

Sisukord

1	Üldkogumi kirjeldamine	2
	Ülesanded 1	8
2	Valimhinnangud	8
	Ülesanded 2	11
3	Normaaljaotus	11
	Ülesanded 3	16
4	Ruutvormi jaotus	17
	Ülesanded 4	19
5	Normaaljaotuse parameetrite hindamine	19
6	Wisharti jaotus	22
	Ülesanded 6	27
7	Wisharti jaotuse karakteristik funktsioon ja tõenäosustihedus	27
	Ülesanded 7	29
8	Hotellingi T^2-jaotus	29
	Ülesanded 8	30
9	Hüpoteesid üldkogumi keskväärtuse kohta	31
	Ülesanded 9	34
10	Keskvväärtusvektori usalduspiirkond	34
	Ülesanded 10	38
11	Hüpoteesid dispersioonimaatriksi kohta	39
12	Kahe üldkogumi võrdlemine	43
13	Korrelatsioonimaatriks	47
14	Mitmene korrelatsioonikordaja	50
	Ülesanded 14	56
15	Osakorrelatsioon	56
16	Statistikute asümptootilisest jaotusest	58
LISA 1		64
	L1.1 Maatriksalgebra mõisteid ja tulemusi	64
	L1.2 Plokkmaatriksid	72
LISA 2		78
	L2.1 Vajalikku tõenäosusteooriast	78

1 Üldkogumi kirjeldamine

Statistilise lähenemise korral uuritavale nähtusele või objektile määratletakse esmalt uurijat huvitav üldkogum (näiteks Eesti kooliõpilased, teatud loomaliigi täiskasvanud isendid, teatavat tüüpi ettevõtted jms). Seejärel valitakse välja tunnused, mis kannavad uurimuse eesmärgi saavutamiseks vajalikku informatsiooni. Mitmemõõtmelise analüüsi korral on neid tunnuseid mitu. Kirjeldagu uuritavat üldkogumit p tunnusest moodustatud juhuslik vektor $\vec{X}$:

$$\vec{X} = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_p \end{pmatrix}.$$

Selle juhusliku vektori tõenäosusjaotust nimetatakse *üldkogumi tõenäosusjaotuseks*. Nagu ikka, on tõenäosusjaotuse esitamiseks mitu võimalust. Alati leiduvad *jaotusfunktsioon* $F_{\vec{X}}(\vec{x})$,

$$F_{\vec{X}}(\vec{x}) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_p \leq x_p) \quad (1.1)$$

ja *karakteristlik funktsioon* $\varphi_{\vec{X}}(\vec{t})$,

$$\varphi_{\vec{X}}(\vec{t}) = E(e^{i\vec{t}'\vec{X}}). \quad (1.2)$$

Kui jaotusfunktsioon on absoluutselt pidev, eksisteerib ka *tõenäosustihedus* $f_{\vec{X}}(\vec{x})$,

$$f_{\vec{X}}(\vec{x}) = \frac{\partial^p F_{\vec{X}}(\vec{x})}{\partial x_1 \partial x_2 \dots \partial x_p}. \quad (1.3)$$

Tõenäosustiheduse jaoks kasutame edaspidi ka sünonüümina sõna *tihedus*. Loomulikult saab etteantud tõenäosustiheduse põhjal leida ka talle vastava jaotusfunktsiooni $F_{\vec{X}}(\vec{x})$,

$$F_{\vec{X}}(\vec{x}) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_p} f_{\vec{X}}(\vec{y}) dy_1 dy_2 \dots dy_p.$$

Nagu ühemõõtmelisel juhul juhuslikku suurust, nii iseloomustab ka mitmemõõtmelist juhuslikku vektorit teatav hulk arvkarakteristikuid, *üldkogumi arvkarakteristikuid*. Olulisemad neist on keskväärtuse ja dispersiooni analoogid – paiknemise ja hajuvuse karakteristikud. Juhusliku vektori *keskväärtusvektor* $E\vec{X}$ on konstantne vektor, mille komponentideks on juhusliku vektori komponentide keskväärtused,

$$E\vec{X} = \begin{pmatrix} EX_1 \\ EX_2 \\ \dots \\ EX_p \end{pmatrix}.$$

Eeldame, et keskväärtused EX_i eksisteerivad iga $i \in \{1, \dots, p\}$ korral. Vastasel juhul öeldakse, et $\vec{X}$ keskväärtus puudub. Dispersiooni analoog – juhusliku vektori *dispersioonimaatriks* $D\vec{X}$ – määratakse võrdusega

$$D\vec{X} = \begin{pmatrix} DX_1 & \dots & Cov(X_1, X_p) \\ \dots & \dots & \dots \\ Cov(X_p, X_1) & \dots & DX_p \end{pmatrix}. \quad (1.4)$$

Ka vektori $\vec{X}$ dispersioonimaatriksist saame rääkida vaid siis, kui eksisteerivad kõik kovariatsioonid $Cov(X_i, X_j)$ ja dispersioonid DX_i . Maatriksi $D\vec{X}$ peadiagonaalil on juhusliku vektori komponentide dispersioonid

$$DX_i = E(X_i - EX_i)^2,$$

väljaspool peadiagonaali aga asuvad juhusliku vektori komponentide vahelised kovariatsioonid

$$Cov(X_i, X_j) = E[(X_i - EX_i)(X_j - EX_j)].$$

Edaspidises kasutame ka tähistusi $E\vec{X} = \vec{\mu}$ ja $D\vec{X} = \Sigma$. Dispersioonimaatriksi (1.4) saab esitada keskvväärtusena teatavast juhuslikust maatriksist,

$$D\vec{X} = E[(\vec{X} - \vec{\mu})(\vec{X} - \vec{\mu})'].$$

Korrutades sulgudes olevad tegurid läbi ja koondades ühesugused liikmed, saame dispersioonimaatriksi jaoks teistsuguse avaldise

$$D\vec{X} = E(\vec{X}\vec{X}') - \vec{\mu}\vec{\mu}'.$$

Dispersioonimaatriks on sümmeetriline mittenegatiivselt määratud maatriks.

Kuna mitmemõõtmeline statistika lähtub väga sageli järelduste tegemisel tunnuste vahelistest seostest, on neile kahele karakteristikule lisaks väga oluline informatsioonikandja korrelatsioonimaatriks. *Korrelatsioonimaatriks* $\mathbf{P}$ määratakse võrdusega

$$\mathbf{P} = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1p} \\ \rho_{21} & 1 & \dots & \rho_{2p} \\ \dots & \dots & \dots & \dots \\ \rho_{p1} & \rho_{p2} & \dots & 1 \end{pmatrix}, \quad (1.5)$$

kus ρ_{ij} on tunnuste X_i ja X_j vaheline korrelatsioonikordaja,

$$\rho_{ij} = \frac{Cov(X_i, X_j)}{\sqrt{DX_i DX_j}}.$$

Nagu näha, saab korrelatsioonimaatriksi leida dispersioonimaatriksi elemente kasutades. Arvutuseeskirja esitamiseks maatrikskujul diagonaliseerime dispersioonimaatriksi Σ

$$\Sigma_d = \begin{pmatrix} DX_1 & 0 & \dots & 0 \\ 0 & DX_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & DX_p \end{pmatrix}$$

ning saame maatriksi (1.5) arvutuseeskirja

$$\mathbf{P} = \Sigma_d^{-1/2} \Sigma \Sigma_d^{-1/2}.$$

Siin ja edaspidi eeldame, et diagonaliseerimine rakendatakse maatriksile enne teisi maatriksoperatsioone. Korrelatsioonimaatriks on sümmeetriline mittenegatiivselt määratud maatriks, tema peadiagonaalil asuvad alati ühed.

Klassikaline mitmemõõtmeline statistika on üles ehitatud eeldusel, et üldkogumi jaotuseks on mitmemõõtmeline normaaljaotus. Seda jaotust vaatame lähemalt ülejärgmises paragrahvis, siin aga tutvume ühemõõtmelise jaotuse kahe arvkarakteristiku mitmemõõtmelise üldistusega. Normaaljaotus on keskvväärtuse suhtes sümmeetriline, tegelikkuses ei ole aga sugugi alati tegemist sümmeetriliste jaotustega. Jaotuse ebasümmeetriat mõõdab ühemõõtmelisel juhul *asümmeetriakordaja*, mis määratakse eeskirjaga

$$\beta_1 = \frac{E(X - EX)^3}{(\sqrt{E(X - EX)^2})^3}.$$

Jaotuse järsakust ja sabade raskust mõõdab *ekstsess e. järsakus*, mis ühemõõtmelisel juhul määratakse eeskirjaga

$$\beta_2 = \frac{E(X - EX)^4}{((E(X - EX)^2)^2)}.$$

Nende karakteristikute jaoks on mitmemõõtmelisel juhul välja pakutud erinevaid üldistusi. Enimkasutatavad on artiklis Mardia (1970) esitatud definitsioonid. Vektori $\vec{X}$ *asümmeetriakordaja* on määratud eeskirjaga

$$\beta_{1,p}(\vec{X}) = E[(\vec{X} - \vec{\mu})' \Sigma^{-1} (\vec{Y} - \vec{\mu})]^3,$$

kus $\vec{X}$ ja $\vec{Y}$ on sõltumatud sama jaotusega juhuslikud vektorid. Vektori $\vec{X}$ jäsakus on aga defineeritud võrdusega

$$\beta_{2,p}(\vec{X}) = E[(\vec{X} - \vec{\mu})' \Sigma^{-1} (\vec{X} - \vec{\mu})]^2.$$

Kui $p = 1$, saame need juhusliku vektori karakteristikud siduda vastavate karakteristikutega juhusliku suuruse jaoks, $\beta_{1,1} = \beta_1^2$ ja $\beta_{2,1} = \beta_2$.

Mardia karakteristikute olemus avaldub paremini järgmise esituse kaudu. Tähistame

$$\vec{Y} = \Sigma^{-1/2}(\vec{X} - \vec{\mu}).$$

Siis $E(\vec{Y}) = \vec{0}$ ja $D(\vec{Y}) = \mathbf{I}_p$. Kursusest "Maatriksid statistikas" teame, et juhusliku vektori momendid avalduvad otsekorrutise abil:

$$m_3(\vec{Y}) = E(\vec{Y} \otimes \vec{Y}' \otimes \vec{Y})$$

ja

$$m_4(\vec{Y}) = E(\vec{Y} \otimes \vec{Y}' \otimes \vec{Y} \otimes \vec{Y}').$$

Saab näidata (Kollo and Srivastava (2004)), et

$$\beta_{1,p}(\vec{X}) = \text{tr}[m_3(\vec{Y})m_3(\vec{Y})']$$

ja

$$\beta_{2,p}(\vec{X}) = \text{tr}[m_4(\vec{Y})].$$

Seega avaldub Mardia asümmeetriakordaja vektori $\vec{X}$ tsentreerimisel ja normeerimisel saadud vektori $\vec{Y}$ koordinaatide kolmandat järku segamomentide ruutude summana, jäsakus $\beta_{2,p}(\vec{X})$ on aga vektori $\vec{Y}$ neljandate segamomentide maatriksi jälg. Kui tähistame ühemõõtmelisel juhul

$$Y = (DX)^{-1/2}(X - \mu),$$

siis

$$\beta_1 = E(Y)^3$$

ja

$$\beta_2 = E(Y)^4.$$

Mardia kordajad $\beta_{1,p}$ ja $\beta_{2,p}$ on invariantseid lineaarteisenduse $\vec{Y} = A\vec{X} + \vec{b}$ suhtes.

Tuletame meelde veel mõned tõenäosusteooria tulemused. Kui juhusliku vektori tihedus $f_{\vec{X}}(x_1, x_2, \dots, x_p)$ on teada, siis on võimalik arvutada suvalise hulga B jaoks tõenäosus juhusliku vektori sattumiseks sellesse hulka,

$$P(\vec{X} \in B) = \underbrace{\int \dots \int}_B f_{\vec{X}}(\vec{y}) d\vec{y}. \quad (1.6)$$

Võtame hulgaks B p -tahuka B_Δ ,

$$B_\Delta = [x_1, x_1 + \Delta x_1] \times [x_2, x_2 + \Delta x_2] \times \dots \times [x_p, x_p + \Delta x_p],$$

millesse sattumise tõenäosus on

$$P(\vec{X} \in B_\Delta) = \int_{x_1}^{x_1 + \Delta x_1} \dots \int_{x_p}^{x_p + \Delta x_p} f_{\vec{X}}(y_1, y_2, \dots, y_p) dy_1 \dots dy_p.$$

Kui juurdekasvud Δx_i on väikesed, saame siit ligikaudse võrduse

$$P(\vec{X} \in B_\Delta) \approx f_{\vec{X}}(x_1, x_2, \dots, x_p) \Delta x_1 \Delta x_2 \dots \Delta x_p.$$

Mitmemõõtmelise jaotuse korral on olulised marginaaljaotused ja tinglikud jaotused.

Definitsioon 1.1. Juhusliku p -vektori $\vec{X}$ k -mõõtmeliseks ($k < p$) marginaaljaotuseks nimetatakse selle vektori k koordinaadi k -mõõtmelist ühisjaotust.

Erinevaid k -mõõtmelisi marginaaljaotusi saab konstrueerida nii palju, kui palju on võimalik valida erinevaid k -elemendilisi kombinatsioone p elemendi hulgast, seega on erinevate k -mõõtmeliste marginaaljaotuste arv C_p^k . Oletame, et oleme välja valinud k esimest komponenti ja vaatame, kuidas leida nende marginaaljaotust ühisjaotuse kaudu. Leiame esmalt k -mõõtmelise marginaaljaotusfunktsiooni $F_{X_1, \dots, X_k}(x_1, x_2, \dots, x_k)$,

$$\begin{aligned} F_{X_1, \dots, X_k}(x_1, \dots, x_k) &= P(X_1 \leq x_1, \dots, X_k \leq x_k) \\ &= P(X_1 \leq x_1, \dots, X_k \leq x_k, X_{k+1} \leq \infty, \dots, X_p \leq \infty) \\ &= F_{\vec{X}}(x_1, \dots, x_k, \infty, \dots, \infty). \end{aligned}$$

Ka k -mõõtmelist marginaaltihedust $f_{X_1, \dots, X_k}(x_1, \dots, x_k)$ saab leida vektori $\vec{X}$ tihedusest,

$$\begin{aligned} &f_{X_1, \dots, X_k}(x_1, \dots, x_k) \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{\vec{X}}(x_1, \dots, x_k, x_{k+1}, \dots, x_p) dx_{k+1} \dots dx_p. \end{aligned}$$

Juhusliku vektori komponendid võivad olla sõltumatud. Sõltumatus on defineeritud jaotusfunktsiooni kaudu.

Definitsioon 1.2. Juhusliku vektori $\vec{X}$ komponendid on vastastikku sõltumatud, kui vektori jaotusfunktsioon (1.1) avaldub korrutisena

$$F_{\vec{X}}(x_1, \dots, x_p) = F_{X_1}(x_1) \times \dots \times F_{X_p}(x_p).$$

Definitsioonist järeldub, et vastastikku sõltumatute komponentide korral on ka vektori tihe-
dus (1.3) avalduv analoogilise korrutisena.

Samamoodi saab määrata ka juhusliku vektori alamvektorite sõltumatuse. Juhusliku vektori komponendid $X_1, \dots, X_k$ on sõltumatud komponentidest $X_{k+1}, \dots, X_p$, kui

$$F_{\vec{X}}(x_1, \dots, x_p) = F_{X_1 \dots X_k}(x_1, \dots, x_k) F_{X_{k+1} \dots X_p}(x_{k+1}, \dots, x_p).$$

Järgmine oluline juhusliku vektoriga seotud mõiste on *tinglik jaotus*. Käsitleme seda mõistet pidevate juhuslike vektorite korral ja defineerime tingliku jaotuse tõenäosustiheduse kaudu. Vaatame lihtsuse mõttes esmalt kahemõõtmelist juhtu ($p = 2$). Lähtume tingliku tõenäosuse valemist

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad (1.7)$$

kus $P(B) \neq 0$. Valime sündmuse A , $A = \{a \leq X_1 \leq b\}$, ja sündmuse B , $B = \{c \leq X_2 \leq d\}$. Siis sündmuse A tingliku tõenäosuse tingimusel B saame leida valemi (1.7) abil,

$$P(A|B) = \frac{P(a \leq X_1 \leq b, c \leq X_2 \leq d)}{P(c \leq X_2 \leq d)} = \frac{\int_a^b \int_c^d f_{X_1 X_2}(u, v) du dv}{\int_c^d f_{X_2}(v) dv}.$$

Kui B määrame lühikese lõigu abil, $B = \{c \leq X_2 \leq c + \Delta c\}$, siis pidevuse tõttu kehtib keskväärtusteoreem

$$\int_c^{c+\Delta c} f_{X_2}(v) dv = f_{X_2}(v_*) \Delta c,$$

kus $v_* \in [c, c + \Delta c]$. Samal põhjusel rahuldab murrujoone peal olev integraal tingimust

$$\int_c^{c+\Delta c} f_{X_1 X_2}(u, v) dv = f_{X_1 X_2}(u, v_*(u)) \Delta c,$$

kus $v_*(u) \in [c, c + \Delta c]$. Seega oleme saanud võrduse

$$P(A|B) = \int_a^b \frac{f_{X_1 X_2}(u, v_*(u))}{f_{X_2}(v_*)} du.$$

Kui c ja Δc on fikseeritud, siis on integreeritav funktsioon viimases võrduses ühemuutuja funktsioon.

Olgu c argumendi väärtus, mille korral $f_{X_2}(c) > 0$. Defineerime tõenäosuse

$$P(a \leq X_1 \leq b | X_2 = c) = \lim_{\Delta c \rightarrow 0} P(a \leq X_1 \leq b | c \leq X_2 \leq c + \Delta c) = \int_a^b \frac{f_{X_1 X_2}(u, c)}{f_{X_2}(c)} du.$$

Saadud integraalialust funktsiooni kasutatakse X_1 tingliku tiheduse defineerimisel, tingimusel $X_2 = c$. Asendame arvud u ja c arvudega x_1 ja x_2 ning esitame nendes tähistustes tingliku tõenäosustiheduse definitsiooni.

Definitsioon 1.3. Olgu juhusliku vektori $(X_1, X_2)'$ tõenäosustihedus $f_{X_1 X_2}(x_1, x_2)$. Juhusliku suuruse X_1 tinglikuks tõenäosustiheduseks tingimusel $X_2 = x_2$ nimetatakse funktsiooni

$$f_{X_1}(x_1 | x_2) = \frac{f_{X_1 X_2}(x_1, x_2)}{f_{X_2}(x_2)} = \frac{f_{X_1 X_2}(x_1, x_2)}{\int_{-\infty}^{\infty} f_{X_1 X_2}(x_1, x_2) dx_1}.$$

Seega saame $(X_1, X_2)'$ ühisjaotuse tiheduse esitada korrutisena

$$f_{X_1 X_2}(x_1, x_2) = f_{X_1}(x_1 | x_2) f_{X_2}(x_2).$$

Mitmemõõtmelisel juhul viib analoogiline mõttekäik meid sarnasele tulemusele pärast kor-duvat rakendamist:

$$f_{X_1 \dots X_k}(x_1, \dots, x_k | x_{k+1}, \dots, x_p) = \frac{f_{\vec{X}}(x_1, \dots, x_p)}{\underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}}_k f_{\vec{X}}(x_1, \dots, x_k, x_{k+1}, \dots, x_p) dx_1 \dots dx_k}.$$

Siit tekib võimalus esitada $(X_1, \dots, X_p)$ ühisjaotuse tihedus korrutisena

$$f_{\vec{X}}(x_1, \dots, x_p) = f_{X_1}(x_1) f_{X_2}(x_2 | x_1) f_{X_3}(x_3 | x_1, x_2) \dots f_{X_p}(x_p | x_1, \dots, x_{p-1}),$$

kusjuures

$$\begin{aligned} f_{X_1}(x_1) &= \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}}_{p-1} f_{\vec{X}}(x_1, \dots, x_p) dx_2 \dots dx_p, \\ f_{X_2}(x_2 | x_1) &= \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}}_{p-2} \frac{f_{\vec{X}}(x_1, \dots, x_p) dx_3 \dots dx_p}{f_{X_1}(x_1)}, \\ f_{X_3}(x_3 | x_1, x_2) &= \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}}_{p-3} \frac{f_{\vec{X}}(x_1, \dots, x_p) dx_4 \dots dx_p}{f_{X_1}(x_1) f_{X_2}(x_2 | x_1)}, \end{aligned}$$

ning lõpuks

$$f_{X_p}(x_p | x_1, \dots, x_{p-1}) = \frac{f_{\vec{X}}(x_1, \dots, x_p)}{f_{X_1}(x_1) \dots f_{X_{p-1}}(x_{p-1} | x_1, \dots, x_{p-2})}.$$

Ühistiheduse $f_{\vec{X}}(\vec{x})$ esitus tinglike tõenäosustiheduste kaudu ei ole ühene. Muutujate vahetamiseks on siin kokku $p!$ võimalust, mis tähendab sama palju võimalusi tõenäosustiheduse esitamiseks.

Nagu ühemõõtmelisel juhul, saab ka p -mõõtmelise juhusliku vektori funktsiooni tõenäosustiheduse esitada argumendiks oleva juhusliku vektori tiheduse kaudu. Tuletame järgnevalt vastava valemi.

Teoreem 1.1. Olgu $\vec{X}$ p -mõõtmeline juhuslik vektor tihedusega $f_{\vec{X}}(\vec{x})$ ja olgu juhuslik p -vektor $\vec{Y}$ selle vektori funktsioon:

$$Y_i = y_i(X_1, X_2, \dots, X_p), \quad i = 1, 2, \dots, p,$$

kus

$$y_i = y_i(x_1, x_2, \dots, x_p), \quad i = 1, 2, \dots, p \quad (1.8)$$

on üksühene teisendus $\mathbb{R}^p \rightarrow \mathbb{R}^p$.

Vektori $\vec{Y}$ tihedus $f_{\vec{Y}}(\vec{y})$ avaldub siis kujul

$$f_{\vec{Y}}(\vec{y}) = f_{\vec{X}}(x_1(\vec{y}), x_2(\vec{y}), \dots, x_p(\vec{y})) \cdot J \left(\frac{x_1, x_2, \dots, x_p}{y_1, y_2, \dots, y_p} \right),$$

kus $J \left(\frac{x_1, x_2, \dots, x_p}{y_1, y_2, \dots, y_p} \right)$ on $\vec{Y}$ komponente määrava teisenduse (1.8) jakobiaan.

Tõestus. Kuna eelduse kohaselt teisendus (1.8) on üksühene, leidub tal kindlasti ka pöördteisendus $x_i = x_i(y_1, y_2, \dots, y_p)$, $i = 1, 2, \dots, p$. Jakobiaani definitsiooni kohaselt

$$J \left(\frac{x_1, x_2, \dots, x_p}{y_1, y_2, \dots, y_p} \right) = \det \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \dots & \frac{\partial x_1}{\partial y_p} \\ \dots & \dots & \dots \\ \frac{\partial x_p}{\partial y_1} & \dots & \frac{\partial x_p}{\partial y_p} \end{vmatrix}.$$

Võtame vaatluse alla suvalise mõõtuva hulga $A \subset \mathbb{R}^p$. Siis tõenäosus vektori $\vec{Y}$ sattumiseks sellesse hulka avaldub järgmiselt

$$P(\vec{Y} \in A) = \underbrace{\int \dots \int}_A f_{\vec{Y}}(\vec{y}) dy_1 \dots dy_p.$$

Samal ajal iga mõõtuva hulga B korral kehtib vektori $\vec{X}$ jaoks analoogiline tulemus

$$P(\vec{X} \in B) = \underbrace{\int \dots \int}_B f_{\vec{X}}(\vec{x}) dx_1 \dots dx_p.$$

Olgu nüüd B_2 ruumi $\mathbb{R}^p$ alamhulk, milleks teisendus (1.8) teisendab mõõtuva hulga B_1 . Matemaatilisest analüüsist teame, et

$$\underbrace{\int \dots \int}_{B_1} f_{\vec{X}}(\vec{x}) d\vec{x} = \underbrace{\int \dots \int}_{B_2} f_{\vec{X}}(\vec{x}(\vec{y})) J \left(\frac{x_1, x_2, \dots, x_p}{y_1, y_2, \dots, y_p} \right) d\vec{y}.$$

Juhuslike vektorite jaoks saame siit võrduste ahela:

$$\begin{aligned} P(\vec{X} \in B_1) &= P(\omega \mid \{X_1(\omega), \dots, X_p(\omega)\} \in B_1) \\ &= P(\omega \mid \vec{X}(\omega) \in B_1) = \underbrace{\int \dots \int}_{B_1} f_{\vec{X}}(\vec{x}) d\vec{x} \\ &= P(\omega \mid \{y_1(\vec{X}(\omega)), \dots, y_p(\vec{X}(\omega))\} \in B_2) \\ &= P(\vec{Y} \in B_2) = \underbrace{\int \dots \int}_{B_2} f_{\vec{Y}}(\vec{y}) d\vec{y} \\ &= \underbrace{\int \dots \int}_{B_2} f_{\vec{X}}(x_1(\vec{y}), \dots, x_p(\vec{y})) J \left(\frac{\vec{x}}{\vec{y}} \right) d\vec{y}, \end{aligned}$$

kus ω tähistab elementaarsündmust. Seega saime valemi, mis suvalise mõõtuva hulga B_2 jaoks võimaldab leida tõenäosuse, et vektor $\vec{Y}$ sattub sellesse hulka. Vastavalt valemile (1.6) peab integreeritav funktsioon olema vektori $\vec{Y}$ tõenäosustiheduseks, mott.

Ülesanded 1

1.1. Näidata, et korrelatsioonimaatriks $\mathbf{P}$ avaldub kujul

$$\mathbf{P} = \mathbf{\Sigma}_d^{-1/2} \mathbf{\Sigma} \mathbf{\Sigma}_d^{-1/2}.$$

1.2. Kontrollida võrdused

$$\beta_{1,p}(\vec{X}) = \text{tr}[m_3(\vec{Y})m_3(\vec{Y})'],$$

$$\beta_{2,p}(\vec{X}) = \text{tr}[m_4(\vec{Y})],$$

kus

$$\vec{Y} = \mathbf{\Sigma}^{-1/2}(\vec{X} - \vec{\mu}).$$

1.3. Näidata, et kordajad $\beta_{1,p}$ ja $\beta_{2,p}$ on invariantse lineaarteisenduse $\vec{Y} = A\vec{X} + \vec{b}$ suhtes.

1.4. Olgu $\vec{X} = (X_1, X_2)'$ tõenäosustihedusega

$$f_{\vec{X}}(x_1, x_2) = \begin{cases} 6x_1, & \text{kui } x_1 + x_2 < 1, \ x_1, x_2 > 0, \\ 0, & \text{mujal.} \end{cases}$$

Leida marginaaljaotuste tihedused ja tinglikud tõenäosustihedused.

1.5. Olgu $\vec{X} = (X_1, X_2)'$ tõenäosustihedusega $f_{\vec{X}}(x_1, x_2)$ ja $\vec{Y} = \mathbf{A}\vec{X}$, kus

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Leida $f_{\vec{Y}}(y_1, y_2)$.

1.6. Olgu $\vec{X} = (X_1, X_2, X_3)'$ tõenäosustihedusega $f_{\vec{X}}(\vec{x})$ ja $\vec{Y} = (3X_1, X_1 - 4X_2, X_3)'$.
Leida $f_{\vec{Y}}(\vec{y})$.

1.7. Olgu $D\vec{X} = \mathbf{I}_2$ ja $\mathbf{A} = \begin{pmatrix} 1 & 1 \end{pmatrix}$; $\mathbf{B} = \begin{pmatrix} 1 & -2 \end{pmatrix}$. Kas $\mathbf{A}\vec{X}$ ja $\mathbf{B}\vec{X}$ on korreleeritud? Kui jah, siis leida korrelatsioonikordaja.

2 Valimhinnangud

Üldkogumi karakteristikud on reeglina tundmatud. Informatsiooni saamiseks üldkogumi kohta tuleb üldkogumi juhuslikult valitud objektidel mõõta uuritava p -mõõtmelise tunnusvektori $\vec{X}$ väärtused. Olgu seda vektorit mõõdetud n korda. Mõõtmistulemused moodustavad *valimi* (konkreetses valimi), milleks on $p \times n$ -maatriks $\mathcal{X}$,

$$\mathcal{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{p1} & x_{p2} & \dots & x_{pn} \end{pmatrix}.$$

Selle maatriksi esimeses reas on esimese mõõdetud tunnuse väärtused, teises reas – teise mõõdetud tunnuse väärtused jne. Maatriksi esimeses veerus aga asuvad kõikide tunnuste väärtused esimesel mõõtmisel, teises veerus – kõikide tunnuste väärtused teisel mõõtmisel jne. Matemaatiline statistika eeldab, et valim on juhuslik: ühest ja samast üldkogumist tehtud erinevad valimid on erinevad. Valimi juhuslikkuse kirjeldamiseks võetakse kasutusele spetsiaalne mudel: *teoreetiline valim*. Teoreetiline valim on juhuslik maatriks,

$$\mathbf{X} = [\vec{X}_1 \vec{X}_2 \dots \vec{X}_n],$$

mille veeruvektorid $\vec{X}_1, \dots, \vec{X}_n$ rahuldavad kahte tingimust

1) $\forall i, j$ korral, $i \neq j$, on veeruvektorid $\vec{X}_i$ ja $\vec{X}_j$ sõltumatud;

2) kõik veeruvektorid on sama jaotusega, mis kirjeldab üldkogumit: $\vec{X}_i \stackrel{d}{=} \vec{X}$.

Teoreetilise valimi mõõtuvaid funktsioone nimetatakse *statistikuteks*. Sobivalt valitud statistikut saab kasutada üldkogumi parameetri hindamiseks. Hinnangu omadusi (nihketus jm) saab kindlaks teha ainult vastavat statistikut uurides. Konkreetse valimi põhjal arvutatud statistiku väärtus on arvuliseks hinnanguks üldkogumi parameetrile.

Nii ühemõõtmelisel kui ka mitmemõõtmelisel juhul on üldkogumi keskvaartuse hinnanguks valimi keskvaartus. *Valimi keskvaartus* on p -mõõtmeline vektor, mille komponentideks on tunnuste aritmeetilised keskmised,

$$\vec{x} = \begin{pmatrix} \frac{1}{n} \sum_{i=1}^n x_{1i} \\ \frac{1}{n} \sum_{i=1}^n x_{2i} \\ \dots \\ \frac{1}{n} \sum_{i=1}^n x_{pi} \end{pmatrix} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \dots \\ \bar{x}_p \end{pmatrix}.$$

Tähistades ühtedest koosneva n -vektori sümboliga $\vec{\mathbb{1}}_n$, saame esitada valimi keskvaartuse matrikskujul

$$\vec{x} = \frac{1}{n} \mathbf{X} \vec{\mathbb{1}}_n.$$

Statistiku $\vec{X}$ üldkogumi keskvaartuse hindamiseks saame, kui arvutuseeskirjas asendame konkreetse valimi $\mathcal{X}$ teoreetilise valimiga $\mathbf{X}$

$$\vec{X} = \frac{1}{n} \mathbf{X} \vec{\mathbb{1}}_n. \quad (2.1)$$

Dispersioonimaatriksi hinnangu saame, kui kõik tema elemendid asendame ühemõõtmelisest statistikast tuntud hinnangutega. Maatrikskujul saame selle hinnangu esitada konkreetse valimi veeruvektorite ja valimi keskvaartuse abil

$$\hat{\mathbf{S}}_* = \frac{1}{n} \sum_{i=1}^n (\vec{x}_i - \vec{x})(\vec{x}_i - \vec{x})',$$

kust pärast sulgude avamist saame

$$\hat{\mathbf{S}}_* = \frac{1}{n} \sum_{i=1}^n \vec{x}_i \vec{x}_i' - \vec{x} \vec{x}'.$$

Statistiku üldkogumi dispersioonimaatriksi hindamiseks saame, kui arvutuseeskirjas asendame konkreetse valimi $\mathcal{X}$ teoreetilise valimiga $\mathbf{X}$

$$\mathbf{S}_* = \frac{1}{n} \sum_{i=1}^n \vec{X}_i \vec{X}_i' - \vec{X} \vec{X}'.$$

Näitamaks, et statistikut $\mathbf{S}_*$ kasutades saame üldkogumi dispersioonimaatriksi $D\vec{X}$ jaoks nihkega hinnangu, arvutame keskvaartuse:

$$E\mathbf{S}_* = E \left(\frac{1}{n} \sum_{i=1}^n \vec{X}_i \vec{X}_i' - \vec{X} \vec{X}' \right) = \frac{1}{n} \sum_{i=1}^n E(\vec{X}_i \vec{X}_i') - E(\vec{X} \vec{X}').$$

Rakendades nüüd viimase liikme lahti kirjutamiseks valemit (2.1) ja arvestades teoreetilise valimi veergude sõltumatust, saame

$$E(\vec{X} \vec{X}') = \frac{1}{n^2} \left[\sum_{i=1}^n E(\vec{X}_i \vec{X}_i') + \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n E(\vec{X}_i) E(\vec{X}_j') \right],$$

kust tuleneb kõikide teoreetilise valimi veeruvektorite sama jaotust arvestades

$$E\mathbf{S}_* = \frac{n-1}{n}D\vec{X} = \frac{n-1}{n}\mathbf{\Sigma}.$$

Kasutades teoreetilist valimit $\mathbf{X}$ saame statistiku $\mathbf{S}_*$ jaoks anda järgmise esituse:

$$\mathbf{S}_* = \frac{1}{n}\mathbf{X}(\mathbf{I}_n - \frac{1}{n}\vec{\mathbb{1}}_n\vec{\mathbb{1}}_n')\mathbf{X}'.$$

Selles avaldises esinevat maatriksit $\mathbf{I}_n - \frac{1}{n}\vec{\mathbb{1}}_n\vec{\mathbb{1}}_n'$ nimetatakse *tsentreerimismaatriksiks*. Tsentreerimismaatriks on idempotentne ja sümmeetriline.

Asendades statistiku $\mathbf{S}_*$ avaldises teoreetilise valimi $\mathbf{X}$ konkreetse valimiga $\mathcal{X}$, saame hinnangu

$$\hat{\mathbf{S}}_* = \frac{1}{n}\mathcal{X}(\mathbf{I}_n - \frac{1}{n}\vec{\mathbb{1}}_n\vec{\mathbb{1}}_n')\mathcal{X}'. \quad (2.2)$$

Paneme tähele, et äsja leitud nihke avaldis on kooskõlas meie seniste teadmistega. Maatriksi $\hat{\mathbf{S}}_*$ elemendid avalduvad dispersiooni ja kovariatsioonide nihkega hinnangutena:

$$(\hat{\mathbf{S}}_*)_{jj} = \widehat{DX_j} = \frac{1}{n} \sum_{i=1}^n (x_{ji} - \bar{x}_j)^2$$

ja

$$\begin{aligned} (\hat{\mathbf{S}}_*)_{ij} &= \widehat{Cov(X_i, X_j)} = \frac{1}{n} \sum_{k=1}^n (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j) \\ &= \frac{1}{n} \sum_{k=1}^n x_{ik}x_{jk} - \bar{x}_i\bar{x}_j. \end{aligned}$$

Arvestades $\mathbf{S}_*$ nihke avaldist saame leida statistiku dispersioonimaatriksi nihketa hinnangute leidmiseks:

$$\mathbf{S} = (S_{ij}) = \frac{1}{n-1} \sum_{i=1}^n (\vec{X}_i - \vec{X})(\vec{X}_i - \vec{X})'. \quad (2.3)$$

Nihketa hinnangut dispersioonimaatriksile nimetatakse *valimi dispersioonimaatriksiks* ja tähistatakse $\hat{\mathbf{S}}$

$$\hat{\mathbf{S}} = (s_{ij}) = \frac{1}{n-1} \sum_{i=1}^n (\vec{x}_i - \vec{x})(\vec{x}_i - \vec{x})'.$$

Korrelatsioonimaatriksi hindamiseks leitakse valimi põhjal korrelatsioonide hinnangud

$$r_{ij} = \frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}},$$

kus s_{ij} on valimi dispersioonimaatriksi elemendid, ja moodustatakse neist *valimi korrelatsioonimaatriks*

$$\hat{\mathbf{R}} = \begin{pmatrix} 1 & r_{11} & \dots & r_{1p} \\ r_{12} & 1 & \dots & r_{2p} \\ \dots & \dots & \dots & \dots \\ r_{1p} & r_{2p} & \dots & 1 \end{pmatrix}.$$

Sarnaselt teoreetilise korrelatsioonimaatriksiga saame ka valimi korrelatsioonimaatriksi esitada maatrikskujul, kasutades valimi dispersioonimaatriksit ja selle diagonaliseeritud kuju $\hat{\mathbf{S}}_d$:

$$\hat{\mathbf{R}} = \hat{\mathbf{S}}_d^{-1/2} \hat{\mathbf{S}} \hat{\mathbf{S}}_d^{-1/2}.$$

Vastav statistik $\mathbf{R}$ avaldub analoogiliselt juhusliku maatriksi $\mathbf{S}$ kaudu:

$$\mathbf{R} = \mathbf{S}_d^{-1/2} \mathbf{S} \mathbf{S}_d^{-1/2}.$$

Mitmemõõtmelise asümmeetriakordaja hinnang määratakse võrdusega

$$\hat{\beta}_{1,p}(\vec{X}) = \frac{1}{n^2} \sum_{i,j=1}^n [(\vec{x}_i - \vec{x})' \hat{\mathbf{S}}^{-1} (\vec{x}_j - \vec{x})]^3.$$

Hinnang mitmemõõtmelisele järsakusele määratakse eeskirjaga

$$\hat{\beta}_{2,p}(\vec{X}) = \frac{1}{n} \sum_{i=1}^n [(\vec{x}_i - \vec{x})' \hat{\mathbf{S}}^{-1} (\vec{x}_i - \vec{x})]^2.$$

Ülesanded 2

2.1. Kontrollida, et tsentreerimismaatriks $\mathbf{I}_n - \frac{1}{n} \vec{\mathbf{1}}_n \vec{\mathbf{1}}_n'$ on idempotentne maatriks.

2.2. Näidata, et

$$\sum_{i=1}^n \vec{X}_i \vec{X}_i' = \mathbf{X} \mathbf{X}'.$$

2.3. Näidata, et

$$\mathbf{S} = \frac{1}{n-1} \sum_{i=1}^n \vec{X}_i \vec{X}_i' - \frac{n}{n-1} \vec{X} \vec{X}'.$$

3 Normaalkaotus

Selles paragrahvis käsitleme mitmemõõtmelise statistika põhilist mudelit – mitmemõõtmelist normaalkaotust. Kahtlemata on mitmemõõtmeline normaalkaotus ühemõõtmelise normaalkaotuse üldistuseks. Ühemõõtmelisel juhul alustatakse tavaliselt standardse normaalkaotusega juhuslikust suurusest Z , mille tihedus määratakse eeskirjaga

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}.$$

Üldise normaalkaotusega juhusliku suuruse X saame sellest lineaarteisenduse teel

$$X = \mu + \sigma Z,$$

kusjuures $EX = \mu$ ja $DX = \sigma^2$. Lühidalt märgime ühemõõtmelise normaalkaotuse kirjutisega $X \sim N(\mu, \sigma^2)$, tema tõenäosustihedus

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}.$$

Kanname selle konstruktsiooni üle mitmemõõtmelisele juhule. Moodustame esmalt standardse normaalkaotuse mitmemõõtmelise analoogi. Selleks on sõltumatute komponentidega juhuslik vektor $\vec{Z}$, mille iga komponent on standardse normaalkaotusega,

$$\vec{Z} = \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_p \end{pmatrix},$$

kus $Z_i \sim N(0, 1)$ ja $\forall i, j$, korral $Z_i \perp Z_j$, kui $i \neq j$. Me teame, et sõltumatute komponentidega juhusliku vektori korral ühistihedus avaldub marginaaltiheduste korrutisena. Järelikult

$$f_{\vec{Z}}(\vec{z}) = f_{Z_1}(z_1) f_{Z_2}(z_2) \dots f_{Z_p}(z_p) = \frac{1}{(\sqrt{2\pi})^p} e^{-\frac{1}{2} \sum_{i=1}^p z_i^2},$$

ehk

$$f_{\vec{Z}}(\vec{z}) = \frac{1}{(\sqrt{2\pi})^p} e^{-\frac{1}{2}\vec{z}'\vec{z}}.$$

Seda jaotust nimetatakse *standardseks p -mõõtmeliseks normaaljaotuseks*, mida tähistame kujul $\vec{Z} \sim N(\vec{0}, \mathbf{I}_p)$. Nüüd defineerime mitmemõõtmelise normaaljaotuse sama võttega nagu ühemõõtmelisel juhul.

Definitsioon 3.1. Olgu $\vec{Z} \sim N(\vec{0}, \mathbf{I}_p)$ ja $\mathbf{A}$ pööratav $p \times p$ -maatriks, kusjuures $\mathbf{A}\mathbf{A}' = \mathbf{\Sigma}$. Siis juhuslik vektor

$$\vec{X} = \vec{\mu} + \mathbf{A}\vec{Z}$$

on p -mõõtmelise normaaljaotusega parameetritega $\vec{\mu}$, ($\vec{\mu} \in \mathbb{R}^p$) ja $\mathbf{\Sigma}$, ($\mathbf{\Sigma}$ on positiivselt määratud $p \times p$ -maatriks).

Niisuguse normaaljaotuse korral kasutame tähistust $\vec{X} \sim N_p(\vec{\mu}, \mathbf{\Sigma})$ või, kui jaotuse dimensiooni tähistamine pole vajalik, $\vec{X} \sim N(\vec{\mu}, \mathbf{\Sigma})$.

MÄRKUS. Definitsioonis 3.1 määratud normaaljaotuse üldistusega on tegemist siis, kui $\mathbf{A}$ on suvaline $p \times p$ -maatriks. Sel juhul on maatriks $\mathbf{\Sigma} = \mathbf{A}\mathbf{A}'$ mittenegatiivselt määratud. Juhul, kui $r(\mathbf{A}) < p$, on tegemist *kõdunud normaaljaotusega*, $r(\mathbf{A}) = p$ korral rakendub definitsioon 3.1.

Lause 3.1. Olgu $\vec{X} \sim N_p(\vec{\mu}, \mathbf{\Sigma})$. Siis $E\vec{X} = \vec{\mu}$ ja $D\vec{X} = \mathbf{\Sigma}$.

Tõestus. Definitsiooni 3.1 tõttu eksisteerivad $\vec{\mu}$, ($\vec{\mu} \in \mathbb{R}^p$) ja $\mathbf{A} : p \times p$ nii, et $\vec{X} = \vec{\mu} + \mathbf{A}\vec{Z}$, kus $\vec{Z}$ on standardse p -mõõtmelise normaaljaotusega, $\vec{Z} \sim N(\vec{0}, \mathbf{I}_p)$. Juhusliku vektori $\vec{Z}$ keskvääruseks on nullvektor, $E\vec{Z} = \vec{0}$, dispersioonimaatriksiks ühikmaatriks, $D\vec{Z} = \mathbf{I}_p$. Keskvääruse omadusi arvestades saame nüüd

$$E\vec{X} = E(\vec{\mu} + \mathbf{A}\vec{Z}) = \vec{\mu} + \mathbf{A}E\vec{Z} = \vec{\mu}.$$

Dispersiooni omadusi kasutades saame

$$D\vec{X} = E[(\vec{X} - \vec{\mu})(\vec{X} - \vec{\mu})'] = E[(\mathbf{A}\vec{Z})(\mathbf{A}\vec{Z})'] = E(\mathbf{A}\vec{Z}\vec{Z}'\mathbf{A}') = \mathbf{\Sigma},$$

mott.

