Ilkka Mellin

Todennäköisyyslaskenta

Osa 2: Satunnaismuuttujat ja todennäköisyysjakaumat

Kertymäfunktio

Kertymäfunktio

\>> **Kertymäfunktio: Määritelmä**
Diskreettien jakaumien kertymäfunktiot
Jatkuvien jakaumien kertymäfunktiot

Kertymäfunktion määritelmä

- Olkoon ξ *satunnaismuuttuja*.
- Satunnaismuuttujan ξ **kertymäfunktio** F on reaaliarvoinen funktio

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

Kertymäfunktion määritelmä: Kommentteja 1/2

- Satunnaismuuttujan ξ kertymäfunktion F määritelmässä

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

on

ξ = *satunnaismuuttuja*

x = *reaaliluku*, kertymäfunktion F *argumentti*

- Kertymäfunktion F arvo pisteessä x on *todennäköisyys* sille, että satunnaismuuttuja ξ saa arvoja, jotka ovat $\leq x$.
- Piste x erottaa *vasemmalle puolelleen todennäköisyys-massan, jonka koko on*

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

Kertymäfunktion määritelmä: Kommentteja 2/2

- Satunnaismuuttujan ξ kertymäfunktio

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

  kuvaa satunnaismuuttujan ξ *todennäköisyysmassan kertymistä*, kun kertymäfunktion argumentti x *kasvaa*.
- Satunnaismuuttujan ξ kertymäfunktio määrää *kaikkien* ko. satunnaisilmiöön liittyvien *tapahtumien todennäköisyydet*.
- Kertymäfunktion määritelmä sopii *kaikille satunnais-muuttujille* olivatpa ne diskreettejä, jatkuvia tai jotakin muuta tyyppiä.
- Kertymäfunktio on keskeinen työväline *matemaattisessa tilastotieteessä*.

Kertymäfunktio ja tapahtumien todennäköisyydet

- Jos satunnaismuuttujan ξ kertymäfunktio F tunnetaan, *kaikkien ko. satunnaisilmiöön liittyvien tapahtumien todennäköisyydet hallitaan.*
- Tämä johtuu seuraavista seikoista:

(i) *Jokaista tapahtumaa* vastaa jokin reaalilukujen joukon R osajoukko, joka voidaan muodostaa muotoa $(-\infty, x]$ olevista reaaliakselin väleistä tavanomaisten *joukko-opin operaatioiden* avulla.

(ii) *Jokaisen tapahtuman todennäköisyys* saadaan tyyppiä $(-\infty, x]$ olevien reaaliakselin välien todennäköisyyksistä *todennäköisyyslaskennan laskusääntöjen* avulla.

Kertymäfunktion ominaisuudet 1/2

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

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

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

(3) F on *ei-vähenevä*:

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

(4) F on *jatkuva oikealta*:

$\lim_{h\rightarrow0+}F(x+h)=F(x)$

Kertymäfunktion ominaisuudet 2/2

- Jos funktio $F : \mathrm{R} \rightarrow [0,1]$ on *kertymäfunktio*, niin:

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

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

Kertymäfunktion ominaisuuksien perustelu

- Käytämme kertymäfunktioiden ominaisuuksien perustelussa mm. seuraavia todennäköisyyslaskennan lauseita (ks. tarkemmin lukua **Todennäköisyyden aksioomat**):

Lause 1: Olkoon $(S, \mathcal{F}, \Pr)$ *todennäköisyyskenttä* ja $A_1, A_2, A_3, \ldots \in \mathcal{F}$.

(i) Jos $A_1 \subset A_2 \subset A_3 \subset \cdots$, niin

$$\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n \to +\infty} \Pr(A_n)$$

(ii) Jos $A_1 \supset A_2 \supset A_3 \supset \cdots$, niin

$$\Pr\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n \to +\infty} \Pr(A_n)$$

Lause 2: Olkoon $(S, \mathcal{F}, \Pr)$ *todennäköisyyskenttä* ja $A_1, A_2, A_3, \ldots \in \mathcal{F}$.

