Todennäköisyyslaskenta

Osa 2: Satunnaismuuttujat ja todennäköisyysjakaumat

- **Satunnaismuuttujat ja todennäköisyysjakaumat**
- Kertymäfunktio

Satunnaismuuttujat

- Jos satunnaisilmiötä halutaan mallintaa matemaattisesti, on ilmiön tulosvaihtoehdot kuvattava *numeerisessa* muodossa.

- Jos satunnaiskokeen tulos ei ole valmiiksi reaaliluku, se voidaan usein luontevasti muuntaa reaaliluvuksi jollakin funktiolla, joka suorittaa kuvauksen otosavaruudesta S reaalilukujen joukkoon $\mathbb{R}$. Tämä kuvaus on **satunnaismuuttuja.**

Satunnaismuuttujan määritelmä:

- Olkoon X *funktio* otosavaruudesta *S* reaalilukujen joukkoon $\mathbb{R}$:

$$X : S \to \mathbb{R}$$

  Jos X on 'mitallinen', niin X on **satunnaismuuttuja**.

- Voidaan osoittaa, että *diskreetit* ja *jatkuvat satunnais-muuttujat* – joita tällä kurssilla pelkästään käsitellään – *ovat mitallisia funktioita*.

Esimerkki: Sukupuolen määräytyminen − 1/2

- Tarkastellaan sukupuolen määräytymistä *satunnaisilmiönä*.
- Alkeistapahtumat:

  Tyttö, Poika
- Otosavaruus:

  $S = \{\text{Tyttö}, \text{Poika}\}$
- Määritellään reaaliarvoinen *funktio* $X : S \to \mathbb{R}$, joka liittää otos-avaruuden S alkioihin *numeerisen koodin* eli *reaaliluvun* seuraavalla tavalla:

  Tyttö $\to$ 1

  Poika $\to$ 0
- *Funktiota* X kutsutaan **satunnaismuuttujaksi**, koska *sattuma* määrää mikä funktion arvoista *realisoituu*, vaikka X on funktiona *täysin määrätty*.

Esimerkki: Sukupuolen määräytyminen − 2/2

- Tehdään seuraava, Suomen väkilukutilastoihin vuosilta 1991 – 95 perustuva *oletus* niistä *todennäköisyyksistä*, joilla X saa arvonsa:

  Pr(X= 1) = 0.4902

  Pr(X= 0) = 0.5098

- Satunnaismuuttujan X *arvot* ja niihin liitetyt *todennäköisyydet* muodostavat *yhdessä* satunnaismuuttujan X **todennäköisyys-jakauman**.
- Satunnaismuuttuja X ja sen todennäköisyysjakauma muodostavat **tilastollisen mallin** eli **todennäköisyysmallin** sukupuolen määräytymiselle satunnaisilmiönä.

- (Koska määritelty satunnaismuuttuja X saa vain erillisiä arvoja, sitä sanotaan **diskreetiksi**. Satunnaismuuttuja X noudattaa **Bernoulli-jakaumaa**.)

Todennäköisyysjakauma: Määritelmä

- Satunnaismuuttujan X **todennäköisyysjakaumalla** (tai **jakaumalla**) tarkoitetaan kuvauksen

  $$X : S \rightarrow \mathbb{R}$$

  reaalilukujen joukkoon indusoimaa todennäköisyysmittaa.

- *Satunnaismuuttujan todennäköisyysjakauma* kuvaa koko otosavaruuden **todennäköisyysmassan** (= 1) *jakautumista* satunnaismuuttujan arvoalueelle.

Kommentteja

- *Satunnaisilmiöön liittyvät todennäköisyydet hallitaan täydellisesti*, jos satunnaisilmiön tulosvaihtoehtoja kuvaava satunnaismuuttuja ja sen todennäköisyys-jakauma *tunnetaan*:

  ***Kaikkien* satunnaisilmiöön liittyvien *tapahtumien todennäköisyydet* voidaan määrätä ilmiön tulosvaihto-ehtoja numeerisessa muodossa kuvaavan satunnais-muuttujan ja sen todennäköisyysjakauman avulla**.

Satunnaismuuttujien tyyppejä

- Satunnaismuuttuja määriteltiin edellä (mitallisena) *funktiona* otosavaruudesta reaalilukujen joukkoon.
- Mitalliset funktiot voivat olla funktioina hyvinkin *monimutkaisia*.
- Kaikissa tilastotieteen *tavanomaisissa* sovelluksissa tullaan kuitenkin yleensä hyvin toimeen seuraavien satunnaismuuttujien tyyppien kanssa:

  (1) **Diskreetit satunnaismuuttujat**.

  (2) **Jatkuvat satunnaismuuttujat**.
- Jatkossa rajoitutaan pelkästään diskreettien ja jatkuvien satunnaismuuttujien käsittelyyn.

Satunnaismuuttujat ja todennäköisyysjakaumat

Satunnaismuuttujat ja niiden todennäköisyysjakaumat

>> **Diskreetit satunnaismuuttujat ja niiden todennäköisyysjakaumat**

Jatkuvat satunnaismuuttujat ja niiden todennäköisyysjakaumat

Diskreetti satunnaismuuttuja: Määritelmä

- Olkoon otosavaruus S *äärellinen* tai *numeroituvasti ääretön*.
- Tällöin reaaliarvoinen funktio

$$X : S \to \mathbb{R}$$

joka saa *äärellisen* tai *numeroituvasti äärettömän* määrän erillisiä arvoja on **diskreetti satunnaismuuttuja**.

Diskreetit suureet

- Diskreetit satunnaismuuttujat liittyvät sellaisiin todennäköisyyslaskennan sovelluksiin, joissa tarkastellaan **diskreettejä suureita**.
- Esimerkkejä:
  - *laatuerot* (koodattuina numeerisiksi)
  - *luokittelut* ja *ryhmittelyt* (koodattuina numeerisiksi)
  - *järjestysluvut*
  - *lukumäärät*

Diskreetit satunnaismuuttujat
Esimerkkejä

- Kone tekee tuotetta n kpl päivässä. Satunnaisesti valittu tuote on viallinen todennäköisyydellä p. *Viallisten tuotteiden lukumäärä päivän aikana tehtyjen tuotteiden joukossa* on diskreetti satunnaismuuttuja, joka noudattaa **binomijakaumaa**.

- Pyydystetään järvestä joukko kaloja, merkitään ne ja päästetään takaisin järveen. Odotetaan niin kauan, että merkityt kalat ovat sekoittuneet merkitsemättömien kalojen joukkoon. Pyydystetään järvestä uusi joukko kaloja. *Merkittyjen kalojen lukumäärä uudessa pyynnissä* on diskreetti satunnaismuuttuja, joka noudattaa **hypergeometrista jakaumaa.**

- Palvelujonoon tulee keskimääriin k asiakasta aikayksikköä kohti. *Jonakin aikavälinä saapuvien asiakkaiden lukumäärä* on diskreetti satunnaismuuttuja, joka noudattaa **Poisson-jakaumaa**.

Pistetodennäköisyysfunktio: Määritelmä 1/2

- Olkoon otosavaruus S *äärellinen* tai *numeroituvasti ääretön*.
- Olkoot *diskreetin satunnaismuuttujan* $X : S \to \mathbb{R}$ saamat arvot eli *numeeriset tulosvaihtoehdot*:

  x_i , $i = 1, 2, \ldots, n$, jos S on äärellinen

  x_i , $i = 1, 2, 3, \ldots$, jos S on numeroituvasti ääretön
- Merkitään diskreetin satunnaismuuttujan X *arvojen joukkoa* kirjaimella T:

  $T = \{x_1, x_2, \ldots , x_n\}$, jos S on äärellinen

  $T = \{x_1, x_2, x_3, \ldots \}$, jos S on numeroituvasti ääretön

Pistetodennäköisyysfunktio: Määritelmä 2/2

- Reaaliarvoinen funktio f määrittelee *diskreetin satunnaismuuttujan* X **pistetodennäköisyysfunktion** , jos

  (1) $f(x_i) = \Pr(X = x_i)$ kaikille $x_i \in T$

  (2) $f(x_i) \geq 0$ kaikille $x_i \in T$

  (3) $\sum_{i \mid x_i \in T} f(x_i) = 1$

- Todennäköisyys

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

  on satunnaismuuttujan X arvoa x_i vastaava **piste-todennäköisyys**.

Pistetodennäköisyysfunktio: Toinen määritelmä

- Olkoot f diskreetin satunnaismuuttujan X *pistetodennäköisyysfunktio*, T sen *tulosvaihtoehtojen joukko* ja

$$f(x_i) = \Pr(X = x_i) = p_i \ , \ x_i \in T$$

satunnaismuuttujan X arvoa x_i vastaava *pistetodennäköisyys*.

- Satunnaismuuttujan X **pistetodennäköisyysfunktio** voidaan määritellä *kaikille reaaliluvuille* kaavalla

$$f(x) = \Pr(X = x) = \begin{cases} p_i, & x \in T \\ 0, & x \notin T \end{cases}$$

Pistetodennäköisyysfunktio: Kommentteja

- Jos *f* on diskreetin satunnaismuuttujan X pistetodennäköisyysfunktio, niin sanomme tavallisesti, että X **noudattaa diskreettiä todennäköisyysjakaumaa** *f* .
- Diskreetin satunnaismuuttujan X *pistetodennäköisyys-funktio f* kertoo, miten otosavaruuden *S todennäköisyys-massa* (= 1) *jakautuu* satunnaismuuttujan X *arvoille*.

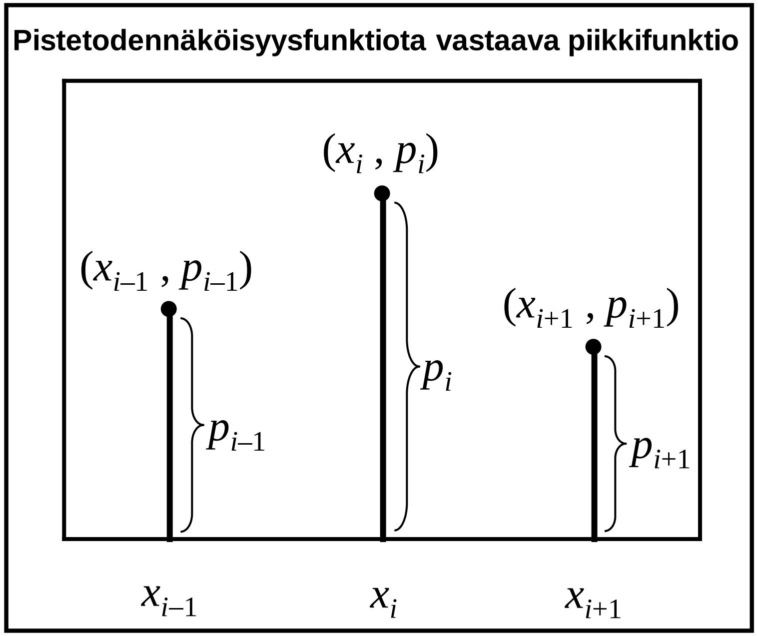
Pistetodennäköisyysfunktion kuvaaja: Havainnollistus

- Satunnaismuuttujan X pistetodennäköisyysfunktiota

  $$f(x_i) = \Pr(X = x_i) = p_i \ , \ x_i \in T$$

  kuvataan graafisesti ns. **piikkifunktiolla**, piirtämällä pisteisiin x_i ”*piikit*”, joiden pituudet vastaavat piste-todennäköisyyksiä

  $\Pr(\text{X} = x_i) = p_i$



Reaaliakselin välien todennäköisyydet

- Diskreetin jakauman tapauksessa välin $[a,b] \subset \mathbb{R}$ *todennäköisyys* on

$$\boxed{\Pr(a \le X \le b) = \sum_{i \,\mid\, x_i \in [a,b]} \Pr(X = x_i) = \sum_{i \,\mid\, x_i \in [a,b]} f(x_i)}$$

- Summassa *lasketaan yhteen* kaikki pistetodennäköisyydet

  $p_i = \Pr(X = x_i)$

  joita vastaavat satunnaismuuttujan X arvot $x_i \in [a, b]$.
- *Geometrisesti* ko. summan määrääminen merkitsee niiden piikkifunktion *piikkien pituuksien yhteenlaskemista*, joita vastaavat satunnaismuuttujan X arvot $x_i \in [a, b]$.

Esimerkki: Diskreetti tasainen jakauma

- Olkoon X *diskreetti satunnaismuuttuja*, jonka mahdolliset arvot ovat $x_1, x_2, \ldots, x_n$.
- Oletetaan, että satunnaismuuttujan X mahdollisiin arvoihin $x_1, x_2, \ldots, x_n$ liittyvät todennäköisyydet ovat *yhtä suuria*:

$$\Pr(X = x_k) = \frac{1}{n}, \qquad k = 1, 2, \ldots, n$$

- Sanomme, että satunnaismuuttuja X noudattaa **diskreettiä tasaista jakaumaa**. Esimerkiksi virheettömän nopan heiton tulos X on diskreetti satunnaismuuttuja, joka noudattaa diskreettiä tasaista jakaumaa.

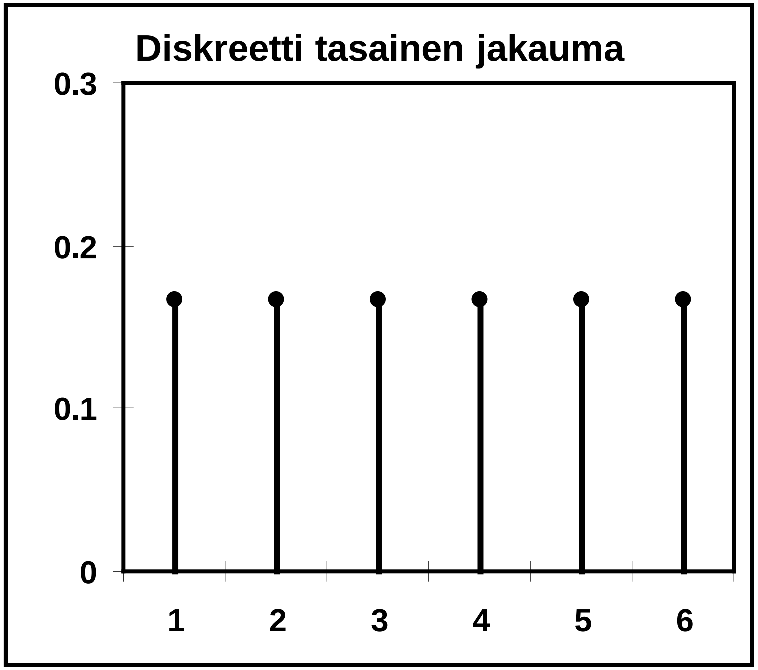
Diskreetti tasainen jakauma

Pistetodennäköisyysfunktion kuvaaja

- Kuva oikealla esittää diskreetin tasaisen jakauman

$$f(x) = \frac{1}{6},\ x = 1, 2, 3, 4, 5, 6$$

pistetodennäköisyysfunktiota.



