Satunnaismuuttujat

- Satunnaismuuttuja kuvaa kokeen tuloksen reaaliakselille
- Satunnaismuuttuja on siis funktio $X(s)$ joka saa reaalisia arvoja
- Koe s kuuluu todennäköisyysavaruuteen S ($s \in S$)
- Esimerkkejä
  - Kolikon heitossa todennäköisyysavaruus $S = \{\text{kruunu,klaava}\}$
    Satunnaismuuttuja voisi olla

$$X(s) = \begin{cases} 1 & (s = \text{kruunu}) \\ -1 & (s = \text{klaava}) \end{cases}$$

– Nopan heitossa $S = \{1, 2, 3, 4, 5, 6\}$.

  ◦ Satunnaismuuttuja voi olla $X(s) = s$, jolloin mahdolliset tulokset ovat 1,2,3,4,5,6

  ◦ Satunnaismuuttuja voi olla myös $X(s) = s^2$, jolloin mahdolliset tulokset ovat 1,4,9,16,25,36

- Usein merkitään yksinkertaisuuden vuoksi $X(s) = X$, mutta on syytä muistaa mistä on kysymys

- Kompleksinen satunnaismuuttuja Z on kahden reaalisen satunnaismuuttujan X ja Y funktio siten että $Z = X + jY$. Muuttujan Z ominaisuuksia tarkasteltaessa on tiedettävä ko. summan omaisuudet

- Satunnaismuuttuja voi saada äärellisen määrän eri arvoja → *diskreetti satunnaismuuttuja* (kuten edellisissä esimerkeissä)
- Satunnaismuuttuja voi myös saada kaikkia mahdollisia arvoja joltain tietyltä väliltä reaaliakselilla → *jatkuva satunnaismuuttuja* (kuten elektronisten laitteiden generoiman kohinan amplitudi)

- Käytännön tutkimustyössä tehdään usein seuraavaa:
  - Näytteistetyn satunnaissignaalin (tiedonsiirtokanavasta vastaanotettu signaali) elementit ovat diskreettejä satunnaismuuttujia, mutta analyysissä niitä kohdellaan jatkuvina, koska jatkuvia muuttujia on helpompi käsitellä
  - Jos sananleveys on riittävä (useita bittejä, ei kahta), niin näin tehty analyysi vastaa kokeellisia tuloksia
  - Pienille sananleveyksille (muutama bitti) näin tehty analyysi ei ole tarkka vaan analyysi ja simulaatiotulokset (joissa käytetään ko. sananleveyttä) poikkevat jonkin verran toisistaan

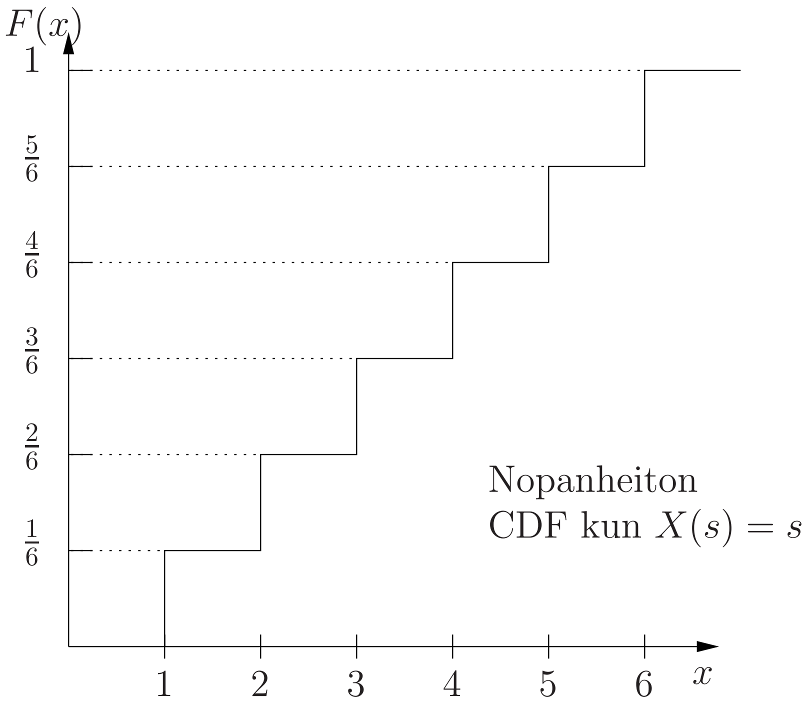
Satunnaismuuttujat ja todennäköisyys

- Tarkastellaan tapahtumaa jossa satunnaismuuttuja X on lukua x pienempi eli tarkastellaan tapahtumaa $\{X \leq x\}$.
- x voi olla mikä tahansa reaaliluku väliltä $(-\infty < x < \infty)$
- Tapahtuman todennäköisyys on $\mathrm{P}(X \leq x)$
- Yleensä se merkitään lyhyesti $F(x)$ (tai $F_X(x)$ jos on vaara sekoittaa johonkin muuhun vastaavaan funktioon) eli

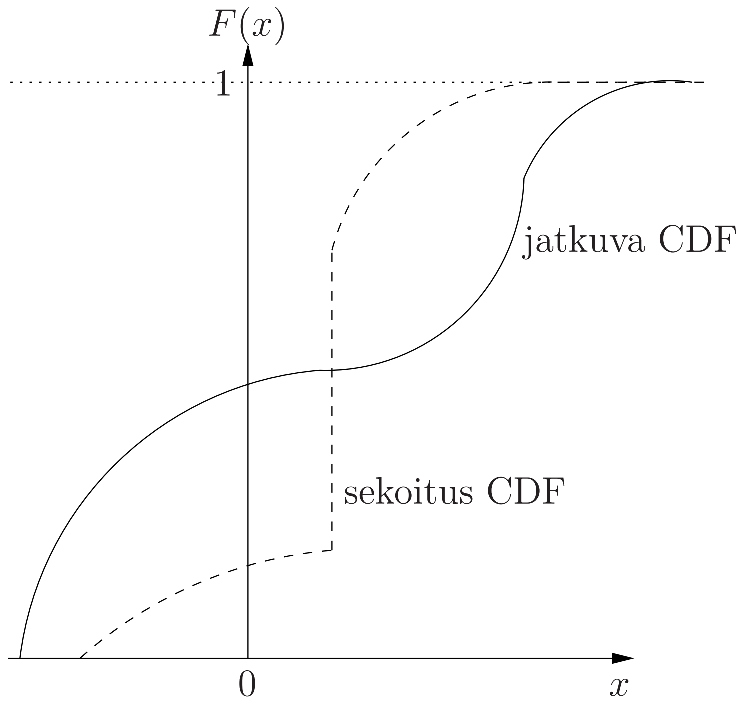
$$F(x) = \mathrm{P}(X \leq x) \quad (-\infty < x < \infty) \tag{22}$$

