Todennäköisyyslaskennan kertausta

1. Joukko-oppi

Alkio, Erotus, Joukko, Komplementti, Leikkaus, Perusjoukko, Pistevieraus, Tyhjä joukko, Unioni, Yhdiste

2. Todennäköisyys ja sen määritteleminen

Alkeistapahtuma, Alkio, Empiirinen todennäköisyys, Frekvenssi, Frekvenssitulkinta, Joukko, Klassinen todennäköisyys, Koetoisto, Lukumääräfunktio, Mahdoton tapahtuma, Mitta, Otosavaruus, Perusjoukko, Sattuma, Satunnaisilmiö, Satunnaiskoe, Suhteellinen frekvenssi, Suhteellinen osuus, Suotuisa alkeistapahtuma, Symmetrisyys, Tapahtuma, Todennäköisyys, Tyhjä joukko, Varma tapahtuma

3. Todennäköisyyslaskennan peruslaskusäännöt

Alkeistapahtuma, Ehdollinen todennäköisyys, Ehtotapahtuma, Erotustapahtuma, Joukko, Komplementtitapahtuma, Leikkaus, Mahdoton tapahtuma, Otanta, Otanta palauttaen, Otanta palauttamatta, Otosavaruus, Pistevieraus, Riippumattomuus, Satunnaisotanta, Tapahtuma, Todennäköisyys, Toisensa poissulkevuus, Tulosääntö, Varma tapahtuma, Yhdiste, Yhdistetty tapahtuma, Yhteenlaskusääntö

4. Klassinen todennäköisyys ja kombinatoriikka

Binomikaava, Binomikerroin, Jono, Joukko, Kertoma, Klassinen todennäköisyys, Kombinaatio, Kombinatoriikka, Kertolaskuperiaate, Lukumääräfunktio, Multinomikerroin, Osajono, Osajoukko, Pascalin kolmio, Permutaatio, Riippumattomuus, Suotuisa alkeistapahtuma, Variaatio, Yhteenlaskuperiaate

5. Todennäköisyyden aksioomat

Boolen algebra, Joukko, Komplementtitapahtuma, Otosavaruus, Perusjoukko, Riippumattomuus, σ-algebra, Tapahtuma, Todennäköisyyden aksioomat, Todennäköisyyskenttä, Todennäköisyysmitta, Toisensa poissulkevuus, Yhdistetty tapahtuma, Yhteenlaskusääntö

6. Kokonaistodennäköisyys ja Bayesin kaava

Bayesin kaava, Ehdollinen todennäköisyys, Kokonaistodennäköisyyden kaava, Ositus

7. Verkot ja todennäköisyyslaskenta

Insidenssikuvaus, Piste, Puu, Puutodennäköisyys, Reitti, Rinnan kytkentä, Sarjaan kytkentä, Särmä, Toimintatodennäköisyys, Toimintaverkko, Tulosääntö, Verkko, Yhteenlaskusääntö

8. Satunnaismuuttujat ja todennäköisyysjakaumat

Diskreetti jakauma, Diskreetti satunnaismuuttuja, Jatkuva jakauma, Jatkuva satunnaismuuttuja, Pistetodennäköisyysfunktio, Satunnaismuuttuja, Tiheysfunktio, Todennäköisyysjakauma

9. Kertymäfunktio

Diskreetti jakauma, Diskreetti satunnaismuuttuja, Jatkuva jakauma, Jatkuva satunnaismuuttuja, Kertymäfunktio, Pistetodennäköisyysfunktio, Satunnaismuuttuja, Tiheysfunktio, Todennäköisyysjakauma

10. Jakaumien tunnusluvut

Diskreetti jakauma, Diskreetti satunnaismuuttuja, Huipukkuus, Jatkuva jakauma, Jatkuva satunnaismuuttuja, Keskusmomentti, Markovin epäyhtälö, Momentti, Odotusarvo, Painopiste, Pistetodennäköisyysfunktio, Satunnaismuuttuja, Standardipoikkeama, Tiheysfunktio, Todennäköisyysjakauma, Todennäköisyysmassa, Tshebyshevin epäyhtälö Tunnusluku, Varianssi, Vinous

11. Diskreettejä jakaumia

Bernoulli-jakauma, Bernoulli-koe, Binomijakauma, Diskreetti tasainen jakauma, Eksponentti-jakauma, Geometrinen jakauma, Hypergeometrinen jakauma, Kertymäfunktio, Negatiivinen binomijakauma, Odotusarvo, Otanta, Otanta palauttaen, Otanta palauttamatta, Otantasuhde, Pistetodennäköisyysfunktio, Poisson-jakauma, Standardipoikkeama, Varianssi

12. Jatkuvia jakaumia

Binomijakauma, Beta-jakauma, Cauchy-jakauma, Eksponenttijakauma, Gamma-jakauma, Jatkuva tasainen jakauma, Kertymäfunktio, Keskeinen raja-arvolause, Log-normaalijakauma, Normaaliapproksimaatio, Normaalijakauma, Odotusarvo, Poisson-jakauma, Standardipoikkeama, Standardointi, Taulukot, Tiheysfunktio, Varianssi, Weibull-jakauma

13. Normaalijakaumasta johdettuja jakaumia

χ^2-jakauma, F-jakauma, Normaalijakauma, Normaalijakaumasta johdetutut jakaumat, Odotusarvo, Standardipoikkeama, t-jakauma, Taulukot, Tiheysfunktio, Vapausasteet, Varianssi

14. Moniulotteiset satunnaismuuttujat ja todennäköisyysjakaumat

Diskreetti jakauma, Ehdollinen jakauma, Ehdollinen odotusarvo, Ehdollinen varianssi, Jatkuva jakauma, Karteesinen tulo, Kertymäfunktio, Korrelaatio, Korreloimattomuus, Korreloituneisuus, Kovarianssi, Odotusarvo, Pistetodennäköisyysfunktio, Regressiofunktio, Reunajakauma, Riippumattomuus, Riippuvuus, Tiheysfunktio, Varianssi, Yhteisjakauma

15. Moniulotteisia jakaumia

Binomijakauma, Diskreetti jakauma, Ehdollinen jakauma, Ehdollinen odotusarvo, Ehdollinen varianssi, Jatkuva jakauma, Kaksiulotteinen normaalijakauma, Korrelaatio, Korreloimattomuus, Korreloituneisuus, Kovarianssi, Kulmakerroin, Multinomijakauma, Odotusarvo, Painopiste, Piste-todennäköisyysfunktio, Regressiofunktio, Regressiosuora, Reunajakauma, Riippumattomuus, Riippuvuus, Suora, Tiheysfunktio, Todennäköisyysmassa, Varianssi, Yhteisjakauma, Yhteis-korrelaatiokerroin

16. Momenttiemäfunktio ja karakteristinen funktio

Binomijakauma, Diskreetti tasainen jakauma, Eksponenttijakauma, Geometrinen jakauma, Jatkuva tasainen jakauma, Karakteristinen funktio, Kolmiojakauma, Momentit, Momenttiemäfunktio, Negatiivinen binomijakauma, Normaalijakauma, Odotusarvo, Poisson-jakauma, Summan jakauma, Taylorin sarja, Varianssi

17. Satunnaismuuttujien muunnokset ja niiden jakaumat

Jacobin determinantti, Maksimi, Minimi, Monotoninen muunnos, Muunnos, Osamäärän jakauma, Summan jakauma, Yhteisjakauma

18. Konvergenssikäsitteet ja raja-arvolauseet

Approksimatiivinen jakauma, Aritmeettinen keskiarvo, Jakaumakonvergenssi, Kertymäfunktio, Keskeinen raja-arvolause, Konvergenssikäsitteet, Kvadraattinen konvergenssi, Melkein varma konvergenssi, Normaalijakauma, Odotusarvo, Standardipoikkeama, Standardointi, Stokastinen konvergenssi, Summan jakauma, Suurten lukujen laki, Varianssi

1. Joukko-oppi

Joukko ja sen alkiot

Äärellinen joukko voidaan aina määritellä luettelemalla sen **alkiot**. Yleisemmin **joukko** määritellään antamalla *ehto*, jonka joukon alkioiden on toteutettava. On syytä huomata, että *joukkoja on aina syytä tarkastella jonkin hyvin määritellyn* **perusjoukon** *osajoukkoina*.

Jos *perusjoukon S alkio x on joukon A alkio* eli *x kuuluu joukkoon A*, niin merkitsemme

$$x \in A$$

Vastaavasti, jos *perusjoukon S alkio x ei ole joukon A alkio* eli *x ei kuulu joukkoon A*, niin merkitsemme

$$x \notin A$$

Jos A on niiden perusjoukon S alkioiden x joukko, jotka *toteuttavat ehdon* $P(x)$ eli joille *lause* $P(x)$ *on tosi*, niin merkitsemme

$$A = \{x \in S \mid P(x)\}$$

Osajoukko

Jos jokainen joukon A alkio on myös joukon B alkio, niin joukko A on joukon B **osajoukko** ja merkitsemme

$$A \subset B \text{ tai } B \supset A$$

Siten $A \subset B$, jos ja vain jos

$$x \in A \Rightarrow x \in B$$

Tyhjä joukko

Joukko on **tyhjä**, jos siinä ei ole yhtään alkiota. Merkitsemme tyhjää joukkoa symbolilla

$$\varnothing$$

Tyhjä joukko on kaikkien joukkojen osajoukko. Siis, jos A on perusjoukon S mielivaltainen osajoukko, niin

$$\varnothing \subset A$$

Joukko-opin perusoperaatiot: yhdiste

Olkoot joukot A ja B perusjoukon S osajoukkoja. Joukkojen A ja B **unioni** eli **yhdiste** $A \cup B$ on niiden perusjoukon S alkioiden joukko, jotka kuuluvat joukkoon A tai joukkoon B (tai molempiin):

$$A \cup B = \{x \in S \mid x \in A \text{ } tai \text{ } x \in B\}$$

Joukko-opin perusoperaatiot: leikkaus

Olkoot joukot A ja B perusjoukon S osajoukkoja. Joukkojen A ja B **leikkaus** $A \cap B$ on niiden perusjoukon S alkioiden joukko, jotka kuuluvat joukkoon A ja joukkoon B:

$$A \cap B = \{x \in S \mid x \in A \text{ } ja \text{ } x \in B\}$$

Jos

$$A \cap B = \varnothing$$

niin sanomme, että joukot A ja B ovat **pistevieraita**.

Joukko-opin perusoperaatiot: komplementti

Olkoon joukko A perusjoukon S osajoukko. Joukon A **komplementti** A^c on niiden perusjoukon S alkioiden joukko, jotka eivät kuulu joukkoon A:

$$A^c = \{x \in S \mid x \notin A\}$$

Joukko-opin perusoperaatiot: erotus

Olkoot joukot A ja B perusjoukon S osajoukkoja. Joukkojen A ja B **erotus** $A \backslash B$ on niiden perusjoukon S alkioiden joukko, jotka kuuluvat joukkoon A, mutta eivät kuulu joukkoon B:

$$A \backslash B = \{x \mid x \in A \; ja \; x \notin B\}$$

Selvästi

$$A \backslash B = A \cap B^c$$

2. Todennäköisyys ja sen määritteleminen

Satunnaisilmiö

Reaalimaailman ilmiö on **stokastinen ilmiö** eli **satunnaisilmiö**, jos sillä on seuraavat ominaisuudet:

(i) Ilmiö voi päätyä *alkutilastaan* useisiin erilaisiin *lopputiloihin* eli ilmiöllä on useita erilaisia *vaihtoehtoisia tuloksia*.

(ii) Ilmiön alkutilan perusteella *ei voida* tarkasti *ennustaa* ilmiön lopputilaa eli sitä, mikä mahdollisista *tulosvaihtoehdoista realisoituu* eli *toteutuu*.

(iii) Vaikka ilmiön lopputilaa ei voida ennustaa tarkasti, tulosvaihtoehtojen *suhteellisten frekvenssien* eli *osuuksien nähdään* ilmiön toistuessa *käyttäytyvän säännönmukaisesti*.

Kutsumme *satunnaisilmiötä* usein **satunnaiskokeeksi** ja *satunnaisilmiön esiintymiskertaa* **koetoistoksi**.

Todennäköisyyslaskennan peruskäsitteet

Todennäköisyyslaskennan *peruskäsitteet* ovat **otosavaruus**, **tapahtuma** ja **alkeistapahtuma**:

(i) Satunnaisilmiön (satunnaiskokeen) kaikkien mahdollisten tulosvaihtoehtojen muodostamaa joukkoa sanotaan *otosavaruudeksi*.

(ii) *Tapahtuma* on jokin otosavaruuden tulosvaihtoehtojen muodostama joukko.

(iii) *Alkeistapahtuma* on satunnaisilmiön (satunnaiskokeen) tulosvaihtoehto, jota alkeellisempiin tulosvaihtoehtoihin satunnaisilmiötä ei voida purkaa.

Kun sanomme, että jokin *tapahtuma sattuu*, tarkoitamme, että *jokin tapahtumaan liittyvistä alkeistapahtumista sattuu*.

Todennäköisyyslaskennan ja joukko-opin peruskäsitteet vastaavat seuraavalla tavalla toisiaan:

Otosavaruus	↔	**Perusjoukko**
Tapahtuma	↔	**Joukko**
Mahdoton tapahtuma	↔	**Tyhjä joukko**
Alkeistapahtuma	↔	**Alkio**

Todennäköisyys ja sen perusominaisuudet

Olkoon A jokin otosavaruuden S *tapahtuma* eli olkoon

$$A \subset S$$

Todennäköisyys $\Pr(\cdot)$ on joukkofunktio, joka liittää tapahtumaan A reaaliluvun:

$$\Pr(A) \in \mathbb{R}$$

Todennäköisyyden *perusominaisuudet*:

(i) Olkoon tapahtuma A jokin otosavaruuden S *tapahtuma*. Tällöin

$$0 \leq \Pr(A) \leq 1$$

(ii) Tyhjä joukko $\varnothing$ ja **mahdoton tapahtuma** samastetaan ja

$$\Pr(\varnothing) = 0$$

(iii) Otosavaruus S ja **varma tapahtuma** samastetaan ja

$$\Pr(S) = 1$$

Empiirinen todennäköisyys

Tarkastellaan *satunnaiskoetta*, jota voidaan *toistaa* siten, että seuraavat ehdot pätevät:

(i) Kokeen olosuhteet *säilyvät muuttumattomina* koetoistosta toiseen.

(ii) Koetoistot *ovat riippumattomia* siinä mielessä, että yhdenkään koetoiston tulos ei riipu siitä mitä tuloksia muista koetoistoista saadaan.

Tarkkaillaan *tapahtuman A* esiintymistä koetoistojen aikana. Jos tapahtuman *A suhteellinen frekvenssi* eli *osuus* lähestyy jotakin kiinteätä lukua koetoistojen lukumäärän kasvaessa rajatta, tätä lukua kutsutaan ko. *tapahtuman* **empiiriseksi todennäköisyydeksi**.

Oletetaan, että satunnaiskoetta *toistetaan n* kertaa. Olkoon f_A **tapahtuman *A* frekvenssi** eli *lukumäärä* koetoistojen joukossa. Tällöin

$$\frac{f_A}{n}$$

on **tapahtuman *A* suhteellinen frekvenssi** eli *suhteellinen osuus* koetoistojen joukossa.

Annetaan koetoistojen lukumäärän n kasvaa rajatta. Oletetaan, että (jossakin mielessä)

$$\frac{f_A}{n} \to p_A$$

kun $n \to \infty$. Tällöin luku p_A on ko. *tapahtuman empiirinen todennäköisyys*.

Todennäköisyyden frekvenssitulkinta

Oletetaan, että *toistamme* jotakin satunnaiskoetta ja tarkkailemme *tapahtuman A suhteellista frekvenssiä* koetoistojen aikana. *Todennäköisyyden frekvenssitulkinnan* mukaan *tapahtuman A suhteellinen frekvenssi vaihtelee satunnaisesti koetoistosta toiseen*, *mutta saa keskimäärin tapahtuman todennäköisyyttä lähellä olevia arvoja*.

Olkoon tapahtuman A todennäköisyys

$$\Pr(A) = p$$

Oletetaan, että sitä satunnaiskoetta, jonka tulosvaihtoehtona tapahtuma A on, toistetaan n kertaa. Tällöin todennäköisyyden frekvenssitulkinnasta seuraa, että on odotettavissa, että *tapahtuman A frekvenssi f* on lähellä lukua

$$np$$

Lukumääräfunktio

Olkoon

$$n(A)$$

funktio, joka kertoo *joukon A alkioiden lukumäärän*. Jos siis

$$A = \{a_1, a_2, \ldots, a_k\}$$

on äärellinen joukko, jonka alkioiden lukumäärä on k, niin

$$n(A) = k$$

Kutsumme funktiota $n(\cdot)$ **lukumääräfunktioksi**.

Klassinen todennäköisyys

Oletetaan, että *äärellisen otosavaruuden*

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

alkeistapahtumat

$$s_i , i = 1, 2, \ldots , n$$

ovat *yhtä todennäköisiä* eli

$$\Pr(s_i) = \frac{1}{n} , i = 1, 2, \ldots , n$$

Tällöin sanomme, että alkeistapahtumat s_i , $i = 1, 2, \ldots , n$ ovat **symmetrisiä**.

