.$$

Taigi

$$\sqrt{n}\Big(T_n - \frac{8}{23}\Big) \to N(0, \sigma^2);$$

čia

$$\sigma^2 = \frac{1}{23^4} \begin{pmatrix} 23 & -8 \end{pmatrix} \begin{pmatrix} 128 & -16 \\ -16 & 23 \end{pmatrix} \begin{pmatrix} 23 \\ -8 \end{pmatrix}$$
$$= \frac{1}{23^4} \begin{pmatrix} 23 & -8 \end{pmatrix} \begin{pmatrix} 3072 \\ -552 \end{pmatrix} = \frac{3264}{23^3} \approx 0.27.$$

b. Aišku, kad

$$T_n = \frac{N_{>1}/n}{N_{>1}/n + \overline{X}/n} = g_n(\overline{1_{\{X>1\}}}, \overline{X});$$

čia

$$g_n(u, v) = \frac{u}{u + v/n}.$$

Kadangi

$$\mathrm{E}\, 1_{\{X>1\}}^2 = \mathrm{E}\, 1_{\{X>1\}} = \int_1^\infty \frac{\mathrm{d}x}{x^4} = -\frac{x^{-3}}{3}\Big|_1^\infty = \frac{1}{3},$$
$$\mathrm{E}\, X = \int_0^1 \frac{2x}{3} \mathrm{d}x + \int 1^\infty \frac{\mathrm{d}x}{x^3} = \frac{x^2}{3}\Big|_0^1 - \frac{x^{-2}}{2}\Big|_1^\infty = \frac{1}{3} + \frac{1}{2} = \frac{5}{6},$$
$$\mathrm{E}\, X^2 = \int_0^1 \frac{2x^2}{3} \mathrm{d}x + \int 1^\infty \frac{\mathrm{d}x}{x^2} = \frac{2x^3}{9}\Big|_0^1 - \frac{1}{x}\Big|_1^\infty = \frac{2}{9} + 1 = \frac{11}{9},$$
$$\mathrm{E}\, X 1_{\{X>1\}} = \int_1^\infty \frac{\mathrm{d}x}{x^4} = -\frac{x^{-3}}{3}\Big|_1^\infty = \frac{1}{3},$$

gaunu

$$\mathrm{D}\, 1_{\{X>1\}} = \frac{2}{9}, \quad \mathrm{D}\, X = \frac{19}{36}, \quad \mathrm{cov}(1_{\{X>1\}}, X) = \frac{1}{18}.$$

Reiškia,

$$\sqrt{n}\Big[\begin{pmatrix} \overline{1_{\{X>1\}}} \\ \overline{X} \end{pmatrix} - \begin{pmatrix} 1/3 \\ 5/6 \end{pmatrix}\Big] \to N(0, \Sigma);$$

čia

$$\Sigma = \begin{pmatrix} 2/9 & 1/18 \\ 1/18 & 19/36 \end{pmatrix}.$$

Pažymiu $U_n = \sqrt{n}(\overline{1_{\{X>1\}}} - 1/3)$ ir $V_n = \sqrt{n}(\overline{X} - 5/6)$; tada

$$\overline{1_{\{X>1\}}} = \frac{1}{3} + \frac{U_n}{\sqrt{n}}, \quad \overline{X} = \frac{5}{6} + \frac{V_n}{\sqrt{n}}.$$

Randu sekos $g_n(1/3+u_n/\sqrt{n}, 5/6+v_n/\sqrt{n})$ asimptotiką, kai $u_n \to u$ ir $v_n \to v$:

$$\begin{aligned}
&g_n(1/3+u_n/\sqrt{n}, 5/6+v_n/\sqrt{n}) \\
&= \frac{1/3+u_n/\sqrt{n}}{1/3+u_n/\sqrt{n}+5/(6n)+v_n/n^{3/2}} \\
&= \Big[1+\frac{3u_n}{\sqrt{n}}\Big]\Big[1+\frac{3u_n}{\sqrt{n}}+\frac{5}{2n}+\frac{3v_n}{n^{3/2}}\Big]^{-1} \\
&= \Big[1+\frac{3u_n}{\sqrt{n}}\Big]\Big[1-\Big(\frac{3u_n}{\sqrt{n}}+\frac{5}{2n}+\frac{3v_n}{n^{3/2}}\Big)+\Big(\frac{3u_n}{\sqrt{n}}+\frac{5}{2n}+\frac{3v_n}{n^{3/2}}\Big)^2 \\
&\qquad -\Big(\frac{3u_n}{\sqrt{n}}+\frac{5}{2n}+\frac{3v_n}{n^{3/2}}\Big)^3+o(n^{-3/2})\Big] \\
&= \Big[1+\frac{3u_n}{\sqrt{n}}\Big]\Big[1-\frac{3u_n}{\sqrt{n}}-\frac{5}{2n}-\frac{3v_n}{n^{3/2}}+\frac{9u_n^2}{n}+\frac{15u_n}{n^{3/2}}-\frac{27u_n^3}{n^{3/2}}+o(n^{-3/2})\Big] \\
&= 1-\frac{5}{2n}+\frac{15u_n-6v_n}{2n^{3/2}}+o(n^{-3/2}).
\end{aligned}$$

Kitaip tariant,

$$n^{3/2}\big[g_n(1/3+u_n/\sqrt{n}, 5/6+v_n/\sqrt{n})-1+5/(2n)\big] \to \frac{15u-6v}{2}.$$

Iš tolydaus atvaizdžio principo dabar išplaukia, kad

$$n^{3/2}\Big[T_n-1+\frac{5}{2n}\Big] \to \frac{3}{2}\begin{pmatrix}5 & -2\end{pmatrix} N(0,\Sigma) = N(0,\sigma^2);$$

čia

$$\sigma^2 = \frac{1}{16}\begin{pmatrix}5 & -2\end{pmatrix}\begin{pmatrix}8 & 2\\ 2 & 19\end{pmatrix}\begin{pmatrix}5\\ -2\end{pmatrix} = \frac{1}{16}\begin{pmatrix}5 & -2\end{pmatrix}\begin{pmatrix}36 & -28\end{pmatrix} = \frac{59}{4}.$$

4 skyrius

Taikymai

4.1 Pasikliautinieji intervalai

4.1.1 Uždavinio formulavimas

Tegu $X_1, \ldots, X_n$ yra nepriklausomi vienodai pasiskirstę atsitiktiniai dydžiai. Apie jų skirstinį yra žinoma tik tiek, kad jis priklauso tam tikrai skirstinių šeimai $(F_\theta \mid \theta \in \Theta)$, priklausančiai nuo realaus parametro θ. Reikia įvertinti nežinomą parametro reikšmę θ, t.y. rasti tokias statistikas $\theta_*(x_1, \ldots, x_n)$ ir $\theta^*(x_1, \ldots, x_n)$, kad

$$\mathrm{P}\{\theta_*(X_1, \ldots, X_n) \leqslant \theta \leqslant \theta^*(X_1, \ldots, X_n)\} \geqslant 1 - \alpha; \tag{4.1}$$

čia α yra tam tikras mažas skaičius. Paprastai įmama $\alpha = 5\% = 0.05$; tada intervalas $[\theta_*; \theta^*]$ vadinamas 95% lygmens pasikliautinuoju intervalu.

Jei paimtume

$$\theta_*(x_1, \ldots, x_n) = \theta^*(x_1, \ldots, x_n) = \theta,$$

(4.1) formulė būtų teisinga ir su $\alpha = 0$. Tačiau $[\theta; \theta]$ nevadinamas pasikliautinu intervalu, nes tokių funkcijų θ_* ir θ^* neįmanoma apskaičiuoti, nežinant tikrosios parametro reikšmės, ir todėl jos nevadinamos statistikomis. Griežtai dėstant matematinę statistiką, tikrosios parametro reikšmės nežinojimas formalizuojamas taip. θ laikomas kintamuoju, o atsitiktiniai dydžiai X_i — priklausančiais nuo to kintamojo: $X_i = X_i(\theta)$. Ieškome tokių statistikų $\theta_*(x_1, \ldots, x_n)$ ir $\theta^*(x_1, \ldots, x_n)$, kad (4.1) būtų teisinga su visais θ. Dabar niekas netrukdo paimti, pavyzdžiui,

$$\theta_*(x_1, \ldots, x_n) = \theta^*(x_1, \ldots, x_n) = 2.5,$$

bet tada (4.1) bus teisinga tik su $\theta = 2.5$. Todėl $[2.5; 2.5]$ nėra pasikliautinas intervalas.

4.1.2 Standartinė procedūra

Konkrečiuose modeliuose pasikliautinieji intervalai paprastai konstruojami taip. Iš pradžių randamas pagrįstas nežinomo parametro įvertis — tokia statistika $\hat{\theta}(x_1, \ldots, x_n)$, kad su visais θ

$$\hat{\theta}(X_1, \ldots, X_n) \to \theta.$$

Po to randamas asimptotinis $\hat\theta = \hat\theta\big(X_1, \dots, X_n\big)$ skirstinys; paprastai normalusis:

$$\sqrt{n}\big(\hat\theta - \theta\big) \to N(0, \sigma^2(\theta)).$$

Pažymima $\hat\sigma = \sigma(\hat\theta)$. Jei funkcija $\sigma(\theta)$ tolydi, tai pagal tolydaus atvaizdžio principą

$$\hat\sigma \to \sigma(\theta)$$

ir

$$\sqrt{n}\frac{\hat\theta - \theta}{\hat\sigma} \to N(0,1).$$

Tegu $z_{\alpha/2}$ yra standartinio normalaus skirstinio $(\alpha/2)$-kritinė reikšmė; tada

$$\mathrm{P}\big\{N(0,1) > z_{\alpha/2}\big\} = \alpha/2,$$
$$\mathrm{P}\big\{\big|N(0,1)\big| \leqslant z_{\alpha/2}\big\} = 1 - \alpha$$

ir

$$\mathrm{P}\Big\{\sqrt{n}\frac{\big|\hat\theta - \theta\big|}{\hat\sigma} \leqslant z_{\alpha/2}\Big\} \to 1 - \alpha.$$

Reiškia, kai n didelis, su $\approx 1 - \alpha$ tikimybe

$$\sqrt{n}\frac{\big|\hat\theta - \theta\big|}{\hat\sigma} \leqslant z_{\alpha/2}$$

ir

$$\hat\theta - z_{\alpha/2}\frac{\hat\sigma}{\sqrt{n}} \leqslant \theta \leqslant \hat\theta + z_{\alpha/2}\frac{\hat\sigma}{\sqrt{n}}.$$

Taigi

$$\Big[\hat\theta - z_{\alpha/2}\frac{\hat\sigma}{\sqrt{n}}; \hat\theta + z_{\alpha/2}\frac{\hat\sigma}{\sqrt{n}}\Big]$$

yra asimptotinis $(1 - \alpha)$ lygmens pasikliautinas intervalas.

4.1.3 Modifikuota procedūra

Vienas iš aukščiau aprašytos procedūros trūkumų — gaunamas pasikliautinas intervalas nebūtinai yra aibės Θ poaibis. Pavyzdžiui, jei θ yra kokio nors įvykio tikimybė, tai $\Theta = (0; 1)$. Įvertis $\hat\theta$ dažniausiai priklauso tam intervalui, bet kairysis pasikliautino intervalo kraštas $\hat\theta - z_{\alpha/2}\hat\sigma/\sqrt{n}$ nebūtinai teigiamas, o dešinysis kraštas gali būti didesnis už 1.