Satunnaismuuttujat ja todennäköisyysjakaumat

Satunnaismuuttujat ja niiden todennäköisyysjakaumat

Diskreetit satunnaismuuttujat ja niiden todennäköisyysjakaumat

>> Jatkuvat satunnaismuuttujat ja niiden todennäköisyysjakaumat

Jatkuva satunnaismuuttuja: Määritelmä

- Satunnaismuuttuja X on **jatkuva**, jos seuraavat kaksi ehtoa pätevät:

  (1) Satunnaismuuttuja X saa *kaikki* reaalilukuarvot joltakin *reaaliakselin väliltä*.

  (2) Todennäköisyys, että satunnaismuuttuja X saa *minkä tahansa yksittäisen arvon* = 0.

Jatkuvat suureet

- Jatkuvat satunnaismuuttujat liittyvät sellaisiin todennäköisyyslaskennan sovelluksiin, joissa tarkastellaan **jatkuvia suureita**.
- Esimerkkejä:
  - *aika*
  - *nopeus*
  - *pituus, pinta-ala, tilavuus*
  - *paino*
  - *lämpötila*
  - *rahamäärä, korko*

Jatkuvat satunnaismuuttujat
Esimerkkejä:

- Palvelujonoon tulee asiakkaita keskimääriin *k* kappaletta aikayksikköä kohden. *Odotusaika seuraavan asiakkaan tulolle jonoon* on jatkuva satunnaismuuttuja, joka noudattaa **eksponenttijakaumaa**.

- Viisivuotiaiden tyttöjen pituus on jatkuva satunnaismuuttuja, jonka voidaan sanoa noudattavan approksimatiivisesti **normaalijakaumaa**.

Tiheysfunktio: Määritelmä

- Reaaliarvoinen funktio *f* määrittelee (**todennäköisyys-**) **tiheysfunktion** *jatkuvalle satunnaismuuttujalle* X, jos

  (1) $f(x)$ on x:n jatkuva funktio

  (2) $f(x) \geq 0$ kaikille x

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

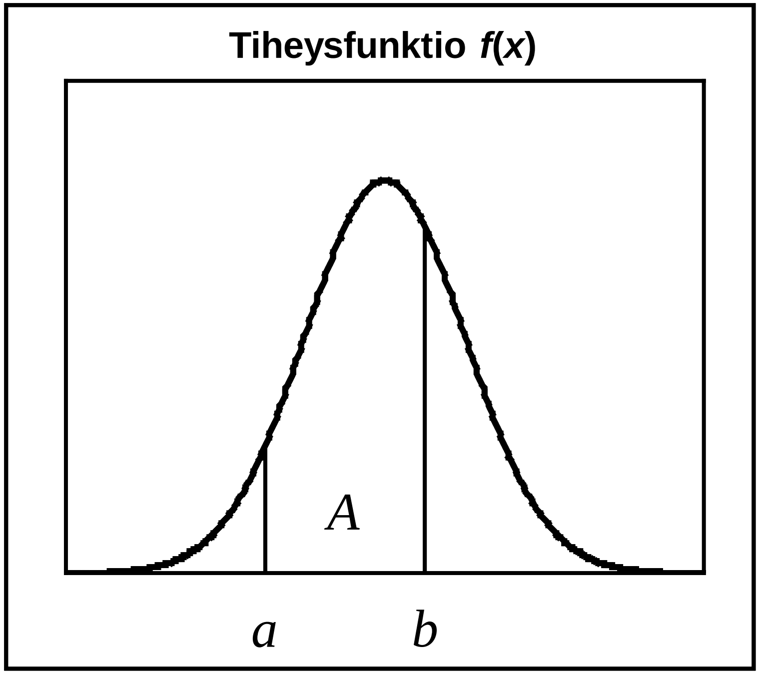
  (4) $\boxed{\Pr(a \leq X \leq b) = \int_a^b f(x)dx}$

- Jatkuvan satunnaismuuttujan X *tiheysfunktio f* määrittelee satunnaismuuttujan X todennäköisyysjakauman.

Tiheysfunktion kuvaaja: Havainnollistus

- Kuva oikealla esittää **normaalijakaumaa** noudattavan *jatkuvan satunnaismuuttujan* X *tiheysfunktiota* $f(x)$.
- Jatkuville satunnaismuuttujille X pätee yleisesti:

$$\Pr(a \leq X \leq b) = \int_a^b f(x)dx = \text{Alueen } A \text{ pinta-ala}$$



Esimerkki: Jatkuva tasainen jakauma

- 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 ja sen tiheysfunktio

- Satunnaismuuttujan X **tiheysfunktio** on

$$f(x) = \begin{cases} 0\,, \; x < a \\ \dfrac{1}{b-a}\,, \; a \le x \le b \\ 0\,, \; x > b \end{cases}$$

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

Jatkuva tasainen jakauma

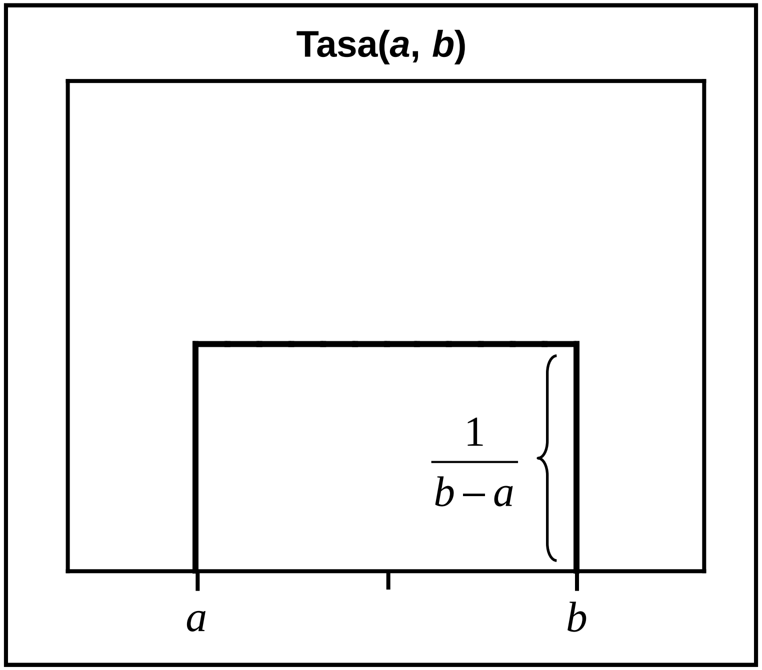
Tiheysfunktion kuvaaja

- Kuva oikealla esittää jatkuvan tasaisen jakauman

  Tasa(a, b)

  tiheysfunktiota

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



Todennäköisyyksien määrääminen jatkuvasta tasaisesta jakaumasta

- Olkoon $X \sim \mathrm{Tasa}(a, b)$.
- Olkoon $[c, d] \subset [a, b]$ jokin välin $[a, b]$ *osaväli*.
- *Välin* $[c, d]$ *todennäköisyys* saadaan integroimalla jatkuvan tasaisen jakauman $\mathrm{Tasa}(a, b)$ tiheysfunktio

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

  välillä $[c, d]$:

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

<u>Diskreetit</u> jakaumat *vs* <u>jatkuvat</u> jakaumat 1/2

- Olkoon X <u>**diskreetti**</u> **satunnaismuuttuja** ja *f* vastaava **pistetodennäköisyysfunktio**.
- Olkoon $\{x_1, x_2, x_3, \ldots\}$ satunnaismuuttujan X *arvojen* joukko.
- Pistetodennäköisyysfunktion *f* arvo pisteessä x_i *on todennäköisyys*:

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

- Koska funktion *f* arvo pisteessä x_i *on* todennäköisyys, niin

$$0 \leq f(x_i) \leq 1$$

Diskreetit jakaumat *vs* jatkuvat jakaumat 2/2

- Olkoon X **jatkuva satunnaismuuttuja** ja *f* vastaava **tiheysfunktio**.

- Reaaliakselin *välien* todennäköisyydet saadaan kaavasta

$$\Pr(a \le X \le b) = \int_a^b f(x)dx$$

- Koska funktion *f* arvo pisteessä *x* *ei ole* todennäköisyys, *on mahdollista*, että $f(x) > 1$.

Todennäköisyyslaskenta

Osa 2: Satunnaismuuttujat ja todennäköisyysjakaumat

- Satunnaismuuttujat ja todennäköisyysjakaumat
- **Kertymäfunktio**

Kertymäfunktion määritelmä

- Satunnaismuuttujan X **kertymäfunktio** *F*

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

kuvaa satunnaismuuttujan X *todennäköisyysmassan kertymistä*, kun funktion argumentti *x kasvaa*.

- Satunnaismuuttujan X kertymäfunktio määrää *kaikkien* ko. satunnaisilmiöön liittyvien *tapahtumien todennäköisyydet*.

Kertymäfunktion ominaisuudet 1/2

- Funktio $F : \mathbb{R} \to [0,1]$ on **kertymäfunktio**, jos ja vain jos se toteuttaa seuraavat ehdot:

(1) $\lim_{x \to -\infty} F(x) = 0$

(2) $\lim_{x \to +\infty} F(x) = 1$

(3) F on *ei-vähenevä*:

$F(x_1) \le F(x_2)$, jos $x_1 \le x_2$

(4) F on *jatkuva oikealta*:

$\lim_{h \to 0+} F(x+h) = F(x)$

Kertymäfunktion ominaisuudet 2/2

- Edellisen lisäksi pätee :

$$(5) \quad \Pr(X > x) = 1 - F(x)$$

$$(6) \quad \Pr(a < X \leq b) = F(b) - F(a)$$

Kertymäfunktio

Kertymäfunktio: Määritelmä

>> Diskreettien jakaumien kertymäfunktiot

Jatkuvien jakaumien kertymäfunktiot

Diskreetin jakauman kertymäfunktion määritelmä

- Olkoon X diskreetti satunnaismuuttuja jonka arvojen joukko on $\{x_1, x_2, x_3, \ldots\}$ ja pistetodennäköisyydet ovat

$$\Pr(X = x_i) = p_i \ , i = 1, 2, 3, \ldots$$

- Määritellään *diskreetin satunnaismuuttujan* X **kertymäfunktio** kaavalla

$$F(x) = \Pr(X \le x) = \sum_{i \,|\, x_i \le x} p_i$$

- Huom. Diskreetin jakauman kertymäfunktion ja pistetodennäköisyysfunktion yhteys:

$$f(x_i) = \Pr(X = x_i) = p_i = F(x_i) - F(x_{i-1})$$

Diskreetin jakauman kertymäfunktio on porrasfunktio

- Diskreetin satunnaismuuttujan kertymäfunktio F on *epäjatkuva ei-vähenevä* funktio, jolla on *epäjatkuvuus-kohta* eli *hyppäys* jokaisessa pisteessä x_i , johon liittyy *positiivinen* todennäköisyys

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

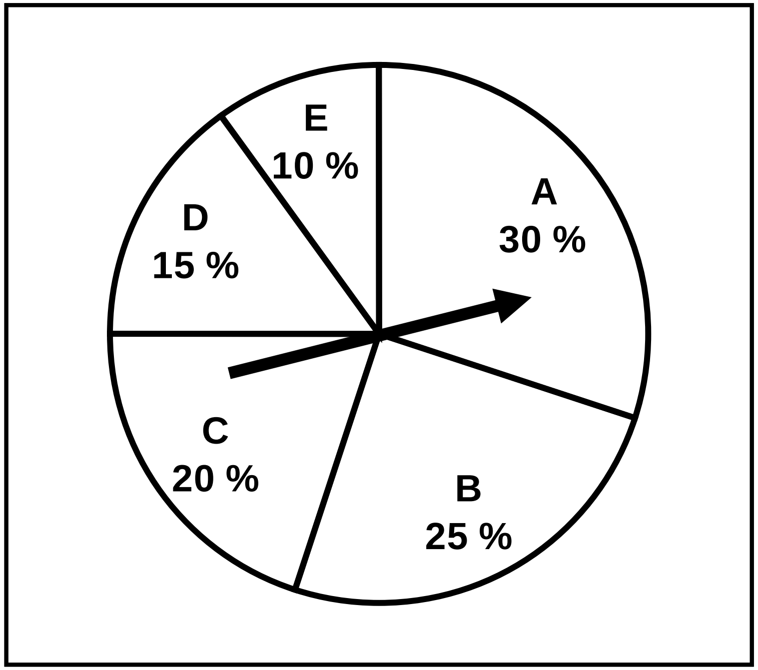
- *Hyppäyksen suuruus* pisteessä x_i on p_i .
- Kertymäfunktio saa *vakioarvon* peräkkäisten pisteiden x_{i-1} ja x_i *välissä*.
- Diskreetin satunnaismuuttujan kertymäfunktio on siten **porrasfunktio**, jossa todennäköisyydet p_i määräävät *askelmien korkeudet* ja erotukset $x_i - x_{i-1}$ määräävät *askelmien syvyydet*.

