
Ilkka Mellin

Todennäköisyyslaskenta

Osa 3: Todennäköisyysjakaumia

Jatkuvia jakaumia

Jatkuvia jakaumia

>> **Jatkuva tasainen jakauma**

Eksponenttijakauma

Normaalijakauma

Keskeinen raja-arvolause

Log-normaali-, Cauchy-, Gamma-, Beta- ja Weibull-jakaumat

Jatkuva tasainen jakauma

Jatkuva tasainen jakauma ja sen tiheysfunktio 1/3

- Olkoon satunnaismuuttujan X arvoalueena *reaaliakselin äärellinen väli* $[a, b]$.
- Olkoot $[c_1, d_1]$ ja $[c_2, d_2]$ välin $[a, b]$ kaksi *mielivaltaista, samanpituista* osaväliä:

$$[c_1, d_1] \subset [a, b]$$

$$[c_2, d_2] \subset [a, b]$$

$$d_1 - c_1 = d_2 - c_2$$

- Oletetaan, että väleihin $[c_1, d_1]$ ja $[c_2, d_2]$ liittyvät todennäköisyydet ovat *yhtä suuria*:

$$\Pr(X \in [c_1, d_1]) = \Pr(X \in [c_2, d_2])$$

Jatkuva tasainen jakauma

Jatkuva tasainen jakauma ja sen tiheysfunktio 2/3

- Satunnaismuuttujan X **tiheysfunktio** on

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

- Funktio $f(x)$ kelpaa tiheysfunktiksi, koska

$$\int_a^b f(x) dx = 1$$

Jatkuva tasainen jakauma

Jatkuva tasainen jakauma ja sen tiheysfunktio 3/3

- Sanomme, että satunnaismuuttuja X noudattaa **jatkuvaa tasaista jakaumaa parametrein a ja b .**
- Merkintä:

$$X \sim \text{Uniform}(a, b)$$

Jatkuva tasainen jakauma

Jatkuva tasainen jakauma ja sen kertymäfunktio

- Olkoon $X \sim \text{Uniform}(a, b)$.
- Satunnaismuuttujan X **kertymäfunktio** on

$$F(x) = \Pr(X \leq x) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x \geq b \end{cases}$$

Jatkuva tasainen jakauma

Jatkuva tasainen jakauma ja sen kertymäfunktio: Johto

- Olkoon $X \sim \text{Uniform}(a, b)$.
- Satunnaismuuttujan X tiheysfunktion lauseke, kun $x \in [a, b]$:

$$f(x) = \frac{1}{b-a}$$

- Siten satunnaismuuttujan X kertymäfunktion lausekkeeksi saadaan, kun $x \in [a, b]$:

$$\begin{aligned} F(x) = \Pr(X \leq x) &= \int_{-\infty}^x f(t) dt = \int_a^x \frac{1}{b-a} dt \\ &= \left[\frac{1}{b-a} \right]_a^x \\ &= \frac{x-a}{b-a} \end{aligned}$$

Jatkuva tasainen jakauma

Odotusarvo, varianssi ja standardipoikkeama

- Olkoon $X \sim \text{Uniform}(a, b)$.

- **Odotusarvo:**

$$E(X) = \frac{a+b}{2}$$

- **Varianssi ja standardipoikkeama:**

$$\text{Var}(X) = D^2(X) = \frac{(b-a)^2}{12}$$

$$D(X) = \frac{b-a}{2\sqrt{3}}$$

Jatkuva tasainen jakauma

Odotusarvon johto

- Olkoon

$$X \sim \text{Uniform}(a, b)$$

- Tällöin

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} xf(x)dx = \int_a^b x \frac{1}{b-a} dx \\ &= \frac{1}{b-a} \left[\frac{1}{2} x^2 \right]_a^b \\ &= \frac{b^2 - a^2}{2(b-a)} \\ &= \frac{a+b}{2} \end{aligned}$$

Jatkuva tasainen jakauma

Tiheysfunktion kuvaaja

- Kuva oikealla esittää jatkuvan tasaisen jakauman

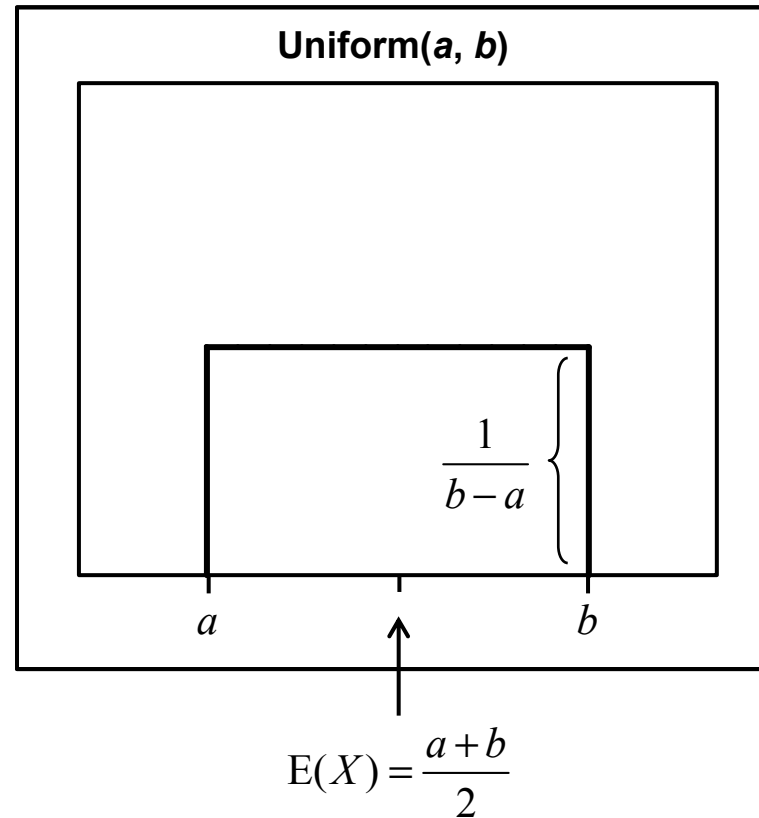
Uniform(a, b)

tiheysfunktioita

$$f(x) = \frac{1}{b-a}, a \leq x \leq b$$

- Jakauman *odotusarvo*:

$$E(X) = \frac{a+b}{2}$$



Jatkuva tasainen jakauma

Tiheysfunktion ja sen kuvaajan ominaisuudet

- Jatkuvan tasaisen jakauman tiheysfunktio

$$f(x) = \frac{1}{b-a}, a \leq x \leq b$$

saa positiivisen *vakioarvon* $1/(b - a)$ välillä $[a, b]$ ja saa arvon 0 välin $[a, b]$ ulkopuolella.

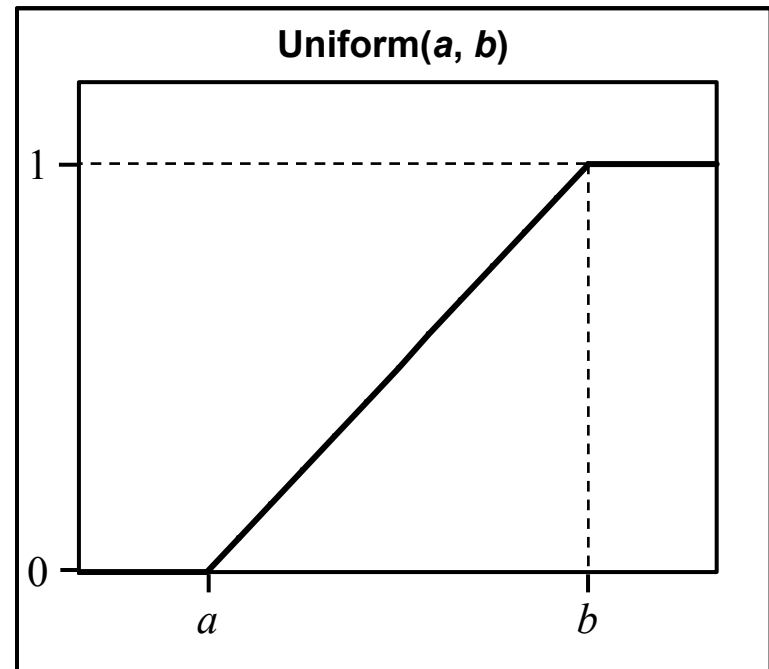
Jatkuva tasainen jakauma

Kertymäfunktion kuvaaja

- Kuva oikealla esittää jatkuvan tasaisen jakauman

Uniform(a, b)
kertymäfunktiota

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



Jatkuva tasainen jakauma

Todennäköisyyksien määrittäminen jatkevasta tasaisesta jakaumasta 1/2

- Olkoon $X \sim \text{Uniform}(a, b)$.

- Olkoon

$$[c, d] \subset [a, b]$$

jokin välin $[a, b]$ osaväli.

- Välin $[c, d]$ todennäköisyys saadaan integroimalla jatkuvan tasaisen jakauman $\text{Uniform}(a, b)$ tiheysfunktio

$$f(x) = \frac{1}{b-a}, a \leq x \leq b$$

välillä $[c, d]$:

$$\Pr(c \leq X \leq d) = \int_c^d f(x) dx = \frac{d-c}{b-a}$$

Jatkuva tasainen jakauma

Todennäköisyyksien määrittäminen jatkuvasta tasaisesta jakaumasta 2/2

- *Kaikkien* muiden jatkuvaan tasaiseen jakaumaan $\text{Uniform}(a, b)$ liittyvien tapahtumien todennäköisyydet saadaan välin $[a, b]$ osavälien todennäköisyyksistä *todennäköisyyslaskennan laskusääntöjen* avulla.

Jatkuvia jakaumia

Jatkuva tasainen jakauma

>> Eksponenttijakauma

Normaalijakauma

Keskeinen raja-arvolause

Log-normaali-, Cauchy-, Gamma-, Beta- ja Weibull-jakaumat

Eksponenttijakauma

Eksponenttijakauma ja sen tiheysfunktio

- Olkoon satunnaismuuttujan X **tiheysfunktio**

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

- Funktio $f(x)$ kelpaa tiheysfunktiksi, koska

$$\int_0^{+\infty} f(x) dx = 1$$

- Sanomme, että satunnaismuuttuja X noudattaa **eksponenttijakaumaa parametrinaan λ** .
- Merkintä:

$$X \sim \text{Exp}(\lambda)$$

Eksponenttijakauma

Eksponenttijakauma ja sen kertymäfunktio

- Olkoon $X \sim \text{Exp}(\lambda)$.
- Satunnaismuuttujan X **kertymäfunktio** on

$$F(x) = \Pr(X \leq x) = 1 - e^{-\lambda x}, \lambda > 0, x \geq 0$$

Eksponttijakauma

Eksponttijakauma ja sen kertymäfunktio: Johto

- Olkoon $X \sim \text{Exp}(\lambda)$.
- Satunnaismuuttujan X tiheysfunktio:

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

- Siten satunnaismuuttujan X kertymäfunktioksi saadaan, kun $x \geq 0$:

$$\begin{aligned} F(x) = \Pr(X \leq x) &= \int_{-\infty}^x f(t) dt = \int_0^x \lambda e^{-\lambda t} dt \\ &= \left[-e^{-\lambda t} \right]_0^x \\ &= 1 - e^{-\lambda x} \end{aligned}$$

Eksponenttijakauma

Odotusarvo, varianssi ja standardipoikkeama

- Olkoon $X \sim \text{Exp}(\lambda)$.

- **Odotusarvo:**

$$E(X) = \frac{1}{\lambda}$$

- **Varianssi ja standardipoikkeama:**

$$\text{Var}(X) = D^2(X) = \frac{1}{\lambda^2}$$

$$D(X) = \frac{1}{\lambda}$$

Eksponttijakauma

Odotusarvon johto

- Olkoon

$$X \sim \text{Exp}(\lambda)$$

- Tällöin *osittaisintegroinnilla* saadaan:

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

Eksponenttijakauma

Tiheysfunktion kuvaaja

- Kuva oikealla esittää eksponenttijakauman

$$\text{Exp}(\lambda)$$

tiheysfunktiota

$$f(x) = \lambda e^{-\lambda x}$$

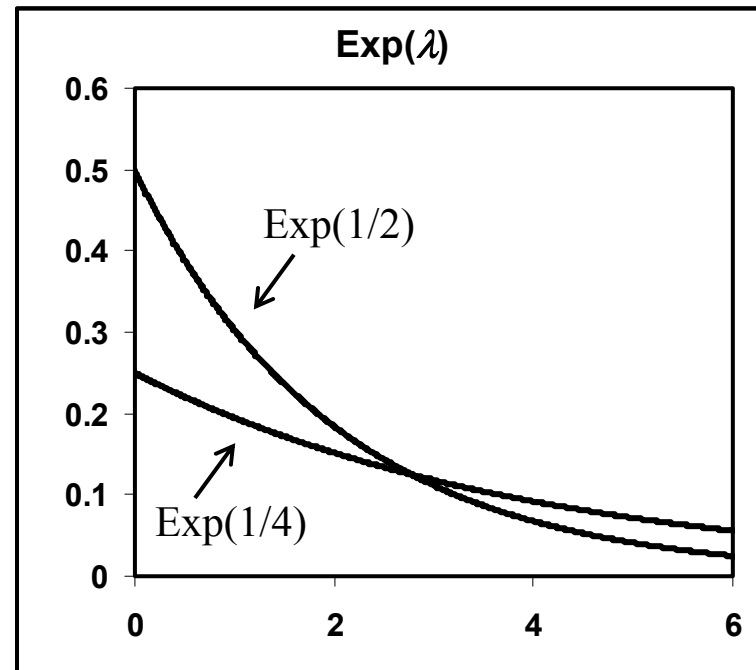
välillä $[0, 6]$, kun

(i) $\lambda = 1/2$

(ii) $\lambda = 1/4$

- Jakauman *odotusarvo*:

$$E(X) = 1/\lambda$$



Tiheysfunktion ja sen kuvaajan ominaisuudet

- Eksponenttijakauman tiheysfunktio

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

on *positiivinen* kaikille ei-negatiivisille argumentin arvoille:

$$f(x) > 0, x > 0$$

- Tiheysfunktiolla on *maksimi* pisteessä

$$x = 0$$

- Tiheysfunktio on *monotonisesti laskeva* kaikille $\lambda > 0$.

