

---

**Ilkka Mellin**  
**Todennäköisyyslaskenta**

**Osa 2: Satunnaismuuttujat ja  
todennäköisyysjakaumat**  
**Jakaumien tunnusluvut**

# Jakaumien tunnusluvut

---

>> **Odotusarvo**

**Varianssi**

**Markovin ja Tshebyshevin epäyhtälöt**

**Momentit**

**Vinous ja huipukkuus**

**Kvantiilit**

**Moodi**

**Suurten lukujen laki**

Odotusarvo

## Johdatteleva esimerkki: Arpajaiset 1/7

---

- Olkoon **arpajaisissa** 1000 *arpaa*.
- *Arpanumerot*: 1, 2, ... , 1000.
- *Voitonjako*:

Voitot (mk)	Voittoja (kpl)
1000	1
100	10
20	100

Odotusarvo

## Johdatteleva esimerkki:

### Arpajaiset 2/7

---

- *Arvotaan* voitonumerot seuraavalla tavalla:
  - (1) Kirjoitetaan arpanumerot lipukkeille.
  - (2) Pannaan lipukkeet urnaan.
  - (3) Poimitaan urnasta *satunnaisesti* 111 arpaa:
    - 100 ensimmäistä saa voittona 20 mk
    - 10 seuraavaa saa voittona 100 mk
    - Viimeinen saa voittona 1000 mk
- *Voitot yhteensä* (mk):
$$1000 \times 1 + 100 \times 10 + 20 \times 100 = 4000$$
- *Voitto yhtä ostettua arpaa kohden eli voitto/arpa* (mk):
$$4000 / 1000 = 4$$

## Odotusarvo

# Johdatteleva esimerkki: Arpajaiset 3/7

---

- *Voitto/arpa* voidaan laskea myös *toisella tavalla*.
- *Arpanumerot*: 1, 2, ..., 1000.
- *Voitonjako*:

Voitot (mk)	Voittoja (kpl)
1000	1
100	10
20	100
0	889

- *Voitto/arpa* (mk):

$$\begin{aligned} & \frac{1 \times 1000 + 10 \times 100 + 100 \times 20 + 889 \times 0}{1000} \\ &= \frac{1}{1000} \times 1000 + \frac{10}{1000} \times 100 + \frac{100}{1000} \times 20 + \frac{889}{1000} \times 0 \\ &= 4 \end{aligned}$$

Odotusarvo

## Johdatteleva esimerkki:

### Arpajaiset 4/7

---

- *Voitto/arpa* saadaan siis laskutoimituksella

$$\frac{1}{1000} \times 1000 + \frac{10}{1000} \times 100 + \frac{100}{1000} \times 20 + \frac{889}{1000} \times 0 = 4$$

jossa *voitto/arpa* on laskettu voittojen *painotettuna summana*, jossa *painoina* on käytetty voittojen *todennäköisyyksiä*:

$$\Pr(\text{Voitto} = 1000) = \frac{1}{1000} = 0.001$$

$$\Pr(\text{Voitto} = 100) = \frac{10}{1000} = 0.01$$

$$\Pr(\text{Voitto} = 20) = \frac{100}{1000} = 0.1$$

$$\Pr(\text{Voitto} = 0) = \frac{889}{1000} = 0.889$$

## Odotusarvo

### Johdatteleva esimerkki:

### Arpajaiset 5/7

---

- Siten suure *voitto/arpa* on laskettu kaavalla

$$\sum_i x_i p_i$$

jossa

$$x_i = \text{voitto}$$

$$p_i = \text{on voiton } x_i \text{ todennäköisyys}$$

- Suuretta *voitto/arpa* kutsutaan todennäköisyyslaskennassa *voiton odotusarvoksi*.
- Voiton odotusarvo on siis *odotettavissa oleva voitto*, jos ostaa *yhden arvan*.
- Voiton odotusarvolle voidaan antaa seuraava *frekvenssitulkinta*:  
Jos ostetaan useita arpoja, voiton odotusarvo kertoo *keskimääräisen voiton yhtä arpa kohden*.

## Odotusarvo

### Johdatteleva esimerkki:

### Arpajaiset 6/7

---

- Arpominen on *satunnaisilmiö*.
- Määritellään *satunnaismuuttuja*  $X = \text{voitto}$ .
- Satunnaismuuttujan  $X$  mahdolliset *arvot*  $x_i$  (voitot) ja niiden *todennäköisyydet*  $p_i$  :

$x_i$	$\Pr(X = x_i) = p_i$
1000	1/1000
100	10/1000
20	100/1000
0	889/1000

- **Huomautus:**

Huomaa, että tulosvaihtoehto 0 mk ja sen todennäköisyys on otettava mukaan!



## Odotusarvo

### Johdatteleva esimerkki:

### Arpajaiset 7/7

---

- Satunnaismuuttujan  $X$  arvot  $x_i$  ja niiden *todennäköisyydet*

$$\Pr(X = x_i) = p_i$$

määrittelevät *diskreetin todennäköisyysjakauman*.

- Lauseke

$$\sum_i x_i p_i = \sum_i x_i \Pr(X = x_i)$$

määrittelee diskreetin satunnaismuuttujan  $X$  *odotusarvon*.

- **Huomautus:**

Odotusarvo määritellään seuraavassa erikseen *diskreeteille* ja *jatkuville* jakaumille.

## Diskreetin jakauman odotusarvo: Määritelmä

---

- Olkoon  $X$  *diskreetti satunnaismuuttuja*.
- Olkoon  $\{x_1, x_2, x_3, \dots\}$  satunnaismuuttujan  $X$  *tulosvaihtoehtojen* eli arvojen joukko.
- Olkoon satunnaismuuttujan  $X$  *pistetodennäköisyysfunktio*

$$f(x_i) = \Pr(X = x_i) = p_i, \quad i = 1, 2, 3, \dots$$

- Satunnaismuuttujan  $X$  **odotusarvo** on *vakio*

$$E(X) = \mu_X = \sum_i x_i \Pr(X = x_i) = \sum_i x_i f(x_i)$$

- Sanomme, että satunnaismuuttujan  $X$  odotusarvo  $E(X)$  on sen *jakauman odotusarvo*, joka kuvaa satunnaismuuttujaan  $X$  liittyviä todennäköisyyksiä.

## Diskreetin jakauman odotusarvo: Kommentteja

---

- Vaikka satunnaismuuttujan sama arvo vaihtelee satunnaisesti koetoistosta toiseen, satunnaismuuttuja saa *keskimäärin* arvoja, jotka vaihtelevat sen odotusarvon ympärillä.
- *Jos jakaumalla on odotusarvo*, se on jakauman todennäköisyysmassan **painopiste**.
- *Diskreetin jakauman odotusarvon ei tarvitse kuulua ko.* satunnaismuuttujan tulosvaihtoehtojen joukkoon.  
Nopanheiton tuloksen odotusarvo on 3.5 (ks. >), mikä ei esiinny mahdollisten tulosvaihtoehtojen joukossa.

## Odotusarvo

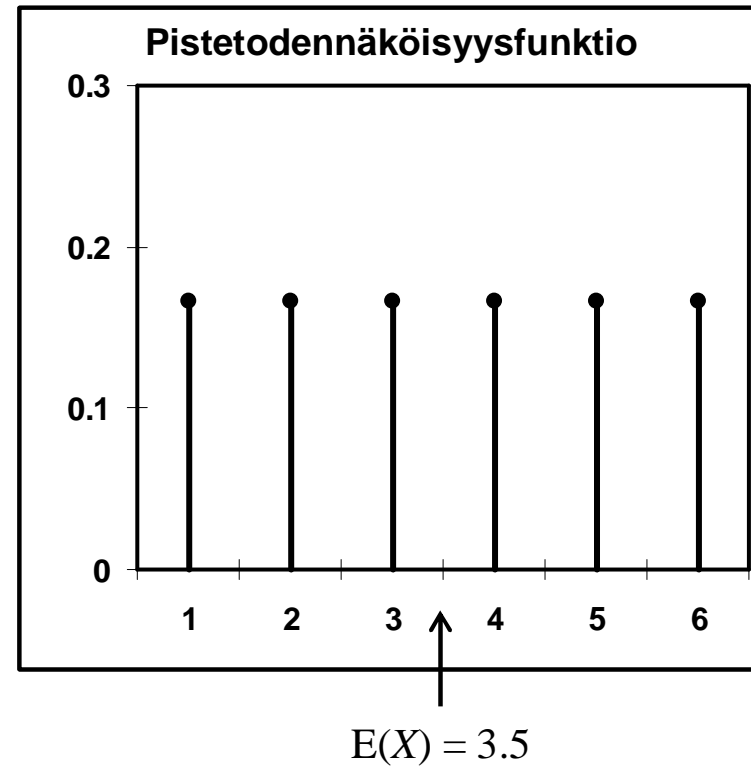
# Diskreetin jakauman odotusarvo: Esimerkki nopanheitosta

- Nopanheittoon liittyvän **diskreetin tasaisen jakauman pistetodennäköisyysfunktio** on muotoa

$$\Pr(X = i) = \frac{1}{6}, i = 1, 2, 3, 4, 5, 6$$

- Satunnaismuuttujan  $X$  *odotusarvo*:

$$\begin{aligned} E(X) &= \sum_{i=1}^6 i \Pr(X = i) = \sum_{i=1}^6 i \frac{1}{6} \\ &= \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = \frac{21}{6} \\ &= 3.5 \end{aligned}$$



Odotusarvo

## Diskreetin jakauman odotusarvo: Esimerkki onnenpyörästä 1/2

- Olkoon *diskreetin satunnaismuuttujan X pistetodennäköisyysfunktio* muotoa

$$\Pr(X = 1) = 0.3$$

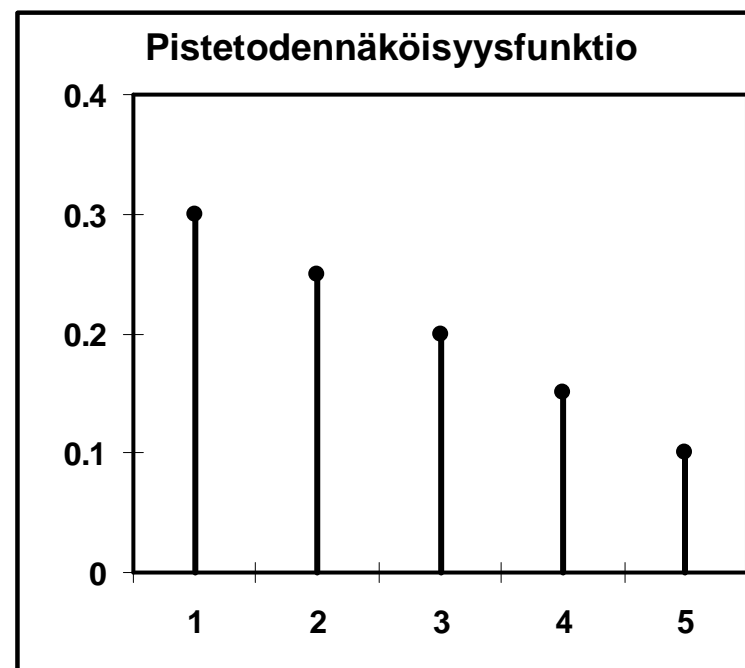
$$\Pr(X = 2) = 0.25$$

$$\Pr(X = 3) = 0.2$$

$$\Pr(X = 4) = 0.15$$

$$\Pr(X = 5) = 0.1$$

- Pistetodennäköisyysfunktio liittyy luvussa **Satunnaismuuttajat ja todennäköisyysjakaumat** käsiteltyyn esimerkkiin onnenpyörästä.



