7 Kaksiulotteinen jakauma

- Diskreetti kaksiulotteinen jakauma — ks. luku 3.
- Uutta asiaa tässä kappaleessa:
  - jatkuva kaksiulotteinen jakauma
  - satunnaisvektorin muunnoksen odotusarvo
  - odotusarvovektori ja kovarianssimatriisi
  - jatkuvasti jakautuneen satunnaisvektorin sileän muunnoksen tiheysfunktio

7.1 Jatkuva kaksiulotteinen jakauma

Määritelmä (Jatkuva (yhteis)jakauma, yhteistiheysfunktio)

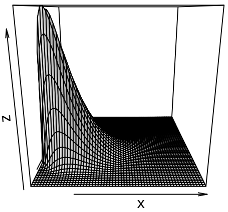
Satunnaisvektorilla (X, Y) on jatkuva jakauma, eli satunnaismuuttujilla X ja Y on jatkuva yhteisjakauma, mikäli on olemassa $\mathbb{R}^2$:lla määritelty funktio $f \geq 0$ siten, että

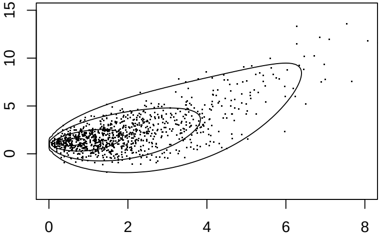
$$P((X, Y) \in B) = \iint_B f(x, y)\,\mathrm{d}x\,\mathrm{d}y, \qquad \text{kaikille } B \subset \mathbb{R}^2. \tag{1}$$

Tällöin $f = f_{X,Y}$ on sv:n (X, Y) tiheysfunktio, eli satunnaismuuttujien X ja Y *yhteistiheysfunktio* (lyh. ytf) (engl. *joint (probability) density function, joint pdf*).

Erilaisia tapoja visualisoida yhteistiheysfunktiota

- Tiheysfunktion kuvaaja perspektiivikuvana
- Tiheysfunktion tasa-arvokäyriä sekä jakaumasta simuloitu pisteparvi





Tasointegraali

- Jos tahdotaan laskea todennäköisyys $P((X, Y) \in B)$, niin jatkuvan jakauman tapauksessa joudutaan laskemaan tiheysfunktion tasointegraali joukon $B \subset \mathbb{R}^2$ yli.
- Joukon B yli laskettu tasointegraali funktiosta g tarkoittaa funktion $1_B\, g$ integraalia koko tason $\mathbb{R}^2$ yli.
- Sille käytetään mm. seuraavia merkintöjä,

$$\begin{aligned}\int_B g = \int_B g(x, y)\,\mathrm{d}(x, y) &= \iint_B g(x, y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \iint_{R^2} 1_B(x, y)\, g(x, y)\,\mathrm{d}x\,\mathrm{d}y.\end{aligned}$$

- Tavallisesti tasointegraalit lasketaan iteroituina integraaleina Fubinin lauseen avulla.

Fubinin lause: tasointegraali iteroituna integraalina

Lause (Fubini, Lebesguen integraalille)

Seuraavissa kahdessa tapauksissa tasointegraali voidaan laskea iteroituna integraalina, eli alla olevat yhtälöt ovat voimassa,

$$\begin{aligned}\int_B g(x,y)\,\mathrm{d}(x,y) &= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} 1_B(x,y)\,g(x,y)\,\mathrm{d}x\right)\mathrm{d}y \\ &= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} 1_B(x,y)\,g(x,y)\,\mathrm{d}y\right)\mathrm{d}x\end{aligned} \tag{2}$$

(a) Jos $g \geq 0$, jolloin integraalien yhteinen arvo voi olla ∞.

(b) Jos $\int_B |g| < \infty$, jolloin integraalien yhteinen arvo on äärellinen. (Huomaa, että integraali $\int_B |g|$ voidaan aina laskea a-kohdan perusteella iteroituna integraalina.)

Tasointegraali iteroituna integraalina: esimerkki 1

- Jos B voidaan esittää muodossa

$$B = \{(x,y) : a < x < b,\ c(x) < y < d(x)\},$$

ja $\int_B |g| < \infty$ tai $g \geq 0$, niin

$$\int_B g = \iint_B g(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_a^b \left(\int_{c(x)}^{d(x)} g(x,y)\,\mathrm{d}y \right) \mathrm{d}x.$$

- Sama tulos saadaan, vaikka joukon B määritelmässä yksi tai useampi relaatioista $<$ korvataan relaatiolla $\leq$.
- Tälle iteroidulle integraalille voidaan käyttää myös muita merkintöjä, kuten esim.

$$\int_B g = \int_{x=a}^b \int_{y=c(x)}^{d(x)} g(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_a^b \mathrm{d}x \int_{c(x)}^{d(x)} g(x,y)\,\mathrm{d}y$$

Tasointegraali iteroituna integraalina: esimerkki 2

- Jos B voidaan esittää muodossa

$$B = \{(x,y) : c < y < d,\ a(y) < x < b(y)\},$$

ja $\int_B |g| < \infty$ tai $g \geq 0$, niin tasointegraali voidaan laskea iteroituna integraalina

$$\iint_B g(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_c^d \left(\int_{a(y)}^{b(y)} g(x,y)\,\mathrm{d}x \right) \mathrm{d}y$$
$$= \int_{y=c}^{d} \int_{x=a(y)}^{b(y)} g(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_c^d \mathrm{d}y \int_{a(y)}^{b(y)} g(x,y)\,\mathrm{d}x$$

- Useimmat sovelluksissa esiintyvät integrointijoukot voidaan osittaa erillisiksi paloiksi, joille voidaan käyttää jompaa kumpaa esitystä. Koko alkuperäisen joukon yli laskettu tasointegraali voidaan laskea summana näiden palojen yli lasketuista tasointegraaleista.

Varoitus merkinnöistä

- Myös merkintä
$$\int_a^b \int_c^d g(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$
on yleinen.
- Ei ole itsestään selvää, kumpaan muuttujaan kumpikin integrointijoukko liitty:
  - onko se joukon $(x,y) \in (a,b) \times (c,d)$ yli laskettu integraali vai
  - joukon $(x,y) \in (c,d) \times (a,b)$ yli laskettu integraali?
- Suositus: käytä jotakin yksikäsitteisempää merkintää!

Tiheysfunktion (epä)yksikäsitteisyydestä

- Kaksiulotteisen jakauman tiheysfunktio ei ole yksikäsitteinen, mutta melkein yksikäsitteinen.
- Jos f on tiheysfunktio, ja g on saman jakauman tiheysfunktio, niin $f = g$ melkein kaikkialla, eli $f(x, y) = g(x, y)$ kaikkialla muualla paitsi nollamittaisessa joukossa.
- Kaksiulotteisessa avaruudessa esim. yksiulotteisten käyrien kuvaajat ovat nollamittaisia.

Reunajakauma yhteisjakaumasta

Lause

Olkoon satunnaisvektorilla (X, Y) jatkuva jakauma. Tällöin reunajakaumat ovat myös jatkuvia, ja niiden tiheysfunktiot saadaan integroimalla toinen muuttuja pois yhteistiheysfunktiosta,

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y, \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x.$$

Todistus: Jos $B \subset \mathbb{R}$, niin

$$P(X \in B) = P((X,Y) \in B \times \mathbb{R}) = \int_B \left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y \right) \mathrm{d}x.$$

Y:n tiheysfunktion lauseke johdetaan vastaavasti.

Onko yhteisjakauma jatkuva, jos reunajakaumat ovat jatkuvia?