Eksponttijakauma

Kertymäfunktion kuvaaja

- Kuva oikealla esittää
eksponenttijakauman

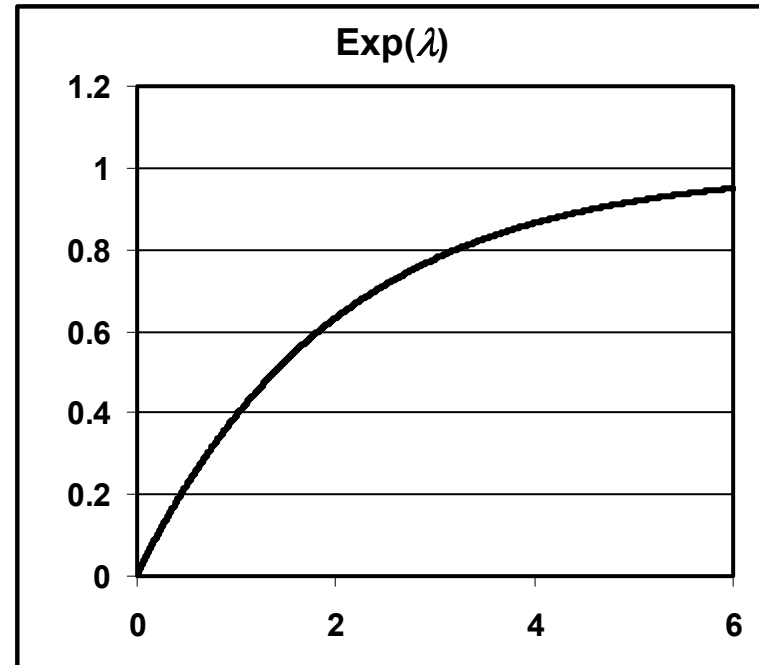
$$\text{Exp}(\lambda)$$

kertymäfunktiota

$$F(x) = 1 - e^{-\lambda x}$$

välillä $[0, 6]$, kun

$$\lambda = 1/2$$



Eksponenttijakauma

Poisson-prosessi

- Tarkastellaan jonkin tapahtuman A sattumista *jatkuvalla* aikavälillä, jonka pituus on t aikayksikköä.
- Määritellään *diskreetti satunnaismuuttuja*
 $Z =$ Niiden tapahtumien lukumäärä, jotka sattuvat aikavälillä $[0, t]$
- Sopivin oletuksin satunnaismuuttuja Z noudattaa **Poisson-jakaumaa parametrinaan νt :**
 $Z \sim \text{Poisson}(\nu t)$
- Parametri νt kuvaa *tapahtumaintensiteettiä* eli *tapahtumien keskimääräistä lukumäärää aikavälillä, jonka pituus on t aikayksikköä.*

Eksponenttijakauma

Poisson-prosessi ja eksponenttijakauma

- Olkoon

$$Z \sim \text{Poisson}(vt)$$

- Määritellään *jatkuva satunnaismuuttuja*

$$\begin{aligned} X &= \text{Ensimmäisen tapahtuman sattumisaika} \\ &= \text{Tapahtumien väliaika} \end{aligned}$$

- Satunnaismuuttuja X noudattaa **eksponenttijakaumaa** parametrinaa ν :

$$X \sim \text{Exp}(\nu)$$

Eksponttijakauman tiheysfunktion johto 1/2

- Olkoon

$$X \sim \text{Exp}(\lambda)$$

- Johdetaan ensin satunnaismuuttujan X kertymäfunktio F_X .
- Kertymäfunktion määritelmän ja komplementtitodennäköisyyden kaavan mukaan kaikille kertymäfunktioille F_X pätee:
(*) $F_X(x) = \Pr(X \leq x) = 1 - \Pr(X > x)$
- Ensimmäinen tapahtuma sattuu ajanhetken x jälkeen, jos ja vain jos aikavälillä $[0, x]$ *ei ole sattunut yhtään tapahtumaa*.
- Siten

$$\Pr(X > x) = \Pr(Z = 0)$$

jossa $Z \sim \text{Poisson}(\lambda x)$.

- Poisson-jakauman pistetodennäköisyysfunktion kaavasta saadaan:

$$\Pr(X > x) = \Pr(Z = 0) = \exp(-\lambda x)$$

Eksponenttijakauma

Eksponenttijakauman tiheysfunktion johto 2/2

- Sijoittamalla tämä satunnaismuuttujan X lausekkeeseen (*) kalvolla 1/2 saadaan

$$F_X(x) = \Pr(X \leq x) = 1 - \Pr(X > x) = 1 - \exp(-\lambda x)$$

josta satunnaismuuttujan X tiheysfunktioksi saadaan derivoimalla

$$f_X(x) = \frac{d}{dx} F_X(x) = \lambda \exp(-\lambda x)$$

- Siten

$$X \sim \text{Exp}(\lambda)$$

Eksponenttijakauman unohtamisominaisuus

- Olkoon $X \sim \text{Exp}(\lambda)$.

- Tällöin

$$\Pr(X \geq a + b | X \geq a) = \Pr(X \geq b)$$

- Siten eksponenttijakaumalla on seuraava *unohtamisominaisuus*:

Se, että tapahtuman sattumista on jouduttu odottamaan ajan a , *ei vaikuta* todennäköisyyteen joutua odottamaan ajan b lisää.

- Poisson-prosessin unohtamisominaisuutta on esimerkki stokastisen prosessin *Markov-ominaisuudesta*.

Eksponenttijakauma

Todennäköisyyksien määrittäminen eksponenttijakaumasta 1/2

- Olkoon $X \sim \text{Exp}(\lambda)$.

- Olkoon

$$[c, d] \subset [0, +\infty)$$

jokin välin $[0, +\infty)$ osaväli.

- Välin $[c, d]$ todennäköisyys saadaan integroimalla eksponenttijakauman $\text{Exp}(\lambda)$ tiheysfunktio

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

välillä $[c, d]$:

$$\Pr(c \leq X \leq d) = \int_c^d f(x) dx = e^{-\lambda c} - e^{-\lambda d}$$

Todennäköisyyksien määrittäminen eksponenttijakaumasta 2/2

- *Kaikkien* muiden eksponenttijakaumaan $\text{Exp}(\lambda)$ liittyvien tapahtumien todennäköisyydet saadaan välin $[a, b]$ osavälien todennäköisyyksistä *todennäköisyyslaskennan laskusääntöjen* avulla.

Jatkuvia jakaumia

Jatkuva tasainen jakauma

Eksponenttijakauma

>> Normaalijakauma

Keskeinen raja-arvolause

Log-normaali-, Cauchy-, Gamma-, Beta- ja Weibull-jakaumat

Normaalijakauma

Normaalijakauma ja sen tiheysfunktio 1/2

- Olkoon satunnaismuuttujan X **tiheysfunktio**

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}, -\infty < \mu < +\infty, \sigma > 0$$

$$-\infty < x < +\infty$$

- Funktio $f(x)$ kelpaa tiheysfunktiksi, koska

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

- Sanomme, että satunnaismuuttuja X noudattaa **normaalijakaumaa parametrein μ ja σ^2** .
- Merkintä:

$$X \sim N(\mu, \sigma^2)$$

Normaalijakauma

Normaalijakauma ja sen tiheysfunktio 2/2

- Normaalijakaumaa kutsutaan kehittäjänsä mukaan usein *Gaussin jakaumaksi* ja sen tiheysfunktion kuvaajaa *Gaussin käyräksi* tai *kellokäyräksi* (engl. *bell curve*).

Normaalijakauma

Normaalijakauma ja sen kertymäfunktio

- Olkoon $X \sim N(\mu, \sigma^2)$.
- Satunnaismuuttujan X **kertymäfunktio** on

$$F(x) = \Pr(X \leq x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

- Koska *normaalijakauman tiheysfunktion integraalifunktiota ei osata esittää alkeisfunktioiden avulla suljetussa muodossa*, niin **normaalijakauman kertymäfunktiole ei voida antaa eksplisiittistä lauseketta**.
- Siten normaalijakauman kertymäfunktion arvojen määräämiseen on käytettävä *numeerista integrointia*.

Normaalijakauma

Odotusarvo, varianssi ja standardipoikkeama

- Olkoon $X \sim N(\mu, \sigma^2)$.

- **Odotusarvo:**

$$E(X) = \mu$$

- **Varianssi ja standardipoikkeama:**

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

$$D(X) = \sigma$$

Normaalijakauma

Odotusarvon johto 1/2

- Olkoon

$$X \sim N(\mu, \sigma^2)$$

- Tällöin

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} xe^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

- *Sijoituksella*

$$z = \frac{x - \mu}{\sigma}$$

saadaan:

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} (\mu + \sigma z) e^{-\frac{1}{2}z^2} dz \\ &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \mu e^{-\frac{1}{2}z^2} dz + \int_{-\infty}^{+\infty} \frac{\sigma}{\sqrt{2\pi}} ze^{-\frac{1}{2}z^2} dz \end{aligned}$$

Normaalijakauma

Odotusarvon johto 2/2

- Nyt

$$E(X) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \mu e^{-\frac{1}{2}z^2} dz + \int_{-\infty}^{+\infty} \frac{\sigma}{\sqrt{2\pi}} z e^{-\frac{1}{2}z^2} dz = \mu$$

- Perustelu:

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \mu e^{-\frac{1}{2}z^2} dz = \mu \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = \mu \cdot 1 = \mu$$

koska integroitava on standardoidun normaalijakauman $N(0,1)$ tiheysfunktio ja

$$\int_{-\infty}^{+\infty} \frac{\sigma}{\sqrt{2\pi}} z e^{-\frac{1}{2}z^2} dz = 0$$

koska integroitava on muuttujan z *pariton funktio*.

Normaalijakauma

Tiheysfunktion kuvaaja

- Kuva oikealla esittää normaalijakauman

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

tiheysfunktiota

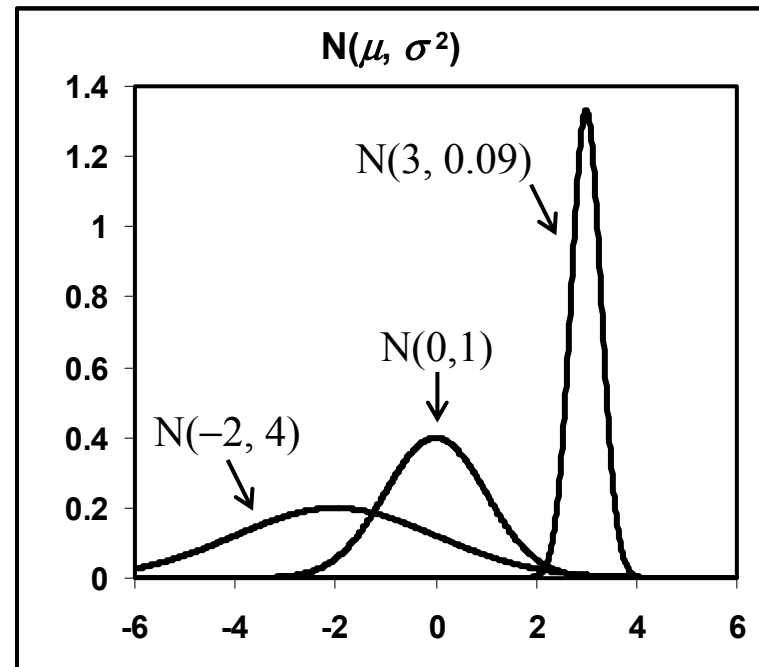
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(x - \mu)^2\right\}$$

välillä $[-6, +6]$, kun

- (i) $\mu = -2$ $\sigma^2 = 4$
- (ii) $\mu = 0$ $\sigma^2 = 1$
- (iii) $\mu = +3$ $\sigma^2 = 0.09$

- Jakauman *odotusarvo*:

$$E(X) = \mu$$



Tiheysfunktion ja sen kuvaajan ominaisuuksia 1/3

- Normaalijakauman tiheysfunktio

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

on kaikkialla *positiivinen*:

$$f(x) > 0 \text{ kaikille } x$$

- Tiheysfunktio on *yksihuippuinen*.
- Tiheysfunktio saa *maksimiarvonsa* pisteessä μ .
- Tiheysfunktio on *symmetrinen* suoran $x = \mu$ suhteen:

$$f(\mu - x) = f(\mu + x) \text{ kaikille } x$$

Tiheysfunktion ja sen kuvaajan ominaisuuksia 2/3

- Tiheysfunktiolla on *käännepisteet* pisteissä

$$\mu - \sigma \text{ ja } \mu + \sigma$$

ja tiheysfunktio on *kupera ylöspäin* välin $[\mu - \sigma, \mu + \sigma]$ *sisäpuolella* ja *kupera alaspäin* välin $[\mu - \sigma, \mu + \sigma]$ *ulkopuolella*.