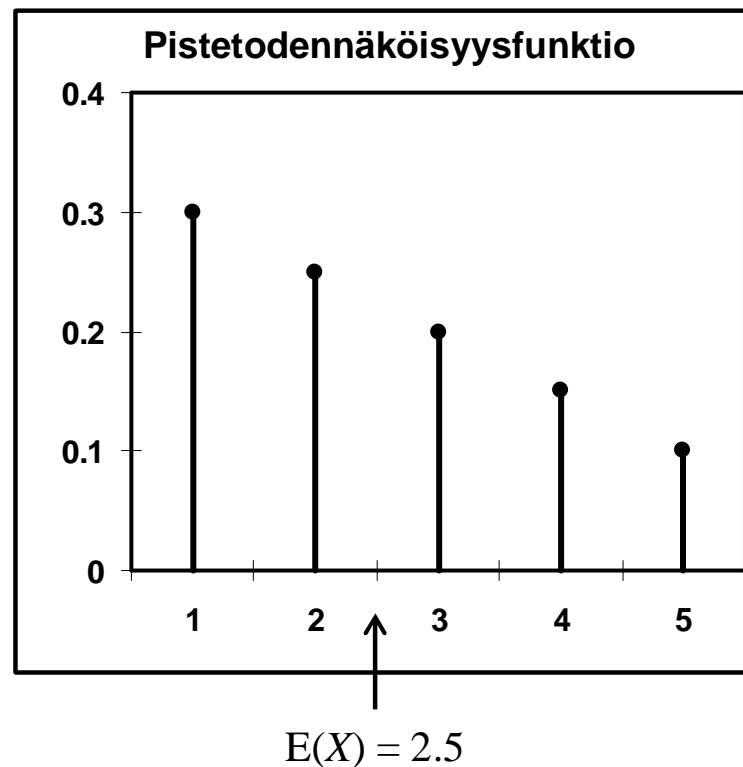
## Odotusarvo

# Diskreetin jakauman odotusarvo:

## Esimerkki onnenpyörästä 2/2

- Satunnaismuuttujan  $X$   
*odotusarvo:*

$$\begin{aligned} E(X) &= \sum_{i=1}^5 i \Pr(X = i) \\ &= 1 \times 0.3 + 2 \times 0.25 + 3 \times 0.2 \\ &\quad + 4 \times 0.15 + 5 \times 0.1 \\ &= 2.5 \end{aligned}$$



## Odotusarvo

# Jatkuvan jakauman odotusarvo: Määritelmä

---

- Olkoon  $X$  on *jatkuva satunnaismuuttuja*.
- Olkoon satunnaismuuttujan  $X$  *tiheysfunktio*  $f(x)$ .
- Satunnaismuuttujan  $X$  **odotusarvo** on *vakio*

$$E(X) = \mu_X = \int_{-\infty}^{+\infty} xf(x)dx$$

- Sanomme, että satunnaismuuttujan  $X$  odotusarvo  $E(X)$  on sen *jakauman odotusarvo*, joka kuvaa satunnaismuuttujaan  $X$  liittyviä todennäköisyyksiä.

## Odotusarvo

# Jatkuvan jakauman odotusarvo: Kommentteja

---

- Vaikka satunnaismuuttujan sama arvo vaihtelee satunnaisesti koetoistosta toiseen, satunnaismuuttuja saa *keskimäärin* arvoja, jotka vaihtelevat sen odotusarvon ympärillä.
- *Jos jakaumalla on odotusarvo*, se on jakauman todennäköisyysmassan **painopiste**.
- *Jatkuvan* jakauman odotusarvo *kuuluu aina* ko. satunnaismuuttujan tulosvaihtoehtojen joukkoon.



## Odotusarvo

# Jatkuvan jakauman odotusarvo: Esimerkki tasaisesta jakaumasta

- Jatkuvan tasaisen jakauman tiheysfunktio on

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{muulloin} \end{cases}$$

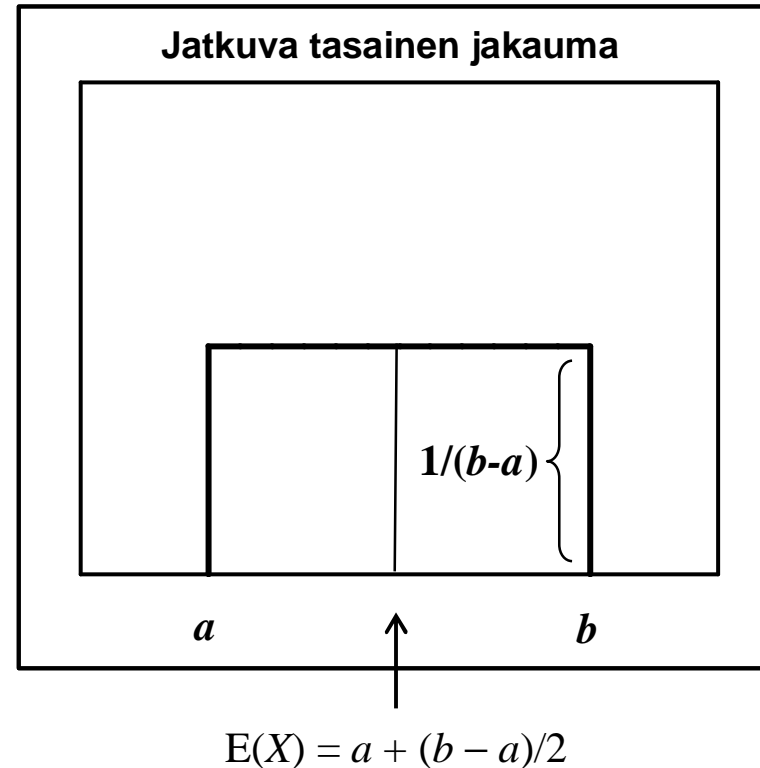
- Jakauman odotusarvo on

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx$$

$$= \int_a^b x \cdot \frac{1}{b-a} dx$$

$$= \frac{1}{b-a} \left[ \frac{1}{2} x^2 \right]_a^b$$

$$= (b+a)/2 = a + (b-a)/2$$



## Odotusarvo

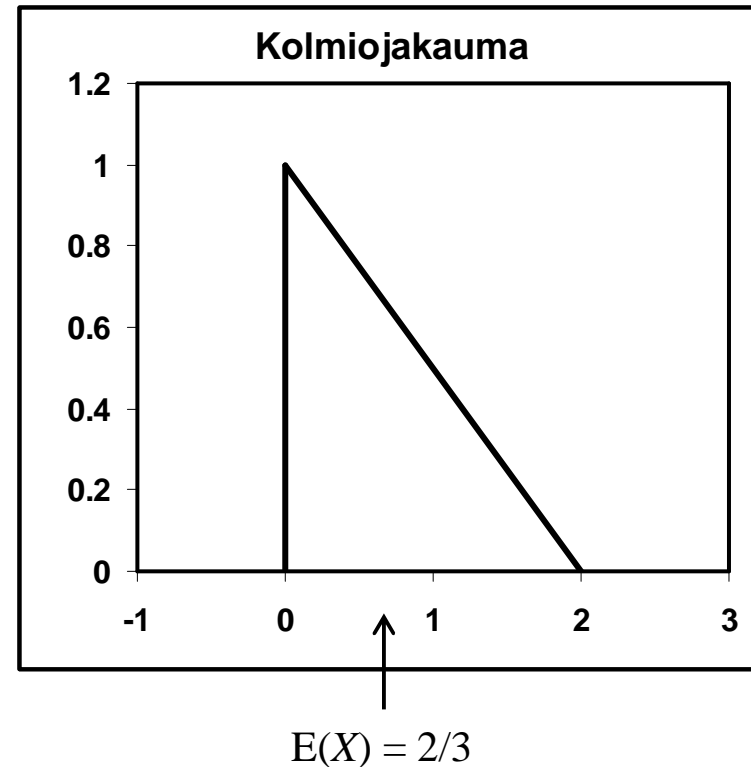
# Jatkuvan jakauman odotusarvo: Esimerkki kolmiojakaumasta

- Erään kolmiojakauman tiheysfunktio on

$$f(x) = \begin{cases} -\frac{1}{2}x + 1, & 0 \leq x \leq 2 \\ 0, & \text{muulloin} \end{cases}$$

- Jakauman odotusarvo on

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} xf(x)dx \\ &= \int_0^2 x\left(-\frac{1}{2}x + 1\right)dx \\ &= \left[-\frac{1}{6}x^3 + \frac{1}{2}x^2\right]_0^2 \\ &= \frac{2}{3} \end{aligned}$$



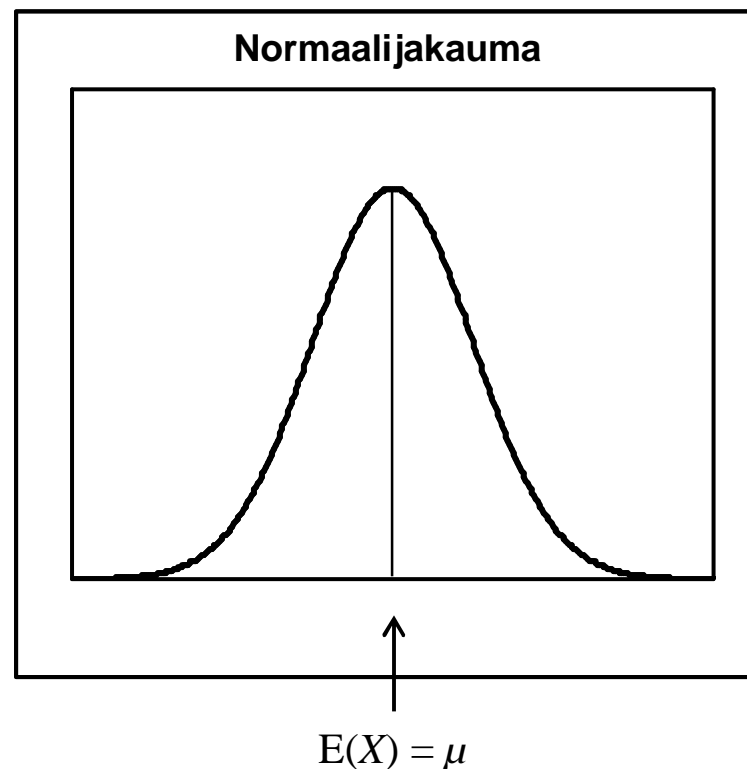
Odotusarvo

## Jatkuvan jakauman odotusarvo: Esimerkki normaalijakaumasta

- **Normaalijakauman tiheysfunktio** on

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\left(\frac{x-\mu}{\sigma}\right)^2\right\}$$

- Normaalijakauman tiheysfunktio on *symmetrinen* suoran  $x = \mu$  suhteen.
- Voidaan osoittaa, että normaalijakauman *odotusarvo*  $E(x) = \mu$   
ks. lukua **Jatkuvia jakaumia**.



## Odotusarvon olemassaolo

---

- Jakaumalla *ei välttämättä ole* odotusarvoa.
- **Odotusarvon olemassaololla** tarkoitetaan diskreetin jakauman tapauksessa sitä, että

$$\sum_i |x_i| f(x_i) < \infty$$

ja jatkuvan jakauman tapauksessa sitä, että

$$\int_{-\infty}^{+\infty} |x| f(x) dx < \infty$$

## Odotusarvo ja todennäköisyysmassan painopiste

---

- Jos jakaumalla on odotusarvo, se yhtyy aina ko. jakauman todennäköisyysmassan **painopisteeseen**.

- Olkoon

$$E(X) = \mu$$

satunnaismuuttujan  $X$  odotusarvo.

- Jos satunnaismuuttujan  $X$  jakauma on *symmetrinen* suoran

$$x = a$$

suhteen, niin

$$E(X) = \mu = a$$

Odotusarvo

## Vakion odotusarvo

---

- Olkoon  $a$  *ei-satunnainen vakio*.
- **Vakion odotusarvo** on vakio itse:

$$E(a) = a$$

- **Kommentti:**

*Vakio ei vaihtelee koetoistosta toiseen.*

## Odotusarvo

# Vakion odotusarvo: Perustelu

---

- Väite: *Vakiolle*  $a$  pätee

$$E(a) = a$$

- Perustelu *jatkuvan jakauman* tapauksessa:

$$E(a) = \int_{-\infty}^{+\infty} af(x)dx = a \int_{-\infty}^{+\infty} f(x)dx = a \cdot 1 = a$$

## Lineaarimuunnoksen odotusarvo

---

- Olkoon satunnaismuuttujan  $X$  odotusarvo  $E(X)$ .
- **Satunnaismuuttujan  $X$  lineaarimuunnoksen**

$$Y = a + bX$$

( $a$  ja  $b$  vakioita) odotusarvo  $E(Y)$  saadaan soveltamalla ko. lineaarimuunnosta odotusarvoon  $E(X)$ :

$$E(Y) = a + b E(X)$$



Odotusarvo

## Lineaarimuunnoksen odotusarvo: Perustelu

---

- Väite: *Lineaarimuunnokselle*

$$Y = a + bX$$

pätee

$$E(Y) = a + bE(X).$$

- Perustelu *jatkuvan jakauman* tapauksessa:

$$\begin{aligned} E(Y) &= E(a + bX) = \int_{-\infty}^{+\infty} (a + bx) f(x) dx \\ &= a \int_{-\infty}^{+\infty} f(x) dx + b \int_{-\infty}^{+\infty} xf(x) dx \\ &= a + bE(X) \end{aligned}$$

## Lineaarimuunnoksen odotusarvo: Kommentteja

---

- Satunnaismuuttujan  $X$  kertominen vakiolla  $b$  merkitsee satunnaismuuttujan  $X$  saamien arvojen *mittakaavan muuttamista*.
- Satunnaismuuttujan  $X$  saamien arvojen mittakaavan muuttaminen *verrannollisuuskertoimella*  $b$  muuttaa satunnaismuuttujan  $X$  jakauman todennäköisyysmassan painopistettä samalla kertoimella.
- Vakion  $a$  lisääminen satunnaismuuttujaan  $X$  merkitsee satunnaismuuttujan  $X$  jakauman todennäköisyysmassan *siirtoa*.
- Todennäköisyysmassan siirtäminen vakion  $a$  verran siirtää todennäköisyysmassan painopistettä *saman* verran.