Olkoon tapahtuma A otosavaruuden S osajoukko. Tällöin **tapahtuman A klassinen todennäköisyys** $\Pr(A)$ saadaan määräämällä *tapahtumalle A suotuisien alkeistapahtumien suhteellinen osuus* kaikista alkeistapahtumista eli

$$\Pr(A) = n(A)/n(S)$$

jossa

$n(A)$ = tapahtumalle A *suotuisien* alkeistapahtumien lukumäärä
= joukkoon A kuuluvien alkeistapahtumien lukumäärä

ja

$n(S)$ = kaikkien mahdollisten alkeistapahtumien lukumäärä
= otosavaruuteen S kuuluvien alkeistapahtumien lukumäärä

Todennäköisyys mittana

Todennäköisyys on **mitta**, joka mittaa satunnaisilmiön *tapahtumien sattumisen mahdollisuutta*.

3. Todennäköisyyslaskennan peruslaskusäännöt

Todennäköisyyslaskennan perusoperaatiot

Todennäköisyyslaskennan **perusoperaatioilla** tarkoitetaan operaatioita, joilla tapahtumista johdetaan uusia tapahtumia.

Todennäköisyyslaskennan **peruslaskusäännöillä** tarkoitetaan laskusääntöjä, joilla alkeellisemmista tapahtumista joukko-opin operaatioilla johdettujen uusien, *yhdistettyjen tapahtumien* todennäköisyydet määrätään alkeellisempien tapahtumien todennäköisyyksien avulla.

Todennäköisyyslaskennan ja joukko-opin *perusoperaatiot* vastaavat seuraavalla tavalla toisiaan:

Tapahtuma A *ei* satu eli tapahtuman A *komplementtitapahtuma* sattuu

$\leftrightarrow$ Joukon A *komplementti* A^c

Tapahtuma A sattuu *tai* tapahtuma B sattuu *tai* molemmat sattuvat

$\leftrightarrow$ Joukkojen A ja B *unioni* eli *yhdiste* $A \cup B$

Tapahtuma A sattuu *ja* tapahtuma B sattuu

$\leftrightarrow$ Joukkojen A ja B *leikkaus* $A \cap B$

Tapahtuma A sattuu *ja* tapahtuma B *ei* satu

$\leftrightarrow$ Joukkojen A ja B *erotus* $A \backslash B$

Komplementtitapahtuman todennäköisyys

Olkoon tapahtuma A otosavaruuden S osajoukko. Joukon A *komplementtitapahtuman*

$$A^c = \{x \in S \mid x \notin A\}$$

todennäköisyys on

$$\Pr(A^c) = 1 - \Pr(A)$$

Yhdisteen todennäköisyys

Olkoot tapahtumat A ja B otosavaruuden S osajoukkoja. Tällöin

$$\Pr(A \cup B)$$

on todennäköisyys sille, että *tapahtuma A sattuu tai tapahtuma B sattuu tai molemmat sattuvat*.

Leikkauksen todennäköisyys

Olkoot tapahtumat A ja B otosavaruuden S osajoukkoja. Tällöin

$$\Pr(A \cap B)$$

on todennäköisyys sille, että *tapahtuma A sattuu ja tapahtuma B sattuu*.

Yleinen yhteenlaskusääntö

Yleisen yhteenlaskusäännön mukaan

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

Toisensa poissulkevien tapahtumien yhteenlaskusääntö

Jos tapahtumat A ja B eivät voi sattua samanaikaisesti eli ovat **toisensa poissulkevia**, niin

$$A \cap B = \varnothing$$

Siten toisensa poissulkevat tapahtumat ovat joukkoina *pistevieraita*. Jos tapahtumat A ja B ovat toisensa poissulkevia eli $A \cap B = \varnothing$, niin pätee *toisensa poissulkevien tapahtumien yhteenlaskusääntö*

$$\Pr(A \cup B) = \Pr(A) + \Pr(B)$$

Olkoot $A_1, A_2, \ldots, A_k$ **pareittain toisensa poissulkevia**, jolloin $A_i \cap A_j = \varnothing$, kun $i \neq j$. Tällöin **yhdisteen**

$$A_1 \cup A_2 \cup \cdots \cup A_k = \text{"}A_1 \textit{ tai } A_2 \textit{ tai } \ldots \textit{ tai } A_k \textit{ sattuu}\text{"}$$

todennäköisyys on

$$\Pr(A_1 \cup A_2 \cup \cdots \cup A_k) = \Pr(A_1) + \Pr(A_2) + \cdots + \Pr(A_k)$$

Ehdollinen todennäköisyys

Olkoot tapahtumat A ja B otosavaruuden S osajoukkoja. Tällöin tapahtuman A **ehdollinen todennäköisyys** sillä ehdolla, että tapahtuma B on sattunut saadaan kaavalla

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

jossa $\Pr(A \cap B)$ on tapahtumien A ja B *leikkauksen* todennäköisyys eli todennäköisyys sille, että tapahtuma A on sattunut ja tapahtuma B on sattunut.

Yleinen tulosääntö

Yleisen tulosäännön mukaan

$$\Pr(A \cap B) = \Pr(A|B)\Pr(B)$$

Tarkastellaan tapahtumia $A_1, A_2, \ldots, A_k$. Tällöin **leikkauksen**

$$A_1 \cap A_2 \cap \cdots \cap A_k = \text{"}A_1 \textit{ ja } A_2 \textit{ ja } \ldots \textit{ ja } A_k \textit{ sattuvat}\text{"}$$

todennäköisyys on

$$\begin{aligned}&\Pr(A_1 \cap A_2 \cap \cdots \cap A_k) \\ &\quad = \Pr(A_1) \times \Pr(A_2 \,|\, A_1) \times \Pr(A_3 \,|\, A_1 \cap A_2) \times \cdots \times \Pr(A_k \,|\, A_1 \cap A_2 \cap \cdots \cap A_{k-1})\end{aligned}$$

Riippumattomuus ja riippumattomien tapahtumien tulosääntö

Tapahtumat A ja B ovat **riippumattomia**, jos ja vain jos *riippumattomien tapahtumien tulosääntö*

$$\Pr(A \cap B) = \Pr(A)\Pr(B)$$

pätee. Riippumattomien tapahtumien tulosääntö on yhtäpitävä sen kanssa, että

$$\Pr(A \mid B) = \Pr(A)$$

Tarkastellaan tapahtumia $A_1, A_2, \ldots, A_k$. Jos tapahtumat $A_1, A_2, \ldots, A_k$ ovat *riippumattomia*, niin tälööin pätee *riippumattomien tapahtumien tulosäännön* yleistys

$$\Pr(A_1 \cap A_2 \cap \cdots \cap A_k) = \Pr(A_1)\Pr(A_2)\Pr(A_3)\cdots\Pr(A_k)$$

Satunnaisotanta ja tulosääntö

Yleistä tulosääntöä sovelletaan tapahtumien todennäköisyyksien määräämisessä *otannassa ilman takaisinpanoa* eli *palauttamatta*.

Riippumattomien tapahtumien tulosääntöä sovelletaan tapahtumien todennäköisyyksien määräämisessä *otannassa takaisinpanolla* eli *palauttaen*.

4. Klassinen todennäköisyys ja kombinatoriikka

Kombinatoriikan perusperiaatteet

(i) **Kertolaskuperiaate**

Oletetaan, että operaatio M voidaan suorittaa m:llä eri tavalla ja operaatio N voidaan suorittaa n:llä eri tavalla ja oletetaan lisäksi, että operaatiot M ja N voidaan suorittaa toisistaan *riippumatta*. Tällöin yhdistetty operaatio

”Suoritetaan operaatio M *ja* operaatio N”

voidaan suorittaa $m \times n$:llä eri tavalla.

(ii) **Yhteenlaskuperiaate**

Oletetaan, että operaatio M voidaan suorittaa m:llä eri tavalla ja operaatio N voidaan suorittaa n:llä eri tavalla ja oletetaan lisäksi, että operaatiot M ja N ovat *toisensa poissulkevia*. Tällöin yhdistetty operaatio

”Suoritetaan operaatio M *tai* operaatio N”

voidaan suorittaa $(m + n)$:llä eri tavalla.

Joukko

Joukko on täysin määrätty, jos sen **alkiot** tunnetaan. Olkoot äärellisen joukon S (erilaiset) alkiot

$$s_1, s_2, \dots , s_n$$

Tällöin merkitään

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

Joukot A ja B ovat **samat**, jos niissä on samat alkiot eli

$$A = B$$

jos ja vain jos

$$x \in A \Leftrightarrow x \in B$$

Olkoon

$$n_S = n(S)$$

lukumääräfunktio, joka kertoo joukon S (erilaisten) alkioiden lukumäärän. Siten joukon S (erilaisten) alkioiden lukumäärä on

$$n_S = n(S) = n$$

Jono

Jono on täysin määrätty, jos sen **alkiot** ja niiden **järjestys** tunnetaan. Olkoon äärellisen jonon s i. alkio

$$s_i \,, i = 1, 2, \dots , n$$

Tällöin merkitään

$$s = (s_1, s_2, \dots , s_n)$$

Jonot $a = (a_1, a_2, \ldots , a_n)$ ja $b = (b_1, b_2, \ldots , b_n)$ ovat **samat**, jos niissä on samat alkiot samassa järjestyksessä eli

$$a = b$$

jos ja vain jos

$$a_i = b_i \,,\ i = 1, 2, \ldots , n$$

Kombinatoriikan perusongelmat

Olkoon S äärellinen joukko, jonka (erilaisten) alkioiden lukumäärä on

$$n = n(S)$$

Kombinatoriikan *perusongelmat*:

(1a) Kuinka monella eri tavalla joukon S alkiot voidaan järjestää *jonoon*?

(1b) Kuinka monella eri tavalla joukon S alkioista voidaan muodostaa k:n alkion *osajono*?

(2) Kuinka monella eri tavalla joukon S alkioista voidaan muodostaa k:n alkion *osajoukko*?

Kombinatoriikan perusongelmien ratkaisut

Olkoon S äärellinen joukko, jonka (erilaisten) alkioiden lukumäärä on

$$n = n(S)$$

Kombinatoriikan *perusongelmien ratkaisut*:

(1a) Kutsumme joukon S *kaikkien alkioiden jonoja* joukon S alkioiden **permutaatioiksi**. Joukon S alkioiden kaikkien mahdollisten permutaatioiden lukumäärä on

$$n!$$

jossa

$$n! = n\times(n-1)\times \cdots \times 2\times 1$$

on *n-kertoma*.

(1b) Kutsumme joukon S *k:n alkion osajonoja* joukon S alkioiden ***k*-permutaatioiksi** eli **variaatioiksi**. Joukon S alkioiden kaikkien mahdollisten k-permutaatioiden lukumäärä on

$$\mathrm{P}(n,k) = \frac{n!}{(n-k)!} = n\times(n-1)\times\cdots\times(n-k-1)$$

(2) Kutsumme joukon S *k:n alkion osajoukkoja* joukon S alkioiden *k alkiota sisältäviksi* **kombinaatioiksi**. Joukon S alkioiden kaikkien mahdollisten k alkiota sisältävien kombinaatioiden lukumäärä on

$$\mathrm{C}(n,k) = \binom{n}{k}$$

jossa

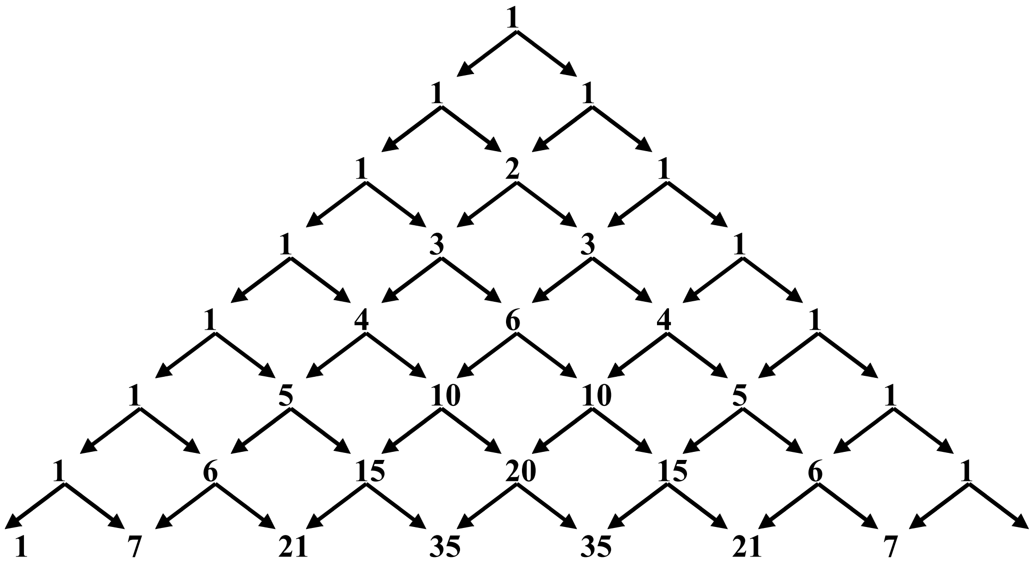
$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

on *binomikerroin*.

Kombinatoriikan perusongelmien (1a), (1b) ja (2) ratkaisut voidaan perustella kombinatoriikan kertolaskuperiaatteen avulla.

Pascalin kolmio

Binomikertoimet saadaan ns. *Pascalin kolmiosta*. Alla on annettu Pascalin kolmion 8 ensimmäistä riviä.



Lukuun ottamatta kolmion reunoilla olevia ykkösiä jokainen kolmion luvuista on saatu laskemalla yhteen kaksi edeltävän rivin lukua nuolten suuntaan.

Pascalin kolmio ja binomikertoimet

Pascalin kolmion (n+1). rivin luvut voidaan ilmaista binomikertoimien avulla seuraavassa muodossa:

$$\binom{n}{0},\binom{n}{1},\binom{n}{2},\ldots,\binom{n}{n-1},\binom{n}{n}$$

Pascalin kolmion muodostamissääntö voidaan ilmaista binomikertoimien avulla seuraavassa muodossa:

$$\binom{n}{k}=\binom{n-1}{k-1}+\binom{n-1}{k}$$

Kaavan mukaan Pascalin kolmion n. rivin k. luku saadaan laskemalla yhteen rivin (n–1) luvut paikoissa (k–1) ja k.

Se, että Pascalin kolmio on symmetrinen kolmion rivien keskikohdan suhteen, voidaan ilmaista binomikertoimien avulla seuraavassa muodossa:

$$\binom{n}{k}=\frac{n!}{k!(n-k)!}=\binom{n}{n-k}$$

Binomikaava

Binomikaavan mukaan binomin

$$x + y$$

k. potenssi voidaan esittää muodossa

$$(x+y)^k = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

Äärellisen joukon osajoukkojen lukumäärä

Olkoon joukon S alkioiden lukumäärä $n = n(S)$. Tällöin joukon S osajoukkojen lukumäärä on

$$N = 2^n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n-1} + \binom{n}{n}$$

Multinomikerroin

Olkoon joukon S alkioiden lukumäärä $n = n(S)$. Oletetaan, että positiiviset kokonaisluvut

$$n_i \,,\ i = 1, 2, \ldots, k$$

toteuttavat ehdon

$$n_1 + n_2 + \cdots + n_k = n$$

Ositetaan joukko S *pistevieraisiin* osajoukkoihin

$$A_i \,,\ i = 1, 2, \ldots, k$$

niin, että joukossa A_i on $n_i = n(A_i)$ alkiota. Kuinka monella erilaisella tavalla tällainen ositus voidaan tehdä?

Vastauksen antaa **multinomikerroin**

$$\binom{n}{n_1 \; n_2 \; \cdots \; n_k} = \frac{n!}{n_1!\, n_2! \cdots n_k!}$$

jossa siis

$$n_1 + n_2 + \cdots + n_k = n$$

Huomaa, että jos $k = 2$, saadaan binomikerroin

$$\binom{n}{n_1 \; n_2} = \frac{n!}{n_1!\, n_2!} = \binom{n}{n_1} = \binom{n}{n_2}$$

jossa

$$n_1 + n_2 = n$$

5. Todennäköisyyden aksioomat

Todennäköisyys äärellisissä otosavaruuksissa

Tarkastellaan ensin todennäköisyyden määrittelemistä *äärellisissä otosavaruuksissa*. Huomattava osa todennäköisyyden peruslaskusäännöistä voidaan todistaa *äärellisten otosavaruuksien aksioomista*.

Boolen algebra

Olkoon S joukko ja jokin $\mathfrak{F}$ joukon S osajoukkojen muodostama *perhe* eli

$$A \in \mathfrak{F} \Rightarrow A \subset S$$

Joukkoperhe $\mathfrak{F}$ on **Boolen algebra**, jos seuraavat 3 ehtoa pätevät:

(i) $\varnothing \in \mathfrak{F}$

(ii) $A \in \mathfrak{F} \Rightarrow A^c \in \mathfrak{F}$

(iii) $A \in \mathfrak{F}, B \in \mathfrak{F} \Rightarrow A \cup B \in \mathfrak{F}$

Kutsumme todennäköisyyslaskennassa perusjoukkoa S *otosavaruudeksi* ja Boolen algebraan $\mathfrak{F}$ kuuluvia otosavaruuden S osajoukkoja A *tapahtumiksi*.