- Kaikki normaalijakaumat ovat *samanmuotoisia*, jos ne piirretään *standardoiduissa yksiköissä*

$$z = \frac{x - \mu}{\sigma}$$

Normaalijakauma

Tiheysfunktion ja sen kuvaajan ominaisuuksia 3/3

- Kuva oikealla esittää normaalijakauman

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

tiheysfunktiota

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

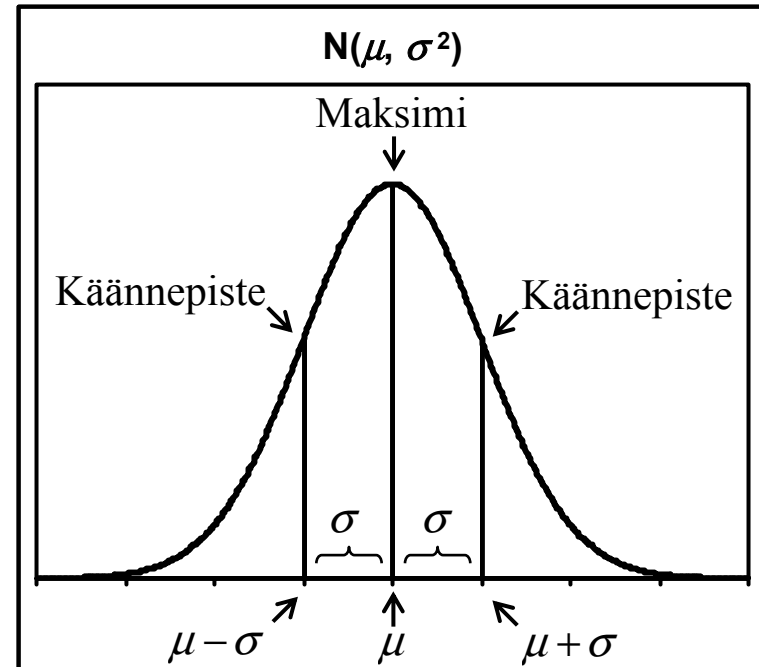
- Tiheysfunktiolla on *maksimi* pisteessä

$$x = \mu$$

- Tiheysfunktiolla on *käännepisteet* pisteissä

$$x = \mu - \sigma$$

$$x = \mu + \sigma$$



Normaalijakauma

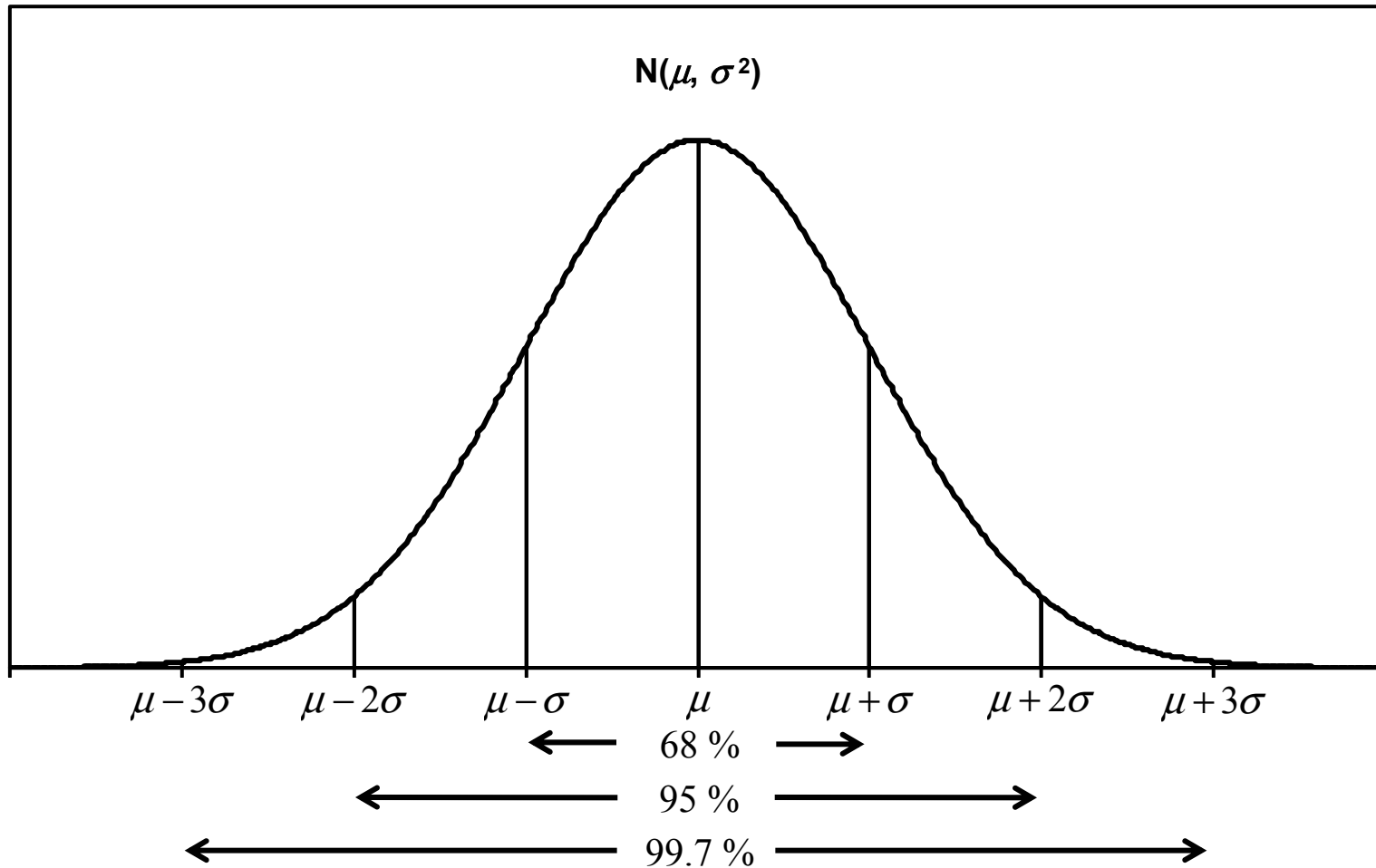
68-95-99.7-sääntö

- *Kaikille* normaalijakaumille pätee (likimäärin):
 - (i) 68% jakauman todennäköisyysmassasta on välillä
 $[\mu - \sigma, \mu + \sigma]$
 - (ii) 95% jakauman todennäköisyysmassasta on välillä
 $[\mu - 2\sigma, \mu + 2\sigma]$
 - (iii) 99.7% jakauman todennäköisyysmassasta on välillä
 $[\mu - 3\sigma, \mu + 3\sigma]$

Normaalijakauma

68-95-99.7-sääntö:

Havainnollistus



Normaalijakauma

Standardoitu normaalijakauma

- Olkoon

$$X \sim N(0,1)$$

jolloin siis

$$E(X) = 0$$

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

- Tällöin sanomme, että satunnaismuuttuja X noudattaa **standardoitua normaalijakaumaa**.

Normaalijakauma

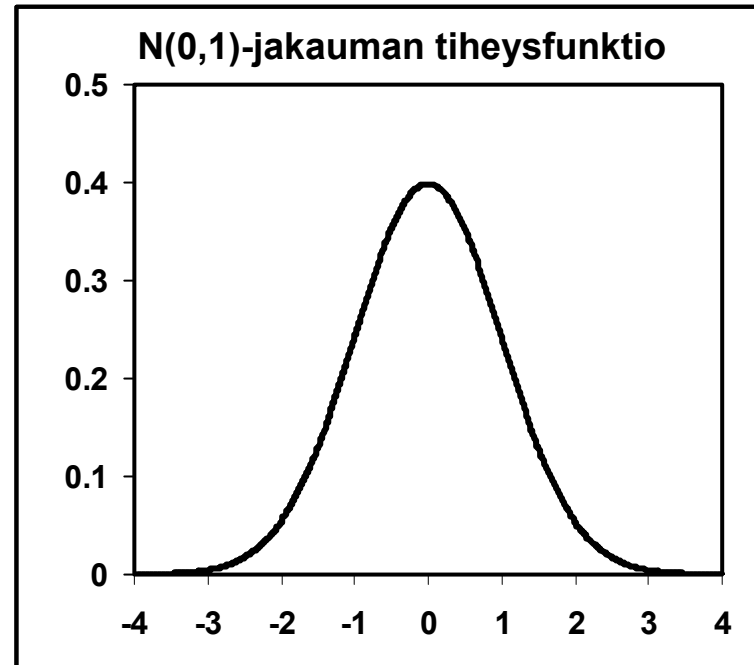
Standardoitu normaalijakauma: Tiheysfunktion kuvaaja

- Kuva oikealla esittää standardoidun normaalijakauman

$N(0,1)$

tiheysfunktio

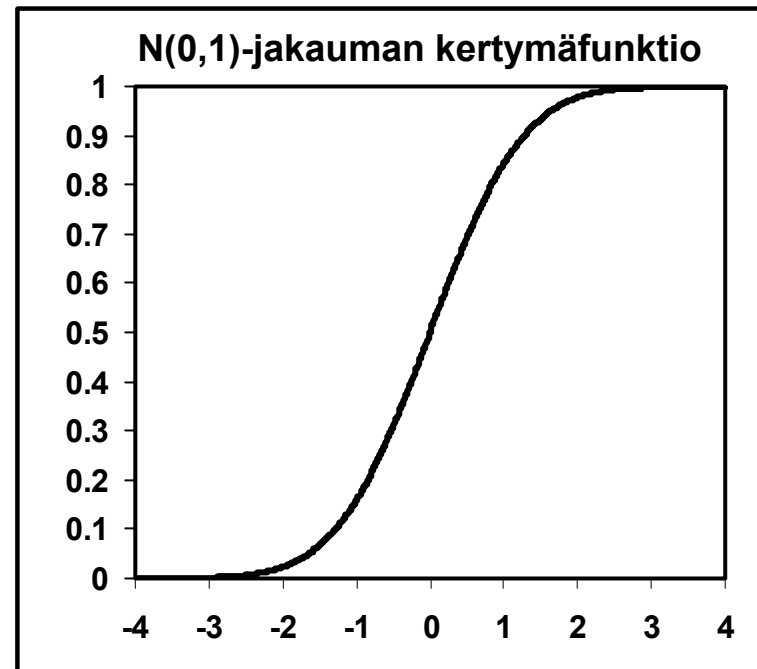
$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\}$$



Normaalijakauma

Standardoitu normaalijakauma: Kertymäfunktion kuvaaja

- Kuva oikealla esittää standardoidun normaalijakauman $N(0,1)$ kertymäfunktiota.
- Standardoidun normaalijakauman kertymäfunktion $F(x)$ määrittelee kaava
$$F(x) = \Pr(X \leq x) = \int_{-\infty}^x f(t) dt$$
jossa $f(x)$ on standardoidun normaalijakauman tiheysfunktio.



Normaalijakauma

Lineaarimuunnoksen jakauma

- Olkoon $X \sim N(\mu, \sigma^2)$.
- Määritellään satunnaismuuttuja

$$Y = a + bX$$

jossa a ja b ovat (ei-satunnaisia) vakioita.

- Tällöin Y on *normaalinen*:

$$Y \sim N(a + b\mu, b^2\sigma^2)$$

- Perustelu:

Ks. lukua **Satunnaismuuttujien muunnosten jakaumat**.

Normaalijakauma

Standardointi

- Olkoon $X \sim N(\mu, \sigma^2)$, jolloin

$$E(X) = \mu$$

$$D(X) = \sigma$$

- *Standardoidaan* satunnaismuuttuja X :

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

- Standardoitu satunnaismuuttuja Z noudattaa *standardoitua normaalijakaumaa* $N(0,1)$:

$$Z \sim N(0,1)$$

Normaalijakauma

Normaalijakauma ja standardoitu normaalijakauma 1/2

- *Kaikki* normaalijakaumat $N(\mu, \sigma^2)$ ovat *samanmuotoisia standardoiduissa yksiköissä*

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

- Siten todennäköisyydet *mielivaltaisesta normaali-jakaumasta* $N(\mu, \sigma^2)$ voidaan aina määrätä *standardoidun normaalijakauman* $N(0,1)$ avulla.