Jos $A_1 \supset A_2 \supset A_3 \supset \cdots \to \varnothing$, niin

$$\lim_{n \to +\infty} \Pr(A_n) = 0$$

Kertymäfunktion ominaisuus (1): Perustelu

- Olkoon $F(x) = \Pr(\xi \le x)$ satunnaismuuttujan ξ kertymäfunktio.
- Tällöin

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

- Todistus:

Olkoon

$$x_1 > x_2 > x_3 > \cdots$$

aleneva lukujono ja lisäksi

$$\lim_{n\to+\infty} x_n = -\infty$$

Tällöin

$$\{\xi \le x_1\} \supset \{\xi \le x_2\} \supset \{\xi \le x_3\} \supset \cdots \to \emptyset$$

Lauseen 2 mukaan

$$\lim_{n\to+\infty} F(x_n) = \lim_{n\to+\infty} \Pr\left(\xi \le x_n\right) = 0$$

Kertymäfunktio: Määritelmä

Kertymäfunktion ominaisuus (2): Perustelu 1/2

- Olkoon $F(x) = \Pr(\xi \le x)$ satunnaismuuttujan ξ kertymäfunktio.
- Tällöin

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

- Todistus:

Olkoon

$$x_1 < x_2 < x_3 < \cdots$$

kasvava lukujono ja lisäksi

$$\lim_{n \to +\infty} x_n = +\infty$$

Tällöin

$$\{\xi \le x_1\} \subset \{\xi \le x_2\} \subset \{\xi \le x_3\} \subset \cdots \to$$

ja

$$\{\xi > x_1\} \supset \{\xi > x_2\} \supset \{\xi > x_3\} \supset \cdots \to \varnothing$$

Kertymäfunktion ominaisuus (2): Perustelu 2/2

Lauseen 2 mukaan

$$\lim_{n\to+\infty} \Pr\left(\xi > x_n\right) = 0$$

joten

$$\begin{aligned}\lim_{n\to+\infty} F(x_n) &= \lim_{n\to+\infty} \Pr\left(\xi \le x_n\right)\\ &= 1 - \lim_{n\to+\infty} \Pr\left(\xi > x_n\right)\\ &= 1\end{aligned}$$

Kertymäfunktio: Määritelmä

Kertymäfunktion ominaisuus (3): Perustelu

- Olkoon $F(x) = \Pr(\xi \leq x)$ satunnaismuuttujan ξ kertymäfunktio.
- Tällöin

(3) $\quad F(x_1) \leq F(x_2)$, jos $x_1 \leq x_2$

- Todistus:

Olkoon

$$x_1 \leq x_2$$

Tällöin

$$\{\xi \leq x_1\} \subset \{\xi \leq x_2\}$$

joten

$$F(x_1) = \Pr(\xi \leq x_1) \leq \Pr(\xi \leq x_2) = F(x_2)$$

Kertymäfunktion ominaisuus (4): Perustelu

- Olkoon $F(x) = \Pr(\xi \le x)$ satunnaismuuttujan ξ kertymäfunktio.
- Tällöin

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

- Todistus:

Olkoon

$$h_1 > h_2 > h_3 > \cdots$$

aleneva lukujono ja lisäksi

$$\lim_{n \to +\infty} h_n = 0$$

Tällöin

$$\{\xi \le x + h_1\} \supset \{\xi \le x + h_2\} \supset \{\xi \le x + h_3\} \supset \cdots \to \{\xi \le x\}$$

Lauseen 1 kohdan (ii) mukaan

$$\lim_{n \to +\infty} F(x + h_n) = \lim_{n \to +\infty} \Pr\left(\xi \le x + h_n\right) = \Pr(\xi \le x) = F(x)$$

Kertymäfunktion ominaisuus (5): Perustelu

- Olkoon $F(x) = \Pr(\xi \le x)$ satunnaismuuttujan ξ kertymäfunktio.
- Tällöin

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

- Todistus:

Komplementtitapahtuman todennäköisyyden kaavan nojalla

$$\begin{aligned}\Pr(\xi > x) &= 1 - \Pr(\xi \le x) \\ &= 1 - F(x)\end{aligned}$$

