10 Moniulotteinen normaalijakauma

- Tässä luvussa tarkastellaan normaalijakauman moniulotteista yleistystä eli moniulotteista (eli monimuuttujaista) normaalijakaumaa (engl. *multivariate normal distribution*).
- Sitä kutsutaan myös multinormaalijakaumaksi.
- Tutustumme tähän sovelluksissa usein esiintyvään moniulotteiseen jakaumaan.
- Näemme kuinka moniulotteisia jakaumia käsitellään vektori- ja matriisimerkinnöillä.

10.1 Standardinormaalijakauma $N_n(\mathbf{0}, \mathbf{I})$

Määritelmä

Sv:lla $\mathbf{U} = (U_1, \dots, U_n)$ on *n-ulotteinen standardinormaalijakauma* eli normaalijakauma $N_n(\mathbf{0}, \mathbf{I})$ täsmälleen silloin, kun sen komponentit ovat riippumattomia $N(0, 1)$-jakaumaa noudattavia satunnaismuuttujia.

Standardinormaalijakauman $N_n(\mathbf{0}, \mathbf{I})$ ominaisuuksia

- Satunnaisvektorin $\mathbf{U} \sim N_n(\mathbf{0}, \mathbf{I})$ tf on

$$\begin{aligned} f_{\mathbf{U}}(\mathbf{u}) &= \prod_{i=1}^{n} f_{U_i}(u_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{-\frac{1}{2}u_i^2} \\ &= (2\pi)^{-n/2} \exp(-\frac{1}{2}(u_1^2 + \cdots + u_n^2)) \\ &= (2\pi)^{-n/2}\, \exp(-\frac{1}{2}\mathbf{u}^T\mathbf{u}). \end{aligned} \tag{1}$$

- Sv:n $\mathbf{U}$ odotusarvovektori on n-dimensionen nollavektori, ja sen kovarianssimatriisi on dimensiota $n \times n$ oleva yksikkömatriisi

$$E\mathbf{U} = \mathbf{0}_n, \qquad \operatorname{Cov} \mathbf{U} = \mathbf{I}_n. \tag{2}$$

Standardinormaalijakauman momenttiemäfunktio

Satunnaisvektorin $\mathbf{U} \sim N_n(\mathbf{0}, \mathbf{I})$ momenttiemäfunktio on

$$
\begin{aligned}
M_{\mathbf{U}}(\mathbf{t}) &= E\exp(\mathbf{t}^T\mathbf{U}) = E\left[\prod_{i=1}^{n}\exp(t_iU_i)\right] \\
&= \prod_{i=1}^{n} E\exp(t_iU_i) && (U_1, \ldots, U_n \text{ riippumattomia}) \\
&= \prod_{i=1}^{n} \mathrm{e}^{\frac{1}{2}t_i^2} && (U_i \sim N(0,1)) \\
&= \exp(\frac{1}{2}\mathbf{t}^T\mathbf{t}).
\end{aligned}
$$

Standardinormaalijakautuneen vektorin affiini muunnos

- Määritellään sv $\mathbf{X}$ kaavalla

$$\mathbf{X} = \mathbf{A}\mathbf{U} + \boldsymbol{\mu}, \tag{3}$$

jossa $\mathbf{A}$ on $m \times n$-vakiomatriisi, $\boldsymbol{\mu} \in \mathbb{R}^m$ on vakiovektori, ja $\mathbf{U} \sim N_n(\mathbf{0}, \mathbf{I}_n)$.

- Tällöin

$$E\mathbf{X} = \boldsymbol{\mu}, \qquad \mathrm{Cov}\,\mathbf{X} = \mathbf{A}\,\mathbf{I}_n\mathbf{A}^T = \mathbf{A}\mathbf{A}^T. \tag{4}$$

- Merkitään $\boldsymbol{\Sigma} = \mathrm{Cov}\,\mathbf{X} = \mathbf{A}\mathbf{A}^T$.
- Sv:n $\mathbf{X}$ momenttiemäfunktio saadaan helpolla laskulla:

$$\begin{aligned} M_{\mathbf{X}}(\mathbf{t}) &= E\exp(\mathbf{t}^T\mathbf{X}) = E\exp(\mathbf{t}^T(\mathbf{A}\mathbf{U} + \boldsymbol{\mu})) = \exp(\mathbf{t}^T\boldsymbol{\mu})\, M_{\mathbf{U}}(\mathbf{A}^T\mathbf{t}) \\ &= \exp(\mathbf{t}^T\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}^T\mathbf{A}\mathbf{A}^T\mathbf{t}) = \exp(\mathbf{t}^T\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}) \end{aligned} \tag{5}$$

Voidaanko edellistä laskua käyttää yleisen multinormaalijakauman määritelmänä?

