TODENNÄKÖISYYSJAKAUMA

Palautetaan mieleen kurssin alussa esiin noussut asia: *Tutkittavalle ilmiölle pyritään löytämään mahdollisimman tarkka matemaattinen malli, jonka avulla saatuja tuloksia sovelletaan käytäntöön ja tehdään ennusteita.*

Todennäköisyyslaskennassa tällaisina malleina käytetään *jakaumia*.

Satunnaismuuttuja ja sen jakauma

Esimerkki Tarkastellaan satunnaiskoetta ”*pakasta vedetään kortti*”. Tällöin funktio

$$f\colon\ f(\text{kortti}) = \text{"kortin numeroarvo”}$$

on satunnaismuuttuja. Esimerkiksi $f(\text{herttajätkä}) = 11$. Myös funktio

$$g\colon\ g(\text{kortti}) = \begin{cases} 0, & \text{jos kortti on punainen} \\ 1, & \text{jos kortti on musta} \end{cases}$$

on satunnaismuuttuja. Esimerkiksi $g(\text{ristiseiska}) = 1$.

Yleisesti satunnaiskokeen (-ilmiön) tulokset voivat olla *lukuja* tai *mittaustuloksia* tai *symboleita* (kuten kruuna/klaava). Jotta saataisiin matemaattinen malli eli jakauma, on satunnaiskokeen jokaiseen alkeistapaukseen liitettävä reaaliluku, kuten edellä käydyssä esimerkissä.

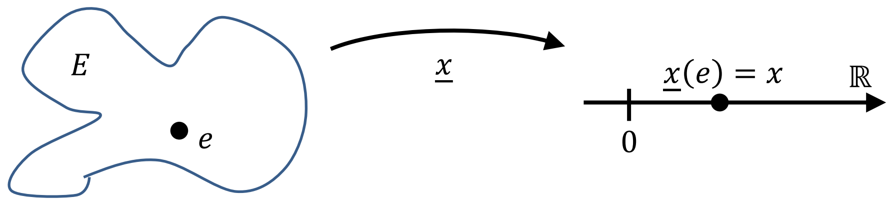
Alkeistapausten ”matematisointi” on tehtävä lisäksi niin, että eri alkeistapauksia vastaa eri reaaliluku. Näin on luotu perusjoukosta E reaaliluvuille $\mathbb{R}$ funktio, jota sanotaan harhaanjohtavasti (historialliset syyt) satunnaismuuttujaksi.

Määritelmä, *satunnaismuuttuja*:
Satunnaismuuttuja on otosavaruudessa E määritelty reaaliarvoinen *funktio*

$$\underline{x}\colon E \to \mathbb{R} \qquad (\text{tai}\ \ X\colon E \to \mathbb{R}).$$

Huomautus: $\underline{x}$ on funktio ja jos alkio $e \in E$, niin satunnaismuuttujan $\underline{x}$ arvo muuttujan arvolla e on reaaliluku, jota merkitään $\underline{x}(e) = x$ ja sanotaan alkion e *reaalisaatioksi*. Siis

$$\underline{x}\colon E \to \mathbb{R}\,; \quad \underline{x}(e) = x \in \mathbb{R}.$$



Määritelmä, *diskreetti satunnaismuuttuja*:
Satunnaismuuttuja, joka saa vain äärellisen määrän arvoja tai jonka arvojoukko on numeroituva ($N(\mathcal{A}_{\underline{x}}) = N\left(\underline{x}(E)\right) = \mathbb{N}$) on *diskreetti satunnaismuuttuja*, lyhyesti *diskreetti*.

Jos diskreetin satunnaismuuttujan $\underline{x}$ arvojoukko $\mathcal{A}_{\underline{x}} = \{x_1, x_2, \dots, x_n\}$, niin luvut

$$p_k = P\underbrace{\left(\underline{x} = x_k\right)}_{\substack{\text{pitäisi olla} \\ \underline{x}(e_k)=x_k}} = \text{todennäköisyys sille, että } \underline{x} \text{ saa arvon } x_k$$

määräävät tämän satunnaismuuttujan *jakauman*.