- Funktiota $F(x)$ kutsutaan *todennäköisyys kertymäfunktioksi* tai *kumulatiiviseksi kertymäfunktioksi* (CDF, cumulative distribution funktion)
- CDF siis kuvaa millä todennäköisyydellä satunnaismuuttuja X on lukua x pienempi

- CDF on todennäköisyys joten $0 \leq F(x) \leq 1$
- On myös helppo havaita että $F(-\infty) = 0$ ja $F(\infty) = 1$
- CDF on kasvava funktio eli $F(x_1) \leq F(x_2)$, $x_1 < x_2$
- Seuraavassa esimerkkejä diskreetistä ja jatkuvasta kertymäfunktiosta sekä näiden sekoituksesta



Diskreetti CDF



Jatkuva ja sekoitus CDF:t

- Kertymäfunktion $F(x)$ derivaattaa $p(x)$ (myös $p_X(x)$) kutsutaan satunnaismuuttujan X *todennäköisyys tiheysfunktioksi* (PDF, probability density function)
- Nyt siis pätee

$$p(x) = \frac{d\,F(x)}{dx} \quad (-\infty < x < \infty) \tag{23a}$$

ja

$$F(x) = \int_{-\infty}^{x} p(u)\,du \quad (-\infty < x < \infty) \tag{23b}$$

- Koska $F(x)$ on kasvava, niin $p(x) \geq 0$
- Selvästi $\int_{-\infty}^{\infty} p(x)\,dx = F(\infty) = 1$ (pinta-ala on 1)

- Diskreettien muuttujien tapauksessa $p(x)$ sisältää impulssin $F(x)$:n epäjatkuvuuskohdassa x_i eli tälläisessa kohdassa $p(x)$ on

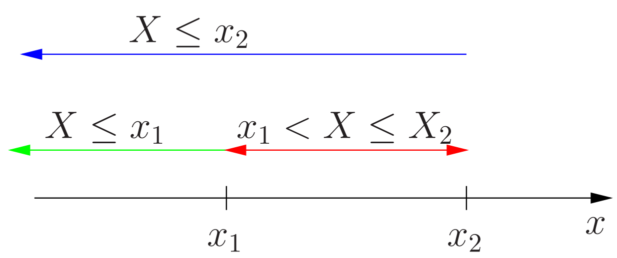
$$p(x_i) = \mathrm{P}(X = x_i)\delta(x - x_i) \tag{24}$$

jossa $\delta(x)$ on impulssifunktio eli

$$\delta(x) = \begin{cases} 1 & x = 0 \\ 0 & x \neq 0 \end{cases} \tag{25}$$

- Satunnaismuuttujalla X sanotaan olevan *jakauma*. Jakaumat on usein nimetty henkilöiden (mahdollisesti keksijä) tai PDF funktion mukaan
- Esimerkkinä ensimmäisestä on Gaussin jakauma ja toisesta eksponentiaalinen jakauma, jossa PDF on eksponentiaalinen.
- Kullakin jakaumalla on PDF ja CDF, joista se on tunnistettavissa.

- Usein halutaan tietää todennäköisyys jolla satunnaismuuttuja on jollain välillä eli tapahtuman $\{x_1 < X \leq x_2\}$ todennäköisyys $(x_1 < x_2)$
- Tapahtuma voidaan aina esittää kahden toisensa poissulkevan tapahtuman unionina (usein käytetty temppu)
- Tapahtuma $\{X \leq x_2\}$ on selvästi tapahtumien $\{X \leq x_1\}$ ja $\{x_1 < X \leq x_2\}$ unioni



- Tällöin

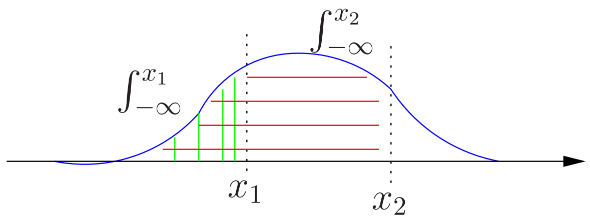
$$\mathrm{P}(X \leq x_2) = \mathrm{P}(X \leq x_1) + \mathrm{P}(x_1 < X \leq x_2)$$

eli

$$F(x_2) = F(x_1) + \mathrm{P}(x_1 < X \leq x_2)$$

- Tästä seuraa

$$\begin{aligned}\mathrm{P}(x_1 < X \leq x_2) &= F(x_2) - F(x_1) \\ &= \int_{-\infty}^{x_2} p(x)\, dx - \int_{-\infty}^{x_1} p(x)\, dx \\ &= \int_{x_1}^{x_2} p(x)\, dx \qquad (26)\end{aligned}$$



- Todennäköisyys että satunaismuuttuja X saa arvoja väliltä $x_1 < X \leq x_2$ on siis PDF:n integraali yli tuon alueen

Useita satunnaismuutujia

- Kun tarkastellaan useita yhtäaikaisia tapahtumia tai peräkkäisiä tapahtumia pitää tietää näiden yhteis kertymä- ja/tai tiheysfunktiot (*yhteisjakauma*)
- Useat satunnaismuuttujat muodostavat periaatteessa moniulotteisen funktion, joka on määritelty yhdistetyssä näyteavaruudessa
- Tarkastellaan ensin kahden muuttujan X_1 ja X_2 yhteisjakaumaa ja yleistetään tulokset sen jälkeen useammille muuttujille

- Yhteiskertymäfunktio (yhteis (joint) CDF) on

$$\begin{aligned} F(x_1, x_2) &= \mathrm{P}(X_1 \leq x_1,\ X_2 \leq x_2) \\ &= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} p(u_1, u_2)\, du_1\, du_2 \end{aligned} \tag{27}$$

jossa $p(x_1, x_2)$ on yhteis todennäköisyys tiheysfunktio (yhteis PDF), joka on

$$p(x_1, x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F(x_1, x_2) \tag{28}$$

- Kun yhteis PDF integroidaan yli toisen satunnaismuuttujan saadaan

$$\int_{-\infty}^{\infty} p(x_1, x_2)\, dx_2 = p(x_1) \tag{29}$$

jota kutsutaan *marginaaliseksi* PDF:ksi (erotukseksi varsinaisesta PDF:stä, jota ei ole saatu integroimalla)