- Multinormaalijakauma voidaan määritellä edellisen laskun avulla, mikäli mielivaltainen kovarianssimatriisi $\boldsymbol{\Sigma}$ ensin osataan esittää tulona $\mathbf{A}\mathbf{A}^T$.
- Eräs (eikä suinkaan ainoa) mahdollisuus on ns. Choleskyn hajotelma (engl. *Cholesky decomposition*).
- Choleskyn hajotelmassa mielivaltainen symmetrinen ja positiivisesti semidefiniitti matriisi $\boldsymbol{\Sigma}$ esitetään tulona

$$\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^T, \tag{6}$$

  jossa $\mathbf{L}$ on alakolmiomatriisi. (Alakolmiomatriisi on neliömatriisi, jonka kaikki yläkolmion alkiot ovat nollia.)
- Choleskyn hajotelma on saatavilla matriisilaskennan ohjelmakirjastoissa (ja tyypillisesti ohjelmat palauttavat jälkimmäisen tekijän eli yläkolmiomatriisin $\mathbf{L}^T$).

10.2 Yleinen multinormaalijakauma

Määritelmä

Sv:lla $\mathbf{X}$ on multinormaalijakauma, jos sillä on sama jakauma kuin satunnaisvektorilla

$$\mathbf{A}\mathbf{U} + \boldsymbol{\mu} \tag{7}$$

jossa $\mathbf{A}$ on $m \times n$-vakiomatriisi, $\boldsymbol{\mu} \in \mathbb{R}^m$ on vakiovektori, ja $\mathbf{U} \sim N_n(\mathbf{0}, \mathbf{I}_n)$ jollakin n.

Määritelmästä sekä kaavoista (4) ja (5) seuraa, että

$$E\mathbf{X} = \boldsymbol{\mu}, \qquad \boldsymbol{\Sigma} = \operatorname{Cov} \mathbf{X} = \mathbf{A}\mathbf{A}^T,$$

$$M_{\mathbf{X}}(\mathbf{t}) = \exp(\mathbf{t}^T \boldsymbol{\mu} + \frac{1}{2}\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t})$$

Yksiulotteinen $N(\mu, \sigma^2)$ on tästä erikoistapaus, sillä

$$X \sim N(\mu, \sigma^2) \quad \Rightarrow \quad (X = \mu + \sigma U, \quad U \sim N(0,1)).$$

Lause

- Määritelmästä 2 seuraa, että $E\mathbf{X} = \boldsymbol{\mu}$ ja $\operatorname{Cov}\mathbf{X} = \mathbf{A}\mathbf{A}^T$. Sv:n $\mathbf{X}$ jakauma riippuu vain sen odotusarvovektorista ja kovarianssimatriisista, ei esitysdimensiosta n eikä esitysmatriisin $\mathbf{A}$ muista ominaisuuksista.
- Kääntäen, jos $\boldsymbol{\mu} \in \mathbb{R}^m$ on vakiovektori ja $\boldsymbol{\Sigma} \in \mathbb{R}^{m \times m}$ on positiivisesti semidefiniitti matriisi, niin on olemassa sv $\mathbf{X}$, jolla on multinormaalijakauma odotusarvovektorilla $\boldsymbol{\mu}$ ja kovarianssimatriisilla $\boldsymbol{\Sigma}$.

Todistus perustuu momenttiemäfunktion käyttöön sekä siihen tietoon, että matriisille $\boldsymbol{\Sigma}$ löytyy hajotelma $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^T$.

Merkintä $N_m(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

- Tästä lähtien multinormaalijakaumaa merkitään tunnuksella $N_m(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, jossa $\boldsymbol{\mu}$ on jakauman odotusarvo ja $\boldsymbol{\Sigma}$ sen kovarianssimatriisi. Alaindeksi m voidaan jättää pois, jos satunnaisvektorin $\mathbf{X}$ dimensiosta ei tarvitse pitää kirjaa.

Multinormaalijakauman $N_m(\boldsymbol{\mu}, \Sigma)$ simulointi

- Määritelmä on samalla simulointiresepti.
- Jos $\boldsymbol{\Sigma}$ on annettu symmetrinen ja positiivisesti semidefiniitti matriisi, niin ensin etsitään hajotelma (esim. Choleskyn hajotelma)
$$\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^T.$$
- Tämän jälkeen simuloidaan (riippumattomasti) m arvoa $(u_1, \ldots, u_m)$ yksiulotteisesta jakaumasta $N(0, 1)$.
- Lopuksi lasketaan
$$\mathbf{x} = \mathbf{A}\mathbf{u} + \boldsymbol{\mu}, \qquad \text{jossa } \mathbf{u} = (u_1, \ldots, u_n).$$
- Tuloksena on yksi simuloitu vektori jakaumasta $N_m(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Reseptiä toistamalla saadaan simuloitua arvoja riippumattomista tätä jakaumaa noudattavista satunnaisvektoreista.