Jei standartinė procedūra dėl aukščiau aprašytų priežasčių mums nepatinka, galima naudoti tokią jos modifikaciją. Pasirenkame kokią nors griežtai didėjančią tolydžiai diferencijuojamą funkciją $g\colon \Theta \to \mathbb{R}$ ir naudodami standartinę procedūrą randame pasikliautiną intervalą $[\gamma_*; \gamma^*]$ parametrui $\gamma = g(\theta)$. Tada suskaičiuojame $\theta_* = g^{-1}(\gamma_*)$, $\theta^* = g^{-1}(\gamma^*)$ ir gauname pasikliautiną intervalą $[\theta_*; \theta^*] \subset \Theta$ parametrui θ.

4.1 PAVYZDYS. Tegu $X_1, \dots, X_n$ yra imtis iš skirstinio

X	1	2	3
	$p/2$	$p/2$	$1 - p$

Reikia įvertinti p.

Sprendimas. Pažymiu

$$\hat{p} = \frac{N_1 + N_2}{n} = \overline{1_{\{X \leqslant 2\}}};$$

čia N_j žymi j reikšmės pasikartojimų skaičių imtyje. Kadangi

$$\mathrm{E}\,1_{\{X \leqslant 2\}} = \mathrm{E}\,1^2_{\{X \leqslant 2\}} = p,$$

gaunu $\mathrm{D}\,1_{\{X \leqslant 2\}} = p - p^2 = p(1-p)$ ir iš centrinės ribinės teoremos

$$\sqrt{n}(\hat{p} - p) \to N\big(0, p(1-p)\big).$$

Tada

$$\sqrt{n}\frac{\hat{p} - p}{\hat{p}(1-\hat{p})} \to N(0,1);$$

todėl

$$\Big[\hat{p} - 1.96\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}; \hat{p} + 1.96\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}\Big]$$

yra 95% lygmens pasikliautinas intervalas.

Jei pasikliautiną intervalą konstruočiau iš duomenų

$$1,\ 2,\ 3,\ 3,\ 3,\ 3,\ 3,\ 3,\ 3,\ 3,$$

(t.y. jei $n = 10$, $\hat{p} = 0.2$), gaučiau intervalą $[-0.25; 0.45]$, kurio kairysis kraštas neigiamas. Kad to išvengčiau, taikysiu modifikuotą procedūrą su funkcija

$$g(p) = \ln\frac{p}{1-p}.$$

Kadangi

$$g'(p) = \big[\ln p - \ln(1-p)\big]' = \frac{1}{p} + \frac{1}{1-p} = \frac{1}{p(1-p)},$$

g funkcija yra griežtai didėjanti. Atvirkštinė funkcija gaunama išsprendus lygtį

$$\gamma = \ln\frac{p}{1-p}$$

kintamojo p atžvilgiu:

$$\begin{aligned} \frac{p}{1-p} &= \mathrm{e}^{\gamma}; \\ p &= (1-p)\mathrm{e}^{\gamma}; \\ p(1+\mathrm{e}^{\gamma}) &= \mathrm{e}^{\gamma}; \\ p &= \frac{\mathrm{e}^{\gamma}}{1+\mathrm{e}^{\gamma}}. \end{aligned}$$

Pažymiu $\hat{\gamma} = g(\hat{p})$. Tada pagal delta metodą

$$\sqrt{n}(\hat{\gamma} - \gamma) \to N(0, \sigma^2)$$

su
$$\sigma^2 = \big[g'(p)\big]^2 p(1-p) = \frac{1}{p(1-p)}.$$
Taigi
$$[\gamma_*;\gamma^*] = \Big[\hat\gamma - \frac{1.96}{\sqrt{n\hat p(1-\hat p)}};\hat\gamma + \frac{1.96}{\sqrt{n\hat p(1-\hat p)}}\Big]$$
yra 95% pasikliautinas intervalas parametrui γ, o
$$\Big[\frac{\mathrm{e}^{\gamma_*}}{1+\mathrm{e}^{\gamma_*}};\frac{\mathrm{e}^{\gamma^*}}{1+\mathrm{e}^{\gamma^*}}\Big] \quad —$$
95% pasikliautinas intervalas parametrui p.

Jei, kaip ir aukščiau, $n = 10$ ir $\hat p = 0.2$, tai
$$\hat\gamma = -1.386, \quad [\gamma_*;\gamma^*] = [-2.936; 0.114] \quad \text{ir} \quad [p_*;p^*] = [0.050; 0.528].$$

4.1.4 Dispersijos stabilizavimo metodas

Taikant modifikuotą pasikliautinų intervalų konstravimo procedūrą, funkciją g kartais galima parinkti taip, kad parametro $\gamma = g(\theta)$ asimptotinė dispersija nebepriklausytų nuo nežinomo parametro. Tokia funkcija vadinama *stabilizuojančia dispersiją*.

Kadangi asimptotinė parametro γ dispersija lygi $\big[g'(p)\big]^2\sigma^2(\theta)$, norėdami rasti dispersiją stabilizuojančią transformaciją, turime išspręsti diferencialinę lygtį
$$\big[g'(\theta)\big]^2\sigma^2(\theta) = 1,$$
t.y. suskaičiuoti integralą
$$g(\theta) = \int \frac{\mathrm{d}\theta}{\sigma(\theta)}.$$
Jei $\gamma = g(\theta)$, tai pasikliautinas intervalas γ parametrui yra
$$[\gamma_*;\gamma^*] = \Big[\hat\gamma - \frac{z_{\alpha/2}}{\sqrt{n}};\hat\gamma + \frac{z_{\alpha/2}}{\sqrt{n}}\Big] \tag{4.2}$$
pavidalo; čia $\hat\gamma = g(\hat\theta)$.

4.2 PAVYZDYS. Tarkime, turime imtį iš to paties skirstinio, kaip ankstesniame pavyzdyje. Reikia įvertinti nežinomą parametrą p, taikant dispersijos stabilizavimo metodą.

Sprendimas. Ankstesniame pavyzdyje suskaičiavau, kad jei $\hat p = (N_1+N_2)/n$, tai
$$\sqrt{n}(\hat p - p) \to N\big(0, p(1-p)\big).$$
Taigi dispersiją stabilizuojanti transformacija
$$g(p) = \int \frac{\mathrm{d}p}{\sqrt{p(1-p)}} = \int \frac{2\mathrm{d}p}{\sqrt{1-(2p-1)^2}} = \Big[t = 2p-1\Big]$$
$$= \int \frac{\mathrm{d}t}{\sqrt{1-t^2}} = \arcsin t = \arcsin(2p-1).$$

Atvirkštinė transformacija gaunama išsprendus lygtį

$$\gamma = \arcsin(2p-1)$$

p atžvilgiu:

$$(2p-1) = \sin\gamma;$$
$$p = \frac{1+\sin\gamma}{2}.$$

Tačiau reikia turėti omenyje, kad ši funkcija yra g funkcijos atvirkštinė tik jei jos apibrėžimo sritį susiaurinsime iki g funkcijos reikšmių srities, t.y. — iki $[-\pi/2;\pi/2]$ intervalo. Kitaip tariant,

$$\gamma_* \leqslant \gamma \leqslant \gamma^* \iff \frac{1+\sin\gamma_*}{2} \leqslant p \leqslant \frac{1+\sin\gamma^*}{2}$$

tik tada, jei $\gamma_*, \gamma^* \in [-\pi/2;\pi/2]$. Todėl jei pagal (4.2) formulę gaunamas rėžis $\gamma_* < -\pi/2$ (atitinkamai, $\gamma^* > \pi/2$), reikia imti $\gamma_* = -\pi/2$ (atitinkamai, $\gamma^* = \pi/2$).

Tegu, pavyzdžiui, $n = 10$ ir $\hat{p} = 0.2$. Tada $\hat{\gamma} = \arcsin(-0.8) = -0.644$ ir pagal (4.2) formulę gaunu

$$[\gamma_*;\gamma^*] = [-1.264;-0.024].$$

Kadangi $-1.264 > -\pi/2$, pasikliautinas intervalas p parametrui yra

$$\Big[\frac{1+\sin(-1.264)}{2};\frac{1+\sin(-0.024)}{2}\Big] = [0.023;0.488].$$

Jei $\hat{p} = 0$, pagal (4.2) formulę gaučiau

$$[\gamma_*;\gamma^*] = [-2.191;-0.951].$$

Šiuo atveju $\gamma_* < \pi/2$; todėl pasikliautinas intervalas p parametrui yra $[0;0.093]$, nors $(1+\sin(-2.191))/2 = 0.093$.

4.2 Hipotezių tikrinimas

4.2.1 Bendra teorija

Tarkime, $X_1,\dots,X_n$ yra imtis iš kokio nors nežinomo skirstinio ir reikia patikrinti tam tikrą statistinę hipotezę apie tą skirstinį. Hipotezės tikrinimo kriterijus apibrėžiamas, nurodant tam tikrą galimų imties reikšmių poaibį A; jei $(X_1,\dots,X_n) \in A$, hipotezė atmetama, priešingu atveju ji priimama.

Minimalus reikalavimas bet kokiam kriterijui — jo pirmos rūšies klaidos tikimybė turi neviršyti tam tikro mažo skaičiaus α, vadinamo kriterijaus reikšmingumo lygmeniu. Kitaip tariant, jei imtis yra iš skirstinio, tenkinančio hipotezę, tai turi būti teisinga nelygybė

$$\mathrm{P}\{(X_1,\dots,X_n) \in A\} \leqslant \alpha.$$

Jei nelygybė teisinga tik asimptotiškai, t.y. jei

$$\overline{\lim_{n\to\infty}}\, \mathrm{P}\{(X_1,\dots,X_n) \in A\} \leqslant \alpha,$$

α vadinamas asimptotiniu reikšmingumo lygmeniu.

Kriterijai paprastai konstruojami taip. Parenkama kokia nors statistika T_n ir randamas jos asimptotinis skirstinys, t.y. įrodomas

$$b_n(T_n - a_n) \to T$$

pavidalo teiginys. Statistika T_n turi būti tokia, kad jei hipotezė teisinga, skaičiai a_n, b_n ir T dydžio skirstinys nepriklausytų nuo nežinomo imties skirstinio. Tada hipotezę galima atmesti, kai $b_n(T_n - a_n) \in A$; čia A aibė parenkama taip, kad

$$\mathrm{P}\{T \in A\} \leqslant \alpha.$$

4.2.2 Nepriklausomumo tikrinimas, remiamtis empiriniu koreliacijos koeficientu

4.3 PAVYZDYS. Tegu $(X_1, Y_1), \dots, (X_n, Y_n)$ yra nepriklausomos atsitiktinio vektoriaus (X, Y) kopijos ir $\mathrm{E}\,X^4 < \infty$, $\mathrm{E}\,Y^4 < \infty$. Reikia patikrinti hipotezę, kad X ir Y nepriklausomi.

