Ilkka Mellin

Todennäköisyyslaskenta

Osa 2: Satunnaismuuttujat ja todennäköisyysjakaumat

Satunnaismuuttujat ja todennäköisyysjakaumat

Satunnaismuuttujat ja todennäköisyysjakaumat

>> **Satunnaismuuttujat ja niiden todennäköisyysjakaumat**
Diskreetit satunnaismuuttujat ja niiden todennäköisyysjakaumat
Jatkuvat satunnaismuuttujat ja niiden todennäköisyysjakaumat

Johdatteleva esimerkki 1: Rahanheitto – 1/2

- Tarkastellaan rahanheittoa *satunnaisilmiönä*.
- Alkeistapahtumat:

  Kruuna, Klaava
- Otosavaruus:

  S = {Kruuna, Klaava}
- Määritellään reaaliarvoinen *funktio* $\xi : S \to \mathrm{R}$, joka liittää otos-avaruuden S alkioihin *reaaliluvun* eli *numeerisen koodin* seuraavalla tavalla:

  Kruuna $\to$ 1

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

Johdatteleva esimerkki 1: Rahanheitto − 2/2

- Tehdään seuraava *oletus* niistä *todennäköisyyksistä*, joilla ξ saa arvonsa:

  $$\Pr(\xi = 1) = 1/2$$
  $$\Pr(\xi = 0) = 1/2$$

- Satunnaismuuttujan ξ *arvot* ja niihin liitetyt *todennäköisyydet* muodostavat *yhdessä* satunnaismuuttujan ξ **todennäköisyysjakauman**.
- Satunnaismuuttuja ξ ja sen todennäköisyysjakauma muodostavat **tilastollisen mallin** eli **todennäköisyysmallin** rahanheitolle satunnaisilmiönä.
- Koska määritelty satunnaismuuttuja ξ saa vain erillisiä arvoja, sitä sanotaan **diskreetiksi**.
- Satunnaismuuttuja ξ noudattaa **Bernoulli-jakaumaa**; ks. lukua **Diskreettejä jakaumia**.

Johdatteleva esimerkki 2: Sukupuolen määräytyminen − 1/2

- Tarkastellaan sukupuolen määräytymistä *satunnaisilmiönä*.
- Alkeistapahtumat:

  Tyttö, Poika
- Otosavaruus:

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

  $$\begin{aligned} \text{Tyttö} &\rightarrow 1 \\ \text{Poika} &\rightarrow 0 \end{aligned}$$
- *Funktiota* ξ kutsutaan **satunnaismuuttujaksi**, koska *sattuma* määrää mikä funktion arvoista *realisoituu*, vaikka ξ on funktiona *täysin määrätty*.

Johdatteleva esimerkki 2: Sukupuolen määräytyminen − 2/2

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

$$\Pr(\xi = 1) = 0.4902$$

$$\Pr(\xi = 0) = 0.5098$$

- Satunnaismuuttujan ξ *arvot* ja niihin liitetyt *todennäköisyydet* muodostavat *yhdessä* satunnaismuuttujan ξ **todennäköisyys-jakauman**.
- Satunnaismuuttuja ξ ja sen todennäköisyysjakauma muodostavat **tilastollisen mallin** eli **todennäköisyysmallin** sukupuolen määräytymiselle satunnaisilmiönä.
- Koska määritelty satunnaismuuttuja ξ saa vain erillisiä arvoja, sitä sanotaan **diskreetiksi**.
- Satunnaismuuttuja ξ noudattaa **Bernoulli-jakaumaa**; ks. lukua **Diskreettejä jakaumia**.

Johdatteleva esimerkki 3: Nopanheitto − 1/2

- Tarkastellaan nopanheittoa *satunnaisilmiönä*.
- Alkeistapahtumat:

  Silmäluvut 1, 2, 3, 4, 5, 6
- Otosavaruus:

  S = {Silmäluvut 1, 2, 3, 4, 5, 6}
- Määritellään reaaliarvoinen *funktio* $\xi : S \to \mathrm{R}$, joka liittää otos-avaruuden S alkioihin *numeerisen koodin* eli *reaaliluvun* siten, että jokaiseen silmälukuun liitetään vastaava kokonaisluku:

  Silmäluku $i \to i$, $i = 1, 2, 3, 4, 5, 6$
- *Funktiota* ξ kutsutaan **satunnaismuuttujaksi**, koska *sattuma* määrää mikä funktion arvoista *realisoituu*, vaikka ξ on funktiona *täysin määrätty*.

Johdatteleva esimerkki 3: Nopanheitto − 2/2

- Tehdään seuraava *oletus* niistä *todennäköisyyksistä*, joilla ξ saa arvonsa:

  $$\Pr(\xi = i) = 1/6,\ i = 1, 2, 3, 4, 5, 6$$

- Satunnaismuuttujan ξ *arvot* ja niihin liitetyt *todennäköisyydet* muodostavat *yhdessä* satunnaismuuttujan ξ **todennäköisyys-jakauman**.
- Satunnaismuuttuja ξ ja sen todennäköisyysjakauma muodostavat **tilastollisen mallin** eli **todennäköisyysmallin** nopanheitolle satunnaisilmiönä.
- Koska määritelty satunnaismuuttuja ξ saa vain erillisiä arvoja, sitä sanotaan **diskreetiksi**.
- Satunnaismuuttuja ξ noudattaa **diskreettiä tasaista jakaumaa**; ks. lukua **Diskreettejä jakaumia**.

Johdatteleva esimerkki 4: Nopanheitto − 1/2

- Heitetään noppaa toistuvasti ja tarkastellaan *satunnaisilmiönä* sen heiton järjestysnumeron määräytymistä, jolla saadaan *ensimmäisen kuutonen*.
- Alkeistapahtumat:

  Järjestysnumerot 1, 2, 3, …
- Otosavaruus:

  $S = \{1, 2, 3, \ldots\}$ on (numeroituvasti) *ääretön*.
- Määritellään reaaliarvoinen *funktio* $\xi : S \rightarrow \mathrm{R}$, joka liittää otos-avaruuden S alkioihin *numeerisen koodin* eli *reaaliluvun* siten, että jokaiseen järjestysnumeroon liitetään vastaava kokonaisluku:

  Järjestysnumero $i \rightarrow i$, $i = 1, 2, 3, \ldots$
- *Funktiota* ξ kutsutaan **satunnaismuuttujaksi**, koska *sattuma* määrää mikä funktion arvoista *realisoituu*, vaikka ξ on funktiona *täysin määrätty*.

Johdatteleva esimerkki 4: Nopanheitto − 2/2

- Määräämme myöhemmin *todennäköisyydet*, joilla ξ saa arvonsa.
- Satunnaismuuttujan ξ saamien arvojen todennäköisyydet noudattavat *diskreettiä jakaumaa*, jota kutsutaan **geometriseksi jakaumaksi**.
- Satunnaismuuttuja ξ ja sen todennäköisyysjakauma muodostavat **tilastollisen mallin** eli **todennäköisyysmallin** sen heiton järjestys-numeron määräytymiselle, jolla saadaan *ensimmäinen kuutonen*.
- *Lisätietoja* geometrisesta jakaumasta:

  (i) Tämän luvun diskreettejä satunnaismuuttujia ja niiden jakaumia käsittelevässä kappaleessa käsitellään *esimerkkiä* geometrisestä jakaumasta.

  (ii) Geometrista jakamaa käsitellään *yleisesti* luvussa **Diskreettejä jakaumia**.

Johdatteleva esimerkki 5: Onnenpyörä – 1/3

- Onnenpyörän keskipisteeseen on asetettu vapaasti pyörivä osoitin, jota pyöräytetään pelissä.
- Tarkastellaan *satunnaisilmiönä* kulmaa, jonka osoitin pysähdyttyään muodostaa lähtöasentoonsa verrattuna.
- Alkeistapahtumat:

  Kulmat välillä $[0°, 360°)$
- Otosavaruus:

  $S = [0°, 360°)$ on (ylinumeroituvasti) *ääretön*.

Johdatteleva esimerkki 5: Onnenpyörä − 2/3

- Määritellään reaaliarvoinen *funktio* $\xi : S \rightarrow \mathbb{R}$, joka liittää otos-avaruuden S alkioihin *numeerisen koodin* eli *reaaliluvun* seuraavasti siten, että jokaiseen kulmaan liitetään vastaava reaaliluku:

  Kulma $x \rightarrow x \in [0, 360)$
- *Funktiota* ξ kutsutaan **satunnaismuuttujaksi**, koska *sattuma* määrää mikä funktion arvoista *realisoituu*, vaikka ξ on funktiona *täysin määrätty*.
- Tarkastelemme myöhemmin *todennäköisyyksiä*, joilla ξ saa arvoja joltakin väliltä $[a, b] \subset [0, 360)$.
- Satunnaismuuttuja ξ noudattaa **jatkuvaa tasaista jakaumaa**; ks. lukua **Jatkuvia jakaumia**.

Johdatteleva esimerkki 5: Onnenpyörä − 3/3

- *Lisätietoja* jatkuvasta tasaisesta jakaumasta:

  (i) Tämän luvun jatkuvia satunnaismuuttujia ja niiden jakaumia käsittelevässä kappaleessa käsitellään *esimerkkiä* jatkuvasta tasaisesta jakaumasta.

  (ii) Jatkuvaa tasaista jakamaa käsitellään *yleisesti* jatkuvia todennäköisyysjakaumia käsittelevässä luvussa.

Satunnaismuuttuja: Määritelmä

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

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

- Tällöin ξ on **satunnaismuuttuja**.
- Jos haluamme korostaa sitä, että satunnaismuuttuja ξ on otosavaruuden S *kuvaus* reaalilukujen joukkoon $\mathbb{R}$, merkitsemme

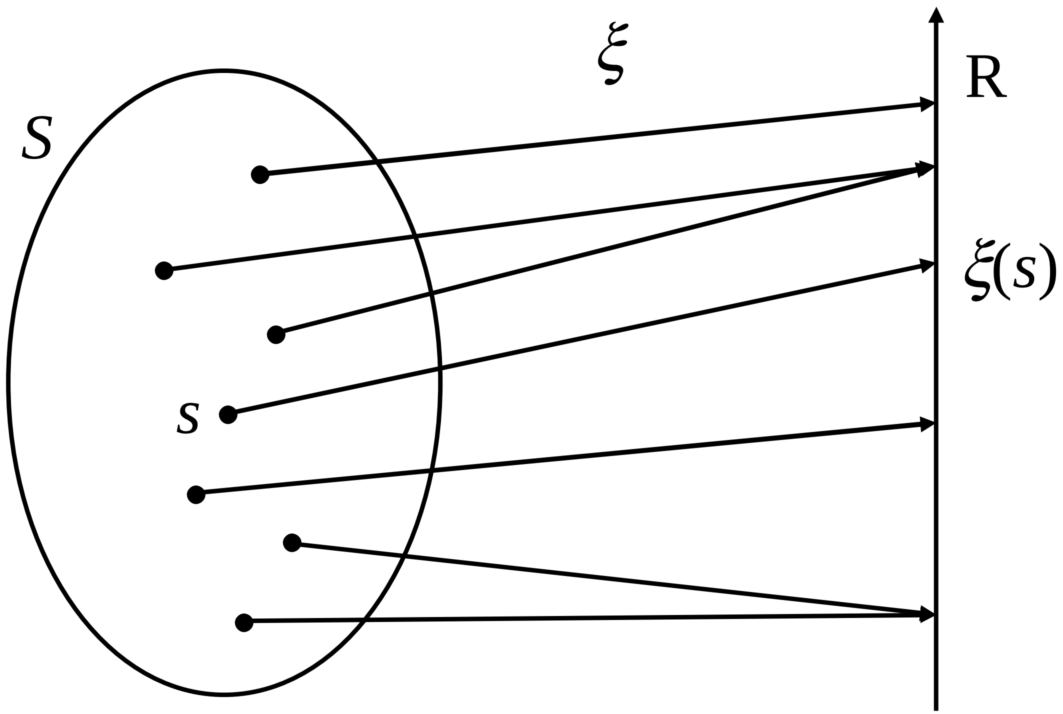
$$\xi(s) \in \mathbb{R}, \, s \in S$$

- Huomautus:

  Merkitsemme *tässä luvussa* satunnaismuuttujia tavallisesti kreikan kielen aakkosten kirjaimella ξ (”ksi”).