10.3 Affiinin muunnoksen jakauma

Lause (Multinormaalisuus säilyy affiinissa muunnoksessa)

Olkoon $\mathbf{X} \sim N_m(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, ja olkoon $\mathbf{B} \in \mathbb{R}^{p \times m}$ vakiomatriisi ja $\mathbf{b} \in \mathbb{R}^p$ vakiovektori. Tällöin

$$\mathbf{BX} + \mathbf{b} \sim N_p(\mathbf{B}\boldsymbol{\mu} + \mathbf{b}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T).$$

Todistus: Käytetään esitystä (7):

$$\mathbf{BX} + \mathbf{b} \stackrel{\mathrm{d}}{=} \mathbf{B}(\mathbf{AU} + \boldsymbol{\mu}) + \mathbf{b} = (\mathbf{BA})\mathbf{U} + (\mathbf{B}\boldsymbol{\mu} + \mathbf{b}).$$

Määritelmän 2 mukaan sv:lla $\mathbf{BX} + \mathbf{b}$ on multinormaalijakauma. Lisäksi

$$\begin{aligned} E(\mathbf{BX} + \mathbf{b}) &= \mathbf{B}E(\mathbf{X}) + \mathbf{b} = \mathbf{B}\boldsymbol{\mu} + \mathbf{b}, \\ \mathrm{Cov}(\mathbf{BX} + \mathbf{b}) &= \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T. \end{aligned}$$

Normaalijakauman reunajakaumat

- Jos määritellään i:s yksikkövektori $\mathbf{e}_i$ siten, että sen i:s koordinaatti on yksi ja muut koordinaatit ovat nollia, niin

$$X_i = \mathbf{e}_i^T \mathbf{X}.$$

- Jos sv $\mathbf{Y}$ koostuu sv:n $\mathbf{X}$ komponenteista $(X_{i_1}, \ldots, X_{i_k})$, niin $\mathbf{Y}$ saadaan kertomalla sv:a $\mathbf{X}$ vasemmalta vakiomatriisilla, jonka j:s vaakarivi on $\mathbf{e}_{i_j}$,

$$\mathbf{Y} = \begin{bmatrix} X_{i_1} \\ \vdots \\ X_{i_k} \end{bmatrix} = \begin{bmatrix} \mathbf{e}_{i_1}^T \\ \vdots \\ \mathbf{e}_{i_k}^T \end{bmatrix} \mathbf{X}.$$

- Koska $\mathbf{Y}$ on affiini muunnos (peräti lineaarinen muunnos) sv:sta $\mathbf{X}$, niin $\mathbf{Y}$llä on multinormaalijakauma.
- Multinormaalijakauman kaikki reunajakaumat ovat multinormaalisia.

Reunajakauman parametrit

- Satunnaisvektorin $\mathbf{Y} = (X_{i_1}, \ldots, X_{i_k})$ jakauman parametrit voidaan tietenkin laskea lauseen 2 avulla.
- Ne saa helpommin selville poimimalla asiaankuuluvan osa $\mathbf{X}$:n jakauman parametreista, sillä

$$E\mathbf{Y} = \begin{bmatrix} EX_{i_1} \\ \vdots \\ EX_{i_k} \end{bmatrix} = \begin{bmatrix} \mu_{i_1} \\ \vdots \\ \mu_{i_k} \end{bmatrix},$$

$$(\mathrm{Cov}\,\mathbf{Y})(p, q) = \mathrm{cov}(X_{i_p}, X_{i_q}) = \mathbf{\Sigma}(i_p, i_q).$$

10.4 Tiheysfunktio

Lause (Multinormaalijakauman tiheysfunktio)

Jos $\boldsymbol{\Sigma}$ on symmetrinen ja positiivisesti definiitti matriisi, niin jakaumalla $N_m(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ on tiheysfunktio

$$f(\mathbf{x}) = (2\pi)^{-m/2} \, (\det(\boldsymbol{\Sigma}))^{-1/2} \, \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right). \tag{8}$$