Sprendimas. Sukonstruosiu kriterijų, besiremiantį empiriniu koreliacijos koeficientu

$$\hat{\rho} = \frac{\overline{(X - \overline{X})(Y - \overline{Y})}}{\sqrt{\overline{(X - \overline{X})^2}}\sqrt{\overline{(Y - \overline{Y})^2}}} = \frac{\overline{XY} - \overline{X}\,\overline{Y}}{\sqrt{\overline{X^2} - \overline{X}^2}\sqrt{\overline{Y^2} - \overline{Y}^2}}.$$

Iš pradžių rasiu asimptotinį $\hat{\rho}$ skirstinį, kai hipotezė teisinga, t.y. kai X ir Y nepriklausomi. Pažymiu $a_{jk} = \mathrm{E}\,X^j Y^k$. Tada

$$\begin{aligned}
\mathrm{D}\,X &= \mathrm{E}\,X^2 - (\mathrm{E}\,X)^2 = a_{20} - a_{10}^2,\\
\mathrm{cov}(X, Y) &= \mathrm{E}\,XY - \mathrm{E}\,X\,\mathrm{E}\,Y = a_{11} - a_{10}a_{01},\\
\mathrm{cov}(X, X^2) &= \mathrm{E}\,X^3 - \mathrm{E}\,X\,\mathrm{E}\,X^2 = a_{30} - a_{10}a_{20},\\
\mathrm{cov}(X, Y^2) &= \mathrm{E}\,XY^2 - \mathrm{E}\,X\,\mathrm{E}\,Y^2 = a_{12} - a_{10}a_{02},\\
\mathrm{cov}(X, XY) &= \mathrm{E}\,X^2Y - \mathrm{E}\,X\,\mathrm{E}\,XY = a_{21} - a_{10}a_{11},\\
\mathrm{D}\,Y &= \mathrm{E}\,Y^2 - (\mathrm{E}\,Y)^2 = a_{02} - a_{01}^2,\\
\mathrm{cov}(Y, X^2) &= \mathrm{E}\,X^2Y - \mathrm{E}\,X^2\,\mathrm{E}\,Y = a_{21} - a_{20}a_{01},\\
\mathrm{cov}(Y, Y^2) &= \mathrm{E}\,Y^3 - \mathrm{E}\,Y\,\mathrm{E}\,Y^2 = a_{03} - a_{01}a_{02},\\
\mathrm{cov}(Y, XY) &= \mathrm{E}\,XY^2 - \mathrm{E}\,Y\,\mathrm{E}\,XY = a_{12} - a_{01}a_{11},\\
\mathrm{D}(X^2) &= \mathrm{E}\,X^4 - (\mathrm{E}\,X^2)^2 = a_{40} - a_{20}^2,\\
\mathrm{cov}(X^2, Y^2) &= \mathrm{E}\,X^2Y^2 - \mathrm{E}\,X^2\,\mathrm{E}\,Y^2 = a_{22} - a_{20}a_{02},\\
\mathrm{cov}(X^2, XY) &= \mathrm{E}\,X^3Y - \mathrm{E}\,X^2\,\mathrm{E}\,Y = a_{31} - a_{20}a_{01},\\
\mathrm{D}(Y^2) &= \mathrm{E}\,Y^4 - (\mathrm{E}\,Y^2)^2 = a_{04} - a_{02}^2,\\
\mathrm{cov}(Y^2, XY) &= \mathrm{E}\,XY^3 - \mathrm{E}\,Y^2\,\mathrm{E}\,XY = a_{13} - a_{02}a_{11},\\
\mathrm{D}(XY) &= \mathrm{E}\,X^2Y^2 - (\mathrm{E}\,XY)^2 = a_{22} - a_{11}^2.
\end{aligned}$$

Pradėsiu nuo atskiro atvejo, kai

$$\mathrm{E}\,X = \mathrm{E}\,Y = 0, \quad \mathrm{E}\,X^2 = \mathrm{E}\,Y^2 = 1. \tag{4.3}$$

Vėliau matysime, kad bendrą atvejį nesunku suvesti į išnagrinėtąjį. Taigi $a_{10} = a_{01} = 0$, $a_{20} = a_{02} = 1$. Be to, jei hipotezė teisinga, $a_{jk} = a_{j0}a_{0k}$. Taigi iš centrinės ribinės teoremos

$$\sqrt{n}\left[\begin{pmatrix} \overline{X} \\ \overline{Y} \\ \overline{X^2} \\ \overline{Y^2} \\ \overline{XY} \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}\right] \to N(0, \Sigma);$$

čia

$$\Sigma = \begin{pmatrix} 1 & 0 & a_{30} & 0 & 0 \\ 0 & 1 & 0 & a_{03} & 0 \\ a_{30} & 0 & a_{40}-1 & 0 & 0 \\ 0 & a_{03} & 0 & a_{04}-1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Dabar taikau vektorinį delta metodą. Aišku, kad $\hat{\rho} = g(\overline{X}, \overline{Y}, \overline{X^2}, \overline{Y^2}, \overline{XY})$; čia

$$g(u) = \frac{u_{11} - u_{10}u_{01}}{\sqrt{u_{20} - u_{10}^2}\sqrt{u_{02} - u_{01}^2}},$$

kai $u = (u_{10}, u_{01}, u_{20}, u_{02}, u_{11})$. Pažymiu

$$a = (a_{10}, a_{01}, a_{20}, a_{02}, a_{11}) = (0, 0, 1, 1, 0).$$

Kadangi

$$\frac{\partial g}{\partial u_{10}} = -\frac{u_{01}}{\sqrt{u_{20} - u_{10}^2}\sqrt{u_{02} - u_{01}^2}} - \frac{u_{11} - u_{10}u_{01}}{2\sqrt{u_{02} - u_{01}^2}}(u_{20} - u_{10}^2)^{-3/2}(-2u_{10}),$$

$$\frac{\partial g}{\partial u_{20}} = -\frac{u_{11} - u_{10}u_{01}}{2\sqrt{u_{02} - u_{01}^2}}(u_{20} - u_{10}^2)^{-3/2},$$

$$\frac{\partial g}{\partial u_{11}} = \frac{1}{\sqrt{u_{20} - u_{10}^2}\sqrt{u_{02} - u_{01}^2}}$$

gaunu

$$\frac{\partial g}{\partial u_{10}}(a) = 0, \quad \frac{\partial g}{\partial u_{20}}(a) = 0, \quad \frac{\partial g}{\partial u_{11}}(a) = 1$$

ir analogiškai

$$\frac{\partial g}{\partial u_{01}}(a) = 0, \quad \frac{\partial g}{\partial u_{02}}(a) = 0.$$

Reiškia,

$$\sqrt{n}(\hat{\rho} - 0) \to N(0, \sigma^2);$$

čia

$$\sigma^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & a_{30} & 0 & 0 \\ 0 & 1 & 0 & a_{03} & 0 \\ a_{30} & 0 & a_{40}-1 & 0 & 0 \\ 0 & a_{03} & 0 & a_{04}-1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$
$$= 1.$$

Kitaip tariant, jei hipotezė teisinga ir teisingos (4.3) lygybės, tai

$$\sqrt{n}\hat{\rho} \to N(0,1).$$

Dabar išnagrinėsiu bendrą atvejį: X ir Y nepriklausomi, bet (4.3) lygybės nebūtinai teisingos. Pažymiu

$$U_i = \frac{X_i - \mathrm{E}\,X}{\sqrt{\mathrm{D}\,X}}, \quad V_i = \frac{Y_i - \mathrm{E}\,Y}{\sqrt{\mathrm{D}\,Y}}.$$

Aišku, kad (U_i, V_i) yra nepriklausomos vektoriaus (U, V) kopijos; čia

$$U = \frac{X - \mathrm{E}\,X}{\sqrt{\mathrm{D}\,X}}, \quad V = \frac{Y - \mathrm{E}\,Y}{\sqrt{\mathrm{D}\,Y}}.$$

Jei hipotezė teisinga, t.y. X ir Y nepriklausomi, tai dydžiai U ir V taip pat nepriklausomi; be to,

$$\mathrm{E}\,U = \frac{\mathrm{E}(X - \mathrm{E}\,X)}{\sqrt{\mathrm{D}\,X}} = \frac{\mathrm{E}\,X - \mathrm{E}\,X}{\sqrt{\mathrm{D}\,X}} = 0,$$
$$\mathrm{E}\,U^2 = \frac{\mathrm{E}(X - \mathrm{E}\,X)^2}{\mathrm{D}\,X} = \frac{\mathrm{D}\,X}{\mathrm{D}\,X} = 1$$

ir analogiškai

$$\mathrm{E}\,V = 0, \quad \mathrm{E}\,V^2 = 1.$$

Taigi U ir V tenkina (4.3) sąlygas; todėl pagal išnagrinėtą atvejį

$$\sqrt{n}\frac{\overline{(U - \overline{U})(V - \overline{V})}}{\sqrt{\overline{(U - \overline{U})^2}}\sqrt{\overline{(V - \overline{V})^2}}} \to N(0,1).$$

Kita vertus, jei su kokiomis nors konstantomis $a > 0$, $c > 0$, b ir d

$$U = aX + b, \quad V = cY + d$$

(mūsų atveju taip ir yra: $a = 1/\sqrt{\mathrm{D}\,X}$, $b = -\mathrm{E}\,X/\sqrt{\mathrm{D}\,X}$, $c = 1/\sqrt{\mathrm{D}\,Y}$, $d = -\mathrm{E}\,Y/\sqrt{\mathrm{D}\,Y}$), tai

$$\mathrm{E}\,U = a\,\mathrm{E}\,X + b, \quad \mathrm{E}\,V = c\,\mathrm{E}\,Y + d;$$

todėl

$$\frac{\overline{(U - \overline{U})(V - \overline{V})}}{\sqrt{\overline{(U - \overline{U})^2}}\sqrt{\overline{(V - \overline{V})^2}}} = \frac{ac\overline{(X - \overline{X})(Y - \overline{Y})}}{\sqrt{a^2\overline{(X - \overline{X})^2}}\sqrt{c^2\overline{(Y - \overline{Y})^2}}} = \hat{\rho}.$$

Reiškia,

$$\sqrt{n}\hat{\rho} \to N(0,1) \tag{4.4}$$

ir bendruoju atveju.

Nepriklausomumo hipotezės tikrinimo kriterijus dabar gali būti aprašytas taip: hipotezė atmetama, kai $\sqrt{n}|\hat{\rho}| > c$; čia c — tokia konstanta, kad

$$\mathrm{P}\big\{\big|N(0,1)\big| > c\big\} = \alpha.$$

Jei hipotezė teisinga, tai iš (4.4) išplaukia, kad jos atmetimo tikimybė

$$\mathrm{P}\{\sqrt{n}|\hat{\rho}| > c\} \to \mathrm{P}\big\{\big|N(0,1)\big| > c\big\} = \alpha.$$

Reiškia, α asimptotinis tokio kriterijaus reikšmingumo lygmuo.

4.3 Klaidingos modelio specifikacijos efektas

4.3.1 Hipotezės apie dispersiją tikrinimas

Uždavinio formulavimas. Tegu $X_1, \ldots, X_n$ yra imtis iš normalaus skirstinio su nežinomu vidurkiu μ ir dispersija σ^2 ir reikia patikrinti hipotezę $\sigma = \sigma^*$; čia σ^* — koks nors fiksuotas skaičius. Paprastai tai daroma taip: apskaičiuojamos statistika

$$Q_n = \frac{1}{\sigma^*2}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

ir hipotezė atmetama, jei $Q_n > \chi^2_\alpha(n-1)$; čia $\chi^2_\alpha(n-1)$ yra $\chi^2(n-1)$ skirstinio α lygmens kritinė reikšmė, t.y. toks skaičius c, kad

$$\mathrm{P}\{\chi^2(n-1) > c\} = \alpha.$$

Aprašytas kriterijus pagrindžiamas taip: jei hipotezė teisinga, tai Q_n statistika pasiskirsčiusi pagal $\chi^2(n-1)$ dėsnį; todėl hipotezės atmetimo tikimybė lygi $\mathrm{P}\{\chi^2(n-1) > \chi^2_\alpha(n-1)\} = \alpha$. Taigi kriterijaus pirmos rūšies klaidos tikimybė yra α.