Normaalijakauma

Normaalijakauma ja standardoitu normaalijakauma 2/2

- Olkoon siis

$$X \sim N(\mu, \sigma^2)$$

$$Z \sim N(0,1)$$

- Tällöin

$$\Pr(a \leq X \leq b)$$

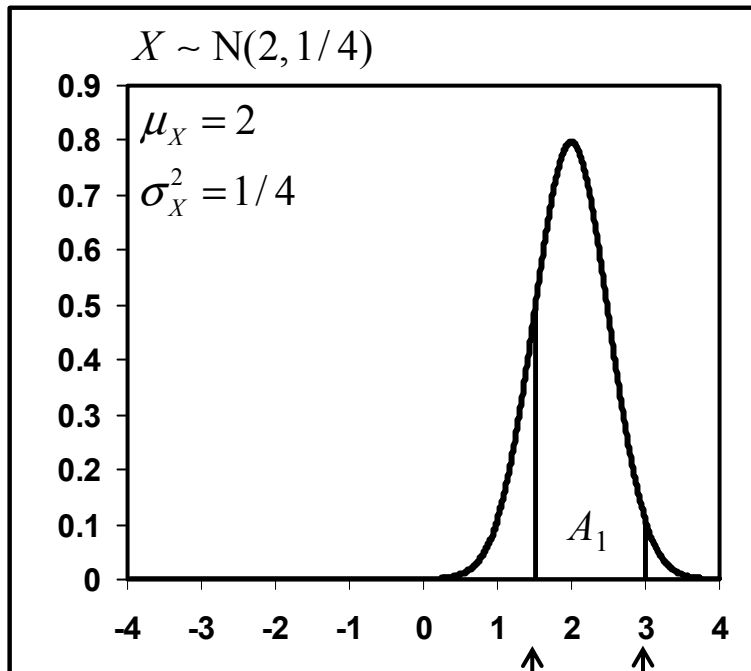
$$= \Pr\left(\frac{a - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{b - \mu}{\sigma}\right)$$

$$= \Pr\left(\frac{a - \mu}{\sigma} \leq Z \leq \frac{b - \mu}{\sigma}\right)$$

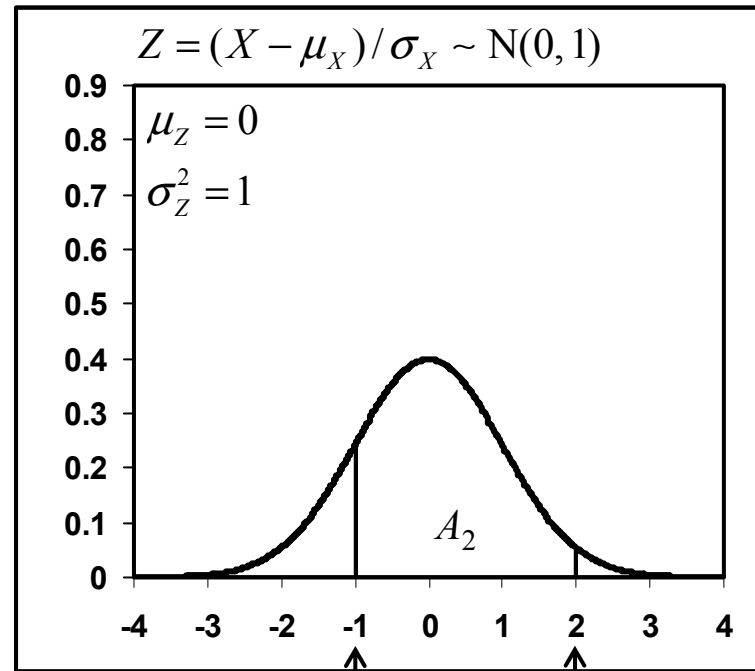
Normaalijakauma

Normaalijakauma ja

standardoitu normaalijakauma: Esimerkki



$$a = 1.5 \quad b = 3$$



$$\frac{a - \mu_X}{\sigma_X} = -1 \quad \frac{b - \mu_X}{\sigma_X} = 2$$

Standardoidun normaalijakauman $N(0,1)$ taulukoista saadaan:

$$\text{Alueen } A_1 \text{ pinta-ala} = \Pr(1.5 \leq X \leq 3)$$

$$= \text{Alueen } A_2 \text{ pinta-ala} = \Pr(-1 \leq Z \leq 2) = 0.8185$$

Normaalijakauma

Todennäköisyyksien määrittäminen standardoidusta normaalijakaumasta 1/2

- Todennäköisyydet *standardoidusta normaalijakaumasta* $N(0,1)$ voidaan määrätä jakauman *kertymäfunktion* avulla.
- Olkoon $Z \sim N(0,1)$.
- Olkoon satunnaismuuttujan Z *kertymäfunktio*

$$\Phi(z) = \Pr(Z \leq z)$$

- Huomautus:

Koska *normaalijakauman tiheysfunktion integraalifunktiota ei osata esittää alkeisfunktioiden avulla suljetussa muodossa*, normaalijakauman kertymäfunktion määrittämiseen on käytettävä jotakin *numeerista menetelmää*.

Siksi useimmissa alan oppikirjoissa on valmis *taulukko*, jossa on taulukoituna normaalijakauman kertymäfunktion arvoja ja niihin liittyviä todennäköisyyksiä.

Normaalijakauma

Todennäköisyyksien määrittäminen standardoidusta normaalijakaumasta 2/2

- *Kaikkien* standardoituun normaalijakaumaan liittyvien tapahtumien todennäköisyydet saadaan todennäköisyyksistä

$$\Pr(Z \leq z) = \Phi(z)$$

todennäköisyyslaskennan laskusääntöjen avulla.

- Esimerkiksi

$$\Pr(a \leq Z \leq b) = \Phi(b) - \Phi(a)$$

Todennäköisyyksien määrittäminen standardoidusta normaalijakaumasta: Taulukot 1/2

- Standardoidun normaalijakauman *taulukot* sisältävät standardoidun normaalijakauman $N(0,1)$ *kertymäfunktion* $\Phi(z)$ *arvoja* taulukoituna useille eri argumentin z arvoille.
- Siten taulukot mahdollistavat seuraavien tehtävien ratkaisemisen:

(i) Määrää todennäköisyys

$$\Pr(Z \leq z) = \Phi(z)$$

kun z on annettu.

(ii) Määrää z , kun todennäköisyys

$$\Pr(Z \leq z) = \Phi(z)$$

on annettu.

Normaalijakauma

Todennäköisyyksien määrittäminen standardoidusta normaalijakaumasta: Taulukot 2/2

- Monissa normaalijakauman taulukoissa on taulukoitu todennäköisyyksiä

$$\Pr(Z \leq z) = \Phi(z)$$

vain, kun $z \geq 0$.

- Tällöin todennäköisyydet $\Pr(Z \leq -z) = \Phi(-z)$ saadaan soveltamalla standardoidun normaalijakauman tiheysfunktion *symmetrisyyttä* pisteen $z = 0$ suhteen:

$$\begin{aligned}\Phi(-z) &= \Pr(Z \leq -z) \\ &= 1 - \Pr(Z \geq -z) \\ &= 1 - \Pr(Z \leq z) \\ &= 1 - \Phi(z)\end{aligned}$$

Normaalijakauma

Todennäköisyyksien määrittäminen

standardoidusta normaalijakaumasta: Esimerkki 1/2

- Olkoon

$$Z \sim N(0,1)$$

ja olkoon

$$f_Z(z)$$

satunnaismuuttujan Z
tiheysfunktio.

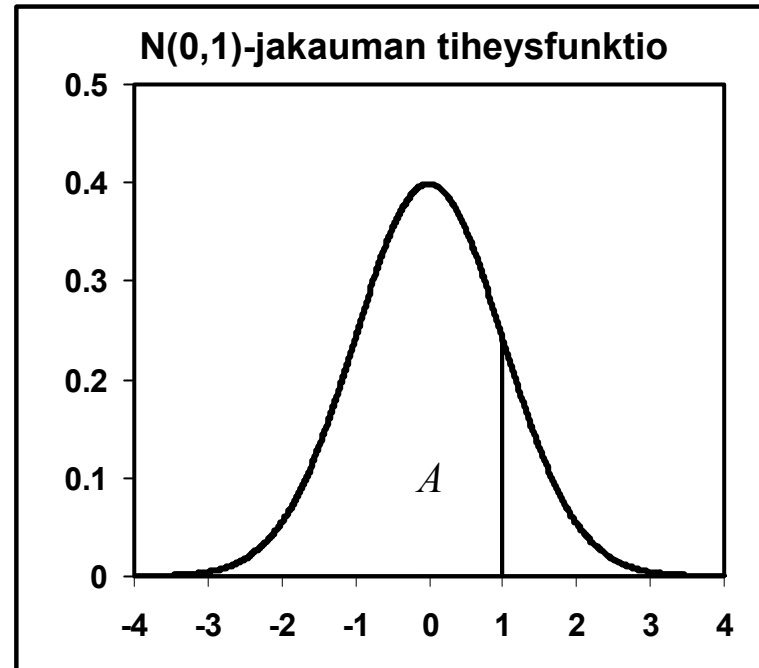
- Standardoidun normaalijakauman $N(0,1)$ taulukoista saadaan:

Alueen A pinta-ala

$$= \int_{-\infty}^1 f_Z(z) dz$$

$$= \Pr(Z \leq 1)$$

$$= 0.8413$$



Normaalijakauma

Todennäköisyyksien määrittäminen

standardoidusta normaalijakaumasta: Esimerkki 2/2

- Olkoon

$$Z \sim N(0,1)$$

ja olkoon

$$\Phi(z)$$

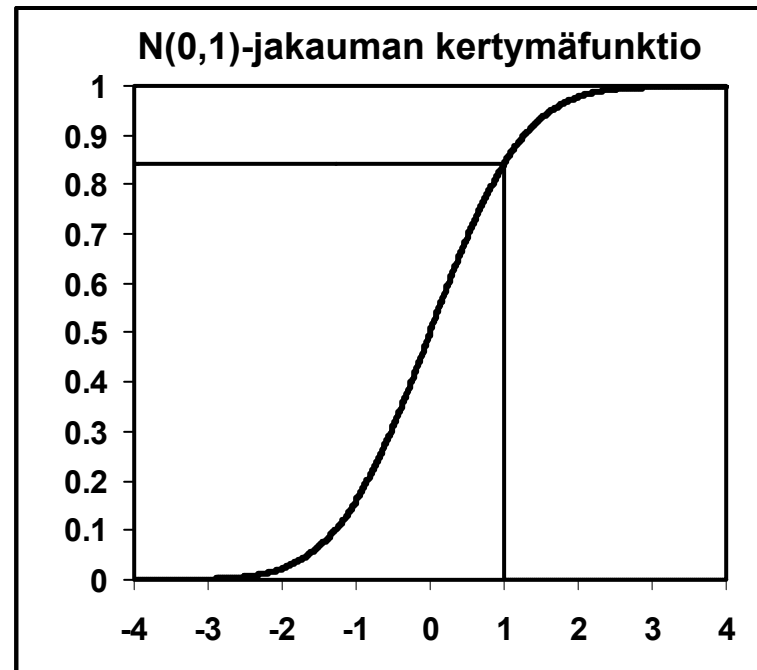
satunnaismuuttujan Z
kertymäfunktio.

- Standardoidun normaalijakauman $N(0,1)$ taulukoista saadaan:

$$\Phi(1)$$

$$= \Pr(Z \leq 1)$$

$$= 0.8413$$



Normaalijakauma

Todennäköisyyksien määrittäminen normaalijakaumasta: Ohjelmat

- Olkoon $X \sim N(\mu, \sigma^2)$.
- Monet *tietokoneohjelmat* mahdollistavat seuraavien tehtävien ratkaisemisen mielivaltaisille parametrien μ, σ^2 arvoille:

- (i) Määrää todennäköisyys

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

kun x on annettu.

- (ii) Määrää x , kun todennäköisyys

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

kun on x annettu.

Normaalijakauma

Kahden normaalijakautuneen satunnaismuuttujan summan jakauma

- Olkoon

$$X \sim N(\mu_X, \sigma_X^2)$$

$$Y \sim N(\mu_Y, \sigma_Y^2)$$

ja olkoot X ja Y lisäksi *riippumattomia*.

- Määritellään satunnaismuuttuja

$$W = X + Y$$

- Tällöin *summa* $W = X + Y$ on *normaalinen*:

$$W \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$$

- Perustelu:

Ks. lukua **Satunnaismuuttujien muunnosten jakaumat**.

Normaalijakauma

Normaalijakautuneiden satunnaismuuttujien summan jakauma 1/2

- Olkoon $X_i, i = 1, 2, \dots, n$ jono *riippumattomia normaalijakautuneita satunnaismuuttujia*.

- Siten

$$X_1, X_2, \dots, X_n \perp$$

$$X_i \sim N(\mu_i, \sigma_i^2), i = 1, 2, \dots, n$$

- Olkoon

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

satunnaismuuttujien $X_i, i = 1, 2, \dots, n$ *summa*.