Kertymäfunktio: Määritelmä
Kertymäfunktion ominaisuus (6): Perustelu

- Olkoon $F(x) = \Pr(\xi \le x)$ satunnaismuuttujan ξ kertymäfunktio.
- Tällöin

(6) $\Pr(a < \xi \le b) = F(b) - F(a)$

- Todistus:

Koska

$$\{\xi \le b\} = \{\xi \le a\} \cup \{a < \xi \le b\}$$

ja

$$\{\xi \le a\} \cap \{a < \xi \le b\} = \varnothing$$

niin *toisensa poissulkevien tapahtumien yhteenlaskusäännön* nojalla

$$\begin{aligned} F(b) &= \Pr(\xi \le b\} \\ &= \Pr(\xi \le a) + \Pr(a < \xi \le b) \\ &= F(a) + \Pr(a < \xi \le b) \end{aligned}$$

Kertymäfunktion tulkinta

- Kertymäfunktion määritelmän

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

  ja kertymäfunktion ominaisuuden

  (3) $F(x_1) \leq F(x_2)$, jos $x_1 \leq x_2$

  perusteella kertymäfunktiolle voidaan antaa seuraava tulkinta:

  Kertymäfunktio F kuvaa *miten satunnaismuuttujan* ξ *todennäköisyysmassaa kumuloituu* eli *kertyy lisää, kun kertymäfunktion argumentti x kasvaa*.

Tilastolliset taulukot ja kertymäfunktio

- *Tavanomaisten tilastollisessa päättelyssä käytettyjen jakaumien* **tilastolliset taulukot** *liittyvät jakaumien kertymäfunktion arvoihin.*
- **Normaalijakauman** taulukoissa on tavallisesti taulukoitu todennäköisyyksiä (kertymäfunktion arvoja)

$$\Pr(\xi \le x) = F(x)$$

useille argumentin x arvoille (ks. lukua **Jatkuvia jakaumia**).
- $\boldsymbol{\chi^2}$**-**, ***F*****-** ja ***t*****-jakaumien** taulukoissa on tavallisesti taulukoitu argumentin x arvoja muutamille todennäköisyyksille

$$\Pr(\xi \ge x) = 1 - F(x)$$

(ks. lukua **Normaalijakaumasta johdettuja jakaumia**).

Diskrettien ja jatkuvien jakaumien kertymäfunktiot

- Tarkastelemme seuraavassa kertymäfunktioita kahdessa erikoistapauksessa:

  (i) **Diskreettien jakaumien kertymäfunktiot**

  (ii) **Jatkuvien jakaumien kertymäfunktiot**

Kertymäfunktio

Kertymäfunktio: Määritelmä
>> Diskreettien jakaumien kertymäfunktiot
Jatkuvien jakaumien kertymäfunktiot

Diskreetin jakauman kertymäfunktion määritelmä

- Olkoon ξ **diskreetti satunnaismuuttuja** ja $\{x_1, x_2, x_3, \ldots\}$ sen *tulosvaihtoehtojen* eli *arvojen* joukko.
- Olkoon satunnaismuuttujan ξ *pistetodennäköisyysfunktio*

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

- Määritellään funktio $F : \mathbb{R} \to [0,1]$ kaavalla

$$F(x) = \Pr(\xi \le x) = \sum_{i \mid x_i \le x} f(x_i)$$

- Tällöin F on *diskreetin satunnaismuuttujan* ξ **kertymä-funktio**.
- Diskreetin satunnaismuuttujan kertymäfunktio F on *epäjatkuva ei-vähenevä* funktio.

Diskreetin jakauman kertymäfunktion määritelmä: Kommentteja

- *Diskreetin jakauman kertymäfunktion F* määritelmän

  $$F(x) = \Pr(\xi \le x) = \sum_{i \mid x_i \le x} f(x_i)$$

  mukaan kertymäfunktion F arvo pisteessä x eli todennäköisyys tapahtumalle $\xi \le x$ saadaan *laskemalla yhteen kaikki pistetodennäköisyydet*

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

  joita vastaavat satunnaismuuttujan ξ arvot $x_i \le x$.
- *Kaikkien* satunnaismuuttujaan ξ liittyvien *tapahtumien todennäköisyydet* voidaan määrätä sen kertymäfunktion avulla.