- Otetaan sm X, jolla on jatkuva jakauma tf:lla f_X. Määritellään sm Y siten, että $Y = X$, ts.

$$Y(\omega) = X(\omega), \qquad \text{kaikilla } \omega \in \Omega.$$

- Sv:n (X, Y) jakauma on keskittynyt lävistäjälle

$$B = \{(x, y) : x = y\},$$

mutta mille tahansa funktiolle $f : \mathbb{R}^2 \to \mathbb{R}$ on $\int_B f = 0$, sillä lävistäjä B on yksiulotteisena joukkona nollamittainen $\mathbb{R}^2$:ssa.

- Tässä esimerkissä reunajakaumat ovat jatkuvia, mutta yhteisjakauma ei ole jatkuva.

Yhteistiheysfunktio yhteiskertymäfunktiosta—johdattelua

- Jatkuvan yhteisjakauman tapauksessa ykf:llä on esitys,

$$\begin{aligned} F_{X,Y}(x,y) &= P((X,Y) \in (-\infty, x] \times (-\infty, y]) \\ &= \int_{-\infty}^{x} \mathrm{d}s \int_{-\infty}^{y} f_{X,Y}(s,t)\, \mathrm{d}t. \end{aligned}$$

- Derivoidaan tämä yhtälö ensin muuttujan x suhteen,

$$\frac{\partial}{\partial x} F_{X,Y}(x,y) = \int_{-\infty}^{y} f_{X,Y}(x,t)\, \mathrm{d}t,$$

  ja sitten vielä muuttujan y suhteen,

$$\frac{\partial^2}{\partial x\, \partial y} F_{X,Y}(x,y) = f_{X,Y}(x,y),$$

  ja tämä tulos pätee ainakin niissä pisteissä (x,y), joissa $f_{X,Y}$ on jatkuva.

Yhteistiheysfunktio yhteiskertymäfunktiosta

Lause

Jos satunnaisvektorilla (X, Y) on jatkuva jakauma, niin jakauman tiheysfunktioksi voidaan valita

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x\,\partial y} F_{X,Y}(x,y).$$

Tämä toisen kertaluvun sekaderivaatta on olemassa melkein kaikilla (x, y), ja niissä pisteissä, joissa sekaderivaatta ei ole määritelty, voidaan käyttää mielivaltaista määrittelyä $f_{X,Y}$:lle.

- Todistus vaatisi geometrista mittateoriaa.
- Mielivaltaisesta annetusta kaksiulotteisen jakauman kertymäfunktiosta voi olla hankala tunnistaa, onko se jatkuvan jakauman kertymäfunktio.

Yhteistiheysfunktion tunnistamisesta

Lause

Jos $f : \mathbb{R}^2 \to \mathbb{R}$ on ei-negatiivinen, ja sen integraali koko tason yli on yksi, niin se on jonkin kaksiulotteisen satunnaisvektorin tiheysfunktio.

Huomautus: Jos $g : \mathbb{R}^2 \to \mathbb{R}$ on sellainen, että $g \geq 0$ ja $0 < \int_{R^2} g < \infty$, niin g:stä saadaan skaalaamalla yhteistiheysfunktio:

$$f(x, y) = \frac{1}{k} g(x, y), \quad \text{jossa } k = \int_{R^2} g$$

on yhteistiheysfunktio.

7.2 Tasajakauma alueessa

Määritelmä (Alueen A tasajakauma)

Olkoon $A \subset \mathbb{R}^2$ joukko, jonka pinta-ala

$$m(A) = \iint 1_A(x, y)\,\mathrm{d}x\,\mathrm{d}y$$

toteuttaa ehdot $0 < m(A) < \infty$. Sv (X, Y) noudattaa tasajakaumaa joukossa A, mikäli sillä on ytf

$$f(x, y) = 1_A(x, y)/m(A).$$

Huomioita alueen tasajakaumasta

- Olkoon vektorilla (X, Y) tasajakauma alueessa A. Jos $B \subset \mathbb{R}^2$, niin

$$P((X, Y) \in B) = \iint_B f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \frac{m(A \cap B)}{m(A)}.$$

  Erityisesti $P((X, Y) \in \mathbb{R}^2) = 1$, kuten pitää ollakin.

Esimerkki: tasajakauma kuvaajan alla

- Olkoon $h(x) = \exp(-\frac{1}{2}x^2)$, ja olkoon sv:lla (X, Y) tasajakauma h:n kuvaajan alla, eli joukossa

$$A = \{(x, y) : 0 < y < h(x)\}.$$

- Joukon A pinta-ala on

$$\begin{aligned} m(A) &= \iint_A 1_A(x, y)\, \mathrm{d}x\, \mathrm{d}y = \int_{-\infty}^{\infty} \mathrm{d}x \int_0^{h(x)} 1\, \mathrm{d}y \\ &= \int_{-\infty}^{\infty} h(x)\, \mathrm{d}x = \sqrt{2\pi}, \end{aligned}$$

joten ytf on

$$f_{X,Y}(x, y) = \frac{1}{\sqrt{2\pi}} 1\{0 < y < h(x)\},$$

Lasketaan reunatiheydet esimerkissä

- X:n reunatiheysfunktio on

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y = \int_0^{h(x)} \frac{1}{\sqrt{2\pi}}\,\mathrm{d}y = \frac{1}{\sqrt{2\pi}}\,\exp(-\frac{1}{2}x^2).$$

- X:n reunajakauma on $N(0,1)$.
- Mikä on Y:n reunatiheysfunktio?

Y:n reunatiheys esimerkissä

- Ratkaistaan ensin ytf:n indikaattorin epäyhtälö $0 < y < h(x)$ muuttujan x suhteen kiinteällä y.
- Kaikilla x on $0 < \exp(-\frac{1}{2}x^2) \le 1$, joten epäyhtälö

$$0 < y < \exp(-\frac{1}{2}x^2)$$

voi toteutua vain, kun $0 < y < 1$.

- Jos $0 < y < 1$, niin

$$0 < y < \exp(-\frac{1}{2}x^2) \quad \Leftrightarrow \quad -\sqrt{-2\ln y} < x < \sqrt{-2\ln y}.$$

- Tämän takia ytf voidaan esittää myös kaavalla

$$f_{X,Y}(x,y) = \frac{1}{\sqrt{2\pi}} 1\{0 < y < 1,\ -\sqrt{-2\ln y} < x < \sqrt{-2\ln y}\}.$$

Y:n reunatiheys

- Saatiin selville, että

$$f_{X,Y}(x,y) = \frac{1}{\sqrt{2\pi}} 1\{0 < y < 1,\ -\sqrt{-2\ln y} < x < \sqrt{-2\ln y}\},$$

- Kun $0 < y < 1$, on

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x = \int_{-\sqrt{-2\ln y}}^{\sqrt{-2\ln y}} \frac{1}{\sqrt{2\pi}}\,\mathrm{d}x \\ &= \frac{2}{\sqrt{\pi}}\sqrt{-\ln y}. \end{aligned}$$

- Muilla y:n arvoilla $f_Y(y) = 0$.

7.3 Riippumattomuus

- Palautetaan mieleen määritelmä. Sm:t X ja Y ovat riippumattomia (merkintä $X \perp\!\!\!\perp Y$), jos

$$P(X \in A, Y \in B) = P(X \in A)\, P(Y \in B), \qquad \text{kaikilla } A, B \subset \mathbb{R}.$$

- Jaksossa 3.5 todettiin, että

$$X \perp\!\!\!\perp Y \quad \iff \quad F_{X,Y}(x, y) = F_X(x)\, F_Y(y) \text{ kaikilla } (x, y).$$

- Nyt osoitetaan, että riippumattomien satunnaismuuttujien yptnf tai ytf voidaan esittää reunajakaumien vastaavien funktioiden tulona, jos yhteisjakauma on joko diskreetti tai jatkuva.