Satunnaismuuttujan määritelmä: Havainnollistus

$$\xi : S \to \mathrm{R}$$



Satunnaismuuttujan määritelmä: Kommentteja 1/2

- *Satunnaismuuttuja* on *funktiona* täysin *määrätty*, mutta *sattuma määrää* mikä funktion arvoista *realisoituu*.
- Satunnaismuuttuja *kuvaa* satunnaisilmiön *tulos-vaihtoehtoja numeerisessa muodossa*.
- Satunnaismuuttuja liittää jokaiseen satunnaisilmiön *tulosvaihtoehtoon reaaliluvun* eli *numeerisen koodin.*
- Huomautus:

  Satunnaismuuttuja on terminä *epäonnistunut*, koska se ei kerro sitä olennaista asiaa, että satunnaismuuttuja on itse asiassa *funktio*.

Satunnaismuuttujan määritelmä: Kommentteja 2/2

- Tässä esitetty satunnaismuuttujan määritelmä *on* siinä mielessä *epätarkka*, että mikä tahansa otosavaruuden reaaliarvoinen funktio *ei kelpaa satunnaismuuttujaksi*.
- Jotta funktio kelpaisi satunnaismuuttujaksi, sen on oltava *mitallinen*.
- Voidaan osoittaa, että *diskreetit* ja *jatkuvat satunnais-muuttujat* – joita tässä esityksessä pelkästään käsitellään – *ovat mitallisia funktioita*.
- Mitallisuuden käsittely *sivuutetaan* tässä esityksessä.

Todennäköisyysjakauma: Määritelmä 1/2

- Olkoon kolmikko (S, $\mathcal{F}$, Pr) otosavaruudessa S määritelty **todennäköisyyskenttä**, jossa on seuraavat elementit:

  S = otosavaruus

  $\mathcal{F}$ = on otosavaruuden S osajoukkojen joukossa määritelty σ-*algebra*

  Pr = on σ-algebran $\mathcal{F}$ alkioille määritelty *todennäköisyysmitta*

- Olkoon ξ otosavaruudessa S määritelty **satunnais-muuttuja** eli (mitallinen) funktio otosavaruudesta S reaalilukujen joukkoon $\mathbb{R}$:

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

Todennäköisyysjakauma: Määritelmä 2/2

- *Satunnaismuuttuja* ξ *määrittelee* otosavaruuden S (mitallisena) kuvauksena *todennäköisyysmitan reaalilukujen joukossa*.
- Satunnaismuuttujan ξ **todennäköisyysjakaumalla** (tai **jakaumalla**) tarkoitetaan kuvauksen

$$\xi : S \rightarrow \mathrm{R}$$

reaalilukujen joukkoon indusoimaa todennäköisyysmittaa.

Todennäköisyysjakauman määritelmä: Kommentteja

- *Todennäköisyysjakauma* kuvaa otosavaruuden **toden-näköisyysmassan** (= 1) *jakautumista* otosavaruudessa määritellyn satunnaismuuttujan arvoalueelle.
- *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**.

Satunnaisilmiöt ja niiden tilastolliset mallit

- Tilastotiede kehittää *menetelmiä* ja *malleja*, joiden avulla reaalimaailmasta pyritään tekemään *johtopäätöksiä* reaalimaailmaa kuvaavien *numeeristen tietojen* perusteella tilanteissa, joissa tietoihin sisältyy *epävarmuutta* ja *satunnaisuutta*.
- Tilastollisten menetelmien ja mallien avulla pyritään *erottamaan* ja *kuvaamaan* reaalimaailman ilmiöiden *säännönmukaiset* ja *satunnaiset* piirteet.
- Koska tilastotieteen tutkimiin ilmiöihin sisältyy epävarmuutta ja satunnaisuutta, tilastolliset menetelmät ja mallit perustuvat *todennäköisyyslaskentaan*.

Todennäköisyysjakaumat tilastollisina malleina

- Satunnaisilmiön *tilastollinen malli* kuvaa ilmiön tulosvaihtoehtoja ja niiden todennäköisyyksiä *matemaattisessa muodossa*.
- Satunnaisilmiön **tilastollisessa mallissa** eli **todennäköisyysmallissa** on seuraavat osat:
  (1) Satunnaisilmiön tulosvaihtoehtoja numeerisessa muodossa kuvaava **satunnaismuuttuja**.
  (2) Otosavaruuden todennäköisyysmassan jakautumista satunnaismuuttujan arvoalueelle kuvaava **todennäköisyysjakauma**.

Tilastollisten mallien muodostaminen

- Kun satunnaisilmiölle *konstruoidaan* tilastollinen malli, vaaditaan *tilastotieteen* ja *todennäköisyyslaskennan* tietojen lisäksi hyviä tietoja ilmiötä selittävästä *taustateoriasta*.
- Taustateorian tuottaa se *tieteenala*, jonka alueeseen ilmiö kuuluu.

  Esimerkiksi *taloustiede* toimii taustateoriana taloudellisten ilmiöiden tilastollisessa analyysissa eli *ekonometriassa*.
- *Tilastollinen tutkimus on parhaimmillaan* tilastotieteen, todennäköisyyslaskennan ja tutkimuksen kohteena olevaa ilmiötä selittävän taustateorian *yhteispeliä*.

Tilastollinen malli *vs* havainnot

- *Teoreettisen tilastotieteen* tehtävänä on konstruoida tutkimuksen kohteena olevalle satunnais-ilmiölle *tilastollinen malli*, joka *selittää* ilmiöstä saatujen *havaintojen käyttäytymisen*.
- *Empiirisen tilastotieteen* tehtävänä on selvittää, onko konstruoitu tilastollinen malli *sopusoinnussa* havaintojen kanssa.
- Huomaa, että tilastollinen malli on *teoreettinen oletus*, joka pitää aina asettaa *testiin havaintojen tuottamaa informaatioita vastaan*.

Satunnaismuuttujien tyyppejä

- **Satunnaismuuttuja** määriteltiin (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.

Satunnaismuuttujien tyyppejä: Diskreetit satunnaismuuttujat

- Satunnaismuuttujaa on **diskreetti**, jos sen arvoalue on *diskreetti joukko* eli sen arvoalue *muodostuu erillisistä reaaliakselin pisteistä*.
- Diskreetin satunnaismuuttujan arvoalue on joko *äärellinen* tai korkeintaan *numeroituvasti ääretön*.
- Diskreetin satunnaismuuttujan todennäköisyysjakauma määrittelee *alkeistapahtumien* todennäköisyydet.
- *Kaikkien* muiden tapahtumien todennäköisyydet saadaan alkeistapahtumien todennäköisyyksistä todennäköisyyden *laskusääntöjen avulla*.

Satunnaismuuttujien tyyppejä: Jatkuvat satunnaismuuttujat

- Satunnaismuuttujaa on **jatkuva**, jos sen arvoalue on jokin *reaaliakselin osaväli*.
- Jatkuvan satunnaismuuttujan arvoalue on reaalilukujen joukon osavälinä *ylinumeroituva*.
- Jatkuvan satunnaismuuttujan todennäköisyysjakauma määrittelee satunnaismuuttujan arvoalueeseen kuuluvien reaaliakselin *välien* todennäköisyydet.
- *Kaikkien* muiden tapahtumien todennäköisyydet saadaan näiden välien todennäköisyyksistä todennäköisyyden *laskusääntöjen avulla*.

Satunnaismuuttujat ja todennäköisyysjakaumat

Satunnaismuuttujat ja niiden todennäköisyysjakaumat

>> Diskreetit satunnaismuuttujat ja niiden todennäköisyysjakaumat

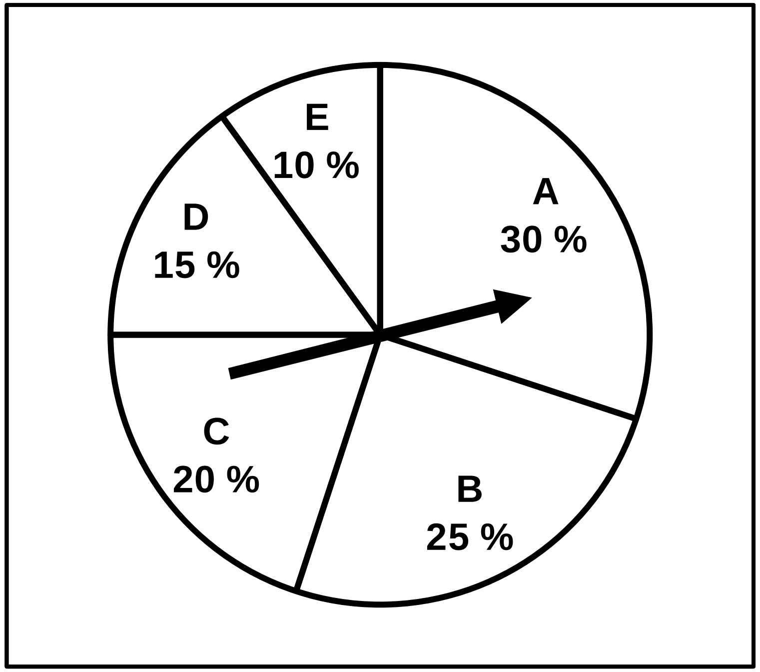
Jatkuvat satunnaismuuttujat ja niiden todennäköisyysjakaumat

Johdatteleva esimerkki: Onnenpyörä − 1/10

- Kuva oikealla esittää onnenpyörää, jonka pinta on jaettu viiteen sektoriin

  A, B, C, D, E
- Alla on esitetty sektoreiden pinta-alojen osuudet onnen-pyörän kokonaispinta-alasta:

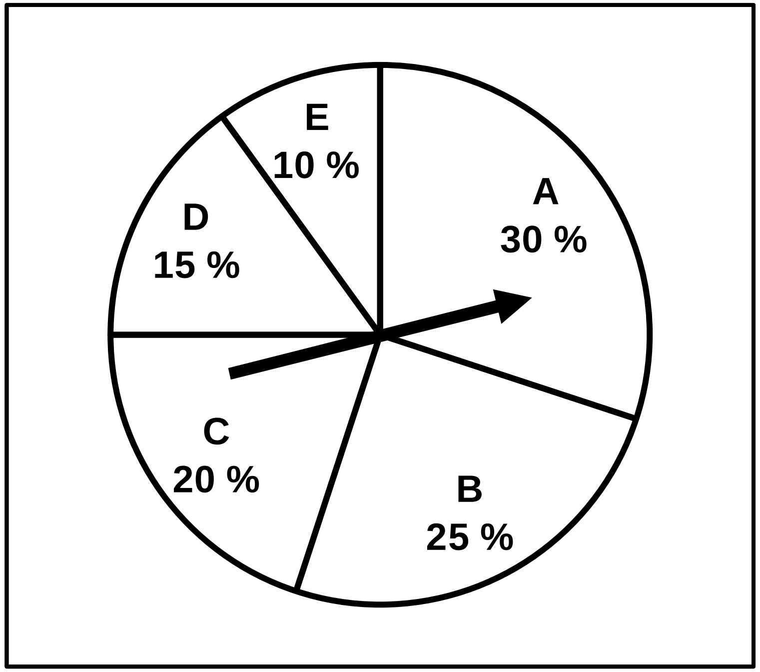
Sektori	%
A	30
B	25
C	20
D	15
E	10
Summa	100



Johdatteleva esimerkki: Onnenpyörä − 2/10

- Onnenpyörän keskipisteeseen on asetettu vapaasti pyörivä osoitin.
- Tarkastellaan peliä, jossa osoitinta pyöräytetään ja pelaaja yrittää *arvata* mihin sektoriin osoitin pysähtyy.
- Peliä voidaan pitää *satunnaisilmiönä*, johon liittyvä **otosavaruus** eli *mahdollisten tulosvaihtoehtojen joukko* on

  S = {Sektorit A, B, C, D, E}



Johdatteleva esimerkki: Onnenpyörä − 3/10

- Oletetaan, että *todennäköisyydet, joilla osoitin pysähtyy sektoreihin* A, B, C, D, E, *suhtautuvat toisiinsa kuten sektoreiden pinta-alat*.
- Tällöin voimme asettaa:

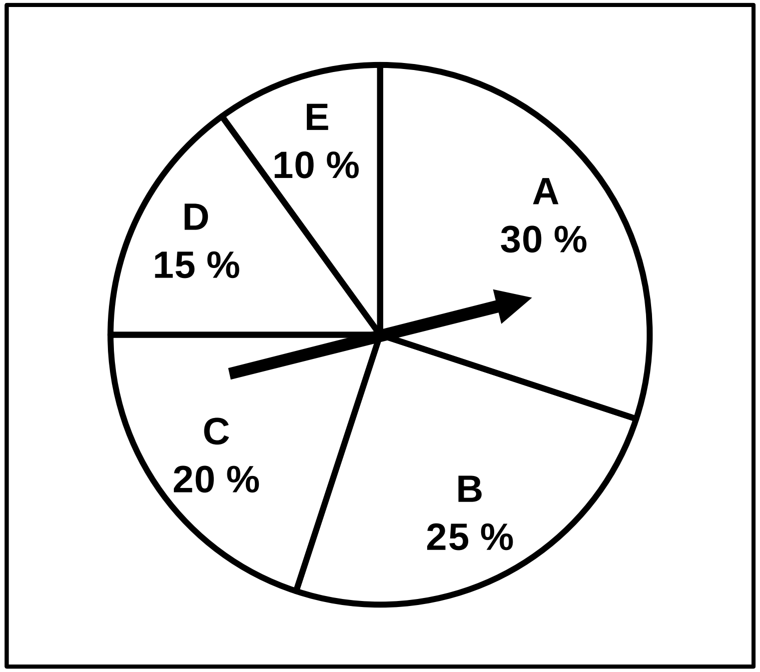
  Pr(A) = 0.30

  Pr(B) = 0.25

  Pr(C) = 0.20

  Pr(D) = 0.15

  Pr(E) = 0.10



Johdatteleva esimerkki: Onnenpyörä − 4/10

- Määritellään **satunnaismuuttuja** ξ, joka liittää tulosvaihtoehtoihin

  A, B, C, D, E

  reaaliluvut seuraavalla tavalla:

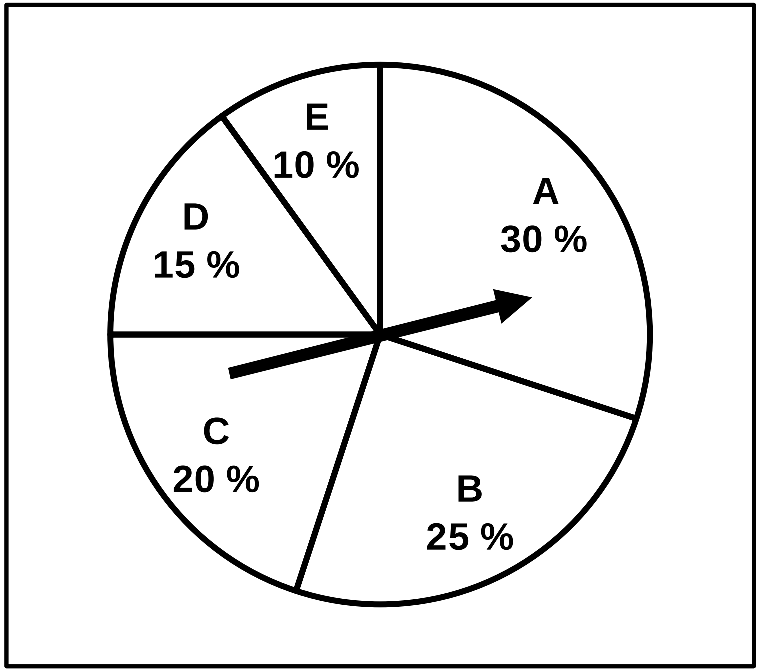
  A $\rightarrow$ 1

  B $\rightarrow$ 2

  C $\rightarrow$ 3

  D $\rightarrow$ 4

  E $\rightarrow$ 5



Johdatteleva esimerkki: Onnenpyörä – 5/10

- Satunnaismuuttuja ξ saa arvonsa seuraavilla todennäköisyyksillä:

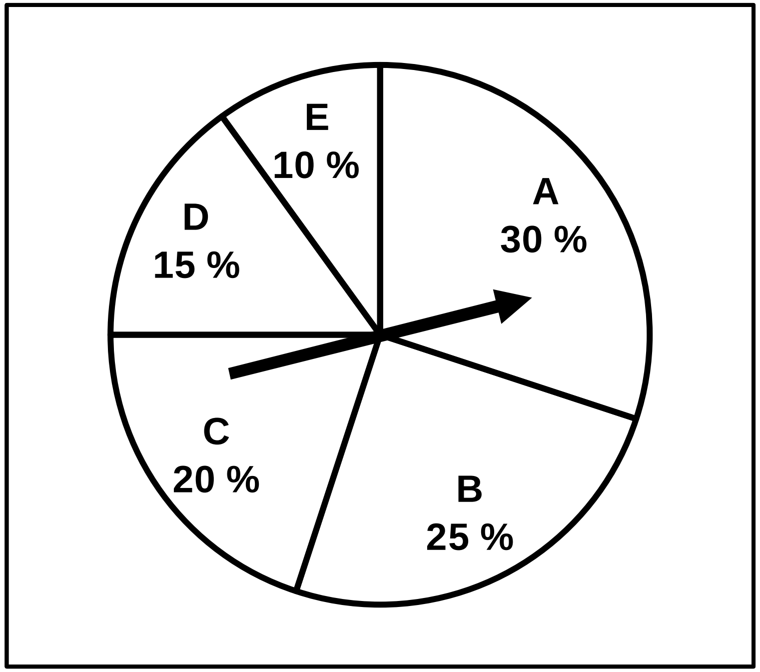
  $\Pr(\xi = 1) = 0.30 = \Pr(A)$

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

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

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

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



Johdatteleva esimerkki: Onnenpyörä − 6/10

- Satunnaismuuttujaa ξ sanotaan **diskreetiksi**, koska se saa vain erillisiä arvoja.
- Satunnaismuuttujan ξ arvoihin liittyviä todennäköisyyksiä sanotaan **pistetodennäköisyyksiksi**, koska ne liittyvät erillisiin pisteisiin reaaliakselilla.
- Diskreetin satunnaismuuttujan ξ arvot ja niihin liittyvät piste-todennäköisyydet muodostavat satunnaismuuttujan ξ **todennäköisyys-jakauman**.

Johdatteleva esimerkki: Onnenpyörä − 7/10

- Diskreetin satunnaismuuttujan ξ todennäköisyysjakaumaa voidaan kuvata sen **pistetodennäköisyysfunktiolla**.
- Diskreetin satunnaismuuttujan ξ pistetodennäköisyysfunktio kertoo miten otosavaruuden S *todennäköisyysmassa* (= 1) *jakautuu* satunnaismuuttujan ξ mahdollisille arvoille.
- Pistetodennäköisyysfunktio f on funktiona *epäjatkuva* ja se saa *positiivisia arvoja* vain *erillisissä pisteissä*.

Johdatteleva esimerkki: Onnenpyörä − 8/10

- Esimerkin tapauksessa satunnaismuuttujan ξ *pistetodennäköisyys-funktio* f voidaan määritellä seuraavalla tavalla:

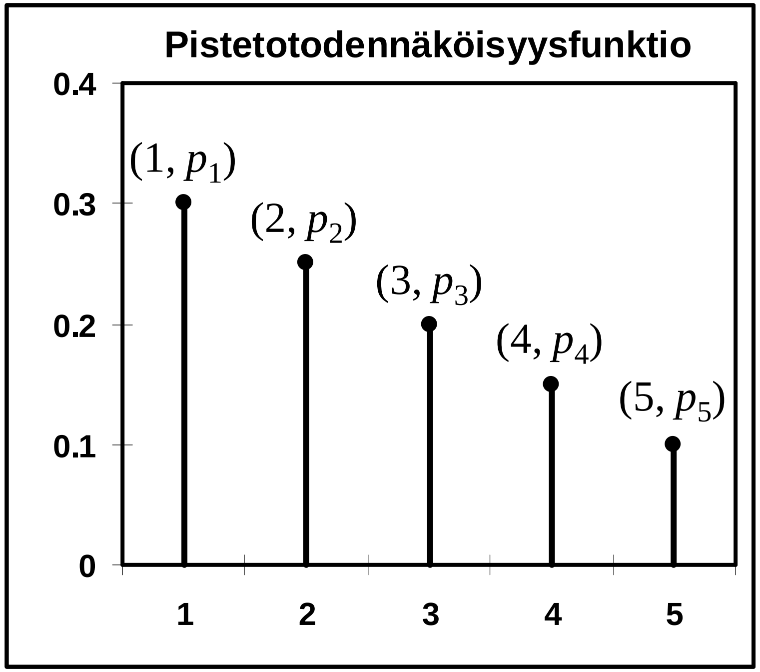
$$
\begin{aligned}
f(1) &= \Pr(\xi = 1) = 0.30 = \Pr(A) \\
f(2) &= \Pr(\xi = 2) = 0.25 = \Pr(B) \\
f(3) &= \Pr(\xi = 3) = 0.20 = \Pr(C) \\
f(4) &= \Pr(\xi = 4) = 0.15 = \Pr(D) \\
f(5) &= \Pr(\xi = 5) = 0.10 = \Pr(E)
\end{aligned}
$$

Johdatteleva esimerkki: Onnenpyörä − 9/10

- Olkoon x diskreetin satunnaismuuttujan ξ mahdollinen arvo ja
  $$\Pr(\xi = x) = p_x$$
  vastaava pistetodennäköisyys.
- Satunnaismuuttujan ξ *pistetodennäköisyysfunktiota voidaan kuvata graafisesti* **piikkifunktioksi** kutsutulla kuviolla, joka saadaan yhdistämällä pisteet
  $$(x,0) \text{ ja } (x, p_x)$$
  toisiinsa janoilla.