Olkoot

$$A \in \mathfrak{F}, B \in \mathfrak{F}$$

Boolen algebran aksioomista seuraa suoraan, että

$$\varnothing \in \mathfrak{F}$$
$$A^c \in \mathfrak{F}$$
$$B^c \in \mathfrak{F}$$
$$A \cup B \in \mathfrak{F}$$

Lisäksi voidaan osoittaa, että

$$S \in \mathfrak{F}$$
$$A \cap B = (A^c \cup B^c)^c \in \mathfrak{F}$$
$$A \setminus B = A \cap B^c \in \mathfrak{F}$$

Todennäköisyyden aksioomat äärellisissä otosavaruuksissa

Olkoon S äärellinen joukko ja $\mathfrak{F}$ jokin joukon S osajoukkojen muodostama *Boolen algebra*. Olkoon lisäksi Pr *joukkofunktio*, joka liittää jokaiseen Boolen algebraan $\mathfrak{F}$ kuuluvaan joukon S osajoukkoon A reaalikuvun eli

$$A \in \mathfrak{F} \Rightarrow A \subset S \Rightarrow \Pr(A) \in \mathbb{R}$$

Joukkofunktio Pr on äärellisen otosavaruuden **todennäköisyysmitta**, jos seuraavat 3 ehtoa pätevät:

(i) $\Pr(S) = 1$

(ii) $0 \le \Pr(A) \le 1$ kaikille $A \in \mathfrak{F}$

(iii) $A \in \mathfrak{F}, B \in \mathfrak{F}, A \cap B = \varnothing \;\Rightarrow \Pr(A \cup B) = \Pr(A) + \Pr(B)$

Äärellinen todennäköisyyskenttä

Kolmikko

$$(S, \mathfrak{F}, \Pr)$$

on **äärellinen todennäköisyyskenttä**, jos S on äärellinen otosavaruus, $\mathfrak{F}$ on otosavaruudessa S määritelty Boolen algebra ja Pr on Boolen algebrassa $\mathfrak{F}$ määritelty todennäköisyysmitta.

Riippumattomuus ja riippumattomien tapahtumien tulosääntö

Tapahtumat A ja B ovat riippumattomia, jos ja vain jos *riippumattomien tapahtumien tulosääntö*

$$\Pr(A \cap B) = \Pr(A)\Pr(B)$$

pätee.

Todennäköisyys mielivaltaisissa otosavaruuksissa

Tarkastellaan todennäköisyyden määrittelemistä *mielivaltaisissa otosavaruuksissa*.

σ-algebra

Olkoon S joukko ja jokin $\mathfrak{F}$ joukon S osajoukkojen muodostama *perhe* eli

$$A \in \mathfrak{F} \;\Rightarrow A \subset S$$

Joukkoperhe $\mathfrak{F}$ on **σ-algebra**, jos seuraavat 3 ehtoa pätevät:

(i) $\varnothing \in \mathfrak{F}$

(ii) $A \in \mathfrak{F} \;\Rightarrow A^c \in \mathfrak{F}$

(iii) $A_1, A_2, A_3, \ldots \in \mathfrak{F} \;\Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathfrak{F}$

Kutsumme perusjoukkoa S *otosavaruudeksi* ja σ-algebraan $\mathfrak{F}$ kuuluvia otosavaruuden S osajoukkoja A *tapahtumiksi*.

Kaikki Boolen algebroille todistetut teoreemat pätevät myös σ-algebroille.

Todennäköisyyden aksioomat mielivaltaisissa otosavaruuksissa

Olkoon S joukko ja $\mathfrak{F}$ jokin joukon S osajoukkojen muodostama *σ-algebra*. Olkoon lisäksi Pr *joukkofunktio*, joka liittää jokaiseen σ-algebraan $\mathfrak{F}$ kuuluvaan joukon S osajoukkoon A reaaliluvun eli

$$A \in \mathfrak{F} \;\Rightarrow A \subset S \;\Rightarrow \Pr(A) \in \mathbb{R}$$

Joukkofunktio Pr on **todennäköisyysmitta**, jos seuraavat 3 ehtoa pätevät:

(i) $\Pr(S) = 1$

(ii) $0 \le \Pr(A) \le 1$ kaikille $A \in \mathfrak{F}$

(iii) $A_1, A_2, A_3, \ldots \in \mathfrak{F}$ ja $A_i \cap A_j = \varnothing$, $i \neq j \;\Rightarrow\; \Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i)$

Todennäköisyyskenttä

Kolmikko

$$(S, \mathfrak{F}, \Pr)$$

on **todennäköisyyskenttä**, jos S on otosavaruus, $\mathfrak{F}$ on otosavaruudessa S määritelty σ-algebra ja Pr on σ-algebrassa $\mathfrak{F}$ määritelty todennäköisyysmitta.

Kaikki äärellisille todennäköisyyskentille todistetut teoreemat pätevät myös äärettömissä todennäköisyyskentissä.

Epämitalliset joukot

On syytä huomata, että jos otosavaruus S on ääretön, sen kaikille osajoukoille ei voida välttämättä määritellä todennäköisyyttä. Niitä otosavaruuden S osajoukkoja, joille todennäköisyys voidaan määritellä, sanotaan *mitallisiksi* ja niitä, joille todennäköisyyttä ei voida määritellä, sanotaan *epämitallisiksi*.

Voidaan osoittaa, että otosavaruuden S mitalliset osajoukot muodostavat aina σ-algebran.

Lause 1.

Olkoon $(S, \mathfrak{F}, \Pr)$ todennäköisyyskenttä ja $A_1, A_2, A_3, \ldots \in \mathfrak{F}$

Tällöin pätee:

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

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

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

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

Lause 2.

Olkoon $(S, \mathfrak{F}, \Pr)$ todennäköisyyskenttä ja $A_1, A_2, A_3, \ldots \in \mathfrak{F}$

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

$$\lim_{i\to\infty} \Pr(A_i) = 0$$

Voidaan osoittaa, että Kolmogorovin aksiooma

(iii) $A_1, A_2, A_3, \ldots \in \mathfrak{F}$ ja $A_i \cap A_j = \varnothing$, $i \neq j \;\Rightarrow\; \Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i)$

on yhtäpitävä aksioomien

(iii)′ $A_1 \cap A_2 = \varnothing \Rightarrow \Pr\left(A_1 \cup A_2\right) = \Pr\left(A_1\right) + \Pr\left(A_2\right)$

(iv) $A_1 \supset A_2 \supset A_3 \supset \cdots \to \varnothing \Rightarrow \lim_{n\to+\infty} \Pr\left(A_n\right) = 0$

kanssa. Aksioomaa (iv) kutsutaan usein todennäköisyyden **jatkuvuusaksioomaksi**.

6. Kokonaistodennäköisyys ja Bayesin kaava

Ositus

Joukon S osajoukot

$$B_1, B_2, \ldots , B_n$$

muodostavat joukon S **osituksen**, jos seuraavat 3 ehtoa pätevät:

(i) $B_i \neq \varnothing$, $i = 1, 2, \ldots, n$

(ii) $B_i \cap B_j = \varnothing$, $i \neq j$

(iii) $S = B_1 \cup B_2 \cup \cdots \cup B_n$

Kokonaistodennäköisyyden kaava

Olkoon A epätyhjä otosavaruuden S osajoukko:

$$A \subset S \text{ , } A \neq \varnothing$$

Oletetaan, että joukot

$$B_1, B_2, \ldots , B_n$$

muodostavat otosavaruuden S *osituksen*. Tällöin pätee **kokonaistodennäköisyyden** *kaava*

$$\Pr(A) = \sum_{i=1}^{n} \Pr(B_i) \Pr(A \mid B_i)$$

Bayesin kaava

Olkoon A epätyhjä otosavaruuden S osajoukko:

$$A \subset S \text{ , } A \neq \varnothing$$

Oletetaan, että joukot

$$B_1, B_2, \ldots , B_n$$

muodostavat otosavaruuden S *osituksen*. Tällöin pätee **Bayesin kaava**

$$\Pr(B_i \mid A) = \frac{\Pr(B_i) \Pr(A \mid B_i)}{\sum_{i=1}^{n} \Pr(B_i) \Pr(A \mid B_i)} \text{ , } i = 1, 2, \ldots, n$$

7. Verkot ja todennäköisyyslaskenta

Verkko

Verkko eli *graafi* muodostuu **pisteiden** *joukosta V*, **särmien** *joukosta A* ja **insidenssikuvauksesta**

$$\Delta : A \to V \times V$$

jossa

$$V \neq \varnothing \,,\; A \neq \varnothing \,,\; A \cap V = \varnothing$$

Insidenssikuvaus Δ kertoo, *mitkä verkon pisteistä ovat särmien yhdistämiä.*

Verkkoja tarkastellaan tässä *suunnattuina verkkoina*, millä tarkoitetaan sitä, että verkon jokaisella särmällä on *suunta*, joka osoittaa särmän *alkupisteestä* särmän *loppupisteeseen*.

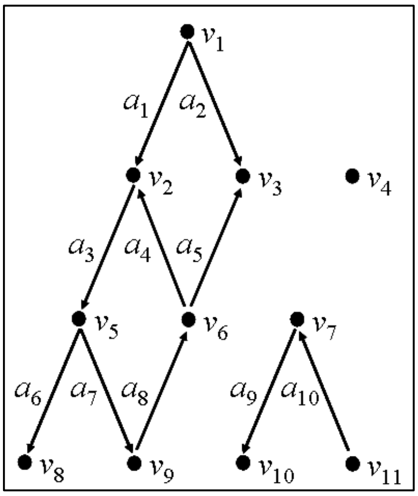
Esimerkki:

Kuviossa oikealla

$$V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8, v_9, v_{10}, v_{11}\}$$
$$A = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10}\}$$

ja esimerkiksi

$$\Delta(a_5) = (v_6, v_3)$$



Reitti

Särmät

$$\{a_1, a_2, \ldots, a_{k-1}\}$$

muodostavat **reitin** pisteestä v_1 pisteeseen v_k , jos on olemassa pisteet

$$v_1,\, v_2,\, \ldots\, ,\, v_k$$

siten, että

$$\Delta(a_i) = (v_i, v_{i+1}) \,,\; i = 1, 2, \ldots, k-1$$

Jos pisteestä v_1 pisteeseen v_k on reitti, sanotaan, että reitti *vie* pisteestä v_1 pisteeseen v_k tai, että pisteestä v_1 *pääsee* pisteeseen v_k .

Esimerkki:

Kuviossa yllä särmät a_1, a_3, a_7, a_8 muodostavat reitin pisteestä v_1 pisteeseen v_6.

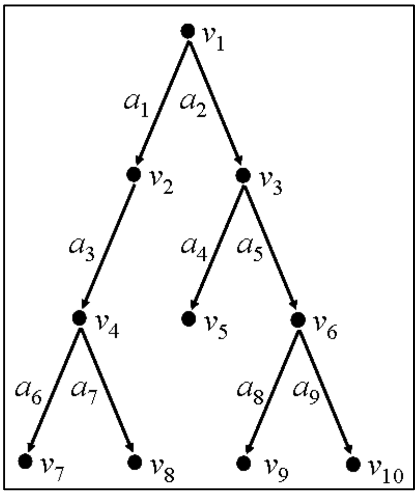
Puu

Verkko on **puu**, jonka *juurena* on piste v_1, jos seuraavat ehdot pätevät:

(i) Verkko on *yhtenäinen*.

(ii) Verkossa ei ole *silmukoita*.

(iii) Jos $w \neq v_1$ on mielivaltainen verkon piste, pisteestä v_1 pisteeseen w pääsee *täsmälleen yhtä reittiä* pitkin.



Esimerkki:

Yllä olevan kuvion verkko *ei ole* puu, koska siinä on silmukoita ja se ei ole yhtenäinen.

Sen sijaan oikealla olevan kuvion verkko *on* puu.

Puudiagrammin konstruointi satunnaisilmiölle

Satunnaisilmiötä *voidaan kuvata puudiagrammilla*, jos ilmiö osataan esittää seuraavassa muodossa:

(i) Ilmiöllä on yksi *alkutila* ja yksi tai useampia *lopputiloja*.

(ii) Ilmiö koostuu *vaihtoehtoisista tapahtumajonoista*.

(iii) Tapahtumajonoissa edetään *vaiheittain tapahtumasta toiseen* lähtien ilmiön alkutilasta ja päätyen johonkin ilmiön lopputiloista.

(iv) Jokaisessa *vaiheessa* kohdataan yksi tai useampia *tapahtumavaihtoehtoja*, joista yksi *realisoituu* ja johtaa *uusin* tapahtumavaihtoehtoihin.

Satunnaisilmiötä vastaava *puudiagrammi konstruoidaan* seuraavalla tavalla:

(i) Asetetaan puun *juuri* vastaamaan ilmiön *alkutilaa*.

(ii) Asetetaan puun *loppupisteet* (”oksien kärjet”) vastaamaan ilmiön *lopputiloja*.

(iii) Asetetaan puun *pisteet* (”oksien haarautumiskohdat”) vastaamaan ilmiön *tapahtumia*.

(iv) Viedään puun jokaisesta pisteestä *särmä* (”oksa”) kaikkiin sellaisiin pisteisiin, joita vastaavat *tapahtumavaihtoehdot* ovat ilmiön *siinä vaiheessa mahdollisia*.

(v) Liitetään jokaiseen pisteestä *lähtevään* särmään *siinä vaiheessa* mahdollisten *tapahtumavaihtoehtojen todennäköisyydet*.

Puutodennäköisyydet

Puutodennäköisyydellä tarkoitetaan todennäköisyyttä päästä *puun alkupisteestä yhden tai useamman muun puun pisteen määräämään yhdistettyyn tapahtumaan*.

Pisteen todennäköisyys saadaan määräämällä alkupisteestä ko. pisteeseen vievän *reitin todennäköisyys*.

Reitin todennäköisyys saadaan soveltamalla reittiin kuuluvien särmien todennäköisyyksiin *tulosääntöä*.

Usean pisteen määräämän yhdistetyn tapahtuman todennäköisyys saadaan soveltamalla ko. pisteisiin vievien reittien todennäköisyyksiin *yhteenlaskusääntöä*.

Puutodennäköisyyksien tulosääntö

Reitin todennäköisyys saadaan määräämällä reittiin kuuluvien särmien todennäköisyyksien *tulo*.

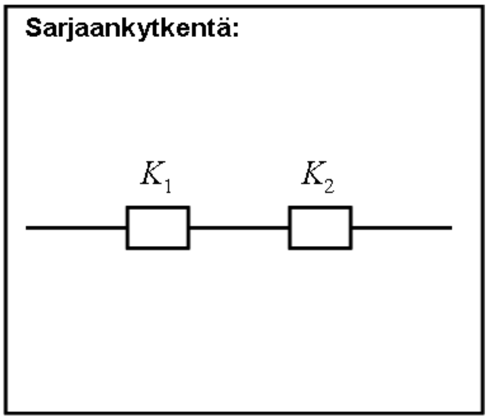
Puutodennäköisyyksien yhteenlaskusääntö

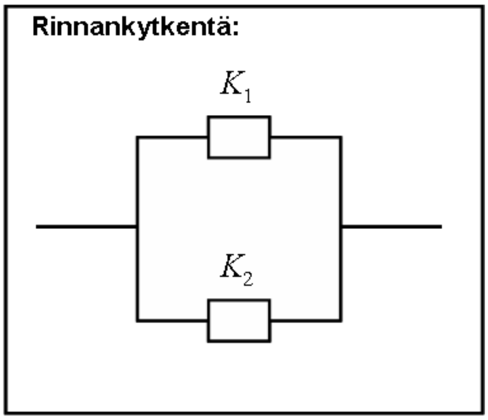
Jos useita (loppu-) tiloja yhdistetään yhdeksi tapahtumaksi, näin saadun *yhdistetyn tapahtuman* todennäköisyys saadaan määräämällä ko. tiloihin vievien reittien todennäköisyyksien *summa*.

Toimintaverkot

Toimintaverkko on *systeemi*, joka koostuu *komponenteista*, jotka on kytketty **rinnan** tai **sarjaan**.

Alla olevat kytkentäkaaviot kuvaavat kahden komponentin K_1 ja K_2 muodostamia sarjaan- ja rinnankytkentöjä.





Sarjaan kytkennän toimintatodennäköisyys

Oletetaan, että komponentit K_1 ja K_2 on kytketty *sarjaan* ja oletetaan lisäksi, että komponentin K_1 toiminta (tai toimimattomuus) *ei riipu* komponentin K_2 toiminnasta (ja kääntäen).

Komponenttien K_1 ja K_2 muodostama *sarjaan kytkentä toimii*, *jos komponentti* K_1 *toimii ja komponentti* K_2 *toimii*.