Riippumattomien sm:ien yptnf/ytf on tulomuotoa

Lause

Olkoon satunnaisvektorilla (X, Y) joko diskreetti tai jatkuva yhteisjakauma, jonka yptnf tai ytf on $f_{X,Y}$.

a) Jos $X \perp\!\!\!\perp Y$, niin yptnf:ksi/ytf:ksi kelpaa $f_{X,Y}(x, y) = f_X(x)\, f_Y(y)$.

b) Jos $f_{X,Y}$ voidaan esittää tulomuodossa

$$f_{X,Y}(x, y) = g(x)\, h(y), \qquad \forall x, y,$$

jossa $g \geq 0$ ja $h \geq 0$ ovat yhden muuttujan funktioita, niin $X \perp\!\!\!\perp Y$.

Todistus taululla.

Kuinka todistetaan, että X ja Y eivät ole riippumattomia?

- Mikäli X ja Y ovat diskreettejä, niin niiden riippuvuuden voi todistaa esittämällä yksi piste (x, y), jossa tulo-ominaisuus ei pidä paikkaansa:

$$f_{X,Y}(x, y) \neq f_X(x)\, f_Y(y).$$

- Jatkuvan yhteisjakauman tilanteessa edellinen ei riitä riippuvuuden todistamiseksi.
- Tämä johtuu siitä, että moni hieman erilainen funktio voi olla saman jakauman tiheysfunktio.

Esimerkki: Riippuvuuden todistaminen tasajakaumalle kolmiossa

- Olkoon vektorilla (X, Y) tasajakauma kolmiossa

$$C = \{(x, y) : x \geq 0, y \geq 0, x + y \leq 1\}.$$

  Ovatko X ja Y riippumattomia?

- Kolmion C pinta-ala on $\frac{1}{2}$, joten ytf on

$$f_{X,Y}(x, y) = 2, \quad \text{kun } (x, y) \in C.$$

- Äkkiseltään voisi näyttää, että ytf on tulomuotoa $g(x)\, h(y)$, mutta totuus paljastuu, kun kirjoitamme ehdon $(x, y) \in C$ indikaattorifunktioiden avulla:

$$f_{X,Y}(x, y) = 2\, 1_{\{x \geq 0\}}\, 1_{\{y \geq 0\}}\, 1_{\{x+y \leq 1\}}$$

  Tässä viimeinen termi ei todellakaan ole tulomuotoa.

Tasajakauma kolmiossa C: jatkoa

- Jos valitaan $A = B = (\frac{1}{2}, 1)$, niin $(A \times B) \cap C = \emptyset$, joten

  $$P((X,Y) \in A \times B) = 0,$$

  mutta $P(X \in A) > 0$ ja $P(Y \in B) > 0$, joten

  $$0 = P(X \in A, Y \in B) \neq P(X \in A)\, P(Y \in B) > 0.$$

- Olemme näyttäneet suoraan riipumattomuuden määritelmän perusteella, että X ja Y eivät ole riippumattomia.

7.4 Satunnaisvektorin muunnoksen odotusarvo

Lause

Olkoon (X, Y) sv, jolla on diskreetti jakauma, ja olkoon $Z = g(X, Y)$ jokin sen reaaliarvoinen muunnos. Tällöin

$$EZ = \sum_{x,y} g(x, y)\, f_{X,Y}(x, y),$$

mikäli summa suppenee itseisesti.

Todistus taululla.

Muunnoksen odotusarvo jatkuvassa tapauksessa

Jatkuvan yhteisjakauman tapauksessa odotusarvo $Eg(X, Y)$ saadaan laskettua integraalina. Sivuutamme tämän tapauksen todistuksen.

Lause

Olkoon (X, Y) sv, jolla on jatkuva yhteisjakauma, ja olkoon $Z = g(X, Y)$ jokin sen reaaliarvoinen muunnos. Tällöin

$$EZ = \iint_{R^2} g(x, y)\, f_{X,Y}(x, y)\, \mathrm{d}x\, \mathrm{d}y$$

mikäli integraali suppenee itseisesti.

Yhteenveto: tiedostamattoman tilastotieteilijän laki

- Mikäli kyseinen odotusarvo on (äärellisenä) olemassa, niin

$$Eg(X,Y) = \begin{cases} \sum_{x,y} g(x,y)\, f_{X,Y}(x,y), & \text{jos diskreetti} \\ \iint_{R^2} g(x,y)\, f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y, & \text{jos jatkuva.} \end{cases}$$

- Jos g riippuu vain yhdestä argumentista, niin esim. diskreetin yhteisjakauman tapauksessa saadaan

$$\begin{aligned} Eg(X) &= \sum_x \sum_y g(x)\, f_{X,Y}(x,y) = \sum_x g(x) \sum_y f_{X,Y}(x,y) \\ &= \sum_x g(x)\, f_X(x). \end{aligned}$$

- Ts. aikaisempi versio tiedostamattoman tilastotieteilijän laista on yhteensopiva tämän uuden version kanssa.

Esimerkki: lineaarisuus seurauksena tiedostamattoman tilastotieteilijän laista

Jos satunnaismuuttujilla X ja Y on jatkuva yhteisjakauma, ja ne ovat integroituvia, niin kaikilla vakioilla $a, b \in \mathbb{R}$

$$
\begin{aligned}
E(aX + bY) &= \iint (ax + by)\, f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= a \int x \left(\int f_{X,Y}(x,y)\,\mathrm{d}y \right) \mathrm{d}x + b \int y \left(\int f_{X,Y}(x,y)\,\mathrm{d}x \right) \\
&= a \int x\, f_X(x)\,\mathrm{d}x + b \int y\, f_Y(y)\,\mathrm{d}y = a\,EX + b\,EY.
\end{aligned}
$$

Tässä integraali tarkoittaa koko reaaliakselin yli laskettua integraalia.

Toinen esimerkki tutun laskusäännön tarkistamisesta

- Jos $X \perp\!\!\!\perp Y$, niin tällöin myös $g(X) \perp\!\!\!\perp h(Y)$, joten

$$E(g(X)\,h(Y)) = E(g(X))\,E(h(Y)),$$

mikäli kyseiset odotusarvot ovat olemassa.

- Jos sv:lla (X, Y) on jatkuva jakauma, niin $f_{X,Y}(x, y) = f_X(x)\,f_Y(y)$, joten asian näkee myös laskulla

$$\begin{aligned} E(g(X)\,h(Y)) &= \iint g(x)\,h(y) f_X(x)\,f_Y(y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \left(\int g(x)\,f_X(x)\,\mathrm{d}x\right)\left(\int h(y)\,f_Y(y)\,\mathrm{d}y\right) \\ &= E(g(X))\,E(h(Y)). \end{aligned}$$

7.5 Kovarianssi ja muita yhteismomentteja

Määritelmä

Olkoot $m, n \geq 0$ kokonaislukuja. Mikäli odotusarvo

$$E[X^m Y^n]$$

on olemassa, sitä kutsutaan kertalukujen (m, n) yhteismomentiksi (tai tulomomentiksi). Mikäli odotusarvo

$$E[(X - EX)^m (Y - EY)^n]$$

on olemassa, sitä sanotaan kertalukujen (m, n) keskusmomentiksi.

Huomioita yhteismomenteista

- Yhteismomenteista tärkeimmät ovat
  - reunajakaumien momentit EX^m sekä EY^n (eli kertalukujen $(m, 0)$ ja $(0, n)$ yhteismomentit)
  - kertaluvun $(1, 1)$ keskusmomentti, joka tunnetaan paremmin nimellä X:n ja Y:n kovarianssi,

$$\mathrm{cov}(X, Y) = E[(X - EX)(Y - EY)].$$

7.6 Paras lineaarinen ennuste

- Joissakin tilanteissa sm:n Y arvoa ei havaita suoraan, mutta jollakin tavalla sm:an Y liittyvän sm:n X arvo on mahdollista havaita.
- Tällöin saatetaan haluta ennustaa Y:n arvo X:n arvon avulla.
- Ennusteelle voidaan esimerkiksi valita lineaarinen muoto $a + bX$, tai yhtäpitävästi muoto $\alpha + \beta(X - EX)$.
- Miten kertoimet pitäisi valita jotta ennuste olisi mahdollisimman hyvä?
- Tarkastelun sivutuotteena saamme tulkinnan korrelaatiokertoimelle $\rho = \mathrm{corr}(X, Y)$.