- On helppo havaita että

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x_1, x_2)\, dx_1\, dx_2 = 1$$

$$F(-\infty, -\infty) = F(-\infty, x_2) = F(x_1, -\infty) = 0$$

- Yleistys useampiulotteisiin tapauksiin on suoraviivaista

$$\begin{aligned} F(x_1, \ldots, x_n) &= \mathrm{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) \\ &= \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} p(u_1, \ldots, u_n)\, du_1 \ldots du_n \\ p(x_1, \ldots, x_n) &= \frac{\partial^n}{\partial x_1 \cdots \partial x_n} F(x_1, \ldots, x_n) \end{aligned}$$

- Marginaaliset PDF:t voidaan generoida mille tahansa määrälle satunnaismuuttujia eli

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x_1, x_2, x_3, x_4, x_5, \ldots, x_n)\, dx_2\, dx_4 = p(x_1, x_3, x_5, \ldots, x_n)$$

- Havaitaan myös että

$$\begin{aligned} F(x_1, \infty, x_3, \infty, x_5, \ldots, x_n) &= F(x_1, x_3, x_5, \ldots, x_n) \\ F(x_1, -\infty, x_2, \ldots, x_n) &= 0 \end{aligned}$$

Ehdollinen tiheysfunktio

- Ehdollinen todennäköisyyshän oli $\mathrm{P}(A|B) = P(A, B)/P(B)$
- Ehdollisen todennäköisyyden, että muuttuja X_1 on pienempi kuin x_1 kun satunnaismuuttuja X_2 on saanut arvon x_2 eli todennäköisyyden $\mathrm{P}(X_1 \leq x_1 \,|\, X_2 = x_2)$ laskeminen ei yleisessä tapauksessa onnistu suoraan, koska tyypillisesti $\mathrm{P}(X_2 = x_2) = 0$ ($\neq 0$ vain jos kyseessä distreetti muuttuja jolla jokin arvo ko. pisteessä)
- Lasketaan siis tapahtuman $\{X_1 \leq x_1 \,|\, x_2 - \triangle < X_2 \leq x_2\}$ todennäköisyys, jossa $\triangle$ on pieni positiivinen luku

- Tällöin

$$\mathrm{P}(X_1 \leq x_1 \,|\, x_2 - \triangle < X_2 \leq x_2) = \frac{\int_{-\infty}^{x_1} \int_{x_2-\triangle}^{x_2} p(u_1, u_2)\, du_1\, du_2}{\int_{x_2-\triangle}^{x_2} p(u_2)\, du_2}$$

$$= \frac{F(x_1, x_2) - F(x_1, x_2 - \triangle)}{F(x_2) - F(x_2 - \triangle)} \tag{30}$$

- Olettaen, että vaaditut PDF:t ovat jatkuvia alueella $(x_2 - \triangle, x_2)$ voidaan osoittaja ja nimittäjä yhtälössä (30) jakaa $\triangle$:lla ja ottaa raja-arvo $\triangle \to 0$
- Derivaatan määritelmä: $\lim_{h \to 0} \big(f(x) - f(x+h)\big)/h$

- Tällöin

$$\begin{aligned} \mathrm{P}(X_1 \leq x_1 \,|\, X_2 = x_2) &\equiv F(x_1 \,|\, x_2) \qquad (31) \\ &= \frac{\partial F(x_1, x_2)/\partial x_2}{\partial F(x_2)/\partial x_2} \\ &= \frac{\partial\big[\int_{-\infty}^{x_1}\int_{-\infty}^{x_2} p(u_1, u_2)\, du_1 du_2\big]/\partial x_2}{\partial\big[\int_{-\infty}^{x_2} p(u_2)\, du_2\big]/\partial x_2} \\ &= \frac{\int_{-\infty}^{x_1} p(u_1, x_2)\, du_1}{p(x_2)} \qquad (32) \end{aligned}$$

- Derivoimalla tämä x_1:sen suhteen saadaan

$$p(x_1 \,|\, x_2) = \frac{p(x_1, x_2)}{p(x_2)} \qquad (33)$$

joka on *ehdollinen* PDF.

- Selvästi meillä on myös

$$p(x_1, x_2) = p(x_1 \,|\, x_2)p(x_2) = p(x_2 \,|\, x_1)p(x_1) \tag{34}$$

- Yhtälön (34) yleistys on joskus hyödyllinen jakaumia määrättäessä. Se on

$$p(x_1, x_2, \ldots, x_n) = p(x_1, \ldots, x_k \,|\, x_{k+1}, \ldots, x_n)p(x_{k+1}, \ldots, x_n)$$

Riippumattomat satunnaismuuttujat

- Riippumattomuushan tarkoitti että eri kokeiden ulostulot eivät millään lailla vaikuta toisiinsa.
- Havaittiin, että tällöin $\mathrm{P}(A_1, \ldots, A_n) = \mathrm{P}(A_1) \cdots \mathrm{P}(A_n)$
- Koska satunnaismuuttujat liittyvät kokeisiin ja $P(X \leq x) = F(x)$ niin

  satunnaismuuttujat ovat *tilastollisesti riippumattomia* jos ja vain jos

$$F(x_1, \ldots, x_n) = F(x_1) \cdots F(x_n) = \prod_{i=1}^{n} F(x_i) \tag{35}$$

tai

$$p(x_1, \ldots, x_n) = p(x_1) \cdots p(x_n) = \prod_{i=1}^{n} p(x_i) \tag{36}$$

- Tämä tarkoittaa siis sitä, että riippumattomien satunnaismuuttujien CDF tai PDF on yksittäisten satunnaismuuttujien CDF:ien tai PDF:ien tulo
- Tämä on hyvin usein käytetty ja hyödyllinen tieto

Satunnaismuuttujien funktiot

- Usein tarvitsee tietää mikä on satunnaismuuttujan/muuttujien funktion jakauma
- Tämä tarkoittaa, että halutaan tietää muuttujan $Y = g(X)$ jakauma, kun tiedetään muuttujan X jakauma
- Tämä on varsin suoraviivaista jos kuvaus g on yksi-yhteen (yksikäsitteinen)
- Jos kuvaus ei ole yksikäsitteinen, esim. $Y = X^2$, niin jakauman määräämisessä täytyy olla hyvin huolellinen
- Tätä tietoutta tullaan käyttämään esim. joidenkin tiedonsiirtomenetelmien suorituskyvyn määräämisessä (DTS)