Johdatteleva esimerkki: Onnenpyörä – 10/10

- Kuvio oikealla esittää esimerkin diskreetin satunnais-muuttujan ξ todennäköisyys-jakauman pistetodennäköisyys-funktiota vastaavaa *piikki-funktiota*.
- Piikkien pituudet vastaavat niitä todennäköisyyksiä, joilla satunnaismuuttuja ξ saa arvonsa.



Äärelliset ja numeroituvasti äärettömät otosavaruudet

- Olkoon otosavaruus S *äärellinen* tai *numeroituvasti ääretön* joukko ja olkoon s_i otosavaruuden alkeistapahtuma.
- Tällöin voimme merkitä

$$S = \{s_1, s_2, \ldots, s_n\}$$

jos S on *äärellinen* ja

$$S = \{s_1, s_2, s_3, \ldots\}$$

jos S on *numeroituvasti ääretön*.

Diskreetti satunnaismuuttuja: Määritelmä

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

$$\xi : S \to \mathrm{R}$$

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

Esimerkki 1: Laadunvalvonta

- Kone tekee erästä tuotetta sarjatuotantona n kpl päivässä.
- Tuotteista osa on viallisia.
- Vialliset tuotteet syntyvät valmistusprosessissa täysin sattumanvaraisesti.
- Oletetaan, että viallisten tuotteiden suhteellinen osuus valmistetuista tuotteista on keskimäärin p.
- Tällöin voimme asettaa:

  p = todennäköisyys, että satunnaisesti valittu tuote on viallinen
- *Viallisten tuotteiden lukumäärä päivän aikana tehtyjen tuotteiden joukossa* on diskreetti satunnaismuuttuja, joka noudattaa **binomi-jakaumaa**; ks. lukua **Diskreettejä jakaumia**.

Esimerkki 2: Laadunvalvonta

- Kone tekee erästä tuotetta sarjatuotantona.
- Tuotteista osa on viallisia.
- Vialliset tuotteet syntyvät valmistusprosessissa täysin sattumanvaraisesti.
- Oletetaan, että viallisten tuotteiden suhteellinen osuus valmistetuista tuotteista on keskimäärin p.
- Tällöin voimme asettaa:

  p = todennäköisyys, että satunnaisesti valittu tuote on viallinen
- Poimitaan tuotteita tarkastettavaksi, *kunnes löydetään ensimmäinen viallinen*.
- *Ensimmäisen viallisen tuotteen järjestysnumero tarkastettujen tuotteiden joukossa* on diskreetti satunnaismuuttuja, joka noudattaa **geometrista jakaumaa**; ks. lukua **Diskreettejä jakaumia**.

Esimerkki 3:
Jonot

- Palvelujonoon tulee asiakkaita keskimääriin k kappaletta aikayksikköä kohti.
- *Jonakin aikavälinä jonoon tulevien asiakkaiden lukumäärä* on diskreetti satunnaismuuttuja, joka noudattaa (tietyin oletuksin) **Poisson-jakaumaa**; ks. lukua **Diskreettejä jakaumia**.
- Huomautus:

    Tällöin 1. asiakkaan *odotusaika* on *jatkuva satunnaismuuttuja*, joka noudattaa **eksponenttijakaumaa**; ks. lukua **Jatkuvia jakaumia**.

Esimerkki 4: Kalakannan koon määrääminen

- Pyydystetään järvestä joukko kaloja elävinä, merkitään pyydystetyt kalat ja lasketaan ne takaisin järveen.
- Pyydystetään järvestä jonkin ajan kuluttua uusi joukko kaloja.
- *Merkittyjen kalojen lukumäärä uudessa pyynnissä* on diskreetti satunnaismuuttuja, joka noudattaa **hypergeometrista jakaumaa**; ks. lukua **Diskreettejä jakaumia**.
- Huomautus:

  Kuvattua *merkintä-takaisinpyynti-menetelmää* voidaan soveltaa riista- ja kalakantojen laskemiseen.

Diskreetit suureet

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

Pistetodennäköisyysfunktio: Määritelmä 1/2

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

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

$$x_i \,,\ i = 1, 2, 3, \ldots\,, \text{ jos } S \text{ on numeroituvasti ääretön}$$

- Merkitään diskreetin satunnaismuuttujan ξ *arvojen joukkoa* kirjaimella T:

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

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

Pistetodennäköisyysfunktio: Määritelmä 2/2

- Reaaliarvoinen funktio f määrittelee **pistetodennäköisyys-funktion** *diskreetille satunnaismuuttujalle* ξ, jos

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

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

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

Pistetodennäköisyysfunktio: Toinen määritelmä

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

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

satunnaismuuttujan ξ arvoa x_i vastaava *pistetodennäköisyys*.

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

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

Pistetodennäköisyysfunktion määritelmä: Kommentteja

- Pistetodennäköisyysfunktio f on *epäjatkuva* funktio.
- Määritelmän ehdon (2) mukaan pistetodennäköisyys-funktio f on kaikkialla *ei-negatiivinen*.
- Määritelmän ehto (2) on välttämätön ehto, koska piste-todennäköisyydet ovat *todennäköisyyksiä*.
- Määritelmän ehdon (3) mukaan *kaikkien* pistetodennäköi-syyksien *summa* = 1.

Diskreetti todennäköisyysjakauma: Määritelmä

- Diskreetin satunnaismuuttujan ξ *pistetodennäköisyys-funktio* f määrittelee satunnaismuuttujan ξ **toden-näköisyysjakauman**.
- Jos f on diskreetin satunnaismuuttujan ξ pistetoden-näköisyysfunktio, sanomme tavallisesti, että ξ **noudattaa diskreettiä todennäköisyysjakaumaa** f .

Diskreetin todennäköisyysjakauman määritelmä: Kommentteja

- Diskreetin satunnaismuuttujan ξ *pistetodennäköisyys-funktio* f kertoo miten otosavaruuden S *todennäköisyys-massa* (= 1) *jakautuu* satunnaismuuttujan ξ *arvoille*.
- Diskreetin satunnaismuuttujan ξ pistetodennäköisyys-funktion avulla voidaan määrätä *kaikki* ko. satunnais-ilmiöön liittyvät *todennäköisyydet*.

Pistetodennäköisyysfunktion kuvaaja: Piikkifunktio 1/2

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

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

satunnaismuuttujan ξ arvoa x_i vastaava *pistetodennäköisyys*.

- Pistetodennäköisyysfunktiota f voidaan *kuvata graafisesti* ns. **piikkifunktiolla**.

Pistetodennäköisyysfunktion kuvaaja: Piikkifunktio 2/2

- Pistetodennäköisyysfunktiota

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

  vastaava **piikkifunktio** piirretään seuraavalla tavalla:

  (i) Merkitään vastinpisteet

  $$(x_i \ , 0) \text{ ja } (x_i \ , p_i)$$

  kaikille $x_i \in T$ tasoon.

  (ii) Yhdistetään vastinpisteet

  $$(x_i \ , 0) \text{ ja } (x_i \ , p_i)$$

  kaikille $x_i \in T$ toisiinsa janalla.

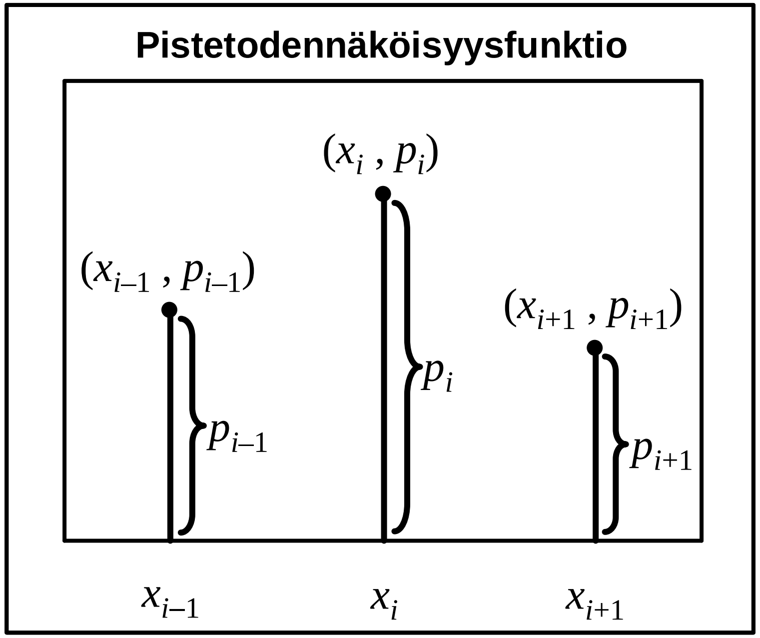
Pistetodennäköisyysfunktion kuvaaja: Havainnollistus

- Satunnaismuuttujan ξ pistetodennäköisyysfunktiota

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

  kuvataan graafisesti piirtämällä pisteisiin x_i ”*piikit*”, joiden pituudet vastaavat pistetodennäköisyyksiä

  $\Pr(\xi = x_i) = p_i$
- ”Piikin” kärkipiste $(x_i\ ,\ p_i)$ piirretään tavallisesti niin, että se erottuu selvästi pisteitä $(x_i\ ,\ 0)$ ja $(x_i\ ,\ p_i)$ yhdistävästä janasta.



Pistetodennäköisyyksien vertailu

- Olkoon

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

- Jos

$$p_1 > p_2$$

niin tulosvaihtoehto x_1 on *todennäköisempi kuin* tulosvaihtoehto x_2.

Reaaliakselin välien todennäköisyydet

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

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

- Summassa *lasketaan yhteen* kaikki pistetodennäköisyydet

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

joita vastaavat satunnaismuuttujan ξ arvot $x_i \in [a, b]$.

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

Tapahtumien todennäköisyydet

- Todennäköisyyden *aksioomia* käsittelevässä luvussa todetaan seuraavaa:
  (1) Jokainen *tapahtuma* reaalilukujen joukossa voidaan muodostaa reaaliakselin väleistä yhdistelemällä välejä sopivasti joukko-opin operaatioilla.
  (2) Jokaisen *tapahtuman todennäköisyys* saadaan reaaliakselin välien todennäköisyyksistä todennäköisyyslaskennan laskusääntöjen avulla.
- Jos siis diskreetin satunnaismuuttujan pistetodennäköisyysfunktio tunnetaan, se *satunnaisilmiö*, *jonka tulosvaihtoehtoja ko. satunnaismuuttuja kuvaa*, *hallitaan täydellisesti*.