## Odotusarvo jakauman sijaintiparametrina 1/2

---

- Koska odotusarvolla on *fysikaalinen tulkinta* todennäköisyysmassan *painopisteenä*, odotusarvoa voidaan kutsua jakauman **sijaintiparametriksi**.
- Oletetaan, että satunnaismuuttujien  $X$  ja  $Y$  *tiheysfunktiot yksihuippuisia ja symmetrisiä painopisteensä suhteen*.
- Tällöin satunnaismuuttujan  $X$  todennäköisyysmassan pääosa *sijaitsee vasemmalla* satunnaismuuttujan  $Y$  todennäköisyysmassan pääosasta, jos ja vain jos

$$E(X) < E(Y)$$

ks. havainnollistusta 1 >.

## Odotusarvo jakauman sijaintiparametrina 2/2

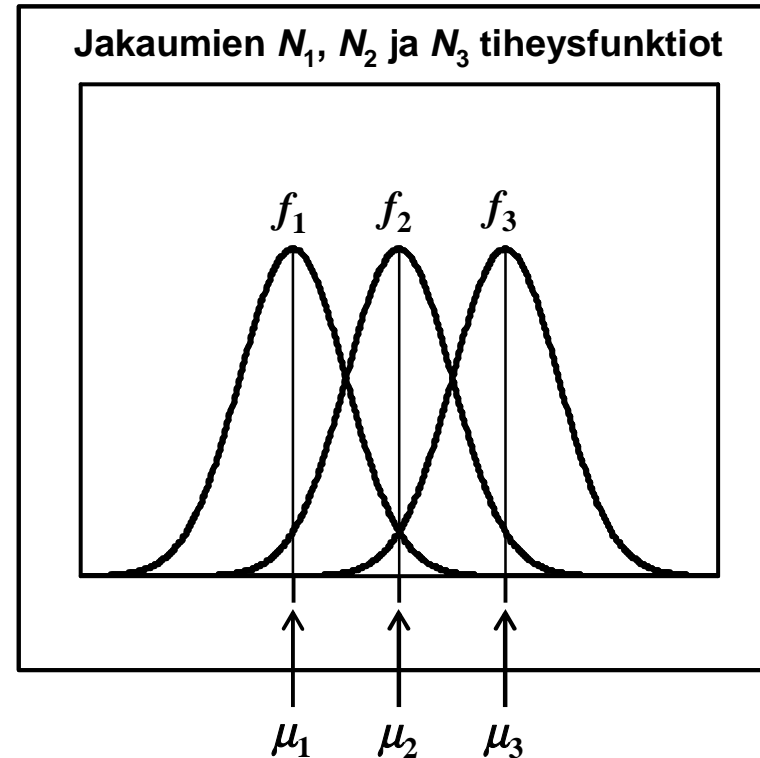
---

- Myös jos jakauma on *yksihuippuinen*, mutta *vino*, odotusarvo kuvaa luontevalla tavalla jakauman todennäköisyysmassan pääosan *sijaintia*; ks. havainnollistusta 2 >.
- Sen sijaan, jos jakauma on *monihuippuinen*, jakauman todennäköisyysmassan pääosien ei tarvitse olla lähellä odotusarvoa; ks. havainnollistusta 3 >.

## Odotusarvo

# Odotusarvo jakauman sijaintiparametrina: Havainnollistus 1

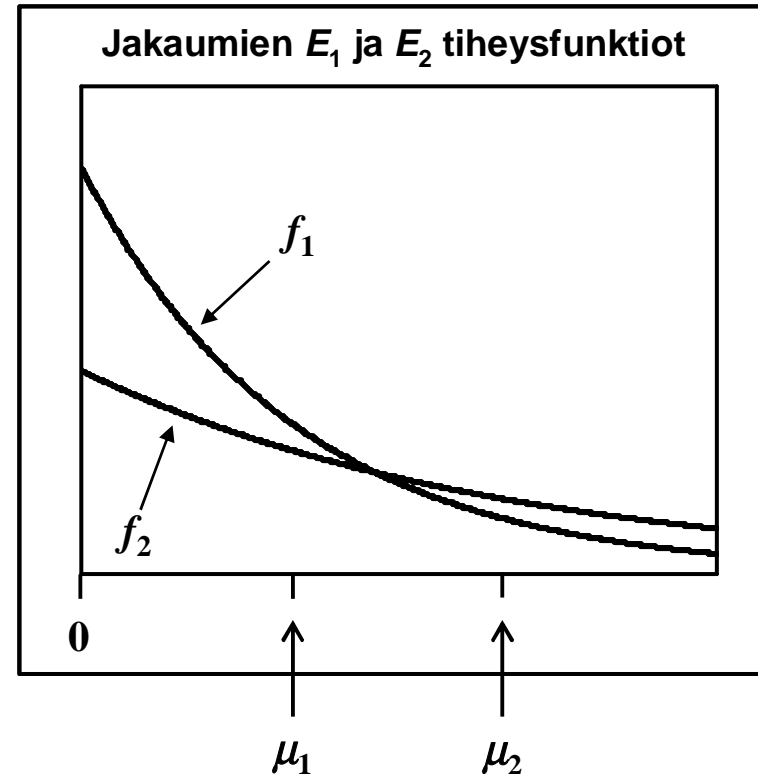
- Kuva oikealla esittää kolmen **normaalijakauman**  $N_1$ ,  $N_2$  ja  $N_3$  tiheysfunktioita  $f_1$ ,  $f_2$  ja  $f_3$ .
- Tiheysfunktiot  $f_1$ ,  $f_2$  ja  $f_3$  ovat *yksihuippuisia* ja *symmetrisiä* suorien  $x = \mu_1$ ,  $x = \mu_2$  ja  $x = \mu_3$  suhteen.
- Jakaumat  $N_2$  ja  $N_3$  saadaan *siirtämällä* jakauman  $N_1$  todennäköisyysmassaa *oikealle*.
- Jakaumien  $N_1$ ,  $N_2$  ja  $N_3$  odotusarvot  $\mu_1$ ,  $\mu_2$  ja  $\mu_3$  toteuttavat epäyhtälöt
$$\mu_1 < \mu_2 < \mu_3$$



## Odotusarvo

# Odotusarvo jakauman sijaintiparametrina: Havainnollistus 2

- Kuva oikealla esittää kahden **eksponenttijakauman**  $E_1$  ja  $E_2$  tiheysfunktioita  $f_1$  ja  $f_2$ .
- Tiheysfunktiot  $f_1$  ja  $f_2$  ovat *yksihuippuisia* ja *epäsymmetrisiä*.
- Jakauman  $E_1$  todennäköisyysmassa *on keskittynyt* jakauman  $E_2$  todennäköisyysmassaa *voimakkaammin origon lähelle*.
- Jakaumien  $E_1$  ja  $E_2$  odotusarvot  $\mu_1$  ja  $\mu_2$  toteuttavat epäyhtälön  $\mu_1 < \mu_2$

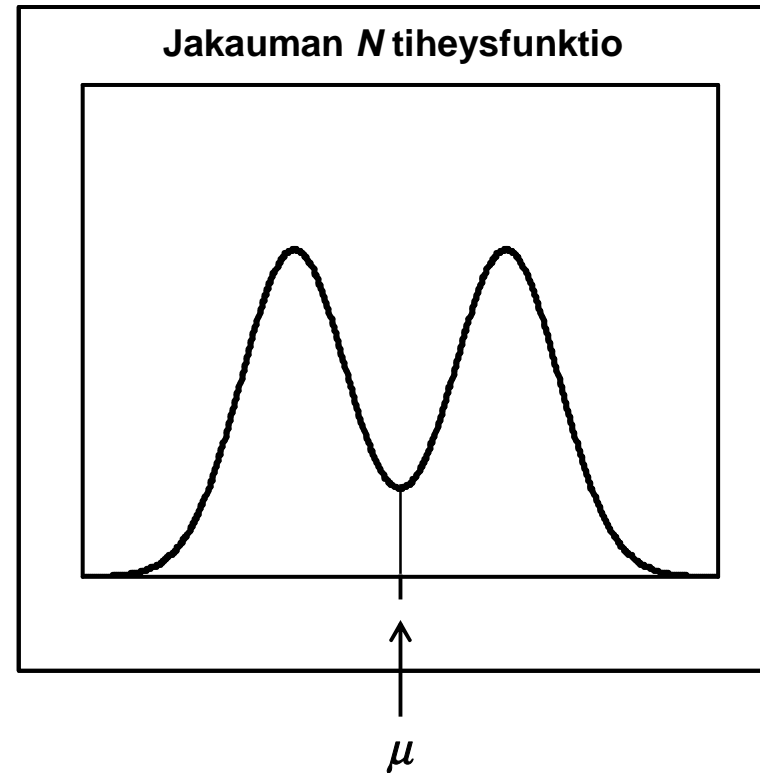


## Odotusarvo

# Odotusarvo jakauman sijaintiparametrina: Havainnollistus 3

---

- Kuva oikealla esittää erään **sekoitetun normaalijakauman**  $N$  tiheysfunktiota  $f$ .
- Tiheysfunktio  $f$  on *kaksi-huippuinen ja symmetrinen* suoran  $x = \mu$  suhteen.
- Jakauman  $N$  todennäköisyysmassalla *on* vaaka-akselilla *kaksi keskittymää*.
- Jakauman  $N$  odotusarvo  $\mu$  on todennäköisyysmassojen keskittymien *välissä*.



## Summan ja erotuksen odotusarvo

---

- Satunnaismuuttujien  $X$  ja  $Y$  **summan  $X + Y$  odotusarvo** on

$$E(X + Y) = E(X) + E(Y)$$

- Satunnaismuuttujien  $X$  ja  $Y$  **erotuksen  $X - Y$  odotusarvo** on

$$E(X - Y) = E(X) - E(Y)$$

- Tämä merkitsee sitä, että odotusarvo on *lineaarinen operaattori*.
- Huomautus:

Todistus vaatii *kaksiulotteisten satunnaismuuttujien* määrittelemistä ja esitetään luvussa **Moniulotteiset satunnaismuuttujat ja jakaumat**.



Odotusarvo

## Summan odotusarvo:

### Yleistys

---

- Olkoot  $X_i$ ,  $i = 1, 2, \dots, n$  satunnaismuuttujia ja  $a_i$ ,  $i = 1, 2, \dots, n$  vakioita.
- Satunnaismuuttujien  $X_i$ ,  $i = 1, 2, \dots, n$  **painotetun summan**

$$\sum a_i X_i$$

**odotusarvo** on

$$E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i E(X_i)$$

## Diskreetin satunnaismuuttujan funktion odotusarvo: Määritelmä

---

- Olkoon  $X$  diskreetti satunnaismuuttuja, jonka pistetodennäköisyysfunktio on

$$f(x_i) = \Pr(X = x_i) = p_i, \quad i = 1, 2, 3, \dots$$

- Olkoon  $g$  reaaliarvoinen funktio.
- Satunnaismuuttujan  $g(X)$  odotusarvo on vakio

$$E(g(X)) = \mu_{g(X)} = \sum_i g(x_i) f(x_i)$$

## Jatkuvan satunnaismuuttujan funktion odotusarvo: Määritelmä

---

- Olkoon  $X$  jatkuva satunnaismuuttuja, jonka tiheysfunktio on

$$f(x)$$

- Olkoon  $g$  reaaliarvoinen funktio.
- Satunnaismuuttujan  $g(X)$  odotusarvo on vakio

$$E(g(X)) = \mu_{g(X)} = \int_{-\infty}^{+\infty} g(x) f(x) dx$$

# Jakaumien tunnusluvut

---

**Odotusarvo**

**>> Varianssi**

**Markovin ja Tshebyshevin epäyhtälöt**

**Momentit**

**Vinous ja huipukkuus**

**Kvantiilit**

**Moodi**

**Suurten lukujen laki**

## Varianssi

### Varianssi:

### Yleinen määritelmä

---

- Olkoon satunnaismuuttujan  $X$  odotusarvo

$$E(X) = \mu_X$$

- Satunnaismuuttujan  $X$  **varianssi** on *vakio*

$$D^2(X) = \text{Var}(X) = \sigma_X^2 = E(X - \mu_X)^2$$

- Satunnaismuuttujan  $X$  varianssi on satunnaismuuttujan  $X$  omasta odotusarvostaan  $\mu_X$  määrätyn *poikkeaman neliön odotusarvo*.

