Jatkuva jakauma

Diskreetti (discrete = erillinen) satunnaismuuttuja voi siis saada vain yksittäisiä, tiettyjä arvoja. Entäpä, jos satunnaismuuttuja voisi saada mitä tahansa arvoja? Millainen voisi olla ko. satunnaismuuttuja?

Esimerkiksi puhelun kestoaika, altaassa kasvatetun kirjolohen massa, joulupukin parran pituus, jne. Tällaisia satunnaismuuttujia sanotaan *jatkuviksi* ja niiden jakaumia *jatkuviksi jakaumiksi*.

Huomautus: On olemassa myös *sekajakaumia*, joissa on sekä diskreettejä että jatkuvia osia. Näihin törmää mm. kaksiulotteisissa jakaumissa, joissa toinen satunnaismuuttuja on jatkuva ja toinen diskreetti. Eli ei lukiomatematiikassa.

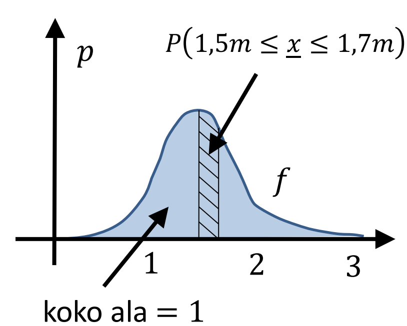
Tarkastellaan jatkuvaa satunnaismuuttujaa. Koska sen arvojoukko on reaalilukujen joukko $\mathbb{R}$ tai jokin sen väli, niin se voi saada minkä tahansa arvon, itse asiassa äärettömän monta arvoa eli

$$N\left(\mathcal{A}_{\underline{x}}\right) = \infty.$$

Nyt jakaumaa ei voida kuvailla pistetodennäköisyyksien avulla, sillä yksittäisen arvon pistetodennäköisyys tulkitaan nollaksi (muuten koko otosavaruuden todennäköisyys olisi ääretön eikä yksi).

Pitää ottaa käyttöön *tiheysfunktio*, joka ei suoraan kuvaa todennäköisyyttä vaan *todennäköisyystiheyttä*, lisäksi tarkastellaan tiettyjen arvojen väliin jäävää ”todennäköisyysmassaa”, $P\left(1{,}5m \le \underline{x} \le 1{,}7m\right)$.
Mitattaessa jatkuvan satunnaismuuttujan arvoja tulokset kuitenkin noudattavat mittaustarkkuudesta riippuvaa diskreettiä jakaumaa.

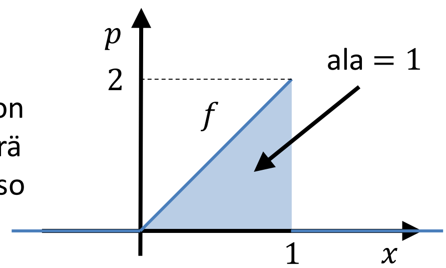
Tiheysfunktio ilmoittaa miten ”todennäköisyysmassa” jakaantuu x-akselin yläpuolelle. Mitä enemmän tätä massaa on tietyn välin yläpuolella, eli mitä suurempia tiheysfunktion f arvot $f(x)$ ovat, niin sitä todennäköisempää on, että satunnaismuuttujan arvot kuuluvat sille välille.



Määritelmä, *jatkuvan satunnaismuuttujan tiheysfunktio*:
Funktio $f: \mathbb{R} \to \mathbb{R}$ on jatkuvan satunnaismuuttujan $\underline{x}$ *tiheysfunktio*, jos
1) $f(x) \geq 0, \quad \forall x \in \mathbb{R}$,
2) f on jatkuva kaikkialla, paitsi mahdollisesti äärellisen monessa kohdassa,
3) käyrän $y = f(x)$ ja x-akselin väliin jäävän alueen pinta-ala on 1.