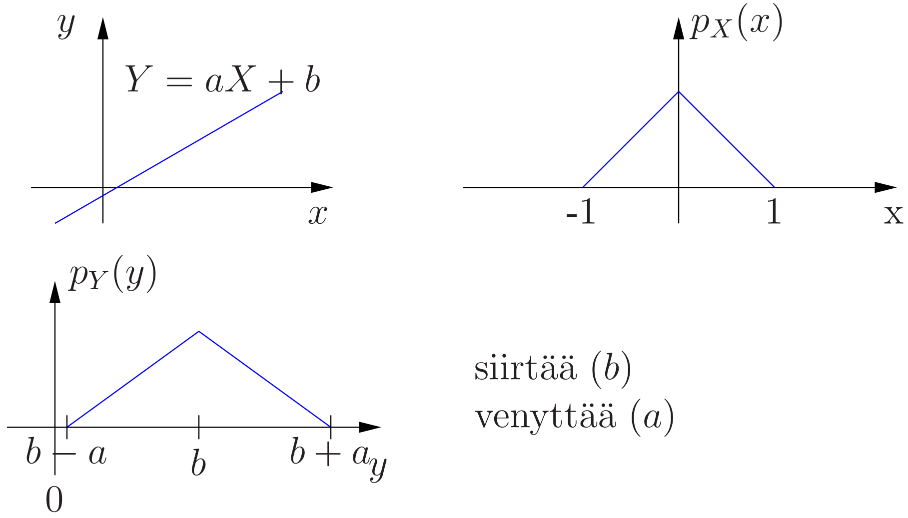
- Esimerkki 1
  - Tarkastellaan muuttujaa $Y = aX + b$, jossa $a > 0$ ja b ovat vakioita
  - Kuvaus on lineaarinen ja monotooninen (= joko kasvava tai vähenevä funktio, nyt kasvava)
  - Kuvaus on siis yksikäsitteinen
  - Olkoon $F_X(x)$ ja $F_Y(y)$ muuttujien X ja Y CDF:t
  - Selvästi

$$F_Y(y) = \mathrm{P}(Y \leq y) = \mathrm{P}(aX + b \leq y) = \mathrm{P}\left(X \leq \frac{y-b}{a}\right)$$
$$= F_X\Big(\frac{y-b}{a}\Big) = \int_{-\infty}^{(y-b)/a} p_X(x)\, dx$$

  - Derivoimalla $F_Y(y)$ saadaan

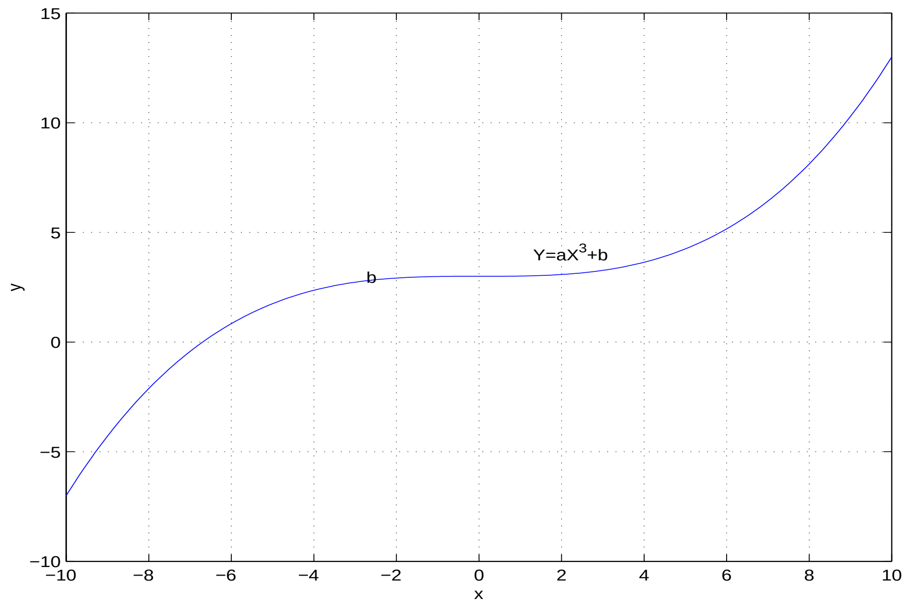
$$p_Y(y) = \frac{1}{a} p_X\Big(\frac{y-b}{a}\Big)$$

– Lineaarisessa muunnoksessa muuttujan $Y = aX + b$ jakauma on helposti esitettävissä muuttujan X jakauman avulla



Esimerkin 1 havainnollistaminen

- Esimerkki 2
  - Tarkastellaan muuttujaa $Y = aX^3 + b$
  - Kuvaus on yksikäsitteinen



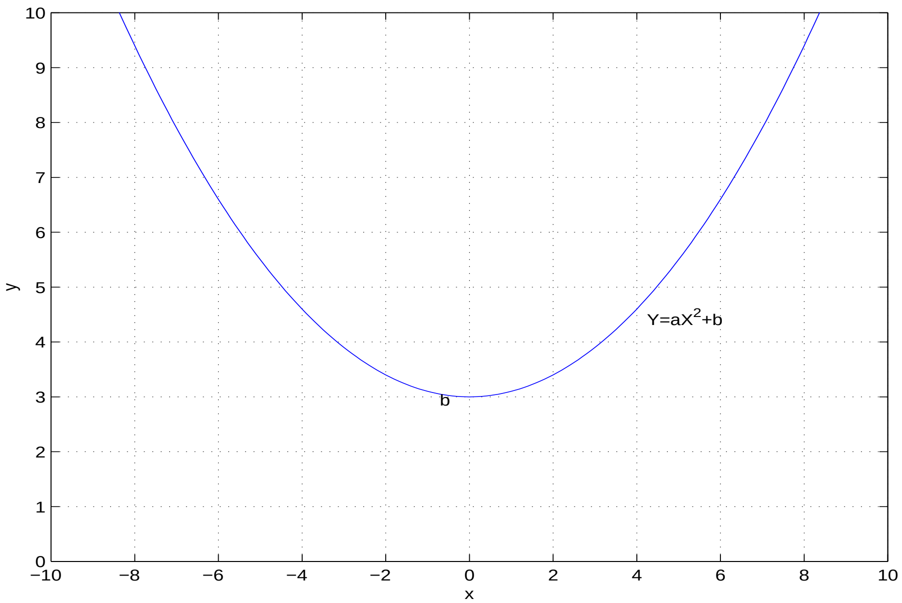
– Nyt on siis

$$\begin{aligned} F_Y(y) &= \mathrm{P}(Y \le y) = \mathrm{P}(aX^3 + b \le y) \\ &= \mathrm{P}\left(X \le \left(\frac{y-b}{a}\right)^{1/3}\right) = F_X\left(\left(\frac{y-b}{a}\right)^{1/3}\right) \end{aligned}$$

josta derivoimalla

$$p_Y(y) = \frac{1}{3a[(y-b)/a]^{2/3}}\, p_X\left(\left(\frac{y-b}{a}\right)^{1/3}\right)$$