Normaalijakauma

Normaalijakautuneiden satunnaismuuttujien summan jakauma 2/2

- Tällöin summa Y_n on *normaalinen*:

$$Y_n \sim N(\mu_1 + \mu_2 + \dots + \mu_n, \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2)$$

- Sanoin:

Riippumattomien, normaalijakautuneiden satunnaismuuttujien summa on normaalinen ja parametrit saadaan yhteenlaskettavien satunnaismuuttujien vastaavien parametrien summina.

- Perustelu:

Ks. lukua **Satunnaismuuttujien muunnosten jakaumat**.

Normaalijakauma

Normaalijakautuneiden satunnaismuuttujien summan jakauma 1/2

- Olkoon $X_i, i = 1, 2, \dots, n$ jono *riippumattomia, samaa normaalijakaumaa noudattavia satunnaismuuttujia*.

- Siten

$$X_1, X_2, \dots, X_n \perp$$

$$X_i \sim N(\mu, \sigma^2), i = 1, 2, \dots, n$$

- Olkoon

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

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

Normaalijakauma

Normaalijakautuneiden satunnaismuuttujien summan jakauma 2/2

- Tällöin summa Y_n on *normaalinen*:

$$Y_n = \sum_{i=1}^n X_i \sim N(n\mu, n\sigma^2)$$

- Siten *riippumattomien, samaa normaalijakaumaa noudattavien satunnaismuuttujien summa on normaalinen* ja parametrit saadaan yhteenlaskettavien satunnaismuuttujien vastaavien parametrien summina.

- Huomautus:

Tulos on *erikoistapaus* riippumattomien, normaalijakautuneiden satunnaismuuttujien summaa koskevasta yleisestä jakaumatuloksesta.

Normaalijakauma

Normaalijakautuneiden satunnaismuuttujien aritmeettisen keskiarvon jakauma 1/2

- Olkoon $X_i, i = 1, 2, \dots, n$ jono *riippumattomia, samaa normaalijakaumaa noudattavia satunnaismuuttujia*.

- Siten

$$X_1, X_2, \dots, X_n \perp$$

$$X_i \sim N(\mu, \sigma^2), i = 1, 2, \dots, n$$

- Olkoon

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

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

Normaalijakauma

Normaalijakautuneiden satunnaismuuttujien aritmeettisen keskiarvon jakauma 2/2

- Tällöin aritmeettinen keskiarvo \bar{X} on *normaalinen*:

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

- Siten *riippumattomien, samaa normaalijakaumaa noudattavien satunnaismuuttujien aritmeettinen keskiarvo on normaalinen.*
- Huomautus:

Ilman normalisuusoletustakin pätee:

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

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

Miksi normaalijakauma on ”normaali”?

- **Normaalijakauma on sekä *teoreettisen* että *soveltavan tilastotieteen* tärkein jakauma.**
- Normaalijakauman keskeinen asema tilastotieteessä perustuu siihen *teoreettiseen* ja *empiiriseen* tosiseikkaan, että moniin satunnaisilmiöihin liittyvät satunnaismuuttujat noudattavat ainakin *approksimatiivisesti* normaalijakaumaa.
- Mikä on tämän tosiseikan selitys?
- Selityksenä on **keskeinen raja-arvolause**; ks. seuraavaa kappaletta.

Jatkuvia jakaumia

Jatkuva tasainen jakauma

Eksponenttijakauma

Normaalijakauma

>> Keskeinen raja-arvolause

Log-normaali-, Cauchy-, Gamma-, Beta- ja Weibull-jakaumat

Keskeinen raja-arvolause

Johdanto 1/2

- Olkoon X_i , $i = 1, 2, \dots, n$ jono *riippumattomia, samaa normaalijakaumaa* $N(\mu, \sigma^2)$ *noudattavia satunnaismuuttujia.*
- Tällöin *satunnaismuuttujien* X_i *summa* Y_n *on normaalin:*

$$Y_n = \sum_{i=1}^n X_i \sim N(n\mu, n\sigma^2)$$

- Kysymys:
Mitä voidaan sanoa *riippumattomien, samaa jakaumaa noudattavien satunnaismuuttujien summan jakaumasta*, jos ko. satunnaismuuttujat *eivät noudata normaali-jakaumaa?*

Keskeinen raja-arvolause

Johdanto 2/2

- *Ei-normaalisten* satunnaismuuttujien summa *ei yleensä ole* normaalin.
- Kuitenkin, jos yhteenlaskettavia on ”tarpeeksi paljon”, *satunnaismuuttujien summa on* (hyvin yleisin ehdoin) *approksimatiivisesti normaalin*.
- Tämä on **keskeisen raja-arvolauseen** olennainen sisältö.
- Koska monia satunnaismuuttujia voidaan pitää *usean riippumattoman tekijän summana*, antaa keskeinen raja-arvolause selityksen *empiiriselle havainnolle* niiden normalisuudesta.

Keskeisen raja-arvolauseen formulointi 1/3

- Olkoon $X_i, i = 1, 2, \dots$ jono *riippumattomia, samoin jakautuneita satunnaismuuttujia*, joiden odotusarvo ja varianssi ovat

$$E(X_i) = \mu, i = 1, 2, \dots$$

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

- Olkoon

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

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

Keskeisen raja-arvolauseen formulointi 2/3

- Summan Y_n odotusarvo ja varianssi ovat

$$E(Y_n) = n\mu$$

$$D^2(Y_n) = n\sigma^2$$

- *Standardoidaan* summa Y_n :

$$Z_n = \frac{Y_n - n\mu}{\sigma\sqrt{n}}$$

- Annetaan $n \rightarrow +\infty$
- Tällöin satunnaismuuttujan Z_n *jakauma lähestyy standardoitua normaalijakaumaa* $N(0,1)$.

Keskeisen raja-arvolauseen formulointi 3/3

- Siten **keskeinen raja-arvolause** sanoo, että

$$\lim_{n \rightarrow +\infty} \Pr \left(\frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \leq z \right) = \Phi(z)$$

jossa Φ on *standardoidun normaalijakauman* $N(0,1)$ *kertymäfunktio*.

- Merkintä:

$$\frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \sim_a N(0,1)$$

Keskeinen raja-arvolause

Kommentteja 1/3

- Keskeiselle raja-arvolauseelle esitetään *todistus* luvussa **Stokastiikan konvergenssikäsitteet ja raja-arvolauseet**.
- Keskeisen raja-arvolauseen mukaan usean satunnaismuuttujan *summa on (tietyin ehdoin) approksimatiivisesti normaalin (lähes) riippumatta yhteenlaskettavien jakaumasta*.
- Huomautus:
Yhteenlaskettavien ei tarvitse olla edes *jatkuvia*, vaan ne voivat olla jopa *diskreettejä*.

Keskeinen raja-arvolause

Kommentteja 2/3

- Approksimaation *hyvyys riippuu* yhteenlaskettavien satunnaismuuttujien lukumäärästä, niiden jakaumasta ja erityisesti niiden jakauman *vinoudesta*.
- Approksimaation *hyvyys paranee*, kun yhteenlaskettavien satunnaismuuttujien *lukumäärä kasvaa*.
- Jos yhteenlaskettavien satunnaismuuttujien jakauma on *symmetrinen*, approksimaatio on hyvä jo suhteellisen pienillä yhteenlaskettavien lukumäärillä.
- Jos yhteenlaskettavien satunnaismuuttujien jakauma on *epäsymmetrinen*, hyvä approksimaatio vaatii enemmän yhteenlaskettavia.

Keskeinen raja-arvolause

Kommentteja 3/3

- Keskeinen raja-arvolause koskee satunnaismuuttujien **asymptoottista käyttäytymistä** samaan tapaan kuin luvussa Jakaumien tunnusluvut esitetty **suurten lukujen laki**.
- Keskeisessä raja-arvolauseessa esiintyvä *rajakäyttäytymisen muoto* on esimerkki ns. *jakauma-konvergenssista* eli *heikosta konvergenssista*.
- Keskeisestä raja-arvolauseesta on olemassa *yleisempiä muotoja*, joissa lievennetään *samoinjakautuneisuus-* ja *riippumattomuusoletuksia*.

Keskeinen raja-arvolause

Aritmeettisen keskiarvon approksimatiivinen jakauma

- Keskeisestä raja-arvolauseesta seuraa:
Riippumattomien samoin jakautuneiden satunnaismuuttujien X_i , $i = 1, 2, \dots, n$ **aritmeettinen keskiarvo**

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

on suurille (mutta äärellisille) n *approksimatiivisesti normaalinen parametreinaan μ ja σ^2/n :*

$$\bar{X}_n \sim_a N\left(\mu, \frac{\sigma^2}{n}\right)$$

Keskeisen raja-arvolauseen seurauksia 1/3

- Keskeisellä raja-arvolauseesta seuraa erikoistapauksina monet yksittäisiä jakaumia koskevat asymptoottiset tulokset.
- Käsittelemme seuraavia erikoistapauksia:
 - (i) **Binomijakauma lähestyy normaalijakaumaa**, kun toistokokeiden lukumäärän n annetaan kasvaa.
 - (ii) **Poisson-jakauma lähestyy normaalijakaumaa**, kun jakauman intensiteettiparametrin λ arvon annetaan kasvaa.

Keskeisen raja-arvolauseen seurauksia 2/3

- Sitä, että *binomijakauma lähestyy* toistokokeiden lukumäärän n kasvaessa *normaalijakaumaa*, kutsutaan tavallisesti **De Moivren ja Laplacen raja-arvolauseeksi**.
- De Moivren ja Laplacen raja-arvolauseen mukaan *binomitodennäköisyyksiä voidaan approksimoida normaalijakaumasta määrätyillä todennäköisyyksillä*, jos toistokokeiden lukumäärä on kyllin suuri.
- Koska hypergeometrinen jakauma muistuttaa tietyin ehdoin binomijakaumaa, myös *hypergeometrisen jakauman todennäköisyyksiä voidaan approksimoida normaalijakaumasta määrätyillä todennäköisyyksillä*.

Keskeisen raja-arvolauseen seurauksia 3/3

- Poisson-jakaumaa koskevan keskeisen raja-arvolauseen muodon mukaan *Poisson-jakauman todennäköisyyksiä voidaan approksimoida normaalijakaumasta määrätyillä todennäköisyyksillä.*

Keskeinen raja-arvolause

De Moivre'n ja Laplacen raja-arvolause

- Olkoon $X \sim \text{Bin}(n, p)$ ja $q = 1 - p$.

- Siten

$$E(X) = np$$

$$\text{Var}(X) = npq$$

- Tällöin

$$\lim_{n \rightarrow +\infty} \Pr \left(\frac{X - np}{\sqrt{npq}} \leq z \right) = \Phi(z)$$

jossa Φ on *standardoidun normaalijakauman* $N(0,1)$ *kertymäfunktio*.

Keskeinen raja-arvolause

De Moivre'n ja Laplacen raja-arvolause: Perustelu 1/2

- Olkoon $X \sim \text{Bin}(n, p)$.
- Tällöin satunnaismuuttuja X voidaan esittää *riippumattomien, samaa Bernoulli-jakaumaa* noudattavien satunnaismuuttujien X_i summana:

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

jossa

$$X_i \sim \text{Bernoulli}(p), i = 1, 2, \dots, n$$

- Koska

$$E(X_i) = p$$

$$\text{Var}(X_i) = npq, q = 1 - p$$

niin

$$E(X) = \sum_{i=1}^n E(X_i) = np$$

$$\text{Var}(X) = \sum_{i=1}^n \text{Var}(X_i) = npq$$

Keskeinen raja-arvolause

De Moivre'n ja Laplacen raja-arvolause: Perustelu 2/2

- Tällöin *keskeisestä raja-arvolauseesta* seuraa, että

$$\lim_{n \rightarrow +\infty} \Pr\left(\frac{X - np}{\sqrt{npq}} \leq z\right) = \Phi(z)$$

jossa Φ on *standardoidun normaalijakauman* $N(0,1)$ *kertymäfunktio*.

Keskeinen raja-arvolause

De Moivre'n ja Laplacen raja-arvolause: Havainnollistus 1/5

- Kalvoilla 2/5-5/5 oleva kuvasarja havainnollistaa De Moivre'n ja Laplacen raja-arvolausetta.
- Kuvasarja näyttää miten satunnaismuuttujien

$$X \sim \text{Bin}(n, p)$$

$$Z \sim \text{N}(\mu, \sigma^2)$$

jakaumat alkavat muistuttaa yhä enemmän toisiaan, kun toistokokeiden lukumäärän n annetaan kasvaa.