Diskreetin jakauman kertymäfunktion ja pistetodennäköisyysfunktion yhteys

- Olkoon ξ *diskreetti satunnaismuuttuja* ja $\{x_1, x_2, x_3, \ldots\}$ sen *tulosvaihtoehtojen* eli *arvojen* joukko.
- Olkoon satunnaismuuttujan ξ *pistetodennäköisyysfunktio*

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

- Olkoon satunnaismuuttujan ξ *kertymäfunktio*

$$F(x) = \Pr(\xi \le x) = \sum_{i \mid x_i \le x} p_i$$

- Tällöin

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

Esimerkki:
Onnenpyörä 1/7

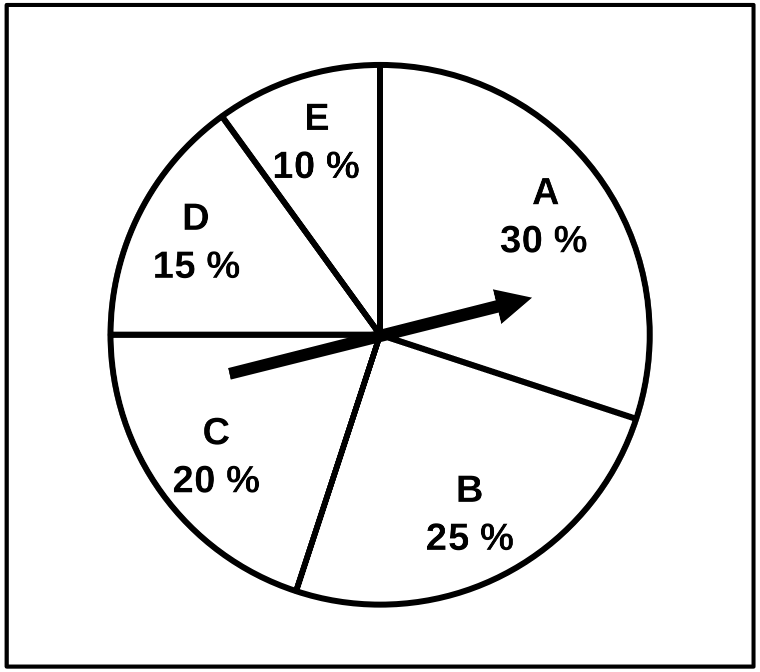
- Luvun

  Satunnaismuuttujat ja todennäköisyysjakaumat

  kappaleen

  Diskreetit satunnaismuuttujat ja niiden todennäköisyysjakaumat

  johdattelevassa esimerkissä käsitellään viereen kuvatun onnenpyörän käyttäytymistä satunnaisilmiönä.

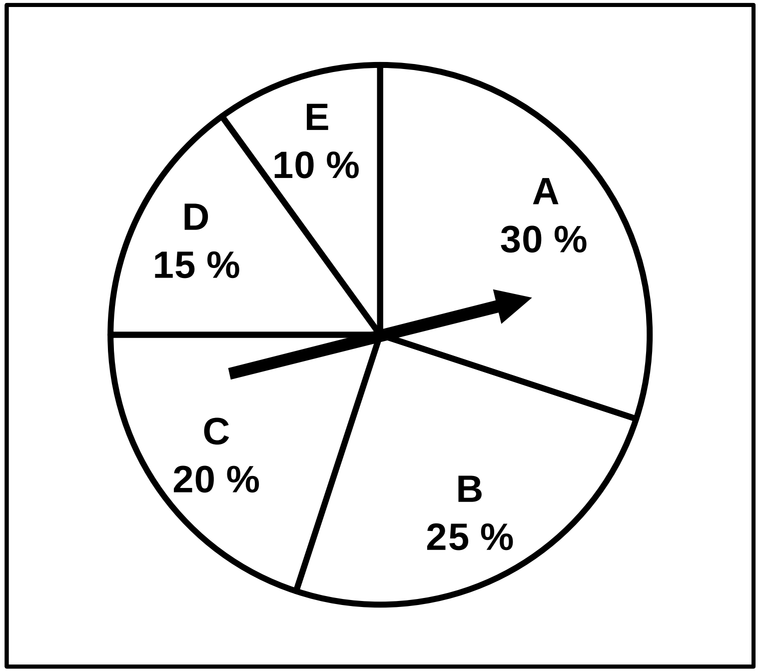


Esimerkki: Onnenpyörä 2/7

- 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 3/7

- Esimerkissä määriteltiin diskreetti satunnaismuuttuja ξ, joka liittää tulosvaihtoehtoihin