Määritellään tapahtumat

A_1 = ”Komponentti K_1 toimii”

A_2 = ”Komponentti K_2 toimii”

Olkoot tapahtumien A_1 ja A_2 todennäköisyydet

$p_1 = \Pr(A_1)$

$p_2 = \Pr(A_2)$

Koska tapahtumat A_1 ja A_2 ovat oletuksen mukaan *riippumattomia*, saadaan komponenttien K_1 ja K_2 muodostama *sarjaan kytkennän toimintatodennäköisyydeksi riippumattomien tapahtumien tulosäännön* mukaan

Pr(Komponentti K_1 toimii *ja* komponentti K_2 toimii)

$= \Pr(A_1 \cap A_2)$

$= \Pr(A_1)\Pr(A_2)$

$= p_1 p_2$

= Pr(Komponentti K_1 toimii)Pr(Komponentti K_2 toimii)

Rinnankytkennän toimintatodennäköisyys

Oletetaan, että komponentit K_1 ja K_2 on kytketty *rinnan* ja oletetaan lisäksi, että komponentin K_1 toiminta (tai toimimattomuus) *ei riipu* komponentin K_2 toiminnasta (ja kääntäen).

Komponenttien K_1 ja K_2 muodostama *rinnan kytkentä toimii*, *jos komponentti* K_1 *toimii tai komponentti* K_2 *toimii* (*tai molemmat toimivat*).

Määritellään tapahtumat

A_1 = ”Komponentti K_1 toimii”

A_2 = ”Komponentti K_2 toimii”

Olkoot tapahtumien A_1 ja A_2 todennäköisyydet

$p_1 = \Pr(A_1)$

$p_2 = \Pr(A_2)$

Koska tapahtumat A_1 ja A_2 ovat oletuksen mukaan *riippumattomia*, saadaan komponenttien K_1 ja K_2 muodostama *rinnan kytkennän toimintatodennäköisyydeksi yleisen yhteenlaskusäännön* ja *riippumattomien tapahtumien tulosäännön* mukaan

Pr(Komponentti K_1 toimii *tai* komponentti K_2 toimii)

$= \Pr(A_1 \cup A_2)$

$= \Pr(A_1) + \Pr(A_2) - \Pr(A_1 \cap A_2)$

$= \Pr(A_1) + \Pr(A_2) - \Pr(A_1)\Pr(A_2)$

$= p_1 + p_2 - p_1 p_2$

= Pr(Komponentti K_1 toimii) + Pr(Komponentti K_2 toimii)
  – Pr(Komponentti K_1 toimii)Pr(Komponentti K_2 toimii)

8. Satunnaismuuttujat ja todennäköisyysjakaumat

Satunnaismuuttuja

Olkoon $(S, \mathfrak{F}, \Pr)$ *todennäköisyyskenttä*, jossa

S = *otosavaruus* (perusjoukko)

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

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

Jos ξ on otosavaruuden S *reaaliarvoinen* (mitallinen) *funktio* eli

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

niin ξ on **satunnaismuuttuja**. Jos siis

$$s \in S$$

niin

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

Todennäköisyysjakauma

Satunnaismuuttujan ξ **todennäköisyysjakaumalla** tarkoitetaan kuvauksen

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

reaalilukujen joukkoon *indusoimaa* todennäköisyysmittaa.

Diskreetti satunnaismuuttuja

Olkoon

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

satunnaismuuttuja. Jos otosavaruus S on *äärellinen* tai *numeroituvasti ääretön*, jolloin myös funktion ξ arvoalue on äärellinen tai numeroituvasti äärellinen, sanotaan satunnaismuuttujaa ξ **diskreetiksi**.

Diskreetin satunnaismuuttujan pistetodennäköisyysfunktio

Olkoon

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

diskreetti satunnaismuuttuja, jonka arvot ovat $x_1, x_2, x_3, \ldots, x_n$ tai $x_1, x_2, x_3, \ldots, x_n, \ldots$

Olkoon

$$T = \{x_1, x_2, x_3, \ldots, x_n\}$$

tai

$$T = \{x_1, x_2, x_3, \ldots, x_n, \ldots\}$$

satunnaismuuttujan ξ arvojen joukko. Joukko T on siis äärellinen tai numeroituvasti ääretön.

Reaaliarvoinen funktio f määrittelee diskreetin satunnaismuuttujan ξ **pistetodennäköisyysfunktion**, jos seuraavat 3 ehtoa pätevät:

(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

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

on satunnaismuuttujan ξ arvoa x_i vastaava **pistetodennäköisyys**.

Diskreetti todennäköisyysjakauma

Jos f on diskreetin satunnaismuuttujan

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

pistetodennäköisyysfunktio, sanomme, että satunnaismuuttuja ξ noudattaa **diskreettiä todennäköisyysjakaumaa**, jonka pistetodennäköisyysfunktio on f.

Reaaliakselin välien todennäköisyydet

Olkoon

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

diskreetti satunnaismuuttuja ja f vastaava *pistetodennäköisyysfunktio*. Tällöin reaaliakselin välin $[a, b]$ todennäköisyys on

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

Jatkuva satunnaismuuttuja

Olkoon

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

satunnaismuuttuja. Satunnaismuuttuja ξ on **jatkuva**, jos seuraavat 2 ehtoa pätevät:

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

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

Jatkuvan satunnaismuuttujan tiheysfunktio

Olkoon

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

jatkuva satunnaismuuttuja.

Reaaliarvoinen funktio f määrittelee satunnaismuuttujan ξ **tiheysfunktion**, jos seuraavat 4 ehtoa pätevät:

(1) $f(x)$ on x:n jatkuva funktio

(2) $f(x) \geq 0$ kaikille x

(3) $\int_{-\infty}^{+\infty} f(x)dx = 1$

(4) $\Pr(a \leq \xi \leq b) = \int_{a}^{b} f(x)dx$

Jatkuva todennäköisyysjakauma

Jos f on jatkuvan satunnaismuuttujan

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

tiheysfunktio, sanomme, että satunnaismuuttuja ξ noudattaa **jatkuvaa todennäköisyysjakaumaa**, jonka tiheysfunktio on f.

Reaaliakselin välien todennäköisyydet

Olkoon

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

jatkuva satunnaismuuttuja ja f vastaava *tiheysfunktio*. Tällöin reaaliakselin välin $[a, b]$ todennäköisyys on

$$\Pr(a \leq \xi \leq b) = \int_{a}^{b} f(x)dx$$

Huomaa, että

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

kaikille $x \in \mathbb{R}$.

9. Kertymäfunktio

Kertymäfunktio

Olkoon

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

satunnaismuuttuja. Satunnaismuuttujan ξ **kertymäfunktio** on reaaliarvoinen funktio

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

Funktio

$$F : \mathbb{R} \to [0,1]$$

on *kertymäfunktio*, jos ja vain jos

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

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

(3) F on *ei-vähenevä*:
$F(x_1) \le F(x_2)$, jos $x_1 \le x_2$

(4) F on *jatkuva oikealta*:
$\lim_{h \to 0+} F(x+h) = F(x)$

Jos funktio

$$F : \mathbb{R} \to [0,1]$$

on kertymäfunktio, niin

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

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

Diskreetin jakauman kertymäfunktio

Olkoon

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

diskreetti satunnaismuuttuja ja *f* vastaava *pistetodennäköisyysfunktio*. Tällöin satunnaismuuttujan ξ *kertymäfunktio* on

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

ja kääntäen

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

Reaaliakselin välien todennäköisyydet

Olkoon

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

diskreetti satunnaismuuttuja, f vastaava *pistetodennäköisyysfunktio* ja F vastaava *kertymäfunktio*. Tällöin reaaliakselin välin $(a, b]$ todennäköisyys on

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

Jatkuvan jakauman kertymäfunktio

Olkoon

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

jatkuva satunnaismuuttuja ja f vastaava *tiheysfunktio*. Tällöin satunnaismuuttujan ξ *kertymäfunktio* on

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

ja kääntäen

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

Reaaliakselin välien todennäköisyydet

Olkoon

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

jatkuva satunnaismuuttuja, f vastaava *tiheysfunktio* ja F vastaava *kertymäfunktio*. Tällöin reaaliakselin välin $(a, b]$ todennäköisyys on

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

Huomaa, että *jatkuville satunnaismuuttujille* pätee:

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

10. Jakaumien tunnusluvut

Diskreetin satunnaismuuttujan odotusarvo

Olkoon

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

satunnaismuuttujan X *pistetodennäköisyysfunktio*. Tällöin satunnaismuuttujan X ja sitä vastaavan todennäköisyysjakauman **odotusarvo** on ei-satunnainen vakio

$$\mathrm{E}(X) = \mu_X = \sum_i x_i f(x_i) = \sum_i x_i \Pr(X = x_i) = \sum_i x_i p_i$$

Diskreetin satunnaismuuttujan X odotusarvo *on olemassa*, jos

$$\sum_i |\, x_i \,|\, f(x_i) = \sum_i |\, x_i \,| \Pr(X = x_i) = \sum_i |\, x_i \,|\, p_i < \infty$$

Jatkuvan satunnaismuuttujan odotusarvo

Olkoon

$$f(x)$$

jatkuvan satunnaismuuttujan X *tiheysfunktio*. Tällöin satunnaismuuttujan X ja sitä vastaavan todennäköisyysjakauman **odotusarvo** on ei-satunnainen vakio

$$\mathrm{E}(X) = \mu_X = \int_{-\infty}^{+\infty} x f(x) dx$$

Jatkuvan satunnaismuuttujan X odotusarvo *on olemassa*, jos

$$\int_{-\infty}^{+\infty} |\, x \,|\, f(x) dx < \infty$$

Odotusarvon ominaisuuksia

Satunnaismuuttujan X odotusarvo on satunnaismuuttujan X todennäköisyysjakauman *todennäköisyysmassan* **painopiste**.

Olkoon a vakio. Tällöin

$$\mathrm{E}(a) = a$$

Olkoot X_i , $i = 1, 2, \ldots , n$ satunnaismuuttujia ja a_i , $i = 1, 2, \ldots , n$ vakioita. Tällöin

$$\mathrm{E}\left(\sum_{i=1}^{n} a_i X_i \right) = \sum_{i=1}^{n} a_i \, \mathrm{E}(X_i)$$

Diskreetin satunnaismuuttujan funktion odotusarvo

Olkoon

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

diskreetin satunnaismuuttujan X *pistetodennäköisyysfunktio* ja olkoon g reaaliarvoinen funktio.

Tällöin satunnaismuuttujan *g*(*X*) **odotusarvo** on ei-satunnainen vakio

$$\mathrm{E}(g(X)) = \mu_{g(X)} = \sum_i g(x_i) f(x_i) = \sum_i g(x_i) \Pr(X = x_i) = \sum_i g(x_i) p_i$$

Jatkuvan satunnaismuuttujan funktion odotusarvo

Olkoon

$$f(x)$$

jatkuvan satunnaismuuttujan *X* *tiheysfunktio* ja olkoon *g* reaaliarvoinen funktio.

Tällöin satunnaismuuttujan *g*(*X*) **odotusarvo** on ei-satunnainen vakio

$$\mathrm{E}(g(X)) = \mu_{g(X)} = \int_{-\infty}^{+\infty} g(x) f(x) dx$$

Varianssi

Olkoon satunnaismuuttujan *X* odotusarvo

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

Tällöin satunnaismuuttujan *X* **varianssi** on ei-satunnainen vakio

$$\mathrm{D}^2(X) = \mathrm{Var}(X) = \sigma_X^2 = \mathrm{E}[(X - \mu_X)^2]$$

Varianssi voidaan laskea myös kaavalla

$$\mathrm{D}^2(X) = \mathrm{Var}(X) = \sigma_X^2 = \mathrm{E}(X^2) - \mu_X^2$$

jossa

$\mathrm{E}(X^2)$ = satunnaismuuttujan *X* 2.momentti

Diskreetin satunnaismuuttujan varianssi

Olkoon

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

diskreetin satunnaismuuttujan *X pistetodennäköisyysfunktio*. Tällöin satunnaismuuttujan *X* **varianssi** on

$$\mathrm{D}^2(X) = \mathrm{Var}(X) = \sigma_X^2 = \mathrm{E}[(X - \mu_X)^2] = \sum_i (x_i - \mu_X)^2 p_i$$

Diskreetin satunnaismuuttujan *X* varianssi *on olemassa*, jos

$$\mathrm{D}^2(X) < \infty$$

Jatkuvan satunnaismuuttujan varianssi

Olkoon

$$f(x)$$

jatkuvan satunnaismuuttujan *X tiheysfunktio*. Tällöin satunnaismuuttujan *X* **varianssi** on

$$\mathrm{D}^2(X) = \mathrm{Var}(X) = \sigma_X^2 = \mathrm{E}[(X-\mu_X)^2] = \int_{-\infty}^{+\infty} (x-\mu_X)^2 f(x)dx$$

Jatkuvan satunnaismuuttujan X varianssi *on olemassa*, jos

$$\mathrm{D}^2(X) < \infty$$

Standardipoikkeama

Satunnaismuuttujan X **standardipoikkeama** on ei-satunnainen vakio

$$\mathrm{D}(X) = \sigma_X = \sqrt{\mathrm{E}[(X-\mu_X)^2]}$$

Varianssin ominaisuuksia

Satunnaismuuttujan X varianssi ja standardipoikkeama kuvaavat satunnaismuuttujan X todennäköisyysjakauman todennäköisyysmassan *hajaantuneisuutta* todennäköisyysmassan painopisteen $\mathrm{E}(X) = \mu_X$ ympärillä.

Olkoon a vakio. Tällöin

$$\mathrm{D}^2(a) = \mathrm{Var}(a) = 0$$

Olkoot X_i , $i = 1, 2, \ldots , n$ *riippumattomia* satunnaismuuttujia ja a_i , $i = 1, 2, \ldots , n$ vakioita. Tällöin

$$\mathrm{D}^2\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 \, \mathrm{D}^2(X_i)$$

Markovin epäyhtälö

Olkoon $g(X)$ satunnaismuuttujan X *positiivinen* reaaliarvoinen funktio, jonka odotusarvo on

$$\mathrm{E}(g(X))$$

Tällöin jokaiselle reaaliselle, ei-satunnaiselle vakiolle $a > 0$ pätee **Markovin epäyhtälö**:

$$\Pr(g(X) \geq a) \leq \frac{\mathrm{E}(g(X))}{a}$$

Tshebyshevin epäyhtälö

Olkoon X satunnaismuuttuja, jonka odotusarvo on

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

ja varianssi on

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

Tällöin pätee **Tshebyshevin epäyhtälö**:

$$\Pr(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Komplementtitapahtuman todennäköisyyden kaavasta seuraa, että

$$\Pr(|X-\mu| < k\sigma) = 1 - \Pr(|X-\mu| \geq k\sigma) \geq 1 - \frac{1}{k^2}$$

Momentit

Olkoon X satunnaismuuttuja. Tällöin satunnaismuuttujan X^k odotusarvo

$$\mathrm{E}(X^k) = \alpha_k \text{ , } k = 0,1,2,\ldots$$

on satunnaismuuttujan X ***k*. momentti** eli ***k*. momentti origon suhteen**.

Erityisesti:

$$\alpha_0 = 1$$
$$\alpha_1 = \mathrm{E}(X) = \mu$$

Siten satunnaismuuttujan X 1. momentti origon suhteen on satunnaismuuttujan X *odotusarvo*.

Olkoon X satunnaismuuttuja, jonka odotusarvo on

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

Tällöin satunnaismuuttujan $(X-\mu)^k$ odotusarvo

$$\mathrm{E}\left[(X-\mu)^k\right] = \mu_k \text{ , } k = 0,1,2,\ldots$$

on satunnaismuuttujan X ***k*. keskusmomentti** eli ***k*. momentti painopisteen μ suhteen**.

Erityisesti:

$$\mu_1 = 0$$
$$\mu_2 = \mathrm{E}\left[(X-\mu)^2\right] = \sigma^2 = \mathrm{Var}(X) = \mathrm{D}^2(X)$$

Siten satunnaismuuttujan X 1. keskusmomentti *häviää* ja 2. keskusmomentti on satunnaismuuttujan X *varianssi*.

Satunnaismuuttujan X k. origomomentti *on olemassa*, jos

$$\mathrm{E}(|X|^k) < \infty$$

Satunnaismuuttujan X k. keskusmomentti *on olemassa*, jos vastaava origomomentti on olemassa.

Voidaan osoittaa, että jos

$$\mathrm{E}(|X|^n) < \infty$$

jollekin $n \in \mathbb{N}$, niin

$$\mathrm{E}(|X|^k) < \infty$$

kaikille $k < n$. Jos siis satunnaismuuttujalla on n. origomomentti, niin sillä on myös kaikki alempien kertalukujen momentit.

Vinous

Tunnuslukua

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}}$$

käytetään *todennäköisyysjakaumien* **vinouden** *mittana*.

Jos todennäköisyysjakauman piste-todennäköisyys- tai tiheysfunktio on *yksihuippuinen*, pätee seuraava:

$\gamma_1 < 0$: Jakauma on **negatiivisesti vino** eli **vino vasemmalle**, jolloin jakauman vasen häntä on pitempi kuin oikea häntä.

$\gamma_1 = 0$: Jakauma on **symmetrinen**.

$\gamma_1 > 0$: Jakauma on **positiivisesti vino** eli **vino oikealle**, jolloin jakauman oikea häntä on pitempi kuin vasen häntä.

Huomautus: *Normaalijakaumalle* $\gamma_1 = 0$.