Diskreetin jakauman parametrointi 1/2

- Oletetaan, että satunnaismuuttuja ξ noudattaa diskreettiä todennäköisyysjakaumaa, jonka *pistetodennäköisyys-funktio* on f .
- Pistetodennäköisyysfunktio f riippuu tavallisesti **para-metreista** eli *vakioista*, jotka määräävät funktion f *muodon*.
- Pistetodennäköisyysfunktion muodon määräviä para-metreja kutsutaan usein sen **todennäköisyysjakauman parametreiksi**, jonka todennäköisyydet ko. pistetoden-näköisyysfunktio määrittelee.

Diskreetin jakauman parametrointi 2/2

- Olkoot pistetodennäköisyysfunktion f muodon määräävät parametrit

$$\theta_1, \theta_2, \ldots , \theta_p$$

- Jos halutaan korostaa pistetodennäköisyysfunktion f riippuvuutta sen muodon määräävistä parametreista, pistetodennäköisyysfunktiota *merkitään*:

$$f(x\,;\,\theta_1, \theta_2, \ldots , \theta_p)$$

Diskreetin jakauman parametrien merkitys

- Diskreetin satunnaismuuttujan ξ *pistetodennäköisyys-funktion* f muodon määräävillä parametreilla $\theta_1, \theta_2, \ldots, \theta_p$ on usein jokin sisällöllinen *tulkinta* siinä satunnais-ilmiössä, jonka malliksi f on konstruoitu.
- Parametrien $\theta_1, \theta_2, \ldots, \theta_p$ arvot ovat sovelluksissa tavallisesti *tuntemattomia* ja niiden arvot on **estimoitava** eli arvioitava havaintojen perusteella.

  Ks. monisteen **Tilastolliset menetelmät** lukua **Estimointi**.
- Parametreja $\theta_1, \theta_2, \ldots, \theta_p$ koskeva *hypoteeseja* eli oletuksia voidaan **testata** tilastollisin testein.

  Ks. monisteen **Tilastolliset menetelmät** lukua **Tilastollinen testaus**.

Esimerkki: Kuutosen odottaminen nopanheitossa Tehtävät (a) − (d)

- Heitetään *virheetöntä* noppaa, kunnes saadaan *kuutonen*.
- Olkoon

  ξ = ”Sen heiton numero, jolla saadaan *ensimmäinen kuutonen*”
- ξ on *diskreetti satunnaismuuttuja*.
- Ratkaistaan seuraavat tehtävät:

  (a) Määrää satunnaismuuttujan ξ *todennäköisyysjakauma*.

  (b) Määrää todennäköisyys, että noppaa joudutaan heittämään *täsmälleen* 6 *kertaa*.

  (c) Määrää todennäköisyys, että noppaa joudutaan heittämään 6 *kertaa tai enemmän*.

  (d) Määrää todennäköisyys, että noppaa joudutaan heittämään 36 *kertaa tai enemmän*.

Esimerkki: Kuutosen odottaminen nopanheitossa Oletukset

- Oletamme, että nopanheitot ovat toisistaan *riippumattomia* siinä mielessä, että yhdenkään heiton tulos ei riipu aikaisempien heittojen tuloksista.
- Määritellään tapahtumat:

  A = ”Noppaa heitettäessä tuloksena *saadaan kuutonen*”

  A^c = ”Noppaa heitettäessä tuloksena *ei saada kuutosta*”
- Koska noppa oletettiin *virheettömäksi*,

$$\Pr(A) = \frac{1}{6}$$

- *Komplementtitapahtuman* todennäköisyyden kaavan mukaan

$$\Pr(A^c) = 1 - \Pr(A) = \frac{5}{6}$$

Esimerkki: Kuutosen odottaminen nopanheitossa Tehtävän (a) ratkaisu

- Tehtävä (a):

  Määrää satunnaismuuttujan.

  ξ = ”Sen heiton numero, jolla saadaan *ensimmäinen kuutonen*”

  todennäköisyysjakauma.

- Koska *ensimmäinen kuutonen* voidaan saada ensimmäisellä, toisella, kolmannella jne. heitolla, ξ on *diskreetti satunnaismuuttuja*, joka saa *kaikki* kokonaislukuarvot 1, 2, 3, …
- Seuraavassa osoitetaan, että satunnaismuuttujan ξ *todennäköisyys-jakauman pistetodennäköisyysfunktio* on

$$\Pr(\xi = i) = \frac{1}{6}\cdot\left(\frac{5}{6}\right)^{i-1},\ i = 1, 2, 3, \ldots$$

- Kaava ratkaisee tehtävän (a).

Esimerkki: Kuutosen odottaminen nopanheitossa Pistetodennäköisyysfunktion johto − 1/2

- Satunnaismuuttujan ξ *todennäköisyysjakauman pistetodennäköisyys-funktion* lausekkeen perustelu:

(1) Oletetaan, että *ensimmäinen kuutonen* saadaan *i*. heitolla, $i = 1, 2, 3, \ldots$

(2) Tällöin *ennen i*. heittoa on täytynyt tapahtua $i - 1$ heittoa, joiden tuloksena ei ole saatu kuutosta:

$$\begin{array}{llllll} \text{Heiton numero:} & 1 & 2 & \cdots & i-1 & i \\ \text{Tapahtuma:} & A^c & A^c & \cdots & A^c & A \end{array}$$

Esimerkki: Kuutosen odottaminen nopanheitossa Pistetodennäköisyysfunktion johto − 2/2

(3) Tapahtumajonon

$$\underbrace{A^c A^c \cdots A^c}_{(i-1)\ \text{kpl}} A$$

todennäköisyys on *riippumattomien tapahtumien tulosäännön* nojalla

$$\underbrace{\frac{5}{6}\cdot\frac{5}{6}\cdots\frac{5}{6}}_{(i-1)\ \text{kpl}}\cdot\frac{1}{6} = \frac{1}{6}\cdot\left(\frac{5}{6}\right)^{i-1}$$

- Kaava pätee kaikille $i = 1, 2, 3, \ldots$ ja ratkaisee siis tehtävän (a).

Esimerkki: Kuutosen odottaminen nopanheitossa Todennäköisyysjakauma

- Pistetodennäköisyydet

$$\Pr(\xi = i) = \frac{1}{6} \cdot \left(\frac{5}{6}\right)^{i-1}, \; i = 1, 2, 3, \ldots$$

kelpaavat todennäköisyyksiksi, koska ne muodostavat *suppenevan geometrisen sarjan*, jonka summa = 1:

$$\sum_{i=1}^{\infty} \Pr(\xi = i) = \frac{1}{6} \sum_{i=1}^{\infty} \left(\frac{5}{6}\right)^{i-1} = \frac{1}{6} \sum_{i=0}^{\infty} \left(\frac{5}{6}\right)^{i} = \frac{1}{6} \cdot \frac{1}{1 - \frac{5}{6}} = 1$$

- Koska pistetodennäköisyydet $\Pr(\xi = i)$ muodostavat geometrisen sarjan, satunnaismuuttujan ξ todennäköisyysjakaumaa sanotaan **geometriseksi jakaumaksi** (lisätietoja: ks. lukua **Diskreettejä jakaumia**).

Esimerkki: Kuutosen odottaminen nopanheitossa
Pistetodennäköisyydet − 1/2

- Geometrisen jakauman

$$\Pr(\xi = i) = \frac{1}{6} \cdot \left(\frac{5}{6} \right)^{i-1} \ , \ i = 1, 2, 3, \ldots$$

pistetodennäköisyydet *vähenevät eksponentiaalista vauhtia* eli *kuten suppenevassa geometrisessa sarjassa*, kun *i* kasvaa.

- Tämä merkitsee seuraavaa:

(1) Todennäköisyys, että *ensimmäinen kuutonen* saadaan heti ensimmäisellä heitolla *on* kaikkein *suurin*.

(2) Todennäköisyys, että *ensimmäinen kuutonen* saadaan *i*. heitolla, *pienenee* heittojen lukumäärän *i* funktiona *eksponentiaalista vauhtia*.

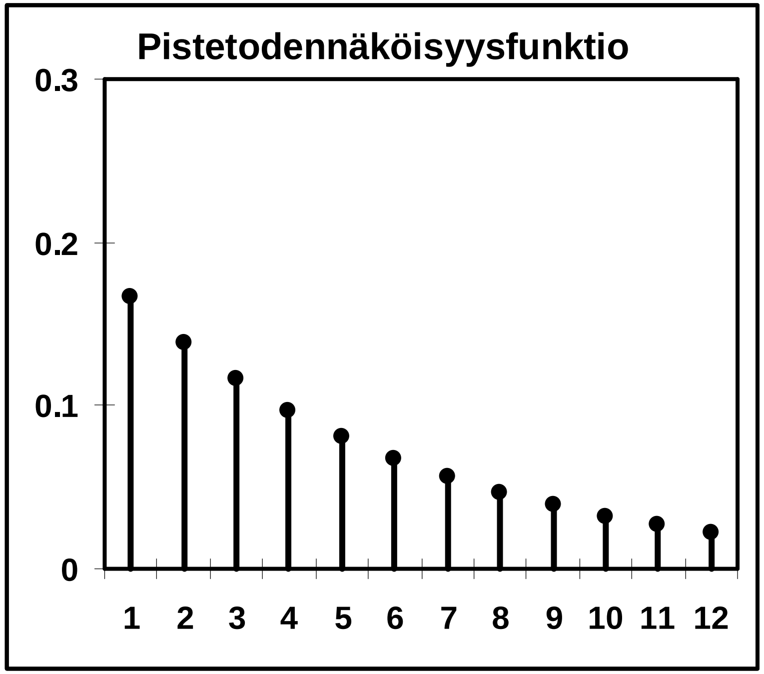
Esimerkki: Kuutosen odottaminen nopanheitossa Pistetodennäköisyydet − 2/2

- Viereinen taulukko ja kuvio esittävät esimerkissä johdetun geometrisen jakauman pistetodennäköisyyksiä

$$\Pr(\xi = i) = \frac{1}{6} \cdot \left(\frac{5}{6}\right)^{i-1}$$

kun $i = 1, 2, \ldots, 12$.

i	$\Pr(\xi = i)$
1	0.1667
2	0.1389
3	0.1157
4	0.0965
5	0.0804
6	0.0670
7	0.0558
8	0.0465
9	0.0388
10	0.0323
11	0.0269
12	0.0224



Esimerkki: Kuutosen odottaminen nopanheitossa Tehtävän (b) ratkaisu

- Tehtävä (b):

  Mikä on todennäköisyys, että *ensimmäinen kuutonen* saadaan 6. heitolla?

- Satunnaismuuttujan ξ pistetodennäköisyysfunktiosta saadaan:

$$\Pr(\xi = 6) = \frac{1}{6} \cdot \left(\frac{5}{6} \right)^5 \approx 0.0670$$

- *Todennäköisyyden frekvenssitulkinnan* mukaan tämä merkitsee seuraavaa:

  Jos 1000 henkilöä heittää noppaa, *keskimäärin* n. 70 saa *ensimmäisen kuutosen* 6. heitolla.