Kriteeri ennusteen hyvyydelle

- Jos ennusteen hyvyyden mittana pidetään mahdollisimman pientä keskineliövirhettä

$$E[(Y - (\alpha + \beta(X - EX)))^2] = \text{min!},$$

niin paras lineaarinen ennuste löytyy yksinkertaisella laskulla.

- Muut tavat mitata ennusteen hyvyyttä voisivat olla yhtä mielekkäitä tai jopa mielekkäämpiä, mutta niille tehtävän ratkaisu ei välttämättä löytyisi helposti.
- Oletamme, että

$$EX^2 < \infty, \quad EY^2 < \infty, \quad \operatorname{var} X > 0, \quad \operatorname{var} Y > 0.$$

- Määritämme kertoimet α ja β siten, että keskineliövirhe minimoituu.

Ennustevirheen odotusarvo ja varianssi

- Merkitään $\mu_X = EX$ ja $\mu_Y = EY$.
- Jos α ja β ovat vakioita, niin ennuste $\alpha + \beta(X - \mu_X)$ poikkeaa totuudesta Y määrän

$$Z = Y - \alpha - \beta(X - \mu_X).$$

- Ennustevirheen odotusarvo ja varianssi ovat

$$\begin{aligned} EZ &= \mu_Y - \alpha - \beta\,(\mu_X - \mu_X) = \mu_Y - \alpha \\ \operatorname{var} Z &= \operatorname{var} Y + \beta^2 \operatorname{var} X - 2\beta \operatorname{cov}(X, Y). \end{aligned}$$

- Keskineliövirheeksi saadaan

$$\begin{aligned} EZ^2 &= (EZ)^2 + \operatorname{var} Z \\ &= (\mu_Y - \alpha)^2 + \operatorname{var} Y + \beta^2 \operatorname{var} X - 2\beta \operatorname{cov}(X, Y). \end{aligned}$$

Keskineliövirheen täydentäminen neliösummaksi

- Seuraavaksi täydennämme kerrointa β sisältävät termit neliöksi,

$$
\begin{aligned}
EZ^2 &= (\mu_Y - \alpha)^2 + \operatorname{var} Y \\
&\quad + \operatorname{var} X \left(\beta^2 - 2\beta \frac{\operatorname{cov}(X, Y)}{\operatorname{var} X} + \frac{(\operatorname{cov}(X, Y))^2}{(\operatorname{var} X)^2} \right) - \frac{(\operatorname{cov}(X, Y))^2}{\operatorname{var} X} \\
&= \operatorname{var} Y + (\mu_Y - \alpha)^2 \\
&\quad + \operatorname{var} X \left(\beta - \frac{\operatorname{cov}(X, Y)}{\operatorname{var} X} \right)^2 - \frac{(\operatorname{cov}(X, Y))^2}{\operatorname{var} X}
\end{aligned}
$$

- Tässä vain toinen ja kolmas termi riippuvat muuttujien α tai β arvoista. Nämä termit ovat molemmat ei-negatiivisia, ja ne saadaan molemmat helposti nollattua.

Tehtävän ratkaisu

- Valitaan α ja β miten tahansa, niin keskineliövirhe on aina suurempi tai yhtä kuin

$$\operatorname{var} Y - \frac{(\operatorname{cov}(X, Y))^2}{\operatorname{var} X} = (1 - \rho^2) \operatorname{var} Y, \tag{3}$$

- Keskineliövirheen minimi saavutetaan valitsemalla

$$\alpha = \mu_Y, \qquad \beta = \frac{\operatorname{cov}(X, Y)}{\operatorname{var} X}.$$

- Keskineliövirheen mielessä paras Y:n lineaarinen ennuste on

$$EY + \frac{\operatorname{cov}(X, Y)}{\operatorname{var}(X)} (X - EX) \tag{4}$$

Korrelaatiokerroin mittaa lineaarisen riippuvuuden voimakkuutta

- Parhaan lineaarisen ennusteen keskineliövirheen kaavasta $(1-\rho^2)\operatorname{var} Y$ sekä epäyhtälöstä $-1 \le \rho \le 1$ nähdään, kuinka korrelaatiokerroin mittaa muuttujien välisen lineaarisen riippuvuuden voimakkuutta.
- Jos korrelaatiokerroin saavuttaa jommankumman ääriarvonsa $\rho = \pm 1$, niin tällöin parhaan lineaarisen ennusteen keskineliövirhe EZ^2 on nolla. Toisaalta

  $$EZ^2 = 0 \quad \Rightarrow \quad Z = 0 \quad (\text{tn:llä } 1),$$

  joten jos $\rho = \pm 1$, niin

  $$Y(\omega) = EY + \frac{\operatorname{cov}(X,Y)}{\operatorname{var} X}(X(\omega) - EX)$$

  kaikilla alkeistapauksilla ω, paitsi sellaisilla, jotka kuuluvat poikkeusjoukkoon, jonka tn on nolla.

Korrelaatiokertoimen tulkinta jatkuu

- Korrelaatikertoimen etumerkki on sama kuin kovarianssin $\operatorname{cov}(X, Y)$ etumerkki, ja se on sama kuin parhaan lineaarisen ennusteen kulmakertoimen $\frac{\operatorname{cov}(X,Y)}{\operatorname{var} X}$ etumerkki.
- Jos $\rho > 0$ eli $\operatorname{cov}(X, Y) > 0$, niin Y:n arvolla on taipumus kasvaa X:n arvon kasvaessa, koska parhaan lineaarisen ennusteen kulmakerroin on tällöin positiivinen.
- Jos $\rho < 0$ eli $\operatorname{cov}(X, Y) < 0$, niin Y:n arvoilla on taipumus vähetä X:n arvon kasvaessa.

Huomautus

- Ratkaisimme parhaan lineaarista muotoa $\alpha + \beta(X - EX)$ olevan Y:n ennusteen X:n funktiona.
- Ratkaisemme myöhemmin vielä kunnianhimoisemman tehtävän, jossa kysytään, mikä on keskineliövirheen mielessä paras Y:n ennuste lausekkeella $m(X)$, jossa funktio m saadaan valita vapaasti.
- Jälkimmäisen tehtävän ratkaisu on ns. regressiofunktio.

7.7 Odotusarvovektori ja kovarianssimatriisi

- Kun satunnaisvektori $\mathbf{V} = (X, Y)$ esiintyy vektori- ja matriisilausekkeissa, pidämme sitä pystyvektorina.
- Varoitus: joissakin muissa lähteissä paria (X, Y) pidetään vaakavektorina.
- Näissä luentomuistiinpanoissa vektorit ja matriisit ovat paksunnettuja ja skalaarit paksuntamattomia.
- Edistyneemmissä esityksissä tällaista konventioita ei välttämättä käytetä, vaan lukijan pitää päätellä asiayhteydestä, minkätyyppistä oliota kukin symboli tarkoittaa.

Satunnaisvektorin odotusarvo

Määritelmä (Satunnaisvektorin odotusarvo)

Satunnaisvektorin (X, Y) odotusarvo(vektori) määritellään ottamalla satunnaisvektorista odotusarvot komponentti komponentilta, eli

$$E\mathbf{V} = E(X, Y) = E\begin{bmatrix} X \\ Y \end{bmatrix} = (EX, EY) = \begin{bmatrix} EX \\ EY \end{bmatrix}$$

mikäli kaikkien komponenttien odotusarvot ovat olemassa.