Esimerkki: Onnenpyörä 1/4

- Onnenpyörän pinta on jaettu viiteen sektoriin

  A, B, C, D, E
- Sektoreiden pinta-alojen osuudet onnenpyörän kokonaispinta-alasta on esitetty alla:

Sektori	%
A	30
B	25
C	20
D	15
E	10
Summa	100



Esimerkki: Onnenpyörä 2/4

- Määritellään diskreetti satunnaismuuttuja X, joka liittää tulosvaihtoehtoihin

  A, B, C, D, E

  reaaliluvut seuraavalla tavalla:

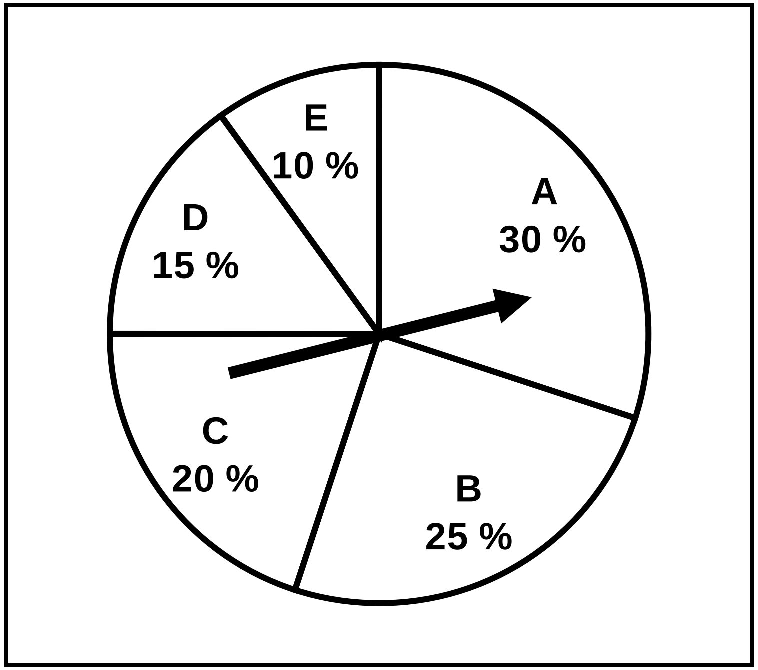
  A $\to$ 1

  B $\to$ 2

  C $\to$ 3

  D $\to$ 4

  E $\to$ 5



Diskreettien jakaumien kertymäfunktiot

Esimerkki: Onnenpyörä 3/4

- Satunnaismuuttujan X piste-todennäköisyysfunktio f voidaan määritellä seuraavasti:

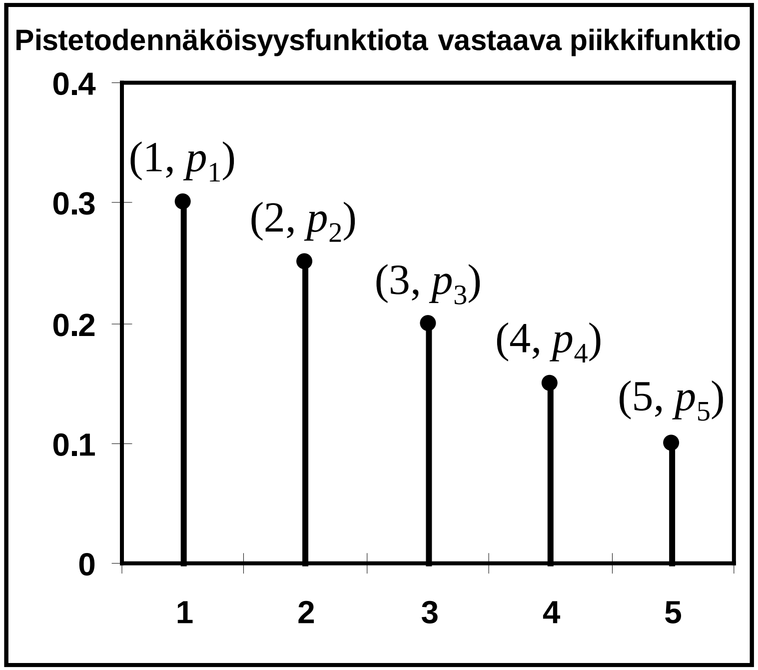
  $f(1) = \Pr(X = 1) = 0.30 = \Pr(A)$

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

  $f(3) = \Pr(X = 3) = 0.20 = \Pr(C)$

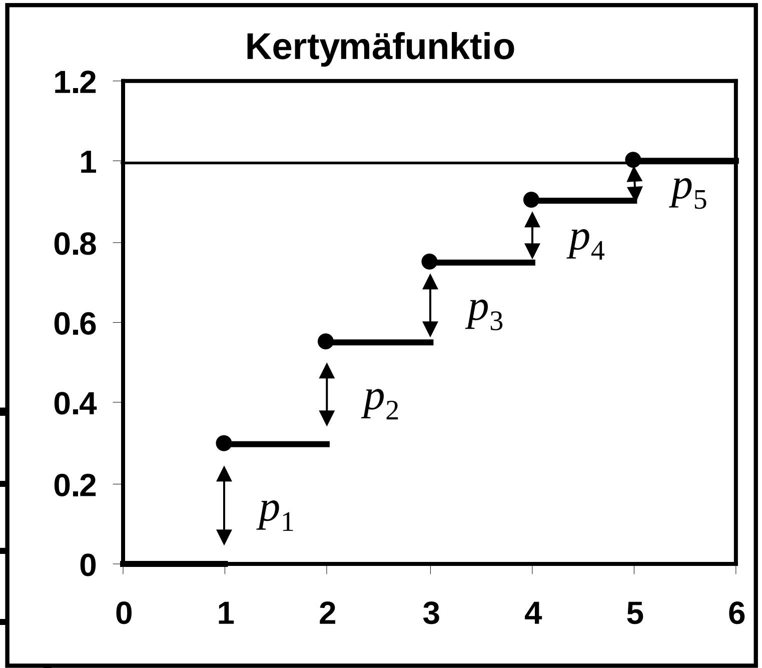
  $f(4) = \Pr(X = 4) = 0.15 = \Pr(D)$

  $f(5) = \Pr(X = 5) = 0.10 = \Pr(E)$



Esimerkki: Onnenpyörä 4/4

- Satunnaismuuttujan X *kertymäfunktio F* voidaan määritellä alla olevan taulukon avulla.
- Kuva oikealla esittää esimerkin kertymäfunktion kuvaajaa.



	$F(x) = \Pr(X \le x)$
$x < 1$	0
$1 \le x < 2$	$p_1 = 0.3$
$2 \le x < 3$	$p_1 + p_2 = 0.55$
$3 \le x < 4$	$p_1 + p_2 + p_3 = 0.75$
$4 \le x < 5$	$p_1 + p_2 + p_3 + p_4 = 0.9$
$5 \le x$	$p_1 + p_2 + p_3 + p_4 + p_5 = 1$

Kertymäfunktio

Kertymäfunktio: Määritelmä

Diskreettien jakaumien kertymäfunktiot

\>> Jatkuvien jakaumien kertymäfunktiot

Jatkuvan jakauman kertymäfunktion määritelmä

- Olkoon X jatkuva satunnaismuuttuja.
- Olkoon satunnaismuuttujan X *tiheysfunktio* $f(x)$.
- Määritellään *jatkuvan satunnaismuuttujan* X **kertymäfunktio** kaavalla

$$F(x) = \Pr(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

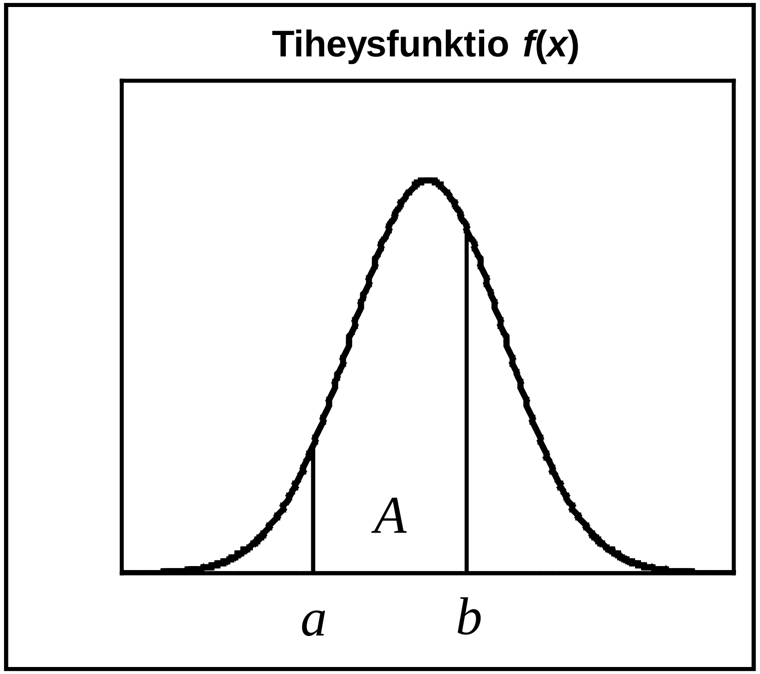
- Huom. Jatkuvan jakauman kertymäfunktion ja tiheysfunktion yhteys:

$$f(x) = F'(x) = \frac{d}{dx}F(x)$$

Jatkuvan jakauman tiheysfunktio ja välien todennäköisyydet: Havainnollistus

- Olkoon $f(x)$ *jatkuvan satunnaismuuttujan X tiheysfunktio*.
- Tällöin:

$$\Pr(a \leq X \leq b)$$
$$= \int_a^b f(x)dx$$
$$= \text{Alueen } A \text{ pinta-ala}$$

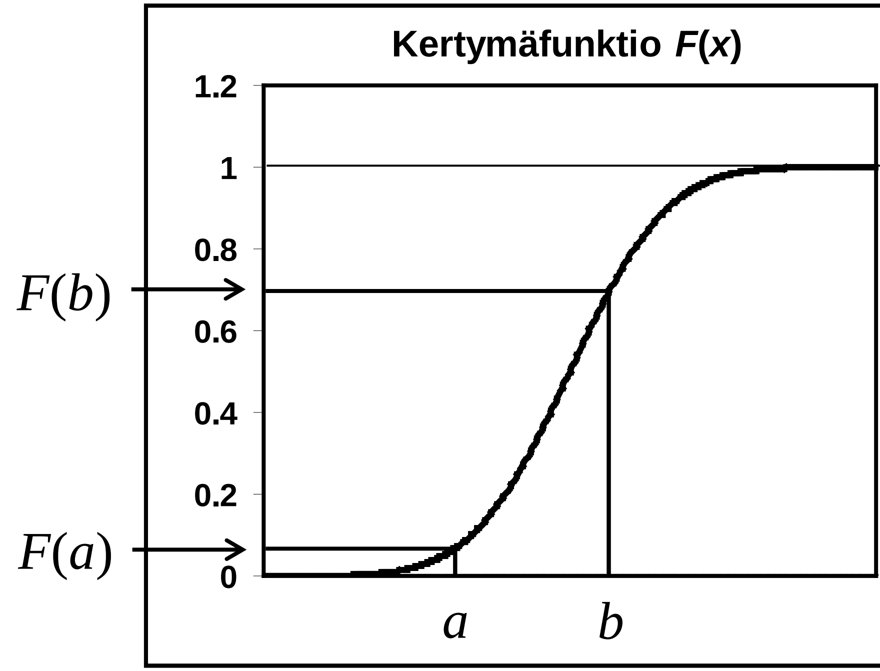


- Kuva oikealla esittää **normaali-jakauman** *tiheysfunktiota*.

Jatkuvan jakauman kertymäfunktio ja välien todennäköisyydet: Havainnollistus

- Olkoon $F(x)$ *jatkuvan satunnaismuuttujan* X *kertymäfunktio* ja $f(x)$ sen *tiheysfunktio*.
- Tällöin:

$$\begin{aligned}&\Pr(a \le X \le b) \\ &= F(b) - F(a) \\ &= \int_a^b f(x)\,dx\end{aligned}$$



- Kuva oikealla esittää **normaali-jakauman** *kertymäfunktiota*.

Todennäköisyyslaskenta

Osa 2: Satunnaismuuttujat ja todennäköisyysjakaumat

- **Jakaumien tunnusluvut**

Jakaumien tunnusluvut

>> Odotusarvo

Varianssi

Odotusarvo

Diskreetin jakauman odotusarvo: Määritelmä

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

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

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

$$\boxed{\mathrm{E}(X) = \sum_i x_i f(x_i)}$$

Odotusarvolle voi myös käyttää merkintää μ_X .

Diskreetin jakauman odotusarvo: Kommentteja

- Vaikka satunnaismuuttujan saama arvo vaihtelee satunnaisesti ko. satunnaisilmiön toistokerrasta toiseen, satunnaismuuttuja saa *keskimäärin* arvoja, jotka vaihtelevat sen odotusarvon ympärillä.

- *Jos jakaumalla on odotusarvo*, se on jakauman toden-näköisyysmassan **painopiste**.

- *Diskreetin* jakauman odotusarvon *ei tarvitse kuulua* ko. satunnaismuuttujan tulosvaihtoehtojen joukkoon.

Diskreetin jakauman odotusarvo: Esimerkki nopanheitosta

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

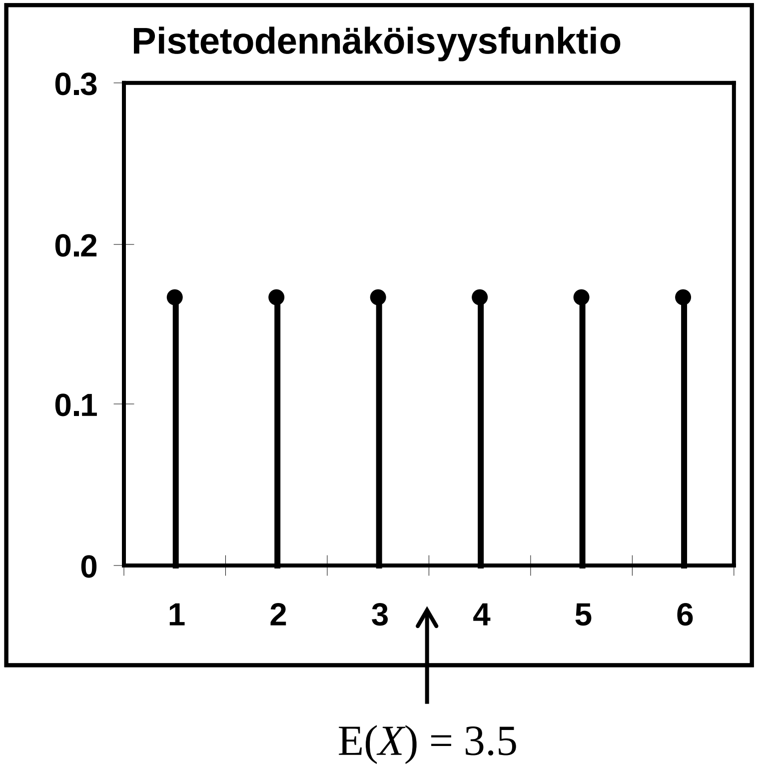
$$\Pr(X = i) = \frac{1}{6}, \; i = 1, 2, 3, 4, 5, 6$$

- Satunnaismuuttujan X *odotusarvo*:

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

$$= \frac{1+2+3+4+5+6}{6} = \frac{21}{6}$$

$$= 3.5$$



Jatkuvan jakauman odotusarvo: Määritelmä

- Olkoon X *jatkuva satunnaismuuttuja*.
- Olkoon satunnaismuuttujan X *tiheysfunktio* $f(x)$.

