Todennäköisyyslaskenta

Osa 3: Todennäköisyysjakaumia

- **Lisää Diskreettejä jakaumia**
- Lisää Jatkuvia jakaumia
- Normaalijakaumasta johdettuja jakaumia

Hypergeometrinen jakauma ja sen pistetodennäköisyysfunktio 1/3

- Olkoon perusjoukon S alkioiden lukumäärä

  $$n(S) = N$$

- Tarkastellaan perusjoukon S *ositusta* joukkoihin A ja A^c .
- Oletetaan, että joukossa $A \subset S$ on

  $$n(A) = r$$

  alkiota.
- Tällöin joukon A komplementissa A^c on

  $$n(A^c) = N - r$$

  alkiota.

Hypergeometrinen jakauma ja sen pistetodennäköisyysfunktio 2/3

- Poimitaan perusjoukosta S *satunnaisesti* osajoukko B, jonka alkioiden lukumäärä on

  $$n(B) = n$$

  käyttämällä poiminnassa *otantaa ilman takaisinpanoa*.

- Määritellään *diskreetti satunnaismuuttuja* X:

  X = Osajoukkoon B tulleiden A:n alkioiden lukumäärä

Hypergeometrinen jakauma ja sen pistetodennäköisyysfunktio 3/3

- Sanomme, että satunnaismuuttuja X noudattaa **hypergeometrista jakaumaa parametrein N, r ja n**.
- Merkintä:

$$X \sim \text{HyperGeom}(N, r, n)$$

- Satunnaismuuttujan X **pistetodennäköisyysfunktio** on

$$f(x) = \Pr(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$$

$$\max[0, n-(N-r)] \le x \le \min(n, r)$$

Pistetodennäköisyysfunktion johto 1/4

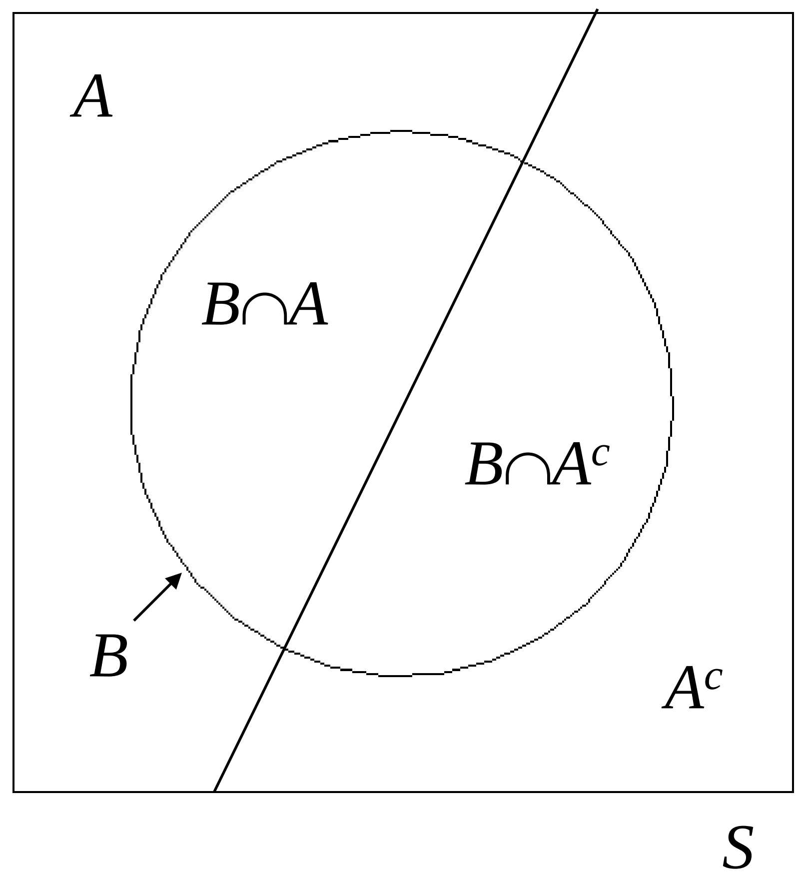
- Olkoon S otosavaruus ja

  $n(S) = N$
- Olkoon $A \subset S$
- Tällöin $\{A, A^c\}$ on otosavaruuden S *ositus*.
- Olkoon $n(A) = r$ ja $n(A^c) = N - r$
- Olkoon $B \subset S$ ja $n(B) = n$
- Otosavaruuden S ositus $\{A, A^c\}$ *indusoi* osituksen joukkoon B:

  $B = (B \cap A) \cup (B \cap A^c)$
- Olkoon

  $n(B \cap A) = x$

  $n(B \cap A^c) = n - x$



Hypergeometrinen jakauma

Pistetodennäköisyysfunktion johto 2/4

- N:n alkion joukosta S voidaan poimia n:n alkion osajoukko B

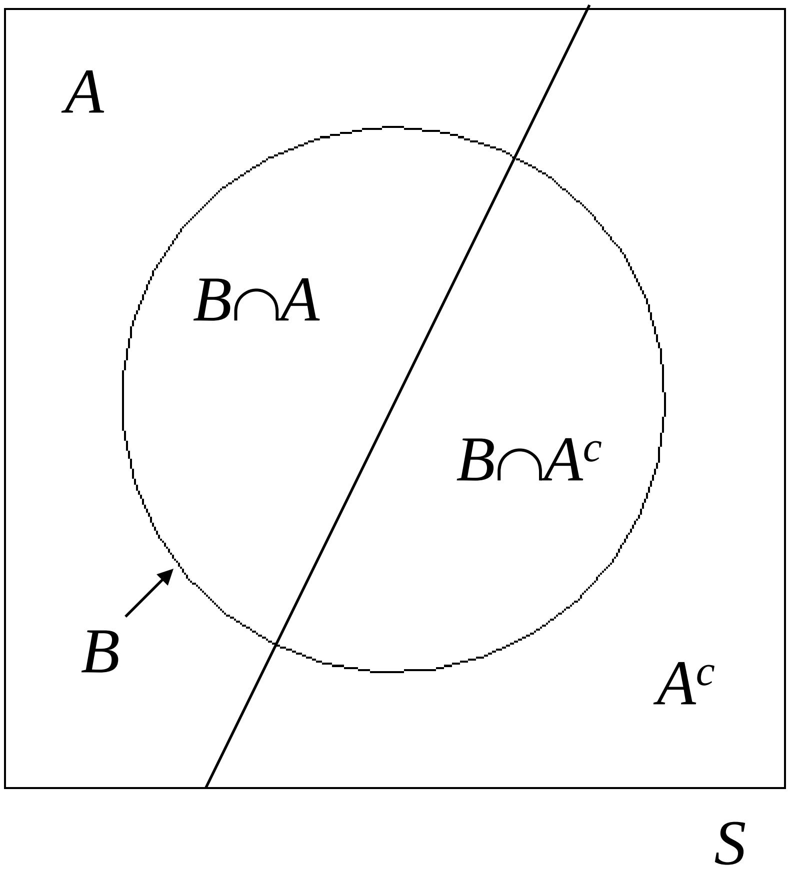
  $\binom{N}{n}$ eri tavalla.

- r:n alkion joukosta A voidaan poimia x alkiota

  $\binom{r}{x}$ eri tavalla.

- $(N-r)$:n alkion joukosta A^c voidaan poimia $n-x$ alkiota

  $\binom{N-r}{n-x}$ eri tavalla.

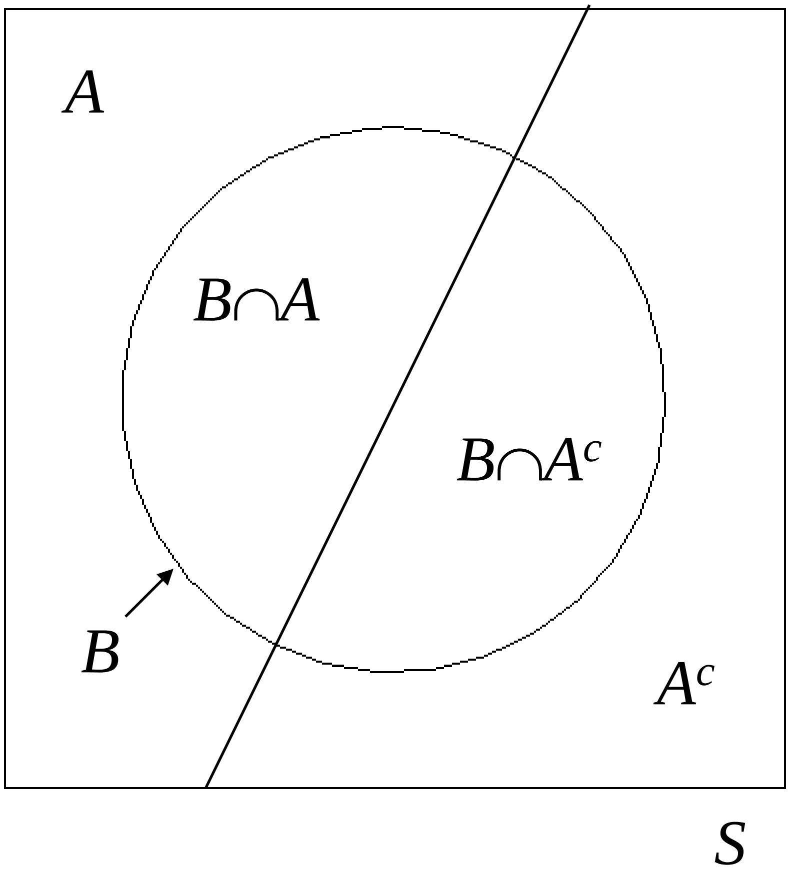


Pistetodennäköisyysfunktion johto 3/4

- r:n alkion joukosta A voidaan poimia x alkiota *riippumatta* siitä, mitkä $n - x$ alkiota poimitaan $(N - r)$:n alkion joukosta A^c.
- *Kertolaskuperiaatteen* nojalla n alkiota voidaan poimia joukosta S niin, että saadaan r alkiota joukosta A ja $(N - r)$ alkiota joukosta A^c

$$\binom{r}{x}\binom{N-r}{n-x}$$

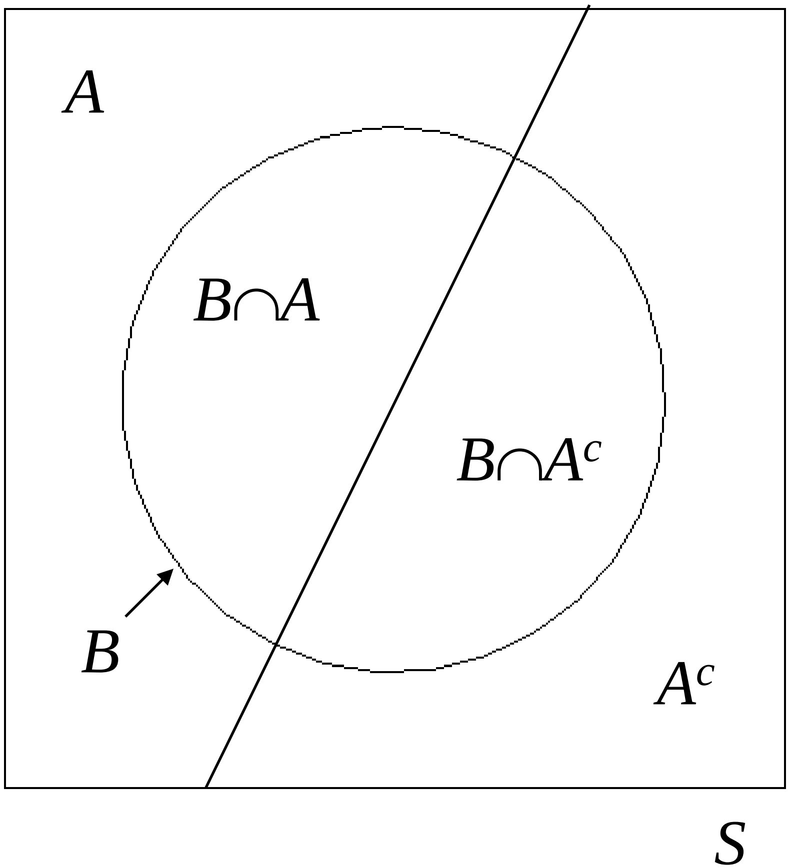
eri tavalla.