- Koska kaksikomponenttista vektoria (X, Y) voidaan pitää myös pitää 2×1-matriisina, niin satunnaisvektorin odotusarvon määritelmä on erikoistapaus seuraavaksi esitettävästä satunnaismatriisin odotusarvon määritelmästä.

Satunnaismatriisin odotusarvo

Määritelmä (Satunnaismatriisi ja sen odotusarvo)

Satunnaismatriisi on matriisi, jonka alkiot ovat satunnaismuuttujia. Sen odotusarvo on se vakiomatriisi, jonka alkio kohdassa (i, j) on ko. satunnaismatriisin alkion (i, j) odotusarvo (mikäli kaikkien alkoiden odotusarvot ovat olemassa).

Esimerkiksi

$$E \begin{bmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \end{bmatrix} = \begin{bmatrix} EZ_{11} & EZ_{12} & EZ_{13} \\ EZ_{21} & EZ_{22} & EZ_{23} \end{bmatrix}$$

Määritelmän seurauksia

- Jos $\mathbf{Z}$ on satunnaismatriisi, niin

$$E(\mathbf{Z}^T) = (E\mathbf{Z})^T \tag{5}$$

- Odotusarvon lineaarisuuden avulla nähdään helposti, että

$$E(\mathbf{AZB} + \mathbf{C}) = \mathbf{A}(E\mathbf{Z})\mathbf{B} + \mathbf{C}, \tag{6}$$

  kun $\mathbf{Z}$ on satunnaismatriisi ja $\mathbf{A}$, $\mathbf{B}$ ja $\mathbf{C}$ ovat vakiomatriiseja, joiden dimensiot ovat sellaiset, että lauseke on määritelty.

- Vakiomatriisit saadaan vetää ulos odotusarvon alta, jos ne sijaitsevat matriisitulossa äärimmäisenä vasemmalla tai äärimmäisenä oikealla. (Ei muuten!) Kaava (6) pätee myös satunnaisvektorille $\mathbf{Z}$, sillä d-komponenttinen pystyvektori voidaan tulkita $d \times 1$-matriisiksi.

Kovarianssimatriisi

Määritelmä (Kovarianssimatriisi)

Satunnaisvektorin $\mathbf{V} = (X, Y)$ kovarianssimatriisi määritellään kaavalla

$$\mathrm{Cov}(\mathbf{V}) = E[(\mathbf{V} - E\mathbf{V})(\mathbf{V} - E\mathbf{V})^T]. \tag{7}$$

- Käytän satunnaisvektorin kovarianssimatriisille merkintää $\mathrm{Cov}(\mathbf{V})$ ja satunnaismuuttujien X ja Y kovarianssille merkintää $\mathrm{cov}(X, Y)$. Nämä käsitteet ovat läheistä sukua toisilleen.
- Sv:n $\mathbf{V}$ kovarianssimatriisille käytetään kirjallisuudessa monia muitakin merkintöjä, kuten esim. $\mathrm{var}(\mathbf{V})$.
- Englannin kielellä kovarianssimatriisista käytetään ainakin nimityksiä *covariance matrix*, *variance–covariance matrix*, *variance matrix*, *dispersion matrix*.

Satunnaisvektorin $\mathbf{V} = (X, Y)$ kovarianssimatriisi

$$\begin{aligned}
\mathrm{Cov}(\mathbf{V}) &= E\left\{\begin{bmatrix} X - EX \\ Y - EY \end{bmatrix}\begin{bmatrix} X - EX & Y - EY \end{bmatrix}\right\} \\
&= E\begin{bmatrix} (X - EX)(X - EX) & (X - EX)(Y - EY) \\ (Y - EY)(X - EX) & (Y - EY)(Y - EY) \end{bmatrix} \\
&= \begin{bmatrix} \mathrm{cov}(X, X) & \mathrm{cov}(X, Y) \\ \mathrm{cov}(Y, X) & \mathrm{cov}(Y, Y) \end{bmatrix} = \begin{bmatrix} \mathrm{var}(X) & \mathrm{cov}(X, Y) \\ \mathrm{cov}(X, Y) & \mathrm{var}(Y) \end{bmatrix}
\end{aligned}$$

- Päälävistäjällä on muuttujien varianssit ja muualla muuttujien väliset kovarianssit.
- Kovarianssimatriisia kutsutaan myös kömpelöllä nimellä varianssi-kovarianssimatriisi.

Eräs tapa kirjoittaa vektorin (X, Y) kovarianssimatriisi

- Merkitään $\sigma_X^2 = \mathrm{var}(X)$ ja $\sigma_Y^2 = \mathrm{var}(Y)$.
- Jos $\sigma_X, \sigma_Y > 0$, niin satunnaismuuttujien X ja Y korrelaatiokerroin $\rho = \mathrm{corr}(X, Y)$ on

$$\rho = \mathrm{corr}(X, Y) = \frac{\mathrm{cov}(X, Y)}{\sqrt{\mathrm{var}\, X}\sqrt{\mathrm{var}\, Y}} = \frac{\mathrm{cov}(X, Y)}{\sigma_X\, \sigma_Y}$$

- Tämän takia sv:n $\mathbf{V} = (X, Y)$ kovarianssimatriisi voidaan kirjoittaa myös muodossa

$$\mathrm{Cov}(\mathbf{V}) = \begin{bmatrix} \sigma_X^2 & \rho\, \sigma_X \sigma_Y \\ \rho\, \sigma_X \sigma_Y & \sigma_Y^2 \end{bmatrix}$$

Satunnaisvektorin affiinin muunnoksen kovarianssimatriisi

- Oletetaan, että satunnaisvektori $\mathbf{W}$ saadaan satunnaisvektorista $\mathbf{V}$ affiinilla muunnoksella,

$$\mathbf{W} = \mathbf{A}\mathbf{V} + \mathbf{b},$$

  jossa $\mathbf{A}$ on vakiomatriisi ja $\mathbf{b}$ on vakiovektori.

- Tällöin

$$E\mathbf{W} = \mathbf{A}(E\mathbf{V}) + \mathbf{b}$$

  joten

$$(\mathbf{W} - E\mathbf{W})(\mathbf{W} - E\mathbf{W})^T = \mathbf{A}(\mathbf{V} - E\mathbf{V})(\mathbf{V} - E\mathbf{V})^T\mathbf{A}^T,$$

- Tämän takia

$$\operatorname{Cov}(\mathbf{A}\mathbf{V} + \mathbf{b}) = \mathbf{A}\operatorname{Cov}(\mathbf{V})\,\mathbf{A}^T. \tag{8}$$

Esimerkki kaavan $\mathrm{Cov}(\mathbf{A}\mathbf{V}+\mathbf{b}) = \mathbf{A}\,\mathrm{Cov}(\mathbf{V})\,\mathbf{A}^T$ soveltamisesta

- Olkoon

$$\mathbf{V} = \begin{bmatrix} X \\ Y \end{bmatrix}, \quad \mathbf{A} = \mathbf{u}^T, \quad \text{jossa} \quad \mathbf{u} = \begin{bmatrix} a \\ b \end{bmatrix}$$

- Tällöin saadaan johdettua tuttu kaava, nimittäin

$$\begin{aligned} \mathrm{var}(aX+bY) &= \mathrm{var}(\mathbf{u}^T\mathbf{V}) = \mathbf{u}^T\,\mathrm{Cov}(\mathbf{V})\,\mathbf{u} \\ &= \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} \mathrm{var}(X) & \mathrm{cov}(X,Y) \\ \mathrm{cov}(X,Y) & \mathrm{var}(Y) \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \\ &= a^2\,\mathrm{var}(X) + 2ab\,\mathrm{cov}(X,Y) + b^2\,\mathrm{var}(Y). \end{aligned}$$

7.8 Tiheysfunktion muuntokaava

- Diskreetissä tapauksessa ei tarvita teoriaa.
- Jos sv:illa (X, Y) on diskreetti jakauma, ja satunnaismuuttuja tai -vektori Z määritellään sen muunnoksena, $Z = g(X, Y)$, niin tällöin Z:lla on diskreetti jakauma, jonka ptnf voidaan laskea suoraan määritelmän perusteella, sillä

$$f_Z(z) = P(Z = z) = \sum_{(x,y):g(x,y)=z} f_{X,Y}(x, y).$$

- Jatkuvan jakauman tapauksessa tarvitaan monimutkaisempia laskuja.