Esimerkki: Osoita, että funktio f; $f(x) = \begin{cases} 2x, & \text{kun } 0 \leq x \leq 1 \\ 0, & \text{muulloin} \end{cases}$ on jatkuvan satunnaismuuttujan $\underline{x}$ tiheysfunktio.

Käydään kolme kohtaa läpi:
1) Selvästi voimassa, OK
2) OK, ainoa epäjatkuvuuskohta on $x = 1$ ja 1kpl on äärellinen määrä
3) Pinta-ala on kolmion ala, (katso kuva): $A = \frac{1}{2} \cdot 1 \cdot 2 = 1$, OK



Esimerkki: Olkoon $a < b$. Millä c:n arvolla funktio

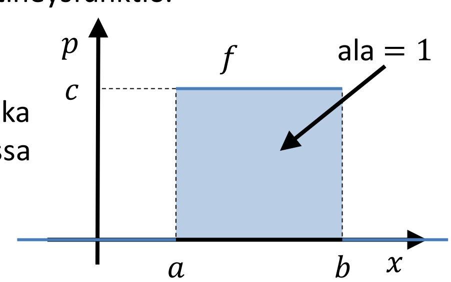
$$f;\ f(x) = \begin{cases} c, & \text{kun } a \leq x \leq b \\ 0, & \text{muulloin} \end{cases}$$

on jatkuvan satunnaismuuttujan $\underline{x}$ tiheysfunktio.

Käydään kolme kohtaa läpi:
1) Selvästi voimassa, OK
2) OK, kunhan $c \geq 0$, mutta koska ehto 3) pitää olla myös voimassa eli
3) Pinta-ala on suorakulmion ala, joten

$$c \cdot (b - a) = 1 \iff c = \frac{1}{b - a}$$

Eli $c = \frac{1}{b-a} > 0$, koska $a < b$.



HUOM! Kohdat a ja/tai b voisivat olla ihan hyvin myös negatiivisia x:n arvoja.

Tämä jatkuva satunnaismuuttuja noudattaa tasaista jakaumaa parametrein a ja b, merkitään $\underline{x} \sim Tas(a, b)$. Huomaa ero diskreetin satunnaismuuttujan tasaiseen jakaumaan $\underline{x} \sim Tas(x_1, x_1, \ldots, x_n)$.

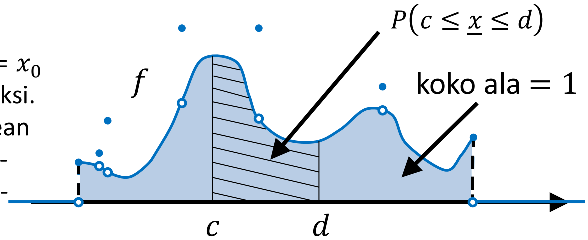
Miten ne todennäköisyydet sitten määritetään/lasketaan?

Lause, tiheysfunktion ja todennäköisyyden välinen yhteys:
Olkoon f satunnaismuuttujan $\underline{x}$ tiheysfunktio ja $c \leq d$. Tällöin todennäköisyys

$$P(c \leq \underline{x} \leq d)$$

on sen alueen pinta-ala, jota rajoittavat käyrä $y = f(x)$, x-akseli ja suorat $x = c$ ja $x = d$.

Muista: Yksittäisen kohdan $x = x_0$ todennäköisyys tulkitaan nollaksi. Pisteellä ei ole pinta-alaa. Oikean puolen kuvaesimerkissä tiheysfunktiolla f on 7 epäjatkuvuuskohtaa



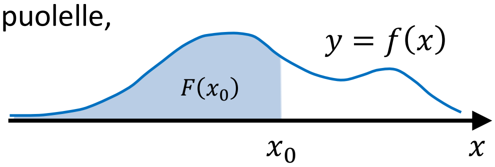
Huomautus: Välin päätepisteiden kuuluminen tai kuulumattominen ei vaikuta todennäköisyyteen

$$P(c \leq \underline{x} \leq d) = P(c < \underline{x} \leq d) = P(c \leq \underline{x} < d) = P(c < \underline{x} < d)$$

Lisäksi jos $c = d$, niin $P(c \leq \underline{x} \leq d) = P(\underline{x} = c = d) = 0$.