Jei $X_1, \ldots, X_n$ skirstinys nenormalus, Q skirstinys jau nebus $\chi^2(n-1)$, net jei hipotezė $\sigma = \sigma^*$ teisinga (t.y. jei X_i dispersija lygi σ^{*2}). Todėl kriterijaus pirmos rūšies klaidos tikimybė jau nėra tiksliai α. Bet galima tikėtis, kad ji bus apytiksliai lygi α, jei n pakankamai didelis. Išsiaiškinkime, ar tikrai.

Kad būtų paprasčiau, laikysiu $\sigma^* = 1$. Tarkime, hipotezė teisinga, t.y. $X_1, \ldots, X_n$ yra imtis iš skirstinio su dispersija 1. Reikia apskaičiuoti ribą

$$\lim_{n\to\infty} \mathrm{P}\{Q_n > \chi^2_\alpha(n-1)\};$$

čia Q_n yra ta pati statistika, kaip aukščiau. Norėdamas rasti tą ribą, iš pradžių rasiu Q_n asimptotiką, o po to — skaičių sekos $\chi^2_\alpha(n-1)$ asimptotiką, kai $n \to \infty$.

Q_n asimptotika. Aišku, kad

$$Q_n/n = \overline{(X-\overline{X})^2} = \overline{X^2 - 2X\overline{X} + \overline{X}^2}$$
$$= \overline{X^2} - 2\overline{X}\overline{X} + \overline{X}^2 = \overline{X^2} - \overline{X}^2 = g(u,v);$$

čia $g(u,v) = v - u^2$.

Su $j \geqslant 1$ pažymiu $a_j = \mathrm{E}\, X^j$. Tada

$$\begin{aligned} \mathrm{D}\, X &= \mathrm{E}\, X^2 - (\mathrm{E}\, X)^2 = a_2 - a_1^2; \\ \mathrm{D}\, X^2 &= \mathrm{E}\, X^4 - (\mathrm{E}\, X^2)^2 = a_4 - a_2^2; \\ \mathrm{cov}(X, X^2) &= \mathrm{E}\, X^3 - \mathrm{E}\, X\, \mathrm{E}\, X^2 = a_3 - a_1 a_2. \end{aligned}$$

Jei hipotezė teisinga, tai $\mathrm{D}\, X = a_2 - a_1^2 = 1$, t.y.

$$a_2 = a_1^2 + 1.$$

Taigi iš centrinės ribinės teoremos

$$\sqrt{n}\Big[\begin{pmatrix} \overline{X} \\ \overline{X^2} \end{pmatrix} - \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}\Big] \to N(0, \Sigma);$$

čia

$$\Sigma = \begin{pmatrix} a_2 - a_1^2 & a_3 - a_1 a_2 \\ a_3 - a_1 a_2 & a_4 - a_2^2 \end{pmatrix} = \begin{pmatrix} 1 & a_3 - a_1^3 - a_1 \\ a_3 - a_1^3 - a_1 & a_4 - a_1^4 - 2a_1^2 - 1 \end{pmatrix}.$$

Dabar taikau vektorinį delta metodą. Kadangi $g(a_1, a_2) = a_2 - a_1^2 = 1$ ir

$$\frac{\partial g}{\partial u} = -2u, \quad \frac{\partial g}{\partial v} = 1,$$

gaunu

$$\sqrt{n}\Big[Q_n/n - 1\Big] \to N(0, \sigma^2);$$

čia

$$\begin{aligned} \sigma^2 &= \begin{pmatrix} -2a_1 & 1 \end{pmatrix} \begin{pmatrix} 1 & a_3 - a_1^3 - a_1 \\ a_3 - a_1^3 - a_1 & a_4 - a_1^4 - 2a_1^2 - 1 \end{pmatrix} \begin{pmatrix} -2a_1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} -2a_1 & 1 \end{pmatrix} \begin{pmatrix} a_3 - a_1^3 - 3a_1 \\ a_4 - 2a_1 a_3 + a_1^4 - 1 \end{pmatrix} \\ &= a_4 - 4a_1 a_3 + 3a_1^4 + 6a_1^2 - 1 \end{aligned}$$

$\chi^2_\alpha(n-1)$ asimptotika. Pažymiu $c_{n-1} = \chi^2_\alpha(n-1)$. Pagal apibrėžimą c_n yra toks skaičius, kad

$$\mathrm{P}\{\chi^2(n) > c_n\} = \alpha.$$

Kita vertus, $\chi^2(n)$ yra sumos $Y_1^2 + \cdots + Y_n^2 = n\overline{Y^2}$ skirstinys; čia $Y_1, \ldots, Y_n$ — nepriklauomi atsitiktiniai dydžiai, turintys standartinį normalų pasiskirstymą.

Kadangi $\mathrm{E}\, Y = 0$, $\mathrm{E}\, Y^2 = \mathrm{D}\, Y = 1$ ir

$$\begin{aligned} \mathrm{E}\, Y^4 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^4 \mathrm{e}^{-x^2/2} \mathrm{d}x \\ &= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x^4 \mathrm{e}^{-x^2} \mathrm{d}x & \begin{bmatrix} x^2/2 = y \\ x = \sqrt{2y} \\ dx = \frac{\mathrm{d}y}{\sqrt{2y}} \end{bmatrix} \\ &= \frac{4}{\sqrt{\pi}} \int_0^{\infty} y^{3/2} \mathrm{e}^{-y} \mathrm{d}y \end{aligned}$$

$$\begin{aligned}
&= \frac{4}{\sqrt{\pi}}\Gamma(5/2) \\
&= \frac{4}{\sqrt{\pi}}\frac{3}{2}\frac{1}{2}\sqrt{\pi} \\
&= 3,
\end{aligned}$$

gaunu $\mathrm{D}\,Y^2 = 3 - 1 = 2$ ir iš centrinės ribinės teoremos

$$\sqrt{n}(\overline{Y^2} - 1) \to N(0, 2).$$

Kitaip tariant,

$$\sqrt{n/2}(\overline{Y^2} - 1) \to N(0, 1).$$

Toliau:

$$\begin{gathered}
\mathrm{P}\{n\overline{Y^2} > c_n\} = \alpha; \\
\mathrm{P}\{\overline{Y^2} > c_n/n\} = \alpha; \\
\mathrm{P}\{\sqrt{n/2}(\overline{Y^2} - 1) > \sqrt{n/2}(c_n/n - 1)\} = \alpha.
\end{gathered}$$

Dydis $\sqrt{n/2}(\overline{Y^2} - 1)$ konverguoja į $N(0, 1)$. Tam, kad tikimybė išliktų pastovi, dydis $\sqrt{n/2}(c_n/n - 1)$ turi artėti į tokį skaičių c, su kuriuo

$$\mathrm{P}\{N(0, 1) > c\} = \alpha.$$

Aišku, kad c yra standartinio normalaus skirstinio α lygmens kritinė reikšmė. Paprastai ji žymima z_α. Taigi

$$\sqrt{n/2}(c_n/n - 1) \to z_\alpha,$$

t.y.

$$\begin{gathered}
\sqrt{n/2}(c_n/n - 1) = z_\alpha + o(1); \\
c_n/n = 1 + \frac{\sqrt{2}z_\alpha}{\sqrt{n}} + o(n^{-1/2}); \\
c_n = n + z_\alpha\sqrt{2n} + o(n^{1/2}).
\end{gathered}$$

Iš čia

$$c_{n-1} = n - 1 + z_\alpha\sqrt{2n-2} + o(n^{1/2}) = n + z_\alpha\sqrt{2n} + o(n^{1/2}).$$

Klaidos tikimybės asimptotika. Jei hipotezė teisinga, tai pirmos rūšies klaidos tikimybė

$$\begin{aligned}
\mathrm{P}\{Q_n > c_{n-1}\} &= \mathrm{P}\Big\{\sqrt{n}\big(Q_n/n - 1\big) > \sqrt{n}\big(c_{n-1}/n - 1\big)\Big\} \\
&= \mathrm{P}\Big\{\sqrt{n}\big(Q_n/n - 1\big) > z_\alpha\sqrt{2} + o(1)\Big\} \\
&= \mathrm{P}\Big\{\sqrt{n}\big(Q_n/n - 1\big) + o(1) > z_\alpha\sqrt{2}\Big\} \\
&\to \mathrm{P}\{N(0, \sigma^2) > z_\alpha\sqrt{2}\} \\
&= \mathrm{P}\{N(0, 1) > z_\alpha\sqrt{2}/\sigma\}.
\end{aligned}$$

Reiškinys dešinėje pusėje lygus α tik tada, kai

$$z_\alpha\sqrt{2}/\sigma = z_\alpha,$$

t.y. $\sigma^2 = 2$, t.y.

$$a_4 - 4a_1a_3 + 3a_1^4 + 6a_1^2 - 1 = 2.$$

Normalios imties atveju taip ir yra (pavyzdžiui, jei $X_1, \ldots, X_n$ yra imtis iš standartinio normalaus skirstinio, tai $a_1 = 0$ ir $a_4 = 3$). Tačiau apskritai σ^2 gali būti bet koks. Reiškia, nagrinėjamo kriterijaus nenormalioms imtims taikyti negalima.

5 skyrius

M-įvertiniai

5.1 Maksimalaus tikėtinumo įvertiniai

Bendra teorija. Tegu $X_1, \dots, X_n$ yra imtis iš skirstinio su tankiu $p_\theta(x)$ (kokio nors mato μ atžvilgiu), priklausančio nuo nežinomo parametro θ, kurį reikia įvertinti. Tegu Θ yra galimų parametro reikšmių aibė,

$$l_\theta(x) = \ln p_\theta(x), \qquad L(\theta) = \frac{1}{n}\sum_{i=1}^{n} l_\theta(X_i) = \overline{l_\theta(X)}$$

ir $\hat\theta = \hat\theta(X_1, \dots, X_n)$ — toks aibės Θ taškas, kad

$$L(\hat\theta) = \max_{\theta\in\Theta} L(\theta).$$

Tada $\hat\theta$ vadinamas θ parametro *maksimalaus tikėtinumo įvertiniu*.

Išvestinę θ parametro atžvilgiu žymėsiu tašku virš funkciją žyminčios raidės; pavyzdžiui, $\dot L(\theta)$. Jei Θ aibė atvira, o L funkcija diferencijuojama, maksimalaus tikėtinumo įvertis yra lygties

$$\dot L(\theta) = 0$$

šaknis.Ši lygtis vadinama *maksimalaus tikėtinumo lygtimi*.

5.1 PAVYZDYS. Tegu $X_1, \dots, X_n$ yra imtis iš skirstinio su tankiu

$$p_\theta(x) = \theta \mathrm{e}^{-\theta x}, \quad \text{kai } x > 0;$$

čia $\theta > 0$ — nežinomas parametras. Raskite jo maksimalaus tikėtinumo įvertinį.

Sprendimas. Šiuo atveju

$$l_\theta(x) = \ln\theta - \theta x,$$
$$L(\theta) = \overline{\ln\theta - \theta X} = \ln\theta - \theta\overline{X},$$
$$\dot L(\theta) = \frac{1}{\theta} - \overline{X}.$$

Sprendžiu maksimalaus tikėtinumo lygtį:

$$\frac{1}{\theta} - \overline{X} = 0;$$

$$\frac{1}{\theta} = \overline{X};$$
$$\theta = \frac{1}{\overline{X}}.$$

Taigi $\hat{\theta} = 1/\overline{X}$.