Perusoletuksia jatkuvassa tapauksessa

- Olkoon (X, Y) jatkuvasti jakautunut satunnaisvektori, jolla on tf $f_{X,Y}$.
- Olkoon $\mathbf{g}$ vektoriargumentin vektoriarvoinen funktio $\mathbf{g} : A \to B$, jossa $A, B \subset \mathbb{R}^2$.
- Olkoon A on sellainen avoin joukko, että

$$P((X, Y) \in A) = 1. \tag{9}$$

Satunnaisvektorin muunnoksen jakauma

- Määrittelemme kaksiulotteisen satunnaisvektorin (U, V) kaavalla
$$\begin{bmatrix} U \\ V \end{bmatrix} = \mathbf{g}(X, Y) = \begin{bmatrix} g_1(X, Y) \\ g_2(X, Y) \end{bmatrix}$$
- Mikä on vektorin (U, V) jakauma? Onko se jatkuva?
- Miten saamme laskettua yhteistiheysfunktion $f_{U,V}$?

Diffeomorfismi

- Oletamme, että

$$\mathbf{g} : A \to B \quad \text{on diffeomorfismi} \tag{10}$$

eli että

  - joukot A ja B ovat avoimia,
  - funktio $\mathbf{g} : A \to B$ on jatkuvasti derivoituva bijektio,
  - sen käänteisfunktio $\mathbf{h} = \mathbf{g}^{-1}$ on jatkuvasti derivoituva bijektio $B \to A$.

- Osoittautuu, että tällöin $f_{U,V}$ saadaan yksinkertaisella kaavalla $f_{X,Y}$:stä.

Diffeomorfisuuden tarkistaminen käytännössä

1. Johda käänteisfunktiolle $\mathbf{h} = (h_1, h_2)$ kaava ratkaisemalla yhtälöistä

$$g_1(x, y) = u, \quad g_2(x, y) = v, \qquad (u, v) \in B$$

muuttujat x ja y muuttujien u ja v funktiona.

2. Mikäli ratkaisu

$$(x, y) = (h_1(u, v), h_2(u, v))$$

on yksikäsitteinen kaikilla $(u, v) \in B$ ja ratkaisu $(x, y) \in A$, niin tällöin ollaan tarkistettu, että kyseessä on bijektiivinen vastaavuus.

3. Tarkista, ovatko lausekkeet $g_1(x, y)$ ja $g_2(x, y)$ jatkuvasti derivoituvia A:lla ja lausekkeet $h_1(u, v)$ ja $h_2(u, v)$ jatkuvasti derivoituvia B:llä.

Jatkuvasti derivoituva—mitä se tarkoittaa?

- Esim. lausekkeen $g_1(x, y)$ kohdalla mieti, ovatko molemmat ensimmäisen kertaluvun osittaisderivaatat $\partial g_1(x, y)/\partial x$ ja $\partial g_2(x, y)/\partial y$ olemassa ja jatkuvia funktioita joukossa A.
- Toista tarkastelu lausekkeille $g_2(x, y)$, $h_1(u, v)$ ja $h_2(u, v)$.
- Derivaattojen jatkuvuus on tyypillisesti ilmiselvää, jos lausekkeet on saatu selville.

Derivaattamatriisi

- Nyt ajattelemme bijektiivistä vastaavuutta

$$(u, v) = (g_1(x, y), g_2(x, y)) \quad \Longleftrightarrow \quad (x, y) = (h_1(u, v), h_2(u, v)), \tag{11}$$

jota on mukavaa merkitä myös

$$(u, v) = (u(x, y), v(x, y)) \quad \Longleftrightarrow \quad (x, y) = (x(u, v), y(u, v)). \tag{12}$$

Kuvauksen **h** derivaatta(matriisi) eli Jacobin matriisi pisteessä (u, v) on

$$\begin{bmatrix} \dfrac{\partial h_1(u, v)}{\partial u} & \dfrac{\partial h_1(u, v)}{\partial v} \\ \dfrac{\partial h_2(u, v)}{\partial u} & \dfrac{\partial h_2(u, v)}{\partial v} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{bmatrix} \tag{13}$$

Jacobiaani

- Jacobin matriisin eli derivaattamatriisin determinanttia kutsutaan kuvauksen **h** *funktionaalideterminantiksi* eli *Jacobin determinantiksi* eli jacobiaaniksi (engl. *Jacobian determinant; Jacobian*).
- Myös nimiä Jakobin determinantti tai jakobiaani käytetään; käsitteen esitti 1800-luvulla saksalainen matemaatikko Carl Gustav Jacobi.
- Varoitus: joissakin lähteissä termi jacobiaani saattaa tarkoittaa derivaattamatriisia eikä sen determinanttia.

Jacobiaanin merkintätapoja

- Me käytämme jacobiaanille sekä modernia merkintää

$$J_{\mathbf{h}}(u, v),$$

että 1800-luvun puolivälistä lähtien käytössä ollutta merkintää

$$\frac{\partial(x, y)}{\partial(u, v)}.$$

- Klassinen merkintä on modernia merkintää paljon kätevämpi silloin, kun funktiot on määritelty lausekkeiden avulla.
- Tässä

$$J_{\mathbf{h}}(u, v) = \frac{\partial(x, y)}{\partial(u, v)} = \det \begin{bmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{bmatrix} = \frac{\partial x}{\partial u}\,\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\,\frac{\partial y}{\partial u}.$$

Tiheysfunktion muuntokaava

Lause

Jos $\mathbf{g} : A \to B$ on diffeomorfisimi, ja $P((X, Y) \in A) = 1$, niin satunnaisvektorilla $(U, V) = \mathbf{g}(X, Y)$ on jatkuva jakauma tf:llä

$$f_{U,V}(u, v) = \begin{cases} f_{X,Y}(h_1(u, v), h_2(u, v)) \, |J_{\mathbf{h}}(u, v)|, & \text{kun } (u, v) \in B \\ 0, & \text{muuten.} \end{cases} \tag{14}$$

Tässä $\mathbf{h} = (h_1, h_2)$ on funktion $\mathbf{g}$ käänteisfunktio.

Todistus seuraa suoraan tasointegraalien muuntokaavasta. (Jotta saataisiin näin yksinkertainen lopputulos, tarvitaan Lebesguen integraalia.)