Todistus onnistuu muuttujanvaihtokaavalla, kun käytetään säännöllistä $m \times m$ esitysmatriisia $\mathbf{A}$, jolle $\boldsymbol{\Sigma} = \mathbf{A}\,\mathbf{A}^T$ ja jonka avulla konstruoidaan $\mathbf{X} \sim N_m(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ muodossa

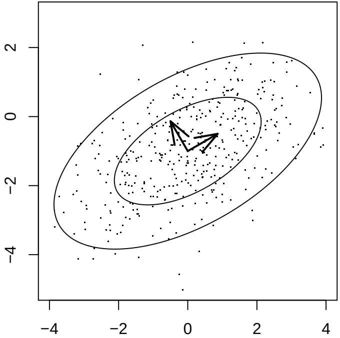
$$\mathbf{X} = \mathbf{A}\,\mathbf{U} + \boldsymbol{\mu}, \qquad \mathbf{U} \sim N_m(\mathbf{0}, \mathbf{I})$$

Huomautuksia tiheysfunktiosta

- Multinormaalijakaumalla $N_m(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ on tiheysfunktio täsmälleen silloin, kun $\boldsymbol{\Sigma}$ on positiivisesti definiitti.
- Jos $\boldsymbol{\Sigma}$ on pelkästään positiivisesti semidefiniitti, mutta ei positiivisesti definiitti, niin (lause 9.2) $\mathbf{X}$:llä ei voi olla tiheysfunktiota.
- Myös tällainen singulaarinen multinormaalijakauma on tärkeä jakauma sovelluksissa.
- Lineaaristen mallien yhteydessä toisaalta sovitevektorin jakauma ja toisaalta residuaalivektorin eli jäännösvektorin jakauma ovat kumpikin singulaarisia multinormaalijakaumia.

10.5 Tiheysfunktion tasa-arvopinnat

- Moniulotteisen normaalijakauman tiheysfunktion (8) tasa-arvopinnat ovat m-ulotteisia ellipsoideja.
- Tämä nähtäisiin soveltamalla kovarianssimatriisiin ominaisarvohajotelmaa (ks. moniste).



10.6 Korreloimattomuus ja riippumattomuus

- Riippumattomat satunnaisvektorit $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}$ eivät korreloi, sillä

$$\begin{aligned}\operatorname{cov}(\mathbf{X}, \mathbf{Y}) &= E[(\mathbf{X} - E\mathbf{X})(\mathbf{Y} - E\mathbf{Y})^T] \\ &= E[\mathbf{X} - E\mathbf{X}]\, E[(\mathbf{Y} - E\mathbf{Y})^T] = \mathbf{0}.\end{aligned}$$

- Yleisesti ottaen korreloimattomuudesta ei seuraa riippumattomuus.
- Jos komponenteista $\mathbf{X}$ ja $\mathbf{Y}$ yhdistetyn satunnaisvektorin $(\mathbf{X}, \mathbf{Y})$ yhteisjakauma on multinormaalinen, niin tässä tapauksessa korreloimattomuudesta seuraa riippumattomuus.

Satunnaisvektorin kokoaminen osista

- Kirjataan myöhempää käyttöä varten, minkälaiset ovat osavektorien $\mathbf{X} = (X_1, \ldots, X_k)$ ja $\mathbf{Y} = (Y_1, \ldots, Y_m)$ multinormaalijakaumien parametrit, jos yhdistetty vektori $(\mathbf{X}, \mathbf{Y})$ noudattaa multinormaalijakaumaa $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.
- Odotusarvovektori $\boldsymbol{\mu}$ koostuu osista

$$\boldsymbol{\mu} = E\mathbf{Z} = \begin{bmatrix} \boldsymbol{\mu}_X \\ \boldsymbol{\mu}_Y \end{bmatrix} = \begin{bmatrix} E\mathbf{X} \\ E\mathbf{Y} \end{bmatrix}, \tag{9}$$

Ositettu kovarianssimatriisi

- Satunnaisvektorin $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ kovarianssimatriisi $\boldsymbol{\Sigma} = \operatorname{Cov} \mathbf{Z}$ voidaan myös osittaa vastaavasti, sillä se on

$$\begin{bmatrix} \operatorname{cov}(X_1, X_1) & \dots & \operatorname{cov}(X_1, X_k) & \operatorname{cov}(X_1, Y_1) & \dots & \operatorname{cov}(X_1, Y_m) \\ \vdots & & \vdots & \vdots & & \vdots \\ \operatorname{cov}(X_k, X_1) & \dots & \operatorname{cov}(X_k, X_k) & \operatorname{cov}(X_k, Y_1) & \dots & \operatorname{cov}(X_k, Y_m) \\ \operatorname{cov}(Y_1, X_1) & \dots & \operatorname{cov}(Y_1, X_k) & \operatorname{cov}(Y_1, Y_1) & \dots & \operatorname{cov}(Y_1, Y_m) \\ \vdots & & \vdots & \vdots & & \vdots \\ \operatorname{cov}(Y_m, X_1) & \dots & \operatorname{cov}(Y_m, X_k) & \operatorname{cov}(Y_m, Y_1) & \dots & \operatorname{cov}(Y_m, Y_m) \end{bmatrix}$$