  A, B, C, D, E

  reaaliluvut seuraavalla tavalla:

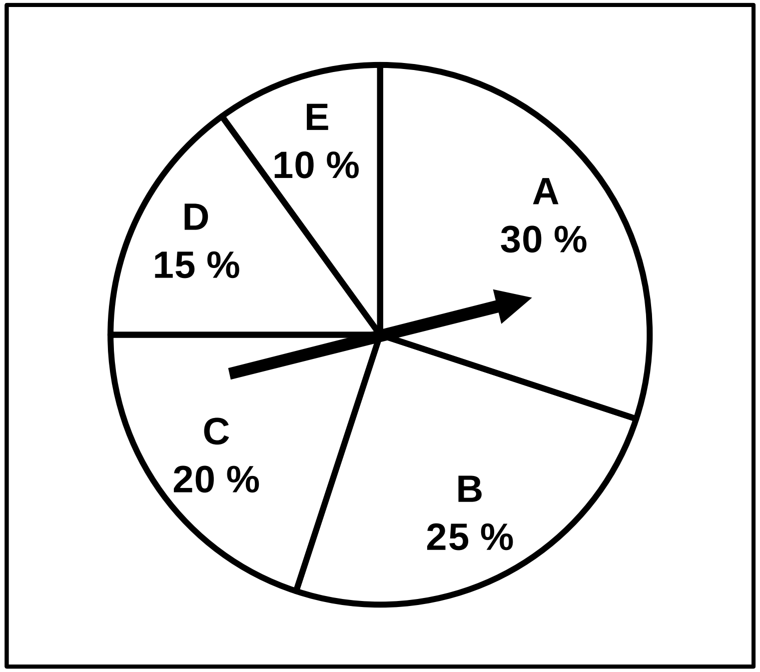
  A $\rightarrow$ 1

  B $\rightarrow$ 2

  C $\rightarrow$ 3

  D $\rightarrow$ 4

  E $\rightarrow$ 5



Esimerkki: Onnenpyörä 4/7

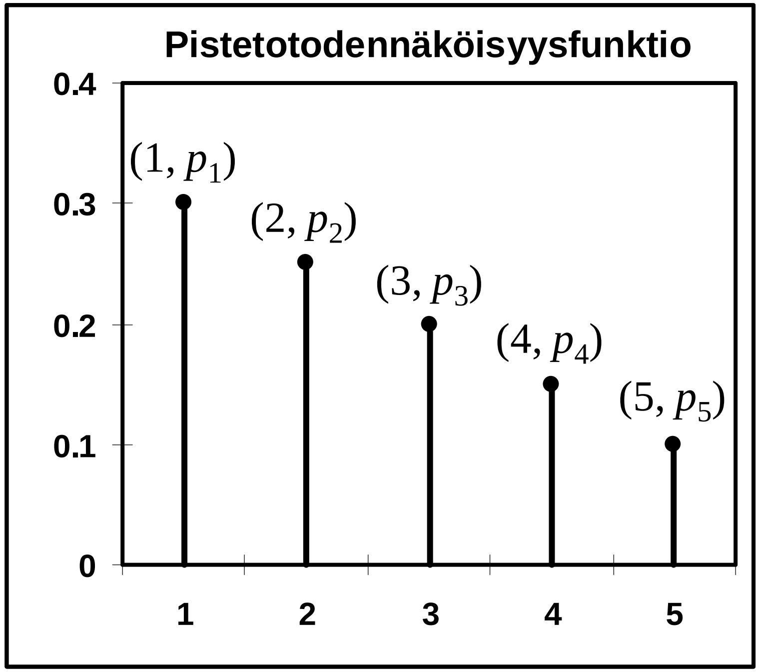
- *Diskreetin satunnaismuuttujan ξ pistetodennäköisyysfunktio f* voidaan yleisesti määritellä kaavalla

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

  jossa

  $$\{x_1, x_2, x_3, \ldots \}$$

  on satunnaismuuttujan ξ saamien *arvojen* joukko.