- Kuvasarjassa

$$p = 0.1$$

$$n = 1, 10, 30, 100$$

$$\mu = np$$

$$\sigma^2 = np(1 - p)$$

Keskeinen raja-arvolause

De Moivre'n ja Laplacen raja-arvolause: Havainnollistus 2/5

- Olkoon

$$X \sim \text{Bin}(n, p)$$

$$n = 1$$

$$p = 0.1$$

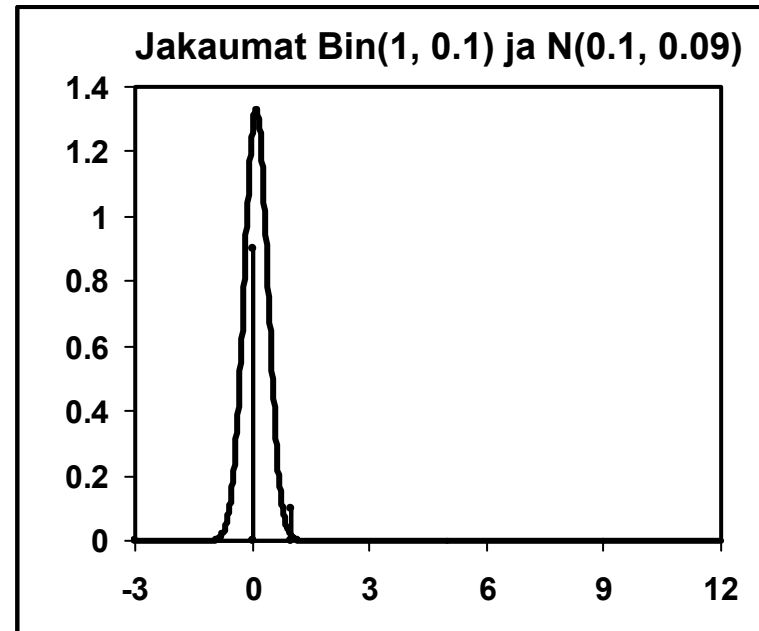
ja

$$Z \sim \text{N}(\mu, \sigma^2)$$

$$\mu = np = 0.1$$

$$\sigma^2 = np(1 - p) = 0.09$$

- Kuva oikealla esittää satunnaismuuttujan X pistetodennäköisyysfunktiota ja satunnaismuuttujan Z tiheysfunktiota välillä $[-3, 12]$.



Keskeinen raja-arvolause

De Moivre'n ja Laplacen raja-arvolause: Havainnollistus 3/5

- Olkoon

$$X \sim \text{Bin}(n, p)$$

$$n = 10$$

$$p = 0.1$$

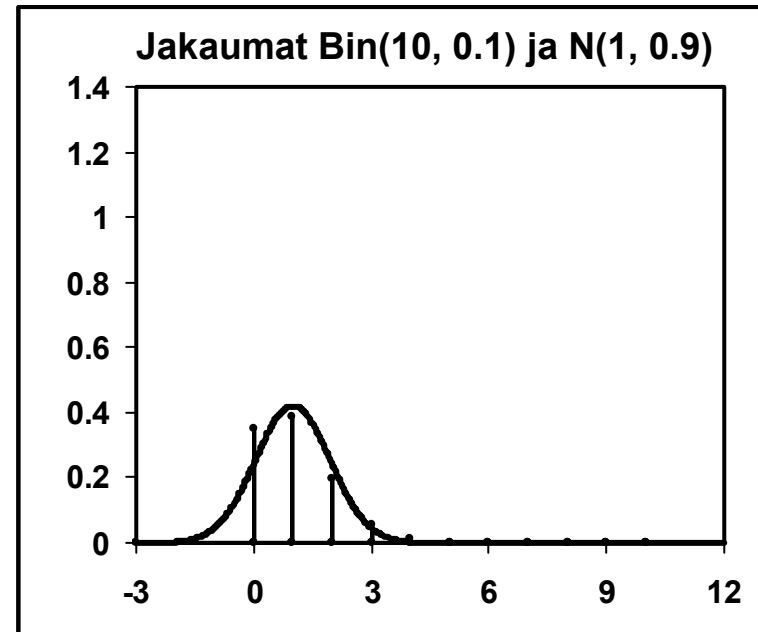
ja

$$Z \sim N(\mu, \sigma^2)$$

$$\mu = np = 1$$

$$\sigma^2 = np(1 - p) = 0.9$$

- Kuva oikealla esittää satunnaismuuttujan X pistetodennäköisyysfunktiota ja satunnaismuuttujan Z tiheysfunktiota välillä $[-3, 12]$.



Keskeinen raja-arvolause

De Moivre'n ja Laplacen raja-arvolause: Havainnollistus 4/5

- Olkoon

$$X \sim \text{Bin}(n, p)$$

$$n = 30$$

$$p = 0.1$$

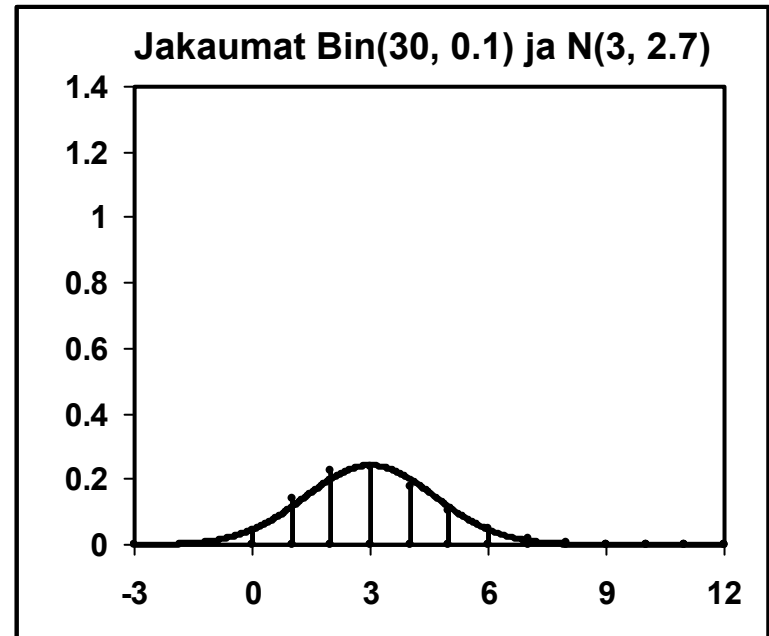
ja

$$Z \sim \text{N}(\mu, \sigma^2)$$

$$\mu = np = 3$$

$$\sigma^2 = np(1 - p) = 2.7$$

- Kuva oikealla esittää satunnaismuuttujan X pistetodennäköisyysfunktioita ja satunnaismuuttujan Z tiheysfunktioita välillä $[-3, 12]$.



Keskeinen raja-arvolause

De Moivre'n ja Laplacen raja-arvolause: Havainnollistus 5/5

- Olkoon

$$X \sim \text{Bin}(n, p)$$

$$n = 100$$

$$p = 0.1$$

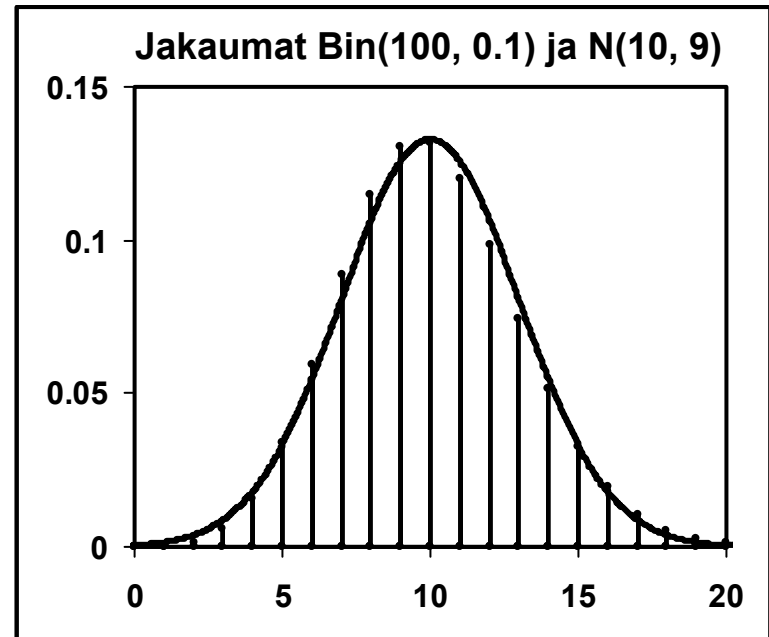
ja

$$Z \sim \text{N}(\mu, \sigma^2)$$

$$\mu = np = 10$$

$$\sigma^2 = np(1 - p) = 9$$

- Kuva oikealla esittää satunnaismuuttujan X pistetodennäköisyysfunktioita ja satunnaismuuttujan Z tiheysfunktioita välillä $[0, 20]$.



Keskeinen raja-arvolause
**Binomitodennäköisyydet ja
normaalijakauma 1/4**

- De Moivre'n ja Laplace'n raja-arvolauseen mukaan
binomijakaumaa

$$\text{Bin}(n, p)$$

voidaan suurille n approksimoida normaalijakaumalla

$$\text{N}(\mu, \sigma^2)$$

jossa

$$\mu = np$$

$$\sigma^2 = npq, q = 1 - p$$

Keskeinen raja-arvolause Binomitodennäköisyydet ja normaalijakauma 2/4

- Jos siis

$$X \sim \text{Bin}(n, p)$$

niin De Moivre'n ja Laplace'n raja-arvolauseen mukaan suurille n

$$\Pr(a < X \leq b) \approx \Phi\left(\frac{b - np}{\sqrt{npq}}\right) - \Phi\left(\frac{a - np}{\sqrt{npq}}\right)$$

jossa Φ on standardoidun normaalijakauman $N(0,1)$ kertymäfunktio.

Keskeinen raja-arvolause

Binomitodennäköisyydet ja normaalijakauma 3/4

- Jos a ja b ovat kokonaislukuja, approksimaatio *on* hieman *parempi*, jos käytetään kaavaa

$$\Pr(a < X \leq b) \approx \Phi\left(\frac{b + 1/2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{a - 1/2 - np}{\sqrt{npq}}\right)$$

- *Korjaustekijän 1/2* tarve perustuu siihen, että *diskreettiä binomijakaumaa approksimoidaan jatkuvalla normaalijakaumalla.*

Keskeinen raja-arvolause

Binomitodennäköisyydet ja normaalijakauma 4/4

- Jos annetaan $a \rightarrow -\infty$, saadaan approksimaatiotulos

$$\Pr(X \leq b) = F_X(b) \approx \Phi\left(\frac{b + 1/2 - np}{\sqrt{npq}}\right)$$

jossa F_X on *binomijakauman kertymäfunktio*.

- Jos $a = b$, saadaan approksimaatiotulos

$$\Pr(X = a) = f_X(a) \approx \Phi\left(\frac{a + 1/2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{a - 1/2 - np}{\sqrt{npq}}\right)$$

jossa f_X on *binomijakauman pistetodennäköisyysfunktio*.

Keskeinen raja-arvolause

Binomitodennäköisyydet ja normaalijakauma: Esimerkki 1/3

- Kuva oikealla esittää jakauman $\text{Bin}(100, 0.1)$ pistetodennäköisyysfunktiota ja jakauman

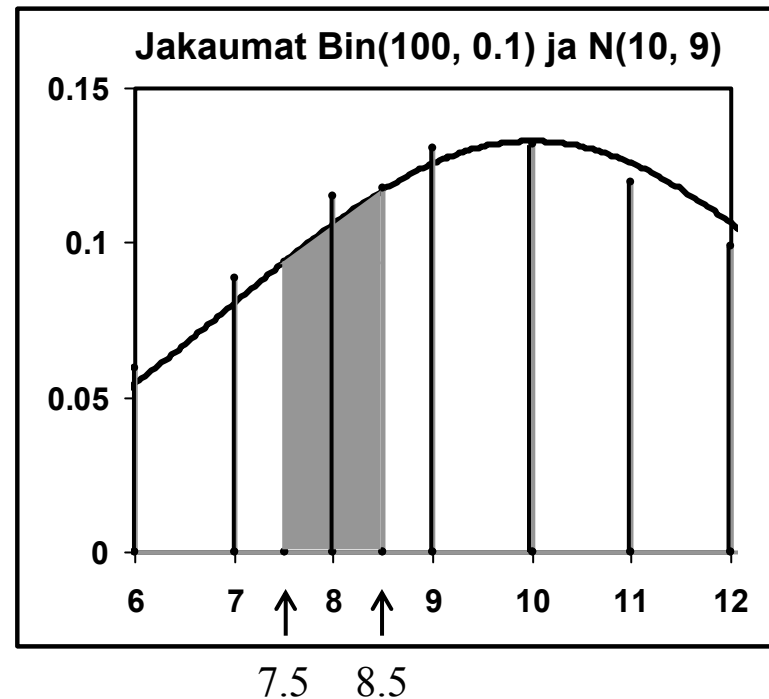
$$N(10, 9)$$

tiheysfunktiota välillä $[6, 12]$.

- De Moivre'n ja Laplace'n raja-arvolauseen mukaan binomitodennäköisyyttä pisteessä

$$x = 8$$

voidaan *approksimoida* varjostetun alueen pinta-alalla; ks. kalvoja 2/3-3/3.