$$= \begin{bmatrix} \boldsymbol{\Sigma}_{XX} & \boldsymbol{\Sigma}_{XY} \\ \boldsymbol{\Sigma}_{YX} & \boldsymbol{\Sigma}_{YY} \end{bmatrix} = \begin{bmatrix} \operatorname{cov}(\mathbf{X}, \mathbf{X}) & \operatorname{cov}(\mathbf{X}, \mathbf{Y}) \\ \operatorname{cov}(\mathbf{Y}, \mathbf{X}) & \operatorname{cov}(\mathbf{Y}, \mathbf{Y}) \end{bmatrix} \tag{10}$$

- Tällöin reunajakaumat ovat

$$\mathbf{X} \sim N(\boldsymbol{\mu}_X, \boldsymbol{\Sigma}_{XX}), \qquad \mathbf{Y} \sim N(\boldsymbol{\mu}_Y, \boldsymbol{\Sigma}_{YY}). \tag{11}$$

Normaaliselle yhteisjakaumalle korreloimattomuus ja riippumattomuus ovat samoja asioita

Lause

Jos vektorilla $(\mathbf{X}, \mathbf{Y})$ on multinormaalijakauma, niin

$$\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \quad \Longleftrightarrow \quad \operatorname{cov}(\mathbf{X}, \mathbf{Y}) = \mathbf{0}.$$

Todistuksessa tarvitaan satunnaisvektorin $(\mathbf{X}, \mathbf{Y})$ momenttiemäfunktiota, ja lausetta 9.6.

Varoitus: normaalisuus ei ole automaattista

- Jos oletetaan multinormaaliset reunajakaumat

$$\mathbf{X} \sim N(\boldsymbol{\mu}_X, \boldsymbol{\Sigma}_{XX}), \qquad \mathbf{Y} \sim N(\boldsymbol{\mu}_Y, \boldsymbol{\Sigma}_{YY}),$$

  niin yhdistetyn vektorin $(\mathbf{X}, \mathbf{Y})$ jakauma ei välttämättä ole multinormaalinen.
- Edes lisäehto $\mathrm{cov}(\mathbf{X}, \mathbf{Y}) = \mathbf{0}$ ei auta asiaa.

Riippumattomien normaalijakautuneitten vektoreitten yhteisjakauma on normaalinen

Lause

Jos

$$\mathbf{X} \sim N(\boldsymbol{\mu}_X, \boldsymbol{\Sigma}_{XX}), \qquad \mathbf{Y} \sim N(\boldsymbol{\mu}_Y, \boldsymbol{\Sigma}_{YY}),$$

ja $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}$, niin vektori $(\mathbf{X}, \mathbf{Y})$ noudattaa multinormaalijakaumaa $N(\mu, \boldsymbol{\Sigma})$, jossa jakauman parametrit ovat

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_X \\ \boldsymbol{\mu}_Y \end{bmatrix}, \qquad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{XX} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{YY} \end{bmatrix}.$$

Todistus sujuu helposti momenttiemäfunktion avulla.

Esimerkki monimutkaisemmasta tilanteesta, jossa yhteisjakauman nähdään olevan normaalinen

- Olkoot $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}$ ja

$$\mathbf{X} \sim N(\boldsymbol{\mu}_X, \boldsymbol{\Sigma}_{XX}), \qquad \mathbf{Y} \sim N(\boldsymbol{\mu}_Y, \boldsymbol{\Sigma}_{YY}),$$

  ja olkoot $\mathbf{A}$ sekä $\mathbf{B}$ vakiomatriiseja siten, että vektorit $\mathbf{A}\,\mathbf{X}$ ja $\mathbf{B}\,\mathbf{Y}$ ovat samanpituisia.

