

## 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)

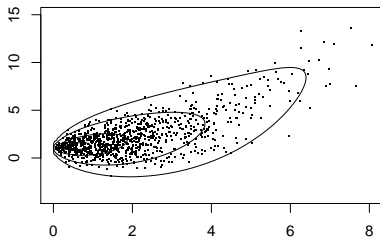
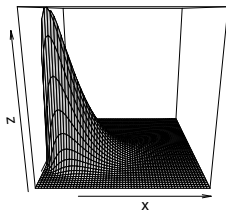
Satunnaisvektorilla  $(X, Y)$  on jatkuva jakauma, eli satunnaismuuttujilla  $X$  ja  $Y$  on jatkuva yhteisjakauma, mikäli on olemassa  $\mathbb{R}^2$ :lla määritelty funktio  $f \geq 0$  siten, että

$$P((X, Y) \in B) = \iint_B f(x, y) \, dx \, dy, \quad \text{kaikille } B \subset \mathbb{R}^2. \quad (1)$$

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,

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

- 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,

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

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

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

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

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

## 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

$$\begin{aligned} \iint_B g(x, y) \, dx \, dy &= \int_c^d \left( \int_{a(y)}^{b(y)} g(x, y) \, dx \right) dy \\ &= \int_{y=c}^d \int_{x=a(y)}^{b(y)} g(x, y) \, dx \, dy = \int_c^d dy \int_{a(y)}^{b(y)} g(x, y) \, dx \end{aligned}$$

- 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.

- Myös merkintä

$$\int_a^b \int_c^d g(x, y) \, dx \, dy.$$

on yleinen.

- Ei ole itsestään selvää, kumpaan muuttujaan kumpikin integrointijoukko liittyy:
  - 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,

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx.$$

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

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

$Y$ :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.

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

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

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

mutta mille tahansa funktiolle  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  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 yhteiskertymäfunktioista—johdattelua

- Jatkuvan yhteisjakauman tapauksessa ykf:llä on esitys,

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

- Derivoidaan tämä yhtälö ensin muuttujan  $x$  suhteen,

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

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 yhteiskertymäfunktioista

## Lause

Jos satunnaisvektorilla  $(X, Y)$  on jatkuva jakauma, niin jakauman tiheysfunktioiksi 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 kertymäfunktioista voi olla hankala tunnistaa, onko se jatkuvan jakauman kertymäfunktio.

# Yhteistiheysfunktion tunnistamisesta

## Lause

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

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

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

on yhteistiheysfunktio.

## 7.2 Tasajakauma alueessa

### Määritelmä (Alueen $A$ tasajakauma)

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

$$m(A) = \iint 1_A(x, y) \, dx \, dy$$

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

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

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

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

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

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

- $X$ :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  $0 < \exp(-\frac{1}{2}x^2) \leq 1$ , joten epäyhtälö

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

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

- Jos  $0 < y < 1$ , niin

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

- Tämän takia ytf voidaan esittää myös kaavalla

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

- Saatiin selville, että

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

- Kun  $0 < y < 1$ , on

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

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

## 7.3 Riippumattomuus

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

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

- Jaksossa 3.5 todettiin, että

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

- Nyt osoitetaan, että riippumattomien satunnaismuuttujien ypdf 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

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

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:

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

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 riippumattomuuden 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

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

mikäli integraali suppenee itseisesti.

# Yhteenveto: tiedostamattoman tilastotieteilijän laki

- Mikäli kyseinen odotusarvo on (äärellisenä) olemassa, niin

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

- 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  $a, b \in \mathbb{R}$

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

Tässä integraali tarkoittaa koko reaaliakselin yli laskettua integraalia.

## Toinen esimerkki tutun laskusäännön tarkistamisesta

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

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

## 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.

- 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,

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

- Ennestään tuttuja **kovarianssin ominaisuuksia**
  - symmetria:  $\text{cov}(X, Y) = \text{cov}(Y, X)$
  - varianssi kovarianssin avulla:  $\text{var } X = \text{cov}(X, X)$ .
  - identiteetti:  $\text{cov}(X, Y) = E(XY) - (EX)(EY)$ .