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

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

Jatkuvan jakauman odotusarvo: Kommentteja

- Vaikka satunnaismuuttujan saama 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.

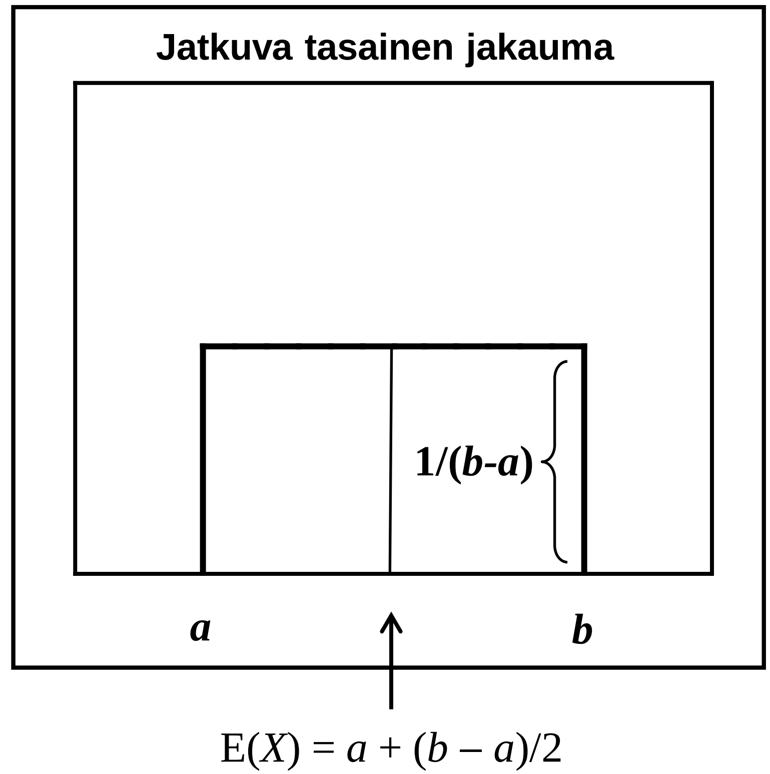
Jatkuvan jakauman odotusarvo: Esimerkki tasaisesta jakaumasta

- **Jatkuvan tasaisen jakauman** *tiheysfunktio* on

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

- Jakauman *odotusarvo* on

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



Jatkuvan jakauman odotusarvo: Esimerkki kolmiojakaumasta

- Erään **kolmiojakauman** *tiheysfunktio* on

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

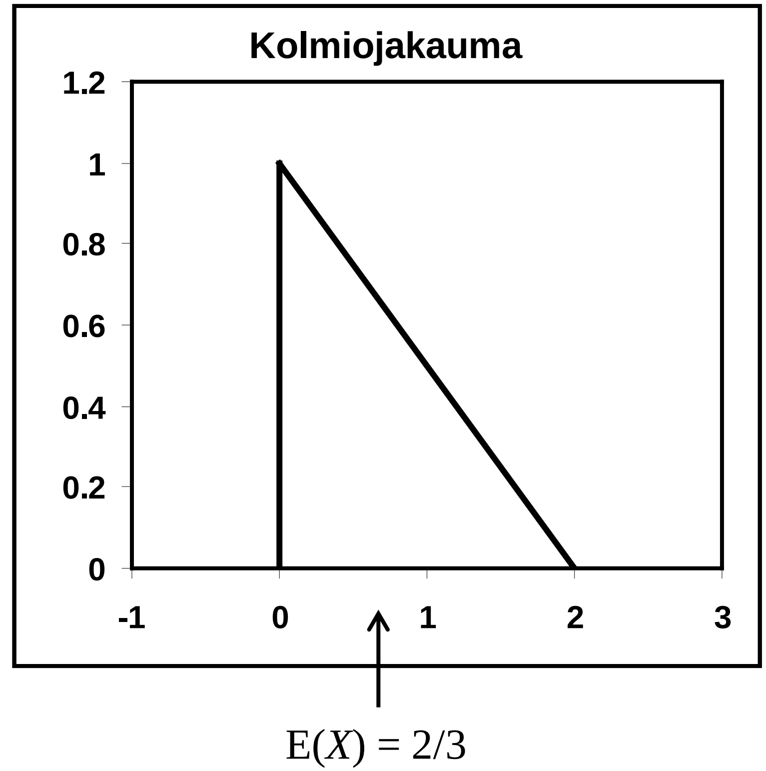
- Jakauman *odotusarvo* on

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

$$= \int_0^2 x\left(-\tfrac{1}{2}x + 1\right)dx$$

$$= \left[-\tfrac{1}{6}x^3 + \tfrac{1}{2}x^2\right]_0^2$$

$$= \frac{2}{3}$$



Jatkuvan jakauman odotusarvo: Esimerkki normaalijakaumasta

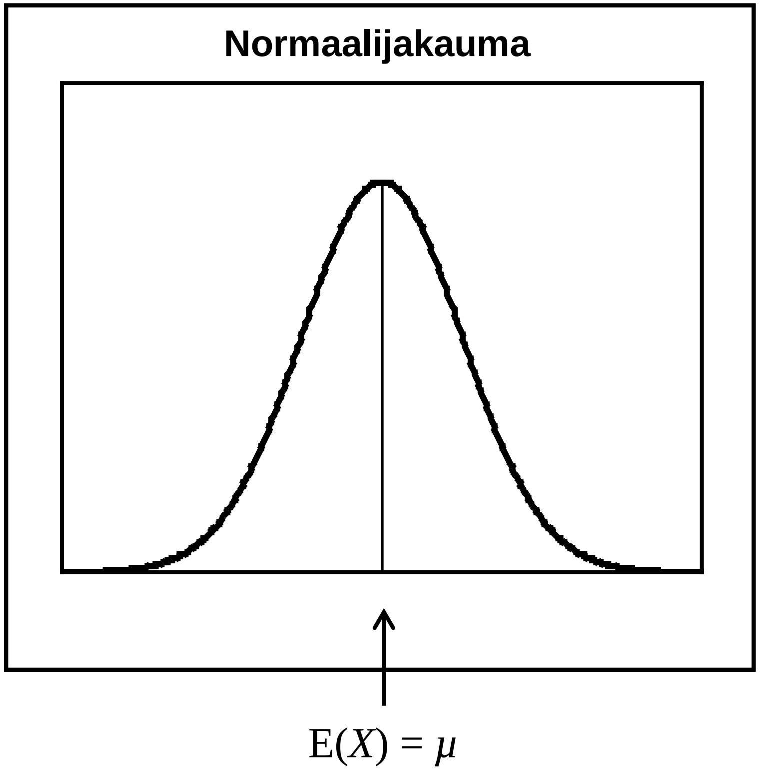
- **Normaalijakauman** *tiheysfunktio* on

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

- Normaalijakauman tiheysfunktio on *symmetrinen* pisteen $x = \mu$ suhteen.
- Voidaan osoittaa, että normaalijakauman *odotusarvo*

$\mathrm{E}(x) = \mu$

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

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

  satunnaismuuttujan X odotusarvo.
- Jos satunnaismuuttujan X jakauma on *symmetrinen* pisteen

  $$x = a$$

  suhteen, niin

  $$\mu = a$$

Odotusarvon ominaisuuksia

- Vakion odotusarvo on vakio itse, koska se ei vaihtele koetoistosta toiseen:

$$\boxed{\mathrm{E}(a) = a}$$

- Olkoon satunnaismuuttujan X odotusarvo $\mathrm{E}(X)$. Satunnaismuuttujan X lineaarimuunnokselle $Y = a + bX$ (a ja b *vakioita*) pätee:

$$\boxed{\mathrm{E}(Y) = a + b\,\mathrm{E}(X)}$$

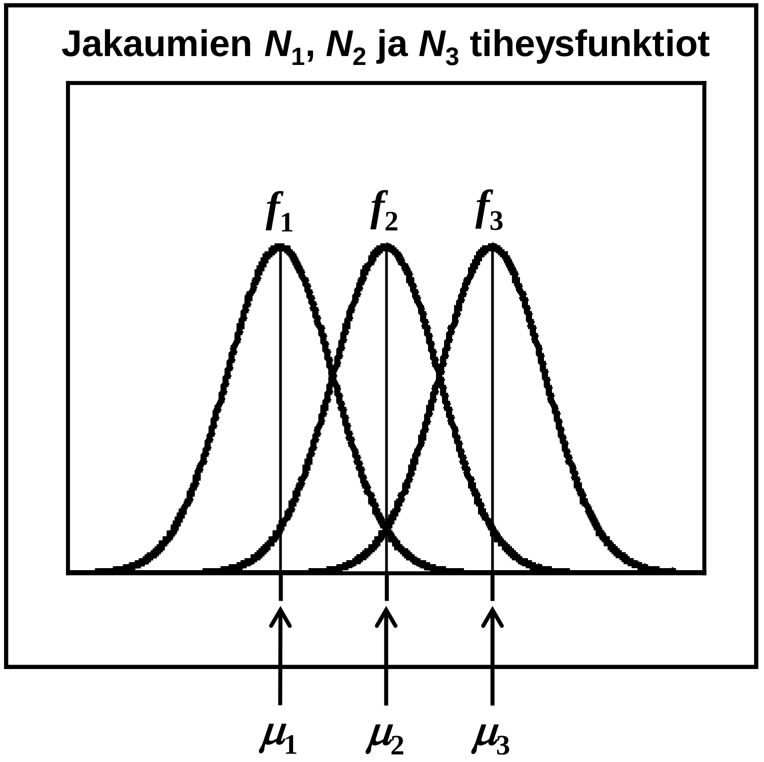
Lineaarimuunnoksen odotusarvo: Kommentteja

- 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: 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* pisteiden $x = \mu_1$, $x = \mu_2$ 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 μ_1, μ_2 ja μ_3 toteuttavat epäyhtälöt

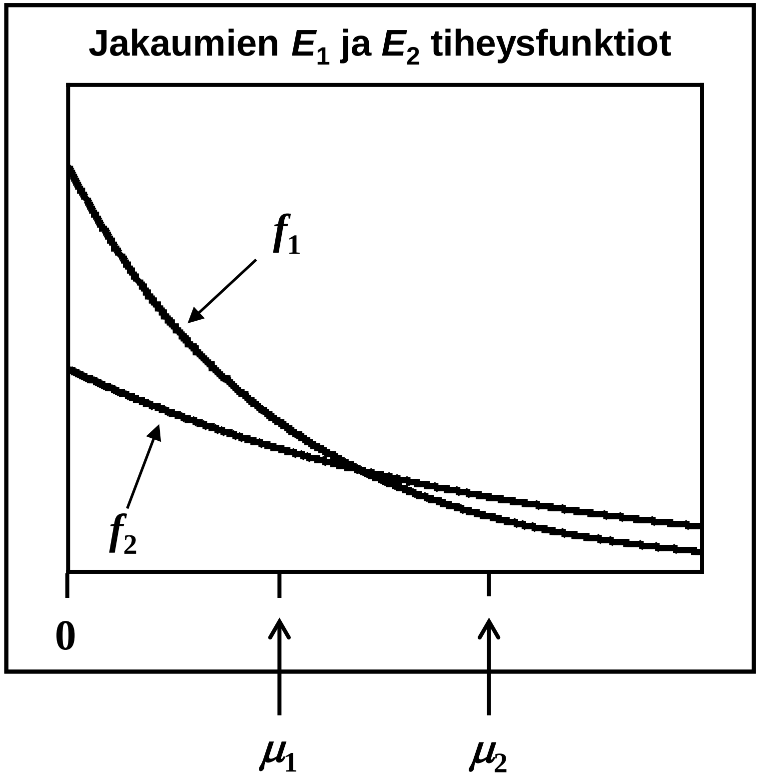
  $\mu_1 < \mu_2 < \mu_3$



Odotusarvo jakauman sijaintiparametrina: Havainnollistus 2

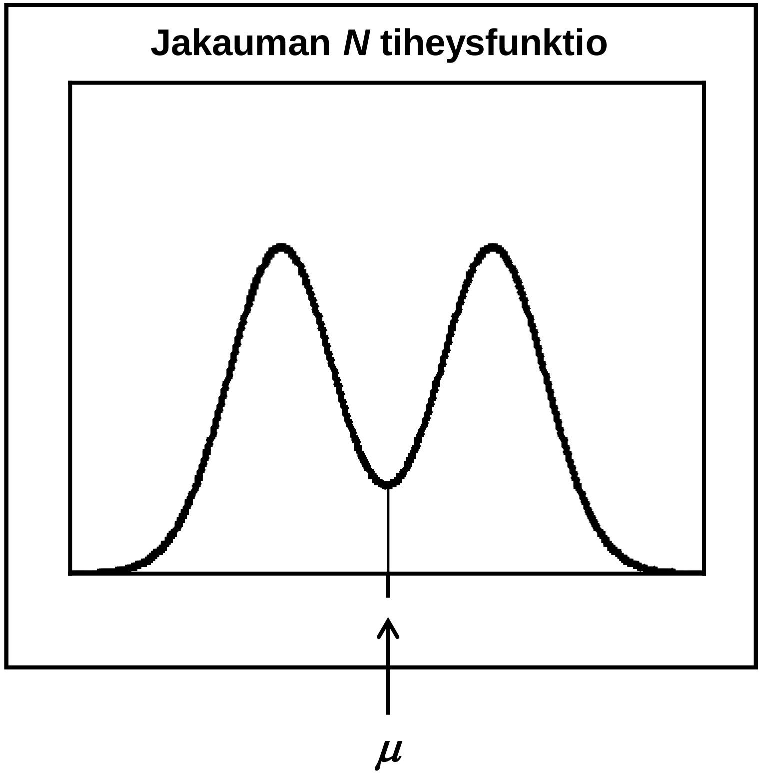
- Kuva oikealla esittää kahden **eksponenttijakauman** E_1 ja E_2 *tiheysfunktioita* f_1 ja f_2 .
- Tiheysfunktiot f_1 ja f_2 ovat *yksihuippuisia* ja *epä-symmetrisiä*.
- Jakauman E_1 todennäköisyys-massa *on keskittynyt* jakauman E_2 todennäköisyysmassaa *voi-makkaammin origon lähelle*.
- Jakaumien E_1 ja E_2 odotusarvot μ_1 ja μ_2 toteuttavat epäyhtälön