## Diskreetin jakauman varianssi

---

- Olkoon  $X$  *diskreetti satunnaismuuttuja*.
- Olkoon  $\{x_1, x_2, \dots\}$  satunnaismuuttujan  $X$  *tulosvaihtoehtojen* eli arvojen joukko.
- Olkoon satunnaismuuttujan  $X$  *pistetodennäköisyysfunktio*

$$f(x_i) = \Pr(X = x_i) = p_i, \quad i = 1, 2, \dots$$

- Tällöin satunnaismuuttujan  $X$  **varianssi** on *vakio*

$$D^2(X) = \text{Var}(X) = \sigma_X^2 = \sum_i (x_i - \mu_X)^2 p_i$$

## Varianssi

# Jatkuvan jakauman varianssi

---

- Olkoon  $X$  on *jatkuva satunnaismuuttuja*.
- Olkoon satunnaismuuttujan  $X$  *tiheysfunktio*  $f(x)$ .
- Tällöin satunnaismuuttujan  $X$  **varianssi** on *vakio*

$$D^2(X) = \text{Var}(X) = \sigma_X^2 = \int_{-\infty}^{+\infty} (x - \mu_X)^2 f(x) dx$$

## Varianssin olemassaolo

---

- Jakaumalla *ei välttämättä ole* varianssia.
- **Varianssin olemassaololla** tarkoitetaan sitä, että varianssin määrittelevä summa (diskreetin jakauman tapauksessa) tai integraali (jatkuvan jakauman tapauksessa) on äärellinen.



## Varianssi

### Varianssin määritelmä:

### Kommentteja

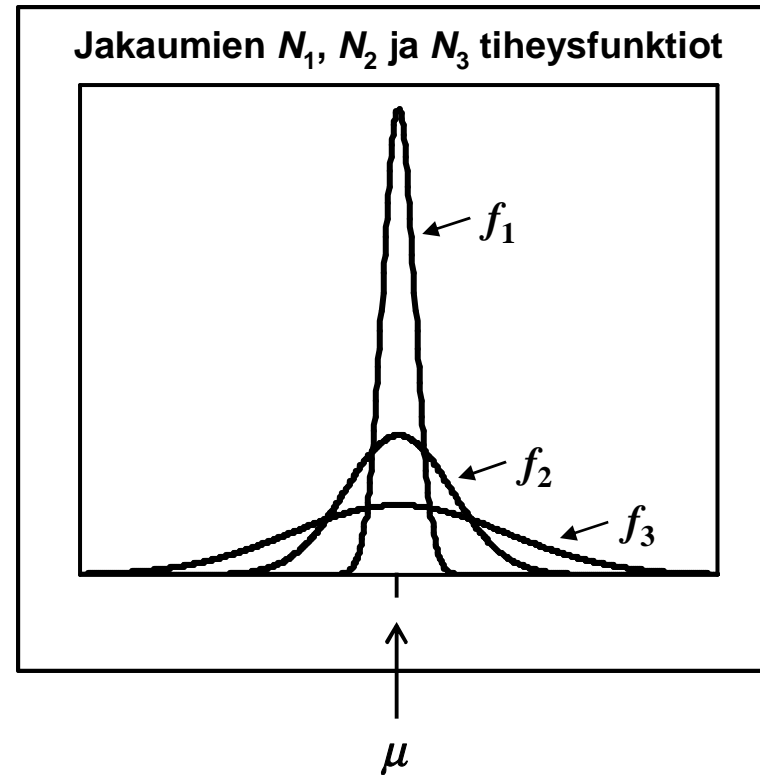
---

- Varianssi kuvaa todennäköisyysmassan **hajaantuneisuutta** tai – mikä on sama asia – **keskittyneisyyttä** jakauman *painopisteen* suhteen.
- Jos
$$\text{Var}(X) > \text{Var}(Y)$$
niin satunnaismuuttujan  $X$  todennäköisyysmassassa on hajaantunut *voimakkaammin oman painopisteeseensä suhteen* kuin satunnaismuuttujan  $Y$  todennäköisyysmassassa *oman painopisteeseensä suhteen*.
- Koska varianssi kuvaa todennäköisyysmassan hajaantuneisuutta, sitä voidaan kutsua **hajonta-parametriksi**.

## Varianssi

# Varianssi jakauman hajaantuneisuuden mittana: Esimerkki normaalijakaumista 1/2

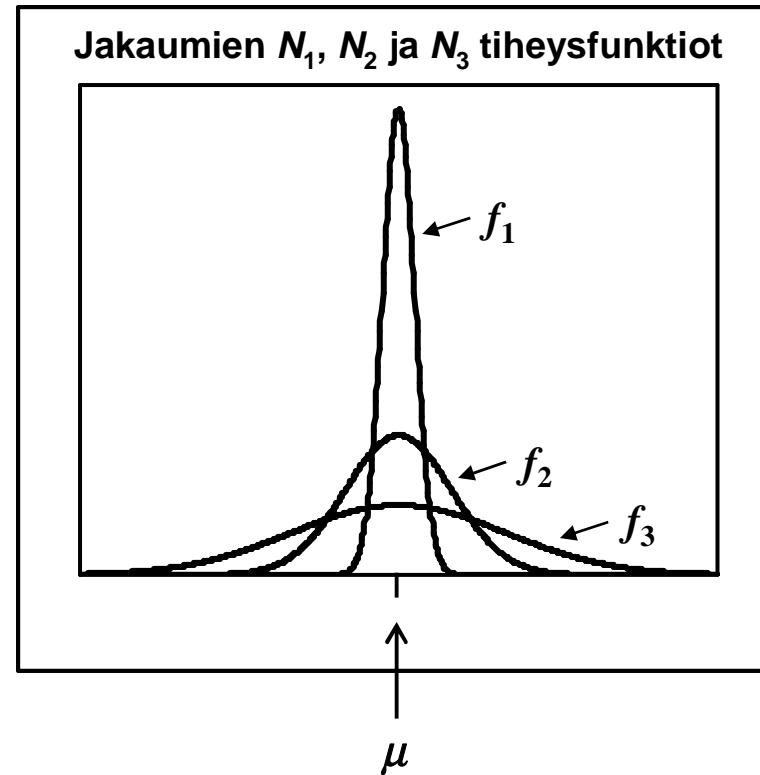
- Kuva oikealla esittää kolmen **normaalijakauman**  $N_1$ ,  $N_2$  ja  $N_3$  tiheysfunktioita  $f_1$ ,  $f_2$  ja  $f_3$ .
- Kaikilla jakaumilla on *sama* odotusarvo  $\mu$ .
- Tiheysfunktiot  $f_1$ ,  $f_2$  ja  $f_3$  ovat *yksihuippuisia* ja *symmetrisiä* suoran  $x = \mu$  suhteen.



## Varianssi

# Varianssi jakauman hajaantuneisuuden mittana: Esimerkki normaalijakaumista 2/2

- Jakauman  $N_1$  todennäköisyysmassa on *keskittynein*, kun taas jakauman  $N_3$  todennäköisyysmassa on *hajaantunein*.
- Jakaumien varianssit toteuttavat epäyhtälöt:  
$$\text{Var}(X_1) < \text{Var}(X_2) < \text{Var}(X_3)$$



Varianssi

**Varianssi:**

**Toinen laskukaava**

---

- Olkoon satunnaismuuttujan  $X$  *odotusarvo*

$$E(X) = \mu$$

- Satunnaismuuttujan  $X$  *varianssi* voidaan laskea myös kaavalla

$$\text{Var}(X) = \alpha_2 - \mu^2$$

jossa

$$\alpha_2 = E(X^2)$$

on satunnaismuuttujan  $X$  *toinen (origo-) momentti*.

## Varianssi

# Varianssin toinen laskukaava: Perustelu

---

- Olkoot satunnaismuuttujan  $X$  odotusarvo

$$E(X) = \mu$$

ja toinen momentti

$$E(X^2) = \alpha_2$$

- Satunnaismuuttujan  $X$  varianssi on

$$\begin{aligned} \text{Var}(X) &= E(X - \mu)^2 \\ &= E(X^2 - 2\mu X + \mu^2) \\ &= E(X^2) - 2\mu E(X) + E(\mu^2) \\ &= \alpha_2 - 2\mu \times \mu + \mu^2 \\ &= \alpha_2 - \mu^2 \end{aligned}$$

Varianssi

## Standardipoikkeama: Määritelmä

---

- Olkoon satunnaismuuttujan  $X$  odotusarvo

$$E(X) = \mu_X$$

- Satunnaismuuttujan  $X$  **standardipoikkeama** on vakio

$$D(X) = \sigma_X = \sqrt{E(X - \mu_X)^2}$$

- Standardipoikkeamaa käytetään samaan tapaan kuin varianssia todennäköisyyssmassan *hajaantuneisuuden* (*keskittyneisyyden*) mittana.
- Standardipoikkeama on – toisin kuin varianssi – samoissa *mittayksiköissä* kuin odotusarvo.

## Varianssi

# Diskreetin jakauman varianssi: Esimerkki nopanheitosta 1/2

---

- Nopanheiton tulosta satunnaisilmiönä kuvaavan **diskreetin tasaisen jakauman pistetodennäköisyysfunktio**:

$$P(X = i) = \frac{1}{6}, i = 1, 2, 3, 4, 5, 6$$

- Satunnaismuuttujan  $X$  odotusarvo:

$$E(X) = \mu = \sum_{i=1}^6 i \Pr(X = i) = \sum_{i=1}^6 i \frac{1}{6} = \frac{21}{6} = 3.5$$

- Satunnaismuuttujan  $X$  toinen momentti:

$$\begin{aligned} E(X^2) = \alpha_2 &= \sum_{i=1}^6 i^2 \Pr(X = i) = \sum_{i=1}^6 i^2 \frac{1}{6} \\ &= \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2}{6} = \frac{91}{6} \end{aligned}$$

## Varianssi

# Diskreetin jakauman varianssi: Esimerkki nopanheitosta 2/2

---

- Satunnaismuuttujan  $X$  varianssi:

$$\text{Var}(X) = D^2(X) = \sigma^2 = \alpha_2 - \mu^2 = \frac{91}{6} - \left(\frac{21}{6}\right)^2 = \frac{35}{12} \approx 2.917$$

- Satunnaismuuttujan  $X$  standardipoikkeama eli keskihajonta:

$$D(X) = \sigma = \sqrt{2.917} \approx 1.708$$



## Varianssi

# Diskreetin jakauman odotusarvo ja varianssi: Laskujen järjestäminen 1/5

---

- Nopanheiton tulosta satunnaisilmionä kuvaavan **diskreetin tasaisen jakauman** *odotusarvon* ja *varianssin* määrittämistä varten tarvittavat laskutoimitukset voidaan järjestää seuraavan taulukon muotoon:

	Keskiarvo			Varianssi 1		Varianssi 2		
1	2	3	4	5	6	7	8	9
$i$	$x_i$	$p_i$	$x_i p_i$	$x_i^2$	$x_i^2 p_i$	$x_i - \mu$	$(x_i - \mu)^2$	$(x_i - \mu)^2 p_i$
1	1	1/6	1/6	1	1/6	-2.5	6.25	25/24
2	2	1/6	2/6	4	4/6	-1.5	2.25	9/24
3	3	1/6	3/6	9	9/6	-0.5	0.25	1/24
4	4	1/6	4/6	16	16/6	+0.5	0.25	1/24
5	5	1/6	5/6	25	25/6	+1.5	2.25	9/24
6	6	1/6	6/6	36	36/6	+2.5	6.25	25/24
$\Sigma$	21	1	21/6	91	91/6	0	17.5	70/24

## Varianssi

# Diskreetin jakauman odotusarvo ja varianssi: Laskujen järjestäminen 2/5

---

- Taulukon rivillä  $\Sigma$  on sarakesummat riveiltä 1-6:

$$\text{Sarake 2: } \sum x_i = 21$$

$$\text{Sarake 3: } \sum p_i = 1$$

$$\text{Sarake 4: } \sum x_i p_i = \mu = 21/6 = 3.5$$

$$\text{Sarake 5: } \sum x_i^2 = 91$$

$$\text{Sarake 6: } \sum x_i^2 p_i = \alpha_2 = 91/6 = 15.167$$

$$\text{Sarake 7: } \sum (x_i - \mu) = 0$$

$$\text{Sarake 8: } \sum (x_i - \mu)^2 = 17.5$$

$$\text{Sarake 9: } \sum (x_i - \mu)^2 p_i = \sigma^2 = 70/24 = 2.917$$

## Varianssi

# Diskreetin jakauman odotusarvo ja varianssi: Laskujen järjestäminen 3/5

---

- Satunnaismuuttujan  $X$  odotusarvon määrittämistä varten tarvittavat laskutoimitukset on suoritettu sarakkeissa 2-4.
- Odotusarvo saadaan rivin  $\Sigma$  sarakkeesta 4:

$$E(X) = \mu = \sum x_i p_i = 21/6 = 3.5$$

## Varianssi

# Diskreetin jakauman odotusarvo ja varianssi: Laskujen järjestäminen 4/5

---

- Satunnaismuuttujan  $X$  varianssi voidaan määrätä kahdella eri tavalla:
- Kaava 1:

$$\text{Var}(X) = \alpha_2 - \mu^2$$

jossa

$$\alpha_2 = E(X^2) = \sum x_i^2 p_i$$

$$\mu = E(X) = \sum x_i p_i$$

- Kaava 2:

$$\text{Var}(X) = E(X - \mu)^2 = \sigma^2 = \sum (x_i - \mu)^2 p_i$$

jossa  $\mu$  on kuten kaavassa 1.