# Kovarianssin bilinearisuus

## Lause

Kovarianssi on **bilineaarinen**, eli molempien argumenttiensa suhteen lineaarinen operaattori. Ts. jos  $a_1, a_2 \in \mathbb{R}$  ovat vakioita ja  $X_1, X_2$  ja  $Z$  ovat satunnaismuuttujia, niin

$$\text{cov}(a_1 X_1 + a_2 X_2, Z) = a_1 \text{cov}(X_1, Z) + a_2 \text{cov}(X_2, Z)$$

$$\text{cov}(Z, a_1 X_1 + a_2 X_2) = a_1 \text{cov}(Z, X_1) + a_2 \text{cov}(Z, X_2)$$

Jos lisäksi  $b_1, b_2 \in \mathbb{R}$  ja  $Y_1, Y_2$  ovat satunnaismuuttujia, niin

$$\text{cov}(a_1 X_1 + a_2 X_2, b_1 Y_1 + b_2 Y_2) = \sum_{i=1}^2 \sum_{j=1}^2 a_i b_j \text{cov}(X_i, Y_j).$$

Nämä kaavat pitävät paikkansa, mikäli kyseiset odotusarvot ovat olemassa.

## Esimerkki: lineaarikombinaation varianssin laskeminen kovarianssin bilineaarisuuden avulla

$$\begin{aligned}\text{var}(aX + bY) &= \text{cov}(aX + bY, aX + bY) \\ &= a \text{cov}(X, aX + bY) + b \text{cov}(Y, aX + bY) \\ &= a^2 \text{cov}(X, X) + ab \text{cov}(X, Y) + ab \text{cov}(Y, X) + b^2 \text{cov}(Y, Y) \\ &= a^2 \text{var}(X) + 2ab \text{cov}(X, Y) + b^2 \text{var}(Y).\end{aligned}$$

## 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  $\rho = \text{corr}(X, Y)$ .

# Kriteeri ennusteen hyvyydelle

- Jos ennusteen hyvyyden mittana pidetään mahdollisimman pientä keskineliövirhettä

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

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

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

- Määritämme kertoimet  $\alpha$  ja  $\beta$  siten, että keskineliövirhe minimoituu.

# Ennustevirheen odotusarvo ja varianssi

- Merkitään  $\mu_X = EX$  ja  $\mu_Y = EY$ .
- Jos  $\alpha$  ja  $\beta$  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

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

- Keskineliövirheeksi saadaan

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

# Keskineliövirheen täydentäminen neliösummaksi

- Seuraavaksi täydennämme kerrointa  $\beta$  sisältävät termit neliöksi,

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

- Tässä vain toinen ja kolmas termi riippuvat muuttujien  $\alpha$  tai  $\beta$  arvoista. Nämä termit ovat molemmat ei-negatiivisia, ja ne saadaan molemmat helposti nollattua.

# Tehtävän ratkaisu

- Valitaan  $\alpha$  ja  $\beta$  miten tahansa, niin keskineliövirhe on aina suurempi tai yhtä kuin

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

- Keskineliövirheen minimi saavutetaan valitsemalla

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

- Keskineliövirheen mielessä paras  $Y$ :n lineaarinen ennuste on

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

# Korrelaatiokerroin mittaa lineaarisen riippuvuuden voimakkuutta

- Parhaan lineaarisen ennusteen keskineliövirheen kaavasta  $(1 - \rho^2) \text{var } Y$  sekä epäyhtälöstä  $-1 \leq \rho \leq 1$  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

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

kaikilla alkeistapauksilla  $\omega$ , paitsi sellaisilla, jotka kuuluvat poikkeusjoukkoon, jonka tn on nolla.



# Korrelaatiokertoimen tulkinta jatkuu

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

- 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.

## 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  $2 \times 1$ -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 alkioiden 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

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

- Odotusarvon lineaarisuuden avulla nähdään helposti, että

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

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.