- Esimerkki 3
  - Nyt tarkastellaan muuttujaa $Y = aX^2 + b$
  - Kuvaus ei ole yksikäsitteinen sillä yhtä Y:n arvoa vastaa kaksi X:n arvoa



– Nyt

$$
\begin{aligned}
F_Y(y) &= \mathrm{P}(aX^2 + b \le y) \\
= \mathrm{P}\left(|X| \le \sqrt{\frac{y-b}{a}}\right) &= \mathrm{P}\left(-\sqrt{\frac{y-b}{a}} < X \le \sqrt{\frac{y-b}{a}}\right) \\
&= F_X\left(\sqrt{\frac{y-b}{a}}\right) - F_X\left(-\sqrt{\frac{y-b}{a}}\right)
\end{aligned}
$$

– Derivoimalla saadaan

$$
p_Y(y) = \frac{p_X\left(\sqrt{(y-b)/a}\right)}{2a\sqrt{(y-b)/a}} + \frac{p_X\left(-\sqrt{(y-b)/a}\right)}{2a\sqrt{(y-b)/a}}
$$

- Esimerkissä 3 havaittiin, että yhtälöllä $g(x) = ax^2 + b$ on kaksi ratkaisua

$$x_1 = \sqrt{(y-b)/a} \text{ ja } x_2 = -\sqrt{(y-b)/a}$$

ja että muuttujan $Y = aX^2 + b$ PDF on muotoa

$$p_Y(y) = \frac{p_X(x_1)}{|g'(x_1)|} + \frac{p_X(x_2)}{|g'(x_2)|}$$

jossa x_1 ja x_2 ovat yhtälön $g(x) = ax^2 + b$ juuret ja $g'(x)$ on $g(x)$:än derivaatta x:n suhteen

- Yleisesti muuttujalle $Y = g(X)$ päteekin

$$p_Y(y) = \frac{p_X(x_1)}{|g'(x_1)|} + \cdots + \frac{p_X(x_n)}{|g'(x_n)|} = \sum_{i=1}^{n} \frac{p_X(x_i)}{|g'(x_i)|} \qquad (37)$$

jossa $x_i,\ i = 1, \ldots, n$ ovat yhtälön $y = g(x)$ reaaliset juuret

- Sovelletaan edellistä tulosta (37) esimerkkiin 1
  - Nyt $Y = aX + b$, mutta ei rajoituta tilanteeseen $a > 0$ kuten esimerkkissä 1
  - Yhtälön $y = g(x)$ juuri on $y = (y - b)/a$ ja funktion $g(x) = ax + b$ derivaatta on a.
  - Täten

$$p_Y(y) = \frac{p_X\big((y-b)/a\big)}{|a|}$$

Usean muuttujan tilanne

- Olkoon $\mathbf{X} = [x_1 \; \ldots \; x_n]$ n dimensioinen satunnaisvektori
- Sen CDF on $F(\mathbf{X}) = F(x_1, \ldots, x_n)$ ja PDF vastaavasti $p(\mathbf{X}) = p(x_1, \ldots, x_n)$. Nämä ovat siis muuttujien $X_1, \ldots, X_n$ yhteis CDF ja PDF.

  Vektorimerkintä on siis oiva tapa yksinkertaistaa merkintöjä
- Meillä on lisäksi yksiarvoiset (yksi ulostulo) funktiot $g_i(\mathbf{X})$ ja satunnaismuuttujat $Y_i = g_i(\mathbf{X})$, jotka riippuvat muuttujista X_i, $i = 1, \ldots, n$
- Myös Y_i:t ja $g_i(\mathbf{X})$:t voidaan vektorisoida eli $\mathbf{Y} = [Y_1 \; \ldots \; Y_n]$ ja $\mathbf{g}(\mathbf{X}) = [g_1(\mathbf{X}) \; \ldots \; g_n(\mathbf{X})]$
- Lisäksi oletetaan että käänteiskuvaukset $x_i = g^{-1}(\mathbf{Y})$ ovat olemassa ja käänteisfunktiot ovat yksiarvoisia. Lisäksi vaaditaan, että funktioiden osittaisderivaatat ovat jatkuvia

(seuraava tulos ei siis sovi joka tilanteeseen)

- Silloin $p(\mathbf{Y})$ on

$$p_Y(\mathbf{Y}) = p_X\big(\mathbf{g}^{-1}(\mathbf{Y})\big)\left|\frac{\partial \mathbf{g}^{-1}(\mathbf{Y})}{\partial \mathbf{Y}}\right| \tag{38}$$

$$= \frac{p_X\big(\mathbf{g}^{-1}(\mathbf{Y})\big)}{\left|\frac{\partial \mathbf{g}(\mathbf{X})}{\partial \mathbf{X}}\right|} \tag{39}$$

- Determinantit $\left|\frac{\partial \mathbf{g}^{-1}(\mathbf{Y})}{\partial \mathbf{Y}}\right|$ ja $\left|\frac{\partial \mathbf{g}(\mathbf{X})}{\partial \mathbf{X}}\right|$ ovat kuvausten $\mathbf{g}^{-1}(\mathbf{Y})$ ja $\mathbf{g}(\mathbf{X})$ Jacobiaanit

- Jacobiaanit ovat

$$\left|\frac{\partial \mathbf{g}^{-1}(\mathbf{Y})}{\partial \mathbf{Y}}\right| = \begin{vmatrix} \frac{\partial g_1^{-1}}{\partial y_1} & \cdots & \frac{\partial g_1^{-1}}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n^{-1}}{\partial y_1} & \cdots & \frac{\partial g_n^{-1}}{\partial y_n} \end{vmatrix} \text{ ja } \left|\frac{\partial \mathbf{g}(\mathbf{X})}{\partial \mathbf{X}}\right| = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_n} \end{vmatrix} \tag{40}$$