Keskeinen raja-arvolause

Binomitodennäköisyydet ja normaalijakauma: Esimerkki 2/3

- Olkoon $X \sim \text{Bin}(n, p)$, jossa

$$n = 100$$

$$p = 0.1$$

- Tällöin

$$f_X(8) = 0.1148$$

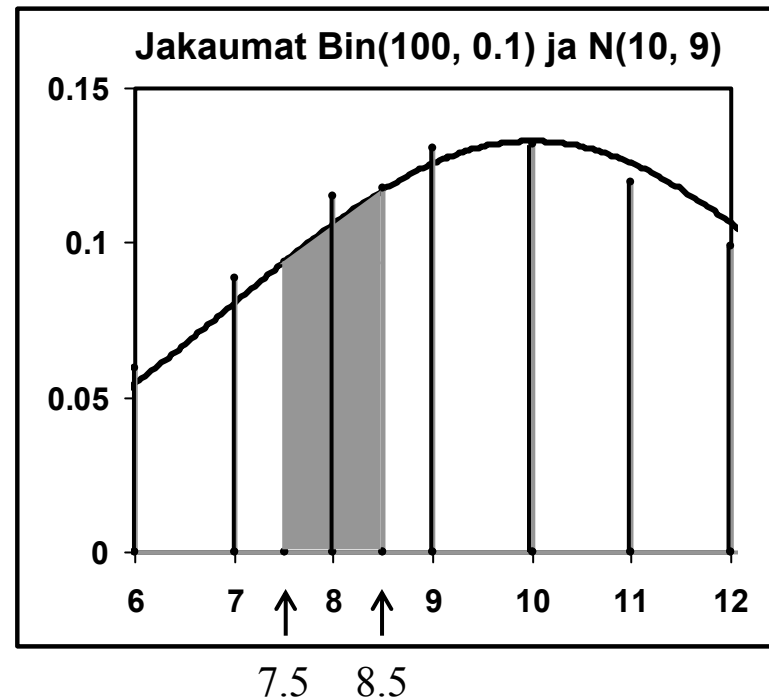
jossa

$$f_X(x)$$

on binomijakauman

$$\text{Bin}(100, 0.1)$$

pistetodennäköisyysfunktio.



Keskeinen raja-arvolause

Binomitodennäköisyydet ja normaalijakauma: Esimerkki 3/3

- Olkoot

$$\mu = np = 10$$

$$\sigma^2 = np(1 - p) = 9$$

jossa

$$n = 100$$

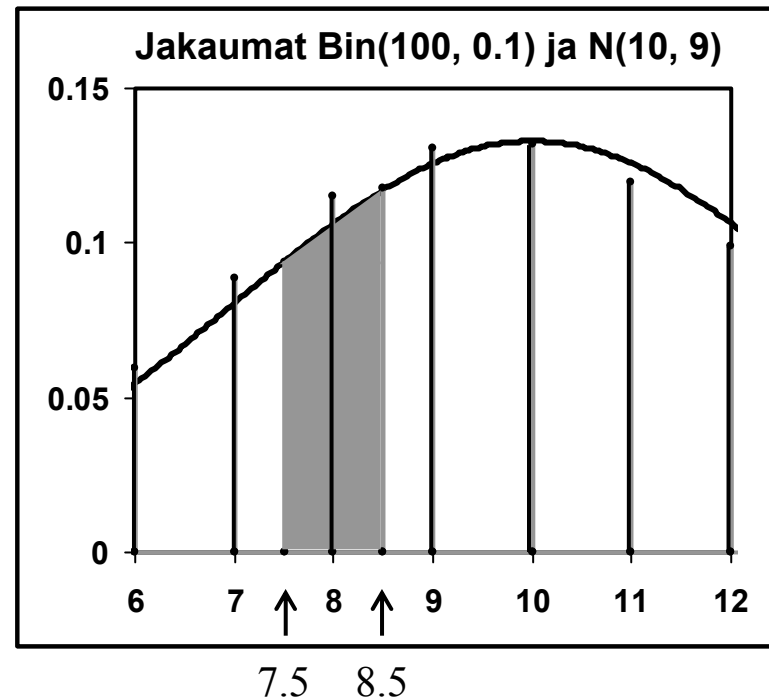
$$p = 0.1$$

- Tällöin

$$\Phi\left(\frac{8 + 1/2 - \mu}{\sigma}\right)$$

$$- \Phi\left(\frac{8 - 1/2 - \mu}{\sigma}\right) = 0.1052$$

jossa Φ on standardoidun normaalijakauman $N(0,1)$ kertymäfunktio.



Keskeinen raja-arvolause

Hypergeometrisen jakauman todennäköisyydet ja normaalijakauma 1/2

- *Hypergeometrinen jakauma*

$$\text{HyperGeom}(N, r, n)$$

lähestyy perusjoukon koon N kasvaessa rajatta *binomijakaumaa*

$$\text{Bin}(n, p)$$

jossa

$$p = r/N$$

Keskeinen raja-arvolause

Hypergeometrisen jakauman todennäköisyydet ja normaalijakauma 2/2

- Siten *hypergeometrista jakaumaa*

$$\text{HyperGeom}(N, r, n)$$

voidaan suurille N approksimoida normaalijakaumalla

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

jossa

$$\mu = n \frac{r}{N}$$

$$\sigma^2 = n \frac{r}{N} \left(1 - \frac{r}{N} \right)$$

Keskeinen raja-arvolause Poisson-jakauma ja normaalijakauma

- Olkoon $X \sim \text{Poisson}(\lambda)$.

- Siten

$$E(X) = \lambda$$

$$\text{Var}(X) = \lambda$$

- Tällöin

$$\lim_{\lambda \rightarrow +\infty} \Pr\left(\frac{X - \lambda}{\sqrt{\lambda}} \leq z\right) = \Phi(z)$$

jossa Φ on *standardoidun normaalijakauman* $N(0,1)$ *kertymäfunktio*.

Keskeinen raja-arvolause

Poisson-jakauman todennäköisyydet ja normaalijakauma 1/4

- Poisson-jakaumaa koskevan raja-arvolauseen mukaan
Poisson-jakaumaa

$$\text{Poisson}(\lambda)$$

voidaan suurille λ approksimoida normaalijakaumalla

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

jossa

$$\mu = \lambda$$

$$\sigma^2 = \lambda$$

Keskeinen raja-arvolause

Poisson-jakauman todennäköisyydet ja normaalijakauma 2/4

- Jos siis

$$X \sim \text{Poisson}(\lambda)$$

niin Poisson-jakaumaa koskevan raja-arvolauseen mukaan suurille λ

$$\Pr(a < X \leq b) \approx \Phi\left(\frac{b - \lambda}{\sqrt{\lambda}}\right) - \Phi\left(\frac{a - \lambda}{\sqrt{\lambda}}\right)$$

jossa Φ on *standardoidun normaalijakauman* $N(0,1)$ *kertymäfunktio*.

Keskeinen raja-arvolause

Poisson-jakauman todennäköisyydet ja normaalijakauma 3/4

- Jos a ja b ovat kokonaislukuja, approksimaatio *on* hieman *parempi*, jos käytetään kaavaa

$$\Pr(a < X \leq b) \approx \Phi\left(\frac{b + 1/2 - \lambda}{\sqrt{\lambda}}\right) - \Phi\left(\frac{a - 1/2 - \lambda}{\sqrt{\lambda}}\right)$$

- *Korjaustekijän 1/2* tarve perustuu siihen, että *diskreettiä Poisson-jakaumaa approksimoidaan jatkuvalla normaalijakaumalla.*

Keskeinen raja-arvolause

Poisson-jakauman todennäköisyydet ja normaalijakauma 4/4

- Jos annetaan $a \rightarrow -\infty$, saadaan approksimaatiotulos

$$\Pr(X \leq b) = F_X(b) \approx \Phi\left(\frac{b + 1/2 - \lambda}{\sqrt{\lambda}}\right)$$

jossa F_X on *Poisson-jakauman kertymäfunktio*.

- Jos $a = b$, saadaan approksimaatiotulos

$$\Pr(X = a) = f_X(a) \approx \Phi\left(\frac{a + 1/2 - \lambda}{\sqrt{\lambda}}\right) - \Phi\left(\frac{a - 1/2 - \lambda}{\sqrt{\lambda}}\right)$$

jossa f_X on *Poisson-jakauman pistetodennäköisyysfunktio*.

Jatkuvia jakaumia

Jatkuva tasainen jakauma

Eksponenttijakauma

Normaalijakauma

Keskeinen raja-arvolause

>> Log-normaali-, Cauchy-, Gamma-, Beta- ja Weibull-jakaumat

Log-normaalijakauma

- Satunnaismuuttuja X noudattaa **log-normaalijakaumaa** parametrein μ ja σ^2 , jos sen *tiheysfunktio* on

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \frac{1}{x} \cdot \exp\left\{-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2\right\}, x > 0$$

- Merkitään:

$$X \sim \text{LogN}(\mu, \sigma^2)$$

- *Odotusarvo*:

$$E(X) = \exp(\mu + \sigma^2 / 2)$$

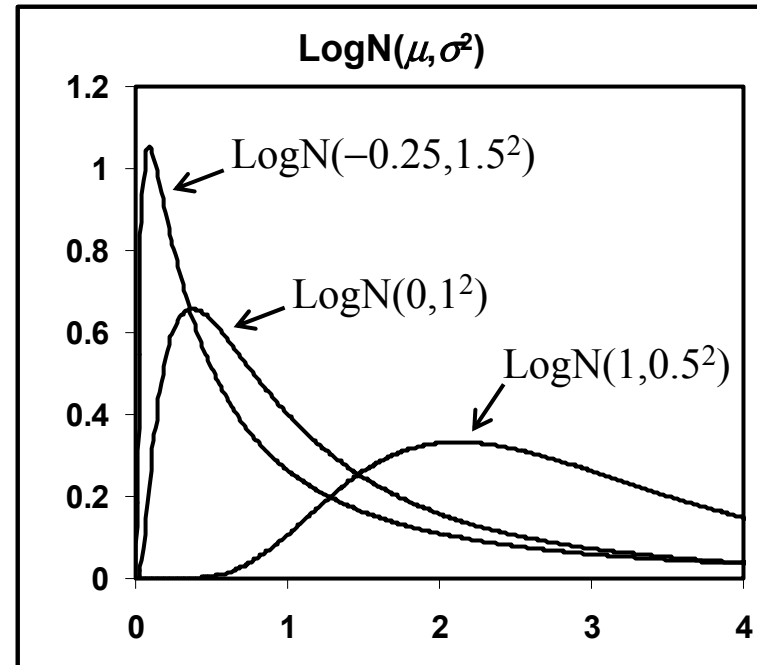
- *Varianssi*:

$$\text{Var}(X) = \exp[2(\mu + \sigma^2)] - \exp[2(\mu + \sigma^2 / 2)]$$

Jakaumia

Log-normaalijakauman tiheysfunktion kuvaajia

- Oikealla on log-normaalijakauman *tiheysfunktion* kuvaajia, kun
 - (i) $\mu = -0.25, \sigma^2 = 1.5^2$
 - (ii) $\mu = 0, \sigma^2 = 1^2$
 - (iii) $\mu = 1, \sigma^2 = 0.5^2$



Log-normaalijakauman yhteydet muihin jakaumiin

- Jos

$$X \sim \text{LogN}(\mu, \sigma^2)$$

niin satunnaismuuttuja

$$Y = \log(X)$$

noudattaa *normaalijakaumaa* parametrein μ ja σ^2 :

$$Y \sim \text{N}(\mu, \sigma^2)$$

ja kääntäen, jos

$$Y \sim \text{N}(\mu, \sigma^2)$$

niin

$$X = \text{Exp}(Y) \sim \text{LogN}(\mu, \sigma^2)$$

Cauchyn jakauma

- Satunnaismuuttuja X noudattaa **Cauchy-jakaumaa** parametrinaan θ , jos sen *tiheysfunktio* on

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1 + (x - \theta)^2}, \quad -\infty < x < +\infty$$

- Merkitään:

$$X \sim \text{Cauchy}(\theta)$$

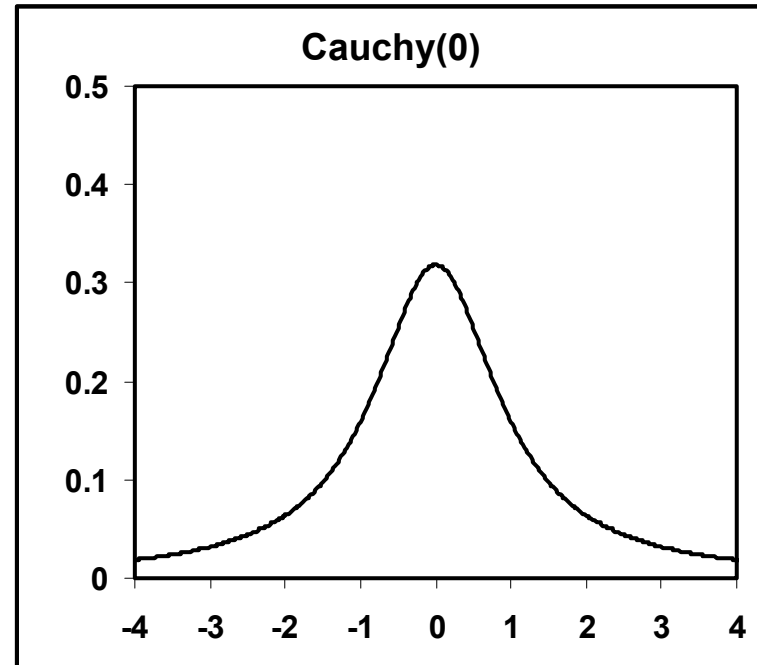
- *Odotusarvo*: Cauchy-jakaumalla *ei ole* odotusarvoa.
- *Varianssi*: Cauchy-jakaumalla *ei ole* varianssia.
- Parametri θ on Cauchy-jakauman *mediaani* ja *moodi*.