Esimerkki: Onnenpyörä 5/7

- Esimerkin tapauksessa satunnaismuuttujan ξ piste-todennäköisyysfunktio f voidaan määritellä seuraavasti:

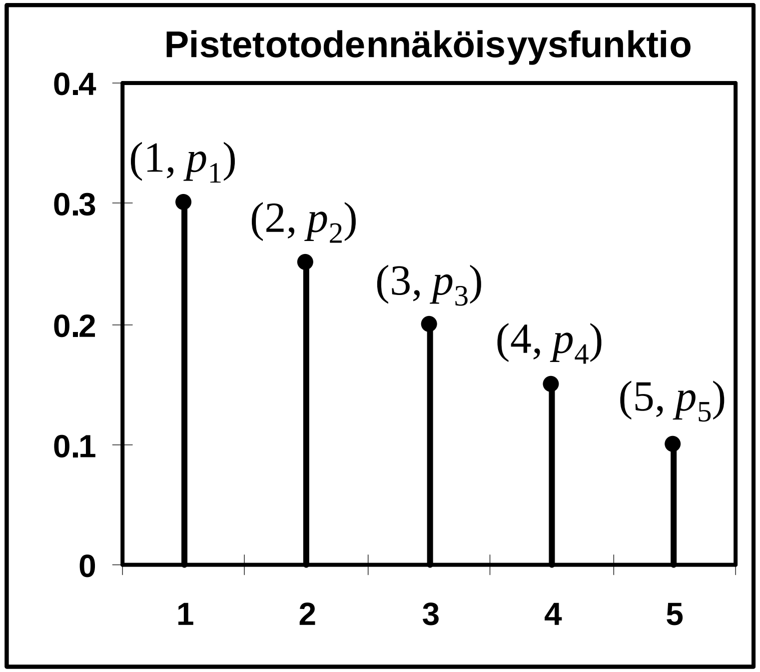
  $f(1) = \Pr(\xi = 1) = 0.30 = \Pr(A)$

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

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

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

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



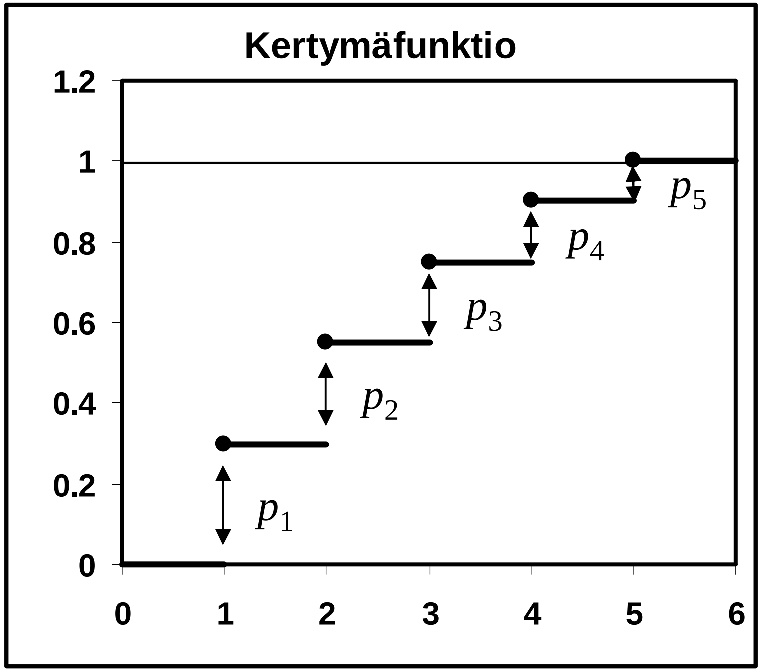
Diskreettien jakaumien kertymäfunktiot

Esimerkki:
Onnenpyörä 6/7

- Satunnaismuuttujan ξ *kertymäfunktio* on

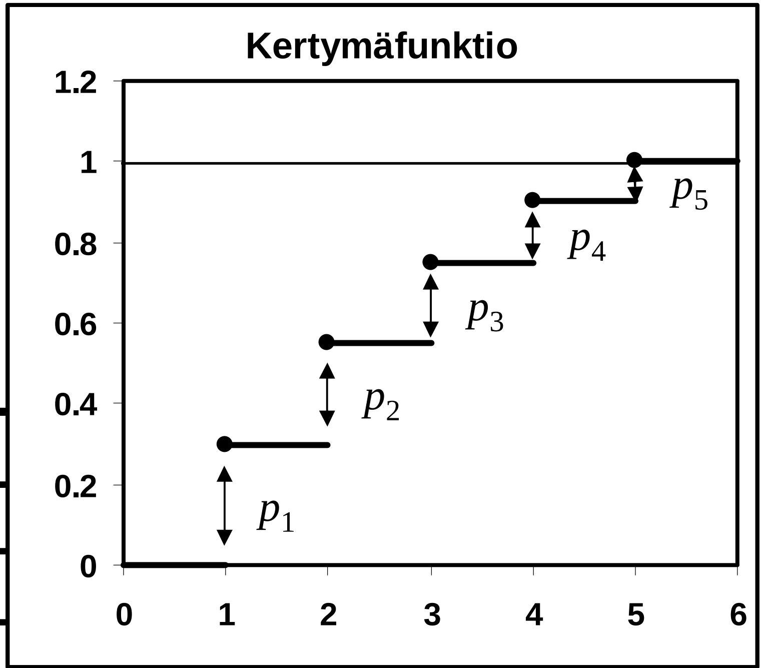
  $F(x) = \Pr(\xi \leq x)$
- *Diskreetin* satunnaismuuttujan *pistetodennäköisyys*- ja *kertymäfunktioiden* välillä on seuraava yhteys:

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



Esimerkki: Onnenpyörä 7/7

- Esimerkin tapauksessa satunnaismuuttujan ξ *kertymäfunktio* F voidaan määritellä alla olevan taulukon avulla.
- Kuva oikealla esittää esimerkin kertymäfunktion kuvaajaa.



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

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

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

- *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*.

Välien todennäköisyydet 1/2