  $\mu_1 < \mu_2$



Odotusarvo jakauman sijaintiparametrina: Havainnollistus 3

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



Summan ja erotuksen odotusarvo

- Kahden satunnaismuuttujien *summalle* ja *erotukselle* pätee:

$$\boxed{\begin{aligned} \mathrm{E}(X+Y) &= \mathrm{E}(X)+\mathrm{E}(Y) \\ \mathrm{E}(X-Y) &= \mathrm{E}(X)-\mathrm{E}(Y) \end{aligned}}$$

- Yleisesti satunnaismuuttujien X_i , $i = 1, 2, \dots , n$ **painotetulle summalle** pätee:

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

Diskreetin satunnaismuuttujan funktion odotusarvo: Määritelmä

- Olkoon X diskreetti satunnaismuuttuja, jonka piste-todennäköisyysfunktio on

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

- Olkoon g reaaliarvoinen funktio.

- Satunnaismuuttujan $g(X)$ odotusarvo on vakio

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

Jatkuvan satunnaismuuttujan funktion odotusarvo: Määritelmä

- Olkoon X jatkuva satunnaismuuttuja, jonka tiheysfunktio on $f(x)$.
- Olkoon g reaaliarvoinen funktio.
- Satunnaismuuttujan $g(X)$ odotusarvo on vakio

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

Jakaumien tunnusluvut

Odotusarvo

>> Varianssi

Varianssi:
Yleinen määritelmä

- Olkoon satunnaismuuttujan *X* odotusarvo

$$\mathrm{E}(X) = \mu_X$$

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

$$\boxed{\mathrm{Var}(X) = \mathrm{E}\left[(X - \mu_X)^2\right]}$$

- Käytetään myös merkintää $\mathrm{D}^2(X)$ ja σ_X^2 .
- Satunnaismuuttujan *X* varianssi mittaa *X*:n arvojen vaihtelun laajuutta odotusarvon ympäri .

Diskreetin jakauman varianssi

- Olkoon X *diskreetti satunnaismuuttuja*.
- Olkoon $\{x_1, x_2, \ldots\}$ satunnaismuuttujan X *tulos-vaihtoehtojen* eli arvojen joukko.
- Olkoon satunnaismuuttujan X *pistetodennäköisyysfunktio*

$$f(x_i) = \Pr(X = x_i) = p_i \text{ , } i = 1, 2, \ldots$$

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

$$\boxed{\mathrm{Var}(X) = \sum_i (x_i - \mu_X)^2 f(x_i)}$$

Jatkuvan jakauman varianssi

- Olkoon X *jatkuva satunnaismuuttuja*.
- Olkoon satunnaismuuttujan X *tiheysfunktio* $f(x)$.

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

$$\boxed{\operatorname{Var}(X) = \int_{-\infty}^{+\infty} (x - \mu_X)^2 f(x)\,dx}$$

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.

Varianssin määritelmä: Kommentteja

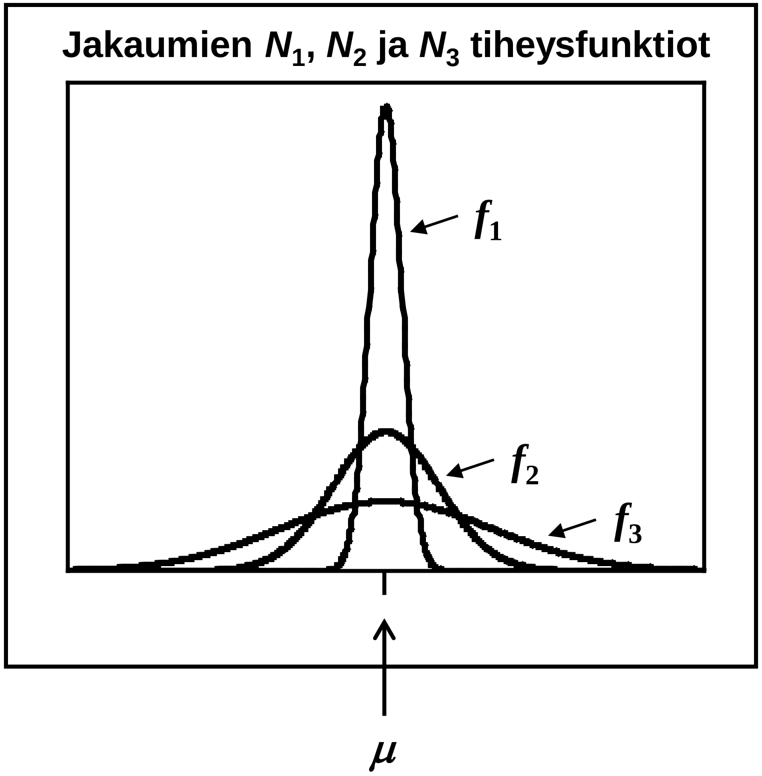
- Varianssi kuvaa todennäköisyysmassan **hajaantuneisuutta** tai – mikä on sama asia – **keskittyneisyyttä** jakauman *painopisteen* suhteen.
- Jos

  $$\mathrm{Var}(X) > \mathrm{Var}(Y)$$

  niin satunnaismuuttujan X todennäköisyysmassa on hajaantunut *voimakkaammin oman painopisteeseensä suhteen* kuin satunnaismuuttujan Y todennäköisyysmassa *oman painopisteeseensä suhteen*.
- Koska varianssi kuvaa todennäköisyysmassan hajaantuneisuutta, sitä voidaan kutsua **hajonta-parametriksi**.

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 μ.
- Tiheysfunktiot f_1, f_2 ja f_3 ovat *yksihuippuisia* ja *symmetrisiä* pisteen $x = \mu$ suhteen.

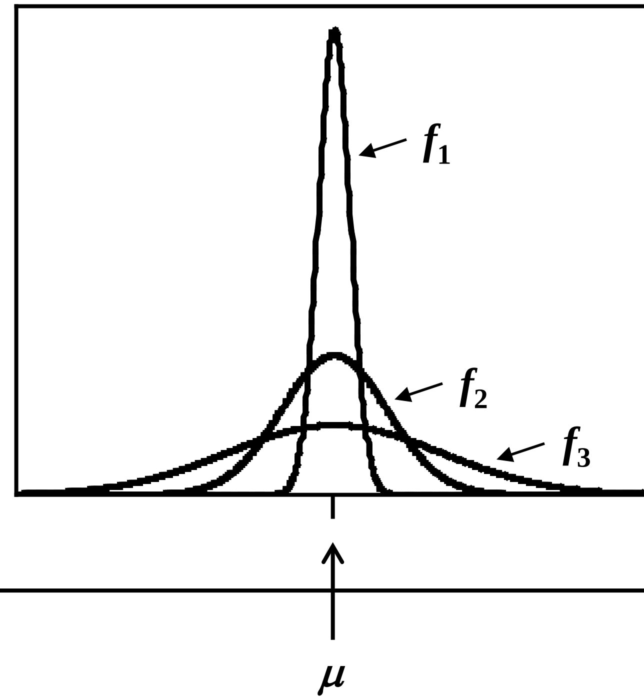


Varianssi jakauman hajaantuneisuuden mittana: Esimerkki normaalijakaumista 2/2

- Jakauman N_1 todennäköisyysmassa *on keskittynein*, kun taas jakauman N_3 todennäköisyysmassa on *hajaantunein*.
- Jakaumien varianssit toteuttavat epäyhtälöt:

  $\mathrm{Var}(X_1) < \mathrm{Var}(X_2) < \mathrm{Var}(X_3)$

Jakaumien N_1, N_2 ja N_3 tiheysfunktiot



Varianssin toinen laskukaava:

- Olkoot satunnaismuuttujan X *odotusarvo* $\mathrm{E}(X) = \mu$ ja *toinen momentti* $\mathrm{E}(X^2)$.

- Satunnaismuuttujan X *varianssi* on

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

Standardipoikkeama: Määritelmä

- Satunnaismuuttujan X **standardipoikkeama** on vakio

$$\boxed{\sqrt{\mathrm{Var}(X)}}$$

- Käytetään myös merkintää $\mathrm{D}(X)$ ja σ_X .
- Standardipoikkeamaa käytetään samaan tapaan kuin varianssia todennäköisyysmassan *hajaantuneisuuden* (*keskittyneisyyden*) mittana.
- Standardipoikkeama on – toisin kuin varianssi – samoissa *mittayksiköissä* kuin odotusarvo.

Varianssin ominaisuuksia

- **Vakion varianssi** on nolla, koska vakio ei vaihtele:

$$\boxed{\mathrm{Var}(a) = 0}$$

- Olkoon satunnaismuuttujan X varianssi Var(X). **Satunnaismuuttujan *X* lineaarimuunnoksen**

$$Y = a + bX$$

(*a* ja *b vakiota*) varianssi on

$$\boxed{\mathrm{Var}(Y) = b^2\,\mathrm{Var}(X)}$$

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.

Standardointi

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

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

odotusarvo

$$\mathrm{E}(Z) = 0$$

ja varianssi

$$Var(Z) = 1$$

Summan ja erotuksen varianssi 2/2

- Kahden riippumattomien* satunnaismuuttujien summalle ja erotukselle pätee:

$$\boxed{\begin{aligned}\mathrm{Var}(X+Y) &= \mathrm{Var}(X)+\mathrm{Var}(Y)\\ \mathrm{Var}(X-Y) &= \mathrm{Var}(X)+\mathrm{Var}(Y)\end{aligned}}$$

- Huomaa: $\mathrm{Var}(X-Y) \neq \mathrm{Var}(X)-\mathrm{Var}(Y)$

$$\mathrm{D}(X+Y) \neq \mathrm{D}(X)+\mathrm{D}(Y)$$

$$\mathrm{D}(X-Y) \neq \mathrm{D}(X)-\mathrm{D}(Y)$$

*Satunnaismuuttujien X ja Y riippumattomuuden käsite täsmennetään luvussa **Moniulotteiset satunnaismuuttujat ja jakaumat**.

Diskreetin jakauman varianssi: Esimerkki nopanheitosta 1/2

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

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

- Satunnaismuuttujan *X odotusarvo*:

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

- Satunnaismuuttujan *X toinen momentti*:

$$\mathrm{E}(X^2) = \sum_{i=1}^{6} i^2 \Pr(X = i) = \sum_{i=1}^{6} i^2 \frac{1}{6}$$

$$= \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2}{6} = \frac{91}{6}$$

Diskreetin jakauman varianssi: Esimerkki nopanheitosta 2/2

- Satunnaismuuttujan *X varianssi*:

$$Var(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{91}{6} - \left(\frac{21}{6}\right)^2 = \frac{35}{12} \approx 2.917$$

- Satunnaismuuttujan *X standardipoikkeama* eli *keskihajonta*:

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

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

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

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

- Satunnaismuuttujan X *odotusarvo*:

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

- Satunnaismuuttujan X *toinen momentti*:

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

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

- Satunnaismuuttujan X *varianssi*:

$$Var(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{b^2}{3} - \left(\frac{b}{2}\right)^2 = \frac{b^2}{12}$$

- Satunnaismuuttujan X *standardipoikkeama*:

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

Todennäköisyyslaskenta

Osa 3: Todennäköisyysjakaumia

- Diskreettejä jakaumia

Diskreettejä jakaumia

>> **Diskreetti tasainen jakauma**
   Bernoulli-jakauma
   Binomijakauma

Diskreetti tasainen jakauma 1/2

- Olkoon X *diskreetti satunnaismuuttuja*, jonka mahdolliset arvot ovat $x_1, x_2, \dots , x_n$.
- Oletetaan, että satunnaismuuttujan X mahdollisiin arvoihin $x_1, x_2, \dots , x_n$ liittyvät todennäköisyydet ovat *yhtä suuria*:

$$\Pr(X = x_k) = \frac{1}{n}, \qquad k = 1, 2, \dots, n$$

- Sanomme, että satunnaismuuttuja X noudattaa **diskreettiä tasaista jakaumaa**. Esimerkiksi virheettömän nopan heiton tulos X on diskreetti satunnaismuuttuja, joka noudattaa diskreettiä tasaista jakaumaa.