Hypergeometrinen jakauma
Pistetodennäköisyysfunktion johto 4/4

- Soveltamalla *klassisen todennäköisyyden* määritelmää saadaan:

$$\Pr(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$$



Hypergeometrinen jakauma
Odotusarvo, varianssi ja standardipoikkeama

- Olkoon

$$X \sim \mathrm{HyperGeom}(N, r, n)$$

- **Odotusarvo**:

$$\mathrm{E}(X) = n\frac{r}{N}$$

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

$$\mathrm{Var}(X) = \mathrm{D}^2(X) = n\frac{r}{N}\left(1-\frac{r}{N}\right)\left(\frac{N-n}{N-1}\right)$$

$$\mathrm{D}(X) = \sqrt{n\frac{r}{N}\left(1-\frac{r}{N}\right)\left(\frac{N-n}{N-1}\right)}$$

Hypergeometrinen jakauma
Pistetodennäköisyysfunktion kuvaaja

- Kuva oikealla esittää hypergeometrisen jakauman

  HyperGeom(100, 12, 20)

  pistetodennäköisyysfunktiota

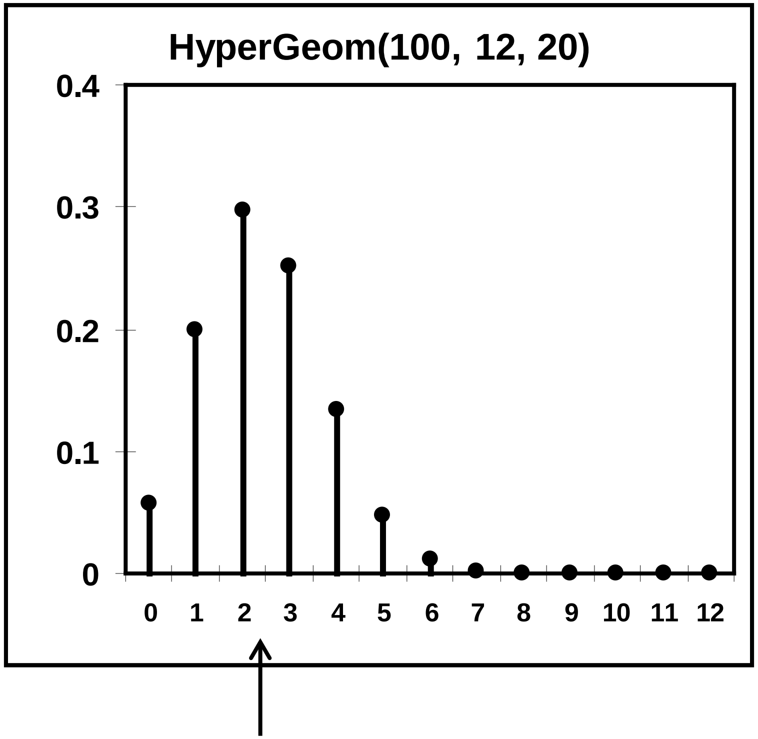
$$f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$$

$$N = 100\,,\; r = 12\,,\; n = 20$$

  pisteissä

  x = 0, 1, 2, … , 12
- Jakauman *odotusarvo*:

$$\mathrm{E}(X) = n\frac{r}{N} = 2.4$$



E(X) = 2.4

Hypergeometrinen jakauma

Esimerkki (Laininen)

- Leipuri leipoo aamulla 40 munkkia ja sekoittaa niiden joukkoon 10 edellisenä päivänä myymättä jäänyttä munkkia.
- Ensimmäinen asiakas ostaa viisi satunnaisesti valittua munkkia.
- Asiakkaan saamien edellisenä päivänä myymättä jääneiden munkkien lukumäärä on X. Kuinka X jakautuu?

- Satunnaismuuttuja X noudattaa hypergeometrista jakaumaa parametrein perusjoukon koko N=50, otoskoko n=5 ja edellispäiväisten munkkien määrä r =10.
- On odotettavissa, että asiakas saa $\mathrm{E}(X) = \left(5 \times 10\right) / 50 = 1$ kpl edellisenä päivänä myymättä jääneitä munkkeja.

Hypergeometrinen jakauma *vs* binomijakauma 1/3

- Hypergeometrisen jakauman todennäköisyydet ovat lähellä binomitodennäköisyyksiä, jos *otantasuhde*

$$\frac{n}{N} \approx 0$$

- Otantasuhde ≈ 0, jos otoskoko n on *pieni* perusjoukon kokoon N nähden.

Hypergeometrinen jakauma *vs* binomijakauma 2/3

- Olkoon

  $$X \sim \text{HyperGeom}(N, r, n)$$

- Merkitään $p = r/N$, jolloin $r = Np$.
- Siten

  $$X \sim \text{HyperGeom}(N, Np, n)$$

- Annetaan $N \to +\infty$.
- Tällöin hypergeometrisen jakauman HyperGeom(N, Np, n) pistetodennäköisyydet *lähestyvät* binomijakauman Bin(n, p) pistetodennäköisyyksiä:

  $$\lim_{N\to+\infty} f_{\text{HyperGeom}(N,Np,n)}(x) = f_{\text{Bin}(n,p)}(x) \ , \ x = 0,1,2,\ldots,n$$

Hypergeometrinen jakauma *vs* binomijakauma 3/3

- Hypergeometrisen jakauman ja binomijakauman yhteys tulee esille siinä, että jakaumilla on *sama* odotusarvo ja varianssit *eroavat vain multiplikatiivisella tekijällä*

$$\frac{N-n}{N-1}$$

jota sanotaan *äärellisen perusjoukon korjaustekijäksi*.