- Tällöin satunnaisvektorilla

$$\mathbf{Z} = \mathbf{A}\,\mathbf{X} + \mathbf{B}\,\mathbf{Y}$$

  on multinormaalijakauma, sillä

$$\mathbf{A}\,\mathbf{X} + \mathbf{B}\,\mathbf{Y} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix}$$

Esimerkki jatkuu

- Satunnaisvektorin $\mathbf{Z} = \mathbf{A}\,\mathbf{X} + \mathbf{B}\,\mathbf{Y}$ jakauman parametrit saadaan helpoimmin selville laskemalla suoraan vektorin odotusarvo ja kovarianssimatriisi,

$$
\begin{aligned}
E\mathbf{Z} &= E(\mathbf{A}\,\mathbf{X} + \mathbf{B}\,\mathbf{Y}) = \mathbf{A}\,\mu_X + \mathbf{B}\,\mu_Y \\
\mathrm{Cov}\,\mathbf{Z} &= \mathrm{Cov}(\mathbf{A}\,\mathbf{X}) + \mathrm{Cov}(\mathbf{B}\,\mathbf{Y}) = \mathbf{A}\,\Sigma_{XX}\,\mathbf{A}^T + \mathbf{B}\,\Sigma_{YY}\,\mathbf{B}^T
\end{aligned}
$$

- Kovarianssin laskussa käytettiin riippumattomuutta sekä esimerkkiä 9.1

$$
\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \quad \Rightarrow \quad \mathbf{A}\,\mathbf{X} \perp\!\!\!\perp \mathbf{B}\,\mathbf{Y}
$$

10.7 Ehdolliset jakaumat

Lause

Olkoon $(\mathbf{X}, \mathbf{Y}) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, ja käytetään odotusarvovektorille $\boldsymbol{\mu}$ ositusta (9), ja kovarianssimatriisille $\boldsymbol{\Sigma}$ ositusta (10). Jos osamatriisi $\boldsymbol{\Sigma}_{XX}$ on säännollinen, niin sv:n $\mathbf{Y}$ jakauma ehdolla $\mathbf{X} = \mathbf{x}$ on multinormaalinen, ja ehdollisen jakauman parametrit ovat

$$\mathbf{Y} \mid (\mathbf{X} = \mathbf{x}) \sim N(\boldsymbol{\mu}_Y + \boldsymbol{\Sigma}_{YX}\boldsymbol{\Sigma}_{XX}^{-1}(\mathbf{x} - \boldsymbol{\mu}_X),\ \boldsymbol{\Sigma}_{YY} - \boldsymbol{\Sigma}_{YX}\boldsymbol{\Sigma}_{XX}^{-1}\boldsymbol{\Sigma}_{XY}).$$

- Huomaa, että ehdollinen odotusarvo $E[\mathbf{Y} \mid \mathbf{X} = \mathbf{x}]$ on argumentin $\mathbf{x}$ affiini funktio, ja että ehdollisen jakauman kovarianssimatriisi ei riipu pisteestä $\mathbf{x}$.
- Todistus käydään läpi taululla.

10.8 Kaksiulotteinen normaalijakauma

- Tarkastellaan kaksiulotteisen normaalijakauman $(X, Y) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ominaisuuksia.
- Olkoon $-1 < \rho = \mathrm{corr}(X, Y) < 1$, ja merkitään jakauman parametrien komponentteja seuraavasti,

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix}, \qquad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}.$$

- Tämän jakauman ominaisuudet (reunajakaumat, ytf, ehdolliset jakaumat jne.) saadaan selville erikoistapauksina n-ulotteisen tapauksen kaavoista.

Joitakin tuloksia

- Reunajakaumat ovat

$$X \sim N(\mu_X, \sigma_X^2), \qquad Y \sim N(\mu_Y, \sigma_Y^2).$$

- Ehdolliset jakaumat ovat

$$Y \mid (X = x) \sim N(\mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X),\ (1 - \rho^2)\sigma_Y^2),$$
$$X \mid (Y = y) \sim N(\mu_X + \rho \frac{\sigma_X}{\sigma_Y}(y - \mu_Y),\ (1 - \rho^2)\sigma_X^2).$$

10.9 Normaalijakauman otoskeskiarvon ja otosvarianssin yhteisjakauma

- Tarkastellaan sm:ia $X_1, \ldots, X_n$.
- Niiden otoskeskiarvo $\bar{X}$ ja otosvarianssi S^2 määritellään kaavoilla

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2. \tag{12}$$

- Jos sm:t X_i ovat riippumattomia, ja niillä on yhteinen odotusarvo μ ja yhteinen varianssi σ^2, niin helpohkoilla laskuilla voidaan näyttää, että

$$E\bar{X} = \mu, \qquad ES^2 = \sigma^2.$$

  Ts. otoskeskiarvo ja otosvarianssi ovat vastaavien populaatioparametrien μ ja σ^2 harhattomia estimaattoreita.