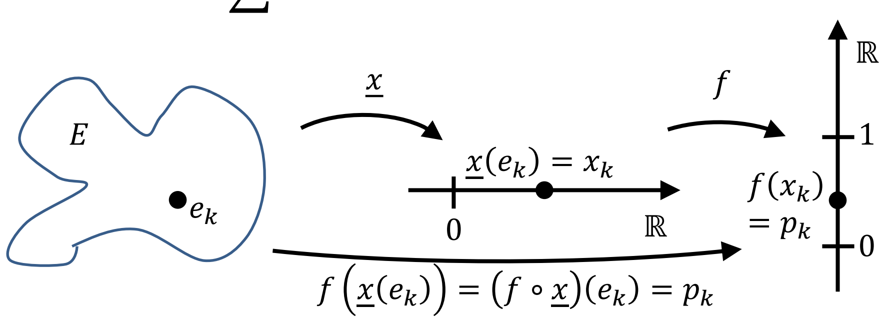
Huomautus: Määritelmässä oleva todennäköisyys on alkeistapauksen e_k pistetodennäköisyys p_k. Nyt voidaan muodostaa ns. *pistetodennäköisyysfunktio* $f\colon \mathbb{R} \to [0,1]$; $f(x_k) = p_k$, koska eri alkeistapauksia vastasi eri reaaliluku ja näin ollen funktio f on hyvin määritelty.

Huomaa, että jokainen $f(x_k) = p_k \in [0,1]$, tarkemmin

$$f(x) \geq 0 \quad \forall x \in \mathbb{R} \quad \text{ja} \quad f(x) = 0, \quad \text{kun } x \notin \{x_1, x_2, \dots, x_n\}$$

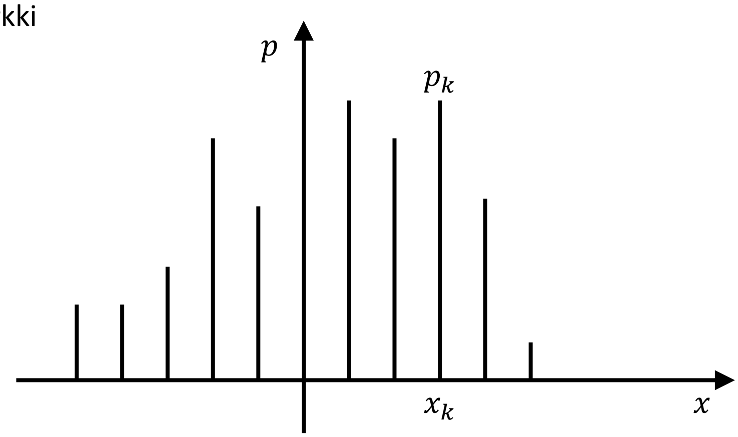
Lisäksi

$$\sum p_k = p_1 + p_2 + \dots + p_n = 1.$$



Määritelmä, *diskreetin satunnaismuuttujan $\underline{x}$ jakauma*:
Diskreetin satunnaismuuttujan $\underline{x}$ jakaumalla eli todennäköisyysjakaumalla tarkoitetaan pistetodennäköisyysfunktion f arvojoukkoa $\mathcal{A}_f$ tai lukuparien (x_k, p_k) joukkoa, missä $f(x_k) = p_k$.

Esimerkki



Yksittäiset todennäköisyydet ovat toki tärkeitä, mutta monesti halutaan tietää todennäköisyys, kun satunnaismuuttujan $\underline{x}$ arvot ovat tietyllä välillä eli esim. $P\left(1 < \underline{x} \leq 5\right)$ tai $P\left(\underline{x} \leq 3\right)$. Näitä todennäköisyyksiä varten tarvitaan kertymäfunktio

Määritelmä, *kertymäfunktio:*
Diskreetin satunnaismuuttujan *kertymäfunktio* on $F: \mathbb{R} \to [0,1]$, jonka arvot määritetään yhtälöllä

$$F(x) = P\left(\underline{x} \leq x\right).$$

Merkintä $\underline{x} \leq x$ tarkoittaa: Satunnaismuuttujan $\underline{x}$ arvo ei ylitä arvoa x.