Teoreem 3.1. Kui juhuslik p -vektor on mitmemõõtmelise normaaljaotusega, $\vec{X} \sim N_p(\vec{\mu}, \mathbf{\Sigma})$, siis tema tihedus avaldub kujul

$$f_{\vec{X}}(\vec{x}) = \frac{1}{(\sqrt{2\pi})^p |\mathbf{\Sigma}|^{1/2}} \exp \left\{ -\frac{1}{2}(\vec{x} - \vec{\mu})' \mathbf{\Sigma}^{-1} (\vec{x} - \vec{\mu}) \right\}.$$

Tõestus. Teoreemi tõestamiseks kasutame ära teoreemi 1.1 juhusliku vektori funktsiooni jaotuse kohta. Olgu $\vec{Z} \sim N(\vec{0}, \mathbf{I}_p)$. Definitsiooni 3.1 kohaselt

$$\vec{X} = \mathbf{A}\vec{Z} + \vec{\mu},$$

kus maatriks $\mathbf{A}$ on pööratav. Seega on meil tegemist üksühese teisendusega ruumis $\mathbb{R}^p$:

$$\vec{x} = \mathbf{A}\vec{z} + \vec{\mu},$$

mis üksühesuse tõttu on pööratav:

$$\vec{z} = \mathbf{A}^{-1}(\vec{x} - \vec{\mu}).$$

Teoreemi 1.1 põhjal

$$f_{\vec{X}}(\vec{x}) = f_{\vec{Z}}(z_1(\vec{x}), z_2(\vec{x}), \dots, z_p(\vec{x})) J \left(\frac{z_1, z_2, \dots, z_p}{x_1, x_2, \dots, x_p} \right).$$

Asendades siia vektori $\vec{Z}$ tiheduse, saame

$$f_{\vec{X}}(\vec{x}) = \frac{1}{(\sqrt{2\pi})^p} \exp \left\{ -\frac{1}{2}(\vec{x} - \vec{\mu})'(\mathbf{A}^{-1})'\mathbf{A}^{-1}(\vec{x} - \vec{\mu}) \right\} J \left(\frac{\vec{z}}{\vec{x}} \right).$$

Leiame jakobiaani kasutades maatrikstuletise omadust ((iii)) Lisas 1 ja determinandi omadust (L1.8)

$$J \left(\frac{\vec{z}}{\vec{x}} \right) = \text{abs} \left| \frac{d\vec{z}}{d\vec{x}} \right| = \text{abs}|\mathbf{A}^{-1}| = \text{abs}|\mathbf{A}|^{-1}.$$

Kuna $\mathbf{A}\mathbf{A}' = \mathbf{\Sigma}$, saame

$$f_{\vec{X}}(\vec{x}) = \frac{1}{(\sqrt{2\pi})^p} \exp \left\{ -\frac{1}{2}(\vec{x} - \vec{\mu})'\mathbf{\Sigma}^{-1}(\vec{x} - \vec{\mu}) \right\} \frac{1}{\text{abs}|\mathbf{A}|}.$$

Arvestades vördust $|\mathbf{A}| = (|\mathbf{A}||\mathbf{A}'|)^{1/2}$, saame tiheduse jaoks avaldise

$$f_{\vec{X}} = \frac{1}{(\sqrt{2\pi})^p |\mathbf{\Sigma}|^{1/2}} \exp \left\{ -\frac{1}{2}(\vec{x} - \vec{\mu})'\mathbf{\Sigma}^{-1}(\vec{x} - \vec{\mu}) \right\},$$

mott.

Kõdunud normaaljaotusel pole tihedust, tema esitamisel kasutatakse karakteristikku funktsiooni.

Teoreem 3.2. Olgu $\vec{X} \sim N_p(\vec{\mu}, \mathbf{\Sigma})$, siis tema karakteristik funktsioon avaldub kujul

$$\varphi_{\vec{X}}(\vec{t}) = \exp(i\vec{t}'\vec{\mu} - \frac{1}{2}\vec{t}'\mathbf{\Sigma}\vec{t}).$$

Tõestus. Leiame esmalt mitmemõõtmelise standardse normaaljaotusega juhusliku vektori $\vec{Z}$ karakteristikku funktsiooni. Kuna $\vec{Z}$ komponendid on vastastikku sõltumatud ja standardse normaaljaotusega, siis tema karakteristik funktsioon avaldub jaotuse $N(0, 1)$ karakteristiklike funktsioonide korrutisena,

$$\varphi_{\vec{Z}}(\vec{t}) = \prod_{j=1}^p \varphi_{Z_j}(t_j) = \prod_{j=1}^p e^{-\frac{1}{2}t_j^2} = e^{-\frac{1}{2}\vec{t}'\vec{t}}.$$

Kasutades vektori $\vec{X}$ esitust

$$\vec{X} = \vec{\mu} + \mathbf{A}\vec{Z},$$

kus $\mathbf{\Sigma} = \mathbf{A}\mathbf{A}'$, saame

$$\varphi_{\vec{X}}(\vec{t}) = Ee^{i\vec{t}'\vec{X}} = Ee^{i\vec{t}'(\vec{\mu} + \mathbf{A}\vec{Z})} = e^{i\vec{t}'\vec{\mu}} Ee^{i\vec{t}'\mathbf{A}\vec{Z}} = e^{i\vec{t}'\vec{\mu}} Ee^{i(\mathbf{A}'\vec{t})'\vec{Z}}.$$

Kuna

$$e^{i\vec{t}'\vec{\mu}} Ee^{i(\mathbf{A}'\vec{t})'\vec{Z}} = e^{i\vec{t}'\vec{\mu}} \varphi_{\vec{Z}}(\mathbf{A}'\vec{t}),$$

siis saame

$$\varphi_{\vec{X}}(\vec{t}) = e^{i\vec{t}'\vec{\mu}} e^{-\frac{1}{2}\vec{t}'\mathbf{A}\mathbf{A}'\vec{t}} = \exp \left\{ i\vec{t}'\vec{\mu} - \frac{1}{2}\vec{t}'\mathbf{\Sigma}\vec{t} \right\},$$

mott.

Näeme, et karakteristikku funktsiooni väljakirjutamiseks on vaja teada normaaljaotuse parameetreid $\vec{\mu}$ ja $\mathbf{\Sigma}$. Kehtib ka vastupidine: kui me teame normaaljaotuse karakteristikku funktsiooni, siis saab sealt välja lugeda ka selle jaotuse parameetrite väärtused.

Veendume, et mitmemõõtmeliste normaaljaotuste klass on lineaarteisenduse suhtes kinnine, st teisendatud juhuslik vektor on ka normaaljaotusega. Muutuvad ainult selle juhusliku vektori jaotuse parameetrid.

Lause 3.2. Olgu $\vec{X} \sim N_p(\vec{\mu}, \Sigma)$, matriiks $\mathbf{B} : q \times p$ täisastakuga ($q \leq p$, $r(\mathbf{B}) = q$) ja kehtigu võrdus $\vec{Y} = \mathbf{B}\vec{X} + \vec{a}$, $\vec{a} \in \mathbb{R}^q$. Siis

$$\vec{Y} \sim N_q(\vec{a} + \mathbf{B}\vec{\mu}, \mathbf{B}\Sigma\mathbf{B}').$$

Tõestus. Kirjutame välja teisendatud vektori $\vec{Y}$ karakteristliku funktsiooni.

$$\begin{aligned}\varphi_{\vec{Y}}(\vec{t}) &= \varphi_{\vec{a} + \mathbf{B}\vec{X}}(\vec{t}) = E \exp(i\vec{t}'(\vec{a} + \mathbf{B}\vec{X})) \\ &= e^{i\vec{t}'\vec{a}} E \exp\{i(\mathbf{B}'\vec{t})'\vec{X}\} = e^{i\vec{t}'\vec{a}} \varphi_{\vec{X}}(\mathbf{B}'\vec{t}).\end{aligned}$$

Kuna vektori $\vec{X}$ karakteristliku funktsiooni avaldis on teada, saame

$$\begin{aligned}\varphi_{\vec{Y}}(\vec{t}) &= e^{i\vec{t}'\vec{a}} \exp\left\{i(\mathbf{B}'\vec{t})'\vec{\mu} - \frac{1}{2}(\mathbf{B}'\vec{t})'\Sigma(\mathbf{B}'\vec{t})\right\} \\ &= \exp\left\{i\vec{t}'(\vec{a} + \mathbf{B}\vec{\mu}) - \frac{1}{2}\vec{t}'\mathbf{B}\Sigma\mathbf{B}'\vec{t}\right\}.\end{aligned}$$

See on aga normaaljaotuse $N_q(\vec{a} + \mathbf{B}\vec{\mu}, \mathbf{B}\Sigma\mathbf{B}')$ karakteristlik funktsioon. Järelikult $\vec{Y} \sim N_q(\vec{a} + \mathbf{B}\vec{\mu}, \mathbf{B}\Sigma\mathbf{B}')$, mott.

MÄRKUS. Normaaljaotusega on tegemist juhul $r(\mathbf{B}) = q$, kui $r(\mathbf{B}) < q$, siis on tegemist kõdunud normaaljaotusega.

Võtame nüüd lähema vaatluse alla mitmemõõtmelise normaaljaotuse marginaal- ja tinglikud jaotused. Nende kirjeldamiseks vaatame ühte konkreetset alamvektorite paari, mille saame jaotades vektori $\vec{X}$ kaheks alamvektoriks

$$\vec{X} = \begin{bmatrix} \vec{X}_1 \\ \vec{X}_2 \end{bmatrix},$$

kus $\vec{X}_1$ on esimesest k komponendist moodustatud k -vektor ja $\vec{X}_2$ ülejäänud komponentidest moodustatud $(p - k)$ -vektor. Samal viisil jaotub alamvektoriteks ka keskväärusvektor

$$E\vec{X} = \vec{\mu} = \begin{bmatrix} \vec{\mu}_1 \\ \vec{\mu}_2 \end{bmatrix}.$$

Vektori $\vec{X}$ dispersioonimatriks jaotub neljast plokist koosnevaks plokkmatriksiks,

$$D\vec{X} = \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}, \quad (3.1)$$

kus Σ_{11} on vektori $\vec{X}_1$ dispersioonimatriks, Σ_{22} on vektori $\vec{X}_2$ dispersioonimatriks ja Σ_{12} on vektorite $\vec{X}_1$ ja $\vec{X}_2$ vaheline kovariatsioonimatriks. Dispersioonimatriksi Σ sümmeetrilisuse tõttu $\Sigma'_{12} = \Sigma_{21}$.

Näitame lauset 3.2 kasutades, et mõlema alamvektori marginaaljaotuseks on normaaljaotus. Moodustame matriksi $\mathbf{B}_1 : k \times p$,

$$\mathbf{B}_1 = [\mathbf{I}_k : \mathbf{0}_{k \times (p-k)}].$$

Siis

$$\vec{X}_1 = \mathbf{B}_1\vec{X},$$

millest järeldubki, et $\vec{X}_1 \sim N(\vec{\mu}_1, \Sigma_{11})$. Vektori $\vec{X}_2$ marginaaljaotuse leidmiseks moodustame matriksi $\mathbf{B}_2 = [\mathbf{0}_{(p-k) \times k} : \mathbf{I}_{p-k}]$, mille korral saame

$$\vec{X}_2 = \mathbf{B}_2\vec{X},$$

millest järeldubki, et $\vec{X}_2 \sim N(\vec{\mu}_2, \Sigma_{22})$. Kuna vektori $\vec{X}$ iga alamvektori saame eraldada sobiva konstruktsiooniga lineaarteisenduse abil, järeldub siit, et mitmemõõtmelise normaaljaotuse korral on iga alamvektori marginaaljaotuseks normaaljaotus.

Kui alamvektorid on sõltumatud, $\vec{X}_1 \perp \vec{X}_2$, on nende kõikide komponentide vahelised kovariatsioonid nullid, st $\Sigma_{12} = \mathbf{0}$, sõltumatusest järeldub mittekorreleeritus. Normaaljaotusel on aga üks unikaalne omadus – kehtib ka vastupidine väide: mittekorreleeritusest järeldub sõltumatus. Tõestame selle.

Teoreem 3.3. *Olgu $\vec{X} \sim N_p(\vec{\mu}, \Sigma)$, kusjuures vektor $\vec{X}$ on jaotatud k - ja $(p - k)$ -mõõtmeteliseks alamvektoriteks $\vec{X}_1$ ja $\vec{X}_2$. Seejuures jaotub dispersioonimaatriks plokkideks*

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

Kui plokk $\Sigma_{12} = \mathbf{0}$, siis need alamvektorid on sõltumatud.

Tõestus. Kirjutame välja vektori $\vec{X}$ tiheduse

$$f_{\vec{X}}(\vec{x}) = \frac{1}{(\sqrt{2\pi})^p |\Sigma|^{1/2}} \exp \left\{ -\frac{1}{2} (\vec{x} - \vec{\mu})' \Sigma^{-1} (\vec{x} - \vec{\mu}) \right\}.$$

Kuna Σ on plokkmaatriks, milles nullist erinevad plokid asuvad ainult peadiagonaalil, kehtivad võrdsused

$$|\Sigma| = |\Sigma_{11}| \cdot |\Sigma_{22}|,$$

$$\Sigma^{-1} = \begin{bmatrix} \Sigma_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \Sigma_{22}^{-1} \end{bmatrix}.$$

Vektori $\vec{X}$ tihedus omandab seetõttu kuju

$$\begin{aligned} f_{\vec{X}}(\vec{x}) &= \frac{1}{(\sqrt{2\pi})^p |\Sigma|^{1/2}} \times \exp \left\{ -\frac{1}{2} (\vec{x}_1 - \vec{\mu}_1)' \Sigma_{11}^{-1} (\vec{x}_1 - \vec{\mu}_1) - \frac{1}{2} (\vec{x}_2 - \vec{\mu}_2)' \Sigma_{22}^{-1} (\vec{x}_2 - \vec{\mu}_2) \right\} \\ &= f_{\vec{X}_1}(\vec{x}_1) f_{\vec{X}_2}(\vec{x}_2). \end{aligned}$$

Niisugusest esitusest aga järeldubki alamvektorite sõltumatus, mott.

Leiame nüüd alamvektori $\vec{X}_1$ tingliku jaotuse tingimusel, et teise alamvektori $\vec{X}_2$ väärtus on fikseeritud. Tingliku jaotuse leidmiseks kasutame lauset 3.2. Lisaks võtame kasutusele spetsiaalsed tähistused plokkmaatriksi (3.1) peadiagonaalil asuvate plokkide Σ_{11} ja Σ_{22} nn. *Schuri täiendite* jaoks. Tähistame maatriksi Σ_{22} Schuri täiendi sümboliga $\Sigma_{11 \cdot 2}$,

$$\Sigma_{11 \cdot 2} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$$

ja maatriksi Σ_{11} Schuri täiendi sümboliga $\Sigma_{22 \cdot 1}$,

$$\Sigma_{22 \cdot 1} = \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}.$$

Teoreem 3.4. *Kui vektor $\vec{X} \sim N_p(\vec{\mu}, \Sigma)$ on jaotatud k - ja $(p - k)$ -mõõtmeteliseks alamvektoriteks*

$$\vec{X} = \begin{bmatrix} \vec{X}_1 \\ \vec{X}_2 \end{bmatrix},$$

millega keskäärtusteks on keskäärtusvektori $\vec{\mu}$ alamvektorid

$$\vec{\mu} = \begin{bmatrix} \vec{\mu}_1 \\ \vec{\mu}_2 \end{bmatrix}$$

ja dispersioonimaatriksiteks maatriksi Σ peadiagonaali plokid,

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix},$$

siis vektori $\vec{X}_1$ tinglikuks jaotuseks tingimusel $\vec{X}_2 = \vec{x}_2$, on normaaljaotus $N_k(\vec{\mu}_1 - \Sigma_{12} \Sigma_{22}^{-1} (\vec{\mu}_2 - \vec{x}_2), \Sigma_{11 \cdot 2})$.

Tõestus. Leiame vektori $\vec{X}_1$ tingliku jaotuse tingimisel, et alamvektori $\vec{X}_2$ väärtus on fikseeritud, $\vec{X}_2 = \vec{x}_2$. Paneme tähele, et suvalise maatriksi $\mathbf{B} : k \times (p - k)$ korral

$$\begin{aligned} \mathbf{C} &= \begin{bmatrix} \mathbf{I}_k & -\mathbf{B} \\ \mathbf{0} & \mathbf{I}_{p-k} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I}_k & \mathbf{0} \\ -\mathbf{B}' & \mathbf{I}_{p-k} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{\Sigma}_{11} - \mathbf{B}\boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{12} - \mathbf{B}\boldsymbol{\Sigma}_{22} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I}_k & \mathbf{0} \\ -\mathbf{B}' & \mathbf{I}_{p-k} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{\Sigma}_{11} - \mathbf{B}\boldsymbol{\Sigma}_{21} - \boldsymbol{\Sigma}_{12}\mathbf{B}' + \mathbf{B}\boldsymbol{\Sigma}_{22}\mathbf{B}' & \boldsymbol{\Sigma}_{12} - \mathbf{B}\boldsymbol{\Sigma}_{22} \\ \boldsymbol{\Sigma}_{21} - \boldsymbol{\Sigma}_{22}\mathbf{B}' & \boldsymbol{\Sigma}_{22} \end{bmatrix}. \end{aligned}$$

Võttes nüüd $\mathbf{B} = \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}$, saame korrutismaatriksile $\mathbf{C}$ anda kuju

$$\mathbf{C} = \begin{bmatrix} \boldsymbol{\Sigma}_{11 \cdot 2} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22} \end{bmatrix}.$$

Moodustame uue juhusliku vektori $\vec{Y}$

$$\vec{Y} = \begin{bmatrix} \mathbf{I}_k & -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ \mathbf{0} & \mathbf{I}_{p-k} \end{bmatrix} \vec{X} = \begin{bmatrix} \vec{X}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\vec{X}_2 \\ \vec{X}_2 \end{bmatrix}.$$

See juhuslik vektor on normaaljaotusega, tema dispersioonimaatriksiks on maatriks $\mathbf{C}$, mille väljaspool peadiagonaali asuvad plokid on nullid. Järelikult on alamvektorid $\vec{Y}_1 = \vec{X}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\vec{X}_2$ ja $\vec{Y}_2 = \vec{X}_2$ sõltumatud.

Leiame vektori $\vec{Y}$ keskäärtuse

$$E\vec{Y} = E \begin{bmatrix} \vec{X}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\vec{X}_2 \\ \vec{X}_2 \end{bmatrix} = \begin{bmatrix} \vec{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\vec{\mu}_2 \\ \vec{\mu}_2 \end{bmatrix}.$$

Siit $\vec{Y}_1 \sim N_k(\vec{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\vec{\mu}_2, \boldsymbol{\Sigma}_{11 \cdot 2})$. Sõltumatuse tõttu on $\vec{Y}_1$ tinglik jaotus tingimisel $\vec{X}_2 = \vec{x}_2$ täpselt samasugune. Kuna selle tingimuse korral

$$\vec{Y}_1 = \vec{X}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\vec{x}_2,$$

saame juhusliku vektori $\vec{X}_1$ jaoks avaldise

$$\vec{X}_1 = \vec{Y}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\vec{x}_2.$$

Siis $\vec{X}_1$ tinglik jaotus tingimisel $\vec{X}_2 = \vec{x}_2$ on normaaljaotus parameetritega $\vec{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\vec{\mu}_2 - \vec{x}_2)$ ja $\boldsymbol{\Sigma}_{11 \cdot 2}$, sest $E\vec{X}_1 = E\vec{Y}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\vec{x}_2$ ja $D\vec{X}_1 = D\vec{Y}_1$, mott.

Kokkuvõttes märgime, et mitmemõõtmelise normaaljaotuse korral on meil tegemist heade omadustega andmemudeliga: normaaljaotusega juhusliku vektori komponentide mittekorreleeritusest järeldub nende sõltumatus ning normaaljaotusega on ka kõik marginaaljaotused ja tinglikud jaotused.

Ülesanded 3

3.1. Olgu $X_i \sim N(0, \sigma_i^2)$, $i = 1, \dots, p$ ja $X_i \perp X_j$, $i \neq j$. Leida $\vec{X} = (X_1, \dots, X_p)'$ dispersioonimaatriks $D\vec{X}$ ja tõenäosustihedus.

3.2. Olgu $\vec{X} = N_3(\vec{\mu}, \boldsymbol{\Sigma})$, kus $\vec{\mu}' = (0, 0, 1)$ ja

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 2 \end{pmatrix}.$$

Leida $(X_1, X_2)'$ marginaaljaotus ja tinglik jaotus $(X_2, X_3)|X_1$.

3.3. Olgu $\vec{X} = N_3(\vec{\mu}, \Sigma)$, kus $\vec{\mu}' = (-3, 1, 2)$ ja

$$\Sigma = \begin{pmatrix} 4 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & 2 \end{pmatrix}.$$

Millised järgnevad juhuslikest suurustest moodustatud paarid on sõltuvad:

- X_1 ja X_3 ;
- X_1 ja X_2 ;
- X_1 ja $X_1 + 3X_2 - 2X_3$;
- (X_1, X_3) ja X_2 ;
- $X_1 + 4X_3$ ja $X_1 - 2X_2$;
- $X_1 + X_2 + X_3$ ja $4X_1 - 2X_2 + 3X_3$?

Leida tinglikud jaotused $(X_1, X_2)|X_3$ ja $X_2|(X_1, X_3)$.

3.4. Olgu normaaljaotuse $\vec{X} \sim N_2(\vec{\mu}, \Sigma)$ tõenäosustihedus antud järgmise võrdusega:

$$f_{\vec{X}}(x_1, x_2) = \frac{1}{4\pi\sqrt{11}} \exp\left(-\frac{3}{44}x_1^2 + \frac{2}{11}x_1 + \frac{1}{22}x_1x_2 - \frac{1}{11}x_2^2 + \frac{17}{22}x_2 - 2\frac{9}{44}\right).$$

Leida parameetrid $\vec{\mu}$ ja Σ .

4 Ruutvormi jaotus

Normaaljaotusega üldkogumi kohta järelduste tegemisel on tihti vajalikud juhuslikust vektorist moodustatud ruutvormid. Ruutvormina tuntakse matemaatikas konstruktsiooni $\vec{x}'\mathbf{A}\vec{x}$, kus $\mathbf{A}$ on sümmeetriline maatriks.

Teoreem 4.1. Olgu $\vec{X} \sim N_p(\vec{0}, \Sigma)$. Siis juhuslik suurus $\vec{X}'\Sigma^{-1}\vec{X} \sim \chi^2(p)$.

Tõestus. Normaaljaotuse definitsiooni tõttu leidub $p \times p$ maatriks $\mathbf{A}$ nii, et $\Sigma = \mathbf{A}\mathbf{A}'$, kusjuures maatriks $\mathbf{A}$ on pööratav. See aga tähendab, et ruutvormi saame esitada kujul

$$\vec{X}'(\mathbf{A}')^{-1}\mathbf{A}^{-1}\vec{X}.$$

Kerge on näha, et $E(\mathbf{A}^{-1}\vec{X}) = 0$ ja

$$D(\mathbf{A}^{-1}\vec{X}) = \mathbf{A}^{-1}D(\vec{X})(\mathbf{A}^{-1})' = \mathbf{A}^{-1}\mathbf{A}\mathbf{A}'(\mathbf{A}')^{-1} = \mathbf{I}_p.$$

Seega vektori $\vec{Y} = \mathbf{A}^{-1}\vec{X}$ komponendid on sõltumatud standardse normaaljaotusega juhuslikud suurused ja χ^2 -jaotuse definitsiooni kohaselt

$$\vec{Y}'\vec{Y} = \sum_{i=1}^p Y_i^2 \sim \chi^2(p),$$

mott.

Teoreemist saame vahetult järelduse ka üldise normaaljaotuse jaoks.

Järeldus 4.1.1. Olgu $\vec{X} \sim N_p(\vec{\mu}, \Sigma)$. Siis juhuslik suurus

$$(\vec{X} - \vec{\mu})'\Sigma^{-1}(\vec{X} - \vec{\mu}) \sim \chi^2(p).$$

Seda tulemust on võimalik üldistada. Teoreemis ja selle järelduses on maatriks Σ^{-1} ruutvormi moodustavaks maatriksiks, kusjuures see maatriks on täisastakuga. Vaatame, milline jaotus tekib, kui ruutvormi moodustava $p \times p$ maatriksi astak on väiksem kui selle maatriksi mõõde p .

Teoreem 4.2. Olgu $\vec{X} \sim N(\vec{0}, \mathbf{I}_p)$ ja $\mathbf{G}$ sümmeetriline $p \times p$ maatriks astakuga r ($r < p$). Siis

$$\vec{X}'\mathbf{G}\vec{X} \sim \chi^2(r)$$

parajasti siis, kui $\mathbf{G}$ on idempotentne.

Tõestus. Piisavus. Olgu $\mathbf{G} = \mathbf{G}^2$. Kuna $\mathbf{G}$ on sümmeetriline, siis ta on esitatav omaväärtuste ja omavektorite kaudu:

$$\mathbf{G} = \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}',$$

kus $\mathbf{\Gamma}$ on ortogonaalne maatriks, mille veergudeks on maatriksi $\mathbf{G}$ omavektorid,

$$\mathbf{\Gamma} = [\vec{\gamma}_1 \ \vec{\gamma}_2 \ \dots \ \vec{\gamma}_p], \quad \vec{\gamma}_i' \vec{\gamma}_i = 1, \quad \vec{\gamma}_i' \vec{\gamma}_j = 0, \quad i \neq j$$

ja $\mathbf{\Lambda}$ on omaväärtuste diagonaalmaatriks:

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \lambda_p \end{pmatrix}.$$

Idempotentse maatriksi omaväärtused on kas nullid või ühed. Tõepoolest:

$$\mathbf{G} = \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}' = \mathbf{G}\mathbf{G} = \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}'\mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}' = \mathbf{\Gamma}\mathbf{\Lambda}^2\mathbf{\Gamma}',$$

millest järeldub, et $\lambda_i^2 = \lambda_i$. See on aga võimalik vaid siis, kui $\lambda_i = 1$ või $\lambda_i = 0$. Kuna maatriksi $\mathbf{G}$ astak on r , saavad nullist erineda ainult r omaväärtust. Tähistame $\vec{Y} = \mathbf{\Gamma}'\vec{X}$. Kuna

$$D\vec{Y} = D(\mathbf{\Gamma}'\vec{X}) = \mathbf{\Gamma}'D\vec{X}\mathbf{\Gamma} = \mathbf{\Gamma}'\mathbf{\Gamma} = \mathbf{I}_p,$$

siis ka $\vec{Y} \sim N_p(\vec{0}, \mathbf{I}_p)$. Seega

$$\vec{X}'\mathbf{G}\vec{X} = \vec{X}'\mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}'\vec{X} = \vec{Y}'\mathbf{\Lambda}\vec{Y} = \sum_{i=1}^p \lambda_i Y_i^2 \sim \chi^2(r).$$

Tarvilikkus. Olgu $\vec{X}'\mathbf{G}\vec{X} \sim \chi^2(r)$. Näitame, et sel juhul $\mathbf{G} = \mathbf{G}^2$.

Tähistame lühidalt $Z = \vec{X}'\mathbf{G}\vec{X}$. Kasutame tõestamisel χ_r^2 -jaotuse karakteristlikku funktsiooni,

$$\varphi_Z(t) = E(e^{itZ}) = (1 - 2it)^{-\frac{r}{2}}.$$

Kuna $\mathbf{G}$ on sümmeetriline maatriks, siis saame ta esitada omaväärtuste ja omavektorite maatriksite kaudu:

$$\mathbf{G} = \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}'.$$

Kasutades nüüd tähistust $\vec{Y} = \mathbf{\Gamma}'\vec{X}$, ($\vec{Y} \sim N_p(\vec{0}, \mathbf{I}_p)$), saame

$$Z = \vec{X}'\mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}'\vec{X} = \vec{Y}'\mathbf{\Lambda}\vec{Y} = \sum_{i=1}^p \lambda_i Y_i^2.$$

Vektori $\vec{Y}$ komponendid on sõltumatud. Seda kasutades leiame juhusliku suuruse Z karakteristliku funktsiooni,

$$E(e^{itZ}) = E(e^{it\vec{Y}'\mathbf{\Lambda}\vec{Y}}) = E\prod_{j=1}^p e^{it\lambda_j Y_j^2} = \prod_{j=1}^p (1 - 2it\lambda_j)^{-\frac{1}{2}}.$$

Viimase võrduse saamisel kasutasime asjaolu, et Y_j , $j = 1, \dots, p$, on sõltumatud juhuslikud suurused ja $Y_j^2 \sim \chi_1^2$. Eelduse kohaselt peab kehtima iga t korral võrdus

$$(1 - 2it)^{-\frac{r}{2}} = \prod_{j=1}^p (1 - 2it\lambda_j)^{-\frac{1}{2}},$$

mis on võimalik vaid siis, kui maatriksil $\mathbf{G}$ on r omaväärtust ühed ja ülejäänud omaväärtused nullid. See aga tähendab, et maatriks $\mathbf{G}$ on idempotentne, $\mathbf{G} = \mathbf{G}^2$, mott.

Ülesanded 4

4.1. Näidata, et juhusliku suuruse $X \sim \chi_1^2$ jaotusfunktsioon on kujul

$$F_X(x) = F_{N(0,1)}(\sqrt{x}) - F_{N(0,1)}(-\sqrt{x})$$

ja tõenäosustihedus

$$f_X(x) = \frac{1}{\sqrt{2\pi x}} e^{-\frac{x}{2}}, \quad x > 0.$$

4.2. Näidata, et juhusliku suuruse $X \sim \chi_1^2$ karakteristik funktsioon on kujul

$$\varphi_X(t) = (1 - 2it)^{-\frac{1}{2}}.$$

Teame, et $X = Y^2$, kus $Y \sim N(0, 1)$.

5 Normaaljaotuse parameetrite hindamine

A. Momentide meetod. Momentide meetod tugineb väga lihtsale ideele – avaldada jaotuse parameetrid teoreetilise jaotuse momentide kaudu ja seejärel asendada valemite esinevad tundmatud teoreetilised momendid valimi põhjal arvutatavate momentide hinnangutega – valimi momentidega. Mitmemõõtmelise normaaljaotuse parameetriteks ongi parajasti teoreetilised momendid – jaotuse keskvaartusvektor $\vec{\mu}$ ja teine tsentraalne moment – dispersioonimaatriks Σ . Seega saame hinnangud:

$$\hat{\vec{\mu}} = \vec{\bar{x}},$$

$$\vec{\bar{x}} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \dots \\ \bar{x}_p \end{pmatrix}.$$

ja

$$\hat{\Sigma} = \hat{\mathbf{S}}_*,$$

kus

$$\hat{\mathbf{S}}_* = \frac{1}{n} \sum_{i=1}^n (\vec{x}_i - \vec{\bar{x}})(\vec{x}_i - \vec{\bar{x}})'$$

Momentide meetodil määratud hinnangud on üldreeglina mõjusad, st suurte valimite korral on nad suure tõenäosusega tõelistele parameetri väärtustele lähedased. Väikeste valimite korral ei ole momentide meetodil määratud hinnangute head omadused garanteeritud.

MÄRKUS. Kuna teame, et $\hat{\mathbf{S}}_*$ on nihkega hinnang, siis on loomulikult mõistlikum kasutada Σ hinnanguna nihketa hinnangut $\hat{\mathbf{S}}$, kus kordaja $1/n$ on asendatud kordajaga $1/(n-1)$.

B. Suurima tõepära meetod. Suurima tõepära meetod lähtub tõepärafunktsioonist $L(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n; \vec{\mu}, \Sigma)$. Parameetrite suurima tõepära hinnanguteks on väärtused, mille korral tõepärafunktsioon on antud valimi korral maksimaalne. Mitmemõõtmelise normaaljaotuse tõepärafunktsioon on teoreetilise valimi $p \times n$ juhusliku maatriksi $\mathbf{X}$ – tõenäosustiheduse väärtus konkreetse valimi $\mathcal{X}$ korral. Teoreetilise valimi kohta tehtud eelduste põhjal on selle maatriksi veerud $\vec{x}_i$ sõltumatute p -mõõtmelise normaaljaotusega vektorite $\vec{X}_i \sim N(\vec{\mu}, \Sigma)$ realisatsioonid. Veeurvektorite sõltumatuse tõttu on lihtne välja kirjutada teoreetilise valimi tihedust — see on veeurvektorite tiheduste korrutis,

$$\begin{aligned} L(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n; \vec{\mu}, \Sigma) &= f(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n; \vec{\mu}, \Sigma) = \prod_{i=1}^n f_{\vec{X}_i}(\vec{x}_i; \vec{\mu}, \Sigma) \\ &= \prod_{i=1}^n \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (\vec{x}_i - \vec{\mu})' \Sigma^{-1} (\vec{x}_i - \vec{\mu}) \right\}. \end{aligned}$$

Siit saame võrduse

$$\begin{aligned} L(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n; \vec{\mu}, \Sigma) \\ = \frac{1}{(2\pi)^{\frac{np}{2}} |\Sigma|^{\frac{n}{2}}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n (\vec{x}_i - \vec{\mu})' \Sigma^{-1} (\vec{x}_i - \vec{\mu}) \right\}. \end{aligned} \quad (5.1)$$

Teisendame eksponentfunktsiooni astendajas olevat avaldist, kasutades maatriksi jälje omadust

$$\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA}).$$

Eksponentfunktsiooni astendajaks oleva summa saame nüüd esitada maatriksite jälgede summana. Tõepoolest, kuna see summa on ühemõõtmeline suurus, võime kirjutada

$$\sum_{i=1}^n \text{tr}[(\vec{x}_i - \vec{\mu})' \Sigma^{-1} (\vec{x}_i - \vec{\mu})] = \sum_{i=1}^n \text{tr}[\Sigma^{-1} (\vec{x}_i - \vec{\mu}) (\vec{x}_i - \vec{\mu})']$$

ja kuna maatriksite summa jälg võrdub maatriksite jälgede summaga,

$$\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}\mathbf{A} + \text{tr}\mathbf{B},$$

saame

$$\sum_{i=1}^n \text{tr}[\Sigma^{-1} (\vec{x}_i - \vec{\mu}) (\vec{x}_i - \vec{\mu})'] = \text{tr}[\Sigma^{-1} \sum_{i=1}^n (\vec{x}_i - \vec{\mu}) (\vec{x}_i - \vec{\mu})'].$$

Suurima tõepära hinnangute leidmiseks muudame viimasena saadud summas liidetavad maatriksid esmalt keerulisemaks. Kasutame samasust

$$\vec{x}_i - \vec{\mu} = \vec{x}_i - \vec{\bar{x}} + \vec{\bar{x}} - \vec{\mu}.$$

Arvestades, et

$$\sum_{i=1}^n (\vec{x}_i - \vec{\bar{x}}) = \vec{0},$$

saame

$$\begin{aligned} &\text{tr}[\Sigma^{-1} \sum_{i=1}^n (\vec{x}_i - \vec{\mu}) (\vec{x}_i - \vec{\mu})'] \\ &= \text{tr}\{\Sigma^{-1} [\sum_{i=1}^n (\vec{x}_i - \vec{\bar{x}}) (\vec{x}_i - \vec{\bar{x}})' + n(\vec{\bar{x}} - \vec{\mu}) (\vec{\bar{x}} - \vec{\mu})']\}. \end{aligned}$$

Kuna summa avaldis on parajasti see, mida kasutasime hinnangu $\hat{\mathbf{S}}_*$ määramisel, saame

$$\sum_{i=1}^n (\vec{x}_i - \vec{\mu})' \Sigma^{-1} (\vec{x}_i - \vec{\mu}) = n \text{tr}\{\Sigma^{-1} [\hat{\mathbf{S}}_* + (\vec{\bar{x}} - \vec{\mu}) (\vec{\bar{x}} - \vec{\mu})']\}.$$

Tõepärafunktsioon omandab kuju

$$L(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n; \vec{\mu}, \Sigma) = \frac{1}{(2\pi)^{\frac{np}{2}} |\Sigma|^{\frac{n}{2}}} \exp \left\{ -\frac{n}{2} \text{tr} \{ \Sigma^{-1} [\hat{\mathbf{S}}_* + (\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})'] \} \right\}. \quad (5.2)$$

Leiame nüüd parameetrite väärtused, mille korral tõepärafunktsioon saavutab maksimumi. Tarviliku tingimuse kohaselt peavad funktsiooni osatuletised maksimumpunktis võrduma nulliga. Enne tuletiste leidmist läheme üle logaritmilisele tõepärafunktsioonile:

$$l(\vec{x}_1, \dots, \vec{x}_n; \vec{\mu}, \Sigma) = -\frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln |\Sigma| - \frac{n}{2} \text{tr} \{ \Sigma^{-1} [\hat{\mathbf{S}}_* + (\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})'] \}. \quad (5.3)$$

Esitades osatuletised maatrikstuletise abil saame võrrandisüsteemi

$$\begin{cases} \frac{dl(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n; \vec{\mu}, \Sigma)}{d\vec{\mu}} = \mathbf{0} \\ \frac{dl(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n; \vec{\mu}, \Sigma)}{d\Sigma} = \mathbf{0}. \end{cases}$$

Kasutades maatrikstuletise omadusi lisast 1 on need tuletised lihtsalt leitavad. Omadust (xi) rakendades saame ruutvormi diferentseerimiseeskirja

$$\frac{d(\vec{x}' \mathbf{A} \vec{x})}{d\vec{x}} = 2\vec{x}' \mathbf{A},$$

kui $\mathbf{A}$ on sümmeetriline maatriks. Kuna logaritmilise tõepärafunktsiooni avaldises (5.3) vaid üks liidetav sõltub parameetrist $\vec{\mu}$, siis saame

$$\begin{aligned} \frac{dl(\vec{x}_1, \dots, \vec{x}_n; \vec{\mu}, \Sigma)}{d\vec{\mu}} &= -\frac{n}{2} \frac{d}{d\vec{\mu}} \text{tr} [\Sigma^{-1} (\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})'] \\ &= -\frac{n}{2} \frac{d(\vec{x} - \vec{\mu})' \Sigma^{-1} (\vec{x} - \vec{\mu})}{d\vec{\mu}} = n(\vec{x} - \vec{\mu})' \Sigma^{-1}. \end{aligned}$$

Esimesele võrrandile saame kuju

$$\Sigma^{-1} (\vec{x} - \vec{\mu}) = \vec{0},$$

mille lahendiks on $\hat{\vec{\mu}} = \vec{x}$.

Teises võrrandis asendame $\vec{\mu}$ tema jaoks leitud hinnanguga $\vec{x}$. Saame

$$\frac{dl(\vec{x}_1, \dots, \vec{x}_n; \vec{\mu} = \vec{x}, \Sigma)}{d\Sigma} = -\frac{n}{2} \frac{d \ln |\Sigma|}{d\Sigma} - \frac{n}{2} \frac{d \text{tr}(\Sigma^{-1} \hat{\mathbf{S}}_*)}{d\Sigma}.$$

Leiame esimese liidetava saadud avaldises kasutades tuletise omadusi (v) ja (xvi) lisast 1:

$$-\frac{n}{2} \frac{d \ln |\Sigma|}{d\Sigma} = \frac{d \ln |\Sigma|}{d|\Sigma|} \cdot \frac{d|\Sigma|}{d\Sigma} = -\frac{n}{2} \frac{1}{|\Sigma|} |\Sigma| \text{vec}' \Sigma^{-1} = -\frac{n}{2} \text{vec}' \Sigma^{-1}.$$

Teise liidetava korral kasutame tuletise omadusi (xvii), (vii) ja (v):

$$\begin{aligned} -\frac{n}{2} \frac{d \text{tr}(\Sigma^{-1} \hat{\mathbf{S}}_*)}{d\Sigma} &= -\frac{n}{2} \frac{d \text{tr}(\Sigma^{-1} \hat{\mathbf{S}}_*)}{d(\Sigma^{-1} \hat{\mathbf{S}}_*)} \cdot \frac{d(\Sigma^{-1} \hat{\mathbf{S}}_*)}{d\Sigma} \\ &= -\frac{n}{2} \text{vec}' \mathbf{I}_p (\hat{\mathbf{S}}_* \otimes \mathbf{I}_p) \frac{d\Sigma^{-1}}{d\Sigma} = \frac{n}{2} \text{vec}' \hat{\mathbf{S}}_* (\Sigma^{-1} \otimes \Sigma^{-1}). \end{aligned}$$

Tuletise avaldis on seega kujul

$$\frac{dl(\vec{x}_1, \dots, \vec{x}_n; \vec{\mu} = \vec{x}, \Sigma)}{d\Sigma} = -\frac{n}{2} (\text{vec}' \Sigma^{-1} - \text{vec}' \hat{\mathbf{S}}_* (\Sigma^{-1} \otimes \Sigma^{-1})),$$

kust pärast tuletise võrdsustamist nulliga saame

$$(\text{vec}' \hat{\mathbf{S}}_*)' (\Sigma^{-1} \otimes \Sigma^{-1}) = (\text{vec}' \Sigma^{-1})'.$$

Korrutades nüüd paremalt võrduse mõlemat poolt otsekorrutisega $\Sigma \otimes \Sigma$ ja arvestades võrdust $\text{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\text{vec}\mathbf{B}$, saame

$$(\text{vec}\hat{\mathbf{S}}_*)' = (\text{vec}\Sigma)'$$

mis aga tähendab, et $\hat{\Sigma} = \hat{\mathbf{S}}_*$. Saime sama tulemuse, mis momentide meetodiga. Sõnastame saadud tulemused järgmise teoreemina.

Teoreem 5.1. *Olgu $\vec{X}$ p -mõõtmelise normaaljaotusega juhuslik vektor, $\vec{X} \sim N_p(\vec{\mu}, \Sigma)$. Siis valimi $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n$ põhjal leitud suurima tõepära hinnangud parameetritele $\vec{\mu}$ ja Σ on järgmised*

$$\hat{\vec{\mu}} = \vec{x} = \frac{1}{n} \sum_{i=1}^n \vec{x}_i,$$

$$\hat{\Sigma} = \hat{\mathbf{S}}_* = \frac{1}{n} \sum_{i=1}^n (\vec{x}_i - \vec{x})(\vec{x}_i - \vec{x})'$$

Vaatame veelkord kuju (5.2), mille saime anda tõepärafunktsioonile. Tuletame meelde tõepärafunktsiooni tähenduse: kui vaatame parameetreid fikseerituna ja loeme muutuvateks argumentideks valimi elemendid, on sama funktsioon teoreetilise valimi tihedusfunktsiooniks. Seega saame teoreetilise valimi tihedusfunktsiooni esitada hinnangute $\vec{x}$ ja $\hat{\mathbf{S}}_*$ kaudu,

$$f_{\mathbf{X}}(\mathcal{X}; \vec{\mu}, \Sigma) = \frac{1}{(2\pi)^{\frac{np}{2}} |\Sigma|^{\frac{n}{2}}} \exp \left\{ -\frac{n}{2} \text{tr} \{ \Sigma^{-1} [\hat{\mathbf{S}}_* + (\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})'] \} \right\}.$$

Sellest esitusest aga järeldub, et vastavad statistikud on piisavad statistikud parameetrite $\vec{\mu}$ ja Σ hindamiseks. Kuna siin tihedusfunktsioon jaguneb kahe teguri korrutiseks,

$$f_{\mathbf{X}}(\mathcal{X}; \vec{\mu}, \Sigma) = \frac{1}{(2\pi)^{\frac{np}{2}} |\Sigma|^{\frac{n}{2}}} \exp \left\{ -\frac{n}{2} \text{tr}(\Sigma^{-1} \hat{\mathbf{S}}_*) \right\} \\ \times \exp \left\{ -\frac{n}{2} \text{tr}[\Sigma^{-1} (\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})'] \right\},$$

tuleneb siit ka nende statistikute sõltumatus, kui vaid õnnestub näidata, et tegurid on statistikute marginaaltihedusfunktsioonid.

6 Wisharti jaotus

A. Statistikud. Normaaljaotuse parameetrite hinnangute omaduste uurimiseks võtame vaatluse alla neid määravad statistikud. Vaatame esmalt jaotuse keskväärtust. Teda hindava statistiku $\vec{X}$ saame, kui hinnangut määravas valemis konkreetse valimi elemendid asendame teoreetilise valimi elementidega,

$$\vec{X} = \frac{1}{n} \sum_{i=1}^n \vec{X}_i.$$

Vastavalt teoreetilise valimi omadustele on liidetavad vektorid sõltumatud ja normaaljaotusega. Kuna sõltumatute normaaljaotusega juhuslike vektorite summa on normaaljaotusega, on ka statistik $\vec{X}$ normaaljaotusega. Leiame tema keskväärtuse,

$$E\vec{X} = \frac{1}{n} \sum_{i=1}^n E\vec{X}_i = \vec{\mu}$$

ja dispersioonimaatriksi,

$$D\vec{X} = \frac{1}{n^2} \sum_{i=1}^n D\vec{X}_i = \frac{1}{n} \Sigma.$$

Seega keskvärtuse hinnang on nihketa ja mõjus, tema jaotuseks on mitmemõõtmeline normaaljaotus, $\vec{X} \sim N_p(\vec{\mu}, \frac{1}{n}\Sigma)$.

Normaaljaotuse dispersioonimaatriksile määrab suurima tõepära hinnangu statistik $\mathbf{S}_*$,

$$\mathbf{S}_* = \frac{1}{n} \sum_{i=1}^n (\vec{X}_i - \vec{X})(\vec{X}_i - \vec{X})'. \quad (6.1)$$

Selle statistiku jaotuse kirjeldamiseks peame tutvuma ühe uue jaotusega.