Esimerkki: Kuutosen odottaminen nopanheitossa Aputulos – 1/2

- Edellä esitettyjen tehtävien (c) ja (d) ratkaisemiseksi johdetaan seuraava aputulos:

  Todennäköisyys sille, että noppaa joudutaan heittämään *k kertaa tai enemmän* ennen *kuutosen* saamista on

$$\Pr(\xi \geq k) = \left(\frac{5}{6}\right)^{k-1}, \quad k = 1, 2, 3, \ldots$$

Esimerkki: Kuutosen odottaminen nopanheitossa
Aputulos – 2/2

- *Komplementtitapahtuman todennäköisyyden* ja *geometrisen sarjan osasumman kaavojen* nojalla ko. todennäköisyydeksi saadaan:

$$
\begin{aligned}
\Pr\left(\xi \geq k\right) &= 1 - \Pr(\xi < k) \\
&= 1 - \sum_{i=1}^{k-1} \Pr(\xi = i) \\
&= 1 - \frac{1}{6} \sum_{i=1}^{k-1} \left(\frac{5}{6}\right)^{i-1} \\
&= 1 - \frac{1}{6} \cdot \frac{1 - \left(5/6\right)^{k-1}}{1 - 5/6} \\
&= \left(\frac{5}{6}\right)^{k-1}
\end{aligned}
$$

Esimerkki: Kuutosen odottaminen nopanheitossa
Tehtävän (c) ratkaisu

- Tehtävä (c):

  Määrää todennäköisyys, että noppaa joudutaan heittämään 6 *kertaa tai enemmän* ennen *ensimmäisen kuutosen* saamista.

- Todennäköisyys, että noppaa joudutaan heittämään 6 *kertaa tai enemmän* ennen *ensimmäisen kuutosen* saamista on aputuloksen nojalla

$$\Pr(\xi \geq 6) = \left(\frac{5}{6}\right)^5 \approx 0.402$$

- *Todennäköisyyden frekvenssintulkinnan* mukaan tämä merkitsee seuraavaa:

  Suuressa joukossa ihmisiä *keskimäärin* 40 % joutuu heittämään noppaa 6 *kertaa tai enemmän* ennen kuin he saavat *ensimmäisen kuutosen*.

Esimerkki: Kuutosen odottaminen nopanheitossa Tehtävän (d) ratkaisu

- Tehtävä (d):

  Määrää todennäköisyys, että noppaa joudutaan heittämään 36 *kertaa tai enemmän* ennen *ensimmäisen kuutosen* saamista.

- Todennäköisyys, että noppaa joudutaan heittämään 36 *kertaa tai enemmän* ennen *ensimmäisen kuutosen* saamista on aputuloksen nojalla

$$\Pr(\xi \geq 36) = \left(\frac{5}{6}\right)^{35} \approx 0.00169$$

- *Todennäköisyyden frekvenssintulkinnan* mukaan tämä merkitsee seuraavaa:

  Jos 1000 henkilöä heittää noppaa, *keskimäärin* 1 – 2 henkilöä joutuu heittämään noppaa 36 *kertaa tai enemmän* ennen kuin he saavat *ensimmäisen kuutosen*.

Esimerkki: Kuutosen odottaminen nopanheitossa Tehtävän (d) ratkaisu: Kommentti

- Edellä todettiin, että jos 1000 henkilöä heittää noppaa, *keskimäärin* 1 – 2 henkilöä joutuu heittämään noppaa 36 *kertaa tai enemmän* ennen kuin he saavat *ensimmäisen kuutosen*.
- Nämä 1 – 2 henkilöä varmasti *ihmettelevät* heittosarjaansa.
- Tässä kohdataan *tilastotieteilijän selitys ihmeille*:

  Harvinaisetkin tapahtumat yleensä sattuvat ennemmin tai myöhemmin, jos satunnaisilmiö toistuu riittävän monta kertaa.

Satunnaismuuttujat ja todennäköisyysjakaumat

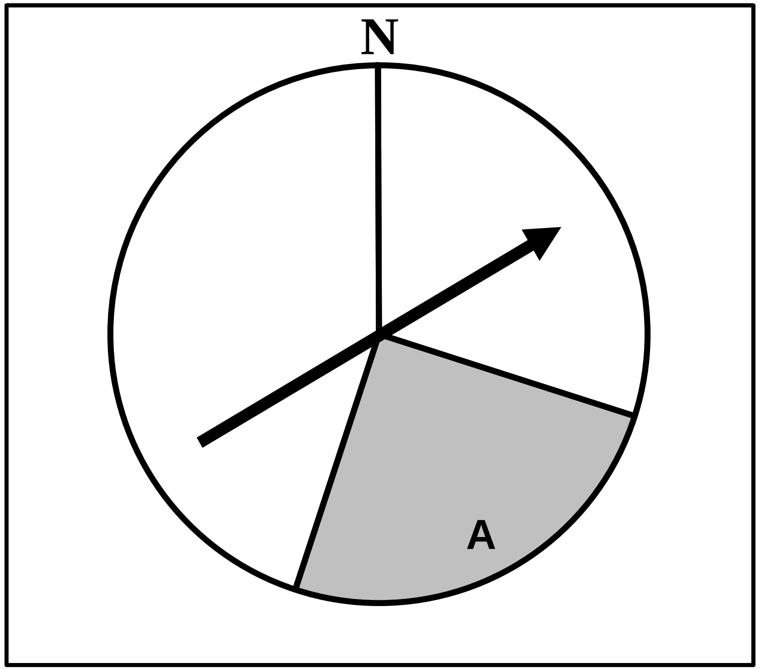
Satunnaismuuttujat ja niiden todennäköisyysjakaumat

Diskreetit satunnaismuuttujat ja niiden todennäköisyysjakaumat

>> Jatkuvat satunnaismuuttujat ja niiden todennäköisyysjakaumat

Johdatteleva esimerkki: Onnenpyörä – 1/6

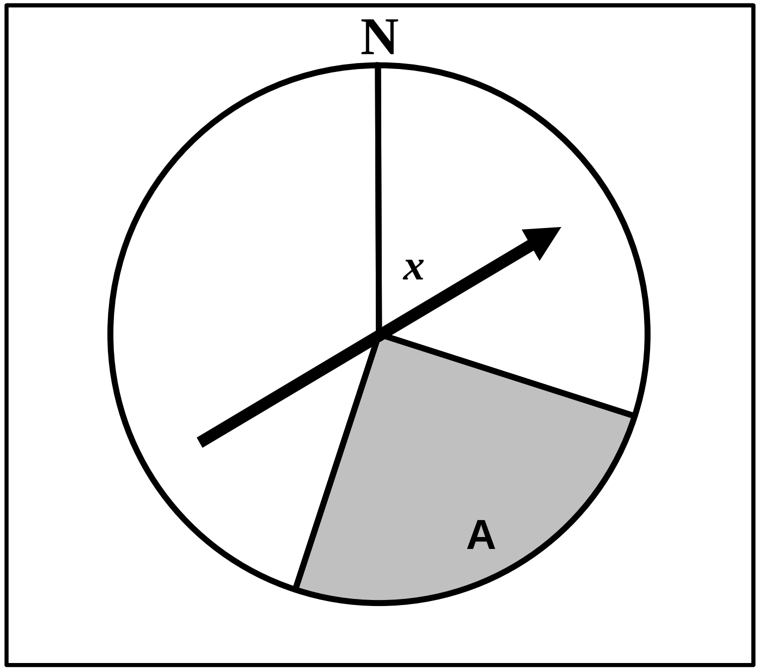
- Kuva oikealla esittää onnenpyörää, jonka keskipisteeseen on asetettu vapaasti pyörivä osoitin.
- Tarkastellaan peliä, jossa pelaaja valitsee onnenpyörästä *mielivaltaisen sektorin A.*
- Osoitinta pyöräytetään ja pelaaja voittaa, jos osoitin pysähtyy valitsemaansa sektoriin.
- Oletetaan, että *todennäköisyys voittaa on suhteessa valitun sektorin A pinta-alaan, mutta ei riipu sektorin A paikasta*.



Johdatteleva esimerkki: Onnenpyörä − 2/6

- Tarkastellaan **satunnais-ilmiönä** sitä kulmaa, jonka osoitin muodostaa pysähdyttyään N:llä (N = *n*orth) merkityn suunnan kanssa.
- Tähän satunnaisilmiöön liittyvä **otosavaruus** eli *mahdollisten tulosvaihtoehtojen joukko* on

  $S = \{\text{Kulma } x \mid x \in [0°, 360°)\}$

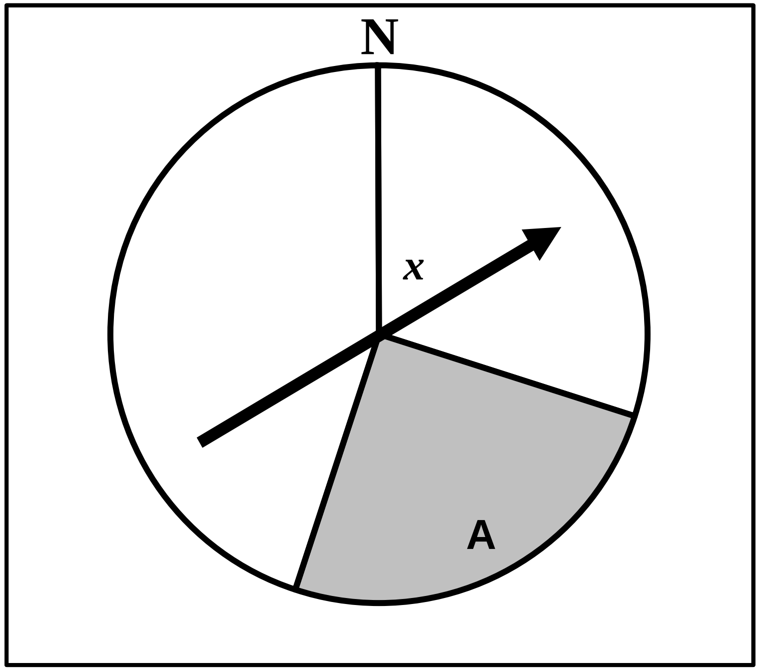


Johdatteleva esimerkki: Onnenpyörä – 3/6

- Määritellään **satunnaismuuttuja** ξ, joka liittää mahdollisiin tulosvaihtoehtoihin *reaaliluvun* seuraavasti:

  Kulma $x \to x \in [0, 360)$

- Satunnaismuuttujaa ξ sanotaan **jatkuvaksi**, koska ξ saa *kaikki arvot* väliltä [0, 360).



Johdatteleva esimerkki: Onnenpyörä – 4/6

- Tehdyn oletuksen mukaan todennäköisyys, että osoitin pysähtyy valitulle sektorille *riippuu vain sektorin leveydestä*, *mutta ei sektorin paikasta*.
- Tämä merkitsee sitä, että onnenpyörän osoittimen pysähtymistodennäköisyydet jakautuvat kaikille onnenpyörän samanlevyisille sektoreille siinä mielessä *tasaisesti*, että

  $$\Pr(\xi \in [a, b]) = \Pr(\xi \in [c, d])$$

  jos välin [0, 360) osavälit $[a, b]$ ja $[c, d]$ ovat *yhtä pitkiä* eli

  $$b - a = d - c$$

- Tällaista todennäköisyyksien jakaumaa kutsutaan **jatkuvaksi tasaiseksi jakaumaksi**; ks. lukua **Jatkuvia jakaumia**.