- Jacobiaanit ovat siis muunnosten osittaisderivaatoista muodostettujen matriisien determinatteja
- Tulos (38) on esitetty oppikirjassa
- Tulosta (39) ei ole esitetty oppikirjassa mutta se voi olla hyödyllinen, sitähän käytettiin jo yhden muuttujan funktion jakauman löytämisessä eli yksikäsitteisessä tapauksessa tulos (37) on suora seuraus tuloksesta (39)

- Tulos (37) on hieman yleisempi (yhdelle muuttujalle) kuin (39), sillä siinä voi yhtälöllä $y = g(x)$ voi olla useita ratkaisuja, mutta ratkaisussa (39) on oletettu yksi ratkaisu yhtälöryhmälle $\mathbf{Y} = \mathbf{g}(\mathbf{X})$. Tulos (39) on toki yleistettävissä
- Otetaan seuraava lineaarisen muunnoksen yleistys esimerkiksi kummankin tavan käytöstä. Kyseistä muunnosta käytetään hyvin usein, joten esimerkki on hyödyllinen
  - Olkoon $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$, jossa $\mathbf{A}$ on ei-singulaarinen $n \times n$ muunnosmatriisi, $\mathbf{Y}$ ja $\mathbf{X}$ ovat $n \times 1$ satunnaisvektoreita ja $\mathbf{b}$ on $n \times 1$ vakiovektori
  - Tällöin matriisin $\mathbf{A}$ käänteismatriisi $\mathbf{A}^{-1}$ on olemassa
  - Käytetään ensin tapaa (38)
  - Nyt $\mathbf{X} = \mathbf{A}^{-1}(\mathbf{Y} - \mathbf{b}) = \mathbf{g}^{-1}(\mathbf{Y})$
  - Käänteiskuvauksen $\mathbf{g}^{-1}(\mathbf{Y})$ Jacobiaani on $|\mathbf{A}^{-1}|$

- Näin on koska $\partial \mathbf{A}^{-1}(\mathbf{Y}-\mathbf{b})/\partial \mathbf{Y} = \partial\{\mathbf{A}^{-1}\mathbf{Y}-\mathbf{A}^{-1}\mathbf{b}\}/\partial \mathbf{Y} = \mathbf{A}^{-1}$
- Miksi näin? Tarkastellaan matriisi vektoritulon $\mathbf{Ay}$ i:nettä elementtiä $x_i = \sum_j a_{ij}y_j$, jossa a_{ij} ja y_j ovat $m \times n$ matriisin $\mathbf{A}$ ja $n \times 1$ vektorin $\mathbf{y}$ elementit. Silloin

$$\frac{\partial x_i}{\partial y_k} = \frac{\partial\left\{\sum_j a_{ij}y_j\right\}}{\partial y_k}$$
$$= \frac{\partial\left\{a_{i1}y_1 + \ldots + a_{ik}y_k + \ldots + a_{in}y_n\right\}}{\partial y_k} = a_{ik}$$

  eli matriisin $\mathbf{A}$ (i,k) :s elementti
- Osittaisderivaattoja $\frac{\partial x_i}{\partial y_k}$ on $m \times n$ kpl jotka järjestettynä matriisiin antavat halutun tuloksen eli

$$\frac{\partial \mathbf{x}}{\partial \mathbf{y}} = \frac{\partial\{\mathbf{Ay}\}}{\partial \mathbf{y}} = \mathbf{A}$$

– Tuloksen (38) mukaan saadaan siis

$$p(\mathbf{Y}) = p_X\big(\mathbf{A}^{-1}(\mathbf{Y}-\mathbf{b})\big)|\mathbf{A}^{-1}| = \frac{p_X\big(\mathbf{A}^{-1}(\mathbf{Y}-\mathbf{b})\big)}{|\mathbf{A}|}$$

sillä determinanteille pätee $|\mathbf{A}^{-1}| = 1/|\mathbf{A}|$

– Käytetään sitten tapaa (39)

– Tarkasteltava muunnos on $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b} = \mathbf{g}(\mathbf{X})$

– Muunnoksen $\mathbf{g}(\mathbf{X})$ Jacobiaani on $|\mathbf{A}|$

– Tuloksen (39) mukaan

$$p(\mathbf{Y}) = \frac{p_X\big(\mathbf{A}^{-1}(\mathbf{Y}-\mathbf{b})\big)}{|\mathbf{A}|}$$

joka on sama kuin tuloksen (38) avulla laskettu tulos

Tilastolliset keskiarvot

- Satunnaismuutujien tilastolliset keskiarvot ovat usein tarvittu ominaisuus ja ne ilmenevät useissa eri yhteyksissä, esim. demodulaattorien ja kanavakorjainten suunnittelussa
- Tärkeimpinä ovat ehkäpä ensimmäinen ja toinen momentti ja 2. keskeismomentti.
- Satunnaismuuttujan *n:s momentti* on

$$\mathrm{E}\{X^n\} = \int_{-\infty}^{\infty} x^n p(x)\, dx \tag{41}$$

- 1. momenttia kutsutaan satunnaismuuttujan *tilastolliseksi keskiarvoksi* tai *odotusarvoksi* (mean or expected value) ja se on

$$\mathrm{E}\{X\} \equiv m_x = \int_{-\infty}^{\infty} x p(x)\, dx \tag{42}$$

- Olkoon satunnaismuuttuja $Y = g(X)$. Sen odotusarvo on

$$\mathrm{E}\{Y\} = \int_{-\infty}^{\infty} g(x)p(x)\,dx \tag{43}$$

- Miksi tämä kuvaa (yleisesti) keskiarvoa?
  - Diskreetissä tapauksessa $\mathrm{E}\{Y\} = \sum_i g(x_i)P(x = x_i)\delta(x - x_i)$ selvästi kuvaa muuttujien $g(x_i)$ keskimääräistä arvoa
  - Esim. nopanheitossa, jos $g(X) =$ nopan silmäluku, saadaan $1 \times 1/6 + 2 \times 1/6 + \cdots + 6 \times 1/6 = 3,5$

    eli 1 esiintyy frekvenssillä 1/6, 2 frekvenssillä 1/6 jne. ja keskimääräinen arvo on 3,5
  - jatkuville muuttujille keskimääräinen arvo vastaa ylläolevaa integraalia