Kertymäfunktion F arvo pisteessä x saadaan nimensä mukaisesti laskemalla yhteen kaikki ne pistetodennäköisyydet p_k, joita vastaavat satunnaismuuttujan arvot x_k eivät ylitä lukua x. Kuinka paljon *kertyy* todennäköisyyttä tiettyyn rajaan x asti.

Symbolisesti sama asia ilmaistuna. Diskreetin satunnaismuuttujan $\underline{x}$ kertymäfunktion F määrää yksikäsitteisesti $\underline{x}$:n pistetodennäköisyysfunktio f, sillä

$$F(x) = \sum_{k:\, x_k < x} f(x_k) = \sum_{k:\, x_k < x} p_k\,.$$

Huomautus:

$$0 \le F(x) \le 1 \qquad \forall x \in \mathbb{R}$$

$$x_1 \le x_2 \quad \Longrightarrow \quad F(x_1) \le F(x_2)$$

Eli kertymäfunktion arvot ovat positiivisia ja funktio on aina kasvava. Lisäksi tapahtuman, jolle $a < \underline{x} \le b$, todennäköisyys saadaan kertymäfunktion arvojen erotuksena, eli $P\big(a < \underline{x} \le b\big) = F(b) - F(a)$.

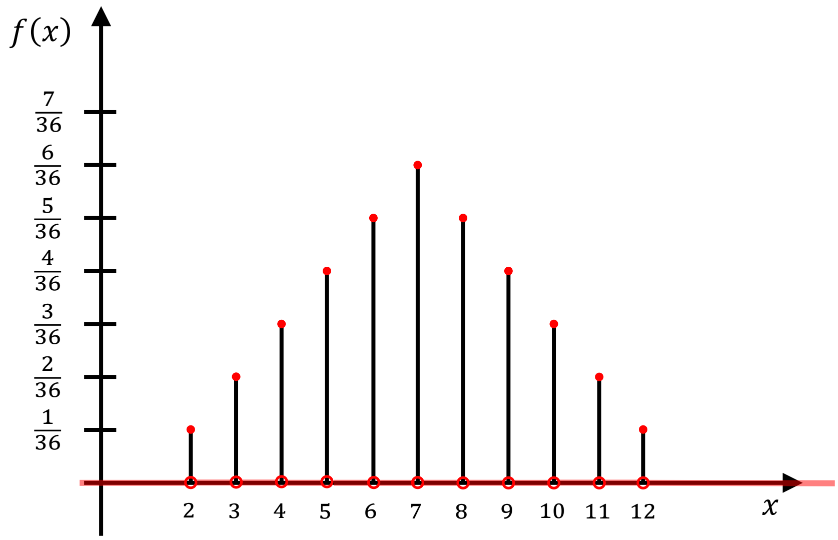
Esimerkki: Heitetään kahta noppaa. Olkoon satunnaismuuttujana $\underline{x} =$ silmälukujen summa.
Tällöin satunnaismuuttujan $\underline{x}$ arvojoukko on

$$\mathcal{A}_{\underline{x}} = \{2,3,4,5,6,7,8,9,10,11,12\}$$

ja pätee $f(x) > 0$, kun $x \in \mathcal{A}_{\underline{x}}$. Pistetodennäköisyydet ovat

$$\begin{aligned}
f(2) &= P(\{1,1\}) = \tfrac{1}{36}\\
f(3) &= P(\{1,2\},\{2,1\}) = \tfrac{2}{36}\\
f(4) &= P(\{1,3\},\{3,1\},\{2,2\}) = \tfrac{3}{36}\\
f(5) &= P(\{1,4\},\{4,1\},\{2,3\},\{3,2\}) = \tfrac{4}{36}\\
f(6) &= P(\{1,5\},\{5,1\},\{2,4\},\{4,2\},\{3,3\}\,) = \tfrac{5}{36}\\
f(7) &= P(\{1,6\},\{6,1\},\{2,5\},\{5,2\},\{3,4\},\{4,3\}) = \tfrac{6}{36}\\
f(8) &= P(\{2,6\},\{6,2\},\{3,5\},\{5,3\},\{4,4\}) = \tfrac{5}{36}\\
f(9) &= P(\{3,6\},\{6,3\},\{4,5\},\{5,4\}) = \tfrac{4}{36}\\
f(10) &= P(\{4,6\},\{6,4\},\{5,5\}) = \tfrac{3}{36}\\
f(11) &= P(\{5,6\},\{6,5\}) = \tfrac{2}{36}\\
f(12) &= P(\{6,6\}) = \tfrac{1}{36}\\
f(x) &= 0, \qquad \text{muualla}
\end{aligned}$$