Huipukkuus

Tunnuslukua

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3$$

käytetään *todennäköisyysjakaumien* **huipukkuuden** *mittana*.

Jos todennäköisyysjakauman pistetodennäköisyys- tai tiheysfunktio on *yksihuippuinen*, pätee seuraava:

$\gamma_2 > 0$: Jakauma on **huipukas** (normaalijakaumaan verrattuna).

$\gamma_2 = 0$: Jakauma on **yhtä huipukas kuin normaalijakauma**.

$\gamma_2 < 0$: Jakauma on **laakea** (normaalijakaumaan verrattuna).

Huomautus: *Normaalijakaumalle* $\gamma_2 = 0$.

Kvantiilit

Olkoon X satunnaismuuttuja. Olkoon lisäksi

$$0 < p < 1$$

Jos luku x_p toteuttaa ehdot

$$\Pr(X \leq x_p) \geq p$$

$$\Pr(X \geq x_p) \geq 1 - p$$

sanomme, että x_p on satunnaismuuttujan X ja sen jakauman **kvantiili** kertalukua p. Kvantiili x_p toteuttaa epäyhtälöt

$$\Pr(X < x_p) \leq p \leq \Pr(X \leq x_p)$$

Kvantiilit voidaan määrätä myös sellaisille satunnaismuuttujille, joilla *ei ole* momentteja.

Kvantiilit *eivät välttämättä ole yksikäsitteisiä*:

(i) *Diskreettien* satunnaismuuttujien kvantiilit *ovat usein monikäsitteisiä*.

(ii) *Jatkuvien* satunnaismuuttujien kvantiilit *ovat yksikäsitteisiä*.

Olkoon

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

jatkuvan satunnaismuuttujan X kertymäfunktio. Tällöin satunnaismuuttujan X kvantiili x_p toteuttaa yhtälön

$$F(x_p) = p$$

Kvantiili x_p jakaa satunnaismuuttujan X jakauman todennäköisyysmassan *kahteen osaan* niin, että massasta

$$p \times 100\ \%$$

on kvantiilista x_p *vasemmalla* ja

$$(1 - p) \times 100\ \%$$

on kvantiilista x_p *oikealla*.

Kvantiilit ja tilastolliset taulukot

Useimmissa *todennäköisyyslaskennan ja tilastotieteen oppikirjoissa on taulukoituna* keskeisten *tilastollisessa päättelyssä* käytettävien jatkuvien jakaumien (so. *normaalijakauman*, χ^2*-jakauman*, *t-jakauman* ja *F-jakauman*) kvantiileja x_p ja niitä vastaavia todennäköisyyksiä p ja useimmissa *tilastollisissa tietokoneohjelmissa on aliohjelmia*, jotka laskevat tavallisimpien jatkuvien jakaumien kvantiileja x_p ja niitä vastaavia todennäköisyyksiä p.

Prosenttipisteet

Jos p on muotoa

$$p = q/100\ ,\ q = 1, 2, \ldots , 99$$

kvantiilia x_p kutsutaan ***q*. prosenttipisteeksi**.

Jatkuvan satunnaismuuttujan tapauksessa q. prosenttipiste jakaa jakauman todennäköisyysmassan *kahteen osaan* niin, että massasta

$$q\ \%$$

on q. prosenttipisteestä *vasemmalla* ja

$$(100 - q)\ \%$$

on q. prosenttipisteestä *oikealla*.

Desiilit

Jos p on muotoa

$$p = 10 \times q/100\ ,\ q = 1, 2, \ldots , 9$$

kvantiilia x_p kutsutaan ***q*. desiiliksi**.

Jatkuvan satunnaismuuttujan tapauksessa q. desiili jakaa jakauman todennäköisyysmassan *kahteen osaan* niin, että massasta

$$10 \times q\ \%$$

on q. desiilistä *vasemmalla* ja

$$(100 - 10 \times q)\ \%$$

on q. desiilistä *oikealla*.

Kvartilit

Jos p on muotoa

$$p = 25 \times q/100 \text{ , } q = 1, 2, 3$$

kvantiilia x_p kutsutaan ***q*. kvartiiliksi**.

Jatkuvan satunnaismuuttujan tapauksessa q. kvartiili jakaa jakauman todennäköisyysmassan *kahteen osaan* niin, että massasta

$$25 \times q \text{ \%}$$

on q. kvartiilista *vasemmalla* ja

$$(100 - 25 \times q) \text{ \%}$$

on q. kvartiilista *oikealla*.

Kvartiileja merkitään tavallisesti symboleilla Q_1, Q_2, Q_3 ja sanotaan, että

Q_1 = **alakvartiili**

Q_2 = **keskikvartiili**

Q_3 = **yläkvartiili**

Jatkuvan satunnaismuuttujan tapauksessa kvartiilit jakavat jakauman todennäköisyysmassan *neljään yhtä suureen osaan*:

25 % massasta on kvartiilista Q_1 *vasemmalle*

25 % massasta on kvartiilien Q_1 ja Q_2 *välissä*

25 % massasta on kvartiilien Q_2 ja Q_3 *välissä*

25 % massasta on kvartiilista Q_3 *oikealle*

Mediaani

Jos

$$p = 0.5$$

kvantiilia x_p kutsutaan **mediaaniksi**. Mediaania merkitään tavallisesti symbolilla Me.

Jatkuvan satunnaismuuttujan tapauksessa mediaani Me jakaa jakauman todennäköisyysmassan *kahteen yhtä suureen osaan* niin, että massasta

50 %

on mediaanista *vasemmalla* ja

50 %

on mediaanista *oikealla*.

Jakauman mediaani *ei välttämättä ole* yksikäsitteinen. Jakauman mediaani yhtyy jakauman 50. *prosenttipisteeseen*, 5. *desiiliin* ja *keskikvartiiliin* Q_2. Mediaani voidaan määrätä myös sellaisille satunnaismuuttujille, joilla *ei ole* odotusarvoa.

Jos satunnaismuuttujan X jakauma on *symmetrinen* suoran $x = a$ suhteen, niin jakauman mediaani yhtyy pisteeseen a:

$$Me = a$$

Jos symmetrisellä jakaumalla *on odotusarvo* $\mathrm{E}(X) = \mu$, niin jakauman mediaani yhtyy pisteeseen μ:

$$Me = \mu$$

Moodi

Olkoon X *diskreetti* satunnaismuuttuja, jonka *pistetodennäköisyysfunktio* on

$$f(x) = \Pr(X = x)$$

Piste Mo on *diskreetin satunnaismuuttujan X* ja sen jakauman **moodi**, jos pistetodennäköisyys-funktio $f(x)$ saavuttaa maksiminsa pisteessä $x = Mo$:

$$f(Mo) = \max_x f(x)$$

Olkoon X *jatkuva* satunnaismuuttuja, jonka *tiheysfunktio* on

$$f(x)$$

Piste Mo on *jatkuvan satunnaismuuttujan X* ja sen jakauman **moodi**, jos tiheysfunktio $f(x)$ saavuttaa maksiminsa pisteessä $x = Mo$:

$$f(Mo) = \max_x f(x)$$

Jakauman moodi *ei välttämättä ole* yksikäsitteinen. Moodi voidaan määrätä myös sellaisille satunnaismuuttujille, joilla *ei ole* odotusarvoa.

11. Diskreettejä jakaumia

Diskreetti tasainen jakauma

Olkoon X *diskreetti satunnaismuuttuja*, jonka mahdolliset arvot ovat

$$x_1, x_2, \dots , x_n$$

Oletetaan, että satunnaismuuttujan X mahdollisiin arvoihin $x_1, x_2, \dots , x_n$ liittyvät todennäköisyydet ovat yhtä suuria:

$$\Pr(X = x_i) = \frac{1}{n}, \, i = 1, 2, \dots, n$$

Tällöin satunnaismuuttuja X noudattaa **diskreettiä tasaista jakaumaa**, jonka *pistetodennäköisyysfunktio* on

$$f(x_i) = \Pr(X = x_i) = p_i = \frac{1}{n}, \, i = 1, 2, \dots, n$$

Diskreetin tasaisen jakauman tunnusluvut

Odotusarvo:

$$\mathrm{E}(X) = \mu_X = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

2. *momentti*:

$$\mathrm{E}(X^2) = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

Varianssi:

$$\mathrm{D}^2(X) = \mathrm{Var}(X) = \sigma_X^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2 = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$$

Standardipoikkeama:

$$\mathrm{D}(X) = \sigma_X = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

Bernoulli-jakauma

Olkoon A otosavaruuden S *tapahtuma* ja olkoon

$$\Pr(A) = p$$

Tällöin tapahtuman A *komplementtitapahtuman* (tapahtuma A ei satu) A^c todennäköisyys on

$$\Pr(A^c) = 1 - p = q$$

Määritellään *diskreetti satunnaismuuttuja* X seuraavalla tavalla:

$$X = \begin{cases} 1, & \text{jos tapahtuma } A \text{ sattuu} \\ 0, & \text{jos tapahtuma } A \text{ ei satu} \end{cases}$$

Tällöin satunnaismuuttujan X *jakauma* on

$$\Pr(X = 1) = p$$
$$\Pr(X = 0) = 1 - p = q$$

ja satunnaismuuttujan X *pistetodennäköisyysfunktio* on

$$f(x) = p^x q^{1-x} \text{ , } 0 < p < 1 \text{ , } q = 1 - p \text{ , } x = 0,1$$

Sanomme, että satunnaismuuttuja X noudattaa **Bernoulli-jakaumaa** parametrilla p ja käytämme tästä merkintää:

$$X \sim \text{Bernoulli}(p)$$

Bernoulli-kokeiden yhteys eräisiin diskreetteihin todennäköisyysjakaumiin

Toistetaan toisistaan riippumatta samaa Bernoulli-koetta ja tarkastellaan tapahtuman A sattumista toistojen aikana.

(i) **Binomijakauma** saadaan määräämällä todennäköisyys sille, että tapahtuma A sattuu x kertaa, kun koetta toistetaan n kertaa.

(ii) **Geometrinen jakauma** saadaan määräämällä todennäköisyys sille, että tapahtuma A sattuu ensimmäisen kerran x. koetoistossa.

(iii) **Negatiivinen binomijakauma** saadaan määräämällä todennäköisyys sille, että tapahtuma A sattuu r. kerran x. koetoistossa.

(iv) **Poisson-jakauma** voidaan johtaa binomijakauman raja-arvona, kun koetoistojen lukumäärän annetaan tiettyjen ehtojen vallitessa kasvaa rajatta. Siten Poisson-jakauma kuvaa *harvinaisten* tapahtumien todennäköisyyksiä *pitkissä* toistokoesarjoissa.

Binomijakauma

Olkoon A otosavaruuden S *tapahtuma* ja olkoon

$$\Pr(A) = p$$

Tällöin tapahtuman *A komplementtitapahtuman* (tapahtuma A ei satu) A^c todennäköisyys on

$$\Pr(A^c) = 1 - p = q$$

Toistetaan sitä satunnaiskoetta, jonka tulosvaihtoehtoja otosavaruus S kuvaa, n kertaa, jossa n on kiinteä (ei-satunnainen), ennen koetoistojen tekemistä päätetty luku. Oletetaan lisäksi, että koetoistot ovat toisistaan riippumattomia.

Olkoon X *diskreetti satunnaismuuttuja*, joka kuvaa tapahtuman A esiintymiskertojen lukumäärää koetoistojen joukossa.

Tällöin satunnaismuuttuja X noudattaa **binomijakaumaa** parametrein n ja p:

$$X \sim \text{Bin}(n, p)$$

ja sen *pistetodennäköisyysfunktio* on

$$f(x) = \Pr(X = x) = \binom{n}{x} p^x q^{n-x} \text{ , } 0 < p < 1 \text{ , } q = 1 - p \text{ , } x = 0,1,2,\ldots,n$$

Binomijakauman tunnusluvut

Odotusarvo:

$$\mathrm{E}(X) = \mu_X = np$$

Varianssi:

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

Standardipoikkeama:

$$\mathrm{D}(X) = \sigma_X = \sqrt{npq}$$

Binomijakaumaa noudattavien satunnaismuuttujien summan jakauma

Olkoot $X_1, X_2, \ldots, X_n$ ovat *riippumattomia* satunnaismuuttujia, jotka noudattavat *binomijakaumia* parametrein (n_1, p), (n_2, p), … , (n_k, p):

$$X_1, X_2, \ldots, X_k \perp$$
$$X_i \sim \mathrm{Bin}(n_i, p)\,,\ i = 1, 2, \ldots, k$$

Tällöin *diskreetti satunnaismuuttuja*

$$X = \sum_{i=1}^{k} X_i$$

noudattaa *binomijakaumaa* parametrein $n = n_1 + n_2 + \cdots + n_k$ ja p:

$$X \sim \mathrm{Bin}(n, p)$$

Binomijakauman ja Bernoulli-jakauman yhteys

Olkoot $X_1, X_2, \ldots, X_n$ ovat *riippumattomia*, samaa *Bernoulli-jakaumaa* Bernoulli(p) noudattavia *diskreettejä satunnaismuuttujia*:

$$X_1, X_2, \ldots, X_n \perp$$
$$X_i \sim \mathrm{Bernoulli}(p)\,,\ i = 1, 2, \ldots, n$$

Tällöin *diskreetti satunnaismuuttuja*

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

noudattaa *binomijakaumaa* parametrein n ja p:

$$X \sim \mathrm{Bin}(n, p)$$

Geometrinen jakauma

Olkoon A otosavaruuden S *tapahtuma* ja olkoon

$$\Pr(A) = p$$

Tällöin tapahtuman A *komplementtitapahtuman* (tapahtuma A ei satu) A^c todennäköisyys on

$$\Pr(A^c) = 1 - p = q$$

Toistetaan sitä satunnaiskoetta, jonka tulosvaihtoehtoja otosavaruus *S* kuvaa kunnes tapahtuma *A* havaitaan 1. kerran. Oletetaan lisäksi, että koetoistot ovat toisistaan riippumattomia.

Olkoon *X diskreetti satunnaismuuttuja*, joka kuvaa tehtyjen koetoistojen lukumäärää, kun tapahtuma *A* havaitaan 1. kerran.

Tällöin satunnaismuuttuja *X* noudattaa **geometrista jakaumaa** parametrilla *p*:

$$X \sim \mathrm{Geom}(p)$$

ja sen pistetodennäköisyysfunktio on

$$f(x) = \Pr(X = x) = q^{x-1}p \text{ , } 0 < p < 1 \text{ , } q = 1 - p \text{ , } x = 1, 2, 3, \ldots$$

Satunnaismuuttujan *X kertymäfunktio* on

$$F(x) = \Pr(X \le x) = 1 - (1-p)^{[x]}$$

jossa

$[x]$ = suurin kokonaisluku, joka $\le x$

Komplementtitapahtuman todennäköisyyden kaavan mukaan

$$\Pr(X > x) = 1 - \Pr(X \le x) = 1 - F(x) = (1-p)^{[x]}$$

Geometrisen jakauman tunnusluvut

Odotusarvo:

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

Varianssi:

$$\mathrm{D}^2(X) = \mathrm{Var}(X) = \sigma_X^2 = \frac{q}{p^2}$$

Standardipoikkeama:

$$\mathrm{D}(X) = \sigma_X = \frac{\sqrt{q}}{p}$$

Negatiivinen binomijakauma

Olkoon *A* otosavaruuden *S tapahtuma* ja olkoon

$$\Pr(A) = p$$

Tällöin tapahtuman *A komplementtitapahtuman* (tapahtuma *A* ei satu) A^c todennäköisyys on

$$\Pr(A^c) = 1 - p = q$$

Toistetaan sitä satunnaiskoetta, jonka tulosvaihtoehtoja otosavaruus *S* kuvaa kunnes tapahtuma *A* havaitaan *r*. kerran. Oletetaan lisäksi, että koetoistot ovat toisistaan riippumattomia.

Olkoon *X diskreetti satunnaismuuttuja*, joka kuvaa tehtyjen koetoistojen lukumäärää, kun tapahtuma *A* havaitaan *r*. kerran.