## Varianssi

# Diskreetin jakauman odotusarvo ja varianssi: Laskujen järjestäminen 5/5

---

- Kaavan 1 vaatimat laskutoimitukset on tehty sarakkeissa 2-4 ja 5-6:

$$E(X) = \mu = \sum x_i p_i = 21/6 = 3.5$$

$$E(X^2) = \alpha_2 = \sum x_i^2 p_i = 91/6 = 15.167$$

- Kaavan 1 mukaan

$$\text{Var}(X) = \alpha_2 - \mu^2 = \frac{91}{6} - \left(\frac{21}{6}\right)^2 = \frac{35}{12} = 2.917$$

- Kaavan 2 vaatimat laskutoimitukset on tehty sarakkeissa 2-4 ja 7-9:

$$\text{Var}(X) = E(X - \mu)^2 = \sigma^2 = \sum (x_i - \mu)^2 p_i = \frac{70}{24} = 2.917$$

- Kaavan 2 soveltaminen on siinä mielessä monimutkaisempaa kuin kaavan 1 soveltaminen, että kaavassa 2 on erotuksien  $(x_i - \mu)$  määrittämiseksi *ensin* määrättävä odotusarvo  $\mu$ .

## Varianssi

# Jatkuvan jakauman odotusarvo ja varianssi: Esimerkki tasaisesta jakaumasta 1/2

---

- Erään **jatkuvan tasaisen jakauman** *tiheysfunktio*:

$$f(x) = \begin{cases} \frac{1}{b}, & 0 \leq x \leq b \\ 0, & \text{muulloin} \end{cases}$$

- Satunnaismuuttujan  $X$  *odotusarvo*:

$$\mathbf{E}(X) = \mu = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^b x \frac{1}{b} dx = \left[ \frac{1}{2b} x^2 \right]_0^b = \frac{b}{2}$$

- Satunnaismuuttujan  $X$  *toinen momentti*:

$$\mathbf{E}(X^2) = \alpha_2 = \int_{-\infty}^{+\infty} x^2 f(x)dx = \int_0^b x^2 \frac{1}{b} dx = \left[ \frac{x^3}{3b} \right]_0^b = \frac{b^2}{3}$$

## Varianssi

# Jatkuvan jakauman odotusarvo ja varianssi: Esimerkki tasaisesta jakaumasta 2/2

---

- Satunnaismuuttujan  $X$  varianssi:

$$\text{Var}(X) = D^2(X) = \sigma^2 = \alpha_2 - \mu^2 = \frac{b^2}{3} - \left(\frac{b}{2}\right)^2 = \frac{b^2}{12}$$

- Satunnaismuuttujan  $X$  standardipoikkeama:

$$D(X) = \sigma = \sqrt{\frac{b^2}{12}} = \frac{b}{2\sqrt{3}}$$

## Varianssi

# Vakion varianssi

---

- Olkoon  $a$  *ei-satunnainen vakio*.

- **Vakion varianssi** on nolla:

$$\text{Var}(a) = 0$$

- Tulkinta:

*Vakio ei vaihtelee* satunnaiskokeesta toiseen.



## Varianssi

# Vakion varianssi: Perustelu

---

- Väite: Vakiolle  $a$  pätee

$$\text{Var}(a) = 0$$

- Perustelu:

$$\text{Var}(a) = E(a - E(a))^2 = E(a - a)^2 = E(0) = 0$$

koska vakiolle  $a$  pätee:

$$E(a) = a$$

## Varianssi

# Lineaarimuunnoksen varianssi

---

- Olkoon satunnaismuuttujan  $X$  varianssi  $\text{Var}(X)$ .
- **Satunnaismuuttujan  $X$  lineaarimuunnoksen**

$$Y = a + bX$$

( $a$  ja  $b$  vakioita) varianssi on

$$\text{Var}(Y) = b^2 \text{Var}(X)$$

## Varianssi

# Lineaarimuunnoksen varianssi: Perustelu

---

- Väite: *Lineaarimuunnokselle*

$$Y = a + bX$$

pätee

$$\text{Var}(Y) = b^2 \text{Var}(X).$$

- Perustelu:

$$\begin{aligned}\text{Var}(Y) &= \text{Var}(a + bX) = \text{E}[(a + bX) - \text{E}(a + bX)]^2 \\ &= \text{E}[a + bX - a - b\text{E}(X)]^2 \\ &= \text{E}[bX - b\text{E}(X)]^2 \\ &= b^2 \text{E}[X - \text{E}(X)]^2 \\ &= b^2 \text{Var}(X)\end{aligned}$$

## Varianssi

# Lineaarimuunnoksen varianssi: Kommentteja

---

- Satunnaismuuttujan  $X$  kertominen vakiolla  $b$  merkitsee satunnaismuuttujan  $X$  saamien arvojen *mittakaavan muuttamista*.
- Satunnaismuuttujan  $X$  saamien arvojen mittakaavan muuttaminen *verrannollisuuskertoimella*  $b$  muuttaa satunnaismuuttujan  $X$  varianssia kertoimella  $b^2$ .
- Vakion  $a$  lisääminen satunnaismuuttujaan  $X$  merkitsee satunnaismuuttujan  $X$  jakauman todennäköisyysmassan *siirtoa*.
- Todennäköisyysmassan siirtäminen *ei muuta* todennäköisyysmassan hajaantuneisuutta.

## Varianssi

# Standardointi

---

- Olkoon  $X$  satunnaismuuttuja, jonka odotusarvo  $E(X) = \mu$  ja varianssi  $D^2(X) = \sigma^2$ .
- Tällöin **standardoidun satunnaismuuttujan**

$$Z = \frac{X - \mu}{\sigma}$$

odotusarvo

$$E(Z) = 0$$

ja varianssi

$$D^2(Z) = 1$$

## Varianssi

# Standardointi:

## Perustelu

---

- Olkoot  $E(X) = \mu$  ja  $D^2(X) = \sigma^2$ .
- Standardoidaan satunnaismuuttuja  $X$ :

$$Z = \frac{X - \mu}{\sigma}$$

- Tällöin

$$E(Z) = E\left(\frac{1}{\sigma}X - \frac{\mu}{\sigma}\right) = \frac{1}{\sigma}E(X) - \frac{\mu}{\sigma} = \frac{1}{\sigma}\mu - \frac{\mu}{\sigma} = 0$$

$$D^2(Z) = D^2\left(\frac{1}{\sigma}X - \frac{\mu}{\sigma}\right) = \frac{1}{\sigma^2}D^2(X) = \frac{1}{\sigma^2}\sigma^2 = 1$$

## Summan ja erotuksen varianssi 1/2

---

- Oletetaan, että satunnaismuuttujat  $X$  ja  $Y$  ovat *riippumattomia*.
- Tarkastellaan satunnaismuuttujien  $X$  ja  $Y$  *summan*  $X + Y$  ja *erotuksen*  $X - Y$  varianssia.
- Huomautus:

Satunnaismuuttujien  $X$  ja  $Y$  *riippumattomuudella* tarkoitetaan seuraavaa:

Se, mitä arvoja satunnaismuuttuja  $X$  saa, ei saa riippua siitä, mitä arvoja satunnaismuuttuja  $Y$  saa ja kääntäen, se, mitä arvoja satunnaismuuttuja  $Y$  saa, ei saa riippua siitä, mitä arvoja satunnaismuuttuja  $X$  saa; käsite täsmennetään luvussa **Moniulotteiset satunnaismuuttujat ja jakaumat**.

## Summan ja erotuksen varianssi 2/2

---

- Riippumattomien satunnaismuuttujien  $X$  ja  $Y$  **summan  $X + Y$  varianssi** on

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$$

- Riippumattomien satunnaismuuttujien  $X$  ja  $Y$  **erotuksen  $X - Y$  varianssi** on

$$\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y)$$

- Huomautus:

Todistus vaatii *kaksiulotteisen satunnaismuuttujan* määrittelemistä ja esitetään luvussa **Moniulotteiset satunnaismuuttujat ja jakaumat**.



## Summan ja erotuksen varianssi: Kommentteja

---

- Oletetaan, että satunnaismuuttujat  $X$  ja  $Y$  ovat *riippumattomia*.
- Tällöin satunnaismuuttujien  $X$  ja  $Y$  *summan ja erotuksen* varianssille pätee

$$\text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y)$$

- Huomaa:

$$\text{Var}(X - Y) \neq \text{Var}(X) - \text{Var}(Y)$$

$$D(X + Y) \neq D(X) + D(Y)$$

$$D(X - Y) \neq D(X) - D(Y)$$

## Varianssi

# Summan varianssi:

## Yleistys

---

- Olkoot satunnaismuuttujat  $X_i$ ,  $i = 1, 2, \dots, n$  riippumattomia ja  $a_i$ ,  $i = 1, 2, \dots, n$  vakioita.
- Tällöin satunnaismuuttujien  $X_i$ ,  $i = 1, 2, \dots, n$  **painotetun summan**

$$\sum a_i X_i$$

**varianssi on**

$$\text{Var}\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 \text{Var}(X_i)$$

## Empiirinen jakauma 1/3

---

- Oletetaan, että *diskreetin satunnaismuuttujan*  $X$  *mahdolliset arvot* ovat

$$x_i, i = 1, 2, \dots, n$$

- Liitetään satunnaismuuttujan  $X$  arvoihin *symmetriset todennäköisyydet*

$$\Pr(X = x_i) = p_i = \frac{1}{n}, i = 1, 2, \dots, n$$

- Otannan perustyypissä, *yksinkertaisessa satunnaisotannassa*, *havaintoarvot*  $x_i$  noudattavat tätä, ns. **empiiristä jakaumaa**.

- Suoraan diskreetin satunnaismuuttujan odotusarvon ja varianssin määritelmistä saadaan:

$$E(X) = \mu = \sum_{i=1}^n x_i p_i = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$$

$$E(X^2) = \alpha_2 = \sum_{i=1}^n x_i^2 p_i = \frac{1}{n} \sum_{i=1}^n x_i^2$$

$$\text{Var}(X) = D^2(X) = \sigma^2 = \sum_{i=1}^n (x_i - \mu)^2 p_i = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

## Empiirinen jakauma 3/3

---

- Huomaa, että odotusarvo

$$E(X) = \mu = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

on lukujen  $x_i$  **aritmeettinen keskiarvo** ja

$$\text{Var}(X) = D^2(X) = \sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

on lukujen  $x_i$  ns. **otosvarianssi**.

## Varianssi

# Aritmeettisen keskiarvon odotusarvo ja varianssi 1/2

---

- Olkoot  $X_i$ ,  $i = 1, 2, \dots, n$  riippumattomia satunnaismuuttujia.
- Oletetaan lisäksi, että satunnaismuuttujilla  $X_i$ ,  $i = 1, 2, \dots, n$  on sama odotusarvo ja varianssi:

$$E(X_i) = \mu, D^2(X_i) = \sigma^2, i = 1, 2, \dots, n$$

- Olkoon

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

satunnaismuuttujien  $X_i$ ,  $i = 1, 2, \dots, n$  aritmeettinen keskiarvo.

## Varianssi

# Aritmeettisen keskiarvon odotusarvo ja varianssi 2/2

---

- Tällöin

$$E(\bar{X}) = \mu$$

$$D^2(\bar{X}) = \frac{\sigma^2}{n}$$

- Huomautuksia:
  - Satunnaismuuttujien  $X_i$  aritmeettisen keskiarvon odotusarvo on sama kuin yksittäisten muuttujien yhteinen odotusarvo.
  - Satunnaismuuttujien  $X_i$  aritmeettinen keskiarvo vaihtelee varianssilla mitattuna *vähemmän* kuin muuttajat itse.