Atvejis, kai imtis iš diskretaus skirstinio. Tegu $X(\theta)$ žymi atsitiktinį dydį, kurio skirstinys priklauso nagrinėjamai skirstinių šeimai ir atitinka parametro reikšmę θ. Jei visi $X(\theta)$ įgyja tik sveikąsias reikšmes, vietoje μ paprastai imamas skaičiuojantis matas. Tada

$$p_\theta(k) = \mathrm{P}\{X(\theta) = k\},$$

o

$$L(\theta) = \frac{1}{n} \sum_k N_k \ln p_\theta(k);$$

čia N_k yra k reikšmės pasikartojimų skaičius imtyje:

$$N_k = \sum_{i=1}^{n} 1_{\{X_i = k\}}.$$

Taigi

$$L(\theta) = \sum_k \ln p_\theta(k) \overline{1_{\{X_i = k\}}}.$$

5.2 PAVYZDYS. Tegu $X_1, \ldots, X_n$ yra imtis iš skirstinio

X	1	2	3
	p	$2p$	$1 - 3p$

čia $p \in (0; 1/3)$ — nežinomas parametras. Raskite jo maksimalaus tikėtinumo įvertį.

Sprendimas. Šiuo atveju

$$L(p) = N_1 \ln p + N_2 \ln(2p) + N_3 \ln(1 - 3p);$$
$$\dot{L}(p) = \frac{N_1}{p} + \frac{2N_2}{2p} - \frac{3N_3}{1 - 3p}.$$

Sprendžiu maksimalaus tikėtinumo lygtį:

$$\frac{N_1}{p} + \frac{2N_2}{2p} - \frac{3N_3}{1 - 3p} = 0;$$
$$\frac{N_1 + N_2}{p} = \frac{3N_3}{1 - 3p};$$
$$(N_1 + N_2)(1 - 3p) = 3N_3 p;$$
$$3(N_1 + N_2 + N_3)p = N_1 + N_2;$$
$$3np = N_1 + N_2;$$
$$p = \frac{N_1 + N_2}{3n}.$$

Taigi $\hat{p} = (N_1 + N_2)/(3n)$.

Vektorinio parametro atvejis. Kai $\theta = (\theta_1, \ldots, \theta_k)$ yra vektorinis parametras,

$$\dot{L}(\theta) = \begin{pmatrix} \dfrac{\partial L}{\partial \theta_1} & \cdots & \dfrac{\partial L}{\partial \theta_k} \end{pmatrix}.$$

Tuomet vektorinė maksimalaus tikėtinumo lygtis ekvivalenti tokiai skaliarinių lygčių sistemai:

$$\begin{cases} \dfrac{\partial L}{\partial \theta_1} & = 0; \\ \quad\vdots & \\ \dfrac{\partial L}{\partial \theta_k} & = 0. \end{cases}$$

5.3 PAVYZDYS. Tegu $X_1, \ldots, X_n$ yra imtis iš normalaus skirstinio su nežinomu vidurkiu μ ir nežinoma dispersija σ^2. Raskite abiejų parametrų maksimalaus tikėtinumo įvertinius.

Sprendimas. Šiuo atveju

$$p_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-(x-\mu)^2/(2\sigma^2)};$$

$$l_{\mu,\sigma}(X) = -\ln\sqrt{2\pi} - \ln\sigma - \frac{(X-\mu)^2}{2\sigma^2};$$

$$L(\mu,\sigma) = -\ln\sqrt{2\pi} - \ln\sigma - \overline{\frac{(X-\mu)^2}{2\sigma^2}};$$

$$\dot{L}(\mu,\sigma) = \begin{pmatrix} -\overline{\dfrac{-2(X-\mu)}{2\sigma^2}} & -\dfrac{1}{\sigma} + \overline{\dfrac{(X-\mu)^2}{\sigma^3}} \end{pmatrix} = \begin{pmatrix} \overline{\dfrac{X-\mu}{\sigma^2}} & -\dfrac{1}{\sigma} + \overline{\dfrac{(X-\mu)^2}{\sigma^3}} \end{pmatrix}.$$

Sprendžiu maksimalaus tikėtinumo lygtį:

$$\begin{cases} \overline{\dfrac{X-\mu}{\sigma^2}} & = 0; \\ -\dfrac{1}{\sigma} + \overline{\dfrac{(X-\mu)^2}{\sigma^3}} & = 0. \end{cases}$$

Iš pirmos lygties

$$\overline{X-\mu} = 0;$$

$$\overline{X} - \mu = 0;$$

$$\mu = \overline{X}.$$

Tada iš antrosios

$$\frac{1}{\sigma} = \overline{\frac{(X-\mu)^2}{\sigma^3}};$$

$$\sigma^2 = \overline{(X-\mu)^2} = \overline{(X-\overline{X})^2};$$

$$\sigma = \sqrt{\overline{(X-\overline{X})^2}}.$$

Taigi

$$\hat{\mu} = \overline{X}, \quad \sigma = \sqrt{\overline{(X-\overline{X})^2}}.$$

5.4 PAVYZDYS. Tegu $X_1, \ldots, X_n$ yra imtis iš skirstinio

X	1	2	3
	p	$p+q$	$1-2p-q$

čia $p > 0$ ir $q \in (-p; 1-2p)$ — nežinomi parametrai. Raskite jų maksimalaus tikėtinumo įverčius.

Sprendimas. Šiuo atveju

$$nL(p,q) = N_1 \ln p + N_2 \ln(p+q) + N_3 \ln(1-2p-q);$$
$$n\dot{L}(p,q) = \left(\frac{N_1}{p} + \frac{N_2}{p+q} - \frac{2N_3}{1-2p-q} \quad \frac{N_2}{p+q} - \frac{N_3}{1-2p-q} \right).$$

Sprendžiu maksimalaus tikėtinumo lygtį:

$$\begin{cases} \dfrac{N_1}{p} + \dfrac{N_2}{p+q} - \dfrac{2N_3}{1-2p-q} & = 0; \\ \dfrac{N_2}{p+q} - \dfrac{N_3}{1-2p-q} & = 0. \end{cases}$$

Iš antros lygties

$$\frac{N_3}{1-2p-q} = \frac{N_2}{p+q}. \tag{5.1}$$

Įstatęs į pirmąją, gaunu:

$$\frac{N_1}{p} + \frac{N_2}{p+q} - \frac{2N_2}{p+q} = 0;$$
$$\frac{N_1}{p} = \frac{N_2}{p+q};$$
$$N_1(p+q) = N_2 p;$$
$$N_1 q = (N_2 - N_1)p;$$
$$q = \frac{N_2 - N_1}{N_1} p.$$

Iš čia

$$p + q = \frac{N_2}{N_1} p \quad \text{ir} \quad 1 - 2p - q = 1 - 2p - \frac{N_2 - N_1}{N_1} p;$$

todėl iš (5.1) gaunu

$$1 - 2p - q = \frac{N_3}{N_2}(p+q);$$
$$1 - 2p - \frac{N_2 - N_1}{N_1} p = \frac{N_3}{N_1} p;$$
$$\left(1 + \frac{N_2 + N_3}{N_1}\right) p = 1;$$
$$\frac{N_1 + N_2 + N_3}{N_1} p = 1;$$
$$\frac{n}{N_1} p = 1;$$
$$p = \frac{N_1}{n}.$$

Tada

$$q = \frac{N_2 - N_1}{N_1} \cdot \frac{N_1}{n} = \frac{N_2 - N_1}{n}.$$

Taigi

$$\hat{p} = \frac{N_1}{n} \quad \text{ir} \quad \hat{q} = \frac{N_2 - N_1}{n}.$$

5.2 M-įvertinių pagrįstumas. Klasikinė teorema

5.2.1 M-įvertiniai ir Z-įvertiniai

Tegu $X_1, \dots, X_n$ yra imtis iš skirstinio F_{θ_0}, priklausančio skirstinių šeimai $(F_\theta \mid \theta \in \Theta)$ ir reikia įvertinti nežinomą parametro reikšmę θ_0. Vienas iš plačiai naudojamų įvertinių konstravimo būdų yra toks. Pasirenkama funkcijų šeima $m_\theta(x)$, apibrėžiama

$$M_n(\theta) = \overline{m_\theta(X)} = \frac{m_\theta(X_1) + \dots + m_\theta(X_n)}{n}$$

ir randamas toks $\hat{\theta}_n = \hat{\theta}_n(X_1, \dots, X_n)$, kad

$$M_n(\hat{\theta}_n) = \sup_{\theta \in \Theta} M_n(\theta).$$

Tada $\hat{\theta}$ vadinamas *M-įvertiniu*.

Pirmame skyrelyje nagrinėti maksimalaus tikėtinumo įvertiniai yra M-įvertiniai: jie atitinka funkciją

$$m_\theta(x) = \ln p_\theta(x);$$

čia p_θ yra F_θ skirstinio tankis kokio nors mato atžvilgiu.

Kad M-įvertinys tikrai vertintų tai, ką reikia, t.y. nežinomą parametro reikšmę θ_0, funkcijos $m_\theta(x)$ turi tenkinti tam tikrą sąlygą. Tikrai, iš didžiųjų skaičių dėsnio

$$M_n(\theta) = \overline{m_\theta(X)} \to M(\theta);$$

čia

$$M(\theta) = \mathrm{E}\, m_\theta(X).$$

$\hat{\theta}_n$ yra $M_n(\theta)$ maksimumo taškas; todėl natūralu reikalauti, kad θ_0 būtų vienintelis funkcijos $M(\theta)$ maksimumo taškas.

Jei Θ yra atviras $\mathbb{R}^k$ poaibis, o $m_\theta(x)$ funkcijos tolydžiai diferencijuojamos θ atžvilgiu, tai maksimumo taško ieškome prilygindami 0 $M_n(\theta)$ funkcijos išvestinę. Pažymiu

$$\psi_\theta(x) = \dot{m}_\theta(x), \quad \Psi_n(\theta) = \overline{\psi_\theta(X)}, \quad \Psi(\theta) = \mathrm{E}\, \psi_\theta(X).$$

Jei θ_0 yra vienintelis $\Psi(\theta)$ funkcijos nulis, o $\hat{\theta}_n$ — koks nors $\Psi_n(\theta)$ funkcijos nulis, galima tikėtis, kad $\hat{\theta}_n \to \theta_0$. Statistika $\hat{\theta}_n$ vadinama *Z-įvertiniu*.

5.2.2 Klasikinė teorema

5.1 teorema. *Tegu Θ yra atviras intervalas, su visais x funkcija $m_\theta(x)$ tolydi θ atžvilgiu ir θ_0 yra M funkcijos lokalaus maksimumo taškas. Tada egzistuoja tokie įvertiniai $\hat{\theta}_n$ ir tokios aibės A_n, kad*

(i) $\mathrm{P}\{(X_1,\ldots,X_n)\in A_n\}\to 1$;

(ii) kai $(x_1,\ldots,x_n)\in A_n$, $\hat{\theta}_n(x_1,\ldots,x_n)$ yra M_n funkcijos lokalaus maksimumo taškas;

(iii) $\hat{\theta}_n\to\theta_0$.