Tällöin satunnaismuuttuja X noudattaa **negatiivista binomijakaumaa** parametrein r ja p:

$$X \sim \text{NegBin}(r, p)$$

ja sen *pistetodennäköisyysfunktio* on

$$f(x) = \Pr(X = x) = \binom{x-1}{r-1} q^{x-r} p^r \ , 0 < p < 1 \, , q = 1 - p$$

$$r = 1, 2, 3, \ldots \, ; \, x = r, r+1, r+2, \ldots$$

Negatiivisen binomijakauman tunnusluvut

Odotusarvo:

$$\text{E}(X) = \mu_X = \frac{r}{p}$$

Varianssi:

$$\text{D}^2(X) = \text{Var}(X) = \sigma_X^2 = \frac{rq}{p^2}$$

Standardipoikkeama:

$$\text{D}(X) = \sigma_X = \frac{\sqrt{rq}}{p}$$

Geometrinen jakauma negatiivisen binomijakauman erikoistapauksena

Olkoon

$$X \sim \text{NegBin}(r, p)$$

jossa

$$r = 1$$

Tällöin X noudattaa *geometrista jakaumaa* parametrilla p:

$$X \sim \text{Geom}(p)$$

Hypergeometrinen jakauma

Olkoon perusjoukon S alkioiden lukumäärä

$$n(S) = N$$

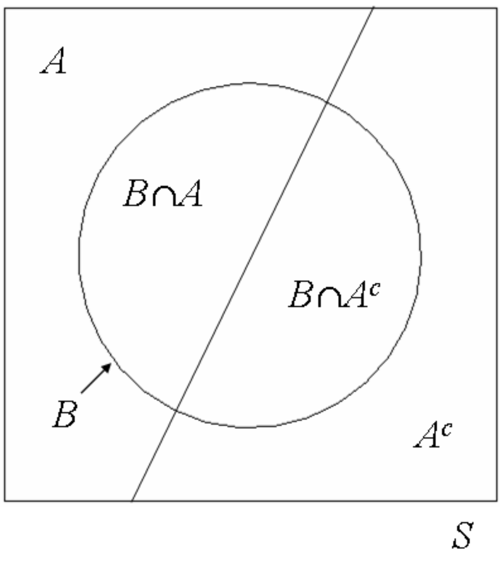
Tarkastellaan perusjoukon S ositusta joukkoon A ja sen komplementtiin A^c ja olkoon

$$n(A) = r$$

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

Valitaan perusjoukosta S satunnaisesti osajoukko B ja olkoon

$$n(B) = n$$



Perusjoukon S ositus joukoiksi A ja A^c indusoi joukon B osituksen joukoiksi $B \cap A$ ja $B \cap A^c$.

Olkoon X *diskreetti satunnaismuuttuja*, joka kuvaa joukossa B olevien joukon A (eli joukon $B \cap A$) alkioiden lukumäärää.

Tällöin satunnaismuuttuja X noudattaa **hypergeometrista jakaumaa** parametrein N, r ja n:

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

ja sen *pistetodennäköisyysfunktio* on

$$f(x) = \Pr(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}, \max[0, n-(N-r)] \leq x \leq \min(n, r)$$

Hypergeometrisen jakauman tunnusluvut

Odotusarvo:

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

Varianssi:

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

Standardipoikkeama:

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

Hypergeometrisen jakauman ja binomijakauman yhteys

Hypergeometrista jakaumaa voidaan *approksimoida* binomijakaumalla, jos *otantasuhde*

$$n/N$$

jossa

$n \;= n(B)\; =$ otoskoko

$N \;= n(S)\; =$ perusjoukon koko

on kyllin pieni. Näin on käytännössä, jos

$$n/N < 0.05$$

Huomaa, että jos perusjoukon S koko N lähestyy *ääretöntä*, otantasuhde konvergoi nollaa kohden ja siten hypergeometrinen jakauma lähestyy binomijakaumaa.

Otanta takaisinpanolla ja ilman takaisinpanoa

Poimitaan perusjoukosta *satunnaisesti otos* (osajoukko) arpomalla alkiot perusjoukosta otokseen yksi kerrallaan.

Otoksen poiminta voidaan toteuttaa joko **takaisinpanolla** tai **ilman takaisinpanoa**:

(i) *Otannassa takaisinpanolla* eli *palauttaen* perusjoukon alkiot *arvotaan* otokseen yksi kerrallaan niin, että alkiot *palautetaan* välittömästi jokaisen arpomisen jälkeen takaisin perusjoukkoon, jolloin sama alkio voi tulla poimituksi otokseen *useita kertoja*.

(ii) *Otannassa ilman takaisinpanoa* eli *palauttamatta* alkiot *arvotaan* otokseen yksi kerrallaan niin, että alkioita *ei palauteta* arpomisen jälkeen takaisin perusjoukkoon, jolloin sama alkio voi tulla poimituksi otokseen *vain kerran*.

Olkoon perusjoukon S koko

$$N = n(S)$$

Tarkastellaan perusjoukon S osajoukkoa A, jonka koko on

$$r = n(A)$$

Poimitaan perusjoukosta S satunnaisesti osajoukko B, jonka koko on

$$n = n(B)$$

Määritellään *diskreetti satunnaismuuttuja*

X = ”A-tyyppisten alkioiden lukumäärä otoksessa B”

Jos otos poimitaan perusjoukosta *palauttaen* eli *takaisinpanolla*, satunnaismuuttuja X noudattaa *binomijakaumaa* parametrein n ja p:

$$X \sim \mathrm{Bin}(n, p)$$

Jos otos poimitaan perusjoukosta *palauttamatta* eli *ilman takaisinpanoa*, satunnaismuuttuja X noudattaa *hypergeometrista jakaumaa* parametrein N, r ja n:

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

Poisson-jakauma

Toistetaan *samaa satunnaiskoetta* ja oletetaan, että toistot ovat toisistaan *riippumattomia*. Tarkastellaan jonkin *tapahtuman A* sattumista toistojen aikana. Oletetaan, että tapahtuman A *tapahtumaintensiteetti* eli *keskimääräinen lukumäärä aika-* (tai *tilavuus-*) *yksikköä kohden* on λ.

Määritellään *diskreetti satunnaismuuttuja X*:

X = Tapahtuman A esiintymisten lukumäärä aika- (tai tilavuus-) yksikköä kohden

Tietyin oletuksin satunnaismuuttuja X noudattaa **Poisson-jakaumaa** parametrilla λs:

$$X \sim \mathrm{Poisson}(\lambda s)$$

jossa

s = ajanjakson pituus aikayksiköissä

λ = tapahtuman A esiintymisten *keskimääräinen lukumäärä aika-* (tai *tilavuus-*) *yksikköä* kohden

ja sen *pistetodennäköisyysfunktio* on

$$f(x) = \Pr(X = x) = \frac{e^{-\lambda s}(\lambda s)^x}{x!}, \ x = 0,1,2,\ldots$$

Poisson-jakauman tunnusluvut

Odotusarvo:

$$\mathrm{E}(X) = \mu_X = \lambda s$$

Varianssi:

$$\mathrm{D}^2(X) = \mathrm{Var}(X) = \sigma_X^2 = \lambda s$$

Standardipoikkeama:

$$\mathrm{D}(X) = \sigma_X = \sqrt{\lambda s}$$

Poisson-jakaumaa noudattavien satunnaismuuttujien summan jakauma

Olkoot $X_1, X_2, \ldots, X_n$ ovat *riippumattomia* satunnaismuuttujia, jotka noudattavat *Poisson-jakaumia* parametrein $\lambda_1, \lambda_2, \ldots, \lambda_k$:

$$X_1, X_2, \ldots, X_k \perp$$
$$X_i \sim \mathrm{Poisson}(\lambda_i), \ i = 1,2,\ldots,k$$

Tällöin *diskreetti satunnaismuuttuja*

$$X = \sum_{i=1}^{k} X_i$$

noudattaa *Poisson-jakaumaa* parametrilla $\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_k$:

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

Poisson-jakauman ja binomijakauman yhteys

Olkoon

$$X_n \sim \mathrm{Bin}(n, p)$$

Annetaan

$$n \to \infty$$

ja

$$p \to 0$$

siten, että

$$np = \lambda$$

Tällöin *satunnaismuuttujan X_n jakauma lähestyy Poisson-jakaumaa parametrilla λ*:

$$\lim_{\substack{n \to \infty \\ p \to 0 \\ np = \lambda}} f_{\mathrm{Bin}(n,p)}(x) = f_{\mathrm{Poisson}(\lambda)}(x), \ x = 0,1,2,\ldots$$

Poisson-jakauman ja eksponenttijakauman yhteys

Olkoon

X = tapahtumien lukumäärä aikayksikköä kohden

Oletetaan, että

$X \sim \text{Poisson}(\lambda)$

ja olkoon

Y = odotusaika 1. tapahtumalle (tai tapahtumien väliaika)

Tällöin Y on *jatkuva satunnaismuuttuja*, joka noudattaa *eksponenttijakaumaa* parametrilla λ:

$Y \sim \text{Exp}(\lambda)$

jolloin

$\text{E}(Y) = 1/\lambda$

Voidaan osoittaa, että jakaumien välinen yhteys toimii molempiin suuntiin: ts. jos satunnaismuuttuja

X = odotusaika 1. tapahtumalle (tai tapahtumien väliaika)

noudattaa *eksponenttijakaumaa* parametrilla λ:

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

niin satunnaismuuttuja

Z = tapahtumien lukumäärä aikayksikköä kohden

noudattaa *Poisson-jakaumaa* parametrilla λ:

$Z \sim \text{Poisson}(\lambda)$

12. Jatkuvia jakaumia

Jatkuva tasainen jakauma

Olkoon *jatkuvan satunnaismuuttujan X tiheysfunktio*

$$f(x) = \begin{cases} 0\,,\ x < a \\ \dfrac{1}{b-a}\,,\ a \le x \le b \\ 0\,,\ x > a \end{cases}$$

Tällöin satunnaismuuttuja X noudattaa **jatkuvaa tasaista jakaumaa** parametrein a ja b.

Merkitään:

$$X \sim \mathrm{Uniform}(a,b)$$

Jatkuvan tasaisen jakauman tunnusluvut

Odotusarvo:

$$\mathrm{E}(X) = \frac{a+b}{2}$$

Varianssi ja standardipoikkeama:

$$\mathrm{D}^2(X) = \mathrm{Var}(X) = \frac{(b-a)^2}{12}$$

$$\mathrm{D}(X) = \frac{b-a}{2\sqrt{3}}$$

Jatkuvan tasaisen jakauman kertymäfunktio

Jatkuvan tasaisen jakauman *kertymäfunktio* on

$$F(x) = \Pr(X \le x) = \begin{cases} 0\,,\ x \le a \\ \dfrac{x-a}{b-a}\,,\ a \le x \le b \\ 1\,,\ x \ge b \end{cases}$$

Eksponenttijakauma

Olkoon *jatkuvan satunnaismuuttujan X tiheysfunktio*

$$f(x) = \lambda \exp(-\lambda x)\,,\ \lambda > 0\,,\ x \ge 0$$

Tällöin satunnaismuuttuja X noudattaa **eksponenttijakaumaa** parametrilla λ.

Merkitään:

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

Eksponenttijakauman tunnusluvut

Odotusarvo:

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

Varianssi ja standardipoikkeama:

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

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

Eksponenttijakauman kertymäfunktio

Eksponenttijakauman *kertymäfunktioksi* saadaan

$$\begin{aligned} F(x) = \int_0^x f(t)dt &= \int_0^x \lambda \exp(-\lambda t)dt \\ &= \left[-\exp(-\lambda t)\right]_0^x \\ &= 1 - \exp(-\lambda x)\,,\ \lambda > 0\,,\ x \geq 0 \end{aligned}$$

Siten

$$\Pr(X > x) = 1 - \mathrm{P}(X \leq x) = 1 - F(x) = \exp(-\lambda x)$$

jossa $F(x)$ on eksponenttijakauman *kertymäfunktio*.

Eksponenttijakauman ja Poisson-jakauman yhteys

Olkoon

X = odotusaika 1. tapahtumalle (tai tapahtumien väliaika)

Oletetaan, että

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

ja olkoon

Z = tapahtumien lukumäärä aikayksikköä kohden

Tällöin Z on *diskreetti satunnaismuuttuja*, joka noudattaa *Poisson-jakaumaa* parametrilla λ:

$Z \sim \mathrm{Poisson}(\lambda)$

jolloin

$\mathrm{E}(Z) = \lambda$

Voidaan osoittaa, että jakaumien välinen yhteys toimii molempiin suuntiin: ts. jos satunnaismuuttuja

Z = tapahtumien lukumäärä aikayksikköä kohden

noudattaa *Poisson-jakaumaa* parametrilla λ:

$Z \sim \mathrm{Poisson}(\lambda)$

niin satunnaismuuttuja

X = odotusaika 1. tapahtumalle (tai tapahtumien väliaika)

noudattaa *eksponenttijakaumaa* parametrilla λ:

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

Normaalijakauma

Olkoon *jatkuvan satunnaismuuttujan X tiheysfunktio*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}, -\infty < \mu < +\infty, \sigma > 0, -\infty < x < +\infty$$

Tällöin satunnaismuuttuja X noudattaa **normaalijakaumaa** parametrein μ ja σ^2.

Merkitään:

$$X \sim \mathrm{N}(\mu, \sigma^2)$$

Normaalijakauman tiheysfunktion ominaisuuksia

(i) Normaalijakauman tiheysfunktio $f(x)$ on kaikkialla *positiivinen*:

$$f(x) > 0, -\infty < x < +\infty$$

(ii) Tiheysfunktio on *yksihuippuinen*.

(iii) Tiheysfunktio saa *maksimiarvonsa* pisteessä μ.

(iv) Tiheysfunktio on *symmetrinen* pisteen $x = \mu$ suhteen:

$$f(\mu - x) = f(\mu + x), -\infty < x < +\infty$$

Normaalijakauman tunnusluvut

Odotusarvo:

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

Varianssi ja standardipoikkeama:

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

Standardoitu normaalijakauma

Jos

$$X \sim \mathrm{N}(0,1)$$

sanomme, että satunnaismuuttuja X noudattaa *standardoitua normaalijakaumaa*.

Standardointi

Jos

$$X \sim \mathrm{N}(\mu, \sigma^2)$$

niin

$$Z = \frac{X-\mu}{\sigma} \sim \mathrm{N}(0,1)$$

Normaalijakauma ja standardoitu normaalijakauma

Olkoon

$$X \sim \mathrm{N}(\mu,\sigma^2)$$

Tällöin

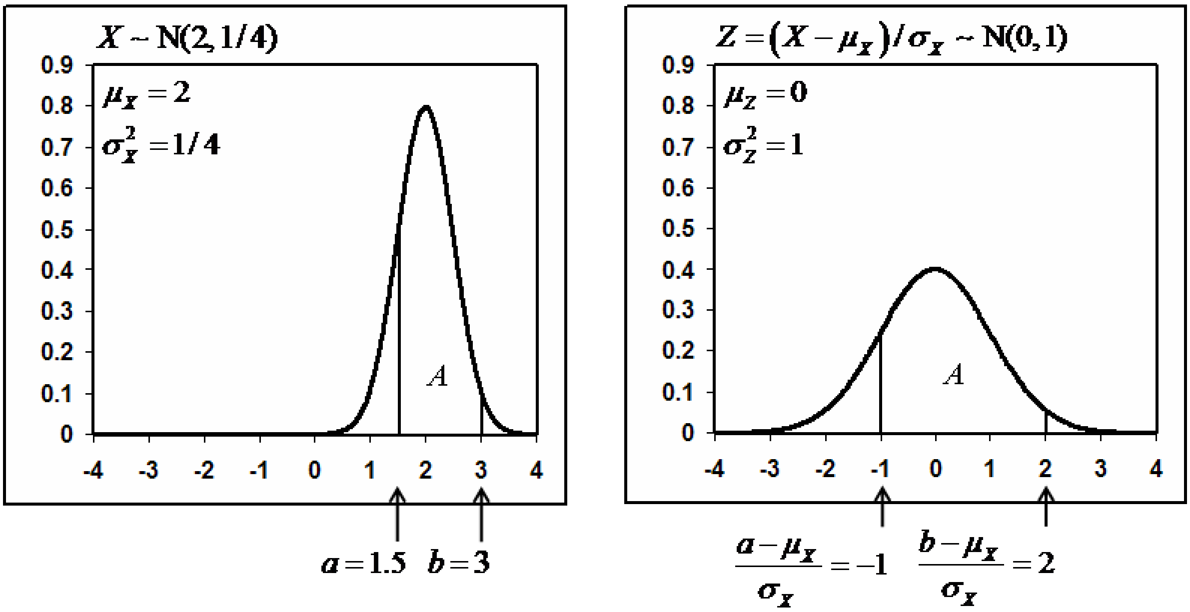
$$Z = \frac{X-\mu}{\sigma} \sim \mathrm{N}(0,1)$$

ja

$$\Pr(a \le X \le b) = \Pr\left(\frac{a-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right) = \Pr\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right)$$

Siten *normaalijakaumaan* $\mathrm{N}(\mu,\sigma^2)$ *liittyvät todennäköisyydet voidaan määrätä standardoidun normaalijakauman* $\mathrm{N}(0,1)$ *avulla*.

