Tõenäosusteooria ja statistika II

Loeng 6

Mitmemõõtmeline normaaljaotus

Märt Möls

mart.mols@ut.ee

Definitsioon

Juhusliku vektori $\mathbf{X} = (X_1,\ X_2,\ \dots,\ X_n)^T$ *keskväärtus on vektor*

$$E\mathbf{X} = (EX_1,\ EX_2\ ,\dots,\ EX_n)^T.$$

Definitsioon

Juhusliku vektori $\mathbf{X} = (X_1,\ X_2,\ \dots,\ X_n)^T$ *kovariatsioonimaatriks (ka dispersioonimaatriks) on järgmine* $n \times n$ *sümmeetriline maatriks:*

$$D(\mathbf{X}) = \begin{pmatrix} DX_1 & cov(X_1,\ X_2) & cov(X_1,\ X_3) & \dots & cov(X_1,\ X_n) \\ cov(X_2,\ X_1) & DX_2 & cov(X_2,\ X_3) & \dots & cov(X_2,\ X_n) \\ cov(X_3,\ X_1) & cov(X_3,\ X_2) & DX_3 & \dots & cov(X_3,\ X_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ cov(X_n,\ X_1) & cov(X_n,\ X_2) & cov(X_n,\ X_3) & \dots & DX_n \end{pmatrix}$$

Definitsioon

Juhusliku vektori $\mathbf{X} = (X_1,\ X_2,\ \dots,\ X_n)^T$ *korrelatsioonimaatriks on järgmine* $n \times n$ *sümmeetriline maatriks:*

$$\rho(\mathbf{X}) = \begin{pmatrix} 1 & \rho(X_1,\ X_2) & \rho(X_1,\ X_3) & \dots & \rho(X_1,\ X_n) \\ \rho(X_2,\ X_1) & 1 & \rho(X_2,\ X_3) & \dots & \rho(X_2,\ X_n) \\ \rho(X_3,\ X_1) & \rho(X_3,\ X_2) & 1 & \dots & \rho(X_3,\ X_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho(X_n,\ X_1) & \rho(X_n,\ X_2) & \rho(X_n,\ X_3) & \dots & 1 \end{pmatrix}$$

Lemma

Kovariatsiooni- ja korrelatsioonimaatriksil on järgmised omadused:

1. *Kovariatsioonimaatriks ja korrelatsioonimaatriks on sümmeetrilised.*
2. *Kui vektori* $\mathbf{X} = (X_1,\ X_2,\ \dots,\ X_n)^T$ *komponendid on sõltumatud, on kovariatsioonimaatriks diagonaalmaatriks ja korrelatsioonimaatriks ühikmaatriks.*
3. *Kui* $\mathbf{\Sigma}$ *on vektori* $\mathbf{X} = (X_1,\ X_2,\ \dots,\ X_n)^T$ *kovariatsioonimaatriks ja* $\boldsymbol{a} = (a_1,\ a_2, \dots,\ a_n)^T$ *on suvaline n-dimensionaalne konstantne vektor, siis*

$$\boldsymbol{a}^t \mathbf{\Sigma} \boldsymbol{a} = D(a_1X_1 + a_2X_2 + \dots + a_nX_n) \geq 0.$$

Mitmemõõtmeline normaaljaotus

Juhuslik vektor $\mathbf{X} = (X_1, X_2, ..., X_n)^T$ on mitmemõõtmelise normaaljaotusega, kui tema tihedusfunktsioon avaldub kujul

$$f_{\mathbf{X}}(\boldsymbol{x}) = \frac{1}{\sqrt{|2\pi \cdot \boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right),$$

kus $\boldsymbol{x} = (x_1, ..., x_n) \in \mathcal{R}^n$, $\boldsymbol{\mu} \in \mathcal{R}^n$ ja $\boldsymbol{\Sigma}$ on $n \times n$ sümmeetriline positiivselt poolmääratud maatriks. Vektor $\boldsymbol{\mu}$ ja maatriks $\boldsymbol{\Sigma}$ on mitmemõõtmelise normaaljaotuse parameetrid ja nende parameetritega mitmemõõtmelist normaaljaotust tähistatakse $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Märkus:

- Ühemõõtmelisel juhul oleme normaaljaotuse teiseks parameetriks kasutanud juhusliku suuruse standardhälvet, näiteks $X \sim N(\mu, \sigma)$ (õpikus Pärna, 2013).
- Sageli kasutatakse teise parameetrina juhusliku suuruse dispersiooni $X \sim N(\mu, \sigma^2)$ (õpikus Traat, 2006).
- Mitmemõõtmelise normaaljaotuse tähistuses on siiski levinum teine variant, ehk teise parameetri rollis on kovariatsioonimaatriks $\boldsymbol{\Sigma}$.