Įrodymas. Fiksuoju δ, su kuriuo $M(\theta_0-\delta)<M(\theta_0)>M(\theta_0+\delta)$. Pažymiu

$$\varepsilon=\min\bigl(M(\theta_0)-M(\theta_0-\delta),M(\theta_0)-M(\theta_0+\delta)\bigr)$$

ir

$$B_n=\bigl\{(x_1,\ldots,x_n)\bigm|\bigl|M_n(\theta)-M(\theta)\bigr|<\varepsilon/2\text{ su }\theta=\theta_0,\theta_0\pm\delta\bigr\}.$$

Tada $\mathrm{P}\{(X_1,\ldots,X_n)\in B_n\}\to 1$ ir, kai $(X_1,\ldots,X_n)\in B_n$,

$$M_n(\theta_0-\delta)<M_n(\theta_0)>M_n(\theta_0+\delta). \tag{5.2}$$

Taigi

$$\mathrm{P}\{M_n(\theta_0-\delta)<M_n(\theta_0)>M_n(\theta_0+\delta)\}\to 1.$$

Iš lemos, kurią įrodysiu žemiau, išplaukia, kad su tam tikra seka δ_n, artėjančia į 0,

$$\mathrm{P}\{M_n(\theta_0-\delta_n)<M_n(\theta_0)>M_n(\theta_0+\delta_n)\}\to 1.$$

Pažymiu

$$A_n=\{(x_1,\ldots,x_n)\mid M_n(\theta_0-\delta_n)<M_n(\theta_0)>M_n(\theta_0+\delta_n)\}.$$

Funkcija $M_n(\theta)$ yra tolydi; todėl $[\theta_0-\delta_n;\theta_0+\delta_n]$ intervale įgyja didžiausią reikšmę. Jei $(x_1,\ldots,x_n)\in A_n$, tai ta didžiausia reikšmė įgyjama intervalo viduje ir todėl yra lokalus funkcijos maksimumas. Pažymiu jį $\hat{\theta}_n(x_1,\ldots,x_n)$. Jei $(x_1,\ldots,x_n)\notin A_n$, $\hat{\theta}_n$ apibrėžiu bet kaip. Tada, kai n pakankamai didelis,

$$\mathrm{P}\{|\hat{\theta}_n-\theta_0|\geqslant\delta\}\leqslant\mathrm{P}\{|\hat{\theta}_n-\theta_0|\geqslant\delta_n\}\leqslant\mathrm{P}\{(X_1,\ldots,X_n)\notin A_n\}\to 0, \tag{5.3}$$

t.y. $\hat{\theta}_n\to\theta_0$. $\square$

5.1 lema. *Tegu $h_n(\delta)$ yra tokia realiųjų funkcijų seka, kad $h_n(\delta)\xrightarrow[n\to\infty]{}1$ su kiekvienu δ. Tada egzistuoja tokia seka $\delta_n\to 0$, kad $h_n(\delta_n)\to 1$.*

Įrodymas. Paimu $n_0=1$ ir su kiekvienu $k\geqslant 1$ randu tokį $n_k>n_{k-1}$, kad $|h_n(1/k)-1|<1/k$ su $n\geqslant n_k$. Apibrėžiu

$$\delta_n=\begin{cases}2, & \text{kai } n<n_1;\\ 1, & \text{kai } n_1\leqslant n<n_2;\\ \vdots & \\ 1/k, & \text{kai } n_k\leqslant n<n_{k+1};\\ \vdots & \end{cases}$$

Aišku, kad $\delta_n \to 0$. Jei $n \geqslant n_k$, tai $n_l \leqslant n < n_{l+1}$ su tam tikru $l \geqslant k$, o tada

$$|h_n(\delta_n) - 1| = |h_n(1/l) - 1| < 1/l \leqslant 1/k. \tag{5.4}$$

Reiškia, $h_n(\delta_n) \to 1$. □

Kad 5.1 teoremą būtų galima pritaikyti maksimalaus tikėtinumo įvertiniams, reikalinga dar viena teorema.

5.2 teorema. *Tegu skirstinių šeima* $(F_\theta \mid \theta \in \Theta)$ *identifikuojama, t.y. skirtingus* θ *atitinka skirtingi skirstiniai* F_θ. *Tada* θ_0 *yra vienintelis funkcijos* $M(\theta) = \mathrm{E}\ln p_\theta(X)$ *maksimumo taškas.*

Įrodymas. Iš nelygybės $\ln x \leqslant 2(\sqrt{x} - 1)$, teisingos su $x > 0$, gaunu

$$\begin{aligned} M(\theta) - M(\theta_0) &= \mathrm{E}\ln\frac{p_\theta(X)}{p_{\theta_0}(X)} \leqslant 2\,\mathrm{E}\Big(\sqrt{\frac{p_\theta(X)}{p_{\theta_0}(X)}} - 1\Big) \\ &= 2\Big(\int(\sqrt{p_\theta p_{\theta_0}}\mathrm{d}\mu - 1\Big) = -\int(\sqrt{p_\theta} - \sqrt{p_{\theta_0}})^2\mathrm{d}\mu \leqslant 0. \end{aligned} \tag{5.5}$$

Taigi $M(\theta) \leqslant M(\theta_0)$ su visais θ. Be to, nelygybė virsta lygybe tik tada, kai $p_\theta = p_{\theta_0}$ beveik visur μ atžvilgiu, t.y. kai $\theta = \theta_0$. □

5.2.3 Įrodymų paaiškinimai

5.1 teorema.

1. Kodėl $\mathrm{P}\{(X_1, \ldots, X_n) \in B_n\} \to 1$?

$B_n^c = B_n^1 \cup B_n^2 \cup B_n^3$; čia

$$B_n^1 = \{\big|M_n(\theta_0) - M(\theta_0)\big| \geqslant \varepsilon/2\}, \quad B_n^2 = \{\big|M_n(\theta_0 - \delta) - M(\theta_0 - \delta)\big| \geqslant \varepsilon/2\},$$
$$B_n^3 = \{\big|M_n(\theta_0 + \delta) - M(\theta_0 + \delta)\big| \geqslant \varepsilon/2\}.$$

Iš didžiųjų skaičių dėsnio išplaukia, kad su visais θ

$$M_n(\theta) = \overline{m_\theta(X)} \to \mathrm{E}\,m_\theta(X) = M(\theta);$$

todėl $\mathrm{P}\{(X_1, \ldots, X_n) \in B_n^j\} \to 0$ su $i = 1, 2, 3$. Bet tada

$$\begin{aligned} &\mathrm{P}\{(X_1, \ldots, X_n) \notin B_n\} \\ &\leqslant \mathrm{P}\{(X_1, \ldots, X_n) \in B_n^1\} + \mathrm{P}\{(X_1, \ldots, X_n) \in B_n^2\} + \mathrm{P}\{(X_1, \ldots, X_n) \in B_n^3\} \\ &\qquad \to 0, \end{aligned}$$

t.y. $\mathrm{P}\{(X_1, \ldots, X_n) \in B_n\} \to 1$.

2. Kodėl teisingos (5.2) nelygybės?

Jei $(x_1, \ldots, x_n) \in B_n$, tai

$$M_n(\theta_0) \geqslant M(\theta_0) - \varepsilon/2, \quad M_n(\theta_0 - \delta) < M(\theta_0 - \delta) + \varepsilon/2;$$

todėl

$$M_n(\theta_0) - M_n(\theta_0 - \delta) > M(\theta_0) - \varepsilon/2 - M(\theta_0 - \delta) - \varepsilon/2 = M(\theta_0) - M(\theta_0 - \delta) - \varepsilon.$$

Dešinė pusė neneigiama pagal ε apibrėžimą. Taigi $M_n(\theta_0) > M_n(\theta_0 - \delta)$. Antra nelygybė įrodoma panašiai.

3. Kodėl $M_n(\theta)$ tolydi?

Nes

$$M_n(\theta) = \frac{m_\theta(x_1) + \cdots + m_\theta(x_n)}{n}$$

ir visos $m_\theta(x_i)$ funkcijos tolydžios θ atžvilgiu.

4. Paaiškinkite smulkiau, kodėl $\hat{\theta}_n$ yra lokalaus maksimumo taškas.

Kadangi $M_n(\theta_0 - \delta) < M_n(\theta_0)$, $\hat{\theta} \neq \theta_0 - \delta$. Dėl panašios priežasties $\hat{\theta}_0 \neq \theta_0 + \delta$. Taigi $\hat{\theta} \in (\theta_0 - \delta; \theta_0 + \delta)$. Tada $(\theta_0 - \delta; \theta_0 + \delta)$ yra $\hat{\theta}$ aplinka ir su visais θ iš tos aplinkos $M_n(\theta) \leqslant M_n(\theta_0)$. Tai ir reiškia, kad $\hat{\theta}$ yra lokalaus maksimumo taškas.

5. Kodėl teisingi sąryšiai (5.3) grandinėje?

Iš $\delta_n \to 0$ išplaukia, kad $\delta_n < \delta$, kai n pakankamai didelis. Tada

$$|\hat{\theta}_n - \theta_0| \geqslant \delta \Rightarrow |\hat{\theta}_n - \theta_0| \geqslant \delta_n;$$

todėl $\mathrm{P}\{|\hat{\theta}_n - \theta_0| \geqslant \delta\} \leqslant \mathrm{P}\{|\hat{\theta}_n - \theta_0| \geqslant \delta_n\}$.

Jei $(x_1, \ldots, x_n) \in A_n$, tai pagal apibrėžimą $\hat{\theta}_n \in (\theta_0 - \delta_n; \theta_0 + \delta_n)$, t.y. $|\hat{\theta}_n - \theta_0| < \delta$. Reiškia,

$$|\hat{\theta}_n - \theta_0| \geqslant \delta \Rightarrow (X_1, \ldots, X_n) \notin A_n;$$

todėl $\mathrm{P}\{|\hat{\theta}_n - \theta_0| \geqslant \delta_n\} \leqslant \mathrm{P}\{(X_1, \ldots, X_n) \notin A_n\}$.

Trečias sąryšis išplaukia iš

$$\mathrm{P}\{(X_1, \ldots, X_n) \notin A_n\} = 1 - \mathrm{P}\{(X_1, \ldots, X_n) \in A_n\}$$

ir $\mathrm{P}\{(X_1, \ldots, X_n) \in A_n\} \to 1$.

5.1 lema.

6. Kodėl egzistuoja tokie n_k?

Nes $h_n(1/k) \xrightarrow[n\to\infty]{} 1$.

7. Kodėl $\delta_n \to 0$?

Nes $0 < \delta_n \leqslant 1/k$, kai $n \geqslant n_k$.

8. Kodėl teisingi sąryšiai (5.4) grandinėje?

Pirma lygybė — todėl, kad $\delta_n = 1/l$, kai $n_l \leqslant n < n_{l+1}$. Antra nelygybė — todėl, kad $\big|h_n(1/l) - 1\big| < 1/l$, kai $n \geqslant n_l$. Trečia nelygybė — todėl, kad $l \geqslant k$.

5.2 teorema.

9. Įrodykite, kad $\ln x \leqslant 2(\sqrt{x} - 1)$, kai $x > 0$.

Pažymiu $f(x) = \ln x - 2(\sqrt{x} - 1)$. Tada

$$f'(x) = \frac{1}{x} - \frac{2}{2\sqrt{x}} = \frac{1}{x} - \frac{1}{\sqrt{x}}.$$

Išvestinė lygi 0 tik viename taške — $x = 1$. Kai $x < 1$, išvestinė teigiama; kai $x > 1$, išvestinė neigiama. Reiškia, kritiniame taške yra funkcijos maksimumas. Todėl

$$f(x) \leqslant f(1) = 0$$

su visais x.

10. Kodėl teisingi sąryšiai (5.5) grandinėje?

Pirma lygybė išplaukia iš

$$M(\theta) - M(\theta_0) = \mathrm{E}\ln p_\theta(X) - \mathrm{E}\ln p_{\theta_0}(X)$$
$$= \mathrm{E}\big[\ln p_\theta(X) - \ln p_{\theta_0}(X)\big] = \mathrm{E}\ln\frac{p_\theta(X)}{p_{\theta_0}(X)}.$$

Antras sąryšis — iš nelygybės $\ln x \leqslant 2(\sqrt{x}-1)$.