Johdatteleva esimerkki: Onnenpyörä − 5/6

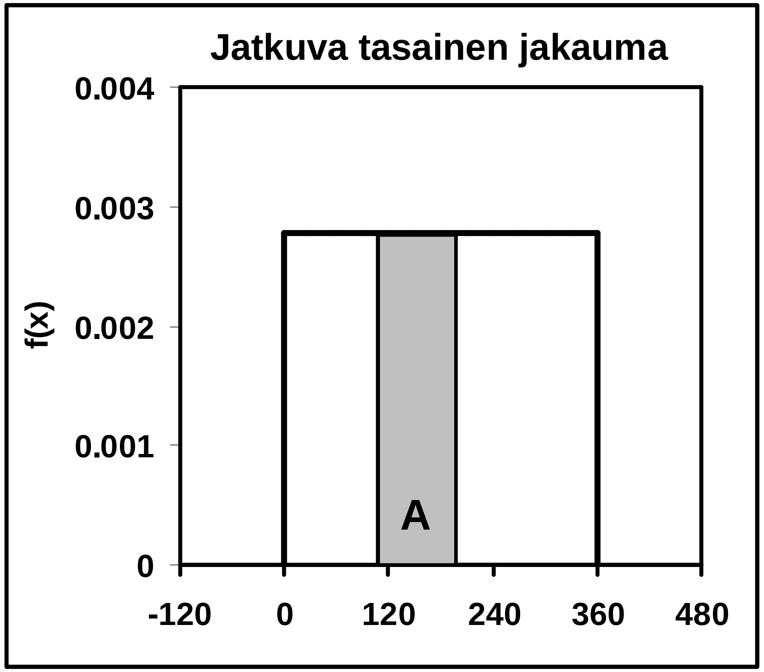
- Esimerkin jatkuvan tasaisen jakauman *todennäköisyyksien jakautumista* välillä [0, 360) voidaan kuvata jatkuvalla käyrällä

$$f(x) = \begin{cases} \dfrac{1}{360} & , \ x \in [0, 360) \\ 0 & , \ x \notin [0, 360) \end{cases}$$

- Käyrää vastaavaa funktiota kutsutaan jatkuvan tasaisen jakauman (*todennäköisyys*-) **tiheysfunktioksi**.

Johdatteleva esimerkki: Onnenpyörä – 6/6

- Kuvio oikealla esittää esimerkin *jatkuvan tasaisen jakauman tiheysfunktiota*.
- Kuvaan on merkitty myös alue, joka vastaa onnenpyörää esittäviin kuviin merkittyä sektoria *A*.



Jatkuva satunnaismuuttuja: Määritelmä

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

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

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

Esimerkki: Jono

- Palvelujonoon tulee asiakkaita keskimääriin k kappaletta aikayksikköä kohden.
- *Seuraavan jonoon tulevan asiakkaan odotusaika* on jatkuva satunnaismuuttuja, joka noudattaa (tietyin oletuksin) **eksponentti-jakaumaa**; ks. lukua **Jatkuvia jakaumia**.
- Huomautus:

  Tällöin *jonkin aikavälin aikana jonoon tulevien asiakkaiden lukumäärä* on *diskreetti satunnaismuuttuja*, joka noudattaa **Poisson-jakaumaa**; ks. lukua **Diskreettejä jakaumia**.

Jatkuvat suureet

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

Tiheysfunktio: Määritelmä

- Reaaliarvoinen funktio f määrittelee (**todennäköisyys-**) **tiheysfunktion** *jatkuvalle satunnaismuuttujalle* ξ, 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) $$\Pr(a \leq \xi \leq b) = \int_a^b f(x)dx$$

Tiheysfunktion määritelmä: Kommentteja

- Määritelmän ehdon (1) mukaan tiheysfunktio on *jatkuva*.
- Määritelmän ehdon (2) mukaan tiheysfunktio on kaikkialla *ei-negatiivinen*.
- Määritelmän ehto (2) on välttämätön ehto, koska ehdon (4) mukaan tiheysfunktion integraalit yli reaaliakselin välien ovat *todennäköisyyksiä*.
- Määritelmän ehdon (3) mukaan tiheysfunktion integraali yli koko reaaliakselin = 1.

Jatkuva todennäköisyysjakauma: Määritelmä

- Jatkuvan satunnaismuuttujan ξ *tiheysfunktio* f määrittelee satunnaismuuttujan ξ **todennäköisyysjakauman**.
- Jatkuvan satunnaismuuttujan ξ tiheysfunktio kertoo, miten otosavaruuden *S todennäköisyysmassa* (= 1) *jakautuu* satunnaismuuttujan ξ saamien *arvojen väleille*.
- Jatkuvan satunnaismuuttujan ξ tiheysfunktion avulla voidaan määrätä *kaikki* ko. satunnaisilmiöön liittyvät *todennäköisyydet*.

Tiheysfunktion kuvaaja

- Jatkuvan satunnaismuuttujan ξ tiheysfunktiota f voidaan *kuvata graafisesti* **jatkuvalla käyrällä**

$$(x, f(x))$$

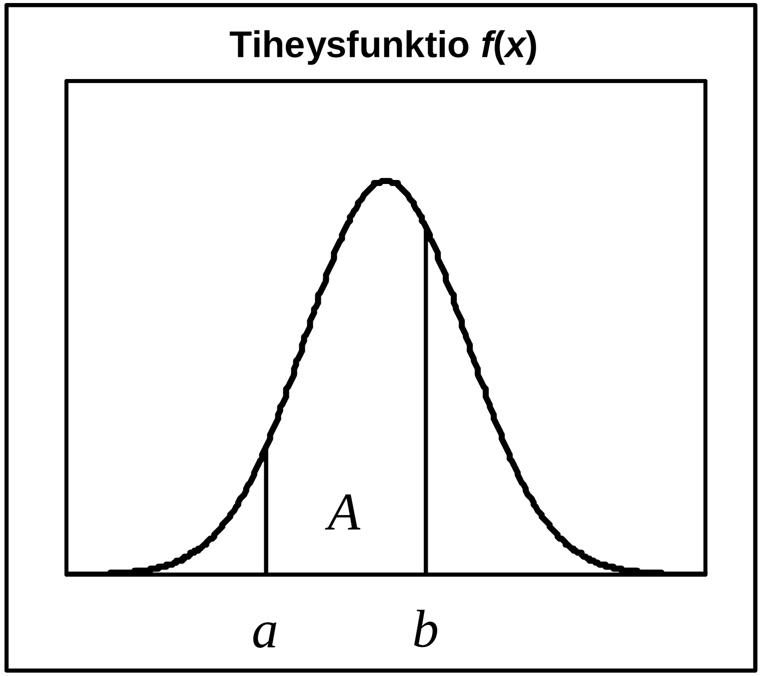
Tiheysfunktion kuvaaja: Kommentteja

- Ehdon (4) mukaan *reaaliakselin väleihin liittyvät todennäköisyydet* saadaan integroimalla tiheysfunktio ko. välillä.
- *Geometrisesti reaaliakselin väliin* $[a, b]$ *liittyvän todennäköisyyden* määrääminen merkitsee tiheysfunktion kuvaajan, x-akselin ja suorien $x = a$ ja $x = b$ rajoittaman tasoalueen *pinta-alan* laskemista.
- Määritelmän ehdon (3) mukaan tiheysfunktion kuvaajan ja x-akselin väliin jäävän tasoalueen *pinta-ala* = 1.

Tiheysfunktion kuvaaja: Havainnollistus

- Kuva oikealla esittää **normaalijakaumaa** (ks. lukua **Jatkuvia jakaumia**) noudattavan *jatkuvan satunnaismuuttujan* ξ *tiheysfunktiota* $f(x)$.
- Jatkuville satunnaismuuttujille ξ pätee yleisesti:

$$\begin{aligned}&\Pr(a \le \xi \le b)\\&= \int_a^b f(x)dx\\&= \text{Alueen } A \text{ pinta-ala}\end{aligned}$$



Tapahtumien todennäköisyydet

- Todennäköisyyden *aksioomia* käsittelevässä luvussa todetaan seuraavaa:
  - (1) Jokainen *tapahtuma* reaalilukujen joukossa voidaan muodostaa reaaliakselin väleistä yhdistelemällä välejä sopivasti joukko-opin operaatioilla.
  - (2) Jokaisen *tapahtuman todennäköisyys* saadaan reaali-akselin välien todennäköisyyksistä todennäköisyys-laskennan laskusääntöjen avulla.
- Jos siis jatkuvan satunnaismuuttujan tiheysfunktio tunnetaan, se *satunnaisilmiö*, *jonka tulosvaihtoehtoja ko. satunnaismuuttuja kuvaa*, *hallitaan täydellisesti*.

Tiheysfunktion määrittelyalue

- Olkoon f *jatkuvan satunnaismuuttujan* ξ *tiheysfunktio*.
- Olkoon

$$\begin{cases} f(x) > 0, & \text{jos } x \in [c,d] \\ f(x) = 0, & \text{jos } x \notin [c,d] \end{cases}$$

- Tällöin

$$\int_c^d f(x)dx = 1$$

Jatkuvat jakaumat ja yhden pisteen todennäköisyydet

- Olkoon *f jatkuvan satunnaismuuttujan* ξ *tiheysfunktio*.
- Tällöin *jokaisen yksittäisen pisteen todennäköisyys on nolla*, koska

$$
\begin{aligned}
\Pr(\xi = b) &= \lim_{a \to b} \Pr(a \le \xi \le b) \\
&= \lim_{a \to b} \int_a^b f(x)dx \\
&= 0
\end{aligned}
$$

Tapahtumien todennäköisyyksien vertailu

- Olkoon ξ *jatkuva satunnaismuuttuja* ja f vastaava *tiheysfunktio*.
- Olkoot $A = \{a \leq \xi \leq b\}$ ja $C = \{c \leq \xi \leq d\}$ kaksi tapahtumaa.
- Olkoon

$$\int_a^b f(x)dx > \int_c^d f(x)dx$$

- Tällöin tapahtuma A on *todennäköisempi kuin* tapahtuma C:

$$\Pr(A) > \Pr(C)$$

Tapahtumien todennäköisyyksien vertailu: Havainnollistus

- Kuva oikealla esittää **eksponenttijakaumaa** (ks. lukua **Jatkuvia jakaumia**) noudattavan *jatkuvan satunnaismuuttujan* ξ *tiheysfunktiota* $f(x)$.
- Olkoot