B. Wisharti jaotus. Defineerime Wisharti jaotuse teatava konstruktsiooni kaudu.

Definitsioon 6.1. Olgu $\mathbf{X}$ juhuslik $p \times n$ -maatriks, mille veerud $\vec{X}_i$ on sõltumatud ja sama tsentreeritud normaaljaotusega, $\vec{X}_i \sim N_p(\vec{0}, \Sigma)$. Siis juhusliku maatriksi

$$\mathbf{V} = \mathbf{X}\mathbf{X}'$$

jaotust nimetatakse Wisharti jaotuseks parameetritega Σ ja n .

Lühidalt tähistame asjaolu, et juhuslik $p \times p$ maatriks $\mathbf{V}$ on Wisharti jaotusega parameetritega Σ ja n , järgmiselt: $\mathbf{V} \sim W_p(\Sigma, n)$. Kui esimeseks parameetriks on ühikmaatriks, $\Sigma = \mathbf{I}_p$, siis ütleme, et tegemist on *standardse Wisharti jaotusega*.

Vastavalt plokk-maatriksite korrutamise eeskirjale avaldub definitsioonis määratud maatriks $\mathbf{V}$ summana

$$\mathbf{V} = \mathbf{X}\mathbf{X}' = [\vec{X}_1 \ \vec{X}_2 \ \dots \ \vec{X}_n] \begin{bmatrix} \vec{X}_1' \\ \vec{X}_2' \\ \dots \\ \vec{X}_n' \end{bmatrix} = \sum_{i=1}^n \vec{X}_i \vec{X}_i'. \quad (6.2)$$

Järgmises teoreemis esitame Wisharti jaotusega juhusliku maatriksi esimesed momendid.

Teoreem 6.1. Olgu $\mathbf{V} \sim W_p(\Sigma, n)$. Siis

$$E\mathbf{V} = n\Sigma,$$

$$D\mathbf{V} = n(\mathbf{I}_{p^2} + \mathbf{K}_{p,p})(\Sigma \otimes \Sigma),$$

kus $\mathbf{K}_{p,p}$ on kommuteerimismaatriks.

Tõestus. Esitus (6.2) võimaldab lihtsalt leida keskvärtuse. Tõepoolest,

$$E\mathbf{V} = E \sum_{i=1}^n \vec{X}_i \vec{X}_i' = \sum_{i=1}^n E(\vec{X}_i \vec{X}_i') = \sum_{i=1}^n D(\vec{X}_i) = n\Sigma.$$

Dispersiooni leidmiseks võtame esmalt $\vec{Z} \sim N(\vec{0}, \mathbf{I}_p)$ ja moodustame korrutise

$$\mathbf{W} = \vec{Z}\vec{Z}' = [Z_1\vec{Z} \ \dots \ Z_p\vec{Z}].$$

Maatriksi $\mathbf{W}$ keskvärtuse saame kujul

$$E\mathbf{W} = D\vec{Z} = \mathbf{I}_p.$$

Standardse normaaljaotusega juhusliku suuruse esimesi momente teame:

$$EZ_i = EZ_i^3 = 0, \quad EZ_i^2 = DZ_i = 1, \quad EZ_i^4 = 3.$$

Dispersioonimaatriks $D\mathbf{W} = D\text{vec}\mathbf{W}$ on plokkmaatriks, mis koosneb $p \times p$ -plokkidest

$$[D\mathbf{W}]_{ij} = [E(Z_i \vec{Z} Z_j \vec{Z}')] \quad i, j = 1, \dots, p.$$

Leiame ij -nda ploki avaldise. Selleks esitame järgmised keskvärtused ühikvektorite $\vec{e}_i$ kaudu

$$E(Z_i \vec{Z}) = \vec{e}_i,$$

$$E(Z_i Z_j \vec{Z} \vec{Z}') = \delta_{ij} \mathbf{I}_p + \vec{e}_i \vec{e}_j' + \vec{e}_j \vec{e}_i',$$

kus δ_{ij} on Kroneckeri delta. Kovariatsiooni definitsiooni kohaselt

$$\text{Cov}(Z_i \vec{Z}, Z_j \vec{Z}) = E[Z_i \vec{Z} (Z_j \vec{Z})'] - E(Z_i \vec{Z}) E(Z_j \vec{Z})' = \delta_{ij} \mathbf{I}_p + \vec{e}_j \vec{e}_i'.$$

Paneme tähele, et maatriksi $\mathbf{W}$ dispersioonimaatriksi esitamiseks saame kasutada kommuterimismaatriksit $\mathbf{K}_{p,p}$ (vt lisa 1, definitsioon L1.9). See maatriks koosneb $p \times p$ -plokkidest, kus ij -ndas plokis ji -s element võrdub ühega ja ülejäänud elemendid on nullid. Juhusliku maatriksi $\mathbf{W}$ dispersioonimaatriks võrdub definitsiooni kohaselt $\text{vec} \mathbf{W}$ dispersioonimaatriksiga. Seega

$$D\mathbf{W} = D(\text{vec} \mathbf{W}) = E[\text{vec} \mathbf{W} (\text{vec} \mathbf{W})'] - E(\text{vec} \mathbf{W}) E(\text{vec} \mathbf{W})'.$$

Kasutades plokkmaatriksi esitust plokkide kaudu, saame

$$D\mathbf{W} = [E(Z_i \vec{Z} Z_j \vec{Z}') - E(Z_i \vec{Z}) E(Z_j \vec{Z})'] = [\delta_{ij} \mathbf{I}_p + \vec{e}_j \vec{e}_i'],$$

$i, j = 1, \dots, p$. Siit aga

$$D\mathbf{W} = \mathbf{I}_{p^2} + \mathbf{K}_{p,p}.$$

Kui $\vec{X} \sim N_p(\vec{0}, \Sigma)$ ja $\mathbf{W}^* = \vec{X} \vec{X}'$, siis saame esitada maatriksi $\mathbf{W}^*$ maatriksi $\mathbf{W}$ kaudu:

$$\mathbf{W}^* = \mathbf{A} \mathbf{W} \mathbf{A}',$$

kus $\mathbf{A} \mathbf{A}' = \Sigma$. Siis kasutades valemit (L1.16)

$$\begin{aligned} D\mathbf{W}^* &= D(\mathbf{A} \mathbf{W} \mathbf{A}') = D(\text{vec}(\mathbf{A} \mathbf{W} \mathbf{A}')) \\ &= D[(\mathbf{A} \otimes \mathbf{A}) \text{vec} \mathbf{W}] = (\mathbf{A} \otimes \mathbf{A})(\mathbf{I}_{p^2} + \mathbf{K}_{p,p})(\mathbf{A} \otimes \mathbf{A})'. \end{aligned}$$

Kuna maatriksid $\mathbf{K}_{p,p}$ ja $\mathbf{A} \otimes \mathbf{A}$ kommuteeruvad otsekorrutise omaduse (vii) tõttu, saame

$$D\mathbf{W}^* = (\mathbf{I}_{p^2} + \mathbf{K}_{p,p})(\mathbf{A} \otimes \mathbf{A})(\mathbf{A} \otimes \mathbf{A})' = (\mathbf{I}_{p^2} + \mathbf{K}_{p,p})(\Sigma \otimes \Sigma).$$

Kasutades esitust (6.2) Wisharti jaotusega maatriksi $\mathbf{V}$ jaoks saame $D\mathbf{W}^*$ avaldisest tõestatava valemi dispersioonimaatriksi $D\mathbf{V}$ jaoks tänu vektorite $\vec{X}_i$ sõltumatusele:

$$D\mathbf{V} = nD(\vec{X}_i \vec{X}_i') = n(\mathbf{I}_{p^2} + \mathbf{K}_{p,p})(\Sigma \otimes \Sigma),$$

mott.

Wisharti jaotuse konstruktsioonist järelduvad omadused esitame järgnevas teoreemis.

Teoreem 6.2. *Wisharti jaotusel on järgmised omadused:*

- (i) *olgu $\mathbf{V}_1 \sim W_p(\Sigma, n_1)$ ja $\mathbf{V}_2 \sim W_p(\Sigma, n_2)$ sõltumatud juhuslikud maatriksid, siis $\mathbf{V}_1 + \mathbf{V}_2 \sim W_p(\Sigma, n_1 + n_2)$;*
- (ii) *olgu $\mathbf{V} \sim W_p(\Sigma, n)$ ja $\mathbf{C} : k \times p$ täisastakuga maatriks ($k \leq p$), siis $\mathbf{C} \mathbf{V} \mathbf{C}' \sim W_k(\mathbf{C} \Sigma \mathbf{C}', n)$;*
- (iii) *olgu $\mathbf{G} : n \times n$ sümmeetriline idempotentne maatriks ja $\mathbf{X} \mathbf{X}' \sim W_p(\Sigma, n)$, siis $\mathbf{X} \mathbf{G} \mathbf{X}' \sim W_p(\Sigma, r)$, kus r on maatriksi $\mathbf{G}$ astak.*

Tõestus. Väide (i). Valemi (6.2) kohaselt on iga Wisharti jaotusega juhuslik maatriks esitatav maatriksite summana. Seega

$$\mathbf{V}_1 = \sum_{i=1}^{n_1} \vec{X}_i^{(1)} \vec{X}_i^{(1)'} \quad \text{ja} \quad \mathbf{V}_2 = \sum_{i=1}^{n_2} \vec{X}_i^{(2)} \vec{X}_i^{(2)'}$$

Sellest esitusest aga järeldub nende maatriksite summa analoogne esitus

$$\mathbf{V}_1 + \mathbf{V}_2 = \sum_{i=1}^{n_1+n_2} \vec{X}_i^* \vec{X}_i^{*'},$$

kus

$$X_i^* = \begin{cases} \vec{X}_i^{(1)}, & \text{kui } i = 1, 2, \dots, n_1, \\ \vec{X}_{i-n_1}^{(2)}, & \text{kui } i = n_1 + 1, \dots, n_1 + n_2, \end{cases}$$

millest järeldub summa esitus maatriksite korrutisena

$$V_1 + V_2 = [\mathbf{X}_1 \ \mathbf{X}_2] \begin{bmatrix} \mathbf{X}'_1 \\ \mathbf{X}'_2 \end{bmatrix},$$

mis rahuldab definitsiooni 6.1 nõudeid. Väide (i) on tõestatud.

Väide (ii). Definitsiooni 6.1 kohaselt on maatriks $\mathbf{V}$ esitatav korrutisena $\mathbf{V} = \mathbf{X}\mathbf{X}'$, järelikult

$$\mathbf{C}\mathbf{V}\mathbf{C}' = \mathbf{C}\mathbf{X}\mathbf{X}'\mathbf{C}' = \mathbf{C}\mathbf{X}(\mathbf{C}\mathbf{X})'.$$

Kuna aga $\vec{X}_i \sim N(0, \mathbf{\Sigma})$, siis $\mathbf{C}\vec{X}_i \sim N(0, \mathbf{C}\mathbf{\Sigma}\mathbf{C}')$. Seega Wisharti jaotuse definitsiooni kohaselt $\mathbf{C}\mathbf{V}\mathbf{C}' \sim W_p(\mathbf{C}\mathbf{\Sigma}\mathbf{C}', n)$.

Väide (iii). Kuna maatriks $\mathbf{G}$ on sümmeetriline, siis on ta esitatav omavektorite ja omaväärtustest kaudu,

$$\mathbf{G} = \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}' = \sum_{i=1}^n \lambda_i \vec{\gamma}_i \vec{\gamma}_i'.$$

Maatriksi $\mathbf{G}$ idempotentsuse tõttu on r omaväärtust ühed, ülejäänud aga nullid. Eeldame, et maatriksi $\mathbf{G}$ omaväärtused on järjestatud mittekasvavalt. Siis

$$\mathbf{G} = \sum_{i=1}^r \vec{\gamma}_i \vec{\gamma}_i'.$$

Nüüd saame

$$\mathbf{X}\mathbf{G}\mathbf{X}' = \mathbf{X} \left(\sum_{i=1}^r \vec{\gamma}_i \vec{\gamma}_i' \right) \mathbf{X}' = \sum_{i=1}^r \mathbf{X} \vec{\gamma}_i \vec{\gamma}_i' \mathbf{X}'.$$

Seega oleme juhusliku maatriksi esitanud r maatriksi summana. Wisharti jaotuse definitsiooni rakendamiseks peame näitama, et liidetavaid maatrikseid moodustavad veeruvektorid $\mathbf{X}\vec{\gamma}_i$ on normaalkaotusega, $\mathbf{X}\vec{\gamma}_i \sim N(\vec{0}, \mathbf{\Sigma})$, ning sõltumatud. Kuna

$$\mathbf{X}\vec{\gamma}_i = \vec{X}_1 \gamma_{1i} + \vec{X}_2 \gamma_{2i} + \dots + \vec{X}_n \gamma_{ni}, \quad (6.3)$$

siis maatriksi $\mathbf{X}$ veeruvektorite sõltumatuse ja omavektorite normeerituse tõttu

$$D(\mathbf{X}\vec{\gamma}_i) = \mathbf{\Sigma} \sum_{j=1}^n \gamma_{ji}^2 = \mathbf{\Sigma}.$$

Kuna sõltumatute normaalkaotusega juhuslike vektorite lineaarkombinatsioon on normaalkaotusega, saamegi siit $\mathbf{X}\vec{\gamma}_i \sim N(\vec{0}, \mathbf{\Sigma})$. Arvutame erinevate veeruvektorite vahelise kovariatsiooni. Maatriksi $\mathbf{X}$ veeruvektorite sõltumatuse ja omavektorite ortogonaalsuse tõttu saame kasutades esitust (6.3)

$$\begin{aligned} \text{Cov}(\mathbf{X}\vec{\gamma}_i, \mathbf{X}\vec{\gamma}_j) &= E[\mathbf{X}\vec{\gamma}_i(\mathbf{X}\vec{\gamma}_j)'] = \sum_{l=1}^n \gamma_{li} \gamma_{lj} E(\vec{X}_l \vec{X}_l') \\ &= \mathbf{\Sigma} \sum_{l=1}^n \gamma_{li} \gamma_{lj} = \mathbf{\Sigma} \vec{\gamma}_i' \vec{\gamma}_j = \mathbf{0}. \end{aligned}$$

Seega saame maatriksi $\mathbf{X}\mathbf{G}\mathbf{X}'$ esitada definitsioonis 6.1 nõutava konstruktsiooniga, $\mathbf{X}\mathbf{G}\mathbf{X}' \sim W_p(\mathbf{\Sigma}, r)$, mott.

Järeldus 6.2.1. Olgu tunnusvektor normaaljaotusega, $\vec{X} \sim N_p(\vec{\mu}, \Sigma)$. Siis statistik $n\mathbf{S}_*$, on Wisharti jaotusega:

$$n\mathbf{S}_* \sim W(\Sigma, n-1),$$

kus $\mathbf{S}_*$ on antud valemiga (6.1).

Tõestus. Vastavalt valemile (2.2) võib statistiku $\mathbf{S}_*$ esitada sümmeetrilise ja idempotentse $n \times n$ -maatriksi $(\mathbf{I}_n - \frac{1}{n}\vec{\mathbf{l}}\vec{\mathbf{l}}')$ abil,

$$\mathbf{S}_* = \frac{1}{n}\mathbf{X}(\mathbf{I}_n - \frac{1}{n}\vec{\mathbf{l}}\vec{\mathbf{l}}')\mathbf{X}'. \quad (6.4)$$

Lihtne on veenduda, et nimetatud maatriksi astak on $\leq n-1$, kuna liites viimasele reale kõik eelnevad read, muutuvad viimase rea elemendid nullideks. Tegelikult on astak võrdne $(n-1)$, sest maatriksil $(\mathbf{I}_n - \frac{1}{n}\vec{\mathbf{l}}\vec{\mathbf{l}}')$ on $(n-1)$ mittenullilist omaväärtust. Siis aga teoreemi 6.2 väite (iii) kohaselt

$$\mathbf{M} = (\mathbf{X} - \vec{\mu}\vec{\mathbf{l}}'_n)(\mathbf{I}_n - \frac{1}{n}\vec{\mathbf{l}}_n\vec{\mathbf{l}}'_n)(\mathbf{X} - \vec{\mu}\vec{\mathbf{l}}'_n)' \sim W_p(\Sigma, n-1).$$

Veendume, et $\mathbf{M} = n\mathbf{S}_*$. Pärast sulgude avamist maatriksi $\mathbf{M}$ avaldises jääb alles ainsa mittenullilise liidetavana korrutis

$$\mathbf{X}(\mathbf{I}_n - \frac{1}{n}\vec{\mathbf{l}}_n\vec{\mathbf{l}}'_n)\mathbf{X}',$$

seega

$$n\mathbf{S}_* \sim W_p(\Sigma, n-1),$$

mott.

Järeldus 6.2.2. Olgu üldkogum normaaljaotusega, $N_p(\vec{\mu}, \Sigma)$. Siis statistiku

$$(n-1)\mathbf{S} = \sum_{i=1}^n (\vec{X}_i - \vec{X})(\vec{X}_i - \vec{X})',$$

jaotuseks on Wisharti jaotus,

$$(n-1)\mathbf{S} \sim W(\Sigma, n-1).$$

Tõestus. Väide järeldub otseselt eelmise järelduse tõestusest.

Eelmise paragrahvi lõpus nägime, et tõepärafunktsiooni kuju viitab statistikute $\vec{X}$ ja $\mathbf{S}$ sõltumatusele. Näitame nüüd, et see on tõepoolest nii.

Teoreem 6.3. Olgu $\mathbf{X} = [\vec{X}_1 \dots \vec{X}_n]$ teoreetiline valim, kus $\vec{X}_i \sim N_p(\vec{\mu}, \Sigma)$. Siis statistikud $\vec{X}$ ja $\mathbf{S}$ on sõltumatud, kus $\vec{X}$ ja $\mathbf{S}_*$ on antud valemitega (2.1) ja (6.4).

Tõestus. Kasutame jälle standardse normaaljaotusega sõltumatutest juhuslikest vektoritest $\vec{Y}_i$ moodustatud $p \times n$ -maatriksit $\mathbf{Y}$. Saame vaatlused $\vec{X}_i$ esitada $\vec{Y}_i$ kaudu

$$\vec{X}_i = \mathbf{A}\vec{Y}_i + \vec{\mu},$$

kus $\Sigma = \mathbf{A}\mathbf{A}'$. Suurima tõepära statistikud vektori $\vec{X}$ jaoks saame esitada vektori $\vec{Y}$ vastavate statistikute kaudu:

$$[\vec{X} \quad \mathbf{S}_*] = [\mathbf{A}\vec{Y} + \vec{\mu} \quad \mathbf{A}\mathbf{S}_{\vec{Y}}\mathbf{A}'].$$

Maatriksis $\mathbf{Y} : p \times n$ on kõik elemendid sõltumatud ja sama jaotusega, $Y_{ij} \sim N(0, 1)$. Võtame kasutusele kaks projektorit

$$\mathbf{G}_1 = \mathbf{I}_n - \frac{1}{n}\vec{\mathbf{l}}\vec{\mathbf{l}}' \quad \text{ja} \quad \mathbf{G}_2 = \frac{1}{n}\vec{\mathbf{l}}\vec{\mathbf{l}}'.$$

Paneme tähele, et need on ortogonaalsed ja nende summa on samasusteisendus:

$$\mathbf{G}_1 \mathbf{G}_2 = \mathbf{0}; \quad \mathbf{G}_1 + \mathbf{G}_2 = \mathbf{I}_n.$$

Seega projekteerivad nad maatriksi $\mathbf{Y}$ ridadest $\mathbf{Y}^{(i)}$ moodustatud vektorid ortogonaalsetele alamruumidele ruumis $\mathbb{R}^n$. Veelgi enam, need projektsioonid on sõltumatud juhuslikud vektorid. Veendume selles. Teame, et vektorid $\mathbf{G}_1(\mathbf{Y}^{(i)})'$ ja $\mathbf{G}_2(\mathbf{Y}^{(j)})'$ on normaalsootusega. Sõltumatuse näitamiseks piisab näidata, et

$$\text{Cov}(\mathbf{G}_1(\mathbf{Y}^{(i)})', \mathbf{G}_2(\mathbf{Y}^{(j)})') = \mathbf{0}, \quad i, j = 1, \dots, p.$$

Tõepoolest:

$$\begin{aligned} & \text{Cov}(\mathbf{G}_1(\mathbf{Y}^{(i)})', \mathbf{G}_2(\mathbf{Y}^{(j)})') \\ &= E(\mathbf{G}_1(\mathbf{Y}^{(i)})'(\mathbf{Y}^{(j)})\mathbf{G}_2) = \mathbf{G}_1 \mathbf{G}_2 = \mathbf{0}. \end{aligned}$$

Siis aga ka $\mathbf{G}_1 \mathbf{Y}'$ ja $\mathbf{G}_2 \mathbf{Y}'$ on sõltumatud. Samas $n\vec{Y} = \vec{\mathbb{I}}_n' \mathbf{G}_2 \mathbf{Y}'$ ning $n\mathbf{S}_{\vec{Y}}^* = \mathbf{Y} \mathbf{G}_1 \mathbf{G}_1 \mathbf{Y}'$ ja seega on ka need statistikud sõltumatud. Sellest omakorda järeldub aga statistikute $\vec{X}$ ja $\mathbf{S}_*$ sõltumatus, kuna nad on saadud statistikute $\vec{Y}$ ja $\mathbf{S}_{\vec{Y}}^*$ lineaarseisendusega, mis sõltumatust ei mõjuta, mott.

Ülesanded 6

- 6.1. Leida maatriksi $\mathbf{I}_n - \frac{1}{n} \vec{\mathbb{I}}_n \vec{\mathbb{I}}_n'$ omaväärtused.
- 6.2. Leida juhusliku suuruse $Z \sim N(0, 1)$ momendid $m_3(Z)$ ja $m_4(Z)$.
- 6.3. Kontrollida, kas $\mathbf{G}_1 = \mathbf{I}_n - \frac{1}{n} \vec{\mathbb{I}}_n \vec{\mathbb{I}}_n'$ ja $\mathbf{G}_2 = \frac{1}{n} \vec{\mathbb{I}}_n \vec{\mathbb{I}}_n'$ on projektorid.

7 Wisharti jaotuse karakteristik funktsioon ja tõenäosustihedus

Teame, et juhusliku vektori karakteristik funktsioon on määratud eeskirjaga

$$\varphi_{\vec{X}}(\vec{t}) = E(e^{i\vec{t}'\vec{X}}).$$

Juhusliku maatriksi karakteristikliku funktsiooni defineerimisel esitame maatriksi vektoriseeritud kujul ja lähtume siis juhusliku vektori karakteristikliku funktsiooni eeskirjast. Rakendades $p \times n$ -maatriksile $\mathbf{X}$ vektoriseerimise operatsiooni vec , saame pn -vektori

$$\text{vec} \vec{X} = \begin{bmatrix} \vec{X}_1 \\ \vec{X}_2 \\ \dots \\ \vec{X}_p \end{bmatrix}.$$

Juhusliku maatriksi $\mathbf{X} : p \times n$ karakteristik funktsioon (mille argumendiks peab olema $p \times n$ -maatriks $\mathbf{T}$) määratakse eeskirjaga

$$\varphi_{\mathbf{X}}(\mathbf{T}) = \varphi_{\text{vec} \mathbf{X}}(\text{vec} \mathbf{T}) = E e^{i(\text{vec}' \mathbf{T} \text{vec} \mathbf{X})}.$$

Arvestades maatriksi jälje omadust

$$\text{tr}(\mathbf{A}' \mathbf{B}) = \text{vec}' \mathbf{A} \text{vec} \mathbf{B},$$

saame maatriksi $\mathbf{X}$ karakteristlikule funktsioonile esituse

$$\varphi_{\mathbf{X}}(\mathbf{T}) = Ee^{i\text{tr}(\mathbf{T}'\mathbf{X})}.$$

Kui juhuslik maatriks on positiivselt määratud (seega sümmeetriline), peab ka argument $\mathbf{T}$ olema positiivselt määratud. Seetõttu Wisharti jaotusega juhusliku maatriksi $\mathbf{V}$ korral

$$\varphi_{\mathbf{V}}(\mathbf{T}) = Ee^{i\text{tr}(\mathbf{T}\mathbf{V})}.$$

Leiame selle keskvärtuse. Vastavalt valemile (6.2) on Wisharti jaotusega maatriks $\mathbf{V} \sim W_p(\mathbf{\Sigma}, n)$ esitatav kujul

$$\mathbf{V} = \sum_{j=1}^n \vec{X}_j \vec{X}_j',$$

kus $\vec{X}_j \sim N_p(\vec{0}, \mathbf{\Sigma})$. Lahutuse $\mathbf{\Sigma} = \mathbf{A}\mathbf{A}'$ tõttu on vektor $\vec{X}_j$ esitatav standardse normaaljaotusega vektori $\vec{Z}_j$ kaudu, $\vec{X}_j = \mathbf{A}\vec{Z}_j$. Karakteristliku funktsiooni saame seetõttu esitada kujul

$$\varphi_{\mathbf{V}}(\mathbf{T}) = Ee^{i\text{tr}(\mathbf{A}'\mathbf{T}\mathbf{A} \sum_{j=1}^n \vec{Z}_j \vec{Z}_j')}.$$

Maatriks $\mathbf{A}'\mathbf{T}\mathbf{A}$ on sümmeetriline, seda saab esitada omaväärtuste ja omavektorite kaudu,

$$\mathbf{A}'\mathbf{T}\mathbf{A} = \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}',$$

kus omavektorite maatriks $\mathbf{\Gamma}$ on ortogonaalne ($\mathbf{\Gamma}\mathbf{\Gamma}' = \mathbf{\Gamma}'\mathbf{\Gamma} = \mathbf{I}_p$), omaväärtuste maatriks $\mathbf{\Lambda}$ aga diagonaalmaatriks. Karakteristliku funktsiooni jaoks saame avaldise

$$\varphi_{\mathbf{V}}(\mathbf{T}) = Ee^{i\text{tr}(\mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}' \sum_{j=1}^n \vec{Z}_j \vec{Z}_j')} = Ee^{i\text{tr}(\mathbf{\Lambda} \sum_{j=1}^n \mathbf{\Gamma}' \vec{Z}_j \vec{Z}_j' \mathbf{\Gamma})}.$$

Kuna aga ortogonaalteisendus ei muuda standardse normaaljaotusega vektori jaotust, $\mathbf{\Gamma}' \vec{Z}_j \sim N(\vec{0}, \mathbf{I}_p)$, võime kirjutada

$$\varphi_{\mathbf{V}}(\mathbf{T}) = Ee^{i\text{tr}(\mathbf{\Lambda} \sum_{j=1}^n \vec{Z}_j \vec{Z}_j')} = Ee^{i \sum_{j=1}^n \sum_{k=1}^p Z_{kj}^2 \lambda_k}.$$

Arvestades, et eksponentfunktsiooni astendajaks olevas summas on liidetavad sõltumatud, saame

$$\varphi_{\mathbf{V}}(\mathbf{T}) = \prod_{j=1}^n \prod_{k=1}^p E(e^{iZ_{kj}^2 \lambda_k}) = \prod_{j=1}^n \prod_{k=1}^p \varphi_{\chi_1^2}(\lambda_k).$$

Nüüd jääb üle veel asendada korrutisse $\chi^2(1)$ -jaotuse karakteristliku funktsiooni avaldis. Saame

$$\varphi_{\mathbf{V}}(\mathbf{T}) = \prod_{j=1}^n \prod_{k=1}^p |1 - 2i\lambda_k|^{-\frac{1}{2}} = \prod_{j=1}^n |\mathbf{I}_p - 2i\mathbf{\Lambda}|^{-\frac{1}{2}} = |\mathbf{I}_p - 2i\mathbf{\Lambda}|^{-\frac{n}{2}}.$$

Seetõttu, et omavektorite ortogonaalmaatriksi determinant võrdub ühega, võime kirjutada

$$\varphi_{\mathbf{V}}(\mathbf{T}) = (|\mathbf{\Gamma}||\mathbf{I}_p - 2i\mathbf{\Lambda}||\mathbf{\Gamma}'|)^{-n/2} = |\mathbf{I}_p - 2i\mathbf{A}'\mathbf{T}\mathbf{A}|^{-n/2}.$$

Kasutades determinandi omadusi jõuame võrdusteni

$$\varphi_{\mathbf{V}}(\mathbf{T}) = (|\mathbf{A}'|(|\mathbf{A}'|^{-1}|\mathbf{A}|^{-1} - 2i\mathbf{T}||\mathbf{A}|)^{-\frac{n}{2}} = (|\mathbf{\Sigma}^{-1} - 2i\mathbf{T}||\mathbf{\Sigma}|)^{-\frac{n}{2}},$$

kust lõppkokkuvõttes saame

$$\varphi_{\mathbf{V}}(\mathbf{T}) = |\mathbf{I}_p - 2i\mathbf{T}\mathbf{\Sigma}|^{-\frac{n}{2}}.$$

Seega oleme tõestanud järgmise teoreemi.

Teoreem 7.1. Olgu juhuslik maatriks $\mathbf{V}$ Wisharti jaotusega, $\mathbf{V} \sim W(\mathbf{\Sigma}, n)$. Siis tema karakteristlik funktsioon on määratud valemiga

$$\varphi_{\mathbf{V}}(\mathbf{T}) = |\mathbf{I}_p - 2i\mathbf{T}\mathbf{\Sigma}|^{-\frac{n}{2}}$$

kus argument $\mathbf{T}$ on positiivselt määratud reaalarvuline maatriks.

Tihedusfunktsiooni tuletamine on tülilikam. Esitame Wisharti jaotuse tiheduse järgmises teoreemis tõestuseta (tõestus vt näiteks Muirhead (1982), Kollo and von Rosen (2005)).

Teoreem 7.2. *Olgu juhuslik maatriks $\mathbf{V}$ Wisharti jaotusega, $\mathbf{V} \sim W(\mathbf{\Sigma}, n)$. Siis tema tihedusfunktsioon on määratud eeskirjaga*

$$f_{\mathbf{V}}(\mathbf{V}) = \frac{|\mathbf{V}|^{(n-p-1)/2}}{2^{pn/2} \pi^{p(p-1)/4} |\mathbf{\Sigma}|^{n/2} \prod_{k=1}^p \Gamma[\frac{1}{2}(n+1-k)]} e^{-\frac{1}{2} \text{tr}(\mathbf{\Sigma}^{-1} \mathbf{V})},$$

kus argument $\mathbf{V}$ on positiivselt määratud reaalarvuline maatriks.

Ülesanded 7

- 7.1. Leida Wisharti jaotusega $W_p(\mathbf{\Sigma}, n)$ juhusliku maatriksi $\mathbf{V}$ keskväärus karakteristliku funktsiooni diferentseerimise teel.
- 7.2. Näidata, et kahe sõltumatu maatriksi $\mathbf{V}_1 \sim W_p(\mathbf{\Sigma}, n_1)$ ja $\mathbf{V}_2 \sim W_p(\mathbf{\Sigma}, n_2)$ summa karakteristlik funktsioon on Wisharti jaotuse $W_p(\mathbf{\Sigma}, n_1 + n_2)$ karakteristlik funktsioon.

8 Hotellingi T^2 -jaotus

Ühemõõtmelisel juhul tuli keskvääruse vahemikhinnangute saamiseks defineerida uus jaotus – t -jaotus. See oli suhte

$$\sqrt{n} \frac{\bar{X} - \mu}{S}$$

jaotus, kus $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$. Samasugune vajadus tekib ka mitmemõõtmelisel juhul. Statistikuks, mille jaotuse leiame, on nüüd ruutvorm

$$n(\vec{X} - \vec{\mu})' \mathbf{S}^{-1} (\vec{X} - \vec{\mu}) \quad (8.1)$$

või kasutades dispersioonimaatriksi suurima tõepära hinnangut, ruutvorm

$$(n-1)(\vec{X} - \vec{\mu})' \mathbf{S}_*^{-1} (\vec{X} - \vec{\mu}).$$

Määrame uue tõenäosusjaotuse teatava konstruktsiooni abil.

Definitsioon 8.1. *Olgu normaaljaotusega juhuslik vektor $\vec{X}$, $\vec{X} \sim N(\vec{0}, \mathbf{I}_p)$, ja Wisharti jaotusega juhuslik maatriks $\mathbf{V}$, $\mathbf{V} \sim W(\mathbf{I}_p, n)$, sõltumatud. Siis juhusliku suuruse*

$$T^2 = n \vec{X}' \mathbf{V}^{-1} \vec{X}$$

jaotust nimetatakse Hotellingi T^2 -jaotuseks parameetritega p ja n .

Lühidalt tähistame selle jaotuse $T^2 \sim T^2(p, n)$.

Vaatame kuidas seda jaotust kasutada ruutvormi (8.1) kirjeldamiseks. Sõnastame teoreemina järgmise väite.

Teoreem 8.1. *Olgu normaaljaotusega üldkogumist, $\vec{X} \sim N_p(\vec{\mu}, \mathbf{\Sigma})$, saadud valim mahuga n . Siis*

$$n(\vec{X} - \vec{\mu})' \mathbf{S}^{-1} (\vec{X} - \vec{\mu}) \sim T^2(p, n-1).$$

Tõestus. Lähtume normaalkaotusega tunnusevektorist $\vec{X}$. Siis on ka statistik $\vec{X}$ normaalkaotusega, $\vec{X} \sim N(\vec{\mu}, \frac{1}{n}\Sigma)$. Vastavalt normaalkaotuse definitsioonile leidub pööratav maatriks $\mathbf{A}$ nii, et $\mathbf{A}\mathbf{A}' = \Sigma$ ja juhuslik vektor

$$\vec{X}_0 = \sqrt{n}\mathbf{A}^{-1}(\vec{X} - \vec{\mu}) \quad (8.2)$$

on standardse normaalkaotusega, $\vec{X}_0 \sim N(\vec{0}, \mathbf{I}_p)$. Järelduse 6.2.2 põhjal $(n-1)\mathbf{S} \sim W(\Sigma, n-1)$. Kasutades Wisharti jaotuse omadust (ii) teoreemis 6.2 saame ka statistiku $\mathbf{S}$ standardiseerida. Määrame statistiku

$$\mathbf{V}_0 = (n-1)\mathbf{A}^{-1}\mathbf{S}\mathbf{A}'^{-1}, \quad (8.3)$$

siis $\mathbf{V}_0 \sim W(\mathbf{I}_p, n-1)$. Kuna statistikud $\vec{X}$ ja $\mathbf{S}$ on sõltumatud, siis on ka $\vec{X}_0$ ja $\mathbf{V}_0$ sõltumatud. Neist aga saame kokku panna juhusliku suuruse, mis on T^2 -jaotusega,

$$(n-1)\vec{X}_0'\mathbf{V}_0^{-1}\vec{X}_0 \sim T^2(p, n-1).$$

Näitame, et selle juhusliku suuruse saame esitada ruutvormina (8.1). Tõepoolest, asendades $\vec{X}_0$ ja $\mathbf{V}_0$ nende määramiseeskirjadest (8.2) ja (8.3), saame

$$\begin{aligned} (n-1)\vec{X}_0'\mathbf{V}_0^{-1}\vec{X}_0 &= (n-1)\sqrt{n}(\vec{X} - \vec{\mu})'(\mathbf{A}')^{-1}\mathbf{A}'\frac{1}{n-1}\mathbf{S}^{-1}\mathbf{A} \\ &\times \sqrt{n}\mathbf{A}^{-1}(\vec{X} - \vec{\mu}) = n(\vec{X} - \vec{\mu})'\mathbf{S}^{-1}(\vec{X} - \vec{\mu}), \end{aligned}$$

mott.

Osutub, et Hotellingi jaotuse saame taandada F -jaotusele. Olgu $T^2 \sim T^2(p, n)$. Siis

$$\frac{n-p+1}{pn}T^2 \sim F(p, n-p+1). \quad (8.4)$$

Seega

$$\frac{n-p}{p(n-1)}n(\vec{X} - \vec{\mu})'\mathbf{S}^{-1}(\vec{X} - \vec{\mu}) \sim F(p, n-p).$$

Järeldus 8.1.1. *Teoreemi 8.1 eeldustel*

$$(n-1)(\vec{X} - \vec{\mu})'\mathbf{S}_*^{-1}(\vec{X} - \vec{\mu}) \sim T^2(p, n-1). \quad (8.5)$$

Tõestus. Väiteni (8.5) jõuame teoreemi mõttekäigu kordamisel $\mathbf{S}_*$ jaoks, mott.

Ülesanded 8

8.1. Tõestada, et

$$(n-1)(\vec{X} - \vec{\mu})'\mathbf{S}_*^{-1}(\vec{X} - \vec{\mu}) \sim T^2(p, n-1).$$

8.2. Tõestada seos (8.4).

9 Hüpoteesid üldkogumi keskvärtuse kohta

Statistiline hüpotees on väide üldkogumi kohta. Otsus väite õigsuse kohta langetatakse valimi põhjal. Kui väide puudutab üldkogumi parameetreid, nimetatakse seda parameetriliseks hüpoteesiks. Enamasti esitatakse statistilised hüpoteesid alternatiivsete väidete H_0 ja H_1 paarina, millest esimest nimetatakse null-hüpoteesiks ja teist sisukaks ehk alternatiivseks hüpoteesiks.

Vaatame parameetriliste hüpoteeside paari parameetri θ kohta. Tähistame selle parameetri võimalike väärtuste ruumi sümboliga Θ . Üldjuhul jaotatakse hüpoteeside esitamisel parameetri võimalike väärtuste ruum osadeks Θ_0 ja Θ_1 , kusjuures

$$\Theta = \Theta_0 \cup \Theta_1 \quad \text{ja} \quad \Theta_0 \cap \Theta_1 = \emptyset.$$

Üldise hüpoteeside paari parameetri θ kohta võime esitada kujul

$$\begin{aligned} H_0 : & \quad \theta \in \Theta_0 \\ H_1 : & \quad \theta \in \Theta_1. \end{aligned} \tag{9.1}$$

Otsuse langetamiseks

- valitakse teststatistik $T(\mathbf{X})$,
- leitakse teststatistiku jaotus H_0 kehtivuse korral,
- määratakse olulisuse nivoo α ,
- leitakse kriitiline piirkond K_α nii, et $P(T(\mathbf{X}) \in K_\alpha \mid H_0) \leq \alpha$,
- arvutatakse statistiku väärtus $T(\mathcal{X})$, antud valimi $\mathcal{X}$ korral
- langetatakse otsus standardse reegli põhjal

$$\begin{aligned} \text{kui} \quad & T(\mathcal{X}) \in K_\alpha \implies H_1, \\ \text{kui} \quad & T(\mathcal{X}) \notin K_\alpha \implies H_0. \end{aligned}$$

Otsust langetades võib tekkida üks kahest võimalikust veast. Võttes vastu sisuka hüpoteesi H_1 võib teha 1. liiki vea. Selle vea tegemise tõenäosus on tõkestatud olulisuse nivooga α . Võttes vastu nullhüpoteesi H_0 võib teha 2. liiki vea. Selle vea tegemise tõenäosus on tõkestatud suurusega $1 - \alpha$. Vastavalt vea tõenäosusele on ka suhtumine vastuvõetud hüpoteesi erinev – vastuvõetud sisukas hüpotees loetakse tõestatuks, vastuvõetud nullhüpoteesi kohta öeldakse, et seda ei õnnestunud kummutada. Seega jääb võimalus, et suurendades valimi mahtu, uurides üldkogumit põhjalikumalt, õnnestub nullhüpotees kummutada.

Esimeseks probleemiks otsuse langetamise eeskirja leidmisel on teststatistiku valik. Siin võib toetuda intuitsioonile, kasutada analoogiat ühemõõtmelise juhuga. Kuid selleks valikuks on olemas ka rida standardseid eeskirju. Väga paljudel juhtudel kasutatakse teststatistikuna *tõepärasuhte statistikut*. Seda statistikut kasutatakse sageli ka ühemõõtmelistes ülesannetes. Defineerime tõepärasuhte statistiku.

Definitsioon 9.1. Olgu $L(\mathcal{X}, \theta)$ tõepärafunktsioon valimi $\mathcal{X}$ korral ja $\mathbf{X}$ vastav teoreetiline valim. Vajagu kontrollimist hüpoteeside paar (9.1). Statistikut

$$\Lambda(\mathbf{X}) = \frac{\max_{\theta \in \Theta_0} L(\mathbf{X}, \theta)}{\max_{\theta \in \Theta} L(\mathbf{X}, \theta)}$$

nimetatakse *tõepärasuhte (TPS) statistikuks*.

Nagu näha, sõltub tõepärasuhte statistiku kuju üldkogumi tõenäosusjaotusest ja esitatud hüpoteesidest. TPS statistikul on ilus sisuline tähendus: ta näitab maksimaalsete tõepärase suhet saada uurija käsutuses olev konkreetne valim. Lugejas olev maksimum arvutatakse eeldusel, et üldkogumi parameeter kuulub nullhüpoteesis näidatud piirkonda, nimetajas olev

maksimum aga eeldusel, et üldkogumi parameeter võib omandada mistahes võimaliku väärtuse. Kuna $\Theta_0 \subset \Theta$, siis võib TPS statistik omada väärtusi lõigul $[0,1]$. Selle statistikuga seotud kriitiline piirkond on standardse kujuga

$$K_\alpha = [0, c],$$

kus konstandi c määramiseks tuleb teada statistiku $\Lambda(\mathbf{X})$ jaotust nullhüpoteesi korral. Seega TPS statistiku kasutamisel langetatakse otsus järgmise reegli põhjal:

$$\text{kui } \Lambda(\mathcal{X}) \leq c \Rightarrow H_1, \quad \text{kui } \Lambda(\mathcal{X}) > c \Rightarrow H_0.$$

Vaatame selle skeemi rakendamist normaaljaotusega üldkogumis, $\vec{X} \sim N(\vec{\mu}, \Sigma)$, kontrollimaks hüpoteese keskväärtusvektori võrdumisest teadaoleva vektoriga $\vec{\mu}_0$,

$$\begin{aligned} H_0 : & \quad \vec{\mu} = \vec{\mu}_0 \\ H_1 : & \quad \vec{\mu} \neq \vec{\mu}_0. \end{aligned}$$

Leiame tõepärasuhte statistiku selle hüpoteesipaari korral. Arvutame esmalt nimetajas oleva maksimumi – tõepärafunktsiooni maksimaalse väärtuse parameetrite kõikvõimalike väärtuste ruumis. Teatavasti omandab tõepärafunktsioon maksimaalse väärtuse, kui parameetriteks valida suurima tõepära hinnangud $\hat{\vec{\mu}} = \vec{x}$ ja $\hat{\Sigma} = \hat{\Sigma}_*$. Kasutades tõepärafunktsiooni esitust (5.2) saame

$$\begin{aligned} L(\mathcal{X}; \vec{x}, \hat{\Sigma}_*) &= \frac{1}{(2\pi)^{\frac{np}{2}} |\hat{\Sigma}_*|^{\frac{n}{2}}} \exp \left\{ -\frac{n}{2} \text{tr}[\hat{\Sigma}_*^{-1}(\hat{\Sigma}_* + \mathbf{0})] \right\} \\ &= \frac{1}{(2\pi)^{\frac{np}{2}} |\hat{\Sigma}_*|^{\frac{n}{2}}} \exp \left\{ -\frac{n}{2} \text{tr} \mathbf{I}_p \right\} = \frac{1}{(2\pi)^{\frac{np}{2}} |\hat{\Sigma}_*|^{\frac{n}{2}}} e^{-\frac{np}{2}}. \end{aligned} \quad (9.2)$$

Lugejas olev tõepärafunktsiooni maksimum tuleb leida eeldusel, et keskväärtusvektori väärtus on $\vec{\mu}_0$. Lähtudes tõepärafunktsiooni avaldisest (5.1) võime kirjutada

$$L(\mathcal{X}; \vec{\mu}_0, \Sigma) = \frac{1}{(2\pi)^{\frac{np}{2}} |\Sigma|^{\frac{n}{2}}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n (\vec{x}_i - \vec{\mu}_0)' \Sigma^{-1} (\vec{x}_i - \vec{\mu}_0) \right\}.$$

Kasutades maatriksi jälje omadust $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$ saame tõepärafunktsiooni esitada kujul

$$L(\mathcal{X}; \vec{\mu}_0, \Sigma) = \frac{1}{(2\pi)^{\frac{np}{2}} |\Sigma|^{\frac{n}{2}}} \exp \left\{ -\frac{1}{2} \text{tr}[\Sigma^{-1} \sum_{i=1}^n (\vec{x}_i - \vec{\mu}_0)(\vec{x}_i - \vec{\mu}_0)'] \right\}.$$

Tähistades

$$n\hat{\Sigma}_0 = \sum_{i=1}^n (\vec{x}_i - \vec{\mu}_0)(\vec{x}_i - \vec{\mu}_0)'$$

võime kirjutada

$$L(\mathcal{X}; \vec{\mu}_0, \Sigma) = \frac{1}{(2\pi)^{\frac{np}{2}} |\Sigma|^{\frac{n}{2}}} \exp \left\{ -\frac{n}{2} \text{tr}(\Sigma^{-1} \hat{\Sigma}_0) \right\}.$$

Korrates viiendas paragrahvis läbi viidud teisendusi saame dispersioonimaatriksi STP hinnanguks eeldusel, et keskväärtusvektoriks on vektor $\vec{\mu}_0$ maatriksi

$$\hat{\Sigma} = \hat{\Sigma}_0$$

ja tõepärafunktsiooni väärtuseks selle hinnangu korral on

$$L(\mathcal{X}; \vec{\mu}_0, \hat{\Sigma}_0) = \frac{1}{(2\pi)^{\frac{np}{2}} |\hat{\Sigma}_0|^{\frac{n}{2}}} e^{-\frac{np}{2}}.$$

Seega tehtud eeldustel omab tõepärasuhte statistik kuju

$$\Lambda(\mathbf{X}) = \left(\frac{|\mathbf{S}_*|}{|\mathbf{S}_0|} \right)^{n/2}.$$

Ülaloodud skeemi järgides peaksime nüüd leidma tõepärasuhte statistiku jaotuse eeldusel, et nullhüpotees kehtib. Selle ülesande lihtsustamiseks püüame üle minna lihtsamale teststatistikule, mis oleks tõepärasuhte statistikuga samaväärne. Vaatame teststatistiku leidmiseks esmalt suhet $\frac{|\mathbf{S}_0|}{|\mathbf{S}_*|}$. Esitame selle suhte lihtsamal kujul.