- Korjaustekijä vaikuttaa hypergeometrisen jakauman varianssiin sitä *vähemmän* mitä *pienempi* on otantasuhde n/N:

$$\frac{N-n}{N-1} \approx 1 \text{ , jos } \frac{n}{N} \approx 0$$

Hypergeometrinen jakauma

Otanta takaisinpanolla *vs* otanta ilman takaisinpanoa

- *Binomijakauma* muodostaa todennäköisyysmallin otannalle *takaisinpanolla* eli *palauttaen*.
- *Hypergeometrinen* jakauma muodostaa todennäköisyysmallin otannalle *ilman takaisinpanoa* eli *palauttamatta*.
- *Ero* otannan takaisinpanolla ja otannan ilman takaisinpanoa välillä *on merkityksetön*, jos *otantasuhde* n/N on *pieni* tai perusjoukko on *ääretön*.
- Käytännössä otanta tehdään lähes aina *ilman takaisinpanoa*, mutta *laskutoimituksissa* käytetään silti usein kaavoja, jotka perustuvat otantaan *takaisinpanolla*.
- Edellä esitetyn mukaan tästä johtuva *virhe on* kuitenkin yleensä *merkityksetön*.

Todennäköisyyslaskenta

Osa 3: Todennäköisyysjakaumia

- Lisää diskreettejä jakaumia
- **Lisää jatkuvia jakaumia**
- Normaalijakaumasta johdettuja jakaumia

Eksponenttijakauma ja sen tiheysfunktio

- Olkoon satunnaismuuttujan X **tiheysfunktio**

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

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

$$\int_0^{+\infty} f(x)dx = 1$$

- Sanomme, että satunnaismuuttuja X noudattaa **eksponenttijakaumaa parametrinaan λ.**
- Merkintä:

$$X \sim \mathrm{Exp}(\lambda)$$

Eksponenttijakauma ja sen kertymäfunktio

- Olkoon $X \sim \text{Exp}(\lambda)$.
- Satunnaismuuttujan X *tiheysfunktio*:

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

- Siten satunnaismuuttujan X **kertymäfunktioksi** saadaan, kun $x \geq 0$:

$$F(x) = \Pr(X \leq x) = \int_{-\infty}^{x} f(t)dt = \int_{0}^{x} \lambda e^{-\lambda t} dt$$

$$= \left[-e^{-\lambda t}\right]_0^x$$

$$= 1 - e^{-\lambda x}$$

Odotusarvo, varianssi ja standardipoikkeama

- Olkoon $X \sim \mathrm{Exp}(\lambda)$.
- **Odotusarvo**:

$$\mathrm{E}(X) = \frac{1}{\lambda}$$

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

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

$$\mathrm{D}(X) = \frac{1}{\lambda}$$

Odotusarvon johto

- Olkoon

$$X \sim \mathrm{Exp}(\lambda)$$

- Tällöin *osittaisintegroinnilla* saadaan:

$$\begin{aligned}\mathrm{E}(X) = \int_0^{+\infty} xf(x)dx &= \int_0^{+\infty} x\lambda e^{-\lambda x}dx \\ &= \left[-xe^{-\lambda x}\right]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x}dx \\ &= \left[-\frac{1}{\lambda}e^{-\lambda x}\right]_0^{+\infty} \\ &= \frac{1}{\lambda}\end{aligned}$$

Eksponenttijakauma

Tiheysfunktion kuvaaja

- Kuva oikealla esittää eksponenttijakauman

  $\mathrm{Exp}(\lambda)$

  tiheysfunktiota

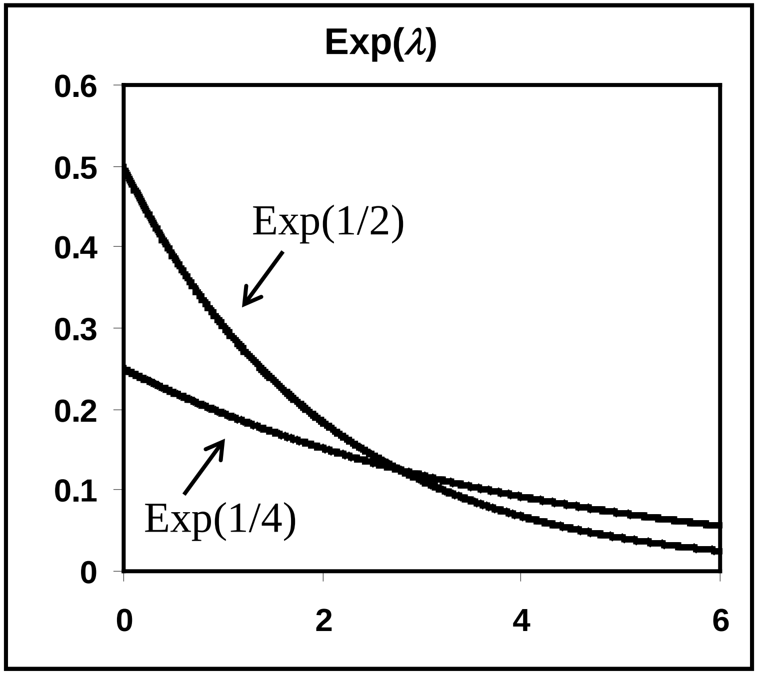
  $f(x) = \lambda e^{-\lambda x}$

  välillä [0, 6], kun

  (i) $\lambda = 1/2$

  (ii) $\lambda = 1/4$
- Jakauman *odotusarvo*:

  $\mathrm{E}(X) = 1/\lambda$



Eksponenttijakauma

Kertymäfunktion kuvaaja

- Kuva oikealla esittää eksponenttijakauman

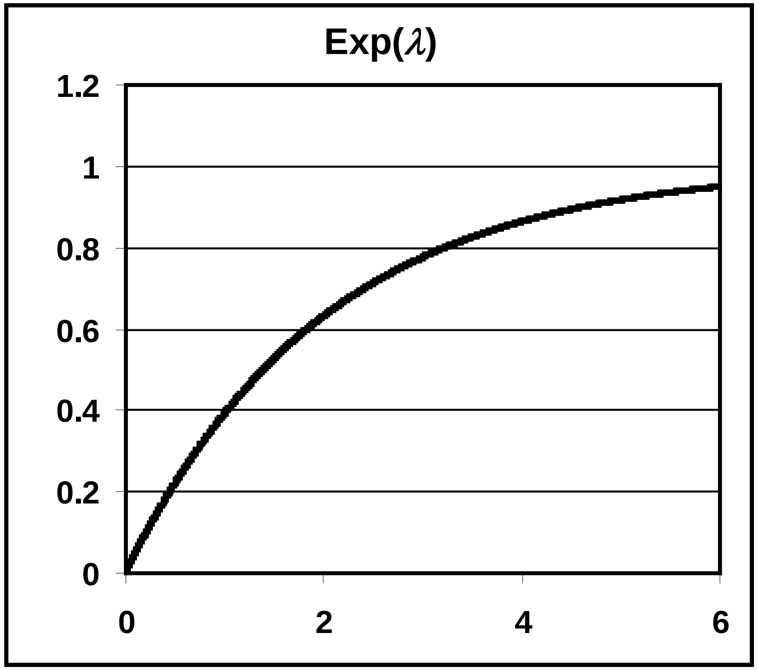
  Exp(λ)

  kertymäfunktiota

  $$F(x) = 1 - e^{-\lambda x}$$

  välillä [0, 6], kun

  $\lambda = 1/2$



Eksponenttijakauma

- Eksponenttijakauma on jonomalleissa tavallinen palvelupisteeseen saapuvien asiakkaiden väliajan X jakauma.

- Jakaumaa käytetään usein myös yksivaiheisen palvelun kestoajan jakautumismallina.

- Luotettavuustekniikassa eksponenttijakaumaa käytetään komponentin eliniän jakautumismallina.

Eksponenttijakauman unohtamisominaisuus

- Olkoon $X \sim \mathrm{Exp}(\lambda)$.
- Tällöin

$$\Pr(X \geq a + b \mid X \geq a) = \Pr(X \geq b)$$

- Siten eksponenttijakaumalla on seuraava *unohtamis-ominaisuus*:

  Se, että tapahtuman sattumista on jouduttu odottamaan ajan a, *ei vaikuta* todennäköisyyteen joutua odottamaan ajan b lisää.

Todennäköisyyksien määrääminen eksponenttijakaumasta

- Olkoon $X \sim \mathrm{Exp}(\lambda)$.
- Olkoon

  $$[c, d] \subset [0, +\infty)$$

  jokin välin $[0, +\infty)$ *osaväli*.
- *Välin* $[c, d]$ *todennäköisyys* saadaan integroimalla eksponenttijakauman $\mathrm{Exp}(\lambda)$ tiheysfunktio

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

  välillä $[c, d]$:

  $$\Pr(c \leq X \leq d) = \int_c^d f(x)dx = e^{-\lambda c} - e^{-\lambda d}$$

Todennäköisyyslaskenta

Osa 3: Todennäköisyysjakaumia

- Lisää diskreettejä jakaumia
- Lisää jatkuvia jakaumia
- **Normaalijakaumasta johdettuja jakaumia**

Normaalijakaumasta johdettuja jakaumia

>> Johdanto

χ^2-jakauma

***t*-jakauma**

Jakaumien määritteleminen normaalijakauman avulla

- Useat *tilastotieteen keskeiset todennäköisyysjakaumat* voidaan *määritellä* normaalijakauman avulla.
- Tällaisia ovat esimerkiksi χ^2- ja t-jakaumat, joilla on keskeinen rooli *otosjakaumien teoriassa*, *estimoinnissa* ja *testauksessa* (ks. monisteen **Tilastolliset menetelmät** lukuja **Otokset ja otosjakaumat, Estimointi** ja **Tilastollinen testaus**).
- Tarkastelemme seuraavien jakaumien määrittelemistä ja ominaisuuksia:
  - **χ^2-jakauma**
  - ***t*-jakauma**

Jakaumien määritteleminen normaalijakauman avulla: Kommentteja

- Tarkastelemme tässä ko. jakaumien määrittelemistä ja keskeisiä ominaisuuksia esittämättä jakaumien tiheysfunktioiden lausekkeita.
- Tämä johtuu siitä, että *emme tarvitse tiheysfunktioiden lausekkeita* niissä tilastotieteellisissä sovelluksissa, joissa käytämme ko. jakaumia.
- Tiheysfunktioiden lausekkeet johdetaan monisteen luvussa **Satunnaismuuttujien muunnokset ja niiden jakaumat.**

Normaalijakaumasta johdettuja jakaumia

Johdanto

>> χ^2-jakauma

***t*-jakauma**

χ^2-jakauman määritelmä 1/2

- Olkoot X_i , i = 1, 2, … , n *riippumattomia, standardoitua normaalijakaumaa* N(0,1) *noudattavia satunnaismuuttujia*.

- Tällöin

$$X_i \sim \mathrm{N}(0,1)\,,\; i = 1, 2, \ldots, n$$

$$X_1, X_2, \ldots, X_n \perp$$

χ^2-jakauman määritelmä 2/2

- Olkoon

$$X = \sum_{i=1}^{n} X_i^2$$

N(0,1)-jakautuneiden, *riippumattomien* satunnais-muuttujien X_i , $i = 1, 2, \ldots , n$ *neliösumma*.