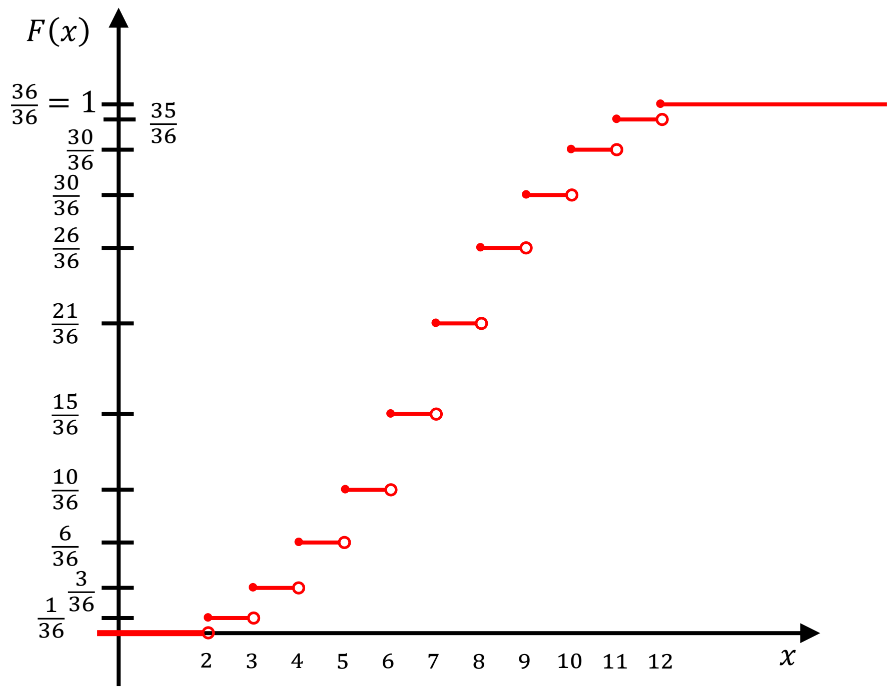
Ja näistä voidaan piirtää jakauma



Kertymäfunktiolle F saadaan pistetodennäköisyysfunktiota f hyödyntäen:

$$F(x) = \begin{cases} 0, & x < 2 \\ \frac{1}{36}, & 2 \leq x < 3 \\ \frac{3}{36}, & 3 \leq x < 4 \\ \frac{6}{36}, & 4 \leq x < 5 \\ \frac{10}{36}, & 5 \leq x < 6 \\ \frac{15}{36}, & 6 \leq x < 7 \\ \frac{21}{36}, & 7 \leq x < 8 \\ \frac{26}{36}, & 8 \leq x < 9 \\ \frac{30}{36}, & 9 \leq x < 10 \\ \frac{33}{36}, & 10 \leq x < 11 \\ \frac{35}{36}, & 11 \leq x < 12 \\ 1, & 12 \leq x \end{cases}$$

Ja kertymäfunktion kuvaaja



Nyt esimerkiksi

$$P(3 < \underline{x} \leq 8) = F(8) - F(3) = \frac{26}{36} - \frac{3}{36} = \frac{23}{36}$$

ja

$$P(\underline{x} \leq 5) = F(5) - \underbrace{\lim_{x\to-\infty} F(x)}_{=0 \text{ aina}} = \frac{10}{36} - 0 = \frac{10}{36}$$

Edelleen

$$P(\underline{x} > 5) = 1 - P(\underline{x} \leq 5) = 1 - \frac{10}{36} = \frac{26}{36}, \qquad \text{mutta}$$

$$P(\underline{x} \geq 5) = 1 - P(\underline{x} < 5) = \left(1 - P(\underline{x} \leq 4)\right) = 1 - \frac{6}{36} = \frac{5}{6}$$