Kasutame determinandi omadust: kui $\mathbf{B} = \mathbf{A} + \vec{a}\vec{a}'$, siis

$$|\mathbf{B}| = |\mathbf{A}|(1 + \vec{a}'\mathbf{A}^{-1}\vec{a}).$$

Tõepoolest, kuna $\mathbf{B} : n \times n$ korral

$$|\mathbf{B}| = |\mathbf{A} + \vec{a}\vec{a}'| = |\mathbf{A}||\mathbf{I}_n + \mathbf{A}^{-1}\vec{a}\vec{a}'|,$$

siis rakendades determinandi omadust

$$|\mathbf{I}_n + \mathbf{A}\mathbf{B}| = |\mathbf{I}_m + \mathbf{B}\mathbf{A}|, \quad \mathbf{A} : n \times m, \quad \mathbf{B} : m \times n,$$

saamegi, et

$$|\mathbf{B}| = |\mathbf{A}|(1 + \vec{a}'\mathbf{A}^{-1}\vec{a}).$$

Teisendame maatriksit $\mathbf{S}_0$,

$$\begin{aligned} n\mathbf{S}_0 &= \sum_{i=1}^n (\vec{X}_i - \vec{\mu}_0)(\vec{X}_i - \vec{\mu}_0)' = \sum_{i=1}^n (\vec{X}_i - \vec{X})(\vec{X}_i - \vec{X})' \\ &\quad + n(\vec{X} - \vec{\mu}_0)(\vec{X} - \vec{\mu}_0)' = n\mathbf{S}_* + n(\vec{X} - \vec{\mu}_0)(\vec{X} - \vec{\mu}_0)'. \end{aligned}$$

Seega

$$|\mathbf{S}_0| = |\mathbf{S}_*|(1 + (\vec{X} - \vec{\mu}_0)'\mathbf{S}_*^{-1}(\vec{X} - \vec{\mu}_0)).$$

Meenutades valemit (8.6) näeme, et sulgudes olev ruutvorm

$$(\vec{X} - \vec{\mu}_0)'\mathbf{S}_*^{-1}(\vec{X} - \vec{\mu}_0)$$

on nullhüpoteesi kehtivuse korral kirjeldatav Hotellingu T^2 -jaotusega, kuna

$$T_*^2(\mathbf{X}) = (n-1)(\vec{X} - \vec{\mu}_0)'\mathbf{S}_*^{-1}(\vec{X} - \vec{\mu}_0) \sim T^2(p, n-1).$$

Kokkuvõttes oleme tõepärasuhte statistiku esitanud kujul

$$\Lambda(\mathbf{X}) = \left(1 + \frac{T_*^2(\mathbf{X})}{n-1}\right)^{-\frac{n}{2}}.$$

Me oleme saanud TPS statistiku esitada ühe statistiku monotoonselt kahaneva funktsioonina, st leiduvad niisugused konstandid c ja c_{kr} , et võrratused

$$\Lambda(\mathcal{X}) \leq c, \Rightarrow T_*^2(\mathcal{X}) \geq c_{kr}$$

kehtivad üheaegselt. Tegemist on väga kasuliku esitusega: statistiku $T_*^2(\mathbf{X})$ abil saame esitada tõepärasuhte statistikuga $\Lambda(\mathbf{X})$ samaväärse otsustuseeskirja, kusjuures kriitiline piirkond $K_\alpha = [c_{kr}, \infty)$. Nullhüpoteesi kehtides on

$$T_*^2(\mathbf{X}) \sim T^2(p, n-1), \quad \text{ehk} \quad \frac{n-p}{p(n-1)}T_*^2(\mathbf{X}) \sim F(p, n-p).$$

Kui oleme fikseerinud olulisuse nivoo α , saame valida kriitilise piirkonna alumiseks otspunktiks F -jaotuse α -täiendkvantiili $\bar{f}_{\alpha, p, n-p}$. Seega oleme jõudnud otsustuseeskirjani

$$\text{kui} \quad \frac{n-p}{p(n-1)}T_*^2(\mathcal{X}) \geq \bar{f}_{\alpha, p, n-p} \Rightarrow H_1,$$

$$\text{kui} \quad \frac{n-p}{p(n-1)}T_*^2(\mathcal{X}) < \bar{f}_{\alpha, p, n-p} \Rightarrow H_0.$$

Ülesanded 9

9.1. Näidata, et fikseeritud keskvaartuse $\vec{\mu}_0$ korral on suurima tõepära hinnanguks normaaljaotuse dispersioonimaatriksile avaldis

$$\hat{\Sigma} = \hat{\mathbf{S}}_0 = \frac{1}{n} \sum_{i=1}^n (\vec{x}_i - \vec{\mu}_0)(\vec{x}_i - \vec{\mu}_0)'.$$

10 Keskvaartusvektori usalduspiirkond

Ühe parameetri θ korral kasutati hinnangu täpsuse kirjeldamiseks usaldusvahemikku. Etteantud tõenäosuse – usaldusnivoo γ – korral oli usaldusvahemikuks teoreetilise valimi abil määratud juhuslik vahemik $(a(\mathbf{X}), b(\mathbf{X}))$, mille korral

$$P(\theta \in (a(\mathbf{X}), b(\mathbf{X}))) \geq \gamma.$$

Tihti esitati usaldusnivoo kujul $1 - \alpha$, kus α on nullile lähedane tõenäosus.

Sama idee realiseerub mitmemõõtmelisel juhul *usaldushulga* e. *usalduspiirkonna* kaudu. Definieerime usalduspiirkonna.

Definitsioon 10.1. Parameetri θ usalduspiirkonnaks usaldusnivooga γ nimetatakse teoreetilise valimi abil määratud juhuslikku piirkonda $R(\mathbf{X})$, kui

$$P(\theta \in R(\mathbf{X})) \geq \gamma.$$

Erinevaid usalduspiirkondi on võimalik konstrueerida palju. Kasulik on piiritleda parameetreid võimalikult täpselt, st konstrueerida see hulk nii väike kui võimalik. Vaatame järgnevas usalduspiirkonna konstrueerimist normaaljaotusega $N(\vec{\mu}, \Sigma)$ üldkogumi keskvaartusvektori $\vec{\mu}$ jaoks.

A. Usaldusellipsoid. Lähtume Hotellingi T^2 -statistikust

$$T^2(\mathbf{X}) = n(\vec{X} - \vec{\mu})' \mathbf{S}^{-1} (\vec{X} - \vec{\mu}),$$

mis on T^2 -jaotusega $T^2(p, n-1)$. Korrutades seda juhuslikku suurust sobivalt valitud konstandiga saame F -jaotusega juhusliku suuruse,

$$\frac{n-p}{p(n-1)} n(\vec{X} - \vec{\mu})' \mathbf{S}^{-1} (\vec{X} - \vec{\mu}) \sim F(p, n-p).$$

Valides nüüd $\gamma = 1 - \alpha$, saame F -jaotuse α -täiendkvantiili $\bar{f}_{\alpha, p, n-p}$ abil välja kirjutada definitsioonile vastava usalduspiirkonna usaldusnivooga $1 - \alpha$,

$$P\left(\frac{n-p}{p(n-1)} n(\vec{X} - \vec{\mu})' \mathbf{S}^{-1} (\vec{X} - \vec{\mu}) \leq \bar{f}_{\alpha, p, n-p}\right) = 1 - \alpha.$$

Seega on usalduspiirkonnaks usaldusnivooga $1 - \alpha$ juhuslik ellipsoid p -mõõtmelises ruumis

$$(\vec{X} - \vec{\mu})' \mathbf{S}^{-1} (\vec{X} - \vec{\mu}) \leq \frac{p(n-1)}{(n-p)n} \bar{f}_{\alpha, p, n-p}. \quad (10.1)$$

Selle ellipsoidi keskpunktiks on valimi keskvaartusvektor, ellipsoidi kuju (pooltelgede pikkus ja suund) on määratud valimi dispersioonimaatriksiga. Erinevate konkreetsete valimite korral ellipsoidi keskpunkt ja kuju mõnevõrra varieeruvad. Usalduspiirkonda kuuluvad kõik

sellised vektori $\vec{\mu}$ hinnangud, mille korral võrratus (10.1) kehtib.

B. Keskvärtusvektori lineaarkombinatsioonide usaldusvahemikud. Olgu tegemist normaaljaotusega üldkogumiga, $\vec{X} \sim N_p(\vec{\mu}, \Sigma)$. Vaatame selle juhusliku vektori lineaarkombinatsiooni $\vec{a}'\vec{X}$. Siis $\vec{a}'\vec{X} \sim N(\vec{a}'\vec{\mu}, \vec{a}'\Sigma\vec{a})$ ja

$$\vec{a}'\vec{X} \sim N(\vec{a}'\vec{\mu}, \frac{1}{n}\vec{a}'\Sigma\vec{a}).$$

Asendades dispersioonimaatriksi Σ statistikuga $\mathbf{S}$ saame konstrueerida t -jaotusega juhusliku suuruse

$$\sqrt{n} \frac{\vec{a}'\vec{X} - \vec{a}'\vec{\mu}}{\sqrt{\vec{a}'\mathbf{S}\vec{a}}} \sim t(n-1).$$

Kasutades nüüd t -jaotuse kvantiile saame leida keskvärtusvektori lineaarkombinatsioonile $\vec{a}'\vec{\mu}$ usaldusvahemiku. Tõepoolest, kvantiilide määratlusest järeldub, et

$$P\left(t_{\frac{\alpha}{2}, n-1} \leq \sqrt{n} \frac{\vec{a}'\vec{X} - \vec{a}'\vec{\mu}}{\sqrt{\vec{a}'\mathbf{S}\vec{a}}} \leq \bar{t}_{\frac{\alpha}{2}, n-1}\right) = 1 - \alpha,$$

kust tuleneb

$$P(\vec{a}'\vec{X} - \bar{t}_{\frac{\alpha}{2}, n-1} \frac{\sqrt{\vec{a}'\mathbf{S}\vec{a}}}{\sqrt{n}} \leq \vec{a}'\vec{\mu} \leq \vec{a}'\vec{X} + \bar{t}_{\frac{\alpha}{2}, n-1} \frac{\sqrt{\vec{a}'\mathbf{S}\vec{a}}}{\sqrt{n}}) = 1 - \alpha.$$

Niisugune ühe lineaarkombinatsiooni uurimine annab üpris piiratud infot keskvärtusvektori kohta. Hoopis olulisem on leida usalduspiirkond mitme lineaarkombinatsiooni jaoks üheaegselt. Püüame seda teha. Võtame vaatluse alla r keskvärtusvektori lineaarkombinatsiooni $\vec{a}'_1\vec{\mu}, \vec{a}'_2\vec{\mu}, \dots, \vec{a}'_r\vec{\mu}$. Võttes kasutusele $p \times r$ maatriksi $\mathbf{A}$, mille veeruvektoriteks on lineaarkombinatsioonide kordajad,

$$\mathbf{A} = [\vec{a}_1 \ \vec{a}_2 \ \dots \ \vec{a}_r],$$

saame moodustada huvipakkuvate lineaarkombinatsioonide vektori $\vec{Y}$,

$$\vec{Y} = \mathbf{A}'\vec{X}.$$

Kuna mitmemõõtmeliste normaaljaotuste klass on kinnine lineaarteisenduse suhtes, on selle vektori jaotuseks r -mõõtmeline normaaljaotus, $\vec{Y} \sim N_r(\mathbf{A}'\vec{\mu}, \mathbf{A}'\Sigma\mathbf{A})$. Tähistades

$$\mathbf{A}'\vec{\mu} = \vec{\nu}, \quad \vec{Y} = \mathbf{A}'\vec{X} \quad \text{ja} \quad \mathbf{A}'\Sigma\mathbf{A} = \mathbf{S}_{\vec{Y}}$$

ning eeldades, et lineaarkombinatsioonide kordajad $\vec{a}_i$ on lineaarselt sõltumatud (sel juhul on maatriks $\mathbf{A}$ täisastakuga), saame välja kirjutada usaldusellipsoidi usaldusnivooga $1 - \alpha$ vektori $\vec{\nu}$ jaoks,

$$(\vec{Y} - \vec{\nu})'(\mathbf{S}_{\vec{Y}})^{-1}(\vec{Y} - \vec{\nu}) \leq \frac{(n-1)r}{(n-r)n} \bar{f}_{\alpha, r, n-r}. \quad (10.2)$$

Usalduspiirkonna abil esitatud informatsioon parameetrite kohta on lihtsamini mõistetav, kui usalduspiirkonnaks on teatud "tahukad", milles iga koordinaadi jaoks konstrueeritakse usaldusvahemik. Niisuguse "tahuka" konstrueerimiseks vajame tõket lineaarkombinatsioonide maksimaalsele väärtusele. Selle leidmiseks kasutame järgmist maatriksalgebra tulemust.

Lemma 10.1. *Olgu $\mathbf{V}$ positiivselt määratud $p \times p$ -maatriks ja $\mathbf{A}$ täisastakuga $p \times r$ -maatriks, $r(\mathbf{A}) = r$, ($r \leq p$) veeruvektoritega $\vec{a}_i$, $i = 1, 2, \dots, r$. Siis kehtivad võrratused*

$$\frac{(\vec{a}_i' \vec{x})^2}{\vec{a}_i' \mathbf{V} \vec{a}_i} \leq \vec{x}' \mathbf{A} (\mathbf{A}' \mathbf{V} \mathbf{A})^{-1} \mathbf{A}' \vec{x} \leq \vec{x}' \mathbf{V}^{-1} \vec{x} \quad (10.3)$$

iga $\vec{x} \in \mathbb{R}^p$ korral.

Tõestus. Näitame esmalt teise võrratuse kehtivuse. Tähistame $\vec{y} = \mathbf{V}^{-\frac{1}{2}}\vec{x}$, kust $\vec{x} = \mathbf{V}^{\frac{1}{2}}\vec{y}$. (Siin $\mathbf{V}^{\frac{1}{2}}$ tähistab sümmeetrilist ruutjuurt maatriksist $\mathbf{V}$). Uutes tähistustes omandab võrratus kuju:

$$\vec{y}'\mathbf{V}^{\frac{1}{2}}\mathbf{A}(\mathbf{A}'\mathbf{V}\mathbf{A})^{-1}\mathbf{A}'\mathbf{V}^{\frac{1}{2}}\vec{y} \leq \vec{y}'\vec{y}.$$

Tähistame

$$\mathbf{V}^{\frac{1}{2}}\mathbf{A}(\mathbf{A}'\mathbf{V}\mathbf{A})^{-1}\mathbf{A}'\mathbf{V}^{\frac{1}{2}} = \mathbf{B}.$$

Rakendades Teoreemi L1.3 peab Rayleigh suhte kohaselt kehtima võrratus

$$\frac{\vec{y}'\mathbf{B}\vec{y}}{\vec{y}'\vec{y}} \leq \lambda_{\max}(\mathbf{B}),$$

kus $\lambda_{\max}(\mathbf{B})$ on maatriksi $\mathbf{B}$ maksimaalne omaväärtus. Näitame, et $\lambda_{\max}(\mathbf{B}) = 1$, sellest järeldub ka tõestatav teine võrratus seoses (10.3). Omaväärtused rahuldavad karakteristlikku võrrandit

$$|\mathbf{B} - \lambda\mathbf{I}_p| = |\mathbf{V}^{\frac{1}{2}}\mathbf{A}(\mathbf{A}'\mathbf{V}\mathbf{A})^{-1}\mathbf{A}'\mathbf{V}^{\frac{1}{2}} - \lambda\mathbf{I}_p| = 0.$$

Determinandi omaduse (L1.6) tõttu

$$|\mathbf{B} - \lambda\mathbf{I}_p| = |\mathbf{A}'\mathbf{V}^{\frac{1}{2}}\mathbf{V}^{\frac{1}{2}}\mathbf{A}(\mathbf{A}'\mathbf{V}\mathbf{A})^{-1} - \lambda\mathbf{I}_r| = |\mathbf{I}_r - \lambda\mathbf{I}_r| = 0.$$

Seega $\mathbf{B}$ kõik mittenullilised omaväärtused võrduvad ühega ja teine võrratus on tõestatud. Võrratusteahela (10.3) esimese võrratuse tõestamiseks näitame esmalt, et võrratus kehtib juhul $i = 1$. Kasutame tähistusi $\vec{z} = \mathbf{A}'\vec{x} = (z_1, \vec{z}_2)'$ ja $\mathbf{C} = \mathbf{A}'\mathbf{V}\mathbf{A}$, mille esitame plokkmaatriksina

$$\mathbf{C} = \begin{bmatrix} c_{11} & \vec{c}_{21}' \\ \vec{c}_{21} & \mathbf{C}_{22} \end{bmatrix}.$$

Maatriksi $\mathbf{C}$ pöördmaatriksi saame esitada kujul (teine esitus teoreemis L1.7):

$$\mathbf{C}^{-1} = \begin{bmatrix} c_{11}^{-1} + c_{11}^{-2}\vec{c}_{21}'\mathbf{C}_{22.1}^{-1}\vec{c}_{21} & -c_{11}^{-1}\vec{c}_{21}'\mathbf{C}_{22.1}^{-1} \\ -c_{11}^{-1}\mathbf{C}_{22.1}^{-1}\vec{c}_{21} & \mathbf{C}_{22.1}^{-1} \end{bmatrix},$$

kus Schuri täiend $\mathbf{C}_{22.1} = \mathbf{C}_{22} - \vec{c}_{21}c_{11}^{-1}\vec{c}_{21}'$. Siis

$$\vec{z}'\mathbf{C}^{-1}\vec{z} = \frac{z_1^2}{c_{11}} + \vec{d}'\vec{d} \geq \frac{z_1^2}{c_{11}} = \frac{(\vec{a}_1'\vec{x})^2}{\vec{a}_1'\mathbf{V}\vec{a}_1},$$

kus

$$\vec{d} = z_1c_{11}^{-1}\vec{c}_{21}'\mathbf{C}_{22.1}^{-\frac{1}{2}} - \vec{z}_2'\mathbf{C}_{22.1}^{-\frac{1}{2}}.$$

Esituse $\vec{d}'\vec{d}$ saame maatriksite vahetu läbikorrutamise teel ruutvormis $\vec{z}'\mathbf{C}^{-1}\vec{z}$. Tulemusena oleme tõestanud väite juhul $i = 1$.

Tõestamiseks väidet suvalise i korral vahetame maatriksis $\mathbf{A}$ esimese ja i -nda veeru ja va- ja on korrata tõestust esimese veeru jaoks teisendatud maatriksi $\mathbf{A}\mathbf{E}_{1i}$ korral, kus $\mathbf{E}_{1i}$ on elementaarmaatriks, mis realiseerib paremalt korrutades veergude vahetamise ja on saadud ühikmaatriksist selle esimese ja i -nda rea ja veeru vahetamise teel. Maatriks $\mathbf{E}_{1i}$ on sümmeetriline ja vahetu kontroll näitab, et ta on iseenda pöördmaatriksiks.

Paneme aga tähele, et teisendatud maatriksi $\mathbf{A}\mathbf{E}_{1i}$ korral kehtib võrdus

$$\mathbf{A}\mathbf{E}_{1i}((\mathbf{A}\mathbf{E}_{1i})'\mathbf{V}\mathbf{A}\mathbf{E}_{1i})^{-1}(\mathbf{A}\mathbf{E}_{1i})' = \mathbf{A}(\mathbf{A}'\mathbf{V}\mathbf{A})^{-1}\mathbf{A}'.$$

Seega ruutvorm tõestatava võrratuse paremal poolel ei muutu maatriksi $\mathbf{A}$ veergude vahetamisega ja tõestuskäiku korrates saame esimese võrratuse ahelas (10.3) ka suvalise vektori $\vec{a}_i$ korral, mott.

Valime nüüd võrratuste ahelas (10.3) $\vec{x} = \vec{X} - \vec{\mu}$, $\mathbf{V} = \frac{1}{n}\mathbf{S}$ ja olgu $\mathbf{A}$ vektori $\vec{X}$ lineaarkombinatsioone määrav maatriks. Siis kasutades tähistust $\vec{Y} = \mathbf{A}'\vec{X}$, saame

$$\begin{aligned} n \frac{(\bar{Y}_i - \nu_i)^2}{\vec{a}_i'\mathbf{S}\vec{a}_i} &\leq n(\vec{X} - \vec{\mu})'\mathbf{A}(\mathbf{S}_{\vec{Y}})^{-1}\mathbf{A}'(\vec{X} - \vec{\mu}) \\ &= n(\vec{Y} - \vec{\nu})'(\mathbf{S}_{\vec{Y}})^{-1}(\vec{Y} - \vec{\nu}) \\ &\leq n(\vec{X} - \vec{\mu})'\mathbf{S}^{-1}(\vec{X} - \vec{\mu}). \end{aligned}$$

Võrratuse (10.2) kohaselt saame vektori $\vec{\nu}$ jaoks usaldusellipsoidi usaldusnivool $1 - \alpha$:

$$P\left(n(\vec{Y} - \vec{\nu})'(\mathbf{S}_{\vec{Y}})^{-1}(\vec{Y} - \vec{\nu}) \leq \frac{(n-1)r}{n-r} \bar{f}_{\alpha, r, n-r}\right) = 1 - \alpha.$$

Maksimaalse lineaarkombinatsiooni tõkestatusest tuleneb, et kõik lineaarkombinatsioonid on sama tõkkega tõkestatud tõenäosusega, mis ei ole väiksem kui $1 - \alpha$. Tähistame

$$k_\alpha = \frac{(n-1)r}{n-r} \bar{f}_{\alpha, r, n-r},$$

siis

$$P\left(n \frac{(\bar{Y}_i - \nu_i)^2}{\vec{a}_i' \mathbf{S} \vec{a}_i} \leq k_\alpha\right) = P\left(\sqrt{n} \frac{|\bar{Y}_i - \nu_i|}{\sqrt{\vec{a}_i' \mathbf{S} \vec{a}_i}} \leq \sqrt{k_\alpha}\right) \geq 1 - \alpha.$$

Seega kehtivad vähemalt tõenäosusega $1 - \alpha$ samaaegselt võrratused

$$\bar{Y}_i - \sqrt{\frac{k_\alpha}{n}} \sqrt{\vec{a}_i' \mathbf{S} \vec{a}_i} \leq \nu_i \leq \bar{Y}_i + \sqrt{\frac{k_\alpha}{n}} \sqrt{\vec{a}_i' \mathbf{S} \vec{a}_i},$$

$i = 1, 2, \dots, r$. Saadud üheaegseid usaldusvahemikke nimetatakse *Roy-Bose üheaegseiks usaldusvahemikeks*.

Analoogilise võttega saab leida üheaegsed usaldusvahemikud keskväärtusvektori kõikvõimalike lineaarkombinatsioonide jaoks. Sel juhul huvitab meid selline konstant c^2 , et

$$\begin{aligned} P\left(\forall \vec{a} : n \frac{(\vec{a}'(\vec{X} - \vec{\mu}))^2}{\vec{a}' \mathbf{S} \vec{a}} \leq c^2\right) \\ = P\left(\max_{\vec{a}} \left[n \frac{(\vec{a}'(\vec{X} - \vec{\mu}))^2}{\vec{a}' \mathbf{S} \vec{a}}\right] \leq c^2\right) = 1 - \alpha. \end{aligned}$$

Kasutame jälle ära võrratused (10.3). Lemmas toodud võrratused kehtivad ka selle lineaarkombinatsiooni korral, mis realiseerib maksimumi

$$\max_{\vec{a}} \left[n \frac{(\vec{a}'(\vec{X} - \vec{\mu}))^2}{\vec{a}' \mathbf{S} \vec{a}}\right] \leq n(\vec{X} - \vec{\mu})' \mathbf{S}^{-1} (\vec{X} - \vec{\mu}).$$

Osutub, et maksimumi korral asendub võrratus võrdusega

$$\max_{\vec{a}} \left[n \frac{(\vec{a}'(\vec{X} - \vec{\mu}))^2}{\vec{a}' \mathbf{S} \vec{a}}\right] = n(\vec{X} - \vec{\mu})' \mathbf{S}^{-1} (\vec{X} - \vec{\mu}),$$

seega

$$\max_{\vec{a}} \left[n \frac{(\vec{a}'(\vec{X} - \vec{\mu}))^2}{\vec{a}' \mathbf{S} \vec{a}}\right] \sim T^2(p, n-1).$$

Siis aga saame välja kirjutada suvaliste lineaarkombinatsioonide jaoks samaaegselt kehtivad usaldusvahemikud

$$\vec{a}' \vec{X} - \sqrt{\frac{c_\alpha}{n}} \sqrt{\vec{a}' \mathbf{S} \vec{a}} \leq \vec{a}' \vec{\mu} \leq \vec{a}' \vec{X} + \sqrt{\frac{c_\alpha}{n}} \sqrt{\vec{a}' \mathbf{S} \vec{a}},$$

kus $c_\alpha = \frac{p(n-1)}{n-p} \bar{f}_{\alpha, p, n-p}$.

Saadud usaldusvahemikke nimetatakse *Scheffe üheaegseiks usaldusvahemikeks*. Selge on see, et Scheffe vahemikud on laiemad kui Roy-Bose omad, kuna viimaste korral on samaaegselt täidetud r tingimust, Scheffe vahemikud aga antakse üheaegselt kõikvõimalike lineaarkombinatsioonide jaoks.

Võib küsida, kas on võimalik konstrueerida täpsemaid usaldusvahemikke kui Roy-Bose omad, kui kitsendusi on kindel arv? Saab, kitsamad võivad olla vahemikud, mis kannavad *Bonferoni usaldusvahemike* nime. Olgu vaatluse all r keskväärtusvektori lineaarkombinatsiooni $\bar{Y}_i = \vec{a}_i' \vec{X}$. Tähistame

$$T_i = \sqrt{n} \frac{\bar{Y}_i - \nu_i}{\sqrt{\vec{a}_i' \mathbf{S}_{\vec{Y}} \vec{a}_i}}, \quad i = 1, 2, \dots, r.$$

Teame, et $T_i \sim t(n-1)$, kuid paraku on need juhuslikud suurused sõltuvad. Defineerime sündmused

$$A_i = \{\omega : |T_i(\omega)| \leq t_0\}.$$

Siis

$$P\left(\bigcap_{i=1}^r A_i\right) = 1 - P\left(\bigcup_{i=1}^r A_i^c\right) \geq 1 - \sum_{i=1}^r P(A_i^c) = 1 - rP(|T_i| > t_0).$$

Vahepealne võrratus

$$1 - P\left(\bigcup_{i=1}^r A_i^c\right) \geq 1 - \sum_{i=1}^r P(A_i^c)$$

on tuntud *Bonferroni võrratusena*. Kuna otsime usalduspiirkonda usaldusnivooga $1 - \alpha$, nõuame et t_0 oleks valitud nii, et

$$1 - rP(|T_i| > t_0) = 1 - \alpha.$$

Siit saame

$$P(|T_i| > t_0) = \frac{\alpha}{r} \Rightarrow P(T_i > t_0) = \frac{\alpha}{2r},$$

kust

$$t_0 = \bar{t}_{\frac{\alpha}{2r}, n-1}.$$

Seega saame Bonferroni üheaegse usaldusvahemiku r lineaarkombinatsiooni jaoks esitada kujul

$$\vec{a}'\vec{X} - \frac{\bar{t}_{\frac{\alpha}{2r}, n-1}}{\sqrt{n}} \sqrt{\vec{a}'\mathbf{S}\vec{a}} \leq \vec{a}'\vec{\mu} \leq \vec{a}'\vec{X} + \frac{\bar{t}_{\frac{\alpha}{2r}, n-1}}{\sqrt{n}} \sqrt{\vec{a}'\mathbf{S}\vec{a}}.$$

Milliseid usaldusvahemikke kasutada – kas Roy-Bose või Bonferroni omi – peab selguma suhtest $\bar{t}_{\frac{\alpha}{2r}, n-1}/k_\alpha$. Loomulikult tuleb eelistada kitsamaid usaldusintervalle. Kui vaatluste arv on suur, pole erinevate üheaegsete usaldusintervallide ulatus oluliselt erinev. Suurte valimimahtude korral saame kasutada ligikaudsete usalduspiiride leidmiseks asümptootilist χ^2 -jaotust:

$$n(\vec{X} - \vec{\mu})'\mathbf{S}^{-1}(\vec{X} - \vec{\mu}) \xrightarrow{D} \chi^2(p).$$

Selle abil saame suvalise lineaarkombinatsiooni korral asendada F -jaotuse täiendkvantiili χ^2 -jaotuse kvantiili ja täiendkvantiiliga:

$$\vec{a}'\vec{X} - \sqrt{\frac{\bar{h}_{\frac{\alpha}{2}, p}}{n}} \sqrt{\vec{a}'\mathbf{S}\vec{a}} \leq \vec{a}'\vec{\mu} \leq \vec{a}'\vec{X} + \sqrt{\frac{\bar{h}_{\frac{\alpha}{2}, p}}{n}} \sqrt{\vec{a}'\mathbf{S}\vec{a}}.$$

Suurte valimimahtude korral on nii saadud usalduspiirid väga lähedased Scheffe üheaegsete usalduspiiridega.

Ülesanded 10

10.1. Teha läbi detailne tuletuskäik lemma 10.1 tõestuses esineva võrduse

$$\vec{z}'\mathbf{C}^{-1}\vec{z} = \frac{z_1^2}{c_{11}} + \vec{d}'\vec{d},$$

saamiseks, kus

$$\vec{d} = z_1 c_{11}^{-1} \vec{c}_{21}' \mathbf{C}_{22 \cdot 1}^{-\frac{1}{2}} - \vec{z}_2' \mathbf{C}_{22 \cdot 1}^{-\frac{1}{2}}.$$

11 Hüpoteesid dispersioonimaatriksi kohta

A. Fikseeritud dispersioonimaatriks. Vaatame normaaljaotusega üldkogumit, $\vec{X} \sim N(\vec{\mu}, \Sigma)$. Olgu Σ_0 (sümmeetriline, positiivselt määratud maatriks) dispersioonimaatriksi oletatav väärtus. Esitame hüpoteesid

$$\begin{aligned} H_0 : \quad & \Sigma = \Sigma_0 \\ H_1 : \quad & \Sigma \neq \Sigma_0. \end{aligned}$$

Tuletame nende hüpoteeside kontrollimiseks tõepärasuhte statistiku

$$\Lambda(\mathbf{X}) = \frac{\max_{\vec{\mu}} L(\mathbf{X}; \vec{\mu}, \Sigma_0)}{\max_{\vec{\mu}, \Sigma} L(\mathbf{X}; \vec{\mu}, \Sigma)}.$$

Kasutame lugejas oleva maksimumi leidmiseks tõepärafunktsiooni kuju (5.1) kohal $\Sigma = \Sigma_0$ ning asendame valemis (5.2) konkreetse valimi ja sellest leitud hinnangud teoreetilise valimi ning vastavate statistikutega:

$$L(\mathbf{X}; \vec{\mu}, \Sigma_0) = \frac{1}{(2\pi)^{\frac{np}{2}} |\Sigma_0|^{\frac{n}{2}}} \exp \left\{ -\frac{n}{2} \left(\text{tr}(\Sigma_0^{-1} \mathbf{S}_*) + (\vec{X} - \vec{\mu})' \Sigma_0^{-1} (\vec{X} - \vec{\mu}) \right) \right\}.$$

Tõepärafunktsioon on maksimaalne, kui eksponentfunktsiooni astendaja on minimaalne, astendaja väärtusi saame muuta ainult teise liidetava abil. Kuna teine liidetav on mittenegeatiivne ruutvorm, on tema vähimaks väärtuseks null, mille saavutamiseks tuleb valida $\hat{\vec{\mu}} = \vec{x}$. Seega

$$\max_{\vec{\mu}} L(\mathbf{X}; \vec{\mu}, \Sigma_0) = \frac{1}{(2\pi)^{\frac{np}{2}} |\Sigma_0|^{\frac{n}{2}}} \exp \left\{ -\frac{n}{2} \text{tr}(\Sigma_0^{-1} \mathbf{S}_*) \right\}.$$

Nimetajas oleva maksimumi leidsime üheksandas paragrahvis, valem (9.2),

$$\max_{\vec{\mu}, \Sigma} L(\mathbf{X}; \vec{\mu}, \Sigma) = L(\mathbf{X}; \vec{X}, \mathbf{S}_*) = \frac{1}{(2\pi)^{\frac{np}{2}} |\mathbf{S}_*|^{\frac{n}{2}}} e^{-\frac{np}{2}}.$$

Seega TPS statistik esitatud hüpoteeside kontrollimiseks omab kuju

$$\Lambda(\mathbf{X}) = |\Sigma_0^{-1} \mathbf{S}_*|^{\frac{n}{2}} \exp \left\{ -\frac{n}{2} \text{tr}(\Sigma_0^{-1} \mathbf{S}_*) + \frac{np}{2} \right\}.$$

Selle statistiku jaotus ei ole teada. Küll on aga teada TPS statistiku käitumine juhul, kui valimimaht kasvab piiramatult, $n \rightarrow \infty$. Nimelt kehtib regulaarsete jaotuste korral üldine tulemus – kui H_0 kehtib, siis

$$-2 \ln \Lambda(\mathbf{X}) \xrightarrow{\mathcal{D}} \chi^2(k),$$

hii-ruut-jaotuse vabadusastmete arv $k = q - v$, kus q on kõikide parameetrite arv selles ülesandes, v aga vabade parameetrite arv H_0 kehtimisel.

Meie TPS statistiku korral

$$\ln \Lambda(\mathbf{X}) = \frac{n}{2} \ln |\Sigma_0^{-1} \mathbf{S}_*| - \frac{n}{2} (\text{tr}(\Sigma_0^{-1} \mathbf{S}_*) - p),$$

kust

$$-2 \ln \Lambda(\mathbf{X}) = n (\text{tr}(\Sigma_0^{-1} \mathbf{S}_*) - p) - n \ln |\Sigma_0^{-1} \mathbf{S}_*|.$$

Tähistame $\mathbf{V} = \Sigma_0^{-1} \mathbf{S}_*$, olgu selle maatriksi omaväärtused λ_i , $i = 1, 2, \dots, p$. Teatavasti $|\mathbf{V}| = \prod_{i=1}^p \lambda_i$ ja $\text{tr} \mathbf{V} = \sum_{i=1}^p \lambda_i$. Tähistame

$$\bar{\lambda} = \frac{1}{p} \sum_{i=1}^p \lambda_i \quad \text{ja} \quad g = p \sqrt{\prod_{i=1}^p \lambda_i}.$$

Siis

$$-2 \ln \Lambda(\mathbf{X}) = np(\bar{\lambda} - \ln g - 1).$$

Leiame TPS statistiku asümptootilise jaotuse vabadusastmete arvu. Keskvärtusvektoril on p koordinaati, dispersioonimaatriksil $\frac{p(p+1)}{2}$ erinevat elementi. Kokku on parameetreid $p + \frac{p(p+1)}{2}$. Nullhüpoteesi kehtides on vabu parameetreid p , kuna Σ_0 on etteantud määratud maatriks. Seega on asümptootilise jaotuse vabadusastmete arv $\frac{p(p+1)}{2}$. TPS statistiku kasutamine on standardne – kui $\Lambda(\mathbf{X}) \leq c \Rightarrow H_1$. Kui kasutame statistikut $-2 \ln \Lambda(\mathbf{X})$, muutub võrratus vastupidiseks. Kui oleme valinud olulisuse nivoo α , saame valida kriitilise piirkonna alumiseks otspunktiks χ^2 -jaotuse α -täiendkvantiili $\bar{h}_{\alpha, \frac{p(p+1)}{2}}$. Seega oleme jõudnud otsustuseeskirjani

$$\text{kui } np(\bar{\lambda} - \ln g - 1) \geq \bar{h}_{\alpha, \frac{p(p+1)}{2}} \Rightarrow H_1,$$

$$\text{kui } np(\bar{\lambda} - \ln g - 1) < \bar{h}_{\alpha, \frac{p(p+1)}{2}} \Rightarrow H_0.$$

B. Sfäärilisuse kontrollimine. Tegelikult on harva võimalik esitada hüpotees dispersiooni-maatriksi täpse kuju kohta. Enamasti tahetakse ikka kontrollida, kas Σ on teatud lihtsama struktuuriga maatriks (näiteks diagonaalmaatriks või plokk-diagonaalmaatriks). Vaatame järgmisena *sfäärilisuse testi*. Sel juhul esitatakse kontrollitavad hüpoteesid kujul

$$\begin{aligned} H_0 : \quad \Sigma &= \gamma \mathbf{I}_p, & (\gamma > 0), \\ H_1 : \quad \Sigma &\neq \gamma \mathbf{I}_p. \end{aligned}$$

Tõepärasuhte statistik omab siis kuju

$$\Lambda(\mathbf{X}) = \frac{\max_{\vec{\mu}, \gamma} L(\mathbf{X}; \vec{\mu}, \gamma \mathbf{I}_p)}{\max_{\vec{\mu}, \Sigma} L(\mathbf{X}; \vec{\mu}, \Sigma)} = \frac{\max_{\vec{\mu}, \gamma} |\gamma \mathbf{I}_p|^{-\frac{n}{2}} e^{-\frac{n}{2} \text{tr}(\gamma^{-1} \mathbf{S}_*)}}{|\mathbf{S}_*|^{-\frac{n}{2}} e^{-\frac{np}{2}}}.$$

Leiame lugeja maksimumi, alustades lugejas oleva avaldise logaritmisest,

$$\ln L_0(\mathbf{X}, \vec{\mu}, \gamma) = -\frac{np}{2} \ln \gamma - \frac{n}{2\gamma} \text{tr} \mathbf{S}_*.$$

Arvutame logaritmi tuletise γ järgi ja võrdsustame selle nulliga,

$$\frac{\partial \ln L_0(\mathbf{X}, \mu, \gamma)}{\partial \gamma} = -\frac{np}{2} \frac{1}{\gamma} + \frac{n}{2\gamma^2} \text{tr} \mathbf{S}_* = 0,$$

kust saame

$$\hat{\gamma} = \frac{1}{p} \text{tr} \mathbf{S}_*.$$

Näeme, et eksponentfunktsiooni astendajad lugejas ja nimetajas osutuvad võrdseteks, saame eksponentfunktsioonid taandada ja TPS statistik omandab kuju

$$\Lambda(\mathbf{X}) = \left(\frac{|\mathbf{S}_*|}{(\frac{1}{p} \text{tr} \mathbf{S}_*)^p} \right)^{\frac{n}{2}}.$$

Olgu $\nu_1, \nu_2, \dots, \nu_p$, maatriksi $\mathbf{S}_*$ omaväärtused. Tähistades $\tilde{\Lambda}(\mathbf{X}) = (\Lambda(\mathbf{X}))^{2/n}$, võime selle statistiku esitada omaväärtuste geomeetrilise keskmise ja aritmeetilise keskmise suhtena,

$$\tilde{\Lambda}(\mathbf{X}) = \left(\frac{\prod_{i=1}^p \nu_i^{1/p}}{\frac{1}{p} \sum_{i=1}^p \nu_i} \right)^p$$

Kui H_0 kehtib, on oodata statistiku $\tilde{\Lambda}(\mathbf{X})$ ühele lähedasi väärtusi. Seega kriitiline piirkond omab kuju $K_\alpha = (0, c_{kr})$. Osutub, et statistiku $\tilde{\Lambda}(\mathbf{X})$ jaoks saab leida täpse jaotuse ja selle abil esimeste momentide avaldised. Kehtib järgmine teoreem (Bilodeau and Brenner (1999), lk.119).

Teoreem 11.1. Olgu $\vec{X} \sim N_p(\vec{\mu}, \Sigma)$. Siis statistiku $\tilde{\Lambda}(\mathbf{X})$ jaotus on määratud järgmise stohhastilise esitusega:

$$\tilde{\Lambda}(\mathbf{X}) = \prod_{i=1}^{p-1} Y_i,$$

kus Y_i on sõltumatud beeta-jaotusega juhuslikud suurused,

$$Y_i \sim \beta\left(\frac{1}{2}(n-1-i), i\left(\frac{1}{2} + \frac{1}{p}\right)\right).$$

Kahjuks ei ole aga TPS statistiku täpse jaotuse kasutamine mugav. Kriitilise piirkonna määramiseks saame kasutada asümptootilist jaotust,

$$-2 \ln \Lambda(\mathbf{X}) \xrightarrow{\mathcal{D}} \chi^2\left(\frac{p(p+1)}{2} - 1\right),$$

mille korral kriitiline piirkond omab kuju (c_{kr}^*, ∞) , kus $c_{kr}^* = \bar{h}_{\alpha, \frac{p(p+1)}{2} - 1}$.