## Varianssi

# Aritmeettisen keskiarvon odotusarvo ja varianssi: Perustelu

---

- Olkoot  $X_i, i = 1, 2, \dots, n$  riippumattomia satunnaismuuttujia, joille

$$E(X_i) = \mu, D^2(X_i) = \sigma^2, i = 1, 2, \dots, n$$

- Tällöin

$$\begin{aligned} E(\bar{X}) &= E\left(\frac{1}{n} \sum_i X_i\right) = \frac{1}{n} E\left(\sum_i X_i\right) = \frac{1}{n} \sum_i E(X_i) \\ &= \frac{1}{n} \sum_i \mu = \frac{1}{n} \cdot n\mu = \mu \end{aligned}$$

$$\begin{aligned} D^2(\bar{X}) &= D^2\left(\frac{1}{n} \sum_i X_i\right) = \frac{1}{n^2} D^2\left(\sum_i X_i\right) = \frac{1}{n^2} \sum_i D^2(X_i) \\ &= \frac{1}{n^2} \sum_i \sigma^2 = \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n} \end{aligned}$$



# Jakaumien tunnusluvut

---

**Odotusarvo**

**Varianssi**

**>> Markovin ja Tshebyshevin epäyhtälöt**

**Momentit**

**Vinous ja huipukkuus**

**Kvantiilit**

**Moodi**

**Suurten lukujen laki**

## Markovin epäyhtälö

---

- Olkoon  $g(X)$  satunnaismuuttujan  $X$  *positiivinen* reaaliarvoinen funktio, jonka *odotusarvo* on  
 $E(g(X))$
- Tällöin jokaiselle reaaliselle vakiolle  $a > 0$  pätee  
**Markovin epäyhtälö**

$$\Pr(g(X) \geq a) \leq \frac{E(g(X))}{a}$$

## Markovin epäyhtälö:

### Todistus

---

- Todistamme Markovin epäyhtälön *jatkuvien* satunnaismuuttujien tapauksessa.
- Olkoon  $g(X)$  satunnaismuuttujan  $X$  *positiivinen* reaaliarvoinen funktio, jonka odotusarvo on  $E(g(X))$ .
- Olkoon satunnaismuuttujan  $X$  *tiheysfunktio*  $f(x)$  ja olkoon  $a > 0$  vakio.
- *Markovin epäyhtälö* saadaan epäyhtälöketjusta

$$\begin{aligned} E(g(X)) &= \int_{-\infty}^{+\infty} g(x) f(x) dx \\ &\geq a \int_S f(x) dx \\ &= a \Pr(g(X) \geq a) \end{aligned}$$

jossa

$$S = \{x \mid g(x) \geq a\}$$

## Markovin epäyhtälö: Kommentteja 1/2

---

- Markovin epäyhtälön

$$\Pr(g(X) \geq a) \leq \frac{E(g(X))}{a}$$

mukaan todennäköisyys sille, että mielivaltaisen satunnaismuuttujan  $X$  (jolle odotusarvo  $E(g(X))$  on olemassa) *positiivinen* funktio  $g(X)$  saa *suurempia* arvoja kuin  $a > 0$ , on *korkeintaan*

$$\frac{E(g(X))}{a}$$

## Markovin epäyhtälö: Kommentteja 2/2

---

- Markovin epäyhtälön erikoistapauksena saadaan *positiivisille* satunnaismuuttujille  $X$  epäyhtälö

$$\Pr(X \geq a) \leq \frac{E(X)}{a}$$

jossa  $a > 0$ .

- Siten mielivaltaisen *positiivisen* satunnaismuuttujan (jonka odotusarvo  $E(X)$  on olemassa) todennäköisyysjakauman todennäköisyysmassasta *korkeintaan*

$$100 \times \frac{E(X)}{a} \%$$

on etäisyyttä  $a > 0$  *kauempana* origosta.

## Markovin ja Tshebyshevin epäyhtälöt

# Tshebyshevin epäyhtälö

---

- Olkoon  $X$  satunnaismuuttuja, jonka *odotusarvo* on

$$E(X) = \mu$$

ja *varianssi* on

$$\text{Var}(X) = \sigma^2$$

- Tällöin pätee **Tshebyshevin epäyhtälö**

$$\Pr(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

# Markovin ja Tshebyshevin epäyhtälöt

## Tshebyshevin epäyhtälö: Todistus

---

- Todistamme Tshebyshevin epäyhtälön *jatkuvien* satunnaismuuttujien tapauksessa.
- Olkoon  $X$  *jatkuva* satunnaismuuttuja.
- Olkoon satunnaismuuttujan  $X$  *tiheysfunktio*  $f(x)$ , sen *odotusarvo*  $E(X) = \mu$  ja *varianssi*  $\text{Var}(X) = \sigma^2$  sekä olkoon  $k > 0$  *vakio*.
- *Tshebyshevin epäyhtälö* seuraa *Markovin epäyhtälöstä*

$$\Pr(g(X) \geq a) \leq \frac{E(g(X))}{a}$$

valitsemalla

$$g(x) = (x - \mu)^2 ; \mu = E(X)$$

$$a = k^2 \sigma^2 ; \sigma^2 = \text{Var}(X) = E[g(X)]$$

## Markovin ja Tshebyshevin epäyhtälöt

# Tshebyshevin epäyhtälö:

### Kommentteja 1/3

---

- Tshebyshevin epäyhtälön

$$\Pr(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

mukaan mielivaltaisen satunnaismuuttujan  $X$  (jonka odotusarvo  $E(X) = \mu$  ja varianssi  $\text{Var}(X) = \sigma^2$  ovat olemassa) todennäköisyydeltään *korkeintaan*

$$100 \times \frac{1}{k^2} \%$$

on etäisyyttä  $k\sigma$  *kauempana* jakauman painopisteestä  $\mu$ .



## Markovin ja Tshebyshevin epäyhtälöt

# Tshebyshevin epäyhtälö:

### Kommentteja 2/3

---

- Jos  $X$  on mielivaltainen satunnaismuuttuja, jolla on odotusarvo ja varianssi, Tshebyshevin epäyhtälö antaa *absoluuttisen ylärajan* satunnaismuuttujan  $X$  todennäköisyysjakauman ”häntäalueiden” todennäköisyysmassan osuudelle.
- Jos satunnaismuuttujan  $X$  jakauma spesifioidaan tarkemmin, ”häntäalueiden” todennäköisyysmassan osuudesta voidaan antaa tarkempia arvioita.

## Markovin ja Tshebyshevin epäyhtälöt

# Tshebyshevin epäyhtälö: Kommentteja 3/3

---

- Esimerkki:

Tshebyshevin epäyhtälön mukaan *kaikille satunnaismuuttujille*  $X$ , joilla on odotusarvo  $E(X) = \mu$  ja varianssi  $\text{Var}(X) = \sigma^2$ , pätee

$$\Pr(|X - \mu| \geq 3\sigma) \leq \frac{1}{9}$$

Jos tiedämme, että  $X$  noudattaa *normaalijakaumaa* (ks. lukua **Jatkuvia jakaumia**), saadaan (esimerkiksi normaalijakaumien taulukoiden avulla) tarkempi tulos:

$$\Pr(|X - \mu| \geq 3\sigma) \approx 0.3 \%$$

# Jakaumien tunnusluvut

---

**Odotusarvo**

**Varianssi**

**Markovin ja Tshebyshevin epäyhtälöt**

**>> Momentit**

**Vinous ja huipukkuus**

**Kvantiilit**

**Moodi**

**Suurten lukujen laki**

## Momentit

# Origomomentit

---

- Olkoon  $X$  satunnaismuuttuja.
- Tällöin satunnaismuuttujan  $X^k$  odotusarvo

$$E(X^k) = \alpha_k$$

on satunnaismuuttujan  $X$   **$k$ . momentti** eli  $k$ . momentti *origon suhteen*.

## Momentit

# Origomomentit: Erikoistapauksia

---

- Olkoon

$$E(X^k) = \alpha_k$$

satunnaismuuttujan  $X$   **$k$ . momentti**

- Erityisesti:

$$\alpha_0 = 1$$

$$\alpha_1 = E(X) = \mu$$

- Siten satunnaismuuttujan  $X$  1. momentti on satunnaismuuttujan  $X$  *odotusarvo*.

## Keskusmomentit

---

- Olkoon  $X$  satunnaismuuttuja, jonka odotusarvo on

$$E(X) = \mu$$

- Tällöin satunnaismuuttujan  $(X - \mu)^k$  odotusarvo

$$E[(X - \mu)^k] = \mu_k$$

on satunnaismuuttujan  $X$   **$k$ . keskusmomentti** eli  $k$ . momentti *painopisteen  $\mu$  suhteen*.

## Momentit

# Keskusmomentit: Erikoistapauksia

---

- Olkoon

$$E[(X - \mu)^k] = \mu_k$$

satunnaismuuttujan  $X$   **$k$ . keskusmomentti** eli  $k$ . momentti *painopisteen  $\mu$  suhteen*.

- Erityisesti:

$$\mu_1 = 0$$

$$\mu_2 = E[(X - \mu)^2] = \sigma^2 = \text{Var}(X) = D^2(X)$$

- Siten satunnaismuuttujan  $X$  1. keskusmomentti *häviää aina* ja 2. keskusmomentti on satunnaismuuttujan  $X$  *varianssi*.

## Momenttien olemassaolo

---

- Satunnaismuuttujan  $X$   $k$ . **origomomentti on olemassa**, jos

$$E(|X|^k) < \infty$$

- Satunnaismuuttujan  $X$   $k$ . **keskusmomentti on olemassa**, jos vastaava origomomentti on olemassa.

- Voidaan osoittaa, että jos

$$E(|X|^n) < \infty$$

jollekin  $n \in \mathbb{N}$ , niin

$$E(|X|^k) < \infty \text{ kaikille } k < n$$

- Jos siis satunnaismuuttujalla on  $n$ . origomomentti, sillä on myös kaikki alempien kertalukujen momentit.



# Jakaumien tunnusluvut

---

**Odotusarvo**

**Varianssi**

**Markovin ja Tshebyshevin epäyhtälöt**

**Momentit**

**>> Vinous ja huipukkuus**

**Kvantiilit**

**Moodi**

**Suurten lukujen laki**

## Vinous ja huipukkuus

# Momentit

---

- Olkoon on

$$\alpha_k = E(X^k), k = 1, 2, 3, \dots, K$$

satunnaismuuttujan  $X$   **$k$ . origomomentti.**

- Olkoon

$$\mu_k = E[(X - \alpha_1)^k], k = 1, 2, 3, \dots, K$$

satunnaismuuttujan  $X$   **$k$ . keskusmomentti.**

- Huomaa:

$$\alpha_1 = E(X) = \mu$$

$$\mu_2 = E[(X - \mu)^2] = \text{Var}(X)$$

## Vinous ja huipukkuus

# Vinous

---

- Tunnuslukua

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}}$$

käytetään *todennäköisyysjakaumien* **vinouden** *mittana*.

## Todennäköisyysjakaumien vinous

---

- Jos todennäköisyysjakauman pistetodennäköisyys- tai tiheysfunktio on *yksihuippuinen*, pätee seuraava:
  - $\gamma_1 < 0$ : Jakauma on **negatiivisesti vino** eli **vino vasemmalle**, jolloin jakauman vasen häntä on pitempi kuin oikea häntä.
  - $\gamma_1 = 0$ : Jakauma on **symmetrinen**.
  - $\gamma_1 > 0$ : Jakauma on **positiivisesti vino** eli **vino oikealle**, jolloin jakauman oikea häntä on pitempi kuin vasen häntä.
- Huomautus:

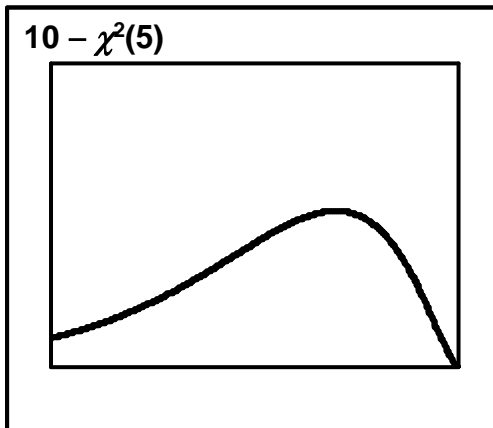
*Normaalijakaumalle  $\gamma_1 = 0$ .*

Vinous ja huipukkuus

## Todennäköisyysjakaumien vinous: Havainnollistus

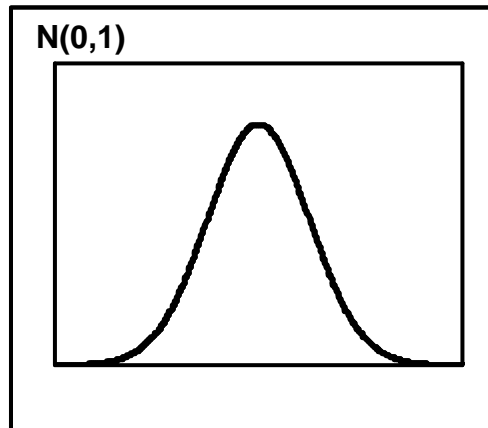
---

Alla on kuvattuna kolme *yksihuippuista* tiheysfunktiota.