Satunnaisotos normaalijakaumasta $N(\mu, \sigma^2)$

- Oletetaan tästä lähtien, että sm:t $X_1, \ldots, X_n$ ovat riippumattomia, ja niillä kaikilla on normaalijakauma $N(\mu, \sigma^2)$.
- Tällöin vektorilla $(X_1, \ldots, X_n)$ on multinormaalijakauma $N_n(\mathbf{m}, \mathbf{\Sigma})$ parametreilla

$$\mathbf{m} = \begin{bmatrix} \mu \\ \vdots \\ \mu \end{bmatrix} = \mu\,\mathbf{1}, \quad \text{jossa} \quad \mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^n,$$

$$\mathbf{\Sigma} = \sigma^2 \mathbf{I}_n.$$

Alustavia huomioita

- Tämän jakson päätavoite on johtaa vektorin $(\bar{X}, S^2)$ yhteisjakauma.
- Otoskeskiarvon reunajakauma on helppo johtaa, nimittäin

$$\bar{X} \sim N(\mu, \frac{1}{n}\sigma^2),$$

  sillä $\bar{X}$ on multinormaalijakautuneen vektorin $\mathbf{X} = (X_1, \ldots, X_n)$ lineaarimuunnos, ja osaamme laskea sm:n $\bar{X}$ odotusarvon ja varianssin.
- Pian nähdään, että otosvarianssilla S^2 on sopivan skaalauksen jälkeen khiin neliön jakauma vapausasteluvulla $n - 1$.

Huomio khiin neliön jakaumasta

- Olkoon satunnaisvektorilla $\mathbf{Z}$ on k-ulotteinen standardinormaalijakauma, ts. $\mathbf{Z} \sim N_k(\mathbf{0}, \mathbf{I})$.
- Sen pituuden neliöllä $\|\mathbf{Z}\|^2 = \mathbf{Z}^T\mathbf{Z}$ on khiin neliön jakauma vapausasteluvulla k, sillä

$$\|\mathbf{Z}\|^2 = \mathbf{Z}^T\mathbf{Z} = \sum_{i=1}^{k} Z_i^2,$$

  jossa $Z_i \sim N(0, 1)$ riippumattomasti.
- Koska $EZ_i^2 = 1$, niin $E\|\mathbf{Z}\|^2 = k$.

Otoskeskiarvon ja -varianssin yhteisjakauma satunnaisotokselle normaalijakaumasta

Lause

Olkoot $X_1, \ldots, X_n \sim N(\mu, \sigma^2) \perp\!\!\!\perp$. Tällöin

a) $\bar{X} \sim N(\mu, \sigma^2/n)$,

b) $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$, joten $ES^2 = \sigma^2$,

c) $\bar{X} \perp\!\!\!\perp S^2$.

Todistuksesta: Kohta a todistettiin jo. Sen jälkeen todistetaan riippumattomuus (kohta c) ja lopuksi otosvarianssin reunajakauma (kohta b).

Riippumattomuuden $\bar{X} \perp\!\!\!\perp S^2$ todistuksen idea (c-kohta)

- Otoskeskiarvo $\bar{X}$ ja jäännösvektori (residuaalivektori) $\mathbf{R}$

$$\mathbf{R} = \begin{bmatrix} X_1 - \bar{X} \\ \vdots \\ X_n - \bar{X} \end{bmatrix}$$

  saadaan normaalijakautuneesta vektorista $\mathbf{X} = (X_1, \ldots, X_n)$ tietyllä lineaarisella muunnoksella.
- $\bar{X}$ ja $\mathbf{R}$ osoitetaan riippumattomiksi tarkistamalla, että ne ovat korreloimattomia.
- Riippumattomuus $\bar{X} \perp\!\!\!\perp S^2$ seuraa siitä, että S^2 voidaan laskea muunnoksena jäännösvektorista:

$$S^2 = \frac{1}{n-1} \|\mathbf{R}\|^2.$$

Todistuksen avainkaavoja (c-kohta)

- Määritellään n-komponenttinen vektori $\mathbf{u}$ kaavalla

$$\mathbf{u} = \frac{1}{\sqrt{n}}\mathbf{1},$$

jolloin $\mathbf{u}$ on ykkösvektori $\mathbf{1}$ normitettuna yksikkövektoriksi.