C. Alamvektorite sõltumatuse kontrollimine. Olgu vaatlusalune tunnusvektor jagatud kaheks alamvektoriks,

$$\vec{X} = \begin{bmatrix} \vec{X}_1 \\ \vec{X}_2 \end{bmatrix},$$

mõõtmetega vastavalt p_1 ja p_2 , $p = p_1 + p_2$. Ka keskvärtusvektor ja dispersioonimaatriks jagunevad siis plokkideks,

$$\vec{\mu} = \begin{bmatrix} \vec{\mu}_1 \\ \vec{\mu}_2 \end{bmatrix} \quad \text{ja} \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix},$$

kus plokk Σ_{12} sisaldab kõikvõimalikke kovariatsioone alamvektorite komponentide vahel. Alamvektorite sõltumatust saame kontrollida järgmiste hüpoteesidega

$$\begin{aligned} H_0 : \Sigma_{12} &= \mathbf{0}, \\ H_1 : \Sigma_{12} &\neq \mathbf{0}. \end{aligned}$$

Leiame jälle tõepärasuhte statistiku. Olgu $\mathcal{S}$ kõigi sümmeetriliste positiivselt määratud maatriksite klass, kus blokk $\Sigma_{12} = \mathbf{0}$. Siis

$$\Lambda(\mathbf{X}) = \frac{\max_{\vec{\mu}; \Sigma \in \mathcal{S}} L(\mathbf{X}; \vec{\mu}, \Sigma)}{\max_{\vec{\mu}, \Sigma} L(\mathbf{X}; \vec{\mu}, \Sigma)}.$$

Nimetaja maksimaalne väärtus saavutatakse STP hinnangute korral,

$$\max_{\vec{\mu}, \Sigma} L(\mathbf{X}; \vec{\mu}, \Sigma) = \frac{1}{(2\pi)^{\frac{np}{2}} |\mathbf{S}_*|^{\frac{n}{2}}} \exp\left\{-\frac{np}{2}\right\}.$$

Lugeja maksimaalne väärtus on aga lihtsalt leitav. Kui dispersioonimaatriks kuulub klassi $\mathcal{S}$, jaguneb tõepärafunktsioon alamvektorite tõepärafunktsioonide korrutiseks,

$$\begin{aligned} L(\mathcal{X}; \vec{\mu}, \Sigma) &= \frac{1}{(2\pi)^{\frac{np}{2}} |\Sigma_{11}|^{\frac{n}{2}} |\Sigma_{22}|^{\frac{n}{2}}} \\ &\times \exp\left\{-\frac{1}{2} \sum_{i=1}^n \left[(\vec{x}_i^{(1)} - \vec{\mu}_1)' \Sigma_{11}^{-1} (\vec{x}_i^{(1)} - \vec{\mu}_1) \right. \right. \\ &\quad \left. \left. + (\vec{x}_i^{(2)} - \vec{\mu}_2)' \Sigma_{22}^{-1} (\vec{x}_i^{(2)} - \vec{\mu}_2) \right] \right\}. \end{aligned}$$

Korrutise maksimiseerimiseks tuleb maksimiseerida mõlemad tegurid. See aga tähendab, et parameetriteks peame valima alamvektorite STP hinnangud,

$$\begin{aligned} \hat{\vec{\mu}}_1 &= \vec{x}_1, \quad \hat{\Sigma}_{11} = \hat{\mathbf{S}}_{*11}, \\ \hat{\vec{\mu}}_2 &= \vec{x}_2, \quad \hat{\Sigma}_{22} = \hat{\mathbf{S}}_{*22}. \end{aligned}$$

Kuna

$$\sum_{i=1}^n (\vec{x}_i^{(1)} - \vec{x}_1)' \hat{\mathbf{S}}_{*11}^{-1} (\vec{x}_i^{(1)} - \vec{x}_1) = n \text{tr}(\hat{\mathbf{S}}_{*11}^{-1} \hat{\mathbf{S}}_{*11}) = np_1$$

ja

$$\sum_{i=1}^n (\vec{x}_i^{(2)} - \vec{x}_2)' \hat{\mathbf{S}}_{*22}^{-1} (\vec{x}_i^{(2)} - \vec{x}_2) = n \text{tr}(\hat{\mathbf{S}}_{*22}^{-1} \hat{\mathbf{S}}_{*22}) = np_2,$$

siis lugeja omandab kuju

$$L(\mathbf{X}; \vec{x}_1, \vec{x}_2, \mathbf{S}_{*11}, \mathbf{S}_{*22}) = \frac{1}{(2\pi)^{\frac{np}{2}} |\mathbf{S}_{*11}|^{\frac{n}{2}} |\mathbf{S}_{*22}|^{\frac{n}{2}}} \exp \left\{ -\frac{np}{2} \right\}.$$

Seega TPS statistik omab kuju

$$\Lambda(\mathbf{X}) = \left(\frac{|\mathbf{S}_*|}{|\mathbf{S}_{*11}| |\mathbf{S}_{*22}|} \right)^{n/2},$$

kust logaritmiline TPS statistik on kujul

$$-2 \ln \Lambda(\mathbf{X}) = -n \ln \frac{|\mathbf{S}_*|}{|\mathbf{S}_{*11}| |\mathbf{S}_{*22}|}.$$

Kasutame nüüd ära plokkmaatriksi determinandi avaldised (L1.13) ja (L1.14)

$$\begin{aligned} |\mathbf{S}_*| &= |\mathbf{S}_{*11}| \cdot |\mathbf{S}_{*22} - \mathbf{S}_{*21} \mathbf{S}_{*11}^{-1} \mathbf{S}_{*12}| \\ &= |\mathbf{S}_{*22}| \cdot |\mathbf{S}_{*11} - \mathbf{S}_{*12} \mathbf{S}_{*22}^{-1} \mathbf{S}_{*21}|, \end{aligned}$$

mis annab logaritmilise TPS statistikule kuju

$$-2 \ln \Lambda(\mathbf{X}) = -n \ln |\mathbf{I}_{p_2} - \mathbf{S}_{22}^{-1} \mathbf{S}_{21} \mathbf{S}_{11}^{-1} \mathbf{S}_{12}|$$

või

$$-2 \ln \Lambda(\mathbf{X}) = -n \ln |\mathbf{I}_{p_1} - \mathbf{S}_{11}^{-1} \mathbf{S}_{12} \mathbf{S}_{22}^{-1} \mathbf{S}_{21}|.$$

Leiame veel selle statistiku asümptootilise χ^2 -jaotuse vabadusastmete arvu. üldse on parameetreid $p + \frac{p(p+1)}{2}$. Kui fikseerida kovariatsioonide plokk nullmaatriksiks, jääb $p_1 + \frac{p_1(p_1+1)}{2} + p_2 + \frac{p_2(p_2+1)}{2}$ vaba parameetrit. Saame

$$q = p + \frac{p(p+1)}{2} - p_1 - \frac{p_1(p_1+1)}{2} - p_2 - \frac{p_2(p_2+1)}{2} = p_1 p_2.$$

Teststatistiku asümptootiliseks jaotuseks on $\chi^2(p_1 p_2)$ -jaotus. Otsuse langetamise reegel on seega järgmine:

$$\begin{aligned} \text{kui} \quad & -n \ln |\mathbf{I}_{p_2} - \mathbf{S}_{22}^{-1} \mathbf{S}_{21} \mathbf{S}_{11}^{-1} \mathbf{S}_{12}| < \bar{h}_{\alpha; p_1 p_2} \Rightarrow H_0, \\ \text{kui} \quad & -n \ln |\mathbf{I}_{p_2} - \mathbf{S}_{22}^{-1} \mathbf{S}_{21} \mathbf{S}_{11}^{-1} \mathbf{S}_{12}| \geq \bar{h}_{\alpha; p_1 p_2} \Rightarrow H_1. \end{aligned}$$

12 Kahe üldkogumi võrdlemine

Vaatame ülesannet, kus kahes üldkogumis uuritakse ühte ja sama tunnusvektorit. Tingimused tunnuste väärtuste kujunemiseks võivad olla erinevad ja seetõttu võivad ka üldkogumite parameetrid erineda. Olgu mõlemad üldkogumid normaalkogumitega, $\vec{X}^{(i)} \sim N_p(\vec{\mu}^{(i)}, \Sigma^{(i)})$, $i = 1, 2$. Teeme veel täiendava eelduse: oletame, et üldkogumite dispersioonimaatriksid on samad ($\Sigma^{(1)} = \Sigma^{(2)} = \Sigma$), st erinevad tingimused ei muuda tunnuste vahelisi seoseid ega tunnuste hajuvust. Järelduste tegemiseks olgu meie käsutuses kummastki üldkogumist üks valim. Kasutame valimite eristamiseks tähistusi $\mathcal{X}^{(1)}$ ja $\mathcal{X}^{(2)}$ (nende elemendid tähistame vastavalt $x_{ij}^{(1)}$ ja $x_{ij}^{(2)}$). Valimimahtusid tähistame sümboolitega n_1 ja n_2 , põhistatistikuid aga valimitega analoogiliselt $\vec{X}^{(1)}$, $\mathbf{S}^{(1)}$ ja $\vec{X}^{(2)}$, $\mathbf{S}^{(2)}$. Eldame, et erinevatest üldkogumitest pärit valimid on sõltumatud ja väikese mahuga.

A. Kahe valimi T^2 -statistik. Vaatame esmalt, kuidas kontrollida keskväärtusvektorite kohta esitatud statistilisi hüpoteese

$$\begin{aligned} H_0 : \vec{\mu}^{(1)} - \vec{\mu}^{(2)} &= \vec{\delta}, \\ H_1 : \vec{\mu}^{(1)} - \vec{\mu}^{(2)} &\neq \vec{\delta}, \end{aligned}$$

kus $\vec{\delta}$ on mingi teadaolev vektor. Loomulikult võiks siin tuletada tõepärasuhte statistiku. Kuid läheme teist teed – lähtume analoogiast ühemõõtmelise juhuga. Hinnanguks keskväärtusvektorite vahele $\vec{\mu}^{(1)} - \vec{\mu}^{(2)}$ on valimite keskväärtusvektorite vahe $\vec{x}^{(1)} - \vec{x}^{(2)}$. Vastava statistiku jaotuseks on normaalkogumite jaotus,

$$\vec{X}^{(1)} - \vec{X}^{(2)} \sim N_p\left(\vec{\mu}^{(1)} - \vec{\mu}^{(2)}, \left(\frac{1}{n_1} + \frac{1}{n_2}\right) \Sigma\right).$$

Dispersioonimaatriksi jaoks on meil olemas kaks hinnangut, $\hat{\mathbf{S}}^{(1)}$ ja $\hat{\mathbf{S}}^{(2)}$. Arvutame nende kaalutud keskmise

$$\hat{\mathbf{S}}_{(2)} = \frac{(n_1 - 1)\hat{\mathbf{S}}^{(1)} + (n_2 - 1)\hat{\mathbf{S}}^{(2)}}{n_1 + n_2 - 2}$$

ja võtame selle üldkogumi parameetri Σ hinnanguks.

Vastav statistik on kujul

$$\mathbf{S}_{(2)} = \frac{(n_1 - 1)\mathbf{S}^{(1)} + (n_2 - 1)\mathbf{S}^{(2)}}{n_1 + n_2 - 2}.$$

Näitasime varem, et ühe valimi korral $(n-1)\mathbf{S} \sim W_p(\Sigma, n-1)$. Rakendades seda meie statistikule saame Wisharti jaotuse aditiivsuse tõttu

$$(n_1 + n_2 - 2)\mathbf{S}_{(2)} \sim W_p(\Sigma, n_1 + n_2 - 2).$$

Siit järeldub, et $E(n_1 + n_2 - 2)\mathbf{S}_{(2)} = (n_1 + n_2 - 2)\Sigma$ ja maatriks $\mathbf{S}_{(2)}$ määrab nihketa hinnangu maatriksile Σ . Esitatud hüpoteeside kontrollimiseks vajaliku teststatistiku leidmise saame taandada järgmisele lausele.

Lause 12.1. Olgu $\vec{X}^{(i)} \sim N_p(\vec{\mu}^{(i)}, \Sigma^{(i)})$, $i = 1, 2$, kusjuures $\Sigma^{(1)} = \Sigma^{(2)} = \Sigma$. Siis

$$\begin{aligned} \frac{n_1 n_2}{n_1 + n_2} (\vec{X}^{(1)} - \vec{X}^{(2)} - (\vec{\mu}^{(1)} - \vec{\mu}^{(2)}))' \mathbf{S}_{(2)}^{-1} (\vec{X}^{(1)} - \vec{X}^{(2)} - (\vec{\mu}^{(1)} - \vec{\mu}^{(2)})) \\ \sim T^2(p, n_1 + n_2 - 2). \end{aligned}$$

Tõestus. Dispersioonimaatriksi Σ saame esitada korrutisena $\Sigma = \mathbf{A}\mathbf{A}'$, kus $|\mathbf{A}| \neq 0$. Järelikult saame vektori $\vec{X}^{(1)} - \vec{X}^{(2)} - (\vec{\mu}^{(1)} - \vec{\mu}^{(2)})$ standardiseerida

$$\vec{Z} = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \mathbf{A}^{-1} (\vec{X}^{(1)} - \vec{X}^{(2)} - (\vec{\mu}^{(1)} - \vec{\mu}^{(2)})) \sim N(\vec{0}, \mathbf{I}_p).$$

Tõepoolest,

$$E\vec{Z} = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \mathbf{A}^{-1} E(\vec{X}^{(1)} - \vec{X}^{(2)} - (\vec{\mu}^{(1)} - \vec{\mu}^{(2)})) = \vec{0},$$

ja

$$\begin{aligned} D\vec{Z} &= \frac{n_1 n_2}{n_1 + n_2} D \left[\mathbf{A}^{-1} (\vec{X}^{(1)} - \vec{X}^{(2)} - (\vec{\mu}^{(1)} - \vec{\mu}^{(2)})) \right] \\ &= \frac{n_1 n_2}{n_1 + n_2} \mathbf{A}^{-1} D \left(\vec{X}^{(1)} - \vec{X}^{(2)} \right) (\mathbf{A}^{-1})' = \mathbf{A}^{-1} \mathbf{\Sigma} (\mathbf{A}^{-1})' = \mathbf{I}_p. \end{aligned}$$

Seega $\vec{Z} \sim N_p(\vec{0}, \mathbf{I}_p)$. Statistiku $\mathbf{S}_{(2)}$ saame siduda standardse Wisharti jaotusega,

$$\mathbf{V} = (n_1 + n_2 - 2) \mathbf{A}^{-1} \mathbf{S}_{(2)} (\mathbf{A}^{-1})' \sim W_p(\mathbf{I}_p, n_1 + n_2 - 2).$$

Aga Hotellingu T^2 -jaotuse definitsiooni põhjal

$$(n_1 + n_2 - 2) \vec{Z}' \mathbf{V}^{-1} \vec{Z} \sim T^2(p, n_1 + n_2 - 2).$$

Asendades $\vec{Z}$ ja $\mathbf{V}$ tagasi ja arvestades, et

$$\mathbf{V}^{-1} = \frac{1}{n_1 + n_2 - 2} \mathbf{A}' \mathbf{S}_{(2)}^{-1} \mathbf{A},$$

saame

$$\begin{aligned} (n_1 + n_2 - 2) \vec{Z}' \mathbf{V}^{-1} \vec{Z} &= \frac{n_1 n_2}{n_1 + n_2} (\vec{X}^{(1)} - \vec{X}^{(2)} - (\vec{\mu}^{(1)} - \vec{\mu}^{(2)}))' \\ &\quad \times \mathbf{S}_{(2)}^{-1} (\vec{X}^{(1)} - \vec{X}^{(2)} - (\vec{\mu}^{(1)} - \vec{\mu}^{(2)})). \end{aligned}$$

Järelikult

$$\begin{aligned} \frac{n_1 n_2}{n_1 + n_2} (\vec{X}^{(1)} - \vec{X}^{(2)} - (\vec{\mu}^{(1)} - \vec{\mu}^{(2)}))' \mathbf{S}_{(2)}^{-1} (\vec{X}^{(1)} - \vec{X}^{(2)} - (\vec{\mu}^{(1)} - \vec{\mu}^{(2)})) \\ \sim T^2(p, n_1 + n_2 - 2), \end{aligned}$$

mott.

Lauses esitatud statistikut nimetatakse *kahe valimi T^2 -statistikuks*, tähistame ta $T_{(2)}^2(\mathbf{X})$. Arvestades T^2 -jaotuse ja F -jaotuse vahelist seost (8.4), saame

$$\frac{n_1 + n_2 - p - 1}{p(n_1 + n_2 - 2)} T_{(2)}^2(\mathbf{X}) \sim F(p, n_1 + n_2 - p - 1).$$

Nüüd on meil olemas statistik, mida saame kasutada eespool esitatud hüpoteeside kontrollimiseks:

$$T_{(2)}^2(\mathbf{X}) = \frac{n_1 n_2}{n_1 + n_2} (\vec{X}^{(1)} - \vec{X}^{(2)} - \vec{\delta})' \mathbf{S}_{(2)}^{-1} (\vec{X}^{(1)} - \vec{X}^{(2)} - \vec{\delta}).$$

Tähistame lühiduse mõttes $n = n_1 + n_2$. Kui nullhüpotees kehtib, siis on

$$\frac{n - p - 1}{p(n - 2)} T_{(2)}^2(\mathbf{X}) \sim F(p, n - p - 1).$$

Otsuse langetamiseks saame järgmise eeskirja

$$\begin{aligned} \text{kui } \frac{n - p - 1}{p(n - 2)} T_{(2)}^2(\mathcal{X}) &\geq \bar{f}_{\alpha; p, n - p - 1} \Rightarrow H_1, \\ \text{kui } \frac{n - p - 1}{p(n - 2)} T_{(2)}^2(\mathcal{X}) &< \bar{f}_{\alpha; p, n - p - 1} \Rightarrow H_0. \end{aligned}$$

B. Keskväertusvektorite vahe usalduspiirkond. Keskväertusvektorite vahe usalduspiirkonna leidmiseks võime kasutada samu võtteid, mis olid kasutusel üldkogumi keskväertusvektori jaoks. Scheffe ühiste usalduspiiride leidmisel, mis sisaldavad kõiki keskväertusvektorite vahe lineaarkombinatsioone, lähtume nõudest

$$P \left(\forall \vec{a} : \frac{n_1 n_2}{n_1 + n_2} \frac{[\vec{a}' (\vec{X}^{(1)} - \vec{X}^{(2)} - (\vec{\mu}^{(1)} - \vec{\mu}^{(2)}))]^2}{\vec{a}' \mathbf{S}_{(2)} \vec{a}} \leq c_{kr} \right) \geq 1 - \alpha.$$

Usalduspiiride määramiseks tuleb määrata konstant c_{kr} . Nagu teame, kehtib võrratus iga $\vec{a}$ korral etteantud tõenäosusega

$$\max_{\vec{a}} \frac{[\vec{a}'(\vec{X}^{(1)} - \vec{X}^{(2)} - (\vec{\mu}^{(1)} - \vec{\mu}^{(2)}))]^2}{\vec{a}'\mathbf{S}_{(2)}\vec{a}} \leq c_{kr} \frac{n_1 + n_2}{n_1 n_2}.$$

Lähtume võrratusest, mis kehtib iga $\vec{a}$ ja $\vec{b}$ korral, kui vaid maatriks $\mathbf{S}_{(2)}$ on positiivselt määratud (vt. lemma 10.1)

$$\frac{(\vec{a}'\vec{b})^2}{\vec{a}'\mathbf{S}_{(2)}\vec{a}} \leq \vec{b}'\mathbf{S}_{(2)}^{-1}\vec{b} \quad \text{ehk} \quad (\vec{a}'\vec{b})^2 \leq (\vec{b}'\mathbf{S}_{(2)}^{-1}\vec{b})(\vec{a}'\mathbf{S}_{(2)}\vec{a}).$$

Siit järeldub, et võttes $\vec{b} = \vec{X}^{(1)} - \vec{X}^{(2)} - (\vec{\mu}^{(1)} - \vec{\mu}^{(2)})$ saame:

$$\begin{aligned} \max_{\vec{a}} \frac{[\vec{a}'(\vec{X}^{(1)} - \vec{X}^{(2)} - (\vec{\mu}^{(1)} - \vec{\mu}^{(2)}))]^2}{\vec{a}'\mathbf{S}_{(2)}\vec{a}} \\ = (\vec{X}^{(1)} - \vec{X}^{(2)} - (\vec{\mu}^{(1)} - \vec{\mu}^{(2)}))'\mathbf{S}_{(2)}^{-1}(\vec{X}^{(1)} - \vec{X}^{(2)} - (\vec{\mu}^{(1)} - \vec{\mu}^{(2)})) \\ = \frac{n_1 + n_2}{n_1 n_2} T_{(2)}^2, \end{aligned}$$

kusjuures võtsime arvesse, et võrratuse asemel kehtib maksimumi korral võrdus. Näeme, et maksimumi määrab kahe valimi T^2 -statistik. Esialgne nõue usalduspiirkonnale saab nüüd kuju:

$$P(T_{(2)}^2 \leq c_{kr}) \geq 1 - \alpha.$$

Kriitilise väärtuse c_{kr} saame määrata F -jaotuse täiendkvantiili abil:

$$P\left(\frac{n-p-1}{p(n-2)} T_{(2)}^2 \leq \bar{f}_{\alpha;p,n-p-1}\right) = 1 - \alpha.$$

Kokkuvõttes saame lineaarkombinatsiooni usalduspiiride määramiseks järgmised eeskirjad

$$\begin{aligned} \underline{u} &= \vec{a}'(\vec{x}^{(1)} - \vec{x}^{(2)}) - c \sqrt{\vec{a}'\hat{\mathbf{S}}_{(2)}\vec{a}}, \\ \bar{u} &= \vec{a}'(\vec{x}^{(1)} - \vec{x}^{(2)}) + c \sqrt{\vec{a}'\hat{\mathbf{S}}_{(2)}\vec{a}}, \end{aligned}$$

kus

$$c = \sqrt{\frac{n_1 + n_2}{n_1 n_2} \frac{p(n-2)}{n-p-1} \bar{f}_{\alpha;p,n-p-1}}.$$

C. Profilide võrdlemine. Keskvärtusvektorite võrdlemisel on veel ühe võimalusena kasutusel nn profilide analüüs (profile analysis). Profiliks nimetatakse teatud joondigrammi. See on murdjoon, mille i -nda tipu vertikaalkoordinaadiks on μ_i , horisontaalkoordinaadiks aga i .

Meetod tuleb kõne alla, kui kõikide tunnuste mõõtmisel on kasutatud samu ühikuid. Tüüpiline on selle meetodi kasutamine kordusmõõtmiste puhul, mida rakendatakse paralleelselt kahes erinevas katsealuste rühmas, mida kirjeldavad juhuslikud vektorid $\vec{X}^{(1)} \sim N_p(\vec{\mu}^{(1)}, \mathbf{\Sigma})$ ja $\vec{X}^{(2)} \sim N_p(\vec{\mu}^{(2)}, \mathbf{\Sigma})$. Küsimuse, kas üldkogumite keskvärtused erinevad, võib asendada kolme järjestikuse küsimusega.

1. Kas üldkogumite profilid on paralleelsed?
2. Kas paralleelsed profilid langevad kokku?
3. Kas ühtivates profilides on kõik keskvärtused võrdsed?

Sõnastame need küsimused keskväärtusvektorite koordinaatide kaudu. Tähistame keskväärtusvektori koordinaadid erinevates üldkogumites

$$\vec{\mu}^{(1)} = \begin{pmatrix} \mu_1^{(1)} \\ \mu_2^{(1)} \\ \dots \\ \mu_p^{(1)} \end{pmatrix} \quad \vec{\mu}^{(2)} = \begin{pmatrix} \mu_1^{(2)} \\ \mu_2^{(2)} \\ \dots \\ \mu_p^{(2)} \end{pmatrix}.$$

Küsimuse, kas profiilid on paralleelsed, saame esitada nullhüpoteesina

$$H_{01} : \mu_j^{(1)} - \mu_j^{(2)} = \mu_{j-1}^{(1)} - \mu_{j-1}^{(2)}, \quad j = 2, \dots, p$$

või samaväärse nullhüpoteesina

$$H_{01} : \mu_j^{(1)} - \mu_{j-1}^{(1)} = \mu_j^{(2)} - \mu_{j-1}^{(2)}, \quad j = 2, \dots, p.$$

Kui nullhüpoteesi ei kummutata, võime küsimuse, kas profiilid langevad kokku, esitada nullhüpoteesina

$$H_{02} : \mu_j^{(1)} = \mu_j^{(2)}, \quad j = 1, 2, \dots, p.$$

Kui ka seda nullhüpoteesi ei kummutata, võime küsimuse, kas ühtivad profiilid on horisontaalsed, esitada hüpoteesina

$$H_{03} : \mu_1^{(1)} = \dots = \mu_p^{(1)} = \mu_1^{(2)} = \dots = \mu_p^{(2)}.$$

Neile küsimustele vastamise saame taandada normaaljaotuse lineaarteisenduse kohta käivate hüpoteeside kontrollimisele. Võtame kasutusele $(p-1) \times p$ maatriksi $\mathbf{C}$,

$$\mathbf{C} = \begin{pmatrix} -1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -1 & 1 \end{pmatrix}. \quad (12.1)$$

Siis saame hüpoteesi H_{01} esitada kujul

$$H_{01} : \mathbf{C}\vec{\mu}^{(1)} = \mathbf{C}\vec{\mu}^{(2)}.$$

Kuna $\mathbf{C}\vec{X}^{(i)} \sim N_{p-1}(\mathbf{C}\vec{\mu}^{(i)}, \mathbf{C}\Sigma\mathbf{C}')$, kus $i = 1, 2$, saame taandada hüpoteesi H_{01} kontrollimise $T_{(2)}^2$ -statistiku kasutamisele. Teisendatud tunnusvektorite jaoks omab kahe valimi T^2 -statistik kuju

$$T_{(2)}^2(\mathbf{C}\mathbf{X}) = \frac{n_1 n_2}{n_1 + n_2} (\vec{X}^{(1)} - \vec{X}^{(2)})' \mathbf{C}' (\mathbf{C}\Sigma_{(2)}\mathbf{C}')^{-1} \mathbf{C} (\vec{X}^{(1)} - \vec{X}^{(2)}).$$

Kui nullhüpotees kehtib, on $T_{(2)}^2(\mathbf{C}\mathbf{X}) \sim T^2(p-1, n_1 + n_2 - 2)$. Hüpoteesi H_{02} kontrollima asudes teame, et profiilid on paralleelsed, st kõik keskväärtusvektorite vastavad komponendid erinevad üksteisest ühepalju. Seetõttu piisab kontrollimaks, et $\vec{\mu}^{(1)} = \vec{\mu}^{(2)}$, kui kontrollime kas keskväärtusvektorite komponentide summad on võrdsed. Valime $1 \times p$ -maatriksiks $\mathbf{C}$ ühtede rea, $\mathbf{C} = \vec{\mathbb{1}}'_p$, siis

$$H_{02} : \vec{\mathbb{1}}'_p \vec{\mu}^{(1)} = \vec{\mathbb{1}}'_p \vec{\mu}^{(2)}.$$

Oleme jõudnud tuttavasse situatsiooni – tegemist on kahe ühemõõtmelise normaaljaotuse keskväärtuse võrdlemisega.

Kolmanda hüpoteesi kontrollimiseni jõudes teame, et $\vec{\mu}^{(1)} = \vec{\mu}^{(2)} = \vec{\mu}$. Valime maatriksi $\mathbf{C}$ samuti nagu esimese hüpoteesi testimisel, kasutades võrdust (12.1). Hüpoteesi H_{03} saame esitada kujul

$$H_{03} : \mathbf{C}\vec{\mu} = \vec{0}.$$

Jälle oleme jõudnud tuttavasse situatsiooni: meie käsutuses on valim mahuga $n_1 + n_2$ ühest üldkogumist. Hüpoteesi kontrollimiseks kasutame ühe valimi T^2 -statistikut.

Olukord on sootuks keerulisem, kui üldkogumite dispersioonimaatriksid on erinevad, st $\Sigma_1 \neq \Sigma_2$. Sel juhul pole võimalik kasutada täpseid jaotusi.

13 Korrelatsioonimaatriks

Mitmemõõtmelisel juhul on hajuvust iseloomustav karakteristik dispersioonimaatriks $D\mathbf{X}$, mis defineeriti p -vektori $\vec{X}$ jaoks $p \times p$ -maatriksina :

$$D\vec{X} = E[(\vec{X} - E\vec{X})(\vec{X} - E\vec{X})'].$$

Kui tähistame $D\vec{X} = \mathbf{\Sigma}$; $E\vec{X} = \vec{\mu}$, siis korrelatsioonikordajad ρ_{ij} avalduvad $\mathbf{\Sigma}$ elementide σ_{ij} kaudu:

$$\rho_{ij} = \frac{Cov(X_i, X_j)}{\sqrt{DX_i DX_j}} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}$$

ja neist moodustub korrelatsioonimaatriks

$$\mathbf{P} = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{p1} & \rho_{p2} & \cdots & 1 \end{pmatrix}.$$

Definitsioonist nähtub, et $\mathbf{P}$ on sümmeetriline maatriks. Veendume, et ta on mittenegatiivselt määratud. Konstruktsiooni tõttu on korrutis $(\vec{X} - E\vec{X})(\vec{X} - E\vec{X})'$ mittenegatiivselt määratud iga elementaarsündmuse $\omega \in \Omega$ korral ja sama omadus on siis ka selle maatriksi keskvaartusel. Seega $D\vec{X} = \mathbf{\Sigma}$ on mittenegatiivselt määratud ja kui ta on pööratav, siis positiivselt määratud. Korrelatsioonimaatriksi $\mathbf{P}$ saime esitada maatriksi $\mathbf{\Sigma}$ kaudu järgmiselt:

$$\mathbf{P} = \mathbf{\Sigma}_d^{-1/2} \mathbf{\Sigma} \mathbf{\Sigma}_d^{-1/2}.$$

Tänu sellele esitusele võime kirjutada

$$\mathbf{P} = \mathbf{\Sigma}_d^{-1/2} \mathbf{\Sigma}^{1/2} \mathbf{\Sigma}^{1/2} \mathbf{\Sigma}_d^{-1/2} = (\mathbf{\Sigma}_d^{-1/2} \mathbf{\Sigma}^{1/2})(\mathbf{\Sigma}_d^{-1/2} \mathbf{\Sigma}^{1/2})',$$

kust järeldub ka $\mathbf{P}$ mittenegatiivne määratus. Kui seejuures $\mathbf{\Sigma}$ on positiivselt määratud, siis on ka $\mathbf{P}$ positiivselt määratud.

Üldjuhul kuuluvad kahe tunnuse vahelise korrelatsioonikordaja väärtused lõiku $[-1, 1]$. Korrelatsioonimaatriksi positiivse määratuse nõue kitsendab võimalike korrelatsioonide suurust. Vaatame sellekohast näidet. Olgu $p = 3$ ja $\rho_{ij} = -\rho$, $\rho > 0$. Positiivse määratuse tõttu $|\mathbf{P}| > 0$, samas

$$|\mathbf{P}| = 1 - 2\rho^3 - 3\rho^2.$$

Kui $\rho = \frac{1}{2}$, siis $|\mathbf{P}| = 0$, järelikult peab kehtima võrratus $\rho < \frac{1}{2}$, ehk teisiti öeldes, niisuguse struktuuriga positiivselt määratud korrelatsioonimaatriksis peab korrelatsioon olema väiksem kui 0.5.

Valimist mahuga n saame leida korrelatsioonimaatriksi hinnangu $\hat{\mathbf{R}}$ dispersioonimaatriksi nihketa hinnangu $\hat{\mathbf{S}}$ kaudu:

$$\hat{\mathbf{R}} = \hat{\mathbf{S}}_d^{-1/2} \hat{\mathbf{S}} \hat{\mathbf{S}}_d^{-1/2}.$$

Vastav statistik $\mathbf{R}$ avaldub kujul

$$\mathbf{R} = \mathbf{S}_d^{-1/2} \mathbf{S} \mathbf{S}_d^{-1/2}. \quad (13.1)$$

Kahjuks muudab see lihtne teisendus korrelatsioonimaatriksi jaotuse keeruliseks, seda ei saa enam esitada Wisharti või mõne teise tuntud jaotuse kaudu. Ka klassikalise normaaljaotusega üldkogumi korral ei ole maatriksi $\mathbf{R}$ jaotus teada, leitud on asümptootiline jaotus kui $n \rightarrow \infty$. Küll aga saame leida maatriksi $\mathbf{R}$ täpse jaotuse ühel olulisel erijuhul.

Teoreem 13.1. *Olgu üldkogum normaaljaotusega $N_p(\vec{\mu}, \mathbf{\Sigma}_d)$ (tunnused on sõltumatud), siis valimi korrelatsioonimaatriksi $\mathbf{R}$ tõenäosustihedus on kujul*

$$f_{\mathbf{R}}(\mathbf{R}) = \frac{[\Gamma(\frac{1}{2}(n-1))]^p}{\Gamma_p(\frac{1}{2}(n-1))} |\mathbf{R}|^{\frac{1}{2}(n-p-2)},$$

kus p -mõõtmeline Γ -funktsioon on määratud eeskirjaga

$$\Gamma_p(a) = \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^p \Gamma[a - \frac{1}{2}(i-1)]. \quad (13.2)$$

Tõestus. Lähtume Wisharti jaotusest ja valemist (13.1). Teame, et

$$\mathbf{V} = (n-1)\mathbf{S} \sim W_p(\Sigma_d, n-1).$$

Teoreemi 7.2 põhjal avaldub Wisharti jaotuse tihedus antud juhul järgmiselt:

$$f_{\mathbf{V}}(\mathbf{V}) = \frac{|\mathbf{V}|^{\frac{1}{2}(n-p-2)}}{2^{\frac{p(n-1)}{2}} |\Sigma_d|^{\frac{n-1}{2}} \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^p \Gamma[\frac{1}{2}(n-i)]} e^{-\frac{1}{2}\text{tr}(\Sigma_d^{-1}\mathbf{V})}.$$

Teeme muutujavahetuse minnes üle maatriksi $\mathbf{V}$ alumise kolmnurga elementidelt

$$v_{11}, v_{22}, \dots, v_{pp}, v_{21}, \dots, v_{p1}, \dots, v_{p,p-1}$$

maatriksi $\mathbf{V}$ diagonaalelementidele ja maatriksi $\mathbf{R}$ allpool peadiagonaali asuvatele elementidele r_{ij} , $i \neq j$:

$$v_{11}, v_{22}, \dots, v_{pp}, r_{21}, \dots, r_{p1}, \dots, r_{p,p-1}.$$

Uued ja vanad muutujad on seotud teisendustega

$$\mathbf{R} = \mathbf{V}_d^{-\frac{1}{2}} \mathbf{V} \mathbf{V}_d^{-\frac{1}{2}}, \quad \mathbf{V} = \mathbf{V}_d^{\frac{1}{2}} \mathbf{R} \mathbf{V}_d^{\frac{1}{2}}.$$

Kasutades üldist seost tiheduste vahel saame

$$f_{\mathbf{V}_d, \mathbf{R}}(\mathbf{V}_d, \mathbf{R}) = f_{\mathbf{V}(\mathbf{R})}(\mathbf{V}(\mathbf{R})) J \left(\frac{\mathbf{V}(\mathbf{R})}{\mathbf{V}_d, \mathbf{R}} \right).$$

Leiame jakobiaani

$$J \left(\frac{\mathbf{V}(\mathbf{R})}{\mathbf{V}_d, \mathbf{R}} \right) = \text{abs} \left| \frac{d(v_{11}, v_{22}, \dots, v_{pp}, v_{21}, \dots, v_{p1}, \dots, v_{p,p-1})}{d(v_{11}, v_{22}, \dots, v_{pp}, r_{21}, \dots, r_{p1}, \dots, r_{p,p-1})} \right|.$$

Saadava tuletismaatriksi diagonaalplokid on kahanevat järku: $p, p-1, \dots, 1$, kusjuures esimene plokk võrdub $\mathbf{I}_p$ ning i -s diagonaalplokk on $(p-i) \times (p-i)$ -maatriks kujul

$$\begin{pmatrix} \sqrt{v_{i-1,i-1}v_{ii}} & 0 & \cdots & 0 \\ 0 & \sqrt{v_{i-1,i-1}v_{i+1,i+1}} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \sqrt{v_{i-1,i-1}v_{pp}} \end{pmatrix},$$

kus $i = 2, \dots, p$. Diagonaalplokkidest ülalpool koosnevad kõik plokid nullidest ja seega saame determinandi arvutamisel kasutada 2×2 -plokkmaatriksi determinandi omadust:

$$\begin{vmatrix} \mathbf{A}_{11} & 0 \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{vmatrix} = |\mathbf{A}_{11}| |\mathbf{A}_{22}|.$$

Seetõttu omab jakobiaan väärtust

$$J \left(\frac{\mathbf{V}(\mathbf{R})}{\mathbf{V}_d, \mathbf{R}} \right) = v_{11}^{\frac{p-1}{2}} \times \cdots \times v_{pp}^{\frac{p-1}{2}} = \prod_{i=1}^p v_{ii}^{\frac{p-1}{2}}.$$

Vaadeldav tihedusfunktsioon omandab kuju

$$\begin{aligned} & f_{\mathbf{V}_d, \mathbf{R}}(\mathbf{V}_d, \mathbf{R}) \\ &= \frac{|\mathbf{R}|^{\frac{n-p-2}{2}} |\mathbf{V}_d|^{\frac{n-p-2}{2}} \prod_{i=1}^p v_{ii}^{\frac{p-1}{2}}}{2^{\frac{p(n-1)}{2}} \pi^{\frac{p(p-1)}{4}} |\Sigma_d|^{\frac{n-1}{2}} \prod_{i=1}^p \Gamma\left(\frac{n-i}{2}\right)} \exp\left(-\frac{1}{2} \sum_{i=1}^p \frac{v_{ii}}{\sigma_{ii}}\right) \\ &= \frac{|\mathbf{R}|^{\frac{n-p-2}{2}} \prod_{i=1}^p v_{ii}^{\frac{n-p-2}{2}} \prod_{i=1}^p v_{ii}^{\frac{p-1}{2}}}{2^{\frac{p(n-1)}{2}} \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^p \sigma_{ii}^{\frac{n-1}{2}} \prod_{i=1}^p \Gamma\left(\frac{n-i}{2}\right)} \exp\left(-\frac{1}{2} \sum_{i=1}^p \frac{v_{ii}}{\sigma_{ii}}\right) \\ &= \frac{|\mathbf{R}|^{\frac{n-p-2}{2}}}{\pi^{\frac{p(p-1)}{4}} \prod_{i=1}^p \Gamma\left(\frac{n-i}{2}\right)} \prod_{i=1}^p \left[\frac{v_{ii}^{\frac{n-3}{2}}}{\sigma_{ii}^{\frac{n-1}{2}}} \exp\left(-\frac{1}{2} \frac{v_{ii}}{\sigma_{ii}}\right) 2^{-\frac{(n-1)}{2}} \right]. \end{aligned}$$

Oleme esitanud tihedusfunktsiooni korruisena, kus murruga antud tegur sõltub vaid korrelatsioonimaatriksist ning teine tegur sõltub muutujatest v_{ii} , $i = 1, \dots, p$. Integreerides üle muutujate v_{ii} , $i = 1, \dots, p$, saame tulemusena korrelatsioonimaatriksi tihedusfunktsiooni. Kuna

$$\begin{aligned} & \int_{\mathbb{R}^p} \prod_{i=1}^p \left[\frac{v_{ii}^{\frac{1}{2}(n-3)}}{\sigma_{ii}^{\frac{n-1}{2}}} \exp\left(-\frac{1}{2} \frac{v_{ii}}{\sigma_{ii}}\right) 2^{-\frac{(n-1)}{2}} \right] dv_{11} \dots dv_{pp} \\ &= \prod_{i=1}^p \int_0^\infty \left[\frac{v_{ii}^{\frac{1}{2}(n-3)}}{\sigma_{ii}^{\frac{n-1}{2}}} \exp\left(-\frac{1}{2} \frac{v_{ii}}{\sigma_{ii}}\right) 2^{-\frac{(n-1)}{2}} \right] dv_{ii}, \end{aligned}$$

siis uurime lähemalt saadud integraali avaldist. Tähistame

$$\frac{v_{ii}}{2\sigma_{ii}} = u_i, \quad \text{kust} \quad dv_{ii} = 2\sigma_{ii} du_i.$$

Saame integraali kujul

$$\int_0^\infty \frac{v_{ii}^{\frac{1}{2}(n-1)-1}}{(2\sigma_{ii})^{\frac{n-1}{2}}} \exp\left(-\frac{1}{2} \frac{v_{ii}}{\sigma_{ii}}\right) dv_{ii} = \int_0^\infty u_i^{\frac{n-1}{2}-1} e^{-u_i} du_i.$$

Arvestades Γ -funktsiooni definitsiooni

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt,$$

esitub meie integraal Γ -funktsioonina $\Gamma\left(\frac{n-1}{2}\right)$. Seega avaldub korrelatsioonimaatriksi $\mathbf{R}$ tõenäosustihedus valemiga

$$f_{\mathbf{R}}(\mathbf{R}) = \frac{|\mathbf{R}|^{\frac{n-p-2}{2}} \left[\Gamma\left(\frac{n-1}{2}\right)\right]^p}{\pi^{\frac{p(p-1)}{4}} \prod_{i=1}^p \Gamma\left(\frac{n-i}{2}\right)}.$$

Kui võrduses (13.2) $a = \frac{n-1}{2}$, siis

$$\begin{aligned} \Gamma_p\left(\frac{n-1}{2}\right) &= \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^p \Gamma\left(\frac{n-1}{2} - \frac{i-1}{2}\right) \\ &= \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^p \Gamma\left(\frac{n-i}{2}\right), \end{aligned}$$

mott.

Järeldus 13.1.1. Kahemõõtmelisel juhul avaldub korrelatsioonikordaja r tõenäosustihedus kujul:

$$f_r(r) = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-2}{2}\right) \sqrt{\pi}} (1-r^2)^{\frac{n}{2}-2},$$

kui teoreetiline korrelatsioonikordaja $\rho = 0$.

14 Mitmene korrelatsioonikordaja

Kui korrelatsioonimaatriksi kirjeldab seoseid juhusliku vektori komponentide vahel, siis mitmene korrelatsioonikordaja kirjeldab seost juhusliku suuruse ja p -mõõtmelise juhusliku vektori vahel. Mitmene korrelatsioonikordajani jõudmiseks vaatame prognoosiülesannet: olgu meie eesmärgiks prognoosida juhuslikku suurust Y juhusliku p -vektori $\vec{X}$ kaudu. Prognoosiks on etteantud funktsioonide klassi $\mathcal{G}$ kuuluv funktsioon $g(\vec{X}) : \mathbb{R}^p \rightarrow \mathbb{R}$. Prognoosi $g(\vec{X})$ prognoosiveaks vähimruutude mõttes nimetatakse suurust

$$Q(g(\vec{X})) = E[Y - g(\vec{X})]^2. \quad (14.1)$$

Eeldame, et juhuslikud suurused Y ja $g(\vec{X})$ on teist järku, st eksisteerivad EY^2 ja $E[g(\vec{X})]^2$. Sel juhul on prognoosiviga määratud ja lõplik. Prognoos $g^*(\vec{X})$ on juhusliku suuruse Y *parim prognoos vähimruutude mõttes*, kui

$$E[Y - g^*(\vec{X})]^2 = \min_{g \in \mathcal{G}} E[Y - g(\vec{X})]^2.$$

Lause 14.1. *Juhusliku suuruse Y parimaks prognoosiks vähimruutude mõttes on tinglik keskvärtus $E(Y|\vec{X})$.*

Tõestus. Teisendame valemiga (14.1) määratud prognoosiviga

$$\begin{aligned} E[Y - g(\vec{X})]^2 &= E[Y - E(Y|\vec{X}) + E(Y|\vec{X}) - g(\vec{X})]^2 \\ &= E[Y - E(Y|\vec{X})]^2 + E[E(Y|\vec{X}) - g(\vec{X})]^2 \\ &\quad + 2E\{[Y - E(Y|\vec{X})][E(Y|\vec{X}) - g(\vec{X})]\}. \end{aligned}$$

Viimane liidetav on null tingliku keskvärtuse omaduste tõttu. Nimelt

$$E[Yh(\vec{X})] = E\{E[Yh(\vec{X})|\vec{X}]\} = E\{h(\vec{X})E(Y|\vec{X})\}.$$

Saame

$$E[Y - g(\vec{X})]^2 = E[Y - E(Y|\vec{X})]^2 + E[E(Y|\vec{X}) - g(\vec{X})]^2.$$

See avaldis on minimaalne, kui $g(\vec{X}) = E(Y|\vec{X})$, niisugune valik muudab teise liidetava nulliks, mott.

Funktsiooni $g^*(\vec{X}) = E(Y|\vec{X})$ nimetatakse juhusliku suuruse Y *mitmeseks regressioonifunktsiooniks* $\vec{X}$ järgi.

Definitsioon 14.1. *Juhuslik suurus Y sõltub juhuslikust vektorist $\vec{X}$ regressiooni mõttes, kui*

$$E(Y|\vec{X}) \neq EY.$$

Parimat prognoosi on lihtne defineerida, kuid tema leidmist praktikas on raske realiseerida. Tingliku keskvärtuse $E(Y|\vec{X})$ jaoks on pea võimatu saada analüütilist avaldist. Lihtsustame oma ülesannet sel teel, et otsime lähendit mitte teist järku juhuslike suuruste klassist $\mathcal{G}$, vaid selle mingist alamklassist $\mathcal{F} \subset \mathcal{G}$. Sel juhul räägitakse $\mathcal{F}$ -regressioonist. Võtame järgnevas vaatluse alla kõige lihtsama mõeldava klassi – lineaarsete funktsioonide klassi $\mathcal{L}$ ja otsime parimat prognoosi kujul

$$L(\vec{X}; a_0, a_1, \dots, a_p) = a_0 + a_1 X_1 + \dots + a_p X_p.$$

Võtame kasutusele tähistused

$$X_0 = 1, \quad \vec{X}^* = \begin{bmatrix} 1 \\ \vec{X} \end{bmatrix}; \quad \vec{a} = \begin{pmatrix} a_1 \\ \dots \\ a_p \end{pmatrix}, \quad \vec{a}^* = \begin{bmatrix} a_0 \\ \vec{a} \end{bmatrix}.$$

Otsitava parima lähendi saab siis esitada kujul

$$L(\vec{X}; a_0, a_1, \dots, a_p) = \vec{a}^{*'} \vec{X}^* = a_0 + \vec{a}' \vec{X}$$

ja meil on tegemist lineaarse regressiooniga, vastavat parimat prognoosi $L(\vec{X}; a_0, a_1, \dots, a_p)$ nimetatakse *lineaarseks regressioonifunktsiooniks*.

Definitsioon 14.2. Juhuslik suurus Y sõltub juhuslikust vektorist $\vec{X}$ korrelatiivselt, kui lineaarne regressioonifunktsioon ei ole konstantne ($\exists a_i \neq 0, i > 0$).