Kadangi p_{θ_0} yra X tankis μ mato atžvilgiu, $\mathrm{E}\,f(X) = \int f p_{\theta_0}\mathrm{d}\mu$ su bet kokia funkcija f. Reiškia,

$$\mathrm{E}\Big(\sqrt{\frac{p_\theta(X)}{p_{\theta_0}(X)}} - 1\Big) = \mathrm{E}\sqrt{\frac{p_\theta(X)}{p_{\theta_0}(X)}} - 1$$
$$= \int\sqrt{\frac{p_\theta}{p_{\theta_0}}}p_{\theta_0}\mathrm{d}\mu - 1 = \int\sqrt{p_\theta p_{\theta_0}}\mathrm{d}\mu - 1.$$

Ketvirta lygybė teisinga, nes

$$\int(\sqrt{p_\theta} - \sqrt{p_{\theta_0}})^2\mathrm{d}\mu = \int\big[p_\theta - 2\sqrt{p_\theta p_{\theta_0}} + p_{\theta_0}\big]\mathrm{d}\mu$$
$$= \int p_\theta\mathrm{d}\mu - 2\int\sqrt{p_\theta p_{\theta_0}}\mathrm{d}\mu + \int p_{\theta_0}\mathrm{d}\mu = 1 - 2\int\sqrt{p_\theta p_{\theta_0}}\mathrm{d}\mu + 1.$$

Paskutinė nelygybė išplaukia iš to, kad neneigiamos funkcijos integralas neneigiamas.

5.2.4 Komentarai

11. Dažnai tvirtinama, kad 5.1 teorema įrodo M-įvertinių pagrįstumą. Deja, taip nėra. Jei $M_n(\theta)$ funkcija turi keletą lokalių maksimumų, iš teoremos neišplaukia, kad bet kokia parametro reikšmių, kuriuose pasiekiamas lokalus maksimumas, seka artėja prie θ_0. Taigi jei $\hat\theta_n$ yra $M_n(\theta)$ globalaus maksimumo taškas, visiškai neaišku, ar $\hat\theta_n \to \theta_0$.

12. Teoremoje sukonstruotų dydžių $\hat\theta_n$ netgi negalima vadinti statistikomis. Pagal konstrukciją $\hat\theta_n$ yra lokalaus maksimumo taškas iš $(\theta_0 - \delta_n; \theta_0 + \delta_n)$ intervalo; taigi norėdami pasirinkti $\hat\theta_n$ mes turime žinoti θ_0.

5.3 Valdo teorema

5.3.1 Valdo teorema

5.3 teorema. *Tarkime, Θ yra kompaktiška metrinė erdvė, su kiekvienu x funkcija $m_\theta(x)$ tolydi θ atžvilgiu ir su kiekvienu $\theta \in \Theta$ egzistuoja toks $r > 0$, kad*

$$\mathrm{E}\big[\sup_{\theta'\in U(\theta,r)} m_{\theta'}(X)\big]^+ < \infty. \tag{5.6}$$

Jei θ_0 yra vienintelis funkcijos $M(\theta)$ maksimumo taškas, o $\hat\theta_n$ — funkcijos $M_n(\theta)$ globalaus maksimumo taškas, tai $\hat\theta_n \to \theta_0$.

Įrodymas. Kad būtų trumpiau, su bet kokia $A \subset \Theta$ pažymiu

$$m_A(x) = \sup_{\theta \in A} m_\theta(x).$$

Iš $\theta \in U(\theta, r)$ išplaukia

$$m_\theta(x) \leqslant m_{U(\theta,r)}(x);$$
$$\big[m_\theta(x)\big]^+ \leqslant \big[m_{U(\theta,r)}(x)\big]^+;$$
$$\mathrm{E}\big[m_\theta(X)\big]^+ \leqslant \mathrm{E}\big[m_{U(\theta,r)}(X)\big]^+ < \infty.$$

Taigi vidurkiai $\mathrm{E}\, m_\theta(X)$ ir $\mathrm{E}\, m_{U(\theta,r)}(X)$ apibrėžti, nors, gal būt, lygūs $-\infty$.

Kadangi $m_{U(\theta,r/n)} \downarrow m_\theta(x)$, iš (5.6) išplaukia, kad su kiekvienu θ

$$\mathrm{E}\, m_{U(\theta,r/n)}(X) \downarrow M(\theta). \tag{5.7}$$

Pažymiu $B = \{\theta \in \Theta \mid d(\theta, \theta_0) \geqslant \varepsilon\}$. Tada su kiekvienu $\theta \in B$ egzistuoja toks $r(\theta)$, kad

$$\mathrm{E}\, m_{U(\theta,r(\theta))}(X) < M(\theta_0).$$

B aibė kompaktiška, o rutuliai $U(\theta, r(\theta))$, $\theta \in \Theta$, ją dengia. Reiškia, egzistuoja tokie $\theta_1, \dots, \theta_k \in \Theta$, kad

$$B \subset U(\theta_1, r(\theta_1)) \cup \dots \cup U(\theta_k, r(\theta_k)).$$

Kad būtų trumpiau, pažymiu $U_j = U(\theta_j, r(\theta_j))$. Tada $B \subset U_1 \cup \dots \cup U_k$ ir $\mathrm{E}\, m_{U_j}(X) < M(\theta_0)$ su visais $j = 1, \dots, k$.

Jei $\mathrm{E}\, m_{U_j}(X) < q_j < M(\theta_0)$, tai

$$\mathrm{P}\{\overline{m_{U_j}(X)} \geqslant q_j\} \xrightarrow[n\to\infty]{} 0 \quad \text{ir} \quad \mathrm{P}\{M_n(\theta_0) \leqslant q_j\} \xrightarrow[n\to\infty]{} 0; \tag{5.8}$$

todėl

$$\mathrm{P}\{\overline{m_{U_j}(X)} < M_n(\theta_0)\} \xrightarrow[n\to\infty]{} 1. \tag{5.9}$$

Jei $\hat\theta_n \in U_j$, tai

$$M_n(\theta_0) \leqslant M_n(\hat\theta_n) \leqslant \overline{m_{U_j}(X)}; \tag{5.10}$$

todėl $\mathrm{P}\{\hat\theta_n \in U_j\} \to 0$. Iš čia $\mathrm{P}\{\hat\theta_n \in B\} \to 0$, t.y. $\hat\theta_n \to \theta_0$. $\square$

5.3.2 Įrodymo paaiškinimai

13. Kodėl $m_{U(\theta,r/n)} \downarrow m_\theta(x)$?

Jei n didėja, tai aibė $U(\theta, r/n)$ mažėja, o kuo mažesnė aibė, tuo mažesnis jos supremumas. Todėl seka $m_{U(\theta,r/n)}$ nedidėjanti.

Fiksuoju ε. Dėl $m_\theta(x)$ funkcijos tolydumo atsiras toks δ, kad su visais $\theta' \in U(\theta, \delta)$

$$\big|m_{\theta'}(x) - m_\theta(x)\big| < \varepsilon.$$

Tada

$$m_{\theta'}(x) \leqslant m_\theta(x) + \varepsilon$$

su visais $\theta' \in U(\theta, \delta)$; todėl

$$m_{U(\theta,\delta)}(x) \leqslant m_\theta(x) + \varepsilon.$$

Iš čia

$$m_\theta(x) \leqslant m_{U(\theta,r/n)}(x) \leqslant m_\theta(x) + \varepsilon,$$

kai n pakankamai didelis (kai $1/n < \delta$). Reiškia, $m_{U(\theta,r/n)}(x) \to m_\theta(x)$.

14. Paaiškinkite smulkiau, kodėl teisingas (5.7) sąryšis.

Pažymiu $Y_n = m_{U(\theta,r/n)}(X)$ ir $Y = m_\theta(X)$. Tada $Y_n \downarrow Y$ ir $\mathrm{E}\,Y_1^+ < \infty$.

Funkcija $y \mapsto y^+$ nemažėjanti ir tolydi; todėl $Y_n^+ \downarrow Y^+$. Be to, $|Y_n^+| = Y_n^+ \leqslant Y_1^+$ ir $\mathrm{E}\,Y_1^+ < \infty$; todėl iš Lebego teoremos apie aprėžtą konvergavimą $\mathrm{E}\,Y_n^+ \to \mathrm{E}\,Y^+$.

Funkcija $y \mapsto y^-$ nedidėjanti ir tolydi; todėl $Y_n^- \uparrow Y^-$. Tada iš Lebego teoremos apie monotonišką konvergavimą išplaukia $\mathrm{E}\,Y_n^- \to \mathrm{E}\,Y^-$. Reiškia,

$$\mathrm{E}\,Y_n = \mathrm{E}\,Y_n^+ - \mathrm{E}\,Y_n^- \to \mathrm{E}\,Y^+ - \mathrm{E}\,Y^- = \mathrm{E}\,Y.$$

15. Kodėl egzistuoja toks $r(\theta)$?

Jei $\theta \in B$, tai $\theta \neq \theta_0$; todėl $M(\theta) < M(\theta_0)$. Tada

$$\mathrm{E}\,m_{U(\theta,r/n)}(X) \to M(\theta) < M(\theta_0);$$

todėl $\mathrm{E}\,m_{U(\theta,r/n)}(X) < M(\theta_0)$, kai n pakankamai didelis. Belieka pažymėti $r(\theta) = r/n$.

16. Kodėl B kompaktiška?

Funkcija $\theta \mapsto d(\theta, \theta_0)$ tolydi, o B yra uždaros aibės $[\varepsilon; \infty)$ pirmavaizdis tos funkcijos atžvilgiu. Todėl B uždara. KOmpaktiškoje erdvėje visi uždari poaibiai kompaktiški; todėl B kompaktiška.

17. Kodėl $U(\theta, r(\theta))$ rutuliai dengia B?

Kiekvienas $\theta \in B$ yra rutulyje $U(\theta, r(\theta))$.

18. Kodėl teisingi (5.8) sąryšiai?

Todėl, kad iš didžiųjų skaičių dėsnio

$$\overline{m_{U_j}(X)} \to \mathrm{E}\,m_{U_j}(X) < q_j \quad \text{ir} \quad M_n(\theta_0) \to M(\theta_0) > q_j.$$

19. Kaip iš (5.8) išplaukia (5.9)?

$$\begin{aligned}
\mathrm{P}\{\overline{m_{U_j}(X)} < M_n(\theta_0)\} &\geqslant \mathrm{P}\{\overline{m_{U_j}(X)} < q_j < M_n(\theta_0)\} \\
&= 1 - \mathrm{P}\{\overline{m_{U_j}(X)} \geqslant q_j \text{ arba } q_j \geqslant M_n(\theta_0)\} \\
&\geqslant 1 - \mathrm{P}\{\overline{m_{U_j}(X)} \geqslant q_j\} - \mathrm{P}\{q_j \geqslant M_n(\theta_0)\} \to 1.
\end{aligned}$$

20. Kodėl teisingos nelygybės (5.10) grandinėje?

Pirma nelygybė — todėl, kad $\hat{\theta}_n$ yra globalaus $M_n(\theta)$ funkcijos maksimumo taškas. Antra nelygybė — todėl, kad iš $\theta_n \in U_j$ išplaukia $m_\theta(x) \leqslant m_{U_j}(x)$ su visais x ir, reiškia,

$$M_n(\theta) \leqslant \overline{m_{U_j}(X)}.$$

21. Kodėl $\mathrm{P}\{\hat{\theta}_n \in B\} \to 0$ ir kodėl iš to daroma išvada, kad $\hat{\theta}_n \to \theta_0$?

Iš $B \subset U_1 \cup \cdots \cup U_k$ išplaukia

$$\mathrm{P}\{\hat{\theta}_n \in B\} \leqslant \sum_{j=1}^{k} \mathrm{P}\{\hat{\theta}_n \in U_j\} \to 0.$$

Iš B aibės apibrėžimo tada gaunu

$$\mathrm{P}\{d(\hat{\theta}_n, \theta_0) \geqslant \varepsilon\} \to 0.$$

Tai teisinga su bet kokiu ε; todėl $\hat{\theta}_n \to \theta_0$.