- Otoskeskiarvo voidaan esittää kaavalla

$$\bar{X} = \frac{1}{n}\mathbf{1}^T\mathbf{X} = \frac{1}{\sqrt{n}}\mathbf{u}^T\mathbf{X}.$$

- Tästä

$$\bar{X}\mathbf{1} = \left(\frac{1}{\sqrt{n}}\mathbf{u}^T\mathbf{X}\right)\sqrt{n}\,\mathbf{u} = \mathbf{u}\mathbf{u}^T\mathbf{X}$$
$$\mathbf{R} = \mathbf{X} - \bar{X}\mathbf{1} = \mathbf{X} - \mathbf{u}\mathbf{u}^T\mathbf{X} = (\mathbf{I}_n - \mathbf{u}\mathbf{u}^T)\mathbf{X}$$

Miksi otosvarianssille tulee skaalauksen jälkeen khiin neliön jakauma (b-kohta)

- Välivaiheiden jälkeen skaalattu otosvarianssi (eli skaalattu jäännösvektorin pituuden neliö) saadaan ilmaistua kaavalla

$$\frac{n-1}{\sigma^2}S^2 = \frac{1}{\sigma^2}\|\mathbf{R}\|^2 = \|\mathbf{Z}\|^2, \quad \text{jossa } \mathbf{Z} \sim N_{n-1}(\mathbf{0}, \mathbf{I}_{n-1}).$$

- Mihin yksi vapausaste katoaa?
- *Vastaus:* Se kuluu odotusarvon μ estimointiin otoskeskiarvolla $\bar{X}$.

Keskeinen idea (b-kohdan todistus)

- Määritellään $n \times n$ neliömatriisi $\mathbf{Q}$ siten, että sen ensimmäinen pystyrivi on $\mathbf{u}$.
- Muut matriisin $\mathbf{Q}$ pystyrivit $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ valitaan siten, että matriisin $\mathbf{Q}$ pystyrivit muodostavat yhdessä $\mathbb{R}^n$:n ortonormeeratun kannan. (Gramin-Shmidtin ortogonalisointi.)
- Tällöin $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_n$, ja koska $\mathbf{Q}$ on neliömatriisi, niin se on ortogonaalinen matriisi, eli $\mathbf{Q}^{-1} = \mathbf{Q}^T$.
- Kun käytetään $\mathbf{Q}$:n ositusta $\mathbf{Q} = \begin{bmatrix} \mathbf{u} & \mathbf{V} \end{bmatrix}$ jossa $n \times (n-1)$-matriisin $\mathbf{V}$ pystyrivit ovat $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$, niin

$$\mathbf{I}_n = \mathbf{Q}\mathbf{Q}^T = \mathbf{u}\mathbf{u}^T + \mathbf{V}\mathbf{V}^T.$$

- Tämän takia $\mathbf{R} = (\mathbf{I}_n - \mathbf{u}\mathbf{u}^T)\mathbf{X} = \mathbf{V}\mathbf{V}^T\mathbf{X}$.

t-jakauman määritelmä taas kerran

- t-jakauma vapausasteluvulla ν määritellään siten, että se on sm:n
$$\frac{Z}{\sqrt{Y/\nu}}$$
jakauma, kun $Z \sim N(0,1)$, $Y \sim \chi^2_\nu$ ja $Z \perp\!\!\!\perp Y$.

Tunnusluvun $(\bar{X}, S^2)$ yhteisjakaumasta t-jakaumaan

- Satunnaismuuttuja $\bar{X} - \mu \sim N(0, \frac{1}{n}\sigma^2)$, joten

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1).$$

- Jos tuntemattoman keskihajontaparametrin σ tilalle sijoitetaan sen otosestimaatti, saadaan ns. t-testisuure

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}.$$

- Tämä t-testisuure voidaan esittää yhtäpitävästi kaavalla

$$T = \frac{(\bar{X} - \mu)/(\sigma/\sqrt{n})}{\sqrt{S^2/\sigma^2}}.$$

- Tästä nähdään, että T:llä on t-jakauma $n - 1$ vapausasteella.

Jakson 10.9 lopuksi

- Edeltävä tarkastelu saataisiin vähällä vaivalla yleistettyä kattamaan vastaava tarkastelu yleisessä lineaarisessa mallissa.
- Siellä varianssiparametrin harhaton estimaattori ja sovitevektori ovat riippumattomia. Sopivasti skaalattuna varianssiparametrin estimaattorilla on khiin neliön jakauma vapausasteluvulla $n - p$, jossa p on (lineaarisesti riippumattomien) kertoimien lukumäärä.
- Äskeisessä lauseessa tuntemattomia kertoimia oli yksi, μ, joten $p = 1$. (Myös varianssiparametri on tuntematon, mutta sitä ei lasketa lineaarisen mallin kertoimeksi).
- Tarkemmin tätä asiaa käsitellään kurssilla Lineaariset mallit.