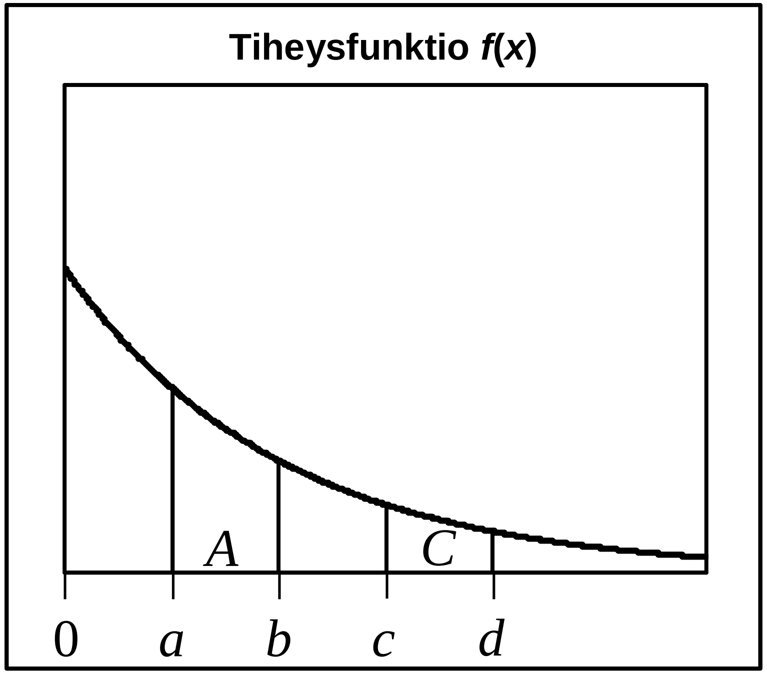
  $A = \{a \le \xi \le b\}$

  $C = \{c \le \xi \le d\}$
- Koska selvästi

  $$\int_a^b f(x)dx > \int_c^d f(x)dx$$

  niin tapahtuma A on *todennäköisempi kuin* tapahtuma C:

  $\Pr(A) > \Pr(C)$



Jatkuvan jakauman parametrointi 1/2

- Oletetaan, että satunnaismuuttuja ξ noudattaa jatkuvaa todennäköisyysjakaumaa, jonka *tiheysfunktio* on f .
- Tiheysfunktio f riippuu tavallisesti **parametreista** eli *vakioista*, jotka määräävät funktion f *muodon*.
- Tiheysfunktion muodon määräviä parametreja kutsutaan usein sen **todennäköisyysjakauman parametreiksi**, jonka todennäköisyydet ko. tiheysfunktio määrittelee.

Jatkuvan jakauman parametrointi 2/2

- Olkoot tiheysfunktion *f* muodon määräävät parametrit

$$\theta_1, \theta_2, \ldots , \theta_p$$

- Jos halutaan korostaa tiheysfunktion *f* riippuvuutta sen muodon määräävistä parametreista, tiheysfunktiota *merkitään*:

$$f(x\,;\,\theta_1, \theta_2, \ldots , \theta_p)$$

Jatkuvan jakauman parametrien merkitys

- Jatkuvan satunnaismuuttujan ξ *tiheysfunktion* f muodon määräävillä parametreilla $\theta_1, \theta_2, \ldots, \theta_p$ on usein jokin sisällöllinen *tulkinta* siinä satunnaisilmiössä, jonka malliksi f on konstruoitu.
- Parametrien $\theta_1, \theta_2, \ldots, \theta_p$ arvot ovat sovelluksissa tavallisesti *tuntemattomia* ja niiden arvot on **estimoitava** eli arvioitava havaintojen perusteella.

  Ks. monisteen **Tilastolliset menetelmät** lukua **Estimointi**.
- Parametreja $\theta_1, \theta_2, \ldots, \theta_p$ koskeva *hypoteeseja* eli oletuksia voidaan **testata** tilastollisin testein.

  Ks. monisteen **Tilastolliset menetelmät** lukua **Tilastolliset testit**.

Diskreetit jakaumat *vs* jatkuvat jakaumat 1/2

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

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

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

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

Diskreetit jakaumat *vs* jatkuvat jakaumat 2/2

- Olkoon ξ **jatkuva satunnaismuuttuja** ja f vastaava **tiheysfunktio**.
- Tiheysfunktion f arvo pisteessä x *ei ole todennäköisyys*.
- Reaaliakselin *välien* todennäköisyydet saadaan kaavasta

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

- Koska $f(x)$ *ei ole* todennäköisyys, *on mahdollista*, että

$$f(x) > 1$$

Esimerkki: Odotusajan jakauma

- Tarkastellaan jonkin tapahtuman *odotusaikaa* satunnaismuuttujana.
- Oletetaan, että tapahtumien *keskimääräinen* lukumäärä jotakin aikayksikköä kohden on λ.
- Tällöin seuraavan tapahtuman *odotusaika* ξ on *jatkuva satunnaismuuttuja*, joka voi saada *kaikki* ei-negatiiviset reaalilukuarvot.
- Tietyin ehdoin satunnaismuuttuja ξ noudattaa *todennäköisyysjakaumaa*, jonka *tiheysfunktio* on

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

- Satunnaismuuttujan ξ todennäköisyysjakaumaa sanotaan **eksponenttijakaumaksi**; ks. lukua **Jatkuvia jakaumia**.

Esimerkki: Odotusajan jakauma
Tehtävät (a) – (d)

- Tarkastellaan sovelluksena erään yrityksen puhelinkeskukseen tulevien puheluiden odotusaikoja.
- Oletetaan, että keskukseen tulee keskimäärin 4 puhelua minuutissa.
- Ratkaistaan seuraavat tehtävät:

(a) Määrää todennäköisyys, että seuraavaa puhelua joudutaan odottamaan korkeintaan 15 sekuntia?

(b) Määrää todennäköisyys, että seuraavaa puhelua joudutaan odottamaan korkeintaan 30 sekuntia?

(c) Kumpi on todennäköisempää:

Se, että seuraavaa puhelua joudutaan odottamaan 15 – 30 sekuntia, vai se, että seuraavaa puhelua joudutaan odottamaan 30 – 60 sekuntia?

(d) Mikä on todennäköisyys, että seuraavaa puhelua joudutaan odottamaan yli 60 sekuntia?

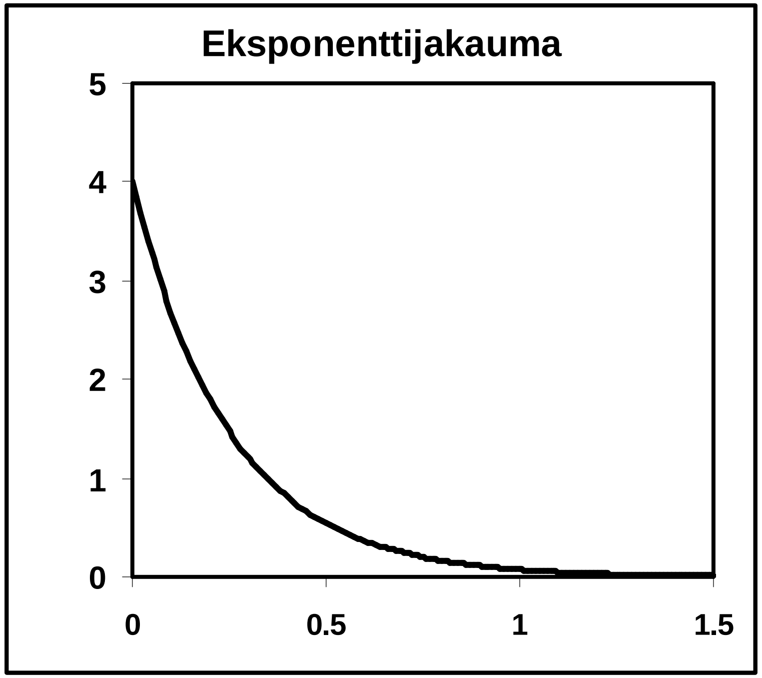
Esimerkki: Odotusajan jakauma
Tiheysfunktio

- Esimerkin tapauksessa $\lambda = 4$, kun aikayksikkönä on 1 minuutti.
- Siten satunnaismuuttujan

  $\xi =$ ”Seuraavan puhelun odotusaika”

  tiheysfunktio on muotoa

  $f(x) = 4e^{-4x}\,,\ x \geq 0$
- Viereinen kuvio esittää esimerkin eksponenttijakauman tiheysfunktion kuvaajaa.



Esimerkki: Odotusajan jakauma
Aputuloksia – 1/2

- Tarvitsemme tehtävien (a) – (d) ratkaisemisessa aputuloksia (1) – (3):

(1) Tapahtuman $\xi \le t$ todennäköisyys on

$$
\begin{aligned}
\Pr(\xi \le t) &= \int_0^t f(x)dx \\
&= 4\int_0^t e^{-4x}dx \\
&= 4\left[-\frac{1}{4}e^{-4x}\right]_0^t \\
&= 1 - e^{-4t}
\end{aligned}
$$

(2) Tapahtuman $\xi > t$ todennäköisyys on komplementtitapahtuman todennäköisyyden kaavan ja aputuloksen (1) nojalla

$$\Pr(\xi > t) = 1 - \Pr(\xi \le t) = e^{-4t}$$

Esimerkki: Odotusajan jakauma
Aputuloksia − 2/2

(3) Tapahtuman $s \le \xi \le t$ todennäköisyys on

$$
\begin{aligned}
\Pr(s \le \xi \le t) &= \int_s^t f(x)dx \\
&= 4\int_s^t e^{-4x}dx \\
&= 4\left[-\frac{1}{4}e^{-4x}\right]_s^t \\
&= e^{-4s} - e^{-4t}
\end{aligned}
$$

Esimerkki: Odotusajan jakauma
Tehtävien (a) ja (b) ratkaisut

- Tapahtuman $\xi \le 1/4$ min todennäköisyys (tehtävä (a)) on aputuloksen (1) nojalla

$$\Pr\left(\xi \le \frac{1}{4}\right) = 1 - e^{-4\times\frac{1}{4}} = 1 - e^{-1} \approx 0.632$$

- Tapahtuman $\xi \le 1/2$ min todennäköisyys (tehtävä (b)) on aputuloksen (1) nojalla

$$\Pr\left(\xi \le \frac{1}{2}\right) = 1 - e^{-4\times\frac{1}{2}} = 1 - e^{-2} \approx 0.865$$

Esimerkki: Odotusajan jakauma
Tehtävän (c) ratkaisu

- Tapahtumien 1/4 min ≤ ξ ≤ 1/2 min ja 1/2 min ≤ ξ ≤ 1 min todennäköisyydet (tehtävä (c)) ovat aputuloksen (3) nojalla

$$\Pr\left(\frac{1}{4} \le \xi \le \frac{1}{2}\right) = e^{-4\times\frac{1}{4}} - e^{-4\times\frac{1}{2}} = e^{-1} - e^{-2} \approx 0.233$$

$$\Pr\left(\frac{1}{2} \le \xi \le 1\right) = e^{-4\times\frac{1}{2}} - e^{-4\times1} = e^{-2} - e^{-4} \approx 0.117$$

- Siten tapahtuma 1/4 min ≤ ξ ≤ 1/2 min on todennäköisempi kuin tapahtuma 1/2 min ≤ ξ ≤ 1 min.

Esimerkki: Odotusajan jakauma
Tehtävän (d) ratkaisu

- Tapahtuman $\xi > 1$ min todennäköisyys (tehtävä (d)) on aputuloksen (2) nojalla

$$\Pr(\xi > 1) = 1 - \Pr(\xi \leq 1) = e^{-4} \approx 0.018$$

- *Todennäköisyyden frekvenssitulkinnan* mukaan tämä merkitsee seuraavaa:

  Keskus joutuu odottamaan seuraavaa puhelua *kauemmin* kuin 1:n minuutin *keskimäärin* melkein 2 kertaa ($100 \times 0.018 \approx 2$) 100:sta odotuskerrasta.