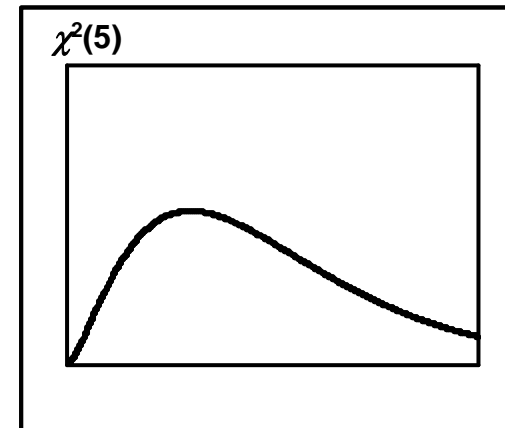
$$\gamma_1 < 0:$$

Jakauma on  
**negatiivisesti vino**  
eli **vino vasemmalle.**



$$\gamma_1 = 0:$$

Jakauma on  
**symmetrinen.**



$$\gamma_1 > 0:$$

Jakauma on  
**positiivisesti vino**  
eli **vino oikealle.**

## Vinous ja huipukkuus

# Huipukkuus

---

- Tunnuslukua

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3$$

käytetään *todennäköisyysjakaumien* **huipukkuuden** *mittana*.

## Todennäköisyysjakaumien huipukkuus

---

- Jos todennäköisyysjakauman pistetodennäköisyys- tai tiheysfunktio on *yksihuippuinen*, pätee seuraava:
  - $\gamma_2 > 0$ : Jakauma on **huipukas** (normaalijakaumaan verrattuna).
  - $\gamma_2 < 0$ : Jakauma on **laakea** (normaalijakaumaan verrattuna).
- Huomautus:
  - Normaalijakaumalle  $\gamma_2 = 0$ .*

# Jakaumien tunnusluvut

---

**Odotusarvo**

**Varianssi**

**Markovin ja Tshebyshevin epäyhtälöt**

**Momentit**

**Vinous ja huipukkuus**

**>> Kvanttiilit**

**Moodi**

**Suurten lukujen laki**



## Kvantiilin määritelmä

---

- Olkoon  $X$  satunnaismuuttuja.

- Olkoon lisäksi

$$0 < p < 1$$

- Jos luku  $x_p$  toteuttaa ehdot

$$\Pr(X \leq x_p) \geq p$$

$$\Pr(X \geq x_p) \geq 1 - p$$

sanomme, että  $x_p$  on satunnaismuuttujan  $X$  ja sen jakauman **kvantiili** kertalukua  $p$ .

- Kvantiili  $x_p$  toteuttaa siis epäyhtälöt

$$\Pr(X < x_p) \leq p \leq \Pr(X \leq x_p)$$

## Kvantiilit

# Kvantiilin määritelmä:

## Kommentteja

---

- Kvantiilit voidaan määrätä myös sellaisille satunnaisuuttujille, joilla *ei ole* momenteja.
- Kvantiilit *eivät välttämättä ole yksikäsitteisiä*:
  - (i) *Diskreettien* satunnaisuuttujien kvantiilit *ovat usein monikäsitteisiä*.
  - (ii) *Jatkuvien* satunnaisuuttujien kvantiilit *ovat yksikäsitteisiä*; ks. seuraavaa kalvoa.

## Jatkuvan satunnaismuuttujan kvantiilit

---

- Olkoon

$$F(x) = \Pr(X \leq x)$$

*jatkuvan satunnaismuuttujan  $X$  kertymäfunktio.*

- Tällöin satunnaismuuttujan  $X$  kvantiili  $x_p$  toteuttaa yhtälön

$$F(x_p) = p$$

- Kvantiili  $x_p$  jakaa satunnaismuuttujan  $X$  jakauman todennäköisyysmassan *kahteen osaan* niin, että massasta

$$p \times 100 \%$$

on kvantiilista  $x_p$  *vasemmalla* ja

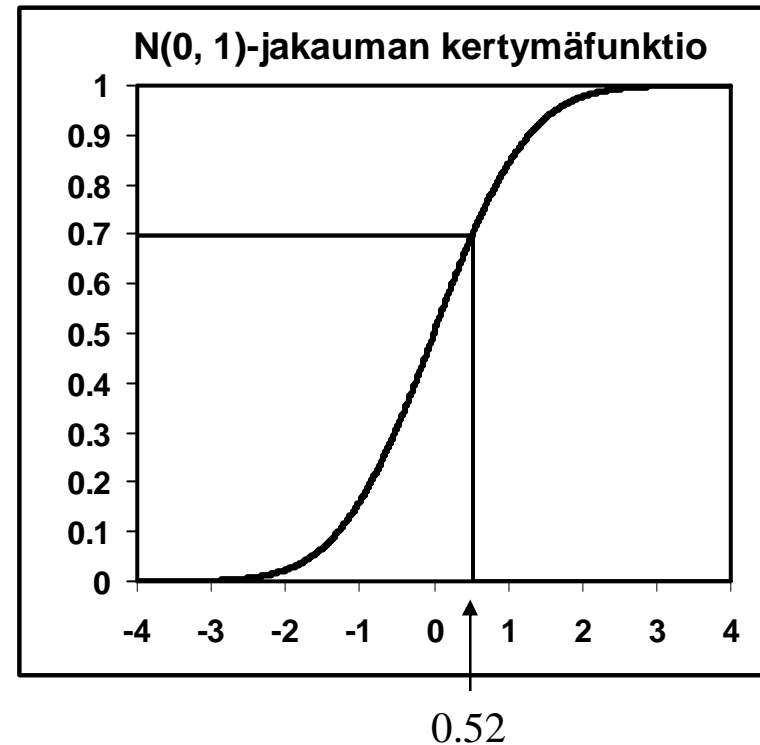
$$(1 - p) \times 100 \%$$

on kvantiilista  $x_p$  *oikealla*.

## Kvantiilit

# Jatkuvan satunnaismuuttujan kvantiilit: Esimerkki 1/2

- Kuva oikealla esittää **standardoidun normaalijakauman**  $N(0, 1)$  *kertymäfunktio*  $\Phi(z)$ .
- Standardoidun normaalijakauman  $N(0, 1)$  taulukoiden mukaan:  
$$\Phi(0.52) = \Pr(Z \leq 0.52) \approx 0.7$$
- Siten  
$$x_{0.7} \approx 0.52$$



## Kvantiilit

# Jatkuvan satunnaismuuttujan kvantiilit: Esimerkki 2/2

- Kuva oikealla esittää **standardoidun normaalijakauman**  $N(0, 1)$  *tiheysfunktio*.
- Standardoidun normaalijakauman  $N(0, 1)$  taulukoiden mukaan:

Alueen A pinta-ala

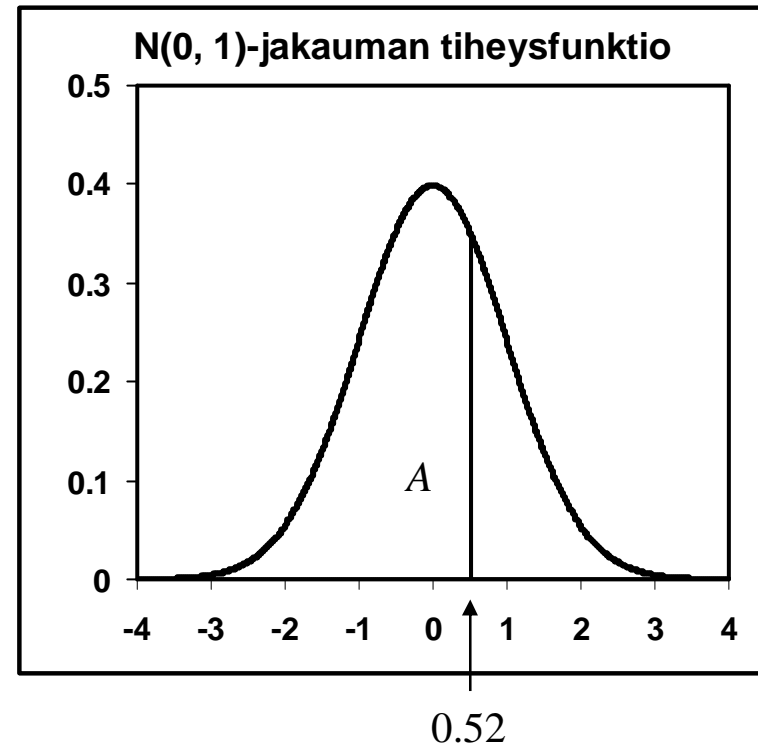
$$= \int_{-\infty}^{0.52} f_Z(z) dz$$

$$= \Pr(Z \leq 0.52)$$

$$\approx 0.7$$

- Siten

$$x_{0.7} \approx 0.52$$



## Kvantiilit ja tilastolliset taulukot

---

- Useimmissa *todennäköisyyslaskennan ja tilastotieteen oppikirjoissa on taulukoituna keskeisten tilastollisessa päättelyssä käytettävien jatkuvien jakaumien (standardoitu normaalijakauma  $N(0,1)$ ,  $\chi^2$ -,  $t$ - ja  $F$ -jakaumat) kvantiileja  $x_p$  ja niitä vastaavia todennäköisyyksiä  $p$ .*
- Useimmissa *tilastollisissa tietokoneohjelmissa on aliohjelmia*, jotka laskevat tavallisimpien jatkuvien jakaumien kvantiileja  $x_p$  ja niitä vastaavia todennäköisyyksiä  $p$ .
- Lisätietoja: ks. lukua **Normaalijakaumasta johdettuja jakaumia**.

## Prosenttipisteet

---

- Jos  $p$  on muotoa

$$p = q/100, q = 1, 2, \dots, 99$$

kvantiilia  $x_p$  kutsutaan  **$q$ . prosenttipisteeksi**.

- Jatkuvan satunnaismuuttujan tapauksessa  $q$ . prosenttipiste jakaa jakauman todennäköisyysmassan *kahteen osaan* niin, että massasta

$$q \%$$

on  $q$ . prosenttipisteestä *vasemmalla* ja

$$(100 - q) \%$$

on  $q$ . prosenttipisteestä *oikealla*.

## Kvantiilit

# Desiilit

---

- Jos  $p$  on muotoa

$$p = 10 \times q / 100, \quad q = 1, 2, \dots, 9$$

kvantiilia  $x_p$  kutsutaan  **$q$ . desiiliksi**.

- Jatkuvan satunnaismuuttujan tapauksessa  $q$ . desiili jakaa jakauman todennäköisyysmassan *kahteen osaan* niin, että massasta

$$10 \times q \%$$

on  $q$ . desiilistä *vasemmalla* ja

$$(100 - 10 \times q) \%$$

on  $q$ . desiilistä *oikealla*.



## Kvartiilit 1/2

---

- Jos  $p$  on muotoa

$$p = 25 \times q / 100, \quad q = 1, 2, 3$$

kvantiilia  $x_p$  kutsutaan  **$q$ . kvartiiliksi**.

- Jatkuvan satunnaismuuttujan tapauksessa  $q$ . kvartiili jakaa satunnaismuuttujan  $X$  jakauman todennäköisyysmassan *kahteen osaan* niin, että massasta

$$25 \times q \%$$

on  $q$ . kvartiilista *vasemmalla* ja

$$(100 - 25 \times q) \%$$

on  $q$ . kvartiilista *oikealla*.

## Kvantiilit 2/2

---

- Kvartiileja merkitään tavallisesti symboleilla  $Q_1$ ,  $Q_2$ ,  $Q_3$  ja sanotaan, että

$$Q_1 = \text{alakovartiili}$$

$$Q_2 = \text{keskikovartiili}$$

$$Q_3 = \text{ylakovartiili}$$

- Jatkuvan satunnaismuuttujan tapauksessa kvantiilit jakavat jakauman todennäköisyysmassan *neljään yhtä suureen osaan*:

25 % massasta on kvartiilista  $Q_1$  vasemmalle

25 % massasta on kvartiilien  $Q_1$  ja  $Q_2$  välissä

25 % massasta on kvartiilien  $Q_2$  ja  $Q_3$  välissä

25 % massasta on kvartiilista  $Q_3$  oikealle

## Kvantiilit

# Mediaani

---

- Jos

$$p = 0.5$$

kvantiilia  $x_p$  kutsutaan **mediaaniksi**.