Diskreetti tasainen jakauma 2/2

- **Pistetodennäköisyysfunktio**:

$$f(x) = \Pr(X = x) = \frac{1}{n}, \qquad x = x_k, \quad k = 1, 2, \ldots, n$$

- **Odotusarvo**:

$$\mathrm{E}(X) = \mu_X = \frac{1}{n}\sum_{k=1}^{n} x_k = \overline{x}$$

- **Varianssi**:

$$\mathrm{Var}(X) = \mathrm{D}^2(X) = \sigma_X^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})^2$$

Diskreetti tasainen jakauma
Pistetodennäköisyysfunktion kuvaaja

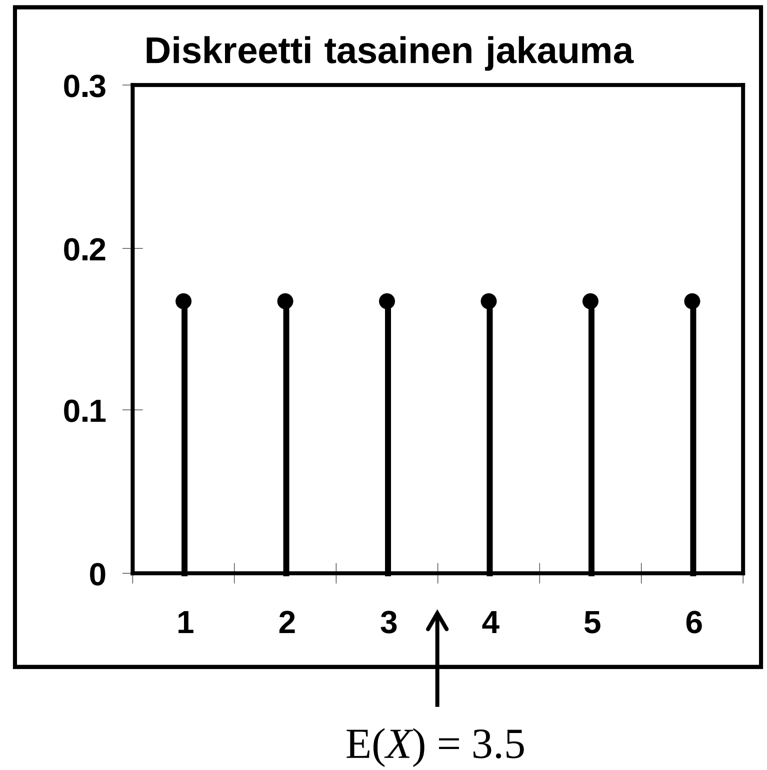
- Kuva oikealla esittää diskreetin tasaisen jakauman

$$f(x) = \frac{1}{6},\ x = 1, 2, 3, 4, 5, 6$$

pistetodennäköisyysfunktiota.

- Jakauman *odotusarvo*:

$$\mathrm{E}(X) = \frac{1}{6}\sum_{k=1}^{6} k = 3.5$$



Diskreettejä jakaumia

Diskreetti tasainen jakauma

>> Bernoulli-jakauma

Binomijakauma

Bernoulli-jakauma 1/2

- Olkoon A otosavaruuden S *tapahtuma* ja olkoon $\Pr(A) = p$.
- Tällöin tapahtuman A *komplementtitapahtuman* A^c *todennäköisyys* on

$$\Pr(A^c) = 1 - \Pr(A) = 1 - p = q$$

- Määritellään *diskreetti satunnaismuuttuja* X:

$$X = \begin{cases} 1, & \text{jos tapahtuma } A \text{ sattuu} \\ 0, & \text{jos tapahtuma } A \text{ ei satu} \end{cases}$$

- Tällöin satunnaismuuttujan X *jakauma* on

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

$$\Pr(X = 0) = 1 - p = q$$

Bernoulli-jakauma 2/2

- Sanomme, että satunnaismuuttuja X noudattaa **Bernoulli-jakaumaa parametrinaan *p***.
- Merkintä:

$$X \sim \text{Bernoulli}(p) = \text{B(p)}$$

- Satunnaismuuttujan X **pistetodennäköisyysfunktio** on

$$f(x) = \Pr(X = x) = p^x q^{1-x}$$

missä $0 < p < 1$, $q = 1 - p$ ja $x = 0,1$.

Odotusarvo, varianssi ja standardipoikkeama

- Olkoon

$$X \sim \mathrm{Bernoulli}(p)$$

- **Odotusarvo**:

$$\mathrm{E}(X) = p$$

- **Varianssi** ja **standardipoikkeama**:

$$\mathrm{Var}(X) = \mathrm{D}^2(X) = pq$$

$$\mathrm{D}(X) = \sqrt{pq}$$

Bernoulli-jakauma

Pistetodennäköisyysfunktion kuvaaja

- Kuva oikealla esittää Bernoulli-jakauman

  B(0.8)

  pistetodennäköisyysfunktiota

  $$f(x) = p^x q^{1-x}$$

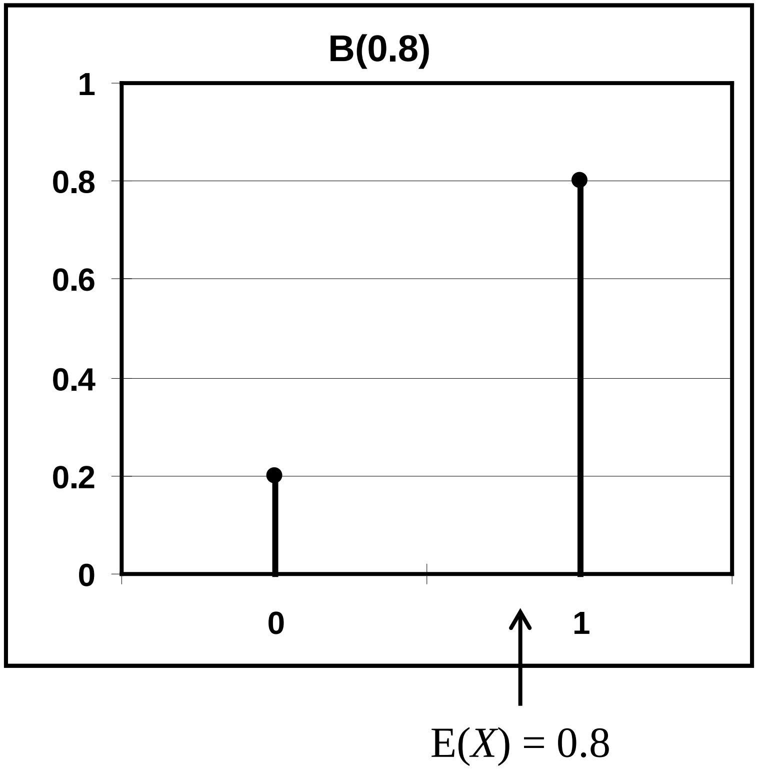
  $$p = 0.8\,,\ q = 1 - p$$

  pisteissä

  $x = 0, 1$

- Jakauman *odotusarvo*:

  $$\mathrm{E}(X) = p = 0.8$$



Bernoulli-kokeet

- Oletetaan että olemme kiinnostuneita vain siitä sattuuko jokin ilmiöön liittyvä tapahtuma A vai ei. Kutsumme tämmöistä satunnaiskoetta **Bernoulli-kokeeksi**.

- Yllä esitetyn mukaan Bernoulli-kokeita voidaan mallintaa Bernoulli-jakaumalla.

Bernoulli-kokeet ja diskreetit todennäköisyysjakaumat

- Toistetaan *samaa Bernoulli-koetta* niin, että toistot ovat toisistaan riippumattomia ja olkoon A se kiinnostuksen kohteena oleva tapahtuma, jonka sattumista toistojen aikana seurataan:

(i) **Binomijakauma** (ks. →) saadaan määräämällä todennäköisyys sille, että tapahtuma A sattuu x kertaa, kun koetta toistetaan n kertaa, jossa n on kiinteä, etukäteen päätetty luku.

(ii) **Geometrinen jakauma** saadaan määräämällä todennäköisyys sille, että tapahtuma A sattuu ensimmäisen kerran x. koetoistossa.

(iii) **Negatiivinen binomijakauma** saadaan määräämällä todennäköisyys sille, että tapahtuma A sattuu r. kerran x. koetoistossa.

Diskreettejä jakaumia

Diskreetti tasainen jakauma

Bernoulli-jakauma

>> Binomijakauma

Binomijakauma 1/2

- Toistetaan *samaa* Bernoulli-koetta n kertaa, jossa n on *kiinteä*, *etukäteen päätetty* luku.
- Oletetaan, että koetoistot ovat *riippumattomia*.
- Olkoon yksittäisessä toistossa tietyn tapahtuman A todennäköisyys

$$\Pr(A) = p.$$

- Määritellään *diskreetti satunnaismuuttuja* X:

X = Tapahtuman A esiintymiskertojen lukumäärä n toistossa.

Binomijakauma 2/2

- Sanomme, että satunnaismuuttuja X noudattaa **binomijakaumaa parametrein *n* ja *p***.
- Merkintä:

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

- Satunnaismuuttujan X **pistetodennäköisyysfunktio** on

$$f(x) = \Pr(X = x) = \binom{n}{x} p^x q^{n-x}, \; 0 < p < 1, \; q = 1 - p$$

$$x = 0, 1, 2, \ldots, n$$

Binomijakauma
Pistetodennäköisyysfunktion johto 1/2

- Toistetaan *samaa* Bernoulli-koetta n kertaa.
- Oletetaan, että koetoistot ovat *riippumattomia* ja tarkastellaan tapahtuman A sattumista koetoistojen aikana.
- Oletetaan, että toistokoesarjan tuloksena saadaan tapahtumajono

$$A\ A\ A^c\ A\ A^c\ \ldots\ A$$

jossa on x kpl tapahtumia A ja $(n - x)$ kpl tapahtumia A^c .

- Koska

$$\Pr(A) = p$$

$$\Pr(A^c) = 1 - \Pr(A) = 1 - p = q$$

tarkasteltavan tapahtumajonon todennäköisyydeksi saadaan *riippumattomien tapahtumien tulosäännön* nojalla

$$ppqpq\ \ldots\ p = p^x q^{n-x}$$

Pistetodennäköisyysfunktion johto 2/2

- *Erilaisia* jonoja, joissa on x kpl tapahtumia A ja $(n - x)$ kpl tapahtumia A^c, on

$$\binom{n}{x} \text{ kpl}$$

- *Erilaiset* tapahtumajonot ovat *toisensa poissulkevia*.
- *Toisensa poissulkevien tapahtumien yhteenlaskusäännön* mukaan todennäköisyys saada sellainen jono, jossa on x kpl tapahtumia A ja $(n - x)$ kpl tapahtumia A^c saadaan laskemalla *erilaisten* tällaisten jonojen todennäköisyydet yhteen.
- Siten kysytyksi todennäköisyydeksi saadaan

$$f(x) = \binom{n}{x} p^x q^{n-x}, \; q = 1 - p$$

$$x = 0, 1, 2, \ldots, n$$

Odotusarvo, varianssi ja standardipoikkeama

- Olkoon

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

- **Odotusarvo**:

  $$\mathrm{E}(X) = np$$

- **Varianssi** ja **standardipoikkeama**:

  $$\mathrm{Var}(X) = \mathrm{D}^2(X) = npq$$

  $$\mathrm{D}(X) = \sqrt{npq}$$

Binomijakauma

Pistetodennäköisyysfunktion kuvaaja

- Kuva oikealla esittää Binomi-jakauman

  Bin(12, 1/3)

  pistetodennäköisyysfunktiota

$$f(x) = \binom{n}{x} p^x q^{n-x}$$

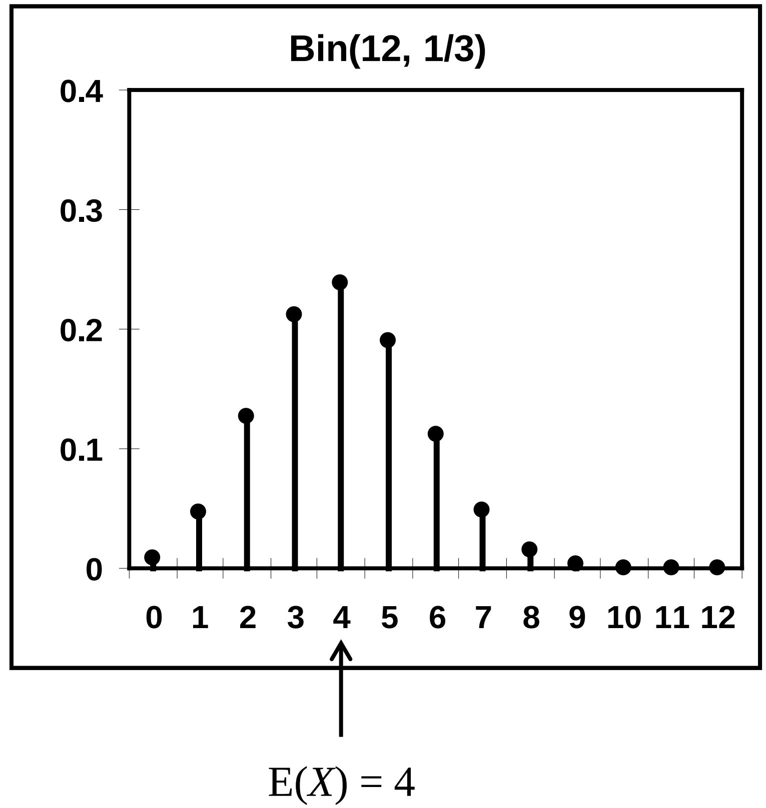
$$n = 12\,,\ p = 1/3\,,\ q = 1 - p$$

  pisteissä

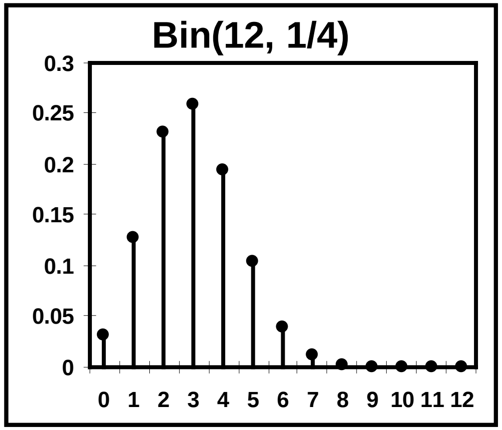
  $x = 0, 1, 2, \ldots, 12$

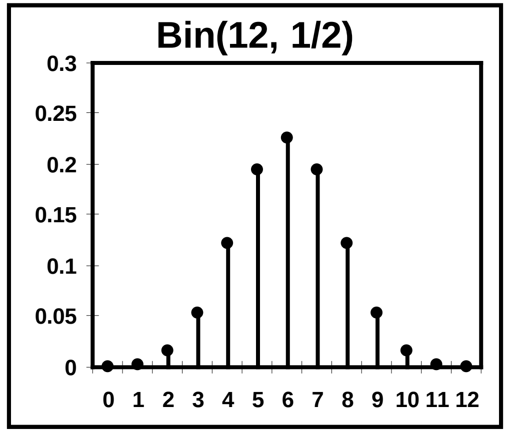
- Jakauman *odotusarvo*:

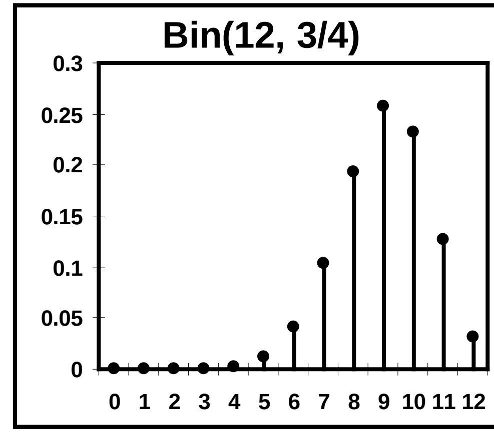
$$\mathrm{E}(X) = np = 4$$



Pistetodennäköisyysfunktion kuvaaja: Tapaukset *p* < 1/2, *p* = 1/2, *p* > 1/2







- $p < 1/2$: Binomijakauma on *vino oikealle*.
- $p = 1/2$: Binomijakauma on *symmetrinen*.
- $p > 1/2$: Binomijakauma on *vino vasemmalle*.