- Olkoon $Y = (X - m_x)^n$ eli X josta on vähennetty keskiarvo (potenssiin n)
- Tämän keskiarvo on

$$\mathrm{E}\{(X - m_x)^n\} = \int_{-\infty}^{\infty} (x - m_x)^n p(x)\, dx \tag{44}$$

jota kutsutaan *n:ksi keskeismomentiksi*

- Se on momentti keskiarvon suhteen
- 2. keskeismomenttia kutsutaan *varianssiksi* ja se on

$$\mathrm{E}\{(X - m_x)^2\} \equiv \sigma_x^2 = \int_{-\infty}^{\infty} (x - m_x)^2 p(x)\, dx \tag{45}$$

ja se kuvaa muuttujan arvon heilumista keskiarvon ympärillä eli hajontaa

- Purkamalla neliöinti saadaan

$$\sigma_x^2 = \int (x^2 - 2xm_x + m_x^2)p(x)\,dx$$
$$= \underbrace{\int x^2 p(x)\,dx}_{\mathrm{E}\{X^2\}} - 2m_x \underbrace{\int xp(x)\,dx}_{m_x} + m_x^2 \underbrace{\int p(x)\,dx}_{1}$$
$$= \mathrm{E}\{X^2\} - m_x^2 \tag{46}$$

joka muodostaa yhteyden varianssin sekä ensimmäisen ja toisen momentin välille

- Nollakeskiarvoiselle satunnaismuuttujalle varianssi ja toinen momentti ovat samat

- Kahden muuttujan tapauksessa saadaan *yhteismomentti*

$$\mathrm{E}\{X_1^k X_2^n\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1^k x_2^n p(x_1, x_2)\, dx_1\, dx_2 \qquad (47)$$

sekä *yhteis keskeismomentti*

$$\mathrm{E}\{(X_1 - m_1)^k (X_2 - m_2)^n\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1 - m_1)^k (x_2 - m_2)^n p(x_1, x_2)\, dx_1\, dx_2 \qquad (48)$$

jossa $m_i = \mathrm{E}\{X_i\}$.

- Kiinnostavin tilanne on kun $n = k = 1$. Silloin muuttujien X_i ja X_j yhteismomenttia kutsutaan *korrelaatioksi* ja yhteiskeskeismomenttia *kovarianssiksi*

- Korrelaatio on siis

$$\mathrm{E}\{X_i X_j\} \tag{49}$$

ja kovarianssi on

$$\mu_{ij} = \mathrm{E}\{(X_i - m_i)(X_j - m_j)\} \tag{50}$$

eli ne ovat muuttujien tulojen odotusarvot

- Kovarianssin sekä korrelaation ja keskiarvojen välillä on yhteys

$$\begin{aligned}\mu_{ij} &= \mathrm{E}\{X_i X_j - X_i m_j - m_i X_j + m_i m_j\} \\ &= \mathrm{E}\{X_i X_j\} - \underbrace{\mathrm{E}\{X_i\}}_{m_i} m_j - m_i \underbrace{\mathrm{E}\{X_j\}}_{m_j} + m_i m_j \\ &= \mathrm{E}\{X_i X_j\} - m_i m_j \end{aligned} \tag{51}$$

- Nollakeskiarvoisille muuttujille korrelaatio ja kovarianssi ovat samat

- Useamman muuttujan tapauksessa voidaan myös määritellä erilaisia yhteismomentteja mutta useimmiten ollaan kiinnostuttu vain korrrelaatiosta ja kovarianssista eri muuttujien välillä.
- Olkoon meillä satunnaisvektori $\mathbf{X} = [X_1 \ X_2 \ \ldots \ X_n]^{\mathrm{T}}$ jossa yläindeksi T tarkoittaa transpoosia. $\mathbf{X}$ on siis $n \times 1$ pystyvektori
- *Korrelaattiomatriisi* sisältää vektorin $\mathbf{X}$ alkioiden väliset korrelaatiot ja *kovarianssimatriisi* niiden väliset kovarianssit
- Korrelaatiomatriisi on siis

$$\mathrm{E}\{\mathbf{X}\mathbf{X}^{\mathrm{T}}\} = \begin{bmatrix} \mathrm{E}\{X_1X_1\} & \mathrm{E}\{X_1X_2\} & \ldots & \mathrm{E}\{X_1X_n\} \\ \mathrm{E}\{X_2X_1\} & \mathrm{E}\{X_2X_2\} & \ldots & \mathrm{E}\{X_2X_n\} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{E}\{X_nX_1\} & \mathrm{E}\{X_nX_2\} & \ldots & \mathrm{E}\{X_nX_n\} \end{bmatrix} \qquad (52)$$

- Diagonaali sisältää kunkin alkion 2. keskeismomentin ja diagonaalin ulkopuoliset elementit alkioiden välisiä yhteismomentteja

- Olkoon $\mathbf{m} = \mathrm{E}\{\mathbf{X}\} = [m_1 \ \ldots \ m_n]^{\mathrm{T}}$ muuttujien X_i keskiarvovektori
- Kovarianssimatriisi on

$$\mathrm{E}\{(\mathbf{X}-\mathbf{m})(\mathbf{X}-\mathbf{m})^{\mathrm{T}}\} = \begin{bmatrix} \mu_{11} & \mu_{12} & \ldots & \mu_{1n} \\ \mu_{21} & \mu_{22} & \ldots & \mu_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n1} & \mu_{n2} & \ldots & \mu_{nn} \end{bmatrix} \tag{53}$$

- Diagonaali sisältää alkioiden varianssit sillä $\mu_{ii} = \sigma_i^2$
- Vastaavasti voidaan määrittää satunnaisvektorien $\mathbf{X}$ ja $\mathbf{Y}$ väliset *ristikorrelaatio-* ja *-kovarianssimatriisi* eli $\mathrm{E}\{\mathbf{X}\mathbf{Y}^{\mathrm{T}}\}$ ja $\mathrm{E}\{(\mathbf{X}-\mathbf{m}_x)(\mathbf{Y}-\mathbf{m}_y)^{\mathrm{T}}\}$

Korreloimattomuus

- Kahta satunnaismuuttujaa X_i ja X_j sanotaan *korreloimattomiksi* jos niiden välinen kovarianssi on nolla eli $\mu_{ij} = 0$
- Tästä seuraa se, että korrelaatio on keskiarvojen tulo, sillä yhtälöstä

$$\mu_{ij} = \mathrm{E}\{X_i X_j\} - m_i m_j = m_i m_j = 0$$

saadaan

$$\mathrm{E}\{X_i X_j\} = m_i m_j = \mathrm{E}\{X_i\}\,\mathrm{E}\{X_j\}$$

- Tästä seuraa että tilastollisesti riippumattomat satunnaismuuttujat ovat korreloimattomia
- Tästä *ei seuraa* että korreloimattomat satunnaismuuttujat ovat välttämättä riippumattomia (näin toki voi olla)
- Satunnaismuuttujien sanotaan olevan *ortogonaalisia* jos

$$\mathrm{E}\{X_i X_j\} = 0$$

- Näin on esim. jos muuttujat ovat korreloimattomia ja jompi kumpi tai molemmat keskiarvoista ovat nollia