- Mediaania merkitään tavallisesti symbolilla  $Me$ .
- Jatkuvan satunnaismuuttujan tapauksessa mediaani  $Me$  jakaa jakauman todennäköisyysmassan *kahteen yhtä suureen osaan* niin, että massasta

50 %

on mediaanista *vasemmalla* ja

50 %

on mediaanista *oikealla*.

- Jakauman mediaani *ei välttämättä ole* yksikäsitteinen.
- Jakauman mediaani yhtyy jakauman 50. prosentti-pisteeseen, 5. desiiliin ja keskikvartiiliin  $Q_2$ .
- Mediaani voidaan määrätä myös sellaisille satunnaismuuttujille, joilla *ei ole* odotusarvoa.
- Jos satunnaismuuttujan  $X$  jakauma on *symmetrinen* suoran  $x = a$  suhteen, niin jakauman mediaani yhtyy pisteeseen  $a$ :

$$Me = a$$

- Jos symmetrisellä jakaumalla *on odotusarvo*  $E(X) = \mu$ , niin jakauman mediaani yhtyy pisteeseen  $\mu$ :

$$Me = \mu$$

## Kvantilit

# Mediaani: Esimerkki

- Kuva oikealla esittää eksponenttijakauman  $\text{Exp}(1)$  tiheysfunktiota

$$f(x) = e^{-x}, x \geq 0$$

välillä  $[0, 4]$ .

- Jakauman *mediaani* saadaan ratkaisemalla yhtälö

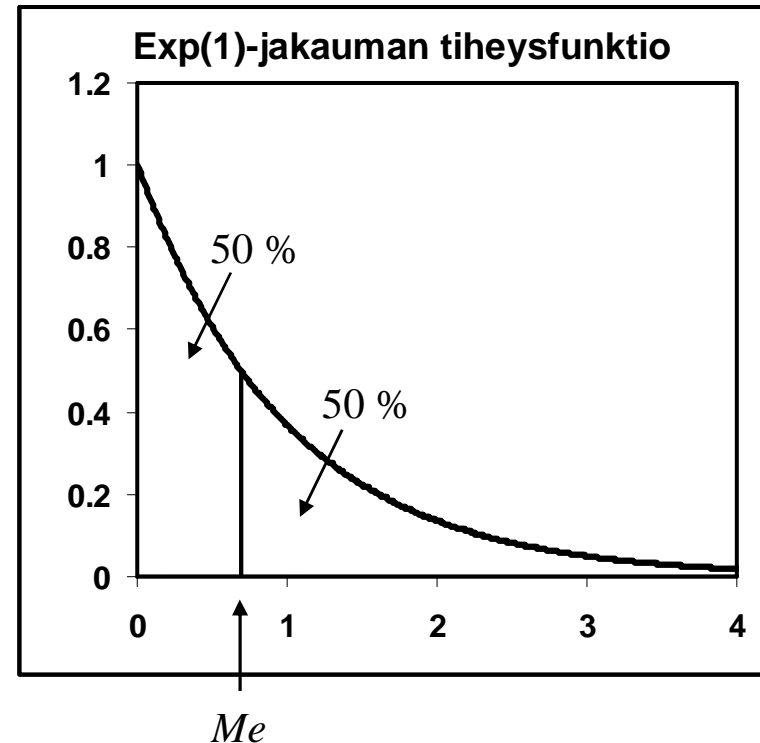
$$\int_0^x e^{-t} dt = \left[ -e^{-t} \right]_0^x$$

$$= 1 - e^{-x} = 0.5$$

$x$ :n suhteen.

- Siten

$$Me = x = \log(2) \approx 0.69$$



# Jakaumien tunnusluvut

---

**Odotusarvo**

**Varianssi**

**Markovin ja Tshebyshevin epäyhtälöt**

**Momentit**

**Vinous ja huipukkuus**

**Kvantiilit**

**>> Moodi**

**Suurten lukujen laki**

## Diskreetin satunnaismuuttujan moodi

---

- Olkoon  $X$  *diskreetti* satunnaismuuttuja, jonka *pistetodennäköisyysfunktio* on

$$f(x) = \Pr(X = x)$$

- Piste  $Mo$  on *diskreetin satunnaismuuttujan*  $X$  ja sen jakauman **moodi**, jos pistetodennäköisyysfunktio  $f(x)$  saavuttaa maksiminsa pisteessä  $x = Mo$ :

$$f(Mo) = \max_x f(x)$$

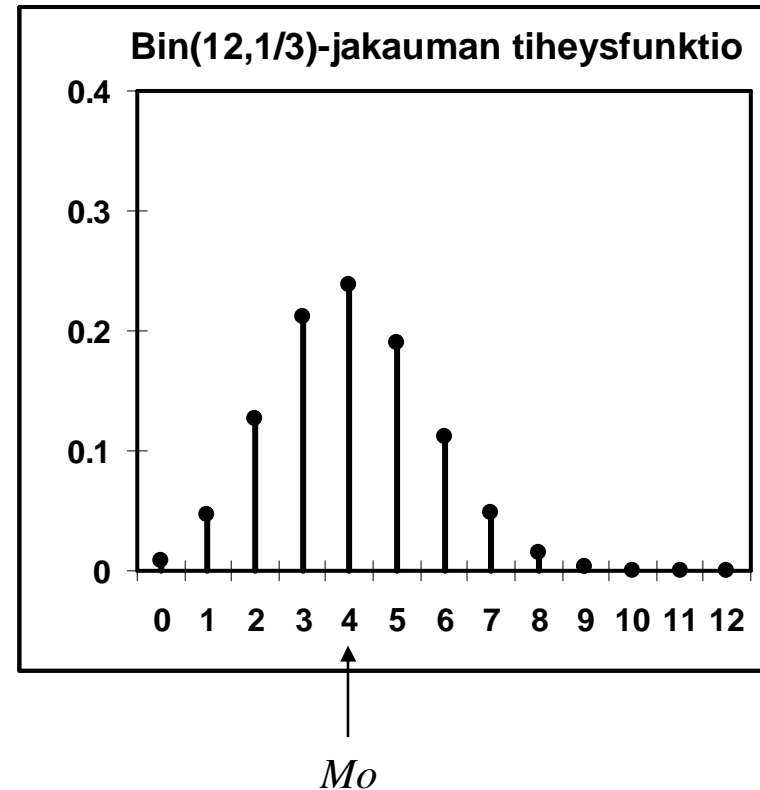
## Moodi

# Diskreetin satunnaismuuttujan moodi: Esimerkki

- Kuva oikealla esittää **binomijakauman**  $\text{Bin}(12, 1/3)$  *pistetodennäköisyysfunktiota*

$$f(x) = \binom{12}{x} \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{12-x}$$

- Jakauman *moodi*  $M_o$  on pisteessä  $x = 4$





## Jatkuvan satunnaismuuttujan moodi

---

- Olkoon  $X$  *jatkuva* satunnaismuuttuja, jonka *tiheysfunktio* on

$$f(x)$$

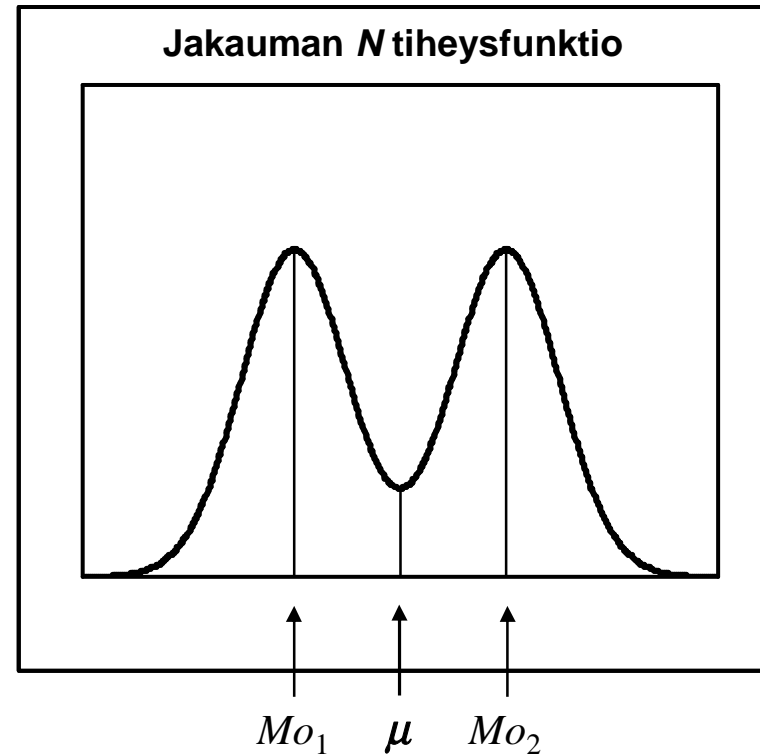
- Piste  $M_0$  on *jatkuvan satunnaismuuttujan*  $X$  ja sen jakauman **moodi**, jos tiheysfunktio  $f(x)$  saavuttaa maksiminsa pisteessä  $x = M_0$ :

$$f(M_0) = \max_x f(x)$$

## Moodi

# Jatkuvan satunnaismuuttujan moodi: Esimerkki

- Kuva oikealla esittää erään **sekoitetun normaalijakauman**  $N$  tiheysfunktiota  $f$ .
- Tiheysfunktio  $f$  on *kaksi-huippuinen* ja *symmetrinen* suoran  $x = \mu$  suhteen.
- Jakaumalla  $N$  on *kaksi lokaalia moodia*  $Mo_1$  ja  $Mo_2$ .



## Moodi

# Satunnaismuuttujan moodi: Kommentteja

---

- Jakauman moodi *ei välttämättä ole* yksikäsitteinen; ks. edellistä kalvoa.
- Moodi voidaan määrätä myös sellaisille satunnaismuuttujille, joilla *ei ole* odotusarvoa.

# Jakaumien tunnusluvut

---

**Odotusarvo**

**Varianssi**

**Markovin ja Tshebyshevin epäyhtälöt**

**Momentit**

**Vinous ja huipukkuus**

**Kvantiilit**

**Moodi**

**>> Suurten lukujen laki**

## Suurten lukujen laki: Formulointi

---

- Olkoon  $X_i$ ,  $i = 1, 2, 3, \dots$  jono *riippumattomia* satunnaismuuttujia, joilla on *sama* odotusarvo ja varianssi:

$$E(X_i) = \mu, D^2(X_i) = \sigma^2, i = 1, 2, 3, \dots$$

- Määritellään satunnaismuuttujien  $X_i$ ,  $i = 1, 2, \dots, n$  *aritmeettinen keskiarvo*:

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

- Tällöin pätee (heikko) **suurten lukujen laki**:

$$\lim_{n \rightarrow +\infty} \Pr(|\bar{X}_n - \mu| > \varepsilon) = 0$$

## Suurten lukujen laki: Kommentteja 1/2

---

- Suurten lukujen laille esitetään *todistus* luvussa **Stokastiikan konvergenssikäsitteet ja raja-arvolauseet**.
- *Suurten lukujen laki* ilmaistaan usein sanoin seuraavasti:  
Samoin jakautuneiden satunnaismuuttujien *aritmeettinen keskiarvo lähestyy muuttujien lukumäärän kasvaessa muuttujien yhteistä odotusarvoa* sellaisella tavalla, että **poikkeamien todennäköisyys satunnaismuuttujien yhteisestä odotusarvosta lähestyy lukua nolla eli poikkeamat tulevat yhä harvinaisemmiksi**.
- Suurten lukujen lakia voidaan pitää matemaattisena formulointina **tilastollisen stabiliteetin** käsitteelle.

## Suurten lukujen laki: Kommentteja 2/2

---

- Tässä formuloitua suurten lukujen lakia kutsutaan *heikoksi suurten lukujen laiksi*.
- Suurten lukujen laki koskee satunnaismuuttujien **asymptoottista käyttäytymistä** samaan tapaan kuin luvussa Jatkuvia jakaumia esitettävä **keskeinen raja-arvolause**.
- Suurten lukujen laissa esiintyvä *rajakäyttäytymisen muoto* on esimerkki **stokastiikan konvergenssikäsitteistä**; ks. lukua **Stokastiikan konvergenssikäsitteet ja raja-arvolauseet**.
- Suurten lukujen laista on olemassa *yleisempiä muotoja*, joissa voidaan lieventää *samoinjakautuneisuus-* ja *riippumattomuusoletuksia*.