Muistisääntö tiheysfunktion muuntokaavalle

- Diffeomorfismin tapauksessa muuttujanvaihtokaava (14) kannattaa pitää mielessään muodossa

$$f_{X,Y}(x,y)\,|\partial(x,y)| = f_{U,V}(u,v)\,|\partial(u,v)|, \quad \text{kun} \tag{15}$$

$$(u,v) = \mathbf{g}(x,y) \quad \Leftrightarrow \quad (x,y) = \mathbf{h}(u,v), \qquad (x,y) \in A, (u,v) \in B. \tag{16}$$

- Tästä sitten tuntemattomat suureet ratkaistaan tunnettujen suureiden avulla.
- Jos $f_{X,Y}$ on annettu, niin

$$\begin{aligned} f_{U,V}(u,v) &= f_{X,Y}(x,y)\left|\frac{\partial(x,y)}{\partial(u,v)}\right| \\ &= f_{X,Y}(h_1(u,v),h_2(u,v))\,|J_{\mathbf{h}}(u,v)|, \qquad (u,v) \in B. \end{aligned}$$

Toinen tapa ratkaista $f_{U,V}$ muistisäännöstä

- Ytf:n voi ratkaista muistisäännöstä myös seuraavasti,

$$f_{U,V}(u,v) = \frac{f_{X,Y}(x,y)}{\left|\dfrac{\partial(u,v)}{\partial(x,y)}\right|} = \frac{f_{X,Y}(h_1(u,v), h_2(u,v))}{|J_{\mathbf{g}}(h_1(u,v), h_2(u,v))|},$$

  kun $(u,v) \in B$.

- Tämäkin kaava on oikein, sillä funktion $\mathbf{g}$ ja sen käänteisfunktion $\mathbf{h}$ derivaattamatriisit ovat toistensa käänteismatriiseja, josta seuraa jacobiaaneille yhteys

$$1 = J_{\mathbf{h}}(u,v)\, J_{\mathbf{g}}(x,y) = \frac{\partial(x,y)}{\partial(u,v)}\, \frac{\partial(u,v)}{\partial(x,y)}. \tag{17}$$

  Tässä kaavassa (x,y) ja (u,v) vastaavat toisiaan bijektiivisesti.

Pätevyysalue

- Tiheysfunktion muuntokaavaa (14) sovellettaessa on tärkeää pitää kirjaa siitä, missä joukossa B johdettu kaava on pätevä.
- Tämän saa usein hoidettua automaattisesti joukkojen indikaattorien avulla (ks. seuraava esimerkki).
- Jatkotarkasteluja varten (esim. reunajakauman johtamista varten) on usein tarpeen esittää alue B esim. muodossa

$$B = \{(u, v) : a < u < b, c(u) < v < d(u)\}$$

tai muodossa

$$B = \{(u, v) : c < v < d, a(v) < u < b(v)\}.$$

Tästä voi aiheutua paljon lisätyötä.

Esimerkki 1

- Olkoon sv:lla (X, Y) jatkuva yhteisjakauma tf:lla

$$f_{X,Y}(x,y) = \begin{cases} k(x,y) & \text{kun } x > 0 \text{ ja } y > 0 \\ 0 & \text{muuten.} \end{cases} \tag{18}$$

- Määritellään

$$U = X + Y, \qquad V = X - Y.$$

- Johda $f_{U,V}$.
- Johda lisäksi f_U.

Indikaattorifunktiotekniikka

- Esitetään ytf

$$f_{X,Y}(x,y) = \begin{cases} k(x,y) & \text{kun } x > 0 \text{ ja } y > 0 \\ 0 & \text{muuten.} \end{cases}$$

  tason ensimmäisen neljänneksen indikaattorin $1\{x > 0, y > 0\}$ ja lausekkeen $k(x,y)$ tulona,

$$f_{X,Y}(x,y) = 1\{x > 0, y > 0\}\, k(x,y).$$

- Tässä lauseke $k(x,y)$ ei ole välttämättä hyvin määritelty, jos $x \leq 0$ tai $y \leq 0$.
- Ylläoleva esitys pitää ymmärtää lyhennemerkintänä pidemmälle kaavalle.
- Käytämme konventiota, jossa joukon indikaattori tarvittaessa nollaa mahdollisesti huonosti määritellyn lausekkeen.

Esim. 1: Diffeomorfinen vastaavuus

- Tehtävänannon mukaan $U = X + Y$ ja $V = X - Y$.
- Tätä vastaa diffeomorfismi

$$\begin{cases} u = x + y \\ v = x - y \end{cases} \qquad \iff \qquad \begin{cases} x = \frac{1}{2}(u + v) \\ y = \frac{1}{2}(u - v) \end{cases}$$

  joka on diffeomorfismi koko tason ja koko tason välillä.
- Käänteismuunnoksen $(x(u, v), y(u, v))$ jacobiaani on

$$\frac{\partial(x, y)}{\partial(u, v)} = -\frac{1}{2}.$$

Esim. 1: $f_{U,V}$ pystytään nyt kirjoittamaan

Satunnaismuuttujien U ja V ytf on

$$
\begin{aligned}
f_{U,V}(u,v) &= f_{X,Y}(x,y) \left| \frac{\partial(x,y)}{\partial(u,v)} \right| \\
&= 1\{u+v>0, u-v>0\}\, k(\frac{1}{2}(u+v), \frac{1}{2}(u-v))\, \frac{1}{2}.
\end{aligned}
$$

Miten tästä johdetaan reunatiheys f_U?

Esim. 1: Reunatiheys f_U

- Saatiin

$$f_{U,V}(u,v) = 1\{u+v>0, u-v>0\}\, k(\frac{1}{2}(u+v), \frac{1}{2}(u-v))\, \frac{1}{2}.$$

- Tietenkin $f_U(u) = \int_{-\infty}^{\infty} f_{U,V}(u,v)\,\mathrm{d}v$.
- Epäyhtälöpari $u+v>0$ ja $u-v>0$ pitää onnistua ratkaisemaan muodossa

$$a<u<b, \quad c(u)<v<d(u).$$

- Hetken pohtimisen jälkeen (piirrä kuva) nähdään, että ratkaisu on

$$u>0, \quad -u<v<u.$$

- Tämän takia U:n reunatiheys on

$$f_U(u) = \int_{-u}^{u} \frac{1}{2}\, k(\frac{1}{2}(u+v), \frac{1}{2}(u-v))\,\mathrm{d}v, \qquad u>0.$$

Kuinka johdetaan skalaarin $g_1(X, Y)$ tiheys?

- Diffeomorfismi on aina kuvaus kahden samandimensioisen euklidisen avaruuden välillä.
- Usein tavoitteena on johtaa jonkin skalaariarvoisen muunnoksen $U = g_1(X, Y)$ tiheysfunktio.
- Tämä voidaan tehdä kahdella erilaisella tekniikalla.
  1. Lasketaan ensin U:n kertymäfunktio

     $$F_U(u) = P(g_1(X, Y) \leq u) = \int_{\{(x,y): g_1(x,y) \leq u\}} f_{X,Y}(x, y)\,\mathrm{d}x\,\mathrm{d}y,$$

     ja sitten U:n tiheysfunktio derivoimalla.
  2. Täydennetään muunnos (keinotekoisesti) bijektioksi valitsemalla $V = g_2(X, Y)$, sitten johdetaan sv:n (U, V) tiheysfunktio, ja muuttujan U tiheysfunktio lasketaan integroimalla ytf:a $f_{U,V}(u, v)$.

Esimerkki 2: summan tiheysfunktio

- Olkoon sv:lla (X, Y) tf $f_{X,Y}$.
- Olkoon
$$U = X + Y.$$
- Johda sm:n U tf.