Karakteristinen funktio

- Määritelmän mukaan satunnaismuuttujan *karakteristinen funktio* on

$$\mathrm{E}\{e^{jvX}\} \equiv \psi(jv) = \int_{-\infty}^{\infty} e^{jvx} p(x)\, dx, \;\; j = \sqrt{-1} \tag{54}$$

- Tämä vastaa Fourier muunnosta (lukuunottamatta eksponentin miinus merkkiä, mutta sillä ei ole väliä)
- Fourier käänteismuunnoksesta seuraa että muuttujan PDF karakteristisen funktion avulla ilmaistuna on

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi(jv) e^{-jvx}\, dv \tag{55}$$

- Karakteristisen funktion avulla voidaan siis määrittää muuttujan PDF

- Toinen hyödyllinen seikka on yhteys momentteihin
- Derivoimalla karakteristien funktio muuttujan v suhteen saadaan

$$\frac{d\psi(jv)}{dv} = j\int_{-\infty}^{\infty} xe^{jvx}p(x)\,dx$$

- Ratkaisemalla tämä pisteessä $v = 0$ saadaan

$$\left.\frac{d\psi(jv)}{dv}\right|_{v=0} = j\int_{-\infty}^{\infty} xp(x)\,dx$$

eli 1. momentti (keskiarvo) on ilmaistavissa myös muodossa

$$\mathrm{E}\{X\} = -j\left.\frac{d\psi(jv)}{dv}\right|_{v=0} \tag{56}$$

- Korkeamman asteen derivaatoista saadaan korkeammille momenteille yhtälö

$$\mathrm{E}\{X^n\} = (-j)^n \frac{d^n \psi(jv)}{dv^n}\bigg|_{v=0} \tag{57}$$

- Karakteristisen funktion avulla voidaan siis määrätä muuttujan momentit

- Karakteristinen funktio voidaan esittää sarjamuodossa
- Karakteristisen funktion Taylorin sarja pisteen $v = 0$ ympäristössä on

$$\psi(jv) = \sum_{k=0}^{\infty} \frac{d^n \psi(jv)}{dv^n}\bigg|_{v=0} \frac{v^n}{n!}$$

- Sijoittamalla tähän derivaatan paikalle yhtälöstä (56) saatu derivaatta saadaan

$$\psi(jv) = \sum_{k=0}^{\infty} \mathrm{E}\{X^n\} \frac{(jv)^n}{n!} \tag{58}$$

- Karakteristinen funktio voidaan siis esittää myös momenttien avulla

- Karakteristisen funktion avulla on helppo löytää riippumattomien satunnaismuuttujien summan PDF
- Tarkastellaan usean muuttujan summaa eli muuttujaa

$$Y = \sum_{i=1}^{n} X_i$$

- Sen karakteristinen funktio on

$$\begin{aligned}
\psi_Y(jv) &= \mathrm{E}\{e^{jvY}\} \\
&= \mathrm{E}\Big\{\exp\big(jv\sum_{i=1}^{n} X_i\big)\Big\} \\
&= \mathrm{E}\Big\{\prod_{i=1}^{n}\big(e^{jvX_i}\big)\Big\} \quad (\text{sillä}\, e^{a+b} = e^a e^b) \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\Big(\prod_{i=1}^{n} e^{jvx_i}\Big)p(x_1,\ldots,x_n)\,dx_1\ldots,dx_n
\end{aligned}$$

- Koska muuttujat X_i ovat riippumattomia, niin $p(x_1, \dots, x_n) = p(x_1) \cdots p(x_n)$ jolloin

$$\begin{aligned} \psi_Y(jv) &= \prod_{i=1}^{n} \Big[\int_{-\infty}^{\infty} e^{jvx_i} p(x_i)\, dx_i \Big] \\ &= \prod_{i=1}^{n} \psi_{X_i}(jv) \end{aligned} \tag{59}$$

- Riippumattomien muuttujien summan karakteristinen funktio on siis yksittäisten muuttujien karakterististen funktioiden tulo
- Summan PDF on tämän käänteinen Fourier muunnos
- Jos muuttujat ovat *identtisesti jakautuneita* eli $p(x_i) = p(x_j)\, \forall i, j$ niin

$$\psi_Y(jv) = \Big[\psi_X(jv)\Big]^n \tag{60}$$

- Fourier muunnoksen ominaisuuksista seuraa, että PDF:iä käyttäen riippumattomien muuttujien summan PDF on yksittäisten PDF:ien konvoluutio
- Tulo on kuitenkin usein helpompi laskea kuin konvoluutio, joten karakteristisen funktion käyttö ko. summan PDF:n määräämiseen on useimmiten järkevää

- Usean muuttujan yhteisjakauman tapauksessa tarvitaan moniulotteista Fourier muunnosta
- n:lle muuttujalle karakteristinen funktio on

$$\begin{aligned}\psi(jv_1, \ldots, jv_n) &= \mathrm{E}\Big\{\exp\big(j\sum_{i=1}^{n} v_i X_i\big)\Big\} \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\big(j\sum_{i=1}^{n} v_i x_i\big) p(x_1, \ldots, x_n)\, dx_1 \ldots dx_n\end{aligned}$$

- Kahden muuttujan tapauksessa meillä on

$$\psi(jv_1, jv_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{j(v_1x_1+v_2x_2)} p(x_1, x_2)\, dx_1\, dx_2$$

- Momentit saadaan generoitua vastaavasti kuten yhden muuttujan tapauksessa ottamalla osittaisderivaatat muuttujien suhteen ja ratkaisemalla se pisteissä $v_i = 0 \; \forall i$.
- Esimerkiksi kahden muuttujan tapauksessa saadaan

$$\underbrace{jj}_{-1} \mathrm{E}\{X_1 X_2\} = \frac{\partial^2 \psi(jv_1, jv_2)}{\partial v_1 \partial v_2}\bigg|_{v_1 = v_2 = 0}$$

- Korkeammat momentit voidaan käsitellä vastaavasti