- Tällöin satunnaismuuttuja X noudattaa **χ^2-jakaumaa** (Khiin neliö -jakaumaa) ***n*:llä vapausasteella**.
- Merkintä:

$$X \sim \chi^2(n)$$

Odotusarvo, varianssi ja standardipoikkeama

- Olkoon $X \sim \chi^2(n)$.
- **Odotusarvo**:

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

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

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

$$\mathrm{D}(X) = \sqrt{2n}$$

χ^2-jakauma
Tiheysfunktion kuvaaja

- Kuva oikealla esittää χ^2-jakauman

  $\chi^2(n)$

  tiheysfunktiota välillä [0, 10], kun vapausasteiden lukumäärällä n on seuraavat arvot:

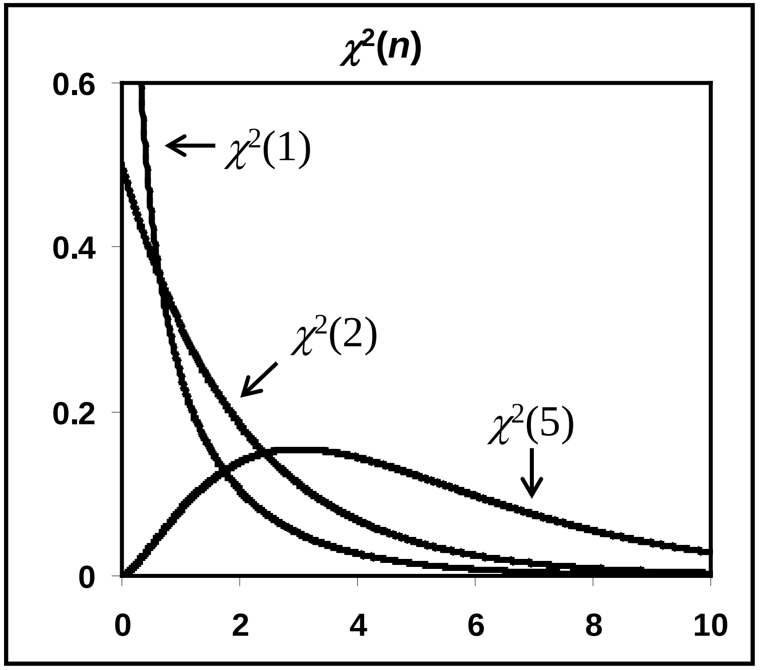
  (i) $n = 1$

  (ii) $n = 2$

  (iii) $n = 5$

- Jakauman *odotusarvo*:

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



χ^2**-jakauma**

Tiheysfunktion ja sen kuvaajan ominaisuuksia

- χ^2-jakauman tiheysfunktio $f(x)$ on *positiivinen* kaikille positiivisille argumentin arvoille:

  $$f(x) > 0 \;,\; x > 0$$

- Jos vapausasteiden lukumäärä

  $$n = 1, 2$$

  niin tiheysfunktio on *monotonisesti laskeva* kaikille $x \geq 0$.

- Jos vapausasteiden lukumäärä

  $$n \geq 3$$

  niin tiheysfunktio on *yksihuippuinen* ja sillä on *maksimi* jossakin pisteessä $x > 0$.

Todennäköisyyksien määrääminen χ^2-jakaumasta 1/2

- Todennäköisyydet voidaan määrätä χ^2*-jakaumasta* jakauman *kertymäfunktion* avulla.
- Olkoon $X \sim \chi^2(n)$.
- Olkoon satunnaismuuttujan X *kertymäfunktio*

$$F_{Chi}(x\,;\,n) = \Pr(X \leq x)$$

- Huomautus 1:

  Merkinnällä $F_{Chi}(x\,;\,n)$ on haluttu korostaa χ^2-jakauman riippuvuutta sen vapausasteiden lukumäärästä n.

- Huomautus 2:

  Koska χ^2*-jakauman tiheysfunktion integraalifunktiota ei osata esittää suljetussa muodossa*, jakauman kertymäfunktion arvojen määräämisessä on käytettävä jotakin *numeerista menetelmää*.

Todennäköisyyksien määrääminen χ^2-jakaumasta 2/2

- *Kaikkien* χ^2-jakaumaan liittyvien tapahtumien todennäköisyydet saadaan todennäköisyyksistä

$$\Pr(X \leq x) = F_{Chi}(x\,;\,n)$$

todennäköisyyslaskennan laskusääntöjen avulla.

- Esimerkiksi

$$\Pr(a \leq X \leq b) = F_{Chi}(b) - F_{Chi}(a)$$

Todennäköisyyksien määrääminen χ^2-jakaumasta: Taulukot 1/2

- χ^2-jakauman *taulukot* sisältävät tavallisesti *argumentin x arvoja* taulukoituna *useille vapausasteiden lukumäärille n*, mutta vain *muutamille kertymäfunktion* F_{Chi} *arvoille*.
- Siten taulukot mahdollistavat seuraavan tehtävän ratkaisemisen (taulukkokohtaisin rajoituksin):

  Määrää x, kun *todennäköisyys*

  $$\Pr(X \leq x) = F_{Chi}(x\,;\,n)$$

  on annettu.

Todennäköisyyksien määrääminen χ^2-jakaumasta: Taulukot 2/2

- Koska χ^2-jakaumaa käytetään tavallisesti *väliestimoinnin* tai *testauksen* yhteydessä, χ^2-jakauman taulukoihin on yleensä taulukoitu sellaisia argumentin x arvoja, jotka vastaavat todennäköisyyden