Esim. 2: ratkaisu bijektioksi täydentämällä

- Täydennämme kuvauksen bijektioksi valitsemalla sm:n $U = X + Y$ rinnalle sm:n $V = X$. (Muut valinnat toimisivat yhtä hyvin.)
- Tällöin

$$\begin{cases} u = x + y \\ v = x \end{cases} \iff \begin{cases} x = v \\ y = u - v \end{cases}$$

- Tästä

$$f_{U,V}(u,v) = f_{X,Y}(x,y) \left| \frac{\partial(x,y)}{\partial(u,v)} \right| = f_{X,Y}(v, u - v).$$

- Sm:n U reunatiheysfunktio saadaan integroimalla v pois,

$$f_U(u) = \int_{-\infty}^{\infty} f_{X,Y}(v, u - v) \, \mathrm{d}v$$

Esim. 2: ratkaisu kertymäfunktiotekniikalla

- Kun lasketaan sm:n $U = X + Y$ kertymäfunktiota pisteessä u, niin integrointijoukko on

$$\{(x, y) : x + y \leq u\} = \{(x, y) : x \in \mathbb{R}, y \leq u - x\}.$$

- Tämän takia

$$F_U(u) = \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{u-x} f_{X,Y}(x, y) \, \mathrm{d}y.$$

- Mikäli derivaatan vienti integraalin alle pystytään perustelemaan jotenkin, niin arvataan, että täytyy olla

$$\begin{aligned} f_U(u) &= F_U'(u) \\ &= \int_{-\infty}^{\infty} \mathrm{d}x \, \frac{\partial}{\partial u} \int_{-\infty}^{u-x} f_{X,Y}(x, y) \, \mathrm{d}y \\ &= \int_{-\infty}^{\infty} f_{X,Y}(x, u - x) \, \mathrm{d}x. \end{aligned}$$

Esim. 2: kommentti ratkaisutavoista

- Jälkimmäisessä laskussa jää arveluttamaan se, mitä (jatkuvuus-)ehtoja ytf:lle pitäisi asettaa, jotta derivaatta saadaan viedä integraalimerkin alle.
- Ensimmäisessä tavassa nähtiin, että tuloksen johtamiseksi ytf:ltä ei vaadita mitään ehtoja.

Konvoluutio

- Jos $X \perp\!\!\!\perp Y$, niin $f_{X,Y}(x,y) = f_X(x)\, f_Y(y)$ identtisesti, ja

$$
\begin{aligned}
f_U(u) &= \int_{-\infty}^{\infty} f_{X,Y}(v, u-v)\,\mathrm{d}v \\
&= \int_{-\infty}^{\infty} f_X(v)\, f_Y(u-v)\,\mathrm{d}v.
\end{aligned}
$$

- Jos $X \perp\!\!\!\perp Y$, niin sm:n $U = X + Y$ tiheysfunktio f_U on tiheysfunktioiden f_X ja f_Y konvoluutio,

$$f_U = f_X * f_Y,$$

joka määritellään yo. kaavalla, ts.

$$(f_X * f_Y)(u) = \int_{-\infty}^{\infty} f_X(v)\, f_Y(u-v)\,\mathrm{d}v.$$

7.9 t-jakauman ominaisuuksia

- Studentin t-jakauma vapausasteluvulla $\nu > 0$ eli t_ν-jakauma määriteltiin niin, että se on sm:n

$$Y = \frac{Z}{\sqrt{X/\nu}} \tag{19}$$

jakauma, kun $X \sim \chi^2_\nu = \mathrm{Gam}(\frac{\nu}{2}, \frac{1}{2})$, $Z \sim N(0,1)$, ja $X \perp\!\!\!\perp Z$.

- Johdamme t-jakauman tiheysfunktion esimerkkinä muuttujanvaihtotekniikan käytöstä.
- Laskemme jakauman odotusarvon ja varianssin käyttämällä yo. jakauman stokastista esitystä.

Ytf:n laskemisen vaiheita

- Koska $X \perp\!\!\!\perp Z$, on

$$f_{X,Z}(x,z) = f_X(x)\, f_Z(z) = \frac{(\frac{1}{2})^{\nu/2}}{\Gamma(\frac{\nu}{2})}\, x^{\frac{\nu}{2}-1}\, \mathrm{e}^{-\frac{1}{2}x}\, \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{-\frac{1}{2}z^2}, \qquad x > 0.$$

- Täydennetään kuvaus bijektioksi valitsemalla $U = X$. Tällöin tarkasteltavana on diffeomorfismi

$$\begin{cases} u = x \\ y = \dfrac{z}{\sqrt{x/\nu}} \end{cases} \qquad \Longleftrightarrow \qquad \begin{cases} x = u \\ z = y\sqrt{u/\nu} \end{cases}$$

jossa $x > 0$ ja $u > 0$.

Muunnoksen ytf ja sen marginalisointi

- Siis

$$f_{U,Y}(u,y) = f_{X,Z}(x,z) \left| \frac{\partial(x,z)}{\partial(u,y)} \right| = f_{X,Z}(u, y\sqrt{\frac{u}{\nu}}) \sqrt{\frac{u}{\nu}}, \qquad u > 0.$$

- Tästä saadaan Y:n reunatiheys integroimalla u pois,

$$\begin{aligned} f_Y(y) &= \int_0^\infty f_{X,Z}(u, y\sqrt{u/\nu}) \sqrt{\frac{u}{\nu}} \, \mathrm{d}u \\ &= \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2}) \sqrt{\nu\pi}} \frac{1}{(1 + y^2/\nu)^{(\nu+1)/2}}. \end{aligned}$$

- Integroitiin tilastotieteilijän tapaan (gammajakauman normalisointivakio!).

Huomioita t_ν-jakauman tiheysfunktion kaavasta

- t_ν-jakauman tiheysfunktioksi saatiin

$$f_Y(y) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\nu\pi}} \frac{1}{(1+y^2/\nu)^{(\nu+1)/2}}.$$

- Vapausasteluvulla $\nu = 1$ saadaan

$$f_Y(y) = \frac{1}{\pi}\frac{1}{1+y^2},$$

joten t_1 on Cauchyn jakauma.

Studentin t-jakauman momentit

- Studentin t-jakaumalla ei ole olemassa kaikkia absoluuttisia momentteja.
- Jos $a > 0$, niin $E|Y|^a < \infty$ täsmälleen silloin kuin $a < \nu$.
- Tämän näkee jakauman stokastisesta esityksestä $Y = Z/\sqrt{X/\nu}$, josta (koska $X \perp\!\!\!\perp Z$)

$$E|Y|^a = \nu^{a/2}\, E|Z|^a\, EX^{-a/2}.$$

- Normaalijakauman absoluuttiset momentit $E|Z|^a$ ovat äärellisiä, mutta $EX^{-a/2}$ on äärellinen jos ja vain jos $a < \nu$ (kirjoita odotusarvo integraalina ja muista että $X \sim \mathrm{Gam}(\nu/2, 1/2)$).

Seurauksia

Siis jos $Y \sim t_\nu$ ja $a \geq \nu$, niin $E|Y|^a = \infty$. Seurauksia:

- Cauchyn jakaumalla (eli t_1:llä) ei ole odotusarvoa.
- t_ν-jakaumalla on olemassa varianssi vain silloin, kun $\nu > 2$.
- t-jakauman momenttiemäfunktio ei ole olemassa missään origon ympäristössä, joten momenttiemäfunktion on hyödytön työkalu tälle jakaumalle.

Studentin t-jakauman odotusarvo ja varianssi

- Käytetään stokastista esitystä $Y = Z/\sqrt{X/\nu} = Z\sqrt{\nu/X}$.
- Jos $\nu > 1$, niin t_ν-jakaumalla on odotusarvo, ja odotusarvo on nolla, sillä (koska $Z \perp\!\!\!\perp X$)

$$EY = EZ\, E\sqrt{\frac{\nu}{X}} = 0 \cdot E\sqrt{\frac{\nu}{X}} = 0.$$

- Jos $\nu > 2$, niin t_ν-jakauman varianssi on

$$\operatorname{var} Y = EY^2 = EZ^2\, E\frac{\nu}{X} = 1 \cdot E\frac{\nu}{X}.$$

- Tästä nähdään muutaman välivaiheen jälkeen (integroi kuten tilastotieteilijä ja käytä gammafunktion funktionaaliyhtälöä), että

$$\operatorname{var} Y = \frac{\nu}{\nu - 2}, \qquad \text{kun } \nu > 2.$$