- Diskreetin jakauman tapauksessa välin $(a,b] \subset \mathrm{R}$
  todennäköisyys on

$$\begin{aligned}
\Pr(a < \xi \leq b) &= F(b) - F(a) \\
&= \sum_{i \mid x_i \in (a,b]} \Pr(\xi = x_i) \\
&= \sum_{i \mid x_i \in (a,b]} p_i
\end{aligned}$$

Välien todennäköisyydet 2/2

- Kaavan

$$\Pr(a < \xi \le b) = F(b) - F(a) = \sum_{i \mid x_i \in (a,b]} p_i$$

mukaan välin $(a,b] \subset \mathbb{R}$ todennäköisyys voidaan määrätä kahdella tavalla:

(i) Jos jakauman *pistetodennäköisyysfunktio* tunnetaan, saadaan välin $(a, b]$ todennäköisyys *laskemalla yhteen* pistetodennäköisyydet p_i , joita vastaavat $x_i \in (a, b]$.

(ii) Jos jakauman *kertymäfunktio F* tunnetaan, saadaan välin $(a, b]$ todennäköisyys laskemalla *kertymäfunktion F arvojen* $F(b)$ ja $F(a)$ *erotus*.

Kertymäfunktio

Kertymäfunktio: Määritelmä

Diskreettien jakaumien kertymäfunktiot

>> Jatkuvien jakaumien kertymäfunktiot

Jatkuvan jakauman kertymäfunktion määritelmä

- Olkoon ξ **jatkuva satunnaismuuttuja**.
- Olkoon satunnaismuuttujan ξ *tiheysfunktio* $f(x)$.
- Määritellään funktio $F : \mathrm{R} \rightarrow [0,1]$ kaavalla

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

- F on *jatkuvan satunnaismuuttujan* ξ **kertymäfunktio**.
- Jatkuvan satunnaismuuttujan *kertymäfunktio* F on *jatkuva ei-vähenevä* funktio.