Kahemõõtmeline normaaljaotus. Vaatleme kahemõõtmelist juhusliku vektorit $(X, Y)^T$. Olgu $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$ ja

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix},$$

kus $\sigma_x^2 = DX$, $\sigma_y^2 = DY$ ja $\sigma_{xy} = cov(X, Y)$.

Näitame (tahvlil), et kahemõõtmelise normaaljaotuse tihedusfunktsioon on

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_x^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_x\sigma_y} + \frac{(y-\mu_2)^2}{\sigma_y^2}\right)\right].$$

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_x^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_x\sigma_y} + \frac{(y-\mu_2)^2}{\sigma_y^2}\right)\right]$$

$$f(x,y) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_x^2}\right] \cdot \frac{1}{\sqrt{2\pi\sigma_y^2(1-\rho^2)}} \cdot \exp\left[\frac{(x-\mu_1)^2}{2\sigma_x^2} - \frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_x^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_x\sigma_y} + \frac{(y-\mu_2)^2}{\sigma_y^2}\right)\right]$$

$$f(x,y) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_x^2}\right] \cdot \frac{1}{\sqrt{2\pi\sigma_y^2(1-\rho^2)}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(\rho^2-1)(x-\mu_1)^2}{\sigma_x^2} + \frac{(x-\mu_1)^2}{\sigma_x^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_x\sigma_y} + \frac{(y-\mu_2)^2}{\sigma_y^2}\right)\right]$$

$$
\begin{aligned}
f(x,y) =& \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_x^2}\right] \cdot \\
& \cdot \frac{1}{\sqrt{2\pi\sigma_y^2(1-\rho^2)}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(\rho^2-1)(x-\mu_1)^2}{\sigma_x^2} + \frac{(x-\mu_1)^2}{\sigma_x^2}\right.\right. \\
& \left.\left. - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_x\sigma_y} + \frac{(y-\mu_2)^2}{\sigma_y^2}\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
f(x,y) =& \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_x^2}\right] \cdot \\
& \cdot \frac{1}{\sqrt{2\pi\sigma_y^2(1-\rho^2)}} \exp\left[-\frac{1}{2(1-\rho^2)} \cdot \right. \\
& \left. \cdot \left(\frac{\rho^2(x-\mu_1)^2}{\sigma_x^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_x\sigma_y} + \frac{(y-\mu_2)^2}{\sigma_y^2}\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
f(x,y) =& \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_x^2}\right] \cdot \\
& \cdot \frac{1}{\sqrt{2\pi\sigma_y^2(1-\rho^2)}} \exp\left[-\frac{1}{2(1-\rho^2)} \cdot \right. \\
& \left. \cdot \left(\frac{\rho^2(x-\mu_1)^2}{\sigma_x^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_x\sigma_y} + \frac{(y-\mu_2)^2}{\sigma_y^2}\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
f(x,y) =& \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_x^2}\right] \cdot \\
& \cdot \frac{1}{\sqrt{2\pi\sigma_y^2(1-\rho^2)}} \exp\left[-\frac{1}{2(1-\rho^2)} \cdot \left(\frac{y-\mu_2}{\sigma_y} - \frac{\rho(x-\mu_1)}{\sigma_x}\right)^2\right]
\end{aligned}
$$

$$
\begin{aligned}
f(x,y) =& \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_x^2}\right] \cdot \\
& \cdot \frac{1}{\sqrt{2\pi\sigma_y^2(1-\rho^2)}} \exp\left[-\frac{1}{2(1-\rho^2)} \cdot \left(\frac{y-\mu_2}{\sigma_y} - \frac{\rho(x-\mu_1)}{\sigma_x}\right)^2\right]
\end{aligned}
$$

$$
\begin{aligned}
f(x,y) =& \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_x^2}\right] \cdot \\
& \cdot \frac{1}{\sqrt{2\pi\sigma_y^2(1-\rho^2)}} \exp\left[-\frac{1}{2} \cdot \frac{(y-(\mu_2+\sigma_y/\sigma_x\cdot\rho(x-\mu_1)))^2}{\sigma_y^2(1-\rho^2)}\right]
\end{aligned}
$$