Uurime, kuidas avalduvad lineaarse regressioonifunktsiooni kordajad $a_0, a_1, \dots, a_p$. Olgu mingi lineaarne prognoos määratud vektoritega $\vec{b}$ ja $\vec{b}^*$:

$$L(\vec{X}; b_0, b_1, \dots, b_p) = \vec{b}^{*'} \vec{X}^* = b_0 + \vec{b}' \vec{X}.$$

Parima lineaarse prognoosi korral peab kordajate jaoks olema täidetud tingimus

$$E(Y - \vec{a}^{*'} \vec{X}^*)^2 = \min_{\vec{b}^*} E(Y - \vec{b}^{*'} \vec{X}^*)^2.$$

Kirjutame välja miinimumi tarviliku tingimuse

$$\frac{dE(Y - \vec{b}^{*'} \vec{X}^*)^2}{d\vec{b}^*} = \vec{0}'.$$

Eeldame, et tegemist on regulaarse jaotusega. Sel juhul võime tuletise võtmise ja keskväärtuse leidmise operatsioonid vahetada. Rakendades maatrikstuletise omadusi (v) ja (vi) lisast 1, saame:

$$\frac{dE(Y - \vec{b}^{*'} \vec{X}^*)^2}{d\vec{b}^*} = -2E[(Y - \vec{b}^{*'} \vec{X}^*) \vec{X}^{*'}] = \vec{0}',$$

millest tuleneb tingimus:

$$E(Y \vec{X}^{*'}) = \vec{a}^{*'} E(\vec{X}^* \vec{X}^{*'}).$$

Saadud tingimuse võime esitada ka transponeeritud kujul

$$E(\vec{X}^* Y) = E(\vec{X}^* \vec{X}^{*'}) \vec{a}^*. \quad (14.2)$$

Kuna vektori $\vec{X}^*$ esimene koordinaat on 1, siis saadud võrrandisüsteemi esimene võrrand omab kuju

$$EY = a_0 + \sum_{j=1}^p a_j EX_j.$$

Avaldame sellest võrrandist kordaja a_0

$$a_0 = EY - \sum_{j=1}^p a_j EX_j. \quad (14.3)$$

Võrrandisüsteemi (14.2) ülejäänud võrrandid on kujul

$$E(X_i Y) = a_0 EX_i + \sum_{j=1}^p a_j E(X_i X_j), \quad i = 1, 2, \dots, p.$$

Asendades nendes võrrandites kordaja a_0 vördusest (14.3) ja muutes liidetavate järjekorda, saame p võrrandist koosneva süsteemi kordajate määramiseks

$$E(X_i Y) - EX_i EY = \sum_{j=1}^p [E(X_i X_j) - EX_i EX_j] a_j, \quad i = 1, 2, \dots, p.$$

Selle võrrandi kordajateks ja vabaliikmeteks on kovariatsioonid ning dispersioonid

$$\begin{aligned} Cov(X_i Y) &= E(X_i Y) - EX_i EY, \\ Cov(X_i X_j) &= E(X_i X_j) - EX_i EX_j, \quad i \neq j \\ DX_i &= E(X_i^2) - (EX_i)^2. \end{aligned}$$

Maatrikskujul saame võrrandisüsteemi parima lineaarse prognoosi kordajate vektori $\vec{a}$ määramiseks esitada kujul

$$(D\vec{X})\vec{a} = Cov(Y, \vec{X}), \quad (14.4)$$

kust kordajate vektor avaldub kujul

$$\vec{a} = (D\vec{X})^{-1}Cov(Y, \vec{X}), \quad (14.5)$$

kusjuures vabaliige a_0 määratakse võrdusega (14.3).

Defineerime nüüd mitmese korrelatsioonikordaja.

Definitsioon 14.3. Juhusliku suuruse Y ja juhusliku vektori $\vec{X}$ vaheliseks mitmeseks korrelatsioonikordajaks $r(Y, \vec{X})$ nimetatakse korrelatsioonikordajat $\rho(Y, \vec{a}^* \vec{X}^*)$ juhusliku suuruse Y ja tema parima lineaarse prognoosi $\vec{a}^* \vec{X}^*$ vahel.

Lause 14.2. Juhusliku suuruse Y ja juhusliku vektori $\vec{X}$ vaheline mitmene korrelatsioonikordaja $r(Y, \vec{X})$ avaldub kujul

$$r(Y, \vec{X}) = \sqrt{\frac{Cov(Y, \vec{X})'(D\vec{X})^{-1}Cov(Y, \vec{X})}{DY}}. \quad (14.6)$$

Tõestus. Definitsiooni kohaselt kehtib võrdus

$$r(Y, \vec{X}) = \rho(Y, \vec{a}^* \vec{X}^*) = \frac{Cov(Y, \vec{a}^* \vec{X}^*)}{\sqrt{DY \cdot D(\vec{a}^* \vec{X}^*)}}.$$

Avaldame lugejas oleva kovariatsiooni

$$\begin{aligned} Cov(Y, \vec{a}^* \vec{X}^*) &= E[(Y - EY)(\vec{a}^* \vec{X}^* - \vec{a}^* E\vec{X}^*)] \\ &= \vec{a}^* E[(Y - EY)(\vec{X}^* - E\vec{X}^*)] \\ &= \vec{a}^* Cov(Y, \vec{X}^*). \end{aligned}$$

Arvestades, et vektori $\vec{X}^*$ esimeseks komponendiks on konstant 1, tuleneb siit võrdus

$$\vec{a}^* Cov(Y, \vec{X}^*) = \vec{a}' Cov(Y, \vec{X}),$$

kuna $Cov(Y, 1) = 0$. Leiame veel parima prognoosi dispersiooni

$$D(\vec{a}^* \vec{X}^*) = \vec{a}^{*'} D\vec{X}^* \vec{a}^* = \vec{a}' D\vec{X} \vec{a} = \vec{a}' Cov(Y, \vec{X}). \quad (14.7)$$

Viimase võrduse saame arvestades seost (14.4). Asendame saadud avaldised korrelatsioonikordaja valemisse

$$r(Y, \vec{X}) = \frac{\vec{a}' Cov(Y, \vec{X})}{\sqrt{DY \cdot \vec{a}' Cov(Y, \vec{X})}} = \sqrt{\frac{\vec{a}' Cov(Y, \vec{X})}{DY}}.$$

Arvestades vektori $\vec{a}$ kuju (14.5), saame

$$r(Y, \vec{X}) = \sqrt{\frac{Cov(Y, \vec{X})'(D\vec{X})^{-1}Cov(Y, \vec{X})}{DY}},$$

mott.

Järeldus 14.0.1. Mitmene korrelatsioonikordaja $r(Y, \vec{X})$ avaldub vektori $\vec{X}$ korrelatsioonimaatriksi $\mathbf{P}$ ja korrelatsioonide vektori $\vec{\rho}(Y, \vec{X})$ kaudu,

$$r(Y, \vec{X}) = \sqrt{\vec{\rho}(Y, \vec{X})' \mathbf{P}^{-1} \vec{\rho}(Y, \vec{X})},$$

kus

$$\vec{\rho}(Y, \vec{X}) = \begin{pmatrix} \rho(Y, X_1) \\ \rho(Y, X_2) \\ \vdots \\ \rho(Y, X_p) \end{pmatrix}.$$

Tõestus. Tähistame diagonaalmaatriksi, mille peadiagonaalil on vektori $\vec{X}$ komponentide dispersioonid, sümboliga $(D\vec{X})_d$. Vastavalt korrelatsioonikordaja definitsioonile saame siis korrelatsioonivektori esitada kujul

$$\vec{\rho}(Y, \vec{X}) = \sqrt{(DY)^{-1}(D\vec{X})_d^{-1}} \text{Cov}(Y, \vec{X}),$$

kust

$$\text{Cov}(Y, \vec{X}) = \sqrt{(DY)(D\vec{X})_d} \vec{\rho}(Y, \vec{X}).$$

Korrelatsioonimaatriks on aga määratud valemiga

$$\mathbf{P} = \sqrt{(D\vec{X})_d^{-1}} D\vec{X} \sqrt{(D\vec{X})_d^{-1}},$$

kust

$$D\vec{X} = \sqrt{(D\vec{X})_d} \mathbf{P} \sqrt{(D\vec{X})_d}.$$

Asendame saadud avaldised valemisse (14.6); vältimaks liigset juurimist avaldame mitmese korrelatsioonikordaja ruudu

$$\begin{aligned} r^2(Y, \vec{X}) &= (\vec{\rho}(Y, \vec{X}))' \sqrt{(D\vec{X})_d} \sqrt{(D\vec{X})_d^{-1}} \mathbf{P}^{-1} \\ &\quad \times \sqrt{(D\vec{X})_d^{-1}} \sqrt{(D\vec{X})_d} \vec{\rho}(Y, \vec{X}), \end{aligned}$$

kust

$$r(Y, \vec{X}) = \sqrt{(\vec{\rho}(Y, \vec{X}))' \mathbf{P}^{-1} \vec{\rho}(Y, \vec{X})},$$

mott.

Mitmene korrelatsioonikordaja on seotud lineaarse prognoosiga. Võtame kasutusele kordaja, mida saab kasutada regressioonisõltuvuse tugevuse kirjeldamiseks üldjuhul. Nägime, et juhusliku suuruse Y parimaks prognoosiks vähimruutude mõttes on tinglik keskväärts $E(Y|\vec{X})$. Tähistame prognoosivea parima prognoosi korral sümboliga $Q^*(E(Y|\vec{X}))$.

Definitsioon 14.4. *Regressioonisuhteks $\eta(Y|\vec{X})$ nimetatakse suurust*

$$\eta(Y|\vec{X}) = \sqrt{1 - \frac{Q^*(E(Y|\vec{X}))}{DY}}.$$

Selle kordaja defineerimise lähtekohaks on dispersiooni lahutus erinevateks liidetavateks:

$$\begin{aligned} DY &= E(Y - EY)^2 = E[Y - E(Y|\vec{X}) + E(Y|\vec{X}) - EY]^2 \\ &= E[Y - E(Y|\vec{X})]^2 + E[E(Y|\vec{X}) - EY]^2 \\ &\quad + 2E\{[Y - E(Y|\vec{X})][E(Y|\vec{X}) - EY]\} \\ &= Q^*(E(Y|\vec{X})) + D(E(Y|\vec{X})). \end{aligned}$$

Definitsioonist järelduvad järgmised regressioonisuhte omadused.

1. Kui regressioonisõltuvus puudub, siis $\eta(Y|\vec{X}) = 0$. Tõepoolest, kui regressioonisõltuvus puudub, siis $E(Y|\vec{X}) = EY$ ja $DY = Q^*(E(Y|\vec{X}))$.
2. Kui Y on $\vec{X}$ funktsioon, siis $\eta(Y|\vec{X}) = 1$. Tõepoolest, kui Y on $\vec{X}$ funktsioon, siis $E(Y|\vec{X}) = Y$ ja $Q^*(E(Y|\vec{X})) = 0$.
3. Kahe juhusliku suuruse X ja Y korral kehtib alati võrratus

$$\eta(Y|X) \geq \rho(X, Y),$$

kus $\rho(X, Y)$ on lineaarne korrelatsioonikordaja. Võrratus kehtib, kuna regressioonisuhe mõõdab üldisema funktsionaalse sõltuvuse tugevust, $\rho(X, Y)$ aga erijuhu, lineaarse sõltuvuse tugevust.

Kokkuvõtlikult öeldes on tegemist normeeritud suurusega, mis iseloomustab regressiooni-sõltuvuse tugevust.

MÄRKUS. Regressioonisuhte arvutamiseks on definitsiooni asemel otstarbekam kasutada esitust

$$\eta(Y|\vec{X}) = \sqrt{\frac{D(E(Y|\vec{X}))}{DY}},$$

mida saame rakendada lihtsate juhuslike suuruste ja klassifitseeritud tunnuste korral üldise funktsionaalse sõltuvuse tugevuse hindamiseks.

Lineaarse prognoosiülesande korral regressioonisuhte leidmine lihtsustub. Tõestame järgmise lause.

Lause 14.3. Lineaarse prognoosi korral võrdub regressioonisuhe $\eta_L(Y|\vec{X})$ mitmese korrelatsioonikordajaga,

$$\eta_L(Y|\vec{X}) = r(Y, \vec{X}).$$

Tõestus. Parima lineaarse prognoosi korral

$$\eta_L(Y|\vec{X}) = \sqrt{1 - \frac{Q^*(\vec{a}^{*'} \vec{X}^*)}{DY}}.$$

Teisendame prognoosivea avaldist

$$Q^*(\vec{a}^{*'} \vec{X}^*) = E(Y - \vec{a}^{*'} \vec{X}^*)^2 = E[Y - EY + EY - \vec{a}^{*'} \vec{X}^*]^2.$$

Võrrandisüsteemi (14.2) esimese võrrandi põhjal $EY = \vec{a}^{*'} EX^*$, seetõttu võime kirjutada

$$Q^*(\vec{a}^{*'} \vec{X}^*) = E[Y - EY - \vec{a}^{*'} (\vec{X}^* - E\vec{X}^*)]^2 = DY - \vec{a}^{*'} D\vec{X}^* \vec{a}^*.$$

Viimase võrduse saamisel kasutasime seoseid (14.7). Siit

$$\eta_L(Y|\vec{X}) = \sqrt{\frac{\vec{a}^{*'} D\vec{X}^* \vec{a}^*}{DY}}.$$

Arvestades valemeid (14.5)-(14.7) saamegi

$$\eta_L(Y|\vec{X}) = r(Y, \vec{X}),$$

mott.

Vaatame, kuidas saab kasutada mitmest korrelatsioonikordajat normaaljaotuse korral. Olgu meil tegemist normaaljaotusega juhusliku vektoriga, mille jaotame kaheks osaks. Esimese komponendi X_1 valime sõltuvaks tunnuseks, ülejäänud komponente aga kasutame vektorina, mille abil seda komponenti prognoosime. Seega

$$\begin{bmatrix} X_1 \\ \vec{X}_2 \end{bmatrix} \sim N_p \left(\begin{bmatrix} \mu_1 \\ \vec{\mu}_2 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \vec{\sigma}_{21}' \\ \vec{\sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix} \right).$$

Esitame kontrollimiseks hüpoteesid

$$\begin{aligned} H_0 : X_1 & \text{ on vektorist } \vec{X}_2 \text{ sõltumatu,} \\ H_1 : X_1 & \text{ ei ole vektorist } \vec{X}_2 \text{ sõltumatu.} \end{aligned}$$

Osutub, et teststatistik nende hüpoteeside kontrollimiseks on valimi mitmese korrelatsioonikordaja funktsioon. Teame, et maksimaalne korrelatsioon juhusliku suuruse X_1 ja juhusliku suuruse $\vec{a}' \vec{X}_2$ vahel saavutatakse, kui $\vec{a} = \mathbf{\Sigma}_{22}^{-1} \vec{\sigma}_{21}$. Sõltuvus puudub, kui kõik regressioonikordajad on nullid. Seega saame esitatud hüpoteesid sõnastada ka teisiti

$$\begin{aligned} H_0 : \vec{\sigma}_{21} &= \vec{0} \text{ ehk } r(X_1, \vec{X}_2) = 0, \\ H_1 : \vec{\sigma}_{21} &\neq \vec{0} \text{ ehk } r(X_1, \vec{X}_2) \neq 0. \end{aligned}$$

Valimi põhjal saame leida valimi kovariatsioonimaatriksi $\hat{\mathbf{S}}$. Vastava statistiku $\mathbf{S}$ jaotus on meile teada,

$$(n-1)\mathbf{S} = \mathbf{V} = \begin{bmatrix} v_{11} & \vec{v}_{21}' \\ \vec{v}_{21} & \mathbf{V}_{22} \end{bmatrix} \sim W_p(\mathbf{\Sigma}, n-1).$$

Valimi mitmese korrelatsioonikordaja ruut on määratud valemiga

$$\widehat{r(X_1, X_2)}^2 = \frac{\vec{v}_{21}' \hat{\mathbf{V}}_{22}^{-1} \vec{v}_{21}}{\hat{v}_{11}},$$

kus võrduse paremal poolel esinevad hinnangud on saadud seosest $(n-1)\hat{\mathbf{S}} = \hat{\mathbf{V}}$. Tähistame vastava statistiku

$$\mathcal{R}(X_1, \vec{X}_2)^2 = \frac{\vec{v}_{21}' \mathbf{V}_{22}^{-1} \vec{v}_{21}}{v_{11}}.$$

Lause 14.4. *Tõepärasuhte statistik hüpoteeside*

$$H_0 : \vec{\sigma}_{21} = \vec{0},$$

$$H_1 : \vec{\sigma}_{21} \neq \vec{0}$$

kontrollimiseks omab kuju

$$\Lambda(\mathbf{X}) = (1 - \mathcal{R}(X_1, \vec{X}_2)^2)^{n/2}.$$

Tõestus. Vastavalt definitsioonile 9.1 avaldub tõepärasuhte statistik kujul

$$\Lambda(\mathbf{X}) = \frac{\max_{\vec{\mu}, \mathbf{\Sigma}, \vec{\sigma}_{21}=\vec{0}} L(\mathbf{X}; \vec{\mu}, \mathbf{\Sigma})}{\max_{\vec{\mu}, \mathbf{\Sigma}} L(\mathbf{X}; \vec{\mu}, \mathbf{\Sigma})}.$$

Teame, et tõepärafunktsioon saavutab maksimaalse väärtuse, kui parameetrite väärtuseks valida nende suurima tõepära hinnangud. Nimetajas peame kasutama tõepärafunktsiooni argumentidega $\vec{x}$ ja $\hat{\mathbf{S}}_*$, $L(\mathcal{X}; \vec{x}, \hat{\mathbf{S}}_*)$. Kui H_0 on tõene, on dispersioonimaatriksi $\mathbf{\Sigma}$ suurima tõepära hinnang

$$\hat{\mathbf{S}}_*^0 = \frac{1}{n} \begin{bmatrix} \hat{v}_{11} & \vec{0}' \\ \vec{0} & \hat{\mathbf{V}}_{22} \end{bmatrix}$$

ja lugejas peame kasutama tõepärafunktsiooni argumentidega $\vec{x}$ ja $\hat{\mathbf{S}}_*^0$, $L(\mathcal{X}; \vec{x}, \hat{\mathbf{S}}_*^0)$. Arvestades üheksandas paragrahvis saadud tulemusi, saame

$$\Lambda(\mathcal{X}) = \frac{|\hat{\mathbf{S}}_*^0|^{-n/2}}{|\hat{\mathbf{S}}_*|^{-n/2}}. \quad (14.8)$$

Kasutades valemit (L1.13) avaldub plokk-maatriksi

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

determinant kujul

$$|\mathbf{A}| = |\mathbf{A}_{22}| |\mathbf{A}_{11} - \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \mathbf{A}_{21}|.$$

Rakendades seda võrdust valemis (14.8) esinevate determinantide korral, saame

$$\begin{aligned} \Lambda(\mathbf{X}) &= \left(\frac{|\mathbf{V}_{22}| |v_{11} - \vec{v}_{21}' \mathbf{V}_{22}^{-1} \vec{v}_{21}|}{v_{11} |\mathbf{V}_{22}|} \right)^{n/2} \\ &= \left(\frac{|v_{11} - \vec{v}_{21}' \mathbf{V}_{22}^{-1} \vec{v}_{21}|}{v_{11}} \right)^{n/2} \\ &= \left(1 - \frac{\vec{v}_{21}' \mathbf{V}_{22}^{-1} \vec{v}_{21}}{v_{11}} \right)^{n/2} = (1 - \mathcal{R}(X_1, \vec{X}_2)^2)^{n/2}, \end{aligned}$$

mott.

Saab näidata, et nullhüpoteesi kehtides

$$\frac{n-p}{p-1} \frac{\mathcal{R}(X_1, \vec{X}_2)^2}{1 - \mathcal{R}(X_1, \vec{X}_2)^2} \sim F(p-1, n-p).$$

Praktika jaoks on oluline ligikaudne asümptootiline jaotus. Esitame selle järgneva lausega.

Lause 14.5. *Normaaljaotusega üldkogumi korral kehtivad valimimahu piiramatul kasvamisel ($n \rightarrow \infty$) järgmised koondumised.*

14.1. Kui $r(X_1, \vec{X}_2) \neq 0$, siis

$$\sqrt{n}(\mathcal{R}(X_1, \vec{X}_2)^2 - r(X_1, \vec{X}_2)^2) \xrightarrow{\mathcal{D}} N(0, 4\omega),$$

kus

$$\omega = r(X_1, \vec{X}_2)^2(1 - r(X_1, \vec{X}_2)^2)^2.$$

14.2. Kui $r(X_1, \vec{X}_2) = 0$, siis

$$n\mathcal{R}(X_1, \vec{X}_2)^2 \xrightarrow{\mathcal{D}} \chi^2(p-1).$$

Ülesanded 14

14.1. Olgu X ja Y lihtsad juhuslikud suurused. Näidata, et

$$EY = E[E(Y|X)].$$

14.2. Olgu X ja Y lihtsad juhuslikud suurused. Tõestada, et

$$E[Yh(X)|X] = h(X)E(Y|X).$$

14.3. Näidata, et

$$E\{[Y - E(Y|\vec{X})][E(Y|\vec{X}) - EY]\} = 0.$$

15 Osakorrelatsioon

Kui juhuslik suurus Y sõltub juhuslikust p -vektorist $\vec{X}$ regressiooni mõttes, siis on võimalik, et nendevaheline seos on suures osas tingitud mõlema sõltuvusest mingist kolmandast q -vektorist $\vec{Z}$. Urimaks Y "puhast" sõltuvust vektorist $\vec{X}$, tuleks vektori $\vec{Z}$ mõju elimineerida. Selleks kasutame järgmist võtet. Olgu juhusliku suuruse Y parim lineaarne prognoos vektori $\vec{Z}$ abil $L(\vec{Z})$. Nimetame juhusliku suuruse Y *prognoosijäägiks* juhuslikku suurust

$$\tilde{Y} = Y - L(\vec{Z}).$$

Analoogiliselt saame määrata ka vektori $\vec{X}$ prognoosijäägi. Olgu $\vec{L}(\vec{Z})$ vektor, mille komponentideks on vektori $\vec{X}$ komponentide parimad lineaarsed prognoosid. Nimetame juhusliku vektori $\vec{X}$ *prognoosijääkide vektoriks* juhuslikku vektorit

$$\tilde{\vec{X}} = \vec{X} - \vec{L}(\vec{Z}).$$

Nende suuruste abil saab defineerida osakorrelatsiooni.

Definitsioon 15.1. Juhusliku suuruse Y osakorrelatsiooniks $r(Y, \vec{X}|\vec{Z})$ juhusliku p -vektoriga $\vec{X}$ juhusliku q -vektori $\vec{Z}$ suhtes nimetatakse mitmest korrelatsioonikordajat prognoosijääkide $\tilde{Y}$ ja $\tilde{\vec{X}}$ vahel,

$$r(Y, \vec{X}|\vec{Z}) = r(\tilde{Y}, \tilde{\vec{X}}).$$

Osakorrelatsiooni arvutuseeskirja tuletamiseks vaatame esmalt üldisemat juhtu. Olgu $\varphi, \psi_1, \psi_2, \dots, \psi_p$ mõõtvad funktsioonid klassist $\mathcal{F}$. Kasutades neid funktsioone juhusliku suuruse Y ja juhusliku vektori $\vec{X}$ prognoosimiseks juhusliku vektori $\vec{Z}$ kaudu, saame prognoosijäägid

$$\tilde{Y} = Y - \varphi(\vec{Z}), \quad \tilde{\vec{X}} = \vec{X} - \vec{\psi}(\vec{Z}),$$

kus $\vec{\psi} = (\psi_1, \dots, \psi_p)'$, $\tilde{X}_i = X_i - \psi_i(\vec{Z})$. Loomulik on nõuda, et parima prognoosi $\varphi(\vec{Z})$ korral kehtiks võrdus $EY = E[\varphi(\vec{Z})]$ ehk $E\tilde{Y} = 0$.

Üldjuhul on regressioonisuhe $\tilde{Y}$ prognoosimisel vektori $\tilde{\vec{X}}$ järgi määratud eeskirjaga

$$\eta_{\mathcal{F}}(\tilde{Y}|\tilde{\vec{X}}) = \sqrt{1 - \frac{Q(\kappa(\tilde{\vec{X}}))}{D\tilde{Y}}}, \quad (15.1)$$

kus $Q(\kappa(\tilde{\vec{X}}))$ on parima prognoosi viga klassis $\mathcal{F}$. Vaatame lähemalt ruutjuure all oleva murru nimetajat,

$$D\tilde{Y} = E(\tilde{Y} - E\tilde{Y})^2 = E[Y - \varphi(\vec{Z}) - (EY - E\varphi(\vec{Z}))]^2 = Q(\varphi(\vec{Z})).$$

Kuna

$$\eta_{\mathcal{F}}(Y|\vec{Z}) = \sqrt{1 - \frac{Q(\varphi(\vec{Z}))}{DY}}$$

siis

$$D\tilde{Y} = DY[1 - \eta_{\mathcal{F}}^2(Y|\vec{Z})]. \quad (15.2)$$

Vaatame nüüd valemis (15.1) ruutjuure all oleva murru lugejat. Moodustame vektoreid $\vec{X}$ ja $\vec{Z}$ ühendades uue $p + q$ -vektori $\vec{U}$,

$$\vec{U} = \begin{bmatrix} \vec{X} \\ \vec{Z} \end{bmatrix}.$$

Prognoosijäägi $Q(\kappa(\tilde{\vec{X}}))$ saame esitada seda uut vektorit kasutades,

$$\begin{aligned} Q(\kappa(\tilde{\vec{X}})) &= E(\tilde{Y} - \kappa(\tilde{\vec{X}}))^2 = E[Y - \varphi(\vec{Z}) - \kappa(\vec{X} - \vec{\psi}(\vec{Z}))]^2 \\ &= E[Y - \tau(\vec{U})]^2 = Q(\tau(\vec{U})). \end{aligned}$$

Regressioonisuhte definitsiooni kohaselt oleme saanud nüüd

$$\eta_{\mathcal{F}}(Y|\vec{U}) = \sqrt{1 - \frac{Q(\tau(\vec{U}))}{DY}}.$$

Avaldame selle kordaja kaudu prognoosijäägi

$$Q(\tau(\vec{U})) = DY[1 - \eta_{\mathcal{F}}^2(Y|\vec{U})]. \quad (15.3)$$

Asendades tulemused (15.2) ja (15.3) regressioonisuhte avaldisse (15.1), saame

$$\eta_{\mathcal{F}}(\tilde{Y}|\tilde{\vec{X}}) = \sqrt{1 - \frac{1 - \eta_{\mathcal{F}}^2(Y|\vec{U})}{1 - \eta_{\mathcal{F}}^2(Y|\vec{Z})}} = \sqrt{\frac{\eta_{\mathcal{F}}^2(Y|\vec{U}) - \eta_{\mathcal{F}}^2(Y|\vec{Z})}{1 - \eta_{\mathcal{F}}^2(Y|\vec{Z})}}.$$

Kuna lineaarse prognoosi korral on regressioonisuhted võrdsed mitmeste korrelatsioonikordajatega, saame lineaarse prognoosi korral valemi

$$r(Y, \vec{X}|\vec{Z}) = \sqrt{\frac{r^2(Y, \vec{U}) - r^2(Y, \vec{Z})}{1 - r^2(Y, \vec{Z})}}. \quad (15.4)$$

Vaatame nüüd kõige lihtsamat võimalikku erijuhtu, kus on tegemist kolme juhusliku suurusega X , Y ja Z . Ühendatud juhuslikuks vektoriks $\vec{U}$ on siis kahedimensionaalne juhuslik vektor

$$\vec{U} = \begin{pmatrix} X \\ Z \end{pmatrix}.$$

Osakorrelatsiooni (15.4) välja kirjutamiseks peame leidma mitmese korrelatsioonikordaja $r(Y, \vec{U})$. Olgu $\vec{\rho}(Y, \vec{U})$ korrelatsioonide vektor Y ja $\vec{U}$ vahel,

$$\vec{\rho}(Y, \vec{U}) = \begin{pmatrix} \rho(Y, X) \\ \rho(Y, Z) \end{pmatrix}$$

ja $\mathbf{P}$ korrelatsioonimaatriks X ja Z vahel,

$$\mathbf{P} = \begin{pmatrix} 1 & \rho(X, Z) \\ \rho(X, Z) & 1 \end{pmatrix}.$$

Siis

$$r^2(Y, \vec{U}) = \vec{\rho}'(Y, \vec{U}) \mathbf{P}^{-1} \vec{\rho}(Y, \vec{U}).$$

Kuna

$$\mathbf{P}^{-1} = \frac{1}{1 - \rho^2(X, Z)} \begin{pmatrix} 1 & -\rho(X, Z) \\ -\rho(X, Z) & 1 \end{pmatrix},$$

siis saame

$$r^2(Y, \vec{U}) = \frac{\rho^2(Y, X) - 2\rho(Y, X)\rho(X, Z)\rho(Y, Z) + \rho^2(Y, Z)}{1 - \rho^2(X, Z)}.$$

Asendades saadud avaldise mitmese korrelatsioonikordaja valemisse (15.4) saame

$$\begin{aligned} r(Y, X|Z) &= \sqrt{\frac{[\rho(Y, X) - \rho(X, Z)\rho(Y, Z)]^2}{(1 - \rho^2(Y, Z))(1 - \rho^2(X, Z))}} \\ &= \frac{\rho(Y, X) - \rho(X, Z)\rho(Y, Z)}{\sqrt{(1 - \rho^2(Y, Z))(1 - \rho^2(X, Z))}}. \end{aligned}$$

Seega kolme juhusliku suuruse korral on X ja Y vaheline osakorrelatsioon Z suhtes avaldatav tavaliste lineaarsete korrelatsioonikordajate kaudu.

16 Statistike asümptootilisest jaotusest

Nagu siamaani oleme näinud, on mitmemõõtmelises statistikas kasutatavad statistikud tavaliselt valimikeskmise $\vec{X}$ ja/või valimi dispersioonimaatriksi $\mathbf{S}$ funktsioonid. Kui üldkogum on normaaljaotusega, $\vec{X} \sim N_p(\vec{\mu}, \mathbf{\Sigma})$, on nende statistikute jaotused ja ka paljude vajalike $\vec{X}$ ja $\mathbf{S}$ funktsioonide jaotus teada. Olukord on täiesti erinev, kui $\vec{X}$ ei ole normaaljaotusega. Sel juhul tavaliselt täpseid jaotusi teada ei ole. Õnneks saab $\vec{X}$ ja $\mathbf{S}$ jaotust kirjeldada ligikaudu laia üldkogumi jaotuste klassi korral, kui valimimaht n on "suur". Mida tähendab - "suur"? Praktikas tähendab see seda, et lähendjaotused muutuvad kasutatavateks, kui $n > 100$, mõnikord ka $n > 50$ korral.

Sellise asümptootilise käitumise taga on asjaolu, et kui $n \rightarrow \infty$, siis nii $\vec{X}$ kui ka $\mathbf{S}$ on kirjeldatavad normaaljaotustega. Nende jaotusteni jõuame järgmise juhuslike vektorite tsentraalse piirteoreemi abil.

Teoreem 16.1 (tsentraalne piirteoreem). *Olgu $\vec{X}_i$, $i = 1, 2, \dots$, sõltumatud sama jaotusega p -vektorid keskväärtusega $E\vec{X}_i = \vec{\mu}$ ja dispersioonimaatriksiga $D\vec{X}_i = \mathbf{\Sigma}$. Kui $n \rightarrow \infty$, siis*

$$\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n \vec{X}_i - \vec{\mu} \right) \xrightarrow{\mathcal{D}} \vec{Z} \sim N_p(\vec{0}, \mathbf{\Sigma}).$$

Tõestus. Teoreemi tõestame karakteristlike funktsioonide abil. Toome sisse järgmise karakteristliku funktsiooni

$$\varphi_n(\vec{t}, u) = E e^{i u \vec{t}' \frac{1}{\sqrt{n}} \sum_{j=1}^n (\vec{X}_j - \vec{\mu})}.$$

Fikseeritud $\vec{t}$ väärtuse korral saame funktsiooni $\varphi_n(\vec{t}, u)$ vaadata kui juhusliku suuruse

$$Y_n = \frac{1}{\sqrt{n}} \sum_{j=1}^n (\vec{t}' \vec{X}_j - \vec{t}' \vec{\mu})$$

karakteristlikku funktsiooni kohal u . Ühemõõtmelise tsentraalse piirteoreemi kohaselt koon-
dub jada $\{Y_n\}$ jaotuse järgi normaaljaotuseks $N(0, \vec{t}' \Sigma \vec{t})$ ja seega

$$\lim_{n \rightarrow \infty} \varphi_n(\vec{t}, u) = e^{-\frac{1}{2} u^2 \vec{t}' \Sigma \vec{t}}$$

iga u ja $\vec{t}$ korral. Võttes $u = 1$, saame

$$\lim_{n \rightarrow \infty} E e^{i \vec{t}' \frac{1}{\sqrt{n}} \sum_{j=1}^n (\vec{X}_j - \vec{\mu})} = e^{-\frac{1}{2} \vec{t}' \Sigma \vec{t}}$$

iga $\vec{t} \in \mathbb{R}^p$ korral. Saime piirväärtuseks normaaljaotuse $N_p(\vec{0}, \Sigma)$ karakteristliku funktsiooni avaldise. Kuna $e^{-\frac{1}{2} \vec{t}' \Sigma \vec{t}}$ on pidev punktis $\vec{t} = \vec{0}$, siis karakteristliku funktsiooni omaduse (iii) põhjal (vt. Lisa 2) on tegemist koondumisega jaotuse järgi ja piirjaotuseks on normaaljaotus $N_p(\vec{0}, \Sigma)$,
mott.

Tõestame nüüd teoreemi, mis kirjeldab valimi keskvaartuse ja valimi dispersioonimaatriksi käitumist suurte valimite korral.

Teoreem 16.2. Olgu $\mathbf{X} = [\vec{X}_1 \ \vec{X}_2 \ \dots \ \vec{X}_n]$ teoreetiline valim üldkogumist, kus tunnusvek-
tori $\vec{X}$ jaotuseks on $P_{\vec{X}}$, $\vec{X} \sim P_{\vec{X}}$. Olgu $E\vec{X} = \vec{\mu}$, $D\vec{X} = \Sigma$ ja eksisteerigu neljandat järku
lõplikud segamomendid, $E(X_i^{r_i} X_j^{r_j} X_k^{r_k} X_l^{r_l}) < \infty$, $r_i, r_j, r_k, r_l = 0, \dots, 4$; $r_i + r_j + r_k + r_l = 4$; $i, j, k, l = 1, \dots, p$. Kui $n \rightarrow \infty$, siis

$$\begin{aligned} \vec{X} &\xrightarrow{P} \vec{\mu}, & \mathbf{S} &\xrightarrow{P} \Sigma; \\ \sqrt{n}(\vec{X} - \vec{\mu}) &\xrightarrow{D} N_p(\vec{0}, \Sigma); \\ \sqrt{n} \text{vec}(\mathbf{S} - \Sigma) &\xrightarrow{D} N_{p^2}(\vec{0}, \mathbf{\Pi}), \end{aligned} \tag{16.1}$$

kus piirjaotuse dispersioonimaatriks $\mathbf{\Pi} : p^2 \times p^2$ koosneb teist ja neljandat järku tsentraal-
setest momentidest,

$$\begin{aligned} \mathbf{\Pi} &= D \text{vec} [(\vec{X} - \vec{\mu})(\vec{X} - \vec{\mu})'] \\ &= E[(\vec{X} - \vec{\mu})(\vec{X} - \vec{\mu})' \otimes ((\vec{X} - \vec{\mu})(\vec{X} - \vec{\mu})')] - \text{vec} \Sigma \text{vec}' \Sigma. \end{aligned}$$

Tõestus. Esimesed kaks väidet on järeldused nõrgast suurte arvude seadusest ja ütlevad, et $\vec{X}$ ja $\mathbf{S}$ on mõjusad hinnangud üldkogumi parameetrite $\vec{\mu}$ ja Σ jaoks.

Tõepoolest, kuna

$$\vec{X} = \frac{1}{n} \sum_{i=1}^n \vec{X}_i,$$

siis nõrga suurte arvude seaduse kohaselt

$$\frac{1}{n} \sum_{i=1}^n \vec{X}_i \xrightarrow{P} E\vec{X} = \vec{\mu}.$$

Dispersioonimaatriksiga on asi pisut keerulisem. Vaatame esmalt statistikut, mis määrab suurima tõepära hinnangu (nihkega hinnangu),

$$\mathbf{S}_* = \frac{1}{n} \sum_{i=1}^n (\vec{X}_i - \vec{X})(\vec{X}_i - \vec{X})'.$$

Muudame sulgudes olevat avaldist liites ja lahutades $\vec{\mu}$:

$$\mathbf{S}_* = \frac{1}{n} \sum_{i=1}^n [(\vec{X}_i - \vec{\mu}) + (\vec{\mu} - \vec{X})][(\vec{X}_i - \vec{\mu}) + (\vec{\mu} - \vec{X})]'$$

Korrutades nüüd kandiilistes sulgudes olevad liikmed, muutes liidetavate järjekorda ja koondades ühesugused liikmed, saame

$$\mathbf{S}_* = \frac{1}{n} \sum_{i=1}^n (\vec{X}_i - \vec{\mu})(\vec{X}_i - \vec{\mu})' - (\vec{X} - \vec{\mu})(\vec{X} - \vec{\mu})'. \quad (16.2)$$

Suurte arvude seaduse tõttu

$$\frac{1}{n} \sum_{i=1}^n (\vec{X}_i - \vec{\mu})(\vec{X}_i - \vec{\mu})' \xrightarrow{P} E[(\vec{X}_i - \vec{\mu})(\vec{X}_i - \vec{\mu})'] = \mathbf{\Sigma}.$$

Samal ajal

$$(\vec{X} - \vec{\mu})(\vec{X} - \vec{\mu})' = \frac{1}{n^2} \sum_{i=1}^n (\vec{X}_i - \vec{\mu}) \sum_{j=1}^n (\vec{X}_j - \vec{\mu})' \xrightarrow{P} \mathbf{0}.$$

Kokkuvõttes saame

$$\mathbf{S}_* \xrightarrow{P} \mathbf{\Sigma}.$$

Siis aga ka

$$\mathbf{S} = \frac{n}{n-1} \mathbf{S}_* \xrightarrow{P} \mathbf{\Sigma},$$

sest $\frac{n}{n-1} \rightarrow 1$, kui $n \rightarrow \infty$.

Nüüd tõestame väited jaotuse järgi koondumisest. Tsentraalse piirteoreemi tõttu

$$\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n \vec{X}_i - \vec{\mu} \right) \xrightarrow{\mathcal{D}} N_p(\vec{0}, \mathbf{\Sigma}).$$

Seega on valimikeskmise käitumine iga lõpliku dispersiooniga jaotuse korral ligikaudu kirjeldatav normaaljaotusega, st $\vec{X}$ on ligikaudu jaotusega $N_p(\vec{\mu}, \frac{1}{n} \mathbf{\Sigma})$.

Ka valimi dispersioonimaatriksi $\mathbf{S}$ asümptootiliseks jaotuseks on normaaljaotus. Vaatame esmalt statistikut $\mathbf{S}_*$. Kasutades esitust (16.2), saame jada $\sqrt{n} \text{vec}(\mathbf{S}_* - \mathbf{\Sigma})$ esitada kujul

$$\begin{aligned} & \sqrt{n} \text{vec} \left[\frac{1}{n} \sum_{i=1}^n (\vec{X}_i - \vec{\mu})(\vec{X}_i - \vec{\mu})' - (\vec{X} - \vec{\mu})(\vec{X} - \vec{\mu})' - \mathbf{\Sigma} \right] \\ &= \sqrt{n} \text{vec} \left[\frac{1}{n} \sum_{i=1}^n (\vec{X}_i - \vec{\mu})(\vec{X}_i - \vec{\mu})' - \mathbf{\Sigma} \right] \\ & \quad - \sqrt{n} \text{vec} [(\vec{X} - \vec{\mu})(\vec{X} - \vec{\mu})']. \end{aligned}$$

Rakendame esimese liidetava korral mitmemõõtmelist tsentraalset piirteoreemi. Tähistame $\mathbf{T}_i = (\vec{X}_i - \vec{\mu})(\vec{X}_i - \vec{\mu})'$, siis teoreemist 16.1

$$\sqrt{n} \text{vec} \left[\frac{1}{n} \sum_{i=1}^n (\vec{X}_i - \vec{\mu})(\vec{X}_i - \vec{\mu})' - \mathbf{\Sigma} \right] \xrightarrow{\mathcal{D}} N_{p^2}(\vec{0}, D\text{vec} \mathbf{T}_i).$$

Asümptootilise jaotuse dispersioonimaatriks $D\text{vec} \mathbf{T}_i$ aga avaldub kujul

$$D\text{vec} \mathbf{T}_i = E[\text{vec} \mathbf{T}_i \text{vec}' \mathbf{T}_i] - E[\text{vec} \mathbf{T}_i] E[\text{vec}' \mathbf{T}_i].$$

Kuna $E\text{vec} \mathbf{T}_i = \text{vec} \mathbf{\Sigma}$ ja vec -operaatori omaduse (L1.16) abil

$$\text{vec} \mathbf{T}_i \text{vec}' \mathbf{T}_i = [(\vec{X}_i - \vec{\mu}) \otimes (\vec{X}_i - \vec{\mu})][(\vec{X}_i - \vec{\mu})' \otimes (\vec{X}_i - \vec{\mu})'],$$

saame otsekorrutise omadust (L1.15) rakendades

$$\begin{aligned} D\text{vec}\mathbf{T}_i &= E\{[(\vec{X}_i - \vec{\mu})(\vec{X}_i - \vec{\mu})'] \otimes [(\vec{X}_i - \vec{\mu})(\vec{X}_i - \vec{\mu})']\} \\ &\quad - \text{vec}\mathbf{\Sigma} \text{vec}'\mathbf{\Sigma} = \mathbf{\Pi}. \end{aligned}$$

Lemma L2.1 põhjal teame, et kui üks juhuslike vektorite jada koondub jaotuse järgi, $\vec{X}_n \xrightarrow{\mathcal{D}} \vec{X}$, ja teine tõenäosuse järgi, $\vec{Y}_n \xrightarrow{P} \vec{0}$, siis summa koondub jaotuse järgi, $\vec{X}_n + \vec{Y}_n \xrightarrow{\mathcal{D}} \vec{X}$. Seega väite (16.1) tõestamiseks tuleb veel näidata, et teine liidetav jada esituses koondub tõenäosuse järgi nulliks,

$$\sqrt{n} \text{vec} [(\vec{X} - \vec{\mu})(\vec{X} - \vec{\mu})'] \xrightarrow{P} \vec{0}. \quad (16.3)$$

Selles jadas on tegemist kahe teguri korrutisega, kusjuures

$$\vec{X} - \vec{\mu} \xrightarrow{P} \vec{0}$$

ja

$$\sqrt{n}(\vec{X} - \vec{\mu}) \xrightarrow{\mathcal{D}} N(\vec{0}, \mathbf{\Sigma}).$$

Jällegi, rakendades lemmat L2.1 saame, et kui üks tegur koondub jaotuse järgi, $\vec{X}_n \xrightarrow{\mathcal{D}} \vec{X}$, ja teine tegur tõenäosuse järgi, $\vec{Y}_n \xrightarrow{P} \vec{0}$, siis sellest järeldub koondumine tõenäosuse järgi, $\vec{X}_n \vec{Y}_n' \xrightarrow{P} \mathbf{0}$, seega kehtib koonduvus (16.3). Siis aga oleme näidanud ka väite (16.1) kehtivuse, mott.

Kokkuvõttes saame öelda, et valimikeskmise ja valimi dispersioonimaatriksi käitumine on ka mittenormaaljaotusega üldkogumi korral kirjeldatav normaaljaotusega, st valimikeskmist kirjeldab asümptootiline normaaljaotus $N_p(\vec{\mu}, \frac{1}{n}\mathbf{\Sigma})$ ja vektorit $\text{vec}\mathbf{S}$ normaaljaotus $N_{p^2}(\text{vec}\mathbf{\Sigma}, \frac{1}{n}\mathbf{\Pi})$, kui üldkogumi jaotusel eksisteerivad neljandat järku segamomendid. Normaaljaotusega üldkogumi korral avaldub dispersioonimaatriksi asümptootiline jaotus üldkogumi dispersioonimaatriksi $\mathbf{\Sigma}$ kaudu.