Jakaumia

Cauchy-jakauman tiheysfunktion kuvaaja

- Oikealla on Cauchy-jakauman *tiheysfunktion* kuvaaja, kun

$$\theta = 0$$



Jakaumia

Cauchyn jakauman yhteydet muihin jakaumiin

- Jos satunnaismuuttuja Y noudattaa *jatkuvaa tasaista jakaumaa* välillä $(-\pi/2, +\pi/2)$ eli

$$Y \sim \text{Uniform}(-\pi/2, +\pi/2)$$

jossa niin tällöin

$$X = \tan(Y) \sim \text{Cauchy}(0)$$

- Jos

$$X \sim \text{Cauchy}(0)$$

niin tällöin

$$X \sim t(1)$$

jossa $t(1)$ on *t-jakauma* yhdellä vapausasteella; ks. lukua

Normaalijakaumasta johdettuja jakaumia.

Gamma-jakauma

- Satunnaismuuttuja X noudattaa **gamma-jakaumaa** parametrein α ja β , jos sen *tiheysfunktio* on

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}, x \geq 0$$

- Merkitään:

$$X \sim \text{Gamma}(\alpha, \beta)$$

- *Odotusarvo*:

$$E(X) = \alpha\beta$$

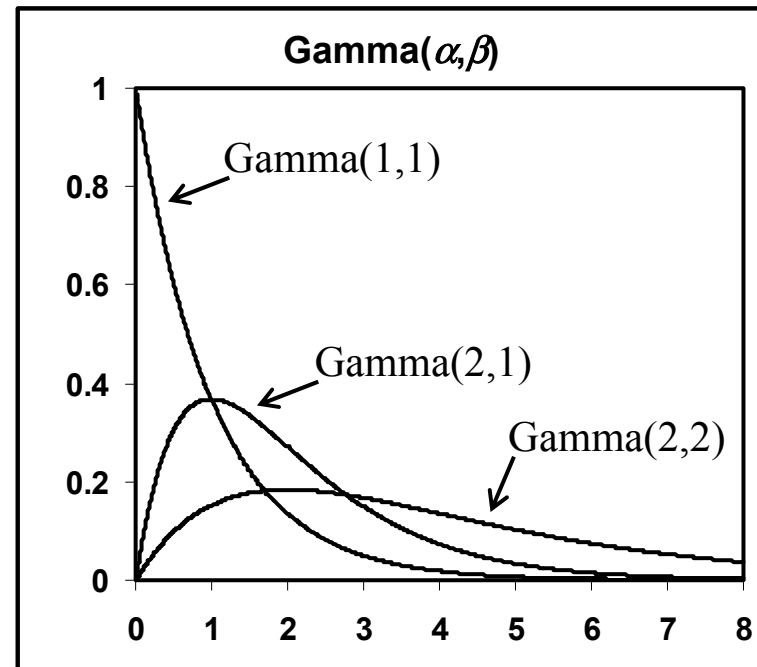
- *Varianssi*:

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

Jakaumia

Gamma-jakauman tiheysfunktion kuvaajia

- Oikealla on gamma-jakauman *tiheysfunktion* kuvaajia, kun
 - (i) $\alpha = 1, \beta = 1$
 - (ii) $\alpha = 2, \beta = 1$
 - (iii) $\alpha = 2, \beta = 2$



Gamma-jakauman yhteydet muihin jakaumiin 1/3

- Jos

$$X \sim \text{Gamma}(\alpha, \beta)$$

jossa α on *kokonaisluku*, niin tällöin

$$\Pr(X \leq x) = \Pr(Y \geq \alpha)$$

jossa satunnaismuuttuja Y noudattaa *Poisson-jakaumaa* parametrilla x/β :

$$Y \sim \text{Poisson}(x / \beta)$$

- Lisätietoja Poisson-jakaumasta: ks. lukua **Diskreettejä jakaumia**.

Gamma-jakauman yhteydet muihin jakaumiin 2/3

- Jos

$$X \sim \text{Gamma}(\alpha, \beta)$$

jossa $\alpha = 1$, niin tällöin satunnaismuuttuja X noudattaa *eksponenttijakaumaa* parametrilla $1/\beta$:

$$X \sim \text{Exp}(1/\beta)$$

Gamma-jakauman yhteydet muihin jakaumiin 3/3

- Jos

$$X \sim \text{Gamma}(\alpha, \beta)$$

jossa $\alpha = p/2$ ja $\beta = 2$, niin tällöin satunnaismuuttuja X noudattaa χ^2 -jakaumaa vapausastein p :

$$X \sim \chi^2(p)$$

- Lisätietoja χ^2 -jakaumasta: ks. lukua **Normaalijakaumasta johdettuja jakaumia**.

Beta-jakauma

- Satunnaismuuttuja X noudattaa **beta-jakaumaa** parametrein α ja β , jos sen *tiheysfunktio* on

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, 0 \leq x \leq 1$$

- Merkitään:

$$X \sim \text{Beta}(\alpha, \beta)$$

- *Odotusarvo*:

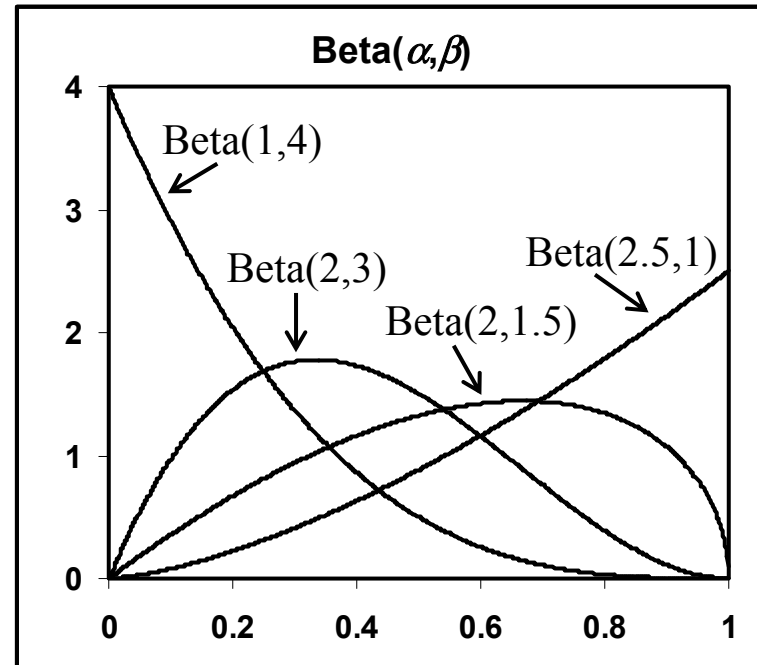
$$E(X) = \alpha / (\alpha + \beta)$$

- *Varianssi*:

$$\text{Var}(X) = \alpha\beta / [(\alpha + \beta)^2 (\alpha + \beta + 1)]$$

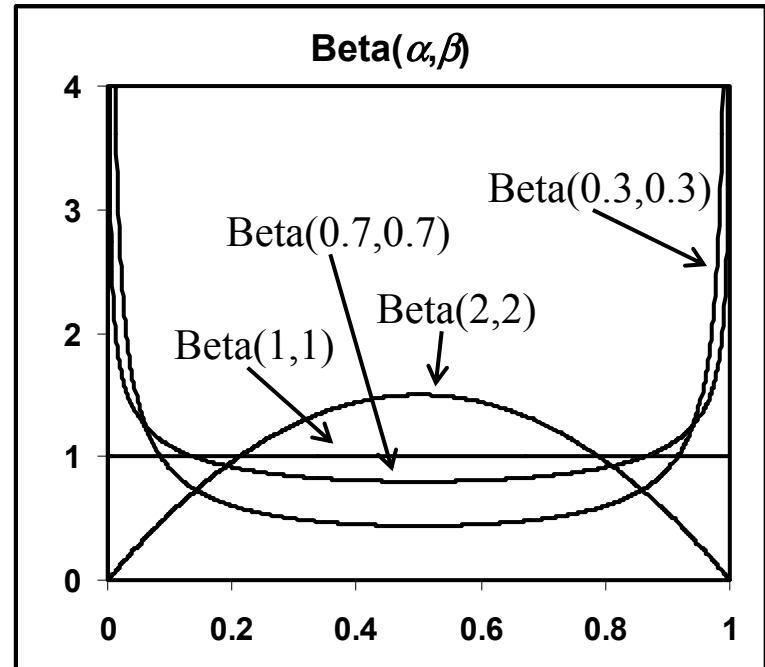
Beta-jakauman tiheysfunktion kuvaajia 1/2

- Oikealla on beta-jakauman *tiheysfunktion* kuvaajia, kun
 - (i) $\alpha = 1, \beta = 4$
 - (ii) $\alpha = 2, \beta = 3$
 - (iii) $\alpha = 2, \beta = 1.5$
 - (iv) $\alpha = 2.5, \beta = 1$



Beta-jakauman tiheysfunktion kuvaajia 2/2

- Oikealla on beta-jakauman *tiheysfunktion* kuvaajia, kun
 - (i) $\alpha = 1, \beta = 1$
 - (ii) $\alpha = 0.7, \beta = 0.7$
 - (iii) $\alpha = 2, \beta = 2$
 - (iv) $\alpha = 0.3, \beta = 0.3$



Beta-jakauman yhteydet muihin jakaumiin

- Jos

$$X \sim \text{Beta}(\alpha, \beta)$$

jossa $\alpha = 1$ ja $\beta = 1$, niin tällöin satunnaismuuttuja X noudattaa *jatkuvaa tasaista jakaumaa* välillä $(0,1)$:

$$X \sim \text{Uniform}(0,1)$$

Weibull-jakauma

- Satunnaismuuttuja X noudattaa **Weibull-jakaumaa** parametrein γ ja β , jos sen *tiheysfunktio* on

$$f(x) = \frac{\gamma}{\beta} x^{\gamma-1} e^{-x^\gamma/\beta}, \quad x \geq 0$$

- Merkitään:

$$X \sim \text{Weibull}(\gamma, \beta)$$

- *Odotusarvo*:

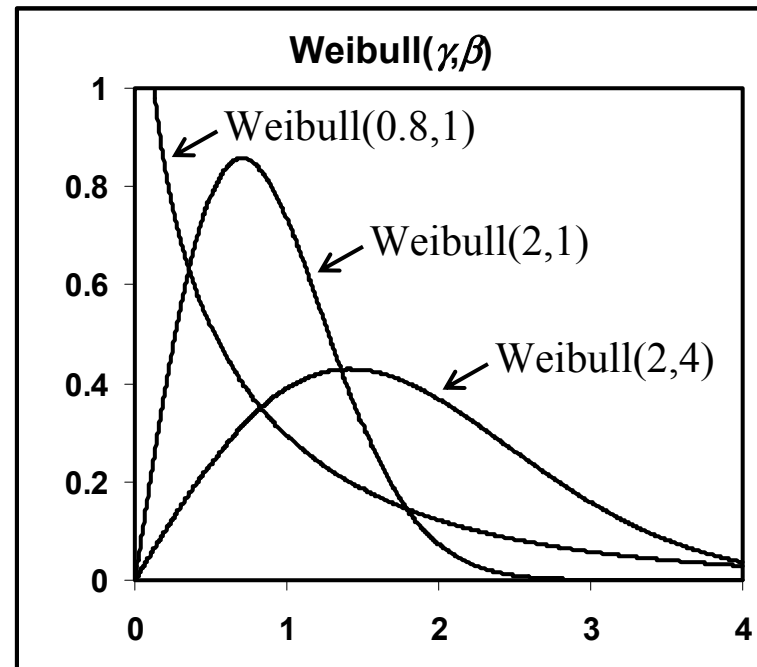
$$E(X) = \beta^{1/\gamma} \Gamma(1 + 1/\gamma)$$

- *Varianssi*:

$$\text{Var}(X) = \beta^{2/\gamma} \Gamma(1 + 2/\gamma) - [E(X)]^2$$

Weibull-jakauman tiheysfunktion kuvaajia

- Oikealla on Weibull-jakauman *tiheysfunktion* kuvaajia, kun
 - (i) $\gamma = 0.8, \beta = 1$
 - (ii) $\gamma = 2, \beta = 1$
 - (iii) $\gamma = 2, \beta = 4$



Jakaumia

Weibull-jakauman yhteydet muihin jakaumiin

- Jos

$$Y \sim \text{Exp}(\beta)$$

niin satunnaismuuttuja

$$X = Y^{1/\gamma}$$

noudattaa *Weibull-jakaumaa* parametrein γ ja β :

$$X \sim \text{Weibull}(\gamma, \beta)$$

ja kääntäen, jos

$$X \sim \text{Weibull}(1, \beta)$$

niin

$$Y = X \sim \text{Exp}(\beta)$$