$$
f_{X,Y}(x,y) = f_X(x) \cdot f_{Y|X=x}(y)
$$

Tinglik jaotus

Olgu $\mathbf{X}$ mitmemõõtmelise normaaljaotusega juhuslik suurus. Jagame vektori $\mathbf{X}$ kaheks osaks,

$$
\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix},
$$

ja vastavalt jaotame osadeks ka keskväärtusvektori ja dispersiooni:

$$
\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}.
$$

Siis

$$
\mathbf{X}_1 | \mathbf{X}_2 = \mathbf{a} \sim N(\boldsymbol{\mu}_{\mathbf{X}_1|\mathbf{X}_2}; \boldsymbol{\Sigma}_{\mathbf{X}_1|\mathbf{X}_2})
$$

$$
\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{a} - \boldsymbol{\mu}_2)
$$

$$
\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}
$$

Eelnevast teame, et kui kovariatsioon kahe juhusliku suuruse vahel on 0, siis ei tähenda see veel juhuslike suuruste sõltumatust. Küll aga kehtib vastupidine seos: kui $cov(X,\ Y) = 0$, siis X ja Y on sõltumatud. Mitmemõõtmeline normaaljaotus on selles mõttes eriline: väide kehtib mõlemas suunas (vt. järgmist lemmat).

Lemma

Mitmemõõtmelise normaaljaotusega vektori $\mathbf{X} = (X_1, X_2, ..., X_N)^T$ komponendid on sõltumatud parajasti siis, kui nendevahelised kovariatsioonid $(cov(X_i,\ X_j),\ j \neq j)$ on kõik nullid. Kovariatsioonimaatriks on sel juhul diagonaalne (väljaspool peadiagonaali on kõik nullid).

Tõestame loengul kahemõõtmelisel juhul.

Mitmemõõtmelise normaaljaotuse lisaomadused (ilma tõestuseta):

1. Sõltumatud normaaljaotusega juhuslikud suurused moodustavad mitmemõõtmelise normaaljaotusega vektori (kovariatsioonimaatriks on diagonaalne).
2. Olgu $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ja $\mathbf{D}:\ l \times n$ maatriks astakuga $l \leq n$ ning olgu $\mathbf{Y} = \mathbf{D}\,\mathbf{X}:\ l \times 1$ (juhuslik vektor, mis on saadud elementide X_i lineaarsete kombinatsioonide abil). Siis $\mathbf{Y} \sim N(\mathbf{D}\boldsymbol{\mu},\ \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^T)$.

Viimasest omadusest järldub:

1. Kui $\mathbf{D} = \boldsymbol{d} = (d_1, d_2, ..., d_n):\ n \times 1$ reavektor, siis $\mathbf{Y} = \boldsymbol{d}\mathbf{X} = \sum_{i=1}^{n} d_i X_i$. See tähendab, et mitmemõõtmelise normaaljaotuse korral vektori elementide lineaarne kombinatsioon on samuti normaaljaotusega.
2. Kui $\mathbf{D} = \boldsymbol{d} = (d_1, d_2, ..., d_k, 0, 0, ..., 0):\ n \times 1$, mille esimest k komponenti erineb nullist ja ülejäänud on nullid, siis $\mathbf{Y} = \boldsymbol{d}\mathbf{X} = \sum_{i=1}^{k} d_i X_i$. See tähendab, et mitmemõõtmelise normaaljaotuse korral vektori elementide suvalise alamhulga lineaarne kombinatsioon on samuti normaaljaotusega.
3. Saame alati valida $\mathbf{D}$ nii, et see "võtaks" vektorist $\mathbf{X}$ suvalise alamhulga elemente, näiteks kui $\mathbf{X} = (X_1, X_2, X_3)^T$ ja
$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$
siis $\mathbf{Y} = \mathbf{D}\,\mathbf{X} = (X_1, X_2)^T$. See tähendab, et mitmemõõtmelise normaaljaotusega vektori suvaline alamhulk on samuti normaaljaotusega.

Juhuslik vektor $(X, Y)^T$ on kahemõõtmelise normaaljaotusega, kusjuures $EX = 0$, $DX = 1$, $EY = -1$, $DY = 4$ ja $\rho_{X,Y} = -1/2$. Leida konstandi a väärtus nii, et suurused $(aX + Y)$ ja $(X + 2Y)$ oleksid sõltumatud. (NB! Panna kirja teisenduste maatriks $\mathbf{D}$.)