Järeldus 16.2.1. *Olgu valim üldkogumist jaotusega $N_p(\mu, \mathbf{\Sigma})$. Siis valimi dispersioonimaatriksi jaoks leiab aset koonduvus*

$$\sqrt{n}\text{vec}(\mathbf{S} - \mathbf{\Sigma}) \xrightarrow{\mathcal{D}} N_{p^2}(\text{vec}\mathbf{0}_{p \times p}, \mathbf{\Pi}_N),$$

kus

$$\mathbf{\Pi}_N = (\mathbf{I}_{p^2} + \mathbf{K}_{p,p})(\mathbf{\Sigma} \otimes \mathbf{\Sigma}),$$

ja $\mathbf{K}_{p,p}$ on kommuteerimismaatriks.

Tõestus. Väide järeldub teoreemi 6.1 rakendamisel maatriksi $\mathbf{\Pi}$ avaldisele valemis (16.1), mott.

Tõestame veel teoreemi, mis võimaldab leida põhistatistikute funktsioonide asümptootilisi jaotusi.

Teoreem 16.3. *Olgu jada $\vec{Z}_n$ asümptootilise normaaljaotusega,*

$$\sqrt{n}(\vec{Z}_n - \vec{a}) \xrightarrow{\mathcal{D}} N_p(\vec{0}, \mathbf{\Sigma}),$$

kusjuures

$$\vec{Z}_n \xrightarrow{P} \vec{a}.$$

Olgu funktsioonil $\vec{g}(\vec{x}) : \mathbb{R}^p \rightarrow \mathbb{R}^q$ pidevad osatuletised mingis punkti $\vec{a}$ ümbruses. Siis

$$\sqrt{n} [\vec{g}(\vec{Z}_n) - \vec{g}(\vec{a})] \xrightarrow{\mathcal{D}} N_q(\vec{0}, \xi \mathbf{\Sigma} \xi'), \quad (16.4)$$

kus $\xi : q \times p$ on maatrikstuletis:

$$\xi = \left. \frac{d\vec{g}(\vec{Z}_n)}{d\vec{Z}_n} \right|_{\vec{Z}_n = \vec{a}} \neq \mathbf{0}.$$

Tõestus. Koonduvuse (16.4) näitamiseks tõestame, et karakteristiklik funktsioon $\varphi_{\sqrt{n}[\vec{g}(\vec{Z}_n) - \vec{g}(\vec{a})]}(\vec{t})$ koondub vastavaks normaaljaotuse karakteristiklikuks funktsiooniks. Eelduse kohaselt

$$\lim_{n \rightarrow \infty} \varphi_{\sqrt{n}(\vec{Z}_n - \vec{a})}(\vec{t}) = \exp(-\frac{1}{2}\vec{t}'\Sigma\vec{t}), \quad \vec{t} \in \mathbb{R}^p.$$

Arendame $\vec{g}(\vec{Z}_n)$ Taylori ritta punktis $\vec{Z}_n = \vec{a}$,

$$\vec{g}(\vec{Z}_n) = \vec{g}(\vec{a}) + \left. \frac{d\vec{g}(\vec{Z}_n)}{d\vec{Z}_n} \right|_{\vec{Z}_n = \vec{a}} (\vec{Z}_n - \vec{a}) + \dots$$

Edasises esituses kasutame reaksarenduses suurust $o_p(\varepsilon_i)$. Tuletame meelde, et kirjpilt $o_p(\varepsilon_i)$ tähistab järgmist vahekorda juhuslike vektorite jada $\{\vec{X}_i\}$ ja arvjada $\{\varepsilon_i\}$ vahel (vt. lisa 2):

$$\frac{\vec{X}_i}{\varepsilon_i} \xrightarrow{P} \vec{0}.$$

Rakendades teoreemi L2.1 $\varepsilon_n = \frac{1}{\sqrt{n}}$ korral saame

$$\begin{aligned} \sqrt{n} [\vec{g}(\vec{Z}_n) - \vec{g}(\vec{a})] &= \left. \frac{d\vec{g}(\vec{Z}_n)}{d\vec{Z}_n} \right|_{\vec{Z}_n = \vec{a}} \sqrt{n}(\vec{Z}_n - \vec{a}) + o_p\left(\frac{1}{\sqrt{n}}\right) \\ &= \vec{\xi}\sqrt{n}(\vec{Z}_n - \vec{a}) + o_p\left(\frac{1}{\sqrt{n}}\right). \end{aligned}$$

Seetõttu

$$\begin{aligned} \lim_{n \rightarrow \infty} \varphi_{\sqrt{n}[\vec{g}(\vec{Z}_n) - \vec{g}(\vec{a})]}(\vec{t}) &= \lim_{n \rightarrow \infty} \varphi_{\vec{\xi}\sqrt{n}(\vec{Z}_n - \vec{a}) + o_p(\frac{1}{\sqrt{n}})}(\vec{t}) \\ &= \lim_{n \rightarrow \infty} \varphi_{\sqrt{n}(\vec{Z}_n - \vec{a})}(\vec{\xi}\vec{t}) = \exp(-\frac{1}{2}\vec{t}'\vec{\xi}\Sigma\vec{\xi}\vec{t}), \end{aligned}$$

mott.

Seega, kui meil on funktsioon valimikeskmisest $\vec{g}(\vec{X})$, mis rahuldab teoreemi eeldusi, siis selle funktsiooni tõenäosuslikku käitumist saame kirjeldada normaaljaotusega. Tähistame

$$\xi_{\vec{X}} = \left. \frac{d\vec{g}(\vec{X})}{d\vec{X}} \right|_{\vec{X} = \vec{\mu}}. \text{ Vastavalt tõestatud teoreemile}$$

$$\sqrt{n}[\vec{g}(\vec{X}) - \vec{g}(\vec{\mu})] \xrightarrow{D} N_q(\vec{0}, \xi_{\vec{X}}\Sigma\xi'_{\vec{X}})$$

ehk

$$\vec{g}(\vec{X}) \approx N_q(\vec{g}(\vec{\mu}), \frac{1}{n}\xi_{\vec{X}}\Sigma\xi'_{\vec{X}}),$$

kui $\xi_{\vec{X}} \neq \vec{0}$.

Sama kehtib ka valimi dispersioonimaatriksi funktsiooni kohta. Kui meil on funktsioon valimi dispersioonimaatriksist, $\vec{h}(\text{vec } \mathbf{S})$, siis selle funktsiooni tõenäosuslikku käitumist saame kirjeldada normaaljaotusega. Tähistame $\xi_{\Sigma} = \left. \frac{d\vec{h}(\mathbf{S})}{d\text{vec } \mathbf{S}} \right|_{\mathbf{S} = \Sigma} \neq \vec{0}$. Vastavalt tõestatud teoreemile

$$\sqrt{n}[\vec{h}(\text{vec } \mathbf{S}) - \vec{h}(\text{vec } \Sigma)] \xrightarrow{D} N_q(\vec{0}, \xi_{\Sigma}\Pi\xi'_{\Sigma})$$

ehk

$$\vec{h}(\text{vec } \mathbf{S}) \approx N_q(\vec{h}(\text{vec } \Sigma), \frac{1}{n}\xi_{\Sigma}\Pi\xi'_{\Sigma}).$$

Illustreerimaks tõestatud tulemust leiame üldistatud dispersiooni $|\Sigma|$ hinnangu jaoks asümptootilise normaaljaotuse. Teoreemi 16.3 põhjal leiab aset koondumine normaaljaotuseks:

$$\sqrt{n}(|\mathbf{S}| - |\Sigma|) \xrightarrow{D} N(0, \xi\Pi\xi'),$$

kus $\xi = \frac{d|\mathbf{S}|}{d\mathbf{S}}|_{\mathbf{S}=\mathbf{\Sigma}}$. Kasutades determinandi tuletise valemit (vt. lisa 1, tuletise omadus (xvi)) saame

$$\xi = |\mathbf{\Sigma}| \text{vec}' \mathbf{\Sigma}^{-1}.$$

Siis asümptootilise normaaljaotuse dispersiooni $\sigma_{|\mathbf{S}|}^2$ jaoks saame võrduse:

$$\sigma_{|\mathbf{S}|}^2 = |\mathbf{\Sigma}|^2 (\text{vec}' \mathbf{\Sigma}^{-1} \mathbf{M} \text{vec} \mathbf{\Sigma}^{-1} - \text{vec}' \mathbf{\Sigma}^{-1} \text{vec} \mathbf{\Sigma} \text{vec}' \mathbf{\Sigma} \text{vec} \mathbf{\Sigma}^{-1}),$$

kus

$$\mathbf{M} = E[(\vec{X} - \vec{\mu})(\vec{X} - \vec{\mu})'] \otimes ((\vec{X} - \vec{\mu})(\vec{X} - \vec{\mu})').$$

Siis

$$\sigma_{|\mathbf{S}|}^2 = |\mathbf{\Sigma}|^2 [\text{vec}' \mathbf{\Sigma}^{-1} \mathbf{M} \text{vec} \mathbf{\Sigma}^{-1} - p^2].$$

Saadud avaldis lihtsustub normaaljaotusega üldkogumi korral. Sel juhul leiab järelduse 16.2.1 põhjal maatriksi $\mathbf{S}$ jaoks aset koonduvus:

$$\sqrt{n} \text{vec}(\mathbf{S} - \mathbf{\Sigma}) \xrightarrow{\mathcal{D}} N_{p^2}(\vec{0}, \mathbf{\Pi}_N),$$

kus

$$\mathbf{\Pi}_N = (\mathbf{I}_{p^2} + \mathbf{K}_{p,p})(\mathbf{\Sigma} \otimes \mathbf{\Sigma}).$$

Üldistatud dispersiooni $|\mathbf{S}|$ asümptootiline dispersioon $\sigma_{|\mathbf{S}|}^2$ on siis kujul

$$\sigma_{|\mathbf{S}|}^2 = 2p|\mathbf{\Sigma}|^2.$$

Põhistatistike asümptootilise normaaljaotuse muudab hüpoteeside kontrolli ülesannete ja usaldushulkade konstrueerimise jaoks eriti väärtuslikuks asjaolu, et asümptootilise normaaljaotuse baasil saab konstrueerida χ^2 -statistikuid. Meetod pärineb artiklist Wald (1943) ning seda tuntakse kui Waldi meetodit. Olgu $\vec{T}_n$ statistik, mille väärtused on hinnanguteks parameetritele $\vec{\theta}$, kusjuures, kui $\vec{\theta} = \vec{\theta}_0$ ja $n \rightarrow \infty$, siis

$$\sqrt{n}(\vec{T}_n - \vec{\theta}_0) \xrightarrow{\mathcal{D}} N_m(\vec{0}, \mathbf{\Omega}). \quad (16.5)$$

Siin $\mathbf{\Omega} : m \times m$ on täisastakuga maatriks, mis võib sõltuda ka suurusest $\vec{\theta}_0$. Kui leidub statistikute jada maatriksi $\mathbf{\Omega}$ hindamiseks $\{\mathbf{\Omega}_n\}$ nii, et $\mathbf{\Omega}_n \xrightarrow{P} \mathbf{\Omega}$, kui $n \rightarrow \infty$ ja $\vec{\theta} = \vec{\theta}_0$, siis

$$n(\vec{T}_n - \vec{\theta}_0)' \mathbf{\Omega}_n^{-1} (\vec{T}_n - \vec{\theta}_0) \xrightarrow{\mathcal{D}} \chi^2(m). \quad (16.6)$$

Seega saab koonduvuses (16.6) esinevat ruutvormi kasutada test-statistikuna testimaks hüpoteesi $H_0 : \vec{\theta} = \vec{\theta}_0$, kuna nullhüpoteesi kehtimisel on tegemist asümptootilise χ^2 -jaotusega. Tulemus (16.6) on järeldus teoreemist 4.1 normaaljaotusega juhusliku vektori ruutvormi jaotusest. Osutub, et Waldi meetod on edasiarendatav ka juhule, kui piirjaotuseks on kõdunud mitmemõõtmeline normaaljaotus (Moore (1977)). Sellise olukorraga oli meil tegemist näiteks teoreemis 16.2 koonduvuse (16.1) korral, kus piirjaotuse dispersioonimaatriksi $\mathbf{\Pi} : p^2 \times p^2$ astak on maatriksite $\mathbf{S}$ ja $\mathbf{\Sigma}$ sümmeetria tõttu maksimaalselt $\frac{1}{2}p(p+1)$. Esitame Waldi meetodi üldistuse järgmise teoreemina.

Teoreem 16.4. *Kehtigu ülaltoodud tähistustes koonduvus (16.5), kusjuures $r(\mathbf{\Omega}) = k$. Eeldame, et $\mathbf{\Omega}_n \xrightarrow{P} \mathbf{\Omega}$, kui $n \rightarrow \infty$ ja $\vec{\theta} = \vec{\theta}_0$ ning et iga n korral $\vec{T}_n - \vec{\theta}_0$ kuulub maatriksi $\mathbf{\Omega}_n$ veeruvektorite poolt määratud alamruumi $\mathcal{C}(\mathbf{\Omega})$. Siis*

(i) *statistik $n(\vec{T}_n - \vec{\theta}_0)' \mathbf{\Omega}_n^{-1} (\vec{T}_n - \vec{\theta}_0)$ on invariantne üldistatud pöördmaatriksi $\mathbf{\Omega}_n^{-1}$ valiku suhtes;*

(ii) *kui $\vec{\theta} = \vec{\theta}_0$, siis*

$$n(\vec{T}_n - \vec{\theta}_0)' \mathbf{\Omega}_n^{-1} (\vec{T}_n - \vec{\theta}_0) \xrightarrow{\mathcal{D}} \chi^2(k).$$

Waldi meetod ja selle üldistus annavad meile võimaluse teoreemide 16.2 ja 16.3 põhjal konstrueerida ligikaudseid usalduspiirkondi ja kontrollida hüpoteese valimikeskmise ja dispersiooni-maatriksi ning nende funktsioonide kohta ka siis, kui üldkogum ei ole normaaljaotusega.

LISA 1

L1.1 Maatriksalgebra mõisteid ja tulemusi

Maatriksalgebra annab vahendid mitmemõõtmelise statistika esitamiseks kompaktsel ja ülevaatlikul kujul. Üleminek maatrikskeelele cá 50 aastat tagasi kiirendas ja süvendas mitmemõõtmelise statistika arengut oluliselt. Samas on olnud tegu vastastikku kasuliku arenguga. Statistikaprobleemid on viinud ka maatriksite teooria uute arendusteni viimastel aastakümnetel. Järgnev lühiülevaade on kokkusurutud esitus olulisematest mõistetest ja tulemustest, mida selles õpikus kasutame. Põhjalikuma käsitlemise huvilistele olgu siinkohal antud mõned viited kirjandusele. Toome välja just statistikasuunitlusega maatriksalgebra raamatud. Suhteliselt lihtsama käsitlemise võib leida raamatust Searle (1982). 1990-ndate lõpus ilmus mitu head ja põhjalikku statistikutele mõeldud maatriksite teooria esitust: Harville (1997), Schott (1997), Rao and Rao (1998), Magnus and Neudecker (1999). Põhjalik käsitus on antud ka raamatutes Kollo (1991) ja Kollo and von Rosen (2005). Kokkuvõtlikult on olulisemad maatriksalgebra mõisted ja tulemused ära toodud klassikalistes mitmemõõtmelise analüüsi monograafiates: Anderson (2003), Srivastava and Khatri (1979), Rao (1973), Muirhead (1982), Siotani et al. (1985) jt. Põhimõistete ja -tulemuste meeldetuletamiseks on kindlasti kasu ka eestikeelsest õpikust Kilp (2005).

Järgnev materjal on jagatud kaheks: esimeses osas on ülevaade maatriksalgebra põhimõistetest ja vajalikest tulemustest, teine osa sisaldab kokkuvõtte plokomaatriksitest.

Põhitehted

Maatriks $\mathbf{A}$ mõõtmetega $m \times n$ on riskülikukujuline tabel elementidest a_{ij} , $i = 1, \dots, m$; $j = 1, \dots, n$:

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix},$$

kus element a_{ij} paikneb tabeli i -ndas reas ja j -ndas veerus. Elemendi a_{ij} tähisena kasutame ka tähistust $(\mathbf{A})_{ij}$. Maatriksi elementideks võivad olla erinevad matemaatilised objektid: reaalarvud, kompleksarvud, funktsioonid jne. Tekstis kasutame paralleelselt järgmisi fraase:

- maatriks $\mathbf{A}$ mõõtmetega $m \times n$,
- $m \times n$ -maatriks $\mathbf{A}$,
- $\mathbf{A} : m \times n$.

Kui $m = n$, siis on $\mathbf{A}$ ruutmaatriks. Kui $\mathbf{A}$ on $m \times 1$ -maatriks, nimetatakse teda vektoriks. Sel juhul jäetakse teine indeks kirjutamata ja kasutatakse kirja pilti

$$\vec{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}.$$

Maatriksid $\mathbf{A}$ ja $\mathbf{B}$ on võrdsed, $\mathbf{A} = \mathbf{B}$, kui nende mõõtmed on võrdsed ja vastavad elemendid on võrdsed. Maatriksit, mille elemendid on võrdsed arvuga 1 tähistame $\mathbf{1}$, vajadusel $\mathbf{1}_{m \times n}$, analoogiliselt ka nullidest koosneva maatriksi jaoks kasutame tähistusi $\mathbf{0}$ või $\mathbf{0}_{m \times n}$.

Maatriksi $\mathbf{A} : m \times n$ korrutis skalaariga c on $m \times n$ maatriks $c\mathbf{A}$, kus

$$(c\mathbf{A})_{ij} = ca_{ij}.$$

Skalaari all mõistetakse maatriksi elementidega sama tüüpi suurust.

Maatriksite $\mathbf{A} : m \times n$ ja $\mathbf{B} : m \times n$ summa $\mathbf{A} + \mathbf{B}$ on $m \times n$ -maatriks elementidega

$$(\mathbf{A} + \mathbf{B})_{ij} = a_{ij} + b_{ij}.$$

Maatriksite liitmise ja skalaariga korrutamise olulisemad omadused on järgmised:

$$\begin{aligned}\mathbf{A} + \mathbf{B} &= \mathbf{B} + \mathbf{A}; \\ (\mathbf{A} + \mathbf{B}) + \mathbf{C} &= \mathbf{A} + (\mathbf{B} + \mathbf{C}); \\ \mathbf{A} + (-1)\mathbf{A} &= \mathbf{0}; \\ (c_1 + c_2)\mathbf{A} &= c_1\mathbf{A} + c_2\mathbf{A}; \\ c(\mathbf{A} + \mathbf{B}) &= c\mathbf{A} + c\mathbf{B}; \\ c_1(c_2\mathbf{A}) &= (c_1c_2)\mathbf{A}.\end{aligned}$$

Maatriksite korrutamine on võimalik, kui esimese maatriksi veergude arv võrdub teise maatriksi ridade arvuga. Maatriksite $\mathbf{A} : m \times n$ ja $\mathbf{B} : n \times r$ korrutis $\mathbf{C} = \mathbf{AB}$ on $m \times r$ -maatriks, kus

$$c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}.$$

Maatriksite korrutamine ei ole kommutatiivne, küll aga kehtivad seosed

$$\begin{aligned}\mathbf{A}(\mathbf{BC}) &= (\mathbf{AB})\mathbf{C}; \\ \mathbf{A}(\mathbf{B} + \mathbf{C}) &= \mathbf{AB} + \mathbf{AC}; \\ (\mathbf{A} + \mathbf{B})\mathbf{C} &= \mathbf{AC} + \mathbf{BC}.\end{aligned}$$

Maatriksi $\mathbf{A} : m \times n$ transponeeritud maatriks $\mathbf{A}'$ on $n \times m$ -maatriks, kus

$$(\mathbf{A}')_{ji} = a_{ij}, \quad i = 1, \dots, m, \quad j = 1, \dots, n.$$

Maatriksi transponeerimine on seotud liitmise ja korrutamisega järgmiselt:

$$\begin{aligned}(\mathbf{A}')' &= \mathbf{A}; \\ (\mathbf{A} + \mathbf{B})' &= \mathbf{A}' + \mathbf{B}'; \\ (\mathbf{AB})' &= \mathbf{B}'\mathbf{A}'.\end{aligned}$$

Erikujulised maatriksid

Ühikelemendi rolli etendab maatriksite hulgas *ühikmaatriks* $\mathbf{I}_n$:

$$\mathbf{I}_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} = (\delta_{ij}), \quad i, j = 1, \dots, n,$$

kus δ_{ij} on Kroneckeri delta:

$$\delta_{ij} = \begin{cases} 1, & \text{kui } i = j; \\ 0, & \text{kui } i \neq j. \end{cases}$$

Ühikmaatriks rahuldab võrdusi $\mathbf{I}_m\mathbf{A} = \mathbf{AI}_n = \mathbf{A}$, kui $\mathbf{A}$ on $m \times n$ -maatriks.

Ruutmaatriks $\mathbf{A}$ on *sümmeetriline*, kui

$$\mathbf{A}' = \mathbf{A}$$

ja *kaldsümmeetriline*, kui

$$\mathbf{A} = -\mathbf{A}'.$$

Reaalarvuline ruutmaatriks $\mathbf{A} : n \times n$ on *ortogonaalne*, kui

$$\mathbf{AA}' = \mathbf{I}_n$$

ja *idempotentne*, kui

$$\mathbf{A}^2 = \mathbf{A}.$$

Ruutmaatriksi $\mathbf{A} : n \times n$ on *ülemine kolmnurkmaatriks*, kui tema allpool peadiagonaali asetsevad elemendid võrduvad nulliga:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Kui ruutmaatriksi $\mathbf{A} : n \times n$ ülalpool peadiagonaali paiknevad elemendid on võrdsed nulliga, on tegemist *alumise kolmnurkmaatriksiga*.

Ruutmaatriksi diagonaliseerimine on operatsioon, mis asendab $\mathbf{A} : n \times n$ väljaspool peadiagonaali asetsevad elemendid nullidega. *Diagonaliseeritud maatriksi* tähistame $\mathbf{A}_d$:

$$\mathbf{A}_d = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Vektorist $\vec{a}$ moodustatud diagonaalmaatriks on sarnase tähistusega:

$$\vec{a}_d = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}.$$

Determinant ja pöördmaatriks

Ruutmaatriksi $\mathbf{A} : n \times n$ korral on tähtsaks karakteristikuks tema *determinant* $|\mathbf{A}|$:

$$|\mathbf{A}| = \sum_{j_1, \dots, j_n} (-1)^{N(j_1, \dots, j_n)} \prod_{i=1}^n a_{ij_i}, \quad (\text{L1.1})$$

kus summeeritakse üle arvude $1, 2, \dots, n$ kõigi erinevate permutatsioonide $(j_1, \dots, j_n)$ ja $N(j_1, \dots, j_n)$ on permutatsiooni $(j_1, \dots, j_n)$ inversioonide arv. Arvud j ja i moodustavad inversiooni, kui j paikneb permutatsioonis eespool kui i ja $j > i$. Determinandi arvutamiseks on valemit (L1.1) tülikas rakendada vähegi suuremate n väärtuste korral. Üks võimalus arvutuste lihtsustamiseks on *miinori* mõiste rakendamine. Elemendi a_{ij} *täiendmiinoriks* nimetatakse $\mathbf{A} : n \times n$ elementidest moodustatud $(n-1) \times (n-1)$ -maatriksi $\mathbf{A}_{(ij)}$ determinanti. Maatriks $\mathbf{A}_{(ij)}$ saadakse maatriksist $\mathbf{A}$ selle i -nda rea ja j -nda veeru eraldamise teel. Täiendmiinorite abil saab maatriksi $\mathbf{A}$ determinandi esitada arendusena i -nda rea või j -nda veeru kaudu:

$$|\mathbf{A}| = \sum_{j=1}^n a_{ij} (-1)^{i+j} |\mathbf{A}_{(ij)}| \quad \text{iga } i \text{ korral}, \quad (\text{L1.2})$$

$$|\mathbf{A}| = \sum_{i=1}^n a_{ij} (-1)^{i+j} |\mathbf{A}_{(ij)}| \quad \text{iga } j \text{ korral}. \quad (\text{L1.3})$$

Maatriksi $\mathbf{A}$ esimesest r reast ja veerust moodustatud $r \times r$ -alammaatriksi determinanti nimetatakse r -järku *nurgamiinoriks*. Esitame järgnevalt mõned determinandi olulisemad omadused:

$$|\mathbf{A}'| = |\mathbf{A}|; \quad (\text{L1.4})$$

$$|\mathbf{AB}| = |\mathbf{A}| \cdot |\mathbf{B}|; \quad (\text{L1.5})$$

$$|\mathbf{I}_m + \mathbf{AB}| = |\mathbf{I}_n + \mathbf{BA}|, \quad (\text{L1.6})$$

kus $\mathbf{A}$ on $m \times n$ -maatriks ja $\mathbf{B}$ on $n \times m$ -maatriks.

Kui $|\mathbf{A}| \neq 0$, siis öeldakse, et maatriks $\mathbf{A} : n \times n$ on *pööratav e. regulaarne*, vastasel juhul, kui $|\mathbf{A}| = 0$, on meil tegemist *singulaarse* maatriksiga. Kui $|\mathbf{A}| \neq 0$, siis leidub maatriksil $\mathbf{A}$ pöördmaatriks $\mathbf{A}^{-1}$, mis on määratud võrdusega

$$\mathbf{AA}^{-1} = \mathbf{I}_n.$$

Pöördmaatriksil on järgmised omadused.

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n;$$

$$(\mathbf{A}^{-1})' = (\mathbf{A}')^{-1}; \quad (\text{L1.7})$$

$$|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1}; \quad (\text{L1.8})$$

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}.$$

Pöördmaatriksi üdelement avaldub valemiga

$$(\mathbf{A}^{-1})_{ij} = \frac{(-1)^{i+j}|\mathbf{A}_{(ji)}|}{|\mathbf{A}|}. \quad (\text{L1.9})$$

Kahe maatriksi summa pöördmaatriksi avaldis on keerulisem.

Teoreem L1.1 (binomiaalne pöördteoreem). *Olgu järgnevas võrduses maatriksid $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ ja $\mathbf{D}$ sobivate dimensioonidega ja eksisteerigu seal esinevad pöördmaatriksid. Siis*

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{DA}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{DA}^{-1}.$$

Astak ja jälg Vektorid $\vec{x}_i$, $i = 1, \dots, r$ on *lineaarselt sõltumatud*, kui

$$\sum_{i=1}^r c_i \vec{x}_i = \vec{0},$$

parajasti siis, kui $c_i = 0$, $i = 1, \dots, r$.

Definitsioon L1.1. *Maatriksi $\mathbf{A} : m \times n$ astakuks $r(\mathbf{A})$ nimetatakse tema lineaarselt sõltumatute veergude maksimaalset arvu.*

Loetleme alljärgnevalt astaku olulisemad omadused:

- (i) $r(\mathbf{A}) = r(\mathbf{A}')$;
- (ii) $r(\mathbf{A}) \leq \min(m, n)$;
- (iii) $r(\mathbf{AB}) \leq \min(r(\mathbf{A}), r(\mathbf{B}))$;
- (iv) $r(\mathbf{A} + \mathbf{B}) \leq r(\mathbf{A}) + r(\mathbf{B})$;
- (v) kui $\mathbf{A}$ ja $\mathbf{C}$ on pööratavad, siis $r(\mathbf{ABC}) = r(\mathbf{B})$;
- (vi) kui $\mathbf{A} : m \times n$ ja $\mathbf{B}$ rahuldavad võrdust $\mathbf{AB} = \mathbf{0}$, siis $r(\mathbf{B}) \leq n - r(\mathbf{A})$.

Definitsioon L1.2. *Maatriksi $\mathbf{A} : n \times n$ jäljeks $\text{tr}\mathbf{A}$ nimetatakse tema peadiagonaali elementide summat*

$$\text{tr}\mathbf{A} = \sum_{i=1}^n a_{ii}.$$

Esitame jälje olulisemad omadused reaalarvuliste maatriksite korral:

- (i) $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$, kus $\mathbf{A} : m \times n$ ja $\mathbf{B} : n \times m$;
- (ii) $\text{tr}(\mathbf{A}) = \text{tr}(\mathbf{B}^{-1}\mathbf{AB})$, kui $\mathbf{B}$ on pööratav;
- (iii) $\text{tr}(\mathbf{A}) = \text{tr}(\mathbf{C}'\mathbf{AC})$, kui $\mathbf{C}$ on ortogonaalne;
- (iv) $\text{tr}(c_1\mathbf{A} + c_2\mathbf{B}) = c_1\text{tr}\mathbf{A} + c_2\text{tr}\mathbf{B}$;
- (v) $\text{tr}(\mathbf{A}) = r(\mathbf{A})$, kui $\mathbf{A}$ on idempotentne;
- (vi) $\text{tr}(\mathbf{A}'\mathbf{A}) = 0$ siis ja ainult siis, kui $\mathbf{A} = \mathbf{0}$;
- (vii) $\text{tr}\mathbf{A} = \text{tr}(\mathbf{A}')$;
- (viii) $\text{tr}(\mathbf{A}'\mathbf{A}) = \text{tr}(\mathbf{AA}') = \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2$;

$$(ix) \quad \vec{x}' \mathbf{A} \vec{x} = \text{tr}(\mathbf{A} \vec{x} \vec{x}');$$

$$(x) \quad r(\mathbf{A}) \geq \frac{(\text{tr} \mathbf{A})^2}{\text{tr}(\mathbf{A}^2)}, \text{ kui } \mathbf{A} \text{ on sümmeetriline.}$$

Omapäärtused, omavektorid

Olgu $\mathbf{A} : n \times n$ reaalarvuline maatriks. Vaatame seda kui lineaarteisendust ruumis $\mathbb{R}^n$, mis teisendab suvalise vektori $\vec{x}$ teiseks vektoriks $\vec{y}$,

$$\mathbf{A} \vec{x} = \vec{y}.$$

Vektorid, mille siht säilib selle teisendusega, on määratud võrrandiga

$$\mathbf{A} \vec{x} = \lambda \vec{x}.$$

Sel võrrandil on mittetriviaalne lahend, kui

$$|\mathbf{A} - \lambda \mathbf{I}_n| = 0. \quad (\text{L1.10})$$

Võrrandit (10) nimetatakse maatriksi $\mathbf{A}$ *karakteristlikuks võrrandiks* ja selle lahendeid λ_i maatriksi $\mathbf{A}$ *omaväärtusteks*. Vektorit $\vec{x}_i \neq \vec{0}$, mis rahuldab võrdust

$$\mathbf{A} \vec{x}_i = \lambda_i \vec{x}_i$$

nimetatakse omaväärtusele λ_i vastavaks *omavektoriks*. Determinandi definitsioonist järeldub, et karakteristlik võrrand on n -astme polünoom λ suhtes, polünoomil on aga n juurt, mis võivad olla nii reaalsed kui kompleksed, kusjuures nende hulgas võib olla kordseid. Seega on reaalarvulisel maatriksil n omaväärtust, mis võivad olla nii reaalarvulised kui kompleksarvud. Omaväärtustele vastavad omavektorid ei ole üheselt määratud. Kui $\vec{x}_i$ on omavektor, siis varieerides tema pikkust ja/või suunda saame jälle vektori, mis rahuldab omavektori nõudeid. Pikkuse osas lepitakse tavaliselt kokku, et vaadeldakse normeeritud omavektoreid pikkusega üks. Suuna fikseerimine on vähem oluline, vajadusel määratakse selleks kindlaks ühe mittenullilise koordinaadi märk (+ või -). Esitame järgnevalt olulisemad omaväärtuste ja omavektorite omadused.

- (i) Reaalarvulise sümmeetrilise maatriksi kõik omaväärtused λ_i on reaalarvulised ja neile vastavad omavektorid $\vec{x}_i$ saab valida reaalarvuliste koordinaatidega.
- (ii) $\mathbf{A}$ ja $\mathbf{A}'$ on samade omaväärtustega.
- (iii) Maatriksid $\mathbf{A}$ ja $\mathbf{C} \mathbf{A} \mathbf{C}^{-1}$ on samade omaväärtustega suvalise pööratava maatriksi $\mathbf{C}$ korral.
- (iv) Kui $\mathbf{A}$ on sümmeetriline maatriks omaväärtustega $\lambda_i \neq \lambda_j$, siis neile omaväärtustele vastavad omavektorid on ortogonaalsed.
- (v) Maatriksi $\mathbf{A}$ erinevatele omaväärtustele vastavad omavektorid on lineaarselt sõltumatud.
- (vi) Olgu $\mathbf{A}$ ja $\mathbf{B}$ $n \times n$ -maatriksid, kusjuures $\mathbf{A}$ on pööratav. Siis $\mathbf{A} \mathbf{B}$ ja $\mathbf{B} \mathbf{A}$ on samade omaväärtustega.
- (vii) Olgu $\mathbf{A} : n \times n$ pööratav omaväärtustega $\lambda_1, \dots, \lambda_n$. Siis $\mathbf{A}^{-1}$ omaväärtused on $\lambda_1^{-1}, \dots, \lambda_n^{-1}$.
- (vii) Olgu maatriks $\mathbf{A} : n \times n$ ortogonaalne. Siis tema omaväärtused on moodulilt võrdsed ühega (reaalarvulised omaväärtused on $+1$ või -1).
- (viii) Idempotentse maatriksi omaväärtused on võrdsed kas nulli või ühega.
- (ix) Maatriksi $\mathbf{A} : m \times n$ korral on maatriksitel $\mathbf{A} \mathbf{A}'$ ja $\mathbf{A}' \mathbf{A}$ ühed ja samad mittenullilised omaväärtused.
- (x) Kolmnurkse maatriksi omaväärtusteks on tema peadiagonaali elemendid.

- (xi) Vähemalt üks singulaarse maatriksi omaväärtustest võrdub nulliga.
 - (xii) Maatriksitel $\mathbf{A}$ ja $\mathbf{A} + c\mathbf{I}_n$ on samad omavektorid iga $c \in \mathbb{R}$ korral.
 - (xiii) Ruutmaatriksi $\mathbf{A} : n \times n$ korral $\sum_{i=1}^n \lambda_i = \text{tr} \mathbf{A}$.
 - (xiv) Olgu $\mathbf{A}$ sümmeetriline $n \times n$ -maatriks omaväärtustega λ_i . Siis $\sum_{i=1}^n \lambda_i^2 = \sum_{i,j=1}^n a_{ij}^2$.
- Mitmemõõtmelise statistika jaoks ülioluline on järgmine tulemus.

Teoreem L1.2. *Olgu $\mathbf{A} : n \times n$ sümmeetriline reaalarvuline maatriks. Siis leidub ortogonaalmaatriks $\mathbf{P}$ nii et*

$$\begin{aligned}\mathbf{P}'\mathbf{A}\mathbf{P} &= \mathbf{\Lambda}; \\ \mathbf{A} &= \mathbf{P}\mathbf{\Lambda}\mathbf{P}',\end{aligned}$$

kus $\mathbf{\Lambda}$ on maatriksi $\mathbf{A}$ omaväärtuste diagonaalmaatriks.

Sümmeetriline maatriks $\mathbf{A}$ on esitatav omaväärtuste ja normeeritud omavektorite kaudu ka nn. *spektraallahutusena*:

$$\mathbf{A} = \sum_{i=1}^n \lambda_i \vec{x}_i \vec{x}_i',$$

kus $\vec{x}_i$ on omaväärtusele λ_i vastav normeeritud omavektor. Kasulikuks osutub ka järgmine tulemus, mida tuntakse *Rayleigh suhte* nime all.

Teoreem L1.3 (Rayleigh suhe). *Olgu $\mathbf{A} : n \times n$ sümmeetriline maatriks omaväärtustega λ_i . Siis iga $\vec{x} \in \mathbb{R}^n$ korral*

$$\lambda_{\min} \leq \frac{\vec{x}'\mathbf{A}\vec{x}}{\vec{x}'\vec{x}} \leq \lambda_{\max}.$$

Positiivne määratus

Definitsioon L1.3. *Sümmeetriline maatriks $\mathbf{A} : n \times n$ on positiivselt (negatiivselt) määratud, kui $\vec{x}'\mathbf{A}\vec{x} > 0$ (< 0) iga vektori $\vec{x} \neq \vec{0} \in \mathbb{R}^n$ korral.*

Kui $\mathbf{A}$ on positiivselt (negatiivselt) määratud, tähistame seda $\mathbf{A} > 0$, ($\mathbf{A} < 0$).

Definitsioon L1.4. *Sümmeetriline maatriks $\mathbf{A} : n \times n$ on positiivselt poolmääratud (negatiivselt poolmääratud), kui $\vec{x}'\mathbf{A}\vec{x} \geq 0$ (≤ 0) iga vektori $\vec{x} \in \mathbb{R}^n$ korral ja leidub vähemalt üks vektor $\vec{x}_0 \neq \vec{0}$ nii, et $\vec{x}_0'\mathbf{A}\vec{x}_0 = 0$.*

Kui $\mathbf{A}$ on positiivselt poolmääratud (negatiivselt poolmääratud), tähistame seda $\mathbf{A} \geq 0$, ($\mathbf{A} \leq 0$).

Definitsioon L1.5. *Sümmeetriline maatriks $\mathbf{A} : n \times n$ on mittenegatiivselt (mittepositiivselt) määratud, kui $\mathbf{A} > 0$ või $\mathbf{A} \geq 0$ ($\mathbf{A} < 0$ või $\mathbf{A} \leq 0$).*

Järgnevalt esitame mõned positiivselt määratud maatriksite omadused:

- (i) maatriks $\mathbf{A} : n \times n$ on positiivselt määratud parajasti siis, kui kõik tema nurgamiinorid on positiivsed;
- (ii) Kui $\mathbf{A} > 0$, siis $\mathbf{A}^{-1} > 0$;
- (iii) Maatriks $\mathbf{A} > 0$ parajasti siis, kui kõik tema omaväärtused λ_i on positiivsed.
- (iv) Olgu $\mathbf{A} : n \times n$, $\mathbf{A} > 0$ ning $\mathbf{B} : n \times m$, $m \leq n$ ja $r(\mathbf{B}) = m$. Siis

$$\mathbf{B}'\mathbf{A}\mathbf{B} > 0.$$

- (v) Olgu $\mathbf{A} > 0$, $\mathbf{A}$ ja $\mathbf{B} : n \times n$ ning $\mathbf{B}$ pööratav maatriks. Siis

$$\mathbf{B}'\mathbf{A}\mathbf{B} > 0.$$

- (vi) Suvalise $\mathbf{A}$ korral on $\mathbf{A}\mathbf{A}'$ ja $\mathbf{A}'\mathbf{A}$ mittenegatiivselt määratud.

- (vii) Kui $\mathbf{A}$ on mittenegatiivselt määratud, siis leidub $\mathbf{A}^{-1}$ parajasti siis, kui $\mathbf{A} > 0$.
- (viii) Kui $\mathbf{A} > 0$, $\mathbf{A} : n \times n$ ning $\mathbf{B} : n \times m$, $m \leq n$ ja $r(\mathbf{B}) = r$, siis $\mathbf{B}'\mathbf{A}\mathbf{B} > 0$ parajasti siis, kui $r = m$. Maatriks $\mathbf{B}'\mathbf{A}\mathbf{B}$ on positiivselt poolmääratud, kui $r < m$.
- (ix) Kui $\mathbf{A} > 0$, $\mathbf{B} > 0$ ja $\mathbf{A} - \mathbf{B} > 0$, siis $\mathbf{B}^{-1} - \mathbf{A}^{-1} > 0$ ja $|\mathbf{A}| > |\mathbf{B}|$.
- (x) Kui $\mathbf{A} > 0$, $\mathbf{B} > 0$, siis $|\mathbf{A} + \mathbf{B}| \geq |\mathbf{A}| + |\mathbf{B}|$.

Projektorid

Vaatame eukleidilist ruumi $\mathbb{R}^n$ ja selles kaht alamruumi $\mathbb{M}$ ja $\mathbb{N}$, mis on teineteise täiendruumid, s.t.

$$\mathbb{R}^n = \mathbb{M} \oplus \mathbb{N}.$$

Siis suvaline vektor $\vec{x} \in \mathbb{R}^n$ avaldub üheselt summana

$$\vec{x} = \vec{x}_1 + \vec{x}_2,$$

kus $\vec{x}_1 \in \mathbb{M}$ ja $\vec{x}_2 \in \mathbb{N}$. Vektorit $\vec{x}_1$ nimetatakse vektori $\vec{x}$ *projektsiooniks ruumile* $\mathbb{M}$ paralleelselt ruumiga $\mathbb{N}$ ja vektorit $\vec{x}_2$ vektori $\vec{x}$ *projektsiooniks ruumile* $\mathbb{N}$ paralleelselt ruumiga $\mathbb{M}$.

Kui iga $\vec{x}_1 \in \mathbb{M}$ ja iga $\vec{x}_2 \in \mathbb{N}$ korral $\vec{x}_1'\vec{x}_2 = 0$, siis alamruumid $\mathbb{M}$ ja $\mathbb{N}$ on teineteise ortogonaalsed täiendid. Sel juhul on tegemist $\vec{x}$ ortogonaalsete projektsioonidega ruumidele $\mathbb{M}$ ja $\mathbb{N}$. Lihtne on kontrollida, et alamruumidele projekteerimine on lineaarne operatsioon. Tähistame sümboliga P lineaarse operaatori, mis projekteerib ruumi $\mathbb{R}^n$ vektorid alamruumi $\mathbb{M}$ paralleelselt ruumiga $\mathbb{N}$,

$$P\vec{x} = \vec{x}_1.$$

Nimetame seda lineaarset operaatorit *projektoriks* ruumile $\mathbb{M}$ paralleelselt ruumiga $\mathbb{N}$. Kui ruumid $\mathbb{M}$ ja $\mathbb{N}$ on teineteise ortogonaalsed täiendid, siis on tegemist *ortogonaalse projektoriga*. Kuna iga lineaarne operaator ruumis $\mathbb{R}^n$ on esitatav $n \times n$ -maatriksina, siis kasutamegi edaspidi maatriksesitust ja vaatame projektorina maatriksit $\mathbf{P}$. Lihtne on kontrollida, et kui $\mathbf{P}$ on projektor ruumile $\mathbb{M}$ paralleelselt ruumiga $\mathbb{N}$, siis $\mathbf{I}_n - \mathbf{P}$ on projektor ruumile $\mathbb{N}$ paralleelselt ruumiga $\mathbb{M}$. Projektorit iseloomustav omadus on idempotentsus.

Teoreem L1.4. *Lineaarne teisendus $\mathbf{P}$ ruumil $\mathbb{R}^n$ on projektor mingile tema alamruumile $\mathbb{M}$ paralleelselt täiendalamruumiga $\mathbb{N}$ parajasti siis, kui ta on idempotentne.*

Vaatame maatriksit $\mathbf{A} : n \times m$ kui teisendust ruumist $\mathbb{R}^m$ ruumi $\mathbb{R}^n$. Kujutised $\mathbf{A}\vec{x}$ moodustavad ruumis $\mathbb{R}^n$ alamruumi, mis erijuhul võib kokku langeda ka kogu ruumiga. Lineaar-teisenduse $\mathbf{A}$ *kujutisruumiks* $\mathcal{C}(\mathbf{A})$ (*column space* inglise k.) nimetatakse hulka

$$\mathcal{C}(\mathbf{A}) = \{\vec{x} : \vec{x} = \mathbf{A}\vec{y}, \vec{y} \in \mathbb{R}^m\}.$$

Teisenduse $\mathbf{A}$ *tuum* e. *nullruum* $\mathcal{N}(\mathbf{A})$ (*null space* inglise k.) on määratud võrdusega

$$\mathcal{N}(\mathbf{A}) = \{\vec{x} : \mathbf{A}\vec{x} = \vec{0}, \vec{x} \in \mathbb{R}^m\}.$$

Teisiti öeldes on lineaarteisenduse $\mathbf{A}$ kujutisruum maatriksi $\mathbf{A}$ veeruvektorite lineaarne kate. Kuidas leida projektor, mis projekteerib etteantud kujutisruumile? Olgu $\mathbf{A} : n \times m$, $m \leq n$ täisastakuga $r(\mathbf{A}) = m$ maatriks. Projektor, mis projekteerib suvalise vektori ruumist $\mathbb{R}^m$ maatriksi $\mathbf{A}$ veeruvektorite $\vec{a}_i$, $i = 1, \dots, m$ poolt määratud kujutisruumi paralleelselt nullruumiga $\mathcal{N}(\mathbf{A})$ on antud järgmises teoreemis.

Teoreem L1.5. *Olgu $\mathbf{A} : n \times m$, $m \leq n$ täisastakuga $r(\mathbf{A}) = m$ maatriks veeruvektoritega $\vec{a}_i$, $i = 1, \dots, m$. Siis projektor kujutisruumile $\mathcal{C}(\mathbf{A})$ on kujul*

$$\mathbf{P}_\mathbf{A} = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'.$$

Üldistatud pöördmaatriks

Üldistatud pöördmaatriksi mõiste on vajalik lineaarsete võrrandisüsteemide lahendamisel, kui võrrandeid on vähem kui tundmatuid või tundmatutele seatud tingimused on lineaarses sõltuvuses.