Esimerkki:



Standardoidun normaalijakauman N(0,1) taulukoista saadaan:

$$\begin{aligned}
&\text{Alueen } A \text{ pinta-ala} \\
&= \Pr(1.5 \le X \le 3) \\
&= \Pr\left(\frac{1.5-2}{1/2} \le \frac{X-2}{1/2} \le \frac{3-2}{1/2}\right) \\
&= \Pr(-1 \le Z \le 2) \qquad\qquad | \, Z \sim \mathrm{N}(0,1) \\
&= \Pr(Z \le 2) - \Pr(Z \le -1) \\
&= 0.9772 - 0.1587 = 0.8185
\end{aligned}$$

Normaalijakaumaa noudattavien satunnaismuuttujien summan jakauma

Olkoon

$$X_i\,,\ i = 1, 2, \ldots , n$$

jono *riippumattomia normaalijakautuneita satunnaismuuttujia*:

$$X_i \sim \mathrm{N}(\mu_i, \sigma_i^2)\,,\ i = 1, 2, \ldots, n$$
$$X_1, X_2, \ldots, X_n \perp$$

Olkoon

$$Y_n = \sum_{i=1}^{n} X_i$$

riippumattomien satunnaismuuttujien X_i , $i = 1, 2, \ldots , n$ *summa*. Tällöin

$$Y_n \sim \mathrm{N}(\mu_1 + \mu_2 + \cdots + \mu_n, \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2)$$

Oletetaan, että *riippumattomat* satunnaismuuttujat X_i , $i = 1, 2, \ldots , n$ noudattavat *samaa normaalijakaumaa*:

$$X_i \sim \mathrm{N}(\mu, \sigma^2)\,,\ i = 1, 2, \ldots, n$$
$$X_1, X_2, \ldots, X_n \perp$$

Tällöin

$$Y_n = \sum_{i=1}^{n} X_i \sim \mathrm{N}(n\mu, n\sigma^2)$$

Normaalijakaumaa noudattavien satunnaismuuttujien aritmeettisen keskiarvon jakauma

Olkoon

$$X_i\,,\ i = 1, 2, \ldots , n$$

jono *riippumattomia*, *samaa normaalijakaumaa noudattavia satunnaismuuttujia*:

$$X_i \sim \mathrm{N}(\mu, \sigma^2)\,,\ i = 1, 2, \ldots, n$$
$$X_1, X_2, \ldots, X_n \perp$$

Olkoon

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

riippumattomien satunnaismuuttujien X_i , $i = 1, 2, \ldots , n$ *aritmeettinen keskiarvo*. Tällöin

$$\bar{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

Keskeinen raja-arvolause

Olkoon

$$X_i\,,\ i = 1, 2, \ldots , n$$

jono *riippumattomia*, *samoin jakautuneita satunnaismuuttujia*, joiden odotusarvo ja varianssi ovat

$$\mathrm{E}(X_i) = \mu \text{ , } i = 1, 2, \ldots, n$$
$$\mathrm{D}^2(X_i) = \sigma^2 \text{ , } i = 1, 2, \ldots, n$$

Olkoon

$$Y_n = \sum_{i=1}^{n} X_i$$

riippumattomien satunnaismuuttujien X_i , $i = 1, 2, \ldots , n$ *summa*. Tällöin

$$\mathrm{E}(Y_n) = n\mu$$
$$\mathrm{D}^2(Y_n) = n\sigma^2$$

Standardoidaan satunnaismuuttuja Y_n:

$$Z_n = \frac{Y_n - n\mu}{\sigma\sqrt{n}}$$

Jos

$$n \to +\infty$$

niin *satunnaismuuttujan* Z_n *jakauma lähestyy rajatta standardoitua normaalijakaumaa* N(0,1):

$$\lim_{n\to+\infty} \Pr\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \le z \right) = \Phi(z)$$

jossa $\Phi(z)$ on *standardoidun normaalijakauman* N(0,1) *kertymäfunktio*.

Merkintä:

$$Z_n \sim_a \mathrm{N}(0,1)$$

Keskeinen raja-arvolause ja binomijakauma

Olkoon

$$X \sim \mathrm{Bin}(n, p)$$

ja

$$q = 1 - p$$

jolloin

$$\mathrm{E}(X) = np$$
$$\mathrm{D}^2(X) = npq$$

Koska binomijakautunut satunnaismuuttuja X voidaan esittää riippumattomien, samaa Bernoulli-jakaumaa Bernoulli(p) noudattavien satunnaismuuttujien summana, keskeisestä raja-arvolauseesta seuraa, että

$$\lim_{n\to+\infty} \Pr\left(\frac{X - np}{\sqrt{npq}} \le z \right) = \Phi(z)$$

jossa Φ on *standardoidun normaalijakauman* N(0,1) *kertymäfunktio*. Tätä keskeisen raja-arvolauseen erikoistapausta kutsutaan tavallisesti *De Moivren ja Laplacen raja-arvolauseeksi*.

Keskeinen raja-arvolause ja Poisson-jakauma

Olkoon

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

jolloin

$$\mathrm{E}(X) = \lambda$$
$$\mathrm{D}^2(X) = \lambda$$

Tällöin *keskeisestä raja-arvolauseesta* seuraa, että

$$\lim_{n\to+\infty} \Pr\left(\frac{X-\lambda}{\sqrt{\lambda}} \le z \right) = \Phi(z)$$

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

Log-normaalijakauma

Olkoon *jatkuvan satunnaismuuttujan X tiheysfunktio*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \frac{1}{x} \exp\left\{ -\frac{1}{2}\left(\frac{\log x - \mu}{\sigma} \right)^2 \right\}, -\infty < \mu < +\infty\,, \sigma > 0\,, 0 < x < +\infty$$

Tällöin satunnaismuuttuja X noudattaa **log-normaalijakaumaa** parametrein μ ja σ^2.

Merkitään:

$$X \sim \mathrm{LogN}(\mu, \sigma^2)$$

Log-normaalijakauman tunnusluvut

Odotusarvo:

$$\mathrm{E}(X) = e^{\mu+\sigma^2/2}$$

Varianssi ja standardipoikkeama:

$$\mathrm{D}^2(X) = \mathrm{Var}(X) = e^{2(\mu+\sigma^2)} - e^{2(\mu+\sigma^2/2)}$$
$$\mathrm{D}(X) = \sqrt{e^{\mu+\sigma^2} - e^{\mu+\sigma^2/2}}$$

Log-normaalijakauman ja normaalijakauman yhteys

Olkoon

$$\log X \sim \mathrm{N}(\mu, \sigma^2)$$

Tällöin

$$X \sim \mathrm{LogN}(\mu, \sigma^2)$$

Cauchy-jakauma

Olkoon *jatkuvan satunnaismuuttujan X tiheysfunktio*

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1+(x-\theta)^2}, -\infty < \theta < +\infty, -\infty < x < +\infty$$

Tällöin satunnaismuuttuja X noudattaa **Cauchy-jakaumaa** parametrilla θ.

Merkitään:

$$X \sim \text{Cauchy}(\theta)$$

Cauchy-jakauman tunnusluvut

Cauchy-jakaumalla *ei ole momentteja* ja siten sillä *ei ole odotusarvoa eikä varianssia*.

Parametri θ on Cauchy-jakauman *mediaani* ja myös *moodi*.

Gamma-jakauma

Olkoon *jatkuvan satunnaismuuttujan X tiheysfunktio*

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}, \alpha > 0, \beta > 0, x \geq 0$$

jossa

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt$$

on ns. *gamma-funktio*.

Tällöin satunnaismuuttuja X noudattaa **gamma-jakaumaa** parametrein α ja β.

Merkitään:

$$X \sim \text{Gamma}(\alpha, \beta)$$

Gamma-funktion ominaisuudet

Gamma-funktio määritellään kaavalla

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt$$

Gamma-funktiolla on seuraavat ominaisuudet:

(i) $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$

(ii) $\Gamma(n) = (n-1)!, n \in \mathbb{N}$

(iii) $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

Gamma-jakauman tunnusluvut

Odotusarvo:

$$\text{E}(X) = \alpha\beta$$

Varianssi ja standardipoikkeama:

$$\mathrm{D}^2(X) = \mathrm{Var}(X) = \alpha\beta^2$$
$$\mathrm{D}(X) = \beta\sqrt{\alpha}$$

Gamma-jakauman ja Poisson-jakauman yhteys

Olkoon

$$X \sim \mathrm{Gamma}(\alpha, \beta)$$

jossa α on *luonnollinen luku* eli

$$\alpha \in \mathbb{N}$$

Tällöin

$$\Pr(X \leq x) = \Pr(Y \geq \alpha)$$

jossa

$$Y \sim \mathrm{Poisson}(x / \beta)$$

Eksponenttijakauma gamma-jakauman erikoistapauksena

Olkoon

$$X \sim \mathrm{Gamma}(\alpha, \beta)$$

jossa

$$\alpha = 1$$

Tällöin X noudattaa *eksponenttijakaumaa* parametrilla $1/\beta$:

$$X \sim \mathrm{Exp}(1 / \beta)$$

χ^2-jakauma gamma-jakauman erikoistapauksena

Olkoon

$$X \sim \mathrm{Gamma}(\alpha, \beta)$$

jossa

$$\alpha = \frac{p}{2}$$
$$\beta = 2$$

Tällöin X noudattaa *χ^2-jakaumaa* (ks. lukua 11) vapausastein p:

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

Beta-jakauma

Olkoon *jatkuvan satunnaismuuttujan X tiheysfunktio*

$$f(x) = \frac{1}{\mathrm{B}(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1} \text{ , } \alpha > 0 \text{ , } \beta > 0 \text{ , } 0 \leq x \leq 1$$

jossa

$$\mathrm{B}(\alpha,\beta)=\int_0^\infty t^{\alpha-1}(1-t)^{\beta-1}\,dt$$

on ns. *beta-funktio*.

Tällöin satunnaismuuttuja X noudattaa **beta-jakaumaa** parametrein α ja β.

Merkitään:

$$X \sim \mathrm{Beta}(\alpha,\beta)$$

Beta-funktion ominaisuuksia

Beta-funktio määritellään kaavalla

$$\mathrm{B}(\alpha,\beta)=\int_0^\infty t^{\alpha-1}(1-t)^{\beta-1}\,dt$$

Beta-funktiolla ja gamma-funktiolla on seuraava yhteys:

$$\mathrm{B}(\alpha,\beta)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

jossa $\Gamma(\cdot)$ ns. *gamma-funktio*.

Beta-jakauman tunnusluvut

Odotusarvo:

$$\mathrm{E}(X)=\frac{\alpha}{\alpha+\beta}$$

Varianssi ja standardipoikkeama:

$$\mathrm{D}^2(X)=\mathrm{Var}(X)=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

$$\mathrm{D}(X)=\frac{1}{\alpha+\beta}\sqrt{\frac{\alpha\beta}{(\alpha+\beta+1)}}$$

Jatkuva tasainen jakauma beta-jakauman erikoistapauksena

Olkoon

$$X \sim \mathrm{Beta}(\alpha,\beta)$$

jossa

$$\alpha=\beta=1$$

Tällöin X noudattaa *jatkuvaa tasaista jakaumaa* välillä $(0,1)$:

$$X \sim \mathrm{Uniform}(0,1)$$

Weibull-jakauma

Olkoon *jatkuvan satunnaismuuttujan X tiheysfunktio*

$$f(x) = \frac{\gamma}{\beta} x^{\gamma-1} e^{-x^{\gamma}/\beta} \ , \gamma > 0 \ , \beta > 0 \ , 0 \le x < \infty$$

Tällöin satunnaismuuttuja X noudattaa **Weibull-jakaumaa** parametrein γ ja β.

Merkitään:

$$X \sim \mathrm{Weibull}(\gamma, \beta)$$

Weibull-jakauman tunnusluvut

Odotusarvo:

$$\mathrm{E}(X) = \beta^{1/\gamma}\Gamma(1+1/\gamma)$$

jossa $\Gamma(\cdot)$ ns. *gamma-funktio*.

Varianssi ja standardipoikkeama:

$$\mathrm{D}^2(X) = \mathrm{Var}(X) = \beta^{2/\gamma}\Gamma(1+2/\gamma) - [\mathrm{E}(X)]^2$$
$$\mathrm{D}(X) = \sqrt{\mathrm{Var}(X)}$$

jossa $\Gamma(\cdot)$ ns. *gamma-funktio*.

Weibull-jakauman ja eksponenttijakauman yhteys

Olkoon

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

Tällöin $X^{1/\gamma}$ noudattaa *Weibullin jakaumaa* parametrein γ ja β:

$$X \sim \mathrm{Weibull}(\gamma, \beta)$$

13. Normaalijakaumasta johdettuja jakaumia

χ^2-jakauma

Olkoot X_i , $i = 1, 2, \ldots , n$ *riippumattomia*, *standardoitua normaalijakaumaa* N(0,1) *noudattavia satunnaismuuttujia*:

$$X_i \sim \mathrm{N}(0,1)\,,\ i = 1,2,\ldots,n$$
$$X_1, X_2, \ldots, X_n \perp$$

Tällöin satunnaismuuttuja

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

noudattaa **χ^2-jakaumaa** vapausastein (eli parametrilla) n:

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

χ^2-jakauman *tiheysfunktio* on muotoa

$$f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}\,,\ 0 \le x < \infty$$

jossa $\Gamma(\cdot)$ ns. *gamma-funktio*.

χ^2-jakauman tunnusluvut

Oletetaan, että $X \sim \chi^2(n)$.

Odotusarvo:

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

Varianssi ja *standardipoikkeama*:

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

χ^2-jakauman taulukot

Koska χ^2-jakauman tiheysfunktion integraalifunktiota ei tunneta, joudutaan χ^2-jakaumaan liittyvät todennäköisyydet määräämään *numeerisesti*. Tämä tapahtuu käytännössä tietokoneohjelmien tai *taulukoiden* avulla.

Sovelletun todennäköisyyslaskennan kurssilla käytetyissä taulukoissa on taulukoituna usealle eri vapausasteiden arvolle $df = n$ pisteitä x, jotka erottavat todennäköisyysmassat

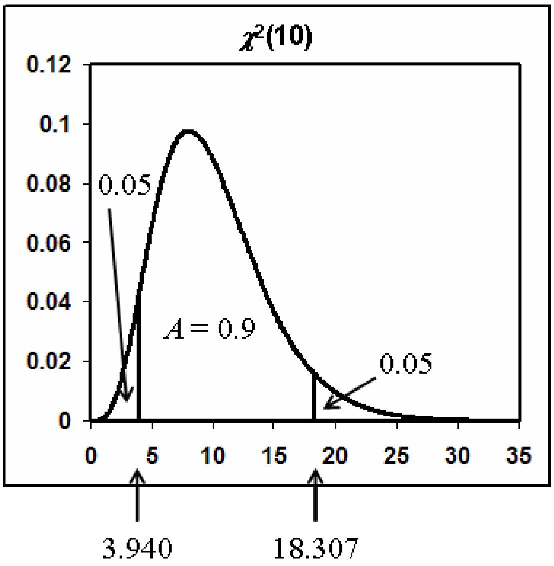
$$p_x = 0.999, 0.99, 0.95, 0.9, 0.1, 0.05, 0.01, 0.001$$

χ^2-jakauman *oikealle hännälle*.

Siten taulukoista löytyy pisteitä x, jotka toteuttavat ehdon

$$p_x = \Pr(X \geq x)$$

Koska eräs taulukoiden tärkeimmistä käyttöalueista on *tilastollinen testaus*, kutsutaan todennäköisyyttä p_x taulukoissa *merkitsevyystasoksi* ja pistettä x vastaavaksi *kriittiseksi rajaksi*.



F-jakauma

Olkoot Y_i , $i = 1, 2, \ldots , m$ ja X_i , $i = 1, 2, \ldots , n$ *riippumattomia*, *standardoitua normaalijakaumaa* N(0,1) *noudattavia satunnaismuuttujia*:

$$Y_i \sim \mathrm{N}(0,1)\,,\ i = 1,2,\ldots,m\ ;\ X_i \sim \mathrm{N}(0,1)\,,\ i = 1,2,\ldots,n$$
$$Y_1, Y_2, \ldots, Y_m, X_1, X_2, \ldots, X_n \perp$$

Tällöin satunnaismuuttujat

$$Y = \sum_{i=1}^{m} Y_i^2\ ;\ X = \sum_{i=1}^{n} X_i^2$$

ovat riippumattomia ja noudattavat χ^2-*jakaumaa* vapausastein m ja n:

$$Y \sim \chi^2(m)\ ;\ X \sim \chi^2(n)$$
$$Y \perp X$$

Määritellään satunnaismuuttuja

$$F = \frac{\frac{1}{m}Y}{\frac{1}{n}X} = \frac{n}{m}\cdot\frac{Y}{X}$$

Satunnaismuuttujat F noudattaa ***F*-jakaumaa** vapausastein (parametrein) m ja n:

$$F \sim F(m,n)$$

F-jakauman *tiheysfunktio* on muotoa

$$f(x) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\left(\frac{m}{n}\right)^{\frac{m}{2}} x^{\frac{m}{2}-1}\left(1+\frac{m}{n}x\right)^{-\frac{m+n}{2}}\ ,\ 0 \leq x < \infty$$

jossa $\Gamma(\cdot)$ ns. *gamma-funktio*.

F-jakauman tunnusluvut

Oletetaan, että $F \sim F(m,n)$.

Odotusarvo:

$$\mathrm{E}(X) = \frac{n}{n-2}\,,\ n > 2$$

Varianssi ja *standardipoikkeama*:

$$\mathrm{D}^2(F) = \mathrm{Var}(F) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}\,,\ n > 4$$

$$\mathrm{D}(F) = \sqrt{\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}}\,,\ n > 4$$

F-jakauman taulukot

Koska *F*-jakauman tiheysfunktion integraalifunktiota ei tunneta, joudutaan *F*-jakaumaan liittyvät todennäköisyydet määräämään *numeerisesti*. Tämä tapahtuu käytännössä tietokoneohjelmien tai *taulukoiden* avulla.