## Määritelmä (Kovarianssimatriisi)

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

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

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

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

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

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



# Eräs tapa kirjoittaa vektorin $(X, Y)$ kovarianssimatriisi

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

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

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

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

# 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

$$\text{Cov}(\mathbf{A}\mathbf{V} + \mathbf{b}) = \mathbf{A} \text{Cov}(\mathbf{V}) \mathbf{A}^T. \quad (8)$$

# Esimerkki kaavan $\text{Cov}(\mathbf{AV} + \mathbf{b}) = \mathbf{A} \text{Cov}(\mathbf{V}) \mathbf{A}^T$ 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

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

## 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 ptmf 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  $\mathbf{g} : A \rightarrow B$ , 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  $f_{U,V}$ ?

- Oletamme, että

$$\mathbf{g} : A \rightarrow B \quad \text{on diffeomorfismi} \quad (10)$$

eli että

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

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

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 vastaavauutta

$$(u, v) = (g_1(x, y), g_2(x, y)) \iff (x, y) = (h_1(u, v), h_2(u, v)), \quad (11)$$

jota on mukavaa merkitä myös

$$(u, v) = (u(x, y), v(x, y)) \iff (x, y) = (x(u, v), y(u, v)). \quad (12)$$

Kuvauksen  $\mathbf{h}$  derivaatta(matriisi) eli **Jacobin matriisi** pisteessä  $(u, v)$  on

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

- 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 ollut merkintää

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

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

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

# Tiheysfunktion muuntokaava

## Lause

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

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

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

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

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

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

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

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

- Ytf:n voi ratkaista muistisäännöstä myös seuraavasti,

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

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

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

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

- 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

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

- Määritellään

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

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

- 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 lyhennemerkinä pidemmälle kaavalle.
- Käytämme konventiota, jossa joukon indikaattori tarvittaessa nolaa mahdollisesti huonosti määritellyn lausekkeen.

## Esim. 1: Diffeomorfinen vastaavuus

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

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

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

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

Miten tästä johdetaan reunatiheys  $f_U$ ?

## Esim. 1: Reunatiheys $f_U$

- Saatiin

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

- Tietenkin  $f_U(u) = \int_{-\infty}^{\infty} f_{U,V}(u, v) dv$ .
- 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

$$f_U(u) = \int_{-u}^u \frac{1}{2} k\left(\frac{1}{2}(u+v), \frac{1}{2}(u-v)\right) dv, \quad u > 0.$$

# 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

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

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,

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



## Esim. 2: ratkaisu kertymäfunktioiteknikalla

- 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

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

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

$$\begin{aligned} f_U(u) &= F'_U(u) \\ &= \int_{-\infty}^{\infty} dx \frac{\partial}{\partial u} \int_{-\infty}^{u-x} f_{X,Y}(x, y) dy \\ &= \int_{-\infty}^{\infty} f_{X,Y}(x, u-x) dx. \end{aligned}$$

## Esim. 2: kommentti ratkaisutavoista

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

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

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

- Jos  $X \perp Y$ , 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.

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

## 7.9 $t$ -jakauman ominaisuuksia

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

$$Y = \frac{Z}{\sqrt{X/\nu}} \quad (19)$$

jakauma, kun  $X \sim \chi_\nu^2 = \text{Gam}(\frac{\nu}{2}, \frac{1}{2})$ ,  $Z \sim N(0, 1)$ , 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

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

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

$$\begin{cases} u = x \\ y = \frac{z}{\sqrt{x/\nu}} \end{cases} \iff \begin{cases} x = u \\ z = y\sqrt{u/\nu} \end{cases}$$

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

# Muunnoksen ytf ja sen marginalisointi

- Siis

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

- Tästä saadaan  $Y$ :n reunatiheys integroimalla  $u$  pois,

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

- Integroitiin tilastotieteilijän tapaan (gammajakauman normalisointivakio!).

# Huomioita $t_\nu$ -jakauman tiheysfunktion kaavasta

- $t_\nu$ -jakauman tiheysfunktiksi 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  $X \sim \text{Gam}(\nu/2, 1/2)$ ).



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.
- $t_\nu$ -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  $t_\nu$ -jakaumalla on odotusarvo, ja odotusarvo on nolla, sillä (koska  $Z \perp X$ )

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

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

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

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

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