$$\Pr(X \le x) = F_{Chi}(x\,;n)$$

komplementtitodennäköisyyttä

$$p = \Pr(X > x) = 1 - F_{Chi}(x\,;n)$$

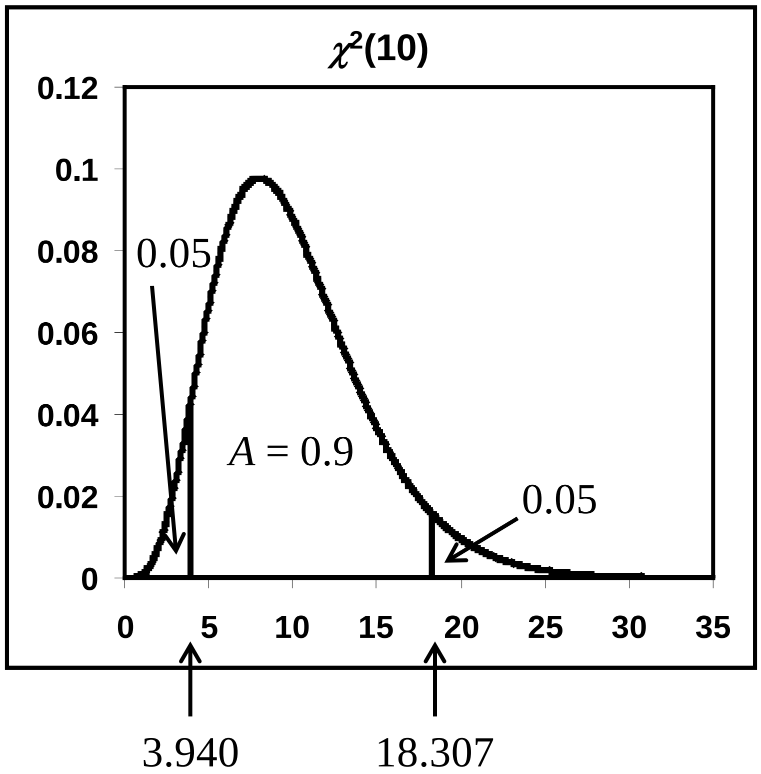
Todennäköisyyksien määrääminen χ^2-jakaumasta: Esimerkki

- Kuva oikealla esittää χ^2-jakauman

  $\chi^2(10)$

  tiheysfunktiota välillä [0, 35].
- χ^2-jakauman taulukoista saadaan:

$$\begin{aligned} &\text{Alueen } A \text{ pinta-ala} \\ &= \Pr(3.940 \le X \le 18.307) \\ &= F_{Chi}(18.307;10) \\ &\qquad - F_{Chi}(3.940;10) \\ &= 0.95 - 0.05 \\ &= 0.9 \end{aligned}$$



Normaalijakaumasta johdettuja jakaumia

Johdanto

χ^2-jakauma

>> *t*-jakauma

t-jakauman määritelmä 1/2

- Olkoot Y ja X_i , i = 1, 2, … , n *riippumattomia*, *standardoitua normaalijakaumaa* N(0,1) *noudattavia satunnaismuuttujia*.
- Tällöin

$$Y \sim \mathrm{N}(0,1)\,,\ X_i \sim \mathrm{N}(0,1)\,,\ i = 1,2,\ldots,n$$

$$Y, X_1, X_2, \ldots, X_n \perp$$

ja edelleen

$$X = \sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$$

$$Y \perp X$$

t-jakauman määritelmä 2/2

- Olkoon

$$t = \frac{Y}{\sqrt{\frac{1}{n}X}}$$

jossa

$$Y \sim \mathrm{N}(0,1)\,,\; X \sim \chi^2(n)\,,\; Y \perp X$$

- Tällöin satunnaismuuttuja t noudattaa **Studentin *t*-jakaumaa *n*:llä vapausasteella**.
- Merkintä:

$$t \sim t(n)$$

t-jakauma

Odotusarvo, varianssi ja standardipoikkeama

- Olkoon $t \sim t(n)$.
- **Odotusarvo**:

$$\mathrm{E}(t) = 0,\ n > 1$$

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

$$\mathrm{Var}(t) = \mathrm{D}^2(t) = \frac{n}{n-2},\ n > 2$$

$$\mathrm{D}(t) = \sqrt{\frac{n}{n-2}},\ n > 2$$

t-jakauma
Tiheysfunktion kuvaaja

- Kuva oikealla esittää t-jakauman

  $t(n)$

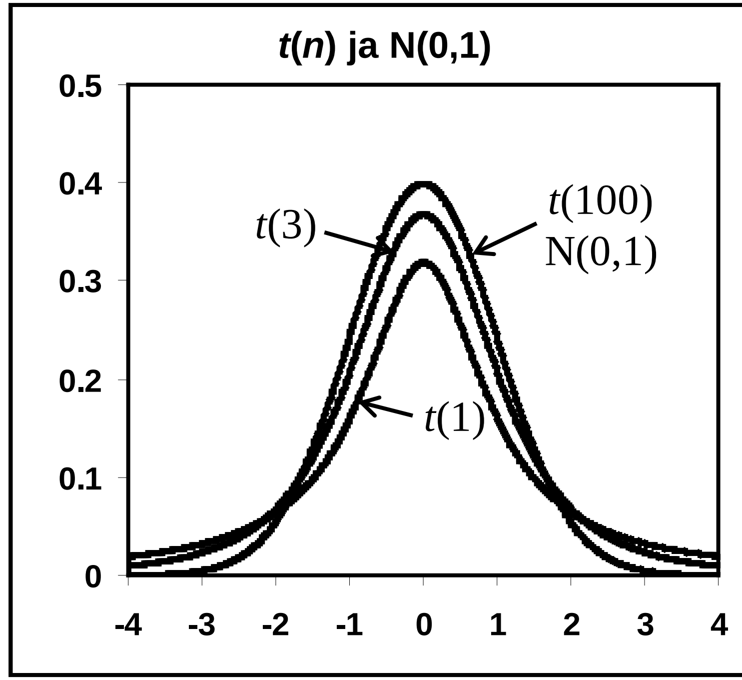
  tiheysfunktiota välillä [–4, +4], kun vapausasteiden lukumäärällä n on seuraavat arvot:

  (i) $n = 1$

  (ii) $n = 3$

  (iii) $n = 100$
- Jakauman *odotusarvo*:

  $$\mathrm{E}(t) = 0 \ , \ n > 1$$
- Kuvaan on piirretty myös standardoidun normaalijakauman N(0,1) tiheysfunktion kuvaaja.