Binomijakauma

Esimerkki (Milton-Arnold) 1/2

- Tutkimukset lennonjohtajien työskentelystä osoittavat, että keskittymistä on vaikea ylläpitää, kun työskennellään pitkiä aikoja ruudulla näkyvän datan parissa.
- Tutkimuksen yllättävä tulos on, että tutkasignaalien havaitseminen muuttuu vaikeammaksi, jos havaittavia signaaleja on vähän.
- Signaalin havaitsemisen todennäköisyys on 0.9, jos keskimäärin 30 minuutin aikana signaaleja on 100. Todennäköisyys on vain 0.5, jos signaaleja on 10.
- Tuloksen arvellaan johtuvan siitä, että ajatukset lähtevät harhailemaan tilanteessa, joka ei pakota keskittymään.

Binomijakauma

Esimerkki (Milton-Arnold) 2/2

- Olkoon X oikein havaittujen signaalien lukumäärä 30 minuutin testijakson aikana. Testijaksossa esiintyy 10 signaalia. Miten X on jakautunut?
- Koe koostuu kymmenestä riippumattomasta Bernoulli-kokeesta, jossa positiivinen tulos on signaalin havaitseminen. Onnistumisen todennäköisyys on 0.5.
- X on binomijakautunut.
- Parametrit ovat n=10 ja p=0.5.
- Pistetodennäköisyysfunktioksi saadaan siis

$$f(x) = \binom{10}{x} (1/2)^x (1/2)^{10-x}$$

x = 1, 2, 3,..., 10.

Todennäköisyyslaskenta

Osa 3: Todennäköisyysjakaumia

- **Jatkuvia jakaumia**

Jatkuvia jakaumia

>> **Jatkuva tasainen jakauma**
Normaalijakauma
Keskeinen raja-arvolause

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 ja sen tiheysfunktio 2/3

- Satunnaismuuttujan *X* **tiheysfunktio** on

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

- Funktio *f*(*x*) kelpaa tiheysfunktioksi, koska

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

Jatkuva tasainen jakauma ja sen tiheysfunktio 3/3

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

$$X \sim \mathrm{Tasa}(a, b)$$

Odotusarvo, varianssi ja standardipoikkeama

- Olkoon $X \sim \mathrm{Tasa}(a, b)$.
- **Odotusarvo**:

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

- **Varianssi** ja **standardipoikkeama**:

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

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

Jatkuva tasainen jakauma
Tiheysfunktion kuvaaja

- Kuva oikealla esittää jatkuvan tasaisen jakauman

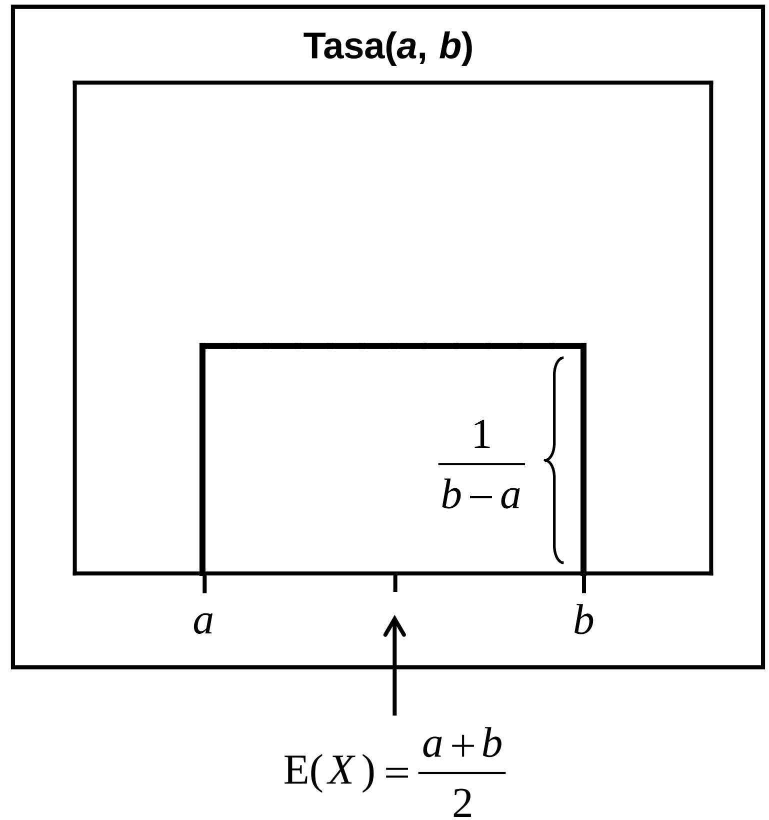
  Tasa(a, b)

  tiheysfunktiota

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

- Jakauman *odotusarvo*:

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



Jatkuvia jakaumia

Jatkuva tasainen jakauma

>> Normaalijakauma

Keskeinen raja-arvolause

Normaalijakauma ja sen tiheysfunktio

- 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 tiheysfunktioksi, koska

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

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

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

Normaalijakauma ja sen kertymäfunktio

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

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

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

Odotusarvo, varianssi ja standardipoikkeama

- Olkoon $X \sim \mathrm{N}(\mu,\ \sigma^2)$.
- **Odotusarvo**:

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

- **Varianssi** ja **standardipoikkeama**:

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

$$\mathrm{D}(X) = \sigma$$

Normaalijakauma
Tiheysfunktion kuvaaja

- Kuva oikealla esittää normaalijakauman

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

  tiheysfunktiota

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

  välillä [–6, +6], kun

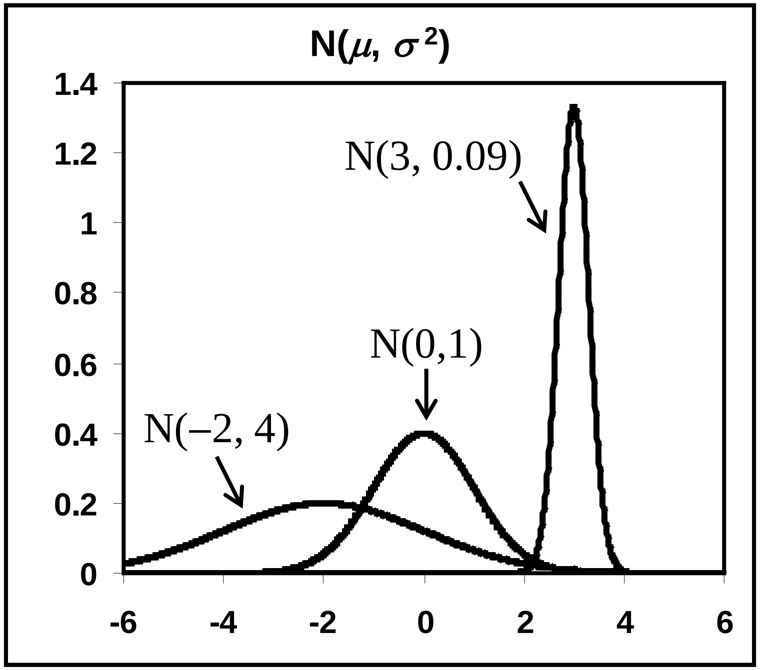
  (i) $\mu = -2 \quad \sigma^2 = 4$

  (ii) $\mu = 0 \quad \sigma^2 = 1$

  (iii) $\mu = +3 \quad \sigma^2 = 0.09$

- Jakauman *odotusarvo*:

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



Tiheysfunktion ja sen kuvaajan ominaisuuksia

- Kuva oikealla esittää normaalijakauman

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

  tiheysfunktiota

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

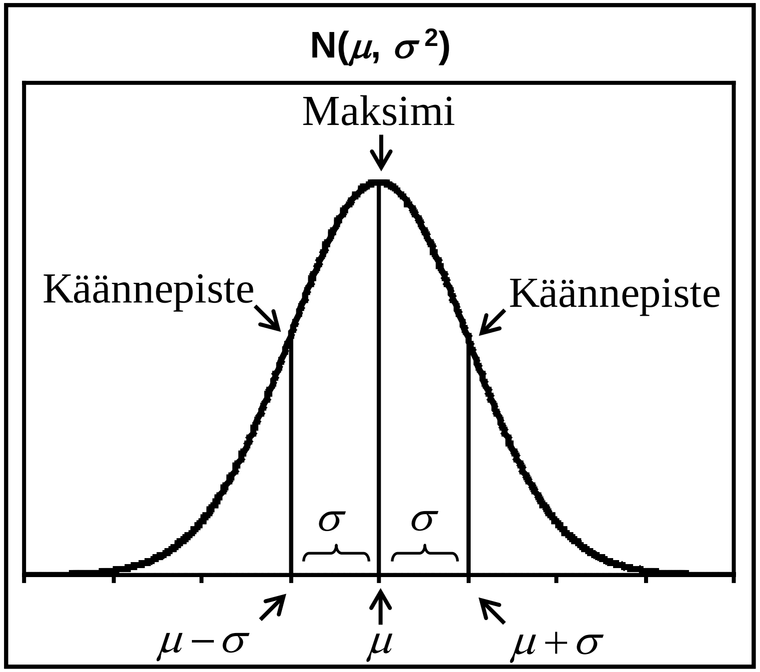
- Tiheysfunktiolla on *maksimi* pisteessä

  $x = \mu$

- Tiheysfunktiolla on *käännepisteet* pisteissä

  $x = \mu - \sigma$

  $x = \mu + \sigma$



68-95-99.7-sääntö

- *Kaikille* normaalijakaumille pätee:

(i) n. 68% jakauman todennäköisyysmassasta on välillä

$$[\mu - \sigma, \mu + \sigma]$$

(ii) n. 95% jakauman todennäköisyysmassasta on välillä

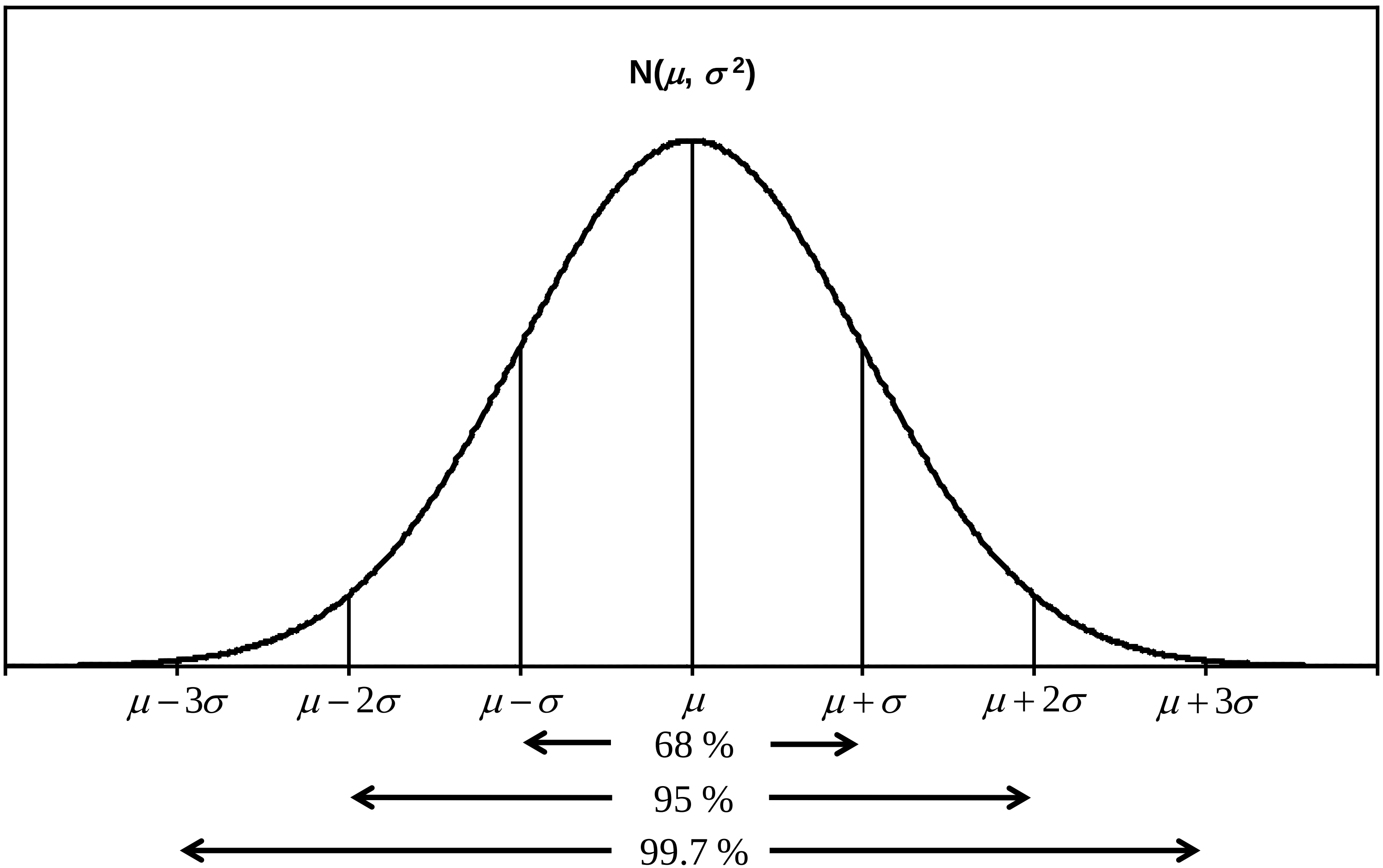
$$[\mu - 2\sigma, \mu + 2\sigma]$$

(iii) n. 99.7% jakauman todennäköisyysmassasta on välillä

$$[\mu - 3\sigma, \mu + 3\sigma]$$

Normaalijakauma

68-95-99.7-sääntö: Havainnollistus



Standardoitu normaalijakauma

- Olkoon

$$X \sim \mathrm{N}(0,1)$$

jolloin siis

$$\mathrm{E}(X) = 0$$

$$\mathrm{D}^2(X) = 1$$

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

Normaalijakauma

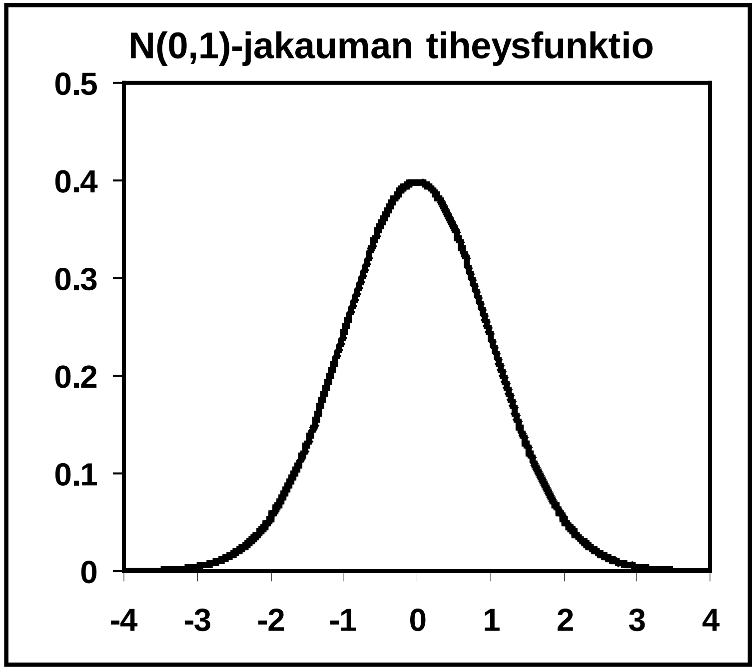
Standardoitu normaalijakauma: Tiheysfunktion kuvaaja

- Kuva oikealla esittää standardoidun normaali-jakauman

  N(0,1)

  tiheysfunktiota

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



Standardoitu normaalijakauma: Kertymäfunktion kuvaaja

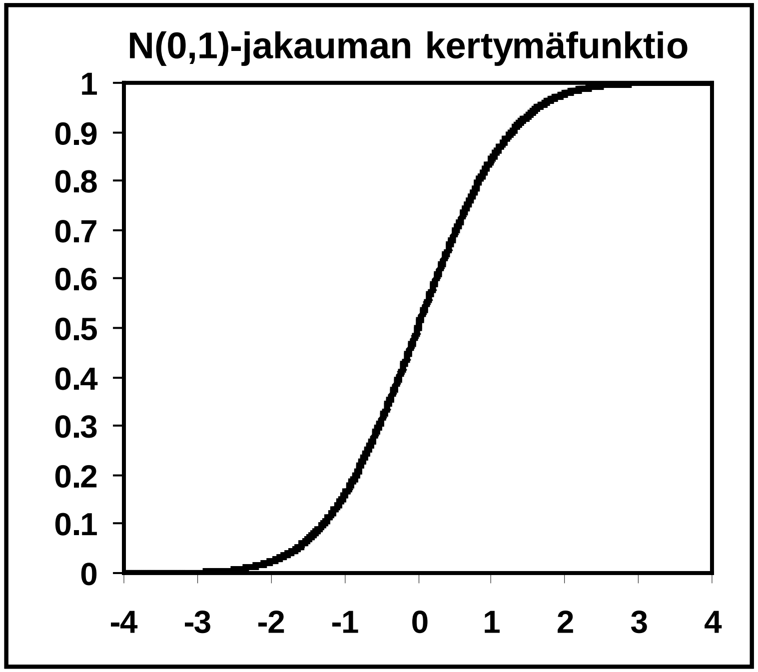
- Kuva oikealla esittää standardoidun normaali-jakauman

  N(0,1)

  kertymäfunktiota.
- Standardoidun normaali-jakauman kertymäfunktion $F(x)$ määrittelee kaava

  $$F(x) = \Pr(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

  jossa $f(x)$ on standardoidun normaalijakauman tiheysfunktio.



Lineaarimuunnoksen jakauma

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

$$Y = a + bX$$

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

- Tällöin Y on *normaalinen*:

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

- Perustelu:

Ks. lukua **Satunnaismuuttujien muunnosten jakaumat.**

Standardointi

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

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

  $\mathrm{D}(X) = \sigma$

- *Standardoidaan* satunnaismuuttuja X:

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

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

  $Z \sim \mathrm{N}(0,1)$

Normaalijakauma ja standardoitu normaalijakauma 1/2

- *Kaikki* normaalijakaumat N(μ, σ^2) ovat *samanmuotoisia standardoiduissa yksiköissä*

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

- Siten todennäköisyydet *mielivaltaisesta normaali-jakaumasta* N(μ, σ^2) voidaan aina määrätä *standardoidun normaalijakauman* N(0,1) avulla.

Normaalijakauma ja standardoitu normaalijakauma 2/2

- Olkoon siis

$$X \sim \mathrm{N}(\mu,\ \sigma^2) \text{ ja } \quad Z = \frac{X-\mu}{\sigma}$$

- Tällöin

$$\begin{aligned}
&\Pr(a \le X \le b) \\
&= \Pr\left(\frac{a-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right) \\
&= \Pr\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right)
\end{aligned}$$

missä $Z \sim \mathrm{N}(0,1)$.

Todennäköisyyksien määrääminen standardoidusta normaalijakaumasta 1/2

- Olkoon $Z \sim \mathrm{N}(0,1)$.
- Olkoon satunnaismuuttujan Z *kertymäfunktio*

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

- 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

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

kun z on annettu.

(ii) Määrää z, kun todennäköisyys

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

on annettu.

Todennäköisyyksien määrääminen standardoidusta normaalijakaumasta 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 tiheys-funktion *symmetrisyyttä* pisteen $z = 0$ suhteen:

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

Todennäköisyyksien määrääminen standardoidusta normaalijakaumasta: Esimerkki 1/2

- Olkoon

  $Z \sim \mathrm{N}(0,1)$

  ja olkoon

  $f_Z(z)$

  satunnaismuuttujan Z *tiheysfunktio*.

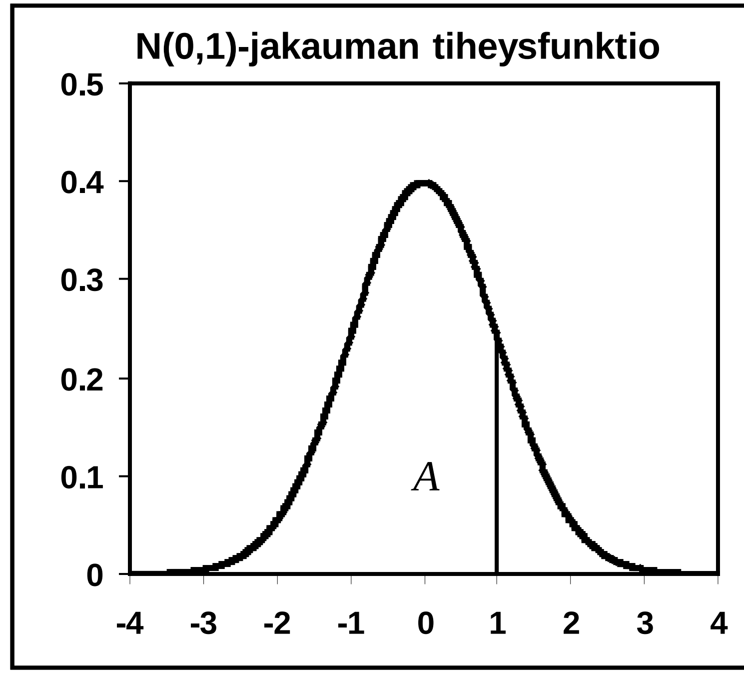
- Standardoidun normaalijakauman N(0,1) taulukoista saadaan:

  Alueen A pinta-ala

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

  $= \Pr(Z \leq 1)$

  $= 0.8413$



Todennäköisyyksien määrääminen standardoidusta normaalijakaumasta: Esimerkki 2/2

- Olkoon

  $$Z \sim \mathrm{N}(0,1)$$

  ja olkoon

  $$\Phi(z)$$

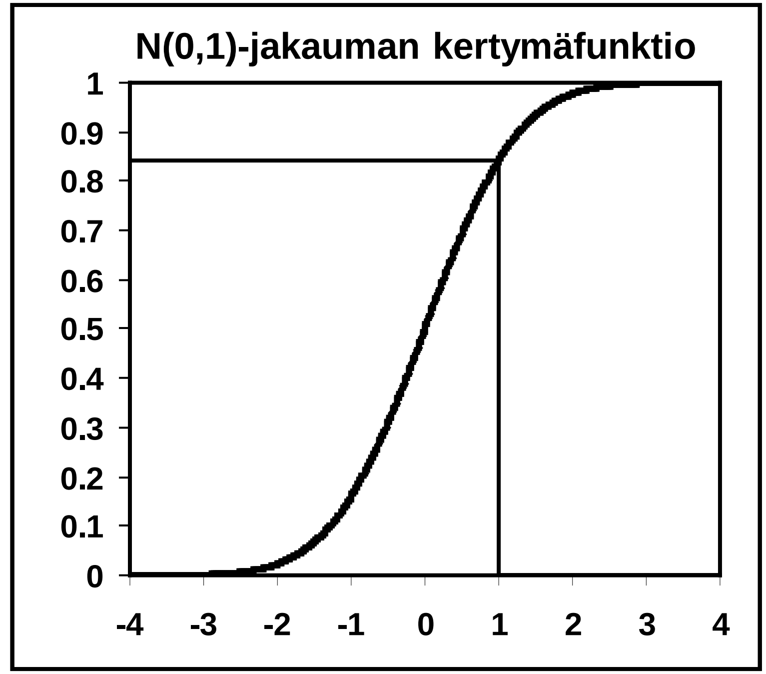
  satunnaismuuttujan Z *kertymäfunktio*.

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

  $$\Phi(1)$$

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

  $$= 0.8413$$



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* normaali-jakaumaa.
- Mikä on tämän tosiseikan selitys?
- Selityksenä on **keskeinen raja-arvolause**; ks. seuraavaa kappaletta.

Jatkuvia jakaumia

Jatkuva tasainen jakauma
Normaalijakauma
>> Keskeinen raja-arvolause

Johdanto 1/2

- Olkoon X_i , i = 1, 2, … , n jono *riippumattomia*, *samaa normaalijakaumaa* N(μ, σ^2) *noudattavia satunnais-muuttujia*.
- Tällöin voimme osoittaa, että *satunnaismuuttujien* X_i *summa* on normaali:

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

- Kysymys:

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

Johdanto 2/2

- *Ei-normaalisten* satunnaismuuttujien summa *ei yleensä ole* normaali.
- Kuitenkin, jos yhteenlaskettavia on ”tarpeeksi paljon”, *satunnaismuuttujien summa on* (hyvin yleisin ehdoin) *approksimatiivisesti normaalinen*.
- Tämä on **keskeisen raja-arvolauseen** olennainen sisältö.

- Koska monia satunnaismuuttujia voidaan pitää *usean riippumattoman tekijän summana*, antaa keskeinen raja-arvolause selityksen *empiiriselle havainnolle* niiden normaalisuudesta.

Keskeisen raja-arvolauseen formulointi

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

$$\mathrm{E}(X_i) = \mu \,,\; i = 1, 2, \ldots$$

$$\mathrm{D}^2(X_i) = \sigma^2 < \infty \,,\; i = 1, 2, \ldots$$

ja olkoon $\bar{X}_n$ **aritmeettinen keskiarvo**

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$, jolle jo tiedetään että $$\mathrm{E}(\bar{X}_n) = \mu$$ $$\mathrm{D}^2(\bar{X}_n) = \frac{\sigma^2}{n}$$

Keskeinen raja-arvolause

Aritmeettisen keskiarvon approksimatiivinen jakauma

- Tällöin **Keskeisen raja-arvolauseen** mukaan **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:

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

Kommentteja 1/2

- Keskeisestä raja-arvolauseesta on olemassa *yleisempiä muotoja*, joissa lievennetään *samoinjakautuneisuus*- ja *riippumattomuusoletuksia*.

- Keskeisen raja-arvolauseen mukaan usean satunnais-muuttujan *summa on* (*tietyin ehdoin*) *approksimatiivisesti normaalinen riippumatta yhteenlaskettavien jakaumasta*.

- Huomautus:

  Yhteenlaskettavien ei tarvitse olla edes *jatkuvia*, vaan ne voivat olla jopa *diskreettejä*.

Kommentteja 2/2

- 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.

Keskeisen raja-arvolauseen seurauksia

- Keskeisestä raja-arvolauseesta seuraa erikoistapauksina monet yksittäisiä jakaumia koskevat asymptoottiset tulokset.
- Erityisesti:

(i) **Binomijakauma lähestyy normaalijakaumaa**, kun toistokokeiden lukumäärän *n* annetaan kasvaa (De Moivren ja Laplacen tulos, ensimmäinen todistettu keskeinen raja-arvolause).