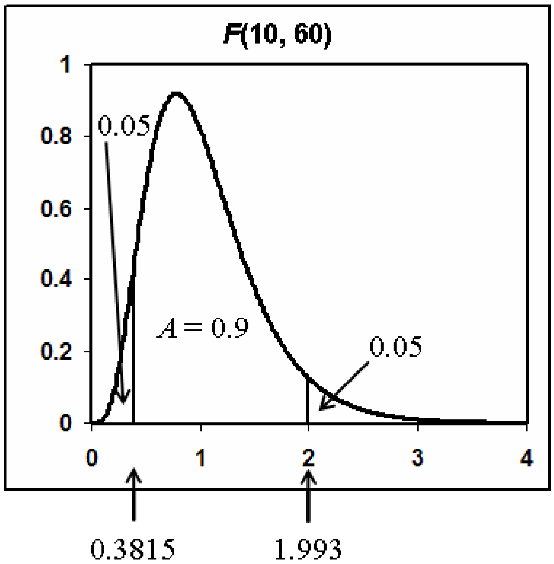
Sovelletun todennäköisyyslaskennan kurssilla käytetyissä taulukoissa on taulukoituna usealle eri vapausasteiden $df_1 = m$ ja $df_2 = n$ arvoille pisteitä x, jotka erottavat todennäköisyysmassat

$$p_x = 0.05,\ 0.025,\ 0.01,\ 0.005$$

F-jakauman *oikealle* hännälle.

Siten taulukoista löytyy pisteitä x, jotka toteuttavat ehdon

$$p_x = \Pr(F \geq x)$$



Koska eräs taulukoiden tärkeimmistä käyttöalueista on *tilastollinen testaus*, kutsutaan todennäköisyyttä p_x taulukoissa *merkitsevyystasoksi* ja pistettä x vastaavaksi *kriittiseksi rajaksi*.

Pisteet x, jotka erottavat todennäköisyysmassat

$$0.05,\ 0.025,\ 0.01,\ 0.005$$

F-jakauman *vasemmalle* hännälle saadaan käyttämällä hyväksi seuraavaa *F*-jakauman ominaisuutta:

Jos

$$G \sim F(n,m)$$

niin

$$F = \frac{1}{G} \sim F(m,n)$$

Määrätään siis piste y siten, että

$$\Pr(G \geq y) = p$$

jossa

$$G \sim F(n,m)$$

Tällöin

$$\Pr(G \geq y) = \Pr\left(\frac{1}{G} \leq \frac{1}{y}\right) = \Pr\left(F \leq \frac{1}{y}\right)$$

jossa

$$F \sim F(m,n)$$

Siten piste

$$x = 1/y$$

toteuttaa ehdon

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

t-jakauma

Olkoot Y ja X_i , $i = 1, 2, \ldots , n$ *riippumattomia*, *standardoitua normaalijakaumaa* N(0,1) *noudattavia satunnaismuuttujia*:

$$Y \sim \mathrm{N}(0,1)\,;\; X_i \sim \mathrm{N}(0,1)\,,\; i = 1,2,\ldots,n$$
$$Y, X_1, X_2, \ldots, X_n \perp$$

Tällöin satunnaismuuttuja

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

noudattaa χ^2-*jakaumaa* vapausastein n:

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

Lisäksi satunnaismuuttujat Y ja X ovat riippumattomia:

$$Y \perp X$$

Määritellään satunnaismuuttuja

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

Satunnaismuuttuja t noudattaa ***t*-jakaumaa** vapausastein (parametrilla) n:

$$t \sim t(n)$$

t-jakauman *tiheysfunktio* on muotoa

$$f(x)=\frac{1}{\sqrt{n\pi}}\cdot\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{1}{n}x^2\right)^{-\frac{n+1}{2}}\ ,\ -\infty<x<+\infty$$

jossa $\Gamma(\cdot)$ ns. *gamma-funktio*.

t-jakauman tunnusluvut

Oletetaan, että $t\sim t(n)$.

Odotusarvo:

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

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-jakauman ja F-jakauman yhteys

Jos

$$t\sim t(n)$$

niin

$$t^2\sim F(1,n)$$

Kääntäen, jos

$$F\sim F(1,n)$$

niin

$$\sqrt{F}\sim t(n)$$

Cauchy-jakauma t-jakauman erikoistapauksena

Olkoon

$$t\sim t(n)$$

jossa

$$n=1$$

Tällöin t noudattaa *Cauchy-jakaumaa* parametrilla 0:

$$t\sim \mathrm{Cauchy}(0)$$

t-jakauman taulukot

Koska t-jakauman tiheysfunktion integraalifunktiota ei tunneta, joudutaan t-jakaumaan liittyvät todennäköisyydet määräämään *numeerisesti*. Tämä tapahtuu käytännössä tietokoneohjelmien tai *taulukoiden* avulla.

Sovelletun todennäköisyyslaskennan kurssilla käytetyissä taulukoissa on taulukoituna usealle eri vapausasteiden $df = n$ arvoille pisteitä x, jotka erottavat todennäköisyysmassat

$$p_x = 0.4, 0.3, 0.2, 0.1, 0.05, 0.025, 0.01, 0.005, 0.001, 0.0005$$

t-jakauman *oikealle* hännälle.

Siten taulukoista löytyy pisteitä x, jotka toteuttavat ehdon

$$p_x = \Pr(t \geq x)$$

Koska eräs taulukoiden tärkeimmistä käyttöalueista on *tilastollinen testaus*, kutsutaan todennäköisyyttä p_x taulukoissa *merkitsevyystasoksi* ja pistettä x vastaavaksi *kriittiseksi rajaksi*.

Pisteet x, jotka erottavat todennäköisyysmassat

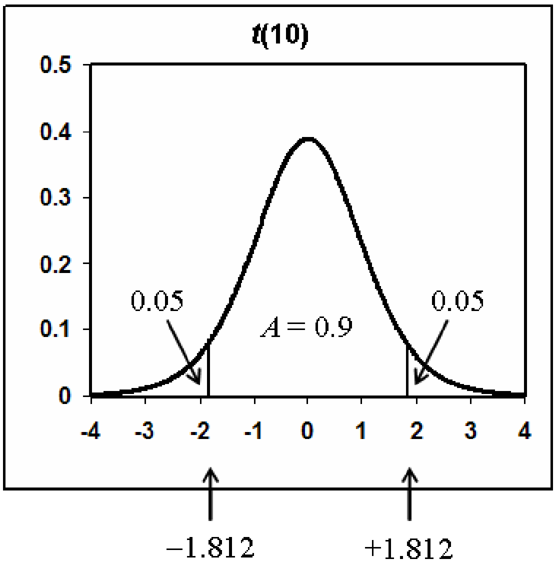
0.4, 0.3, 0.2, 0.1, 0.05,

0.025, 0.01, 0.005, 0.001, 0.0005

t-jakauman *vasemmalle* hännälle saadaan käyttämällä hyväksi sitä, että t-jakauma on *symmetrinen* origon suhteen.

Siten

$$p_x = \Pr(t \geq x) = \Pr(t \leq -x)$$



14. Moniulotteiset satunnaismuuttujat ja todennäköisyysjakaumat

Kaksiulotteiset satunnaismuuttujat

Olkoot X ja Y satunnaismuuttujia, joiden otosavaruudet ovat R ja S. Tällöin

$$X : S \to \mathbb{R}$$
$$Y : R \to \mathbb{R}$$

Olkoon $R \times S$ otosavaruuksien R ja S *karteesinen tulo*:

$$R \times S = \{(r,s) \,|\, r \in R, s \in S\}$$

Satunnaismuuttujien X ja Y *järjestetty pari* (X, Y) määrittelee **kaksiulotteisen satunnaismuuttujan**:

$$(X,Y) : S \times R \to \mathbb{R}^2$$

Diskreetti kaksiulotteinen jakauma

Olkoot X ja Y *diskreettejä* satunnaismuuttujia. Tällöin järjestetty pari

(X, Y)

määrittelee **diskreetin kaksiulotteisen satunnaismuuttujan**. Diskreetti kaksiulotteinen satunnaismuuttuja (X, Y) määrittelee *diskreetin kaksiulotteisen todennäköisyysjakauman*, jota kutsutaan satunnaismuuttujien X ja Y **yhteisjakaumaksi**.

Reaaliarvoinen funktio

$$f_{XY} : \mathbb{R}^2 \to \mathbb{R}$$

määrittelee *diskreettien satunnaismuuttujien* X ja Y **yhteisjakauman pistetodennäköisyysfunktion**, jos seuraavat 3 ehtoa pätevät:

(1) $\quad f_{XY}(x,y) \ge 0$ kaikille x ja y

(2) $\quad \sum_x \sum_y f_{XY}(x,y) = 1$

(3) $\quad \Pr(X = x \text{ ja } Y = y) = f_{XY}(x,y)$

Jatkuva kaksiulotteinen jakauma

Olkoot X ja Y *jatkuvia* satunnaismuuttujia. Tällöin järjestetty pari

(X, Y)

määrittelee **jatkuvan kaksiulotteisen satunnaismuuttujan**. Jatkuva kaksiulotteinen satunnaismuuttuja (X, Y) määrittelee *jatkuvan kaksiulotteisen todennäköisyysjakauman*, jota kutsutaan satunnais-muuttujien X ja Y **yhteisjakaumaksi**.

Reaaliarvoinen funktio

$$f_{XY} : \mathbb{R}^2 \to \mathbb{R}$$

määrittelee *jatkuvien satunnaismuuttujien* X ja Y **yhteisjakauman tiheysfunktion**, jos seuraavat 3 ehtoa pätevät:

(1) $f_{XY}(x,y) \geq 0$ kaikille x ja y

(2) $$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{XY}(x,y)dydx = 1$$

(3) $$\Pr(a \leq X \leq b \text{ ja } c \leq Y \leq d) = \int_a^b\int_c^d f_{XY}(x,y)dydx$$

Kaksiulotteisen jakauman kertymäfunktio

Olkoon (X,Y) satunnaismuuttujien X ja Y muodostama *järjestetty pari*. Satunnaismuuttujien X ja Y **yhteisjakauman kertymäfunktio** F_{XY} määritellään kaavalla

$$F_{XY}(x,y) = \Pr(X \leq x \text{ ja } Y \leq y)$$

Diskreetin kaksiulotteisen jakauman kertymäfunktio

Olkoon $f_{XY}(x,y)$ *diskreetin kaksiulotteisen jakauman pistetodennäköisyysfunktio*. *Jakauman kertymäfunktio* saadaan kaavalla

$$F_{XY}(x,y) = \Pr(X \leq x \text{ ja } Y \leq y) = \sum_{x_i \leq x}\sum_{y_i \leq y} f_{XY}(x_i, y_i)$$

Jatkuvan kaksiulotteisen jakauman kertymäfunktio

Olkoon $f_{XY}(x,y)$ *jatkuvan kaksiulotteisen jakauman tiheysfunktio*. *Jakauman kertymäfunktio* saadaan kaavalla

$$F_{XY}(x,y) = \Pr(X \leq x \text{ ja } Y \leq y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(u,v)dvdu$$

Olkoon $F_{XY}(x,y)$ *jatkuvan kaksiulotteisen jakauman kertymäfunktio*. Jos *derivaatta*

$$\frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y} = f_{XY}(x,y)$$

on olemassa ja on jatkuva, funktio $f_{XY}(x,y)$ on satunnaismuuttujien X ja Y yhteisjakauman *tiheysfunktio*.

Diskreetin kaksiulotteisen jakauman reunajakaumat

Olkoon $f_{XY}(x,y)$ *diskreetin kaksiulotteisen jakauman pistetodennäköisyysfunktio*.

Satunnaismuuttujan X **reunajakauman pistetodennäköisyysfunktio** on

$$f_X(x) = \Pr(X = x) = \sum_y f_{XY}(x,y)$$

Satunnaismuuttujan Y **reunajakauman pistetodennäköisyysfunktio** on

$$f_Y(y) = \Pr(Y = y) = \sum_x f_{XY}(x,y)$$

Satunnaismuuttujien X ja Y reunajakaumat yhtyvät satunnaismuuttujien X ja Y todennäköisyysjakaumiin.

Jatkuvan kaksiulotteisen jakauman reunajakaumat

Olkoon $f_{XY}(x,y)$ *jatkuvan kaksiulotteisen jakauman tiheysfunktio*.

Satunnaismuuttujan X **reunajakauman tiheysfunktio** on

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x,y)dy$$

Satunnaismuuttujan Y **reunajakauman tiheysfunktio** on

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x,y)dx$$

Satunnaismuuttujien X ja Y reunajakaumat yhtyvät satunnaismuuttujien X ja Y todennäköisyysjakaumiin.

Satunnaismuuttujien riippumattomuus

Olkoon $f_{XY}(x,y)$ satunnaismuuttujien X ja Y *yhteisjakauman pistetodennäköisyys*- tai *tiheysfunktio*, $f_X(x)$ satunnaismuuttujan X *reunajakauman pistetodennäköisyys*- tai *tiheysfunktio* ja $f_Y(y)$ satunnaismuuttujan Y *reunajakauman pistetodennäköisyys*- tai *tiheysfunktio*.

Satunnaismuuttujat X ja Y ovat **riippumattomia**, jos ja vain, jos

$$f_{XY}(x,y) = f_X(x)f_Y(y)$$

Satunnaismuuttujien riippumattomuus voidaan määritellä yhtä hyvin niiden kertymäfunktion avulla:

Olkoon $F_{XY}(x,y)$ satunnaismuuttujien X ja Y *yhteisjakauman kertymäfunktio*, $F_X(x)$ satunnaismuuttujan X *reunajakauman kertymäfunktio* ja $F_Y(y)$ satunnaismuuttujan Y *reunajakauman kertymäfunktio*.

Satunnaismuuttujat X ja Y ovat **riippumattomia**, jos ja vain, jos

$$F_{XY}(x,y) = F_X(x)F_Y(y)$$

Diskreetin kaksiulotteisen jakauman yleinen odotusarvo

Olkoon $f_{XY}(x,y)$ diskreettien satunnaismuuttujien X ja Y *yhteisjakauman pistetodennäköisyysfunktio* ja olkoon

$$g : \mathbb{R}^2 \to \mathbb{R}$$

jatkuva funktio. Tällöin satunnaismuuttujan $g(X,Y)$ **odotusarvo** on *vakio*

$$\mathrm{E}(g(X,Y)) = \sum_x \sum_y g(x,y) f_{XY}(x,y)$$

Jatkuvan kaksiulotteisen jakauman yleinen odotusarvo

Olkoon $f_{XY}(x,y)$ jatkuvien satunnaismuuttujien X ja Y *yhteisjakauman tiheysfunktio* ja olkoon

$$g : \mathbb{R}^2 \to \mathbb{R}$$

jatkuva funktio. Tällöin satunnaismuuttujan $g(X,Y)$ **odotusarvo** on *vakio*

$$\mathrm{E}(g(X,Y)) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y) f_{XY}(x,y)\,dydx$$

Diskreetin kaksiulotteisen jakauman odotusarvot

Olkoon satunnaismuuttujien X ja Y *yhteisjakauman pistetodennäköisyysfunktio* $f_{XY}(x,y)$, satunnaismuuttujan X *reunajakauman pistetodennäköisyysfunktio* $f_X(x)$ ja satunnaismuuttujan Y *reunajakauman pistetodennäköisyysfunktio* $f_Y(y)$.

Satunnaismuuttujan X **odotusarvo** $\mathrm{E}(X) = \mu_X$ yhtyy satunnaismuuttujan X *reunajakauman odotusarvoon*:

$$\mathrm{E}(X) = \sum_x \sum_y x f_{XY}(x,y) = \sum_x x \sum_y f_{XY}(x,y) = \sum_x x f_X(x)$$

Satunnaismuuttujan Y **odotusarvo** $\mathrm{E}(Y) = \mu_Y$ yhtyy satunnaismuuttujan Y *reunajakauman odotusarvoon*:

$$\mathrm{E}(Y) = \sum_x \sum_y y f_{XY}(x,y) = \sum_y y \sum_x f_{XY}(x,y) = \sum_y y f_Y(y)$$

Jatkuvan kaksiulotteisen jakauman odotusarvot

Olkoon satunnaismuuttujien X ja Y *yhteisjakauman tiheysfunktio* $f_{XY}(x,y)$, satunnaismuuttujan X *reunajakauman tiheysfunktio* $f_X(x)$ ja satunnaismuuttujan Y *reunajakauman tiheysfunktio* $f_Y(y)$.

Satunnaismuuttujan X **odotusarvo** $\mathrm{E}(X) = \mu_X$ yhtyy satunnaismuuttujan X *reunajakauman odotusarvoon*:

$$\mathrm{E}(X) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x f_{XY}(x,y)\,dydx = \int_{-\infty}^{+\infty} x \int_{-\infty}^{+\infty} f_{XY}(x,y)\,dydx = \int_{-\infty}^{+\infty} x f_X(x)\,dx$$

Satunnaismuuttujan Y **odotusarvo** $\mathrm{E}(Y) = \mu_Y$ yhtyy satunnaismuuttujan Y *reunajakauman odotusarvoon*:

$$\mathrm{E}(Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y f_{XY}(x,y)\,dydx = \int_{-\infty}^{+\infty} y \int_{-\infty}^{+\infty} f_{XY}(x,y)\,dxdy = \int_{-\infty}^{+\infty