t-jakauma

Tiheysfunktion ja sen kuvaajan ominaisuuksia 1/2

- *t*-jakauman tiheysfunktio $f(x)$ on kaikkialla *positiivinen*:

  $f(x) > 0$ kaikille x
- Tiheysfunktio on *yksihuippuinen*.
- Tiheysfunktio saa *maksimiarvonsa* pisteessä 0.
- Tiheysfunktio on *symmetrinen* pisteen $x = 0$ suhteen:

  $f(-x) = f(+x)$ kaikille x

_t_-jakauma

Tiheysfunktion ja sen kuvaajan ominaisuuksia 2/2

- *t*-jakauman tiheysfunktio muistuttaa standardoidun normaalijakauman N(0,1) tiheysfunktiota, mutta on sitä *paksuhäntäisempi*.
- *t*-jakauman tiheysfunktio muistuttaa standardoidun normaalijakauman N(0,1) tiheysfunktiota *sitä voimakkaammin mitä suurempi on vapaus-asteiden lukumäärä n* (ks. tarkemmin >).

t-jakauma ja normaalijakauma 1/2

- *t*-jakauma lähestyy standardoitua normaalijakaumaa, kun *vapausasteiden* lukumäärä *n* kasvaa.
- Olkoon $t \sim t(n)$.
- Tällöin

$$\lim_{n \to +\infty} \Pr(t \le z) = \Phi(z)$$

missä Φ on *standardoidun normaalijakauman* N(0,1) *kertymäfunktio*.

t-jakauma ja normaalijakauma 2/2

- Koska *t*-jakauma lähestyy vapausasteiden lukumäärän *n* kasvaessa standardoitua normaalijakaumaa N(0,1), voidaan *t*-jakaumaan liittyvät todennäköisyydet määrätä suurilla vapausasteiden luvuilla standardoidun normaalijakauman avulla.
- *Normaalijakauma-approksimaatio t*-jakaumalle on kohtuullinen jo, kun $n = 30$, ja riittävä useimpiin tarkoituksiin, kun $n > 100$.
- Esimerkki:

  Edellä esitetyssä kuvassa ei *t*(100)- ja N(0,1)-jakaumien tiheysfunktioiden kuvaajia pysty erottamaan toisistaan (ks. <).

Todennäköisyyksien määrääminen *t*-jakaumasta 1/2

- Todennäköisyyksien määrääminen *t*-jakaumasta voidaan tehdä jakauman *kertymäfunktion* avulla.
- Olkoon $t \sim t(n)$.
- Olkoon satunnaismuuttujan *t kertymäfunktio*

$$F_t(x\,;n) = \Pr(t \le x)$$

- Huomautus 1:

  Merkinnällä $F_t(x\,;n)$ on haluttu korostaa *t*-jakauman riippuvuutta sen vapausasteiden lukumäärästä n.

- Huomautus 2:

  Koska *t-jakauman tiheysfunktion integraalifunktiota ei osata esittää suljetussa muodossa*, jakauman kertymäfunktion arvojen määräämisessä on käytettävä jotakin *numeerista menetelmää*.

Todennäköisyyksien määrääminen *t*-jakaumasta 2/2

- *Kaikkien* tapahtumien todennäköisyydet saadaan todennäköisyyksistä

$$\Pr(t \leq x) = F_t(x\ ;\ n)$$

todennäköisyyslaskennan laskusääntöjen avulla.

- Esimerkiksi

$$\Pr(a \leq t \leq b) = F_t(b) - F_t(a)$$

Todennäköisyyksien määrääminen *t*-jakaumasta: Taulukot 1/3

- *t*-jakauman *taulukot* sisältävät tavallisesti *argumentin x arvoja* taulukoituna *useille vapausasteiden lukumäärille n*, mutta vain *muutamalle kertymäfunktion* F_t *arvolle*.
- Siten taulukot mahdollistavat seuraavan tehtävän ratkaisemisen (taulukkokohtaisin rajoituksin):

  Määrää x, kun todennäköisyys

  $$\Pr(t \leq x) = F_t(x\,;\,n)$$

  on annettu.

Todennäköisyyksien määrääminen *t*-jakaumasta: Taulukot 2/3

- Koska *t*-jakaumaa käytetään tavallisesti *väliestimoinnin* tai *testauksen* yhteydessä, *t*-jakauman taulukoihin on yleensä taulukoitu sellaisia argumentin *x* arvoja, jotka vastaavat todennäköisyyden

  $$\Pr(t \leq x) = F_t(x\ ;\ n)$$

  komplementtitodennäköisyyttä

  $$p = \Pr(t \geq x) = 1 - F_t(x\ ;\ n)$$

Todennäköisyyksien määrääminen *t*-jakaumasta: Taulukot 3/3

- Monissa *t*-jakauman taulukoissa on taulukoitu todennäköisyyksiä

  $$p = \Pr(t \geq x) = 1 - F_t(x\,;n)$$

  vain, kun $x \geq 0$.
- Tällöin todennäköisyydet $\Pr(t \leq -x)$ saadaan soveltamalla *t*-jakauman tiheysfunktion *symmetrisyyttä* pisteen $x = 0$ suhteen:

  $$\begin{aligned}\Pr\left(t \leq -x\right) &= 1 - \Pr(t \geq -x) \\ &= 1 - \Pr(t \leq x) \\ &= \Pr(t \geq x) \\ &= p\end{aligned}$$

Todennäköisyyksien määrääminen *t*-jakaumasta: Esimerkki

- Kuva oikealla esittää *t*-jakauman

  $t(10)$

  tiheysfunktiota välillä [–4, +4].
- *t*-jakauman taulukoista saadaan:

  Alueen A pinta-ala

  $$= \Pr(-1.812 \le t \le +1.812)$$
  $$= F_t(+1.812;10)$$
  $$-F_t(-1.812;10)$$
  $$= 0.95 - 0.05$$
  $$= 0.9$$