Definitsioon L1.6. Matriksi $\mathbf{A} : m \times n$ üldistatud pöördmatriksiks nimetatakse $n \times m$ -matriksit $\mathbf{A}^-$, kui

$$\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}. \quad (\text{L1.11})$$

Üldistatud pöördmatriks ei ole üheselt määratud, selliseid matrikseid, mis rahuldavad võrdust (L1.11), võib olla lõpmata palju. Kahtlemata rahuldab võrdust (L1.11) ruutmatriksi $\mathbf{A}$ pöördmatriks, kui viimane eksisteerib ja on siis ka ühene. Selleks, et üldistatud pöördmatriksite hulgast välja eraldada üks kindlate omadustega esindaja, on vaja lisatingimusi.

Definitsioon L1.7. Matriksi $\mathbf{A} : m \times n$ Moore-Penrose'i üldistatud pöördmatriksiks nimetatakse $n \times m$ -matriksit $\mathbf{A}^+$, kui ta rahuldab järgmist nelja tingimust:

- (i) $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$;
- (ii) $\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+$;
- (iii) $(\mathbf{A}\mathbf{A}^+)' = \mathbf{A}\mathbf{A}^+$;
- (iv) $(\mathbf{A}^+\mathbf{A})' = \mathbf{A}^+\mathbf{A}$.

Moore-Penrose'i üldistatud pöördmatriks on ühene, leidub üks ja ainult üks matriks, mis rahuldab tingimusi (i) – (iv). Moore-Penrose'i üldistatud pöördmatriksi seos lineaarsete võrrandite süsteemi lahendamisega on ära toodud järgmises teoreemis.

Teoreem L1.6. Olgu $\mathbf{A}$ $m \times n$ -matriks, $\vec{x}$ n -vektor ja $\vec{y}$ m -vektor. Võrrand $\mathbf{A}\vec{x} = \vec{y}$ on lahenduv parajasti siis, kui

$$\mathbf{A}\mathbf{A}^+\vec{y} = \vec{y}$$

ning selle üldlahend on kujul

$$\vec{x} = \mathbf{A}^+\vec{y} + (\mathbf{I}_n - \mathbf{A}^+\mathbf{A})\vec{q},$$

kus $\vec{q}$ on suvaline n -vektor.

L1.2 Plokkmaatriksid

Tehted ja tähistused

Matemaatilises statistikas on tihti olukordi, kus andmetes või neid kirjeldavates seosemaatriksites on sisemine struktuur, mida oleks vaja arvesse võtta analüüsi käigus. Tihti saame sisemist struktuuri arvesse võttes tunduvalt lihtsustada ka vajalikke arvutusi. Selles osas vaadeldavad mõisted kannavad statistikale orienteeritud maatriksalgebra esitustes tihti ka üuema maatriksalgebra nimetust (vt. näiteks Magnus and Neudecker (1999)). Selle all peetakse silmas eelkõige mõisteid otsekorrutis, kommutatsioonimaatriks, vec-operaator ja maatrikstuletis ning nende omadusi ja vastastikuseid seoseid.

Definitsioon L1.8.

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1v} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{u1} & \mathbf{A}_{u2} & \cdots & \mathbf{A}_{uv} \end{pmatrix}, \quad \sum_{i=1}^u m_i = m, \quad \sum_{j=1}^v n_j = n.$$

Plokkmaatriksi jaoks kasutame tähistust

$$\mathbf{A} = [\mathbf{A}_{ij}], \quad i = 1, \dots, u; \quad j = 1, \dots, v.$$

Plokkmaatriksi elementide määramisel kasutatakse kahekordseid indekseid, mis võimaldavad arvesse võtta plokkstruktuuri. Öeldakse, et plokkmaatriksi element asub (k, l) -ndas reas ja (g, h) -ndas veerus, kui ta on k -nda plokkide rea l -ndas reas ja g -nda plokkide veeru h -ndas veerus. Sel juhul kasutatakse tähistust $a_{(k,l)(g,h)}$ või $(\mathbf{A})_{(k,l)(g,h)}$. Tavalist tähistust kasutades saame võrduse

$$a_{(k,l)(g,h)} = a_{\sum_{i=1}^{k-1} m_i + l, \sum_{j=1}^{g-1} n_j + h}.$$

Plokkmaatriks $\mathbf{A} = [\mathbf{A}_{ij}]$, $i = 1, \dots, u$; $j = 1, \dots, v$ on plokk-diagonaalne, kui $u = v$ ja $\mathbf{A}_{ij} = 0$, kui $i \neq j$. Eespool vaadeldud tehted maatriksitega esitatakse arvestades plokkstruktuuri:

- (i) $c\mathbf{A} = [c\mathbf{A}_{ij}]$, $i = 1, \dots, u$; $j = 1, \dots, v$;
- (ii) $\mathbf{A} + \mathbf{B} = [\mathbf{A}_{ij} + \mathbf{B}_{ij}]$, $i = 1, \dots, u$; $j = 1, \dots, v$, kui plokiid on samade mõõtmetega;
- (iii) $\mathbf{A} = [\mathbf{A}_{ij}] : m \times n$, $i = 1, \dots, u$; $j = 1, \dots, v$ ja $\mathbf{B} = [\mathbf{B}_{jl}] : n \times r$, $j = 1, \dots, v$; $l = 1, \dots, w$ korrutis $\mathbf{AB}$ on plokkmaatriks, mis koosneb $m_i \times r_l$ -plokkidest

$$[\mathbf{AB}]_{il} = \sum_{j=1}^v \mathbf{A}_{ij} \mathbf{B}_{jl},$$

kui $\mathbf{A}_{ij}$ on $m_i \times n_j$ -plokiid ja $\mathbf{B}_{jl}$ on $n_j \times r_l$ -plokiid.

Paljudes rakendustes on oluline nn. 2×2 -plokkstruktuur:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}. \quad (\text{L1.12})$$

Sel juhul on võimalik esitada determinant ja pöördmaatriks plokkide kaudu.

Teoreem L1.7. Olgu pööratav plokkmaatriks $\mathbf{A}$ antud kujul (L1.12). Kui $\mathbf{A}_{22}$ on pööratav, siis

$$|\mathbf{A}| = |\mathbf{A}_{22}| \cdot |\mathbf{A}_{11} - \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \mathbf{A}_{21}|; \quad (\text{L1.13})$$

kui $\mathbf{A}_{11}$ on pööratav, siis

$$|\mathbf{A}| = |\mathbf{A}_{11}| \cdot |\mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}|. \quad (\text{L1.14})$$

Valemites (L1.13) ja (L1.14) esinevad plokkide vahed kannavad Schuri täiendite nime. Ploki $\mathbf{A}_{ii}$ Schuri täiendiks $\mathbf{A}_{jj \cdot i}$ nimetatakse maatriksit

$$\mathbf{A}_{jj \cdot i} = \mathbf{A}_{jj} - \mathbf{A}_{ji} \mathbf{A}_{ii}^{-1} \mathbf{A}_{ij}, \quad i, j = 1, 2.$$

Teoreem L1.8. Olgu pööratav plokkmaatriks $\mathbf{A}$ antud kujul (L1.12). Siis $\mathbf{A}$ pöördmaatriks on esitatav plokkide kaudu järgmiste võrdustena:

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11 \cdot 2}^{-1} & -\mathbf{A}_{11 \cdot 2}^{-1} \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{22}^{-1} \mathbf{A}_{21} \mathbf{A}_{11 \cdot 2}^{-1} & \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1} \mathbf{A}_{21} \mathbf{A}_{11 \cdot 2}^{-1} \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \end{pmatrix};$$

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22 \cdot 1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22 \cdot 1}^{-1} \\ -\mathbf{A}_{22 \cdot 1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} & \mathbf{A}_{22 \cdot 1}^{-1} \end{pmatrix}.$$

Plokkdiagonaalse maatriksi pöördmaatriks avaldub lihtsamalt:

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22}^{-1} \end{pmatrix}.$$

Kommuteerimismaatriks

Vaadeldav mõiste on kirjanduses tuntud ka kui permutatsioonimaatriks (*permutation matrix*), kuid enamus autoreid kasutab nimetust kommuteerimismaatriks (*commutation matrix*).

Definitsioon L1.9. Plokkmaatriksit $\mathbf{K}_{m,n} : mn \times mn$, mis koosneb $n \times m$ -plokkidest, nimetatakse kommuteerimismaatriksiks, kui

$$(\mathbf{K}_{m,n})_{(i,j)(g,h)} = \begin{cases} 1, & g = j, i = h, \\ 0, & \text{vastasel juhul} \end{cases},$$

kus $i, h = 1, \dots, m$; $j, g = 1, \dots, n$.

Kirjutame näitena välja maatriksi $\mathbf{K}_{2,3}$

$$\mathbf{K}_{2,3} = \begin{pmatrix} 1 & \dots & 0 & \vdots & 0 & 0 & \vdots & 0 & 0 \\ 0 & \dots & 0 & \vdots & 1 & 0 & \vdots & 0 & 0 \\ 0 & \dots & 0 & \vdots & 0 & 0 & \vdots & 1 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \vdots & 0 & 0 & \vdots & 0 & 0 \\ 0 & \dots & 0 & \vdots & 0 & 1 & \vdots & 0 & 0 \\ 0 & \dots & 0 & \vdots & 0 & 0 & \vdots & 0 & 1 \end{pmatrix}.$$

Näeme, et tulemuseks on permuteeritud ühikmaatriks, kus igas reas ja veerus on üks 1 ja ülejäänud elemendid on nullid ning sealjuures järgitakse teatud plokkstruktuuri.

Plokkide kaudu võime defineerida kommuteerimismaatriksi järgmiselt: *kommuteerimismaatriks* $\mathbf{K}_{m,n}$ on $mn \times mn$ -plokkmaatriks, mis koosneb $n \times m$ -plokkidest, kusjuures ij -ndas plokis on ji -s element 1 ja ülejäänud elemendid on nullid, $i = 1, \dots, m$; $j = 1, \dots, n$.

Toome ära mõned $\mathbf{K}_{m,n}$ omadused:

- (i) $\mathbf{K}_{m,n} = \mathbf{K}'_{n,m}$;
- (ii) $\mathbf{K}_{m,n} \mathbf{K}_{n,m} = \mathbf{I}_{mn}$;
- (iii) $\mathbf{K}_{m,1} = \mathbf{K}_{1,m} = \mathbf{I}_m$;
- (iv) $|\mathbf{K}_{m,n}| = \pm 1$;
- (v) $\mathbf{K}_{m,n}$ on ortogonaalmaatriks.

Otsekorrutis

Inglise keeles on selle mõiste kõige levinum nimetus *Kronecker product*, aga ka *direct product* ja *tensor product*. Eesti keeles eelistame *otsekorrutist*.

Definitsioon L1.10. *Maatriksite $\mathbf{A} : m \times n$ ja $\mathbf{B} : r \times s$ otsekorutis $\mathbf{A} \otimes \mathbf{B}$ on $mr \times ns$ -plokomaatriks*

$$\mathbf{A} \otimes \mathbf{B} = [a_{ij}\mathbf{B}], \quad i = 1, \dots, m; \quad j = 1, \dots, n,$$

kus

$$a_{ij}\mathbf{B} = \begin{pmatrix} a_{ij}b_{11} & \cdots & a_{ij}b_{1s} \\ \vdots & \ddots & \vdots \\ a_{ij}b_{r1} & \cdots & a_{ij}b_{rs} \end{pmatrix}.$$

Kui me vaatame $\mathbf{A} \otimes \mathbf{B}$ kui tavalist maatriksit, saame tema üldelemendi jaoks võrduse

$$(\mathbf{A} \otimes \mathbf{B})_{(i,j)(g,h)} = (\mathbf{A} \otimes \mathbf{B})_{(i-1)r+j, (g-1)s+h}.$$

Otsekorutise põhiomadused on:

- (i) $(\mathbf{A} \otimes \mathbf{B})_{(i,j)(g,h)} = a_{ig}b_{jh}$;
- (ii) $(\mathbf{A} \otimes \mathbf{B})' = \mathbf{A}' \otimes \mathbf{B}'$;
- (iii) $(\mathbf{A} + \mathbf{B}) \otimes (\mathbf{C} + \mathbf{D}) = \mathbf{A} \otimes \mathbf{C} + \mathbf{A} \otimes \mathbf{D} + \mathbf{B} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{D}$;
- (iv) $(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C})$;
- (v) maatriksite $\mathbf{A} : m \times n$, $\mathbf{B} : n \times w$, $\mathbf{C} : r \times s$ ja $\mathbf{D} : s \times t$ korral kehtib võrdus

$$(\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D}) = (\mathbf{AB}) \otimes (\mathbf{CD}); \quad (\text{L1.15})$$

- (vi) olgu $\mathbf{A}$ ja $\mathbf{B}$ pööratavad, siis

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1};$$

- (vii) olgu $\mathbf{A} : m \times n$, $\mathbf{B} : r \times s$ siis

$$\mathbf{A} \otimes \mathbf{B} = \mathbf{K}_{m,r}(\mathbf{B} \otimes \mathbf{A})\mathbf{K}_{s,n};$$

- (viii) kui $\mathbf{A} : m \times m$ ja $\mathbf{B} : r \times r$, siis

$$|\mathbf{A} \otimes \mathbf{B}| = |\mathbf{A}|^r |\mathbf{B}|^m;$$

- (ix) vektorite $\vec{a}$ ja $\vec{b}$ korral

$$\vec{a} \otimes \vec{b}' = \vec{a}\vec{b}' = \vec{b}' \otimes \vec{a}.$$

vec-operaator

Definitsioon L1.11. *Olgu $\mathbf{A} = (\vec{a}_1, \dots, \vec{a}_n)$ $m \times n$ -maatriks. Vektoriseerimisoperaator $\text{vec}(\cdot)$ on operaator maatriksite ruumist $\mathbb{R}^{m \times n}$ ruumi $\mathbb{R}^{mn}$ kujul*

$$\text{vec}\mathbf{A} = \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_n \end{pmatrix}.$$

Järgnevad omadused seovad vec-operaatori teiste maatriksoperatsioonidega:

- (i) kui $\mathbf{A}$ on $m \times n$ -maatriks, siis

$$\mathbf{K}_{m,n}\text{vec}\mathbf{A} = \text{vec}\mathbf{A}';$$

- (ii) kui $\mathbf{A}$ on $m \times n$ -maatriks, siis

$$\text{vec}\mathbf{A} = \mathbf{K}_{n,m}\text{vec}\mathbf{A}';$$

(iii) kui $\mathbf{A}$ on $n \times n$ sümmeetriline maatriks, siis

$$\text{vec} \mathbf{A} = \mathbf{K}_{n,n} \text{vec} \mathbf{A};$$

(iv) olgu maatriksid $\mathbf{A} : m \times n$, $\mathbf{B} : n \times r$, $\mathbf{C} : r \times s$, siis

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A}) \text{vec} \mathbf{B}; \quad (\text{L1.16})$$

(v) kui $\mathbf{A}$ ja $\mathbf{B}$ on $m \times n$ -maatriksid, siis

$$\text{tr}(\mathbf{A}'\mathbf{B}) = (\text{vec} \mathbf{A})' \text{vec} \mathbf{B}; \quad (\text{L1.17})$$

(vi) olgu $\mathbf{A} : m \times n$, $\mathbf{B} : n \times r$, $\mathbf{C} : r \times s$ ja $\mathbf{D} : s \times t$ maatriksid, siis

$$\text{tr}(\mathbf{ABCD}) = (\text{vec} \mathbf{A}')' (\mathbf{D}' \otimes \mathbf{B}) \text{vec} \mathbf{C};$$

(vii) olgu $\mathbf{A} : m \times n$ ja $\mathbf{B} : r \times s$, maatriksid, siis

$$\text{vec}(\mathbf{A} \otimes \mathbf{B}) = (\mathbf{I}_n \otimes \mathbf{K}_{s,m} \otimes \mathbf{I}_r) (\text{vec} \mathbf{A} \otimes \text{vec} \mathbf{B}).$$

Maatrikstuletis

Mitmemõõtmelises statistikas on sageli tegemist olukorraga, kus on vaja leida tuletisi mitmemuutuja- või maatriksfunktsioonidest olukorras, kus ka funktsiooni väärtused on vektorid või maatriksid. Selline on olukord näiteks suurima tõepära ja vähimruutude hinnangute leidmisel ja juhuslike vektorite ning maatriksite asümptootilise normaaljaotuse leidmisel. Funktsionaalanalüüsis on kasutusele võetud operaatorkujul mitmeid erinevaid tuletisi selliste olukordade jaoks, nimetame Gâteaux' ehk nõrka tuletist ja Frechet' ehk tugevat tuletist. Neist viimane sobib hästi ülalkirjeldatud statistikaprobleemide lahendamiseks ja tema esitus maatriksite kaudu kannab maatrikstuletise nimetust. Kasutame osatuletiste järjestamiseks definitsiooni, mille võttis kasutusele Heinz Neudecker (Neudecker, 1969).

Definitsioon L1.12. Olgu $r \times s$ -maatriksi $\mathbf{Y}$ elemendid funktsioonid $p \times q$ -maatriksi $\mathbf{X}$ elementidest. Nimetame $rs \times pq$ -maatriksit $\frac{d\mathbf{Y}}{d\mathbf{X}}$ maatriksi $\mathbf{Y}$ tuletiseks maatriksi $\mathbf{X}$ järgi lahtises piirkonnas $\mathcal{D}$, kui kõik osatuletised $\frac{\partial y_{kl}}{\partial x_{gh}}$ eksisteerivad ja on pidevad selles piirkonnas ja

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d}{d\text{vec}' \mathbf{X}} \otimes \text{vec} \mathbf{Y}, \quad (\text{L1.18})$$

kus

$$\frac{d}{d\text{vec}' \mathbf{X}} = \left(\frac{\partial}{\partial x_{11}}, \dots, \frac{\partial}{\partial x_{p1}}, \frac{\partial}{\partial x_{12}}, \dots, \frac{\partial}{\partial x_{p2}}, \dots, \frac{\partial}{\partial x_{1q}}, \dots, \frac{\partial}{\partial x_{pq}} \right)$$

ja osatuletise $\frac{\partial y_{kl}}{\partial x_{gh}}$ paiknemine maatriksis on määratud otsekorrutisega (L1.18).

Kirjutame plokmmaatriksi (L1.18) välja elementide kaudu:

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \begin{pmatrix} \frac{\partial y_{11}}{\partial x_{11}} & \dots & \frac{\partial y_{11}}{\partial x_{p1}} & \vdots & \dots & \vdots & \frac{\partial y_{11}}{\partial x_{1q}} & \dots & \frac{\partial y_{11}}{\partial x_{pq}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{r1}}{\partial x_{11}} & \dots & \frac{\partial y_{r1}}{\partial x_{p1}} & \vdots & \dots & \vdots & \frac{\partial y_{r1}}{\partial x_{1q}} & \dots & \frac{\partial y_{r1}}{\partial x_{pq}} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial y_{12}}{\partial x_{11}} & \dots & \frac{\partial y_{12}}{\partial x_{p1}} & \vdots & \dots & \vdots & \frac{\partial y_{12}}{\partial x_{1q}} & \dots & \frac{\partial y_{12}}{\partial x_{pq}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{r2}}{\partial x_{11}} & \dots & \frac{\partial y_{r2}}{\partial x_{p1}} & \vdots & \dots & \vdots & \frac{\partial y_{r2}}{\partial x_{1q}} & \dots & \frac{\partial y_{r2}}{\partial x_{pq}} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial y_{1s}}{\partial x_{11}} & \dots & \frac{\partial y_{1s}}{\partial x_{p1}} & \vdots & \dots & \vdots & \frac{\partial y_{1s}}{\partial x_{1q}} & \dots & \frac{\partial y_{1s}}{\partial x_{pq}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{rs}}{\partial x_{11}} & \dots & \frac{\partial y_{rs}}{\partial x_{p1}} & \vdots & \dots & \vdots & \frac{\partial y_{rs}}{\partial x_{1q}} & \dots & \frac{\partial y_{rs}}{\partial x_{pq}} \end{pmatrix},$$

Maatrikstuletise omadused esitame loeteluna, maatriksite mõõtmel toome ära, kui need erinevad definitsioonis esinevatest:

(i)

$$\frac{d\mathbf{X}}{d\mathbf{X}} = \mathbf{I}_{pq};$$

(ii) kui $\mathbf{Y} = c\mathbf{X}$, siis

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = c\mathbf{I}_{pq};$$

(iii) kui $\text{vec}\mathbf{Y} = \mathbf{A}\text{vec}\mathbf{X}$, kus $\mathbf{A}$ on konstantne maatriks, siis

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \mathbf{A};$$

(iv) kui $\mathbf{Y} = \mathbf{U} + \mathbf{V}$, siis

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d\mathbf{U}}{d\mathbf{X}} + \frac{d\mathbf{V}}{d\mathbf{X}};$$

(v) olgu $\mathbf{Z} = \mathbf{Z}(\mathbf{Y}) : m \times n$, $\mathbf{Y} = \mathbf{Y}(\mathbf{X}) : r \times s$ ja $\mathbf{X} : p \times q$, siis

$$\frac{d\mathbf{Z}}{d\mathbf{X}} = \frac{d\mathbf{Z}}{d\mathbf{Y}} \frac{d\mathbf{Y}}{d\mathbf{X}};$$

(vi) olgu $\mathbf{Y} = \mathbf{A}\mathbf{X}\mathbf{B}$, kus $\mathbf{A}$ ja $\mathbf{B}$ on konstantsed maatriksid, siis

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \mathbf{B}' \otimes \mathbf{A};$$

(vii) olgu $\mathbf{Z} = \mathbf{A}\mathbf{Y}\mathbf{B}$, kus $\mathbf{A}$ ja $\mathbf{B}$ on konstantsed maatriksid ja $\mathbf{Y} = \mathbf{Y}(\mathbf{X})$, siis

$$\frac{d\mathbf{Z}}{d\mathbf{X}} = (\mathbf{B}' \otimes \mathbf{A}) \frac{d\mathbf{Y}}{d\mathbf{X}};$$

(viii)

$$\frac{d\mathbf{X}'}{d\mathbf{X}} = \mathbf{K}_{p,q}, \quad \frac{d\mathbf{X}}{d\mathbf{X}'} = \mathbf{K}_{q,p};$$

(ix) kui $\mathbf{X}$ on $p \times p$ -maatriks, siis

$$\frac{d\mathbf{X}_d}{d\mathbf{X}} = (\mathbf{K}_{p,p})_d;$$

(x) kui $m \times n$ -maatriks $\mathbf{W} = \mathbf{W}(\mathbf{Y}, \mathbf{Z})$ on $r \times s$ -maatriksi $\mathbf{Y}$ ja $k \times l$ -maatriksi $\mathbf{Z}$ funktsioon, kusjuures $\mathbf{Y}$ ja $\mathbf{Z}$ on $\mathbf{X}$ funktsioonid, siis

$$\frac{d\mathbf{W}}{d\mathbf{X}} = \frac{d\mathbf{W}}{d\mathbf{Y}} \bigg|_{\mathbf{Z}=\text{const}} \frac{d\mathbf{Y}}{d\mathbf{X}} + \frac{d\mathbf{W}}{d\mathbf{Z}} \bigg|_{\mathbf{Y}=\text{const}} \frac{d\mathbf{Z}}{d\mathbf{X}};$$

(xi) kui $\mathbf{Y}$ ja $\mathbf{Z}$ on $\mathbf{X}$ funktsioonid, siis

$$\frac{d(\mathbf{Y}\mathbf{Z})}{d\mathbf{X}} = \frac{d(\mathbf{Y}\mathbf{Z})}{d\mathbf{Y}} \bigg|_{\mathbf{Z}=\text{const}} \frac{d\mathbf{Y}}{d\mathbf{X}} + \frac{d(\mathbf{Y}\mathbf{Z})}{d\mathbf{Z}} \bigg|_{\mathbf{Y}=\text{const}} \frac{d\mathbf{Z}}{d\mathbf{X}};$$

(xii) kui $\mathbf{Y}$ on $r \times r$ -maatriks, siis

$$\frac{d\mathbf{Y}^n}{d\mathbf{X}} = \left(\sum_{i+j=n-1, i,j \geq 0} (\mathbf{Y}')^i \otimes \mathbf{Y}^j \right) \frac{d\mathbf{Y}}{d\mathbf{X}},$$

kus $\mathbf{Y}^0 = \mathbf{I}_r$, $n \geq 1$;

(xiii) olgu $p \times p$ -maatriks $\mathbf{X}$ pööratav, siis

$$\frac{d\mathbf{X}^{-1}}{d\mathbf{X}} = -(\mathbf{X}')^{-1} \otimes \mathbf{X}^{-1};$$

(xiv) kui $\mathbf{Y}$ on pööratav $r \times r$ -maatriks, siis

$$\frac{d\mathbf{Y}^{-n}}{d\mathbf{X}} = - \left(\sum_{i+j=n-1, i,j \geq 0} (\mathbf{Y}')^{-i-1} \otimes \mathbf{Y}^{-j-1} \right) \frac{d\mathbf{Y}}{d\mathbf{X}},$$

kus $\mathbf{Y}^0 = \mathbf{I}_r$, $n \geq 1$;

(xv) olgu $\mathbf{Z} : m \times n$, $\mathbf{Y} : r \times s$ ja $\mathbf{X} : p \times q$, siis

$$\frac{d(\mathbf{Y} \otimes \mathbf{Z})}{d\mathbf{X}} = (\mathbf{I}_s \otimes \mathbf{K}_{n,r} \otimes \mathbf{I}_m) \left[(\mathbf{I}_{rs} \otimes \text{vec} \mathbf{Z}) \frac{d\mathbf{Y}}{d\mathbf{X}} + (\text{vec} \mathbf{Y} \otimes \mathbf{I}_{mn}) \frac{d\mathbf{Z}}{d\mathbf{X}} \right],$$

erijuhul, kui $\mathbf{Y} : 1 \times 1$,

$$\frac{d(\mathbf{Y} \otimes \mathbf{Z})}{d\mathbf{X}} = \text{vec} \mathbf{Z} \frac{d\mathbf{Y}}{d\mathbf{X}} + \mathbf{Y} \frac{d\mathbf{Z}}{d\mathbf{X}};$$

(xvi) kui $\mathbf{X}$ on pööratav $p \times p$ -maatriks, siis

$$\frac{d|\mathbf{X}|}{d\mathbf{X}} = |\mathbf{X}| \text{vec}'(\mathbf{X}')^{-1};$$

(xvii) kui $\mathbf{X}$ on $p \times p$ -maatriks, siis

$$\frac{d(\text{tr} \mathbf{X})}{d\mathbf{X}} = \text{vec}' \mathbf{I}_p.$$

(xviii) kui $\mathbf{X}$ on $p \times p$ -maatriks, siis

$$\frac{d(\mathbf{X}_d)}{d\mathbf{X}} = (K_{p,p})_d.$$

(xix) kui $\mathbf{X}$ on sümmeetriline $p \times p$ -maatriks, siis

$$\frac{d(\mathbf{X})}{d\mathbf{X}} = I_{p^2} + K_{p,p} - (K_{p,p})_d.$$

Kõrgemat järku tuletised defineeritakse rekursiivselt.

Definitsioon L1.13. Olgu $r \times s$ -maatriksi $\mathbf{Y}$ elemendid funktsioonid $p \times q$ -maatriksi $\mathbf{X}$ elementidest. Siis maatriksi $\mathbf{Y}$ k -ndat järku tuletiseks maatriksi $\mathbf{X}$ järgi $\frac{d^k \mathbf{Y}}{d\mathbf{X}^k}$ nimetatakse maatriksi $\mathbf{Y}$ $(k-1)$ -st järku tuletise $\frac{d^{k-1} \mathbf{Y}}{d\mathbf{X}^{k-1}}$ maatrikstuletist maatriksi $\mathbf{X}$ järgi:

$$\frac{d^k \mathbf{Y}}{d\mathbf{X}^k} = \frac{d}{d\mathbf{X}} \left(\frac{d^{k-1} \mathbf{Y}}{d\mathbf{X}^{k-1}} \right), \quad k = 2, 3, \dots,$$

kui see tuletis eksisteerib.

LISA 2

L2.1 Vajalikku tõenäosusteooriast

Selles lisas esitame mõned tõenäosusteooria mõisted ja tulemused, mida raamatus kasutame. Tegu ei ole põhimõtteliselt uute mõistetega aine "Tõenäosusteooria II" läbinute jaoks, samas on nad võib-olla esitatud mõnevõrra teise vaatenurga alt.

Koonduvused

Esitame järgnevas tähistused ja tulemused juhuslike vektorite jada koonduvuse kohta. Olgu $\{\vec{X}_n\}$ juhuslike p -vektorite jada. Selle jada koondumist jaotuse järgi ehk nõrka koondumist tähistame

$$\vec{X}_n \xrightarrow{\mathcal{D}} P_{\vec{X}}$$

või

$$\vec{X}_n \xrightarrow{\mathcal{D}} \vec{X},$$

kui $n \rightarrow \infty$. Jada $\{\vec{X}_n\}$ koondub jaotuse järgi parajasti siis, kui jaotusfunktsioonide jada

$$F_{\vec{X}_n}(\vec{x}) \longrightarrow F_{\vec{X}}(\vec{x})$$

jaotusfunktsiooni $F_{\vec{X}}$ igas pidevuspunktis $\vec{x}$.

Juhuslike maatriksite jada $\{\mathbf{X}_n\}$ koondub jaotuse järgi, kui

$$\text{vec}\mathbf{X}_n \xrightarrow{\mathcal{D}} \text{vec}\mathbf{X}.$$

Juhuslike p -vektorite jada $\{\vec{X}_n\}$ koondumist tõenäosuse järgi tähistame

$$\vec{X}_n \xrightarrow{P} \vec{X},$$

kui $n \rightarrow \infty$. See tähendab, et n lähenemisel lõpmatusele

$$P\{\omega : \rho(\vec{X}_n(\omega), \vec{X}(\omega)) > \varepsilon\} \rightarrow 0$$

iga $\varepsilon > 0$ korral, kus $\rho(\cdot, \cdot)$ on eukleidiline kaugus ruumis $\mathbb{R}^p$ ja ω tähistab elementaarsündmust.

Juhuslike maatriksite tõenäosuse järgi koondumine defineeritakse samuti juhuslike vektorite koondumise abil nagu jaotuse järgi koonduminegi:

$$\mathbf{X}_n \xrightarrow{P} \mathbf{X},$$

kui

$$\text{vec}\mathbf{X}_n \xrightarrow{P} \text{vec}\mathbf{X}.$$

Järgmisena toome sisse tõenäosuslikud analoogid matemaatikas kasutatavate suuruste $o(\cdot)$ ja $O(\cdot)$ jaoks.

Olgu $\{\varepsilon_i\}$ positiivsete arvude jada $\{X_i\}$ juhuslike suuruste jada. Siis järgides Rao (1973), lk. 151–152 tähistusi kasutame kirjalpilti $X_i = o_P(\varepsilon_i)$, kui

$$\frac{X_i}{\varepsilon_i} \xrightarrow{P} 0.$$

Juhuslike vektorite jada $\{\vec{X}_i\}$ jaoks kasutame samasugust tähistust $\vec{X}_i = o_P(\varepsilon_i)$, kui

$$\frac{\vec{X}_i}{\varepsilon_i} \xrightarrow{P} 0.$$

Juhuslike maatriksite jada $\{\mathbf{X}_i\}$ korral on $\mathbf{X}_i = o_P(\varepsilon_i)$, kui

$$\text{vec}\mathbf{X}_i = o_P(\varepsilon_i).$$

Võtame kasutusele ka tähistuse $O(\cdot)$ tõenäosusliku analoogi. Ütleme, et $X_i = O_P(\varepsilon_i)$, kui iga δ korral leiduvad m_δ ja n_δ nii, et

$$P\left(\frac{X_i}{\varepsilon_i} > m_\delta\right) < \delta, \text{ kui } i > n_\delta.$$

Juhuslike vektorite jada $\{\vec{X}_i\}$ korral $\vec{X}_i = O_P(\varepsilon_i)$, kui koordinaadid $(\vec{X}_i)_j = O_P(\varepsilon_i)$, $j = 1, \dots, p$.

Juhuslik maatriks $\mathbf{X}_i = O_P(\varepsilon_i)$, kui $\text{vec}\mathbf{X}_i = O_P(\varepsilon_i)$.

Lemma L2.1. *Olgu $\{\mathbf{X}_n\}$ ja $\{\mathbf{Y}_n\}$ juhuslike maatriksite jadad ja $\mathbf{X}_n \xrightarrow{\mathcal{D}} \mathbf{X}$ ja $\mathbf{Y}_n \xrightarrow{P} \mathbf{0}$. Siis eeldusel, et tehked on kooskõlas maatriksite mõõtmatega*

$$\begin{aligned}\mathbf{X}_n \otimes \mathbf{Y}_n &\xrightarrow{P} \mathbf{0}; \\ \text{vec}\mathbf{X}_n \text{vec}'\mathbf{Y}_n &\xrightarrow{P} \mathbf{0}; \\ \mathbf{X}_n + \mathbf{Y}_n &\xrightarrow{\mathcal{D}} \mathbf{X}.\end{aligned}$$

Lemma on mitmemõõtmeline variant koonduvustest, mida juhuslike suuruste korral tuntakse Slutsky teoreemi või Slutsky lemma nime all. Lemma väidete tõestus kordab kordab ühemõõtmelise juhu mõttekäiku (vt. näiteks Rao (1973), lk. 122–123). Erijuhul, kui $\{\mathbf{X}_n\}$ ja $\{\mathbf{Y}_n\}$ on juhuslike vektorite jadad $\{\vec{X}_n\}$ ja $\{\vec{Y}_n\}$, saame lemmast koonduvused

$$\begin{aligned}\vec{X}_n(\vec{Y}_n)' &\xrightarrow{P} \mathbf{0}; \\ \vec{X}_n + \vec{Y}_n &\xrightarrow{\mathcal{D}} \vec{X}.\end{aligned}$$

Paneme tähele, et jadad $\{\mathbf{X}_n\}$ ja $\{\mathbf{Y}_n\}$ ei pea olema sõltumatud. Nii näiteks, kui $\mathbf{X}_n \xrightarrow{\mathcal{D}} \mathbf{X}$ ja $g(\mathbf{X}_n) \xrightarrow{P} g(\mathbf{X})$, siis $\mathbf{X}_n + g(\mathbf{X}_n)$ asümptootiline jaotus on ka $P_{\mathbf{X}}$. Lemma väited esitame järelalusena kasutades tähistust $o_P(\cdot)$.

Järeldus L2.0.1. *Olgu $\{\mathbf{X}_n\}$, $\{\mathbf{Y}_n\}$ ja $\{\mathbf{Z}_n\}$ juhuslike maatriksite jadad ja $\mathbf{X}_n \xrightarrow{\mathcal{D}} \mathbf{X}$, $\mathbf{Y}_n = o_P(\varepsilon_n)$ ja $\mathbf{Z}_n = o_P(\varepsilon_n)$ ning $\{\varepsilon_i\}$ positiivsete arvude jada. Siis*

$$\begin{aligned}\mathbf{X}_n \otimes \mathbf{Y}_n &= o_P(\varepsilon_n); \\ \text{vec}\mathbf{X}_n \text{vec}'\mathbf{Y}_n &= o_P(\varepsilon_n); \\ \mathbf{Z}_n \otimes \mathbf{Y}_n &= o_P(\varepsilon_n^2).\end{aligned}$$

Järgnev teoreem on aluseks statistikute asümptootiliste jaotuste tuletamisel.

Teoreem L2.1. *Olgu $\{\vec{X}_n\}$ ja $\{\varepsilon_n\}$ vastavalt juhuslike p -vektorite ja positiivsete reaalarvude jadad ning $\vec{X}_n - \vec{a} = o_P(\varepsilon_n)$, kus $\vec{a}$ on konstantne p -vektor ning $\varepsilon_n \rightarrow 0$, kui $n \rightarrow \infty$. Kui funktsioon $\vec{g}: \mathbb{R}^p \rightarrow \mathbb{R}^q$ on arendatav Tayloriga ritta kohal $\vec{a}$:*

$$\vec{g}(\vec{x}) = \vec{g}(\vec{a}) + \sum_{k=1}^m \frac{1}{k!} ((\vec{x} - \vec{a})^{\otimes k-1} \otimes \mathbf{I}_q)' \frac{d^k \vec{g}(\vec{x})}{d\vec{x}^k} \Big|_{\vec{x}=\vec{a}} (\vec{x} - \vec{a}) + \dots,$$

siis

$$\begin{aligned}\vec{g}(\vec{X}_n) - \vec{g}(\vec{a}) \\ - \sum_{k=1}^m \frac{1}{k!} ((\vec{X}_n - \vec{a})^{\otimes k-1} \otimes \mathbf{I}_q)' \frac{d^k \vec{g}(\vec{X}_n)}{d\vec{X}_n^k} \Big|_{\vec{X}_n=\vec{a}} (\vec{X}_n - \vec{a}) &= o_P(\varepsilon_n^m),\end{aligned}$$

kus maatrikstuletis $\frac{d^k \vec{g}(\vec{x})}{d\vec{x}^k}$ on määratud definitsiooniga L1.13.

Karakteristlik funktsioon

Karakteristliku funktsiooniga seotud mõisted ja tulemused on põhjalikumalt esitatud raamatus Kollo and von Rosen (2005).

Definitsioon L2.14. Juhusliku vektori $\vec{X}$ karakteristlik funktsioon on määratud võrdusega

$$\varphi_{\vec{X}}(\vec{t}) = Ee^{i\vec{t}'\vec{X}}, \quad \vec{t} \in \mathbb{R}^p.$$

Juhusliku maatriksi $\mathbf{X}$ karakteristlik funktsiooni mõiste defineeritakse juhusliku vektori $\text{vec}X$ karakteristliku funktsioonina

$$\varphi_{\text{vec}\mathbf{X}}(\text{vec}\mathbf{T}) = Ee^{i\text{vec}\mathbf{T}'\text{vec}\mathbf{X}} = Ee^{i\text{tr}(\mathbf{T}'\mathbf{X})},$$

kus $\mathbf{X}$ ja $\mathbf{T}$ on samade mõõtmetega. Juhusliku vektori karakteristlikul funktsioonil on mitmeid kasulikke omadusi.

(i) Kui $\vec{X}$ ja $\vec{Y}$ on sõltumatud juhuslikud p -vektorid, siis

$$\varphi_{\vec{X}+\vec{Y}}(\vec{t}) = \varphi_{\vec{X}}(\vec{t}) \cdot \varphi_{\vec{Y}}(\vec{t}).$$

(ii) Kui $\vec{X}$ on juhuslik p -vektor ja $\mathbf{A}$ $q \times p$ -maatriks, siis

$$\varphi_{\mathbf{A}\vec{X}}(\vec{t}) = \varphi_{\vec{X}}(\mathbf{A}'\vec{t}), \quad \vec{t} \in \mathbb{R}^q.$$

(iii) Olgu $\{\vec{X}_n\}$ juhuslike p -vektorite jada ja $F_{\vec{X}_n}(\vec{x})$ ning $\varphi_{\vec{X}_n}(\vec{t})$ vastavalt juhusliku vektori jaotusfunktsioon ja karakteristlik funktsioon. Kui iga $\vec{t} \in \mathbb{R}^p$ korral eksisteerib piirväärtus $\lim_n \varphi_{\vec{X}_n}(\vec{t})$ ja funktsioon

$$\varphi(\vec{t}) = \lim_n \varphi_{\vec{X}_n}(\vec{t})$$

on pidev punktis $\vec{t} = \vec{0}$, siis ta on mingi jaotuse $P_{\vec{X}}$ karakteristlik funktsioon ja

$$\vec{X}_n \xrightarrow{\mathcal{D}} P_{\vec{X}}.$$

Definitsioon L2.15. Olgu juhusliku vektori $\vec{X} = (X_1, \dots, X_p)'$ karakteristlik funktsioon $\varphi_{\vec{X}}(\vec{t})$ k korda diferentseeruv ja olgu lõplikud kõik k -ndat järku segamomendid $E(X_1^{k_1} \dots X_p^{k_p})$, $k_i \in \{0, \dots, k\}$, $\sum_{i=1}^p k_i = k$. Siis juhusliku vektori $\vec{X}$ k -ndat järku moment on antud võrdusega

$$m_k(\vec{X}) = \frac{1}{i^k} \frac{d^k \varphi_{\vec{X}}(\vec{t})}{d\vec{t}^k} \Big|_{\vec{t}=\vec{0}}, \quad \vec{t} \in \mathbb{R}^p.$$

Paralleelselt momentidega kirjeldavad mitmemõõtmelist jaotust tema kumulandid, mis defineeritakse kumulantfunktsiooni kaudu.

Definitsioon L2.16. Juhusliku p -vektori $\vec{X}$ kumulantfunktsiooniks $\psi_{\vec{X}}(\vec{t})$ nimetatakse funktsiooni

$$\psi_{\vec{X}}(\vec{t}) = \ln \varphi_{\vec{X}}(\vec{t}), \quad \vec{t} \in \mathbb{R}^p.$$

Kumulandid saadakse kumulantfunktsioonist samuti nagu momendid karakteristlikust funktsioonist.

Definitsioon L2.17. Olgu kumulantfunktsioon $\psi_{\vec{X}}(\vec{t})$ k korda diferentseeruv ja olgu lõplikud kõik k -ndat järku segamomendid $E(X_1^{k_1} \dots X_p^{k_p})$, $k_i \in \{0, \dots, k\}$, $\sum_{i=1}^p k_i = k$. Siis juhusliku vektori $\vec{X}$ k -ndat järku kumulant on antud võrdusega

$$c_k(\vec{X}) = \frac{1}{i^k} \frac{d^k \psi_{\vec{X}}(\vec{t})}{d\vec{t}^k} \Big|_{\vec{t}=\vec{0}}, \quad \vec{t} \in \mathbb{R}^p.$$

Kumulandid on tihti lihtsamalt leitavad kui momendid, samuti on nende puhul oluline fakt, et normaaljaotuse korral $c_k(\vec{X}) = \mathbf{0}$ kui $k > 2$. Kuna kumulantfunktsioon ja karakteristlik funktsioon on omavahel analüütilises seoses, on loomulik ka see, et momendid on avaldatavad kumulantide kaudu ja vastupidi. Esimeste momentide ja kumulantide vahel kehtivad järgmised seosed:

(i) $c_1(\vec{X}) = m_1(\vec{X}) = (E\vec{X})'$;

(ii) $c_2(\vec{X}) = m_2(\vec{X} - E\vec{X}) = D\vec{X}$;

(iii) $c_3(\vec{X}) = m_3(\vec{X} - E\vec{X})$.

Kõrgemat järku kumulantide ja momentide vahelised seosed on keerulisemad.

Viited

- Anderson, T. (2003). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York, third edition.
- Bilodeau, M. and Brenner, D. (1999). *Theory of Multivariate Statistics*. Springer-Verlag, New York.
- Harville, D. (1997). *Matrix Algebra From a Statistician's Perspective*. Springer, New York.
- Kilp, M. (2005). *Algebra I*. Eesti Matemaatika Selts, Tartu.
- Kollo, T. (1991). *Maatrikstuletitis mitmemõõtmelise statistika jaoks (vene keeles)*. Tartu Ülikooli Kirjastus, Tartu.
- Kollo, T. and Srivastava, M. (2004). Estimation and testing of parameters in multivariate laplace distribution. *Comm. Statist.*, 33:2363–2687.
- Kollo, T. and von Rosen, D. (2005). *Advanced Multivariate Statistics with Matrices*. Springer, Dordrecht.
- Magnus, J. and Neudecker, H. (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley, Chichester, second edition.
- Mardia, K. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57:519–530.
- Moore, D. (1977). Generalized inverses, wald's method, and the construction of chi-squared tests of fit. *JASA*, 72:131–137.
- Muirhead, R. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York.
- Rao, C. (1973). *Linear Statistical Inference and Its Applications*. Wiley, New York, second edition.
- Rao, C. and Rao, M. (1998). *Matrix Algebra and Its Applications to Statistics and Econometrics*. World Scientific, Singapore.
- Schott, J. (1997). *Matrix Analysis for Statistics*. Wiley, New York.
- Searle, S. (1982). *Matrix Algebra Useful for Statistics*. Wiley, New York.
- Siotani, M., Hayakawa, T., and Fujikoshi, Y. (1985). *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook*. American Science Press, Columbus, Ohio.
- Srivastava, M. and Khatri, C. (1979). *An Introduction to Multivariate Statistics*. North Holland, New York.
- Wald, A. (1943). Tests of statistical hypothesis concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, 54:426–482.