5.4 Kitos teoremos

5.4.1 Teoremos

5.4 teorema. *Tarkime, Θ yra intervalas, funkcija $\psi_\theta(x)$ tolydi θ atžvilgiu su visais x, $\hat{\theta}_n$ — vienintelis funkcijos $\Psi_n(\theta)$ nulis ir $\Psi(\theta_0 - \varepsilon) < 0 < \Psi(\theta_0 + \varepsilon)$ su visais ε. Tada $\hat{\theta}_n \to \theta_0$.*

Įrodymas. Fiksuoju ε. Iš paskutiniosios teoremos sąlygos išplaukia, kad

$$\mathrm{P}\{\Psi_n(\theta_0 - \varepsilon) < 0 < \Psi_n(\theta_0 + \varepsilon)\} \to 1. \tag{5.11}$$

Bet jei $\Psi_n(\theta_0 - \varepsilon) < 0 < \Psi_n(\theta_0 + \varepsilon)$, tai $(\theta_0 - \varepsilon; \theta_0 + \varepsilon)$ intervale yra $\Psi_n(\theta)$ funkcijos nulis, t.y. $\hat{\theta}_n$ taškas. Taigi

$$\mathrm{P}\{|\hat{\theta}_n - \theta_0| < \varepsilon\} \to 1.$$

□

5.5 teorema. *Tarkime, Θ yra intervalas ir $\psi_\theta(x)$ — nemažėjanti θ atžvilgiu funkcija su visais x. Jei $\Psi(\theta_0 - \varepsilon) < 0 < \Psi(\theta_0 + \varepsilon)$ su visais ε, o $\Psi_n(\hat{\theta}_n) \to 0$, tai $\hat{\theta}_n \to \theta_0$.*

Įrodymas. Fiksuoju ε. Tada

$$\begin{aligned}
&\mathrm{P}\{|\hat{\theta}_n - \theta_0| \geqslant \varepsilon\} \\
&\leqslant \mathrm{P}\{\hat{\theta}_n \leqslant \theta_0 - \varepsilon\} + \mathrm{P}\{\hat{\theta}_n \geqslant \theta_0 + \varepsilon\} \\
&\leqslant \mathrm{P}\{\Psi_n(\hat{\theta}_n) \leqslant \Psi_n(\theta_0 - \varepsilon)\} + \mathrm{P}\{\Psi_n(\hat{\theta}_n) \geqslant \Psi_n(\theta_0 + \varepsilon)\} \\
&\to 0.
\end{aligned} \tag{5.12}$$

□

5.6 teorema. *Jei*

$$\sup_{\theta \in \Theta} \big| M_n(\theta) - M(\theta) \big| \to 0,$$

su visais ε

$$\sup_{d(\theta, \theta_0) \geqslant \varepsilon} M(\theta) < M(\theta_0)$$

ir $\hat{\theta}_n$ yra $M_n(\theta)$ funkcijos maksimumo taškas, tai $\hat{\theta}_n \to \theta_0$.

Įrodymas. Fiksuoju ε ir pažymiu

$$\delta = M(\theta_0) - \sup_{d(\theta,\theta_0)\geqslant\varepsilon} M(\theta).$$

Jei $d(\hat\theta_n,\theta_0)\geqslant\varepsilon$, tai $M(\hat\theta_n) < M(\theta_0)-\delta$. Todėl

$$\begin{aligned}
&\mathrm{P}\{d(\hat\theta_n,\theta_0)\geqslant\varepsilon\}\\
&\leqslant \mathrm{P}\{M(\hat\theta_n) < M(\theta_0)-\delta\}\\
&\leqslant \mathrm{P}\{M_n(\hat\theta_n) < M(\theta_0)-\delta/2\} + \mathrm{P}\{\big|M_n(\hat\theta_n)-M(\hat\theta_n)\big| > \delta/2\} \qquad (5.13)\\
&\leqslant \mathrm{P}\{M_n(\theta_0) < M(\theta_0)-\delta/2\} + \mathrm{P}\{\sup_\theta\big|M_n(\theta)-M(\theta)\big| \geqslant \delta/2\}\\
&\to 0.
\end{aligned}$$

□

5.4.2 Įrodymų paaiškinimai

5.4 teorema.

22. Paaiškinkite smulkiau, kaip gaunamas (5.11) sąryšis.

Jei $Z_n\to a<b$, tai $\mathrm{P}\{Z_n\geqslant b\}\leqslant \mathrm{P}\{|Z_n-a|\geqslant b-a\}\to 0$. Analogiškai įrodoma, kad jei $Z_n\to a>b$, tai $\mathrm{P}\{Z_n\leqslant b\}\to 0$. Taigi

$$\mathrm{P}\{\Psi_n(\theta_0-\varepsilon)\geqslant 0\}\to 0 \quad \text{ir} \quad \mathrm{P}\{\Psi_n(\theta_0+\varepsilon)\leqslant 0\}\to 0,$$

nes $\Psi_n(\theta_0\pm\varepsilon)\to\Psi(\theta_0\pm\varepsilon)$ ir

$$\Psi(\theta_0-\varepsilon) < 0 < \Psi(\theta_0+\varepsilon).$$

Bet tada

$$\begin{aligned}
&\mathrm{P}\{\Psi_n(\theta_0-\varepsilon)\geqslant 0 \text{ arba } \Psi_n(\theta_0+\varepsilon)\leqslant 0\}\\
&\qquad\qquad \leqslant \mathrm{P}\{\Psi_n(\theta_0-\varepsilon)\geqslant 0\} + \mathrm{P}\{\Psi_n(\theta_0+\varepsilon)\leqslant 0\}\to 0,
\end{aligned}$$

t.y. $\mathrm{P}\{\Psi_n(\theta_0-\varepsilon) < 0 < \Psi_n(\theta_0+\varepsilon)\}\to 1$.

23. Paaiškinkite smulkiau, kodėl $\hat\theta_n\in(\theta_0-\varepsilon;\theta_0+\varepsilon)$.

Jei $\psi_\theta(x)$ tolydi θ atžvilgiu su visais x, tai ir

$$\Psi_n(\theta) = \frac{\psi_\theta(X_1)+\cdots+\psi_\theta(X_n)}{n}$$

funkcija tolydi. Jei $\theta_0-\varepsilon$ taške jos reikšmė neigiama, o $\theta_0+\varepsilon$ taške — teigiama, tai kažkokiame taške $\theta^*\in(\theta_0-\varepsilon;\theta_0+\varepsilon)$ jos reikšmė yra 0. Bet $\hat\theta_n$ yra vienintelis taškas, kuriame funkcija lygi 0. Reiškia, $\hat\theta_n=\theta^*\in(\theta_0-\varepsilon;\theta_0+\varepsilon)$.

5.5 teorema.

24. Kodėl teisingi sąryšiai (5.12) grandinėje?

Pirma nelygybė: jei $|\hat\theta_n-\theta_0|\geqslant\varepsilon$, tai arba $\hat\theta_n\leqslant\theta_0-\varepsilon$, arba $\hat\theta_n\geqslant\theta_0+\varepsilon$. Be to, $\mathrm{P}(A\cup B)\leqslant \mathrm{P}(A)+\mathrm{P}(B)$.

Antra nelygybė: kadangi $\Psi_n(\theta)$ nemažėjanti θ atžvilgiu (nes $\psi_\theta(x)$ nemažėjanti su visais x), iš $\hat{\theta}_n \leqslant \theta_0 - \varepsilon$ išplaukia $\Psi_n(\hat{\theta}_n) \leqslant \Psi_n(\theta_0 - \varepsilon)$, o iš $\hat{\theta}_n \geqslant \theta_0 + \varepsilon$ — $\Psi_n(\hat{\theta}_n) \geqslant \Psi_n(\theta_0 + \varepsilon)$. Be to, jei $A \Rightarrow B$, tai $\mathrm{P}(A) \leqslant \mathrm{P}(B)$.

Trečias sąryšis: iš

$$\Psi_n(\hat{\theta}_n) - \Psi_n(\theta_0 - \varepsilon) \to -\Psi(\theta_0 - \varepsilon) > 0$$

išplaukia $\mathrm{P}\{\Psi_n(\hat{\theta}_n) - \Psi_n(\theta_0 - \varepsilon) \leqslant 0\} \to 0$, o iš

$$\Psi_n(\hat{\theta}_n) - \Psi_n(\theta_0 + \varepsilon) \to -\Psi(\theta_0 + \varepsilon) < 0$$

gaunu $\mathrm{P}\{\Psi_n(\hat{\theta}_n) - \Psi_n(\theta_0 + \varepsilon) \geqslant 0\} \to 0$.

5.6 teorema.

25. Kodėl teisingi sąryšiai (5.13) grandinėje?

Pirma nelygybė: paaiškinta įrodyme.

Antra nelygybė: iš $M(\hat{\theta}_n) < M(\theta_0) - \delta$ išplaukia

$$M_n(\hat{\theta}_n) < M(\theta_0) - \delta/2 \text{ arba } \big|M_n(\hat{\theta}_n) - M(\hat{\theta}_n)\big| > \delta/2,$$

nes jei ir $M_n(\hat{\theta}_n) \geqslant M(\theta_0) - \delta/2$, ir $\big|M_n(\hat{\theta}_n) - M(\hat{\theta}_n)\big| \leqslant \delta/2$, tai

$$M(\hat{\theta}_n) = M_n(\hat{\theta}_n) - \big[M_n(\hat{\theta}_n) - M(\hat{\theta}_n)\big] \geqslant M(\theta_0) - \delta/2 - \delta/2 = M(\theta_0) - \delta.$$

Trečia nelygybė: $M_n(\hat{\theta}_n) \geqslant M_n(\theta_0)$, nes $\hat{\theta}_n$ yra $M_n(\theta)$ funkcijos maksimumo taškas. Todėl

$$M_n(\hat{\theta}_n) < M(\theta_0) - \delta/2 \Rightarrow M_n(\theta_0) < M(\theta_0) - \delta/2$$

ir

$$\mathrm{P}\{M_n(\hat{\theta}_n) < M(\theta_0) - \delta/2\} \leqslant \mathrm{P}\{M_n(\theta_0) < M(\theta_0) - \delta/2\}.$$

Be to, akivaizdu, kad

$$\big|M_n(\hat{\theta}_n) - M(\hat{\theta}_n)\big| > \delta/2 \Rightarrow \sup_\theta \big|M_n(\theta) - M(\theta)\big| > \delta/2;$$

todėl

$$\mathrm{P}\{\big|M_n(\hat{\theta}_n) - M(\hat{\theta}_n)\big| > \delta/2\} \leqslant \mathrm{P}\{\sup_\theta \big|M_n(\theta) - M(\theta)\big| > \delta/2\}.$$

Ketvirtas sąryšis: pirmas dėmuo artėja į 0, nes

$$M_n(\theta_0) \to M(\theta_0) > M(\theta_0 - \delta/2);$$

antras dėmuo — dėl to, kad $\sup_\theta \big|M_n(\theta) - M(\theta)\big| \to 0$.