Jatkuvan jakauman kertymäfunktion määritelmä: Kommentteja

- *Jatkuvan jakauman kertymäfunktion* F määritelmän

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

mukaan kertymäfunktion F arvo pisteessä x eli todennäköisyys tapahtumalle $\xi \le x$ määrätään *integroimalla tiheysfunktio* f välillä $(-\infty, x]$.

- *Kaikkien* satunnaismuuttujaan ξ liittyvien *tapahtumien todennäköisyydet* voidaan määrätä sen kertymäfunktion avulla.

Jatkuvan jakauman kertymäfunktion ja tiheysfunktion yhteys

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

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

- Tällöin

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

Välien todennäköisyydet 1/2

- Jatkuvan jakauman tapauksessa välin $(a,b] \subset \mathrm{R}$ a
todennäköisyys on

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

Välien todennäköisyydet 2/2

- Kaavan

$$\Pr(a \le \xi \le b) = F(b) - F(a) = \int_a^b f(x)dx$$

mukaan välin $(a,b] \subset \mathbb{R}$ todennäköisyys voidaan määrätä kahdella tavalla:

(i) Jos jakauman *tiheysfunktio* f tunnetaan, saadaan välin $[a, b]$ todennäköisyys *integroimalla tiheysfunktio* välillä $[a, b]$.

(ii) Jos jakauman *kertymäfunktio* F tunnetaan, saadaan välin $[a, b]$ todennäköisyys laskemalla *kertymäfunktion arvojen* $F(b)$ ja $F(a)$ *erotus*.

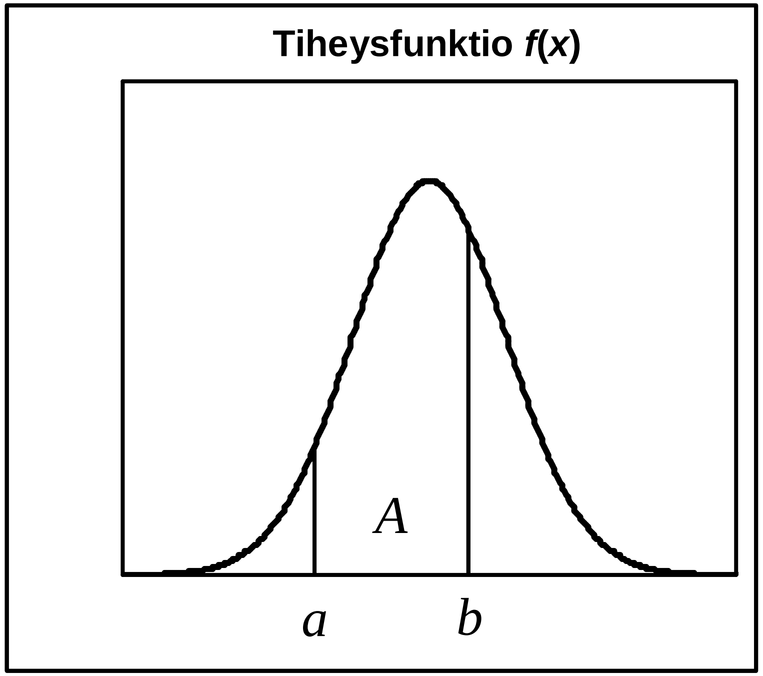
Jatkuvan jakauman tiheysfunktio ja välien todennäköisyydet: Havainnollistus

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

  $\Pr(a \le \xi \le b)$

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

  $=$ Alueen A pinta-ala
- Kuva oikealla esittää **normaali-jakauman** *tiheysfunktiota* (ks. lukua **Jatkuvia jakaumia**).



Jatkuvan jakauman kertymäfunktio ja välien todennäköisyydet: Havainnollistus

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

$$\Pr(a \le \xi \le b)$$
$$= F(b) - F(a)$$
$$= \int_a^b f(x)dx$$

- Kuva oikealla esittää **normaali-jakauman** *kertymäfunktiota* (ks. lukua **Jatkuvia jakaumia**).