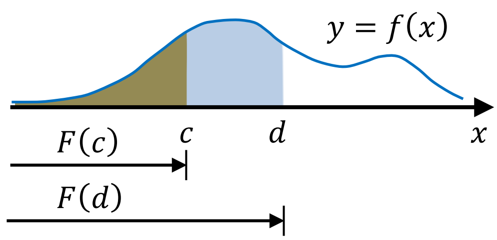
Jatkuvan satunnaismuuttujan $\underline{x}$ kertymäfunktion arvo $F(x)$ kertoo, kuinka paljon "todennäköisyysmassaa" on "kertynyt" kohdan $x = x_0$ negatiiviselle eli vasemmalle puolelle, siis

$$F(x_0) = P(\underline{x} \leq x_0).$$



Tapahtuman $c \leq x \leq d$ todennäköisyys on näin ollen

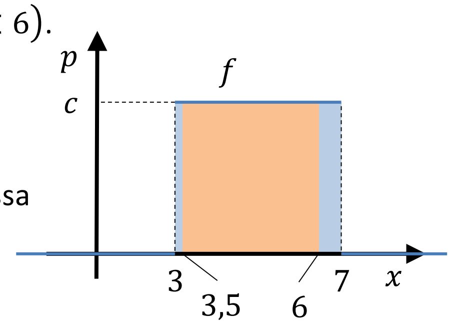
$$P(c \leq \underline{x} \leq d) = P(\underline{x} \leq d) - P(\underline{x} \leq c)$$
$$= F(d) - F(c)$$



Määritelmä, *jatkuvan satunnaismuuttujan kertymäfunktio*:
Jatkuvan satunnaismuuttujan $\underline{x}$ *kertymäfunktio* on $F(x) = P(\underline{x} \leq x)$, tällöin

$$P(c \leq \underline{x} \leq d) = P(c < \underline{x} < d) = F(d) - F(c).$$

Esimerkki: Laske edellisen esimerkin tapauksessa, kun $a = 3$ ja $b = 7$ todennäköisyys $P(3{,}5 \leq \underline{x} < 6)$.



Muodostetaan kertymäfunktio F:
Koska ala on yksi, niin $c = \frac{1}{7-3} = \frac{1}{4}$.
Näin ollen suorakulmion ala kohdassa $x = x_0$ on kanta·korkeus eli

$$\underbrace{c}_{\text{korkeus}} \cdot \underbrace{(x_0 - 3)}_{\text{kanta}} = \frac{1}{4}(x_0 - 3).$$

$$\Longrightarrow P(3{,}5 \leq \underline{x} < 6) = F(6) - F(3{,}5) = \frac{1}{4}(6 - 3) - \frac{1}{4}(3{,}5 - 3) = \frac{5}{8}.$$

Määritelmä, *jatkuvan satunnaismuuttujan odotusarvo, keskihajonta ja varianssi*:
Jatkuvan satunnaismuuttujan $\underline{x}$ *odotusarvoa, keskihajontaa ja varianssia* tarkastellaan ja käsitellään kurssilla 13. Lyhyesti (lähinnä tiedoksi tässä vaiheessa), jos jatkuvan satunnaismuuttujan $\underline{x}$ tiheysfunktio on f, niin

odotusarvo on

$$\mu = \mathbb{E}\underline{x} = \int_{-\infty}^{\infty} x f(x)dx$$

varianssi on

$$\sigma^2 = \mathbb{D}^2\underline{x} = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx$$

ja *keskihajonta* varianssin neliöjuurena

$$\sigma = \mathbb{D}\underline{x} = \sqrt{\mathbb{D}^2\underline{x}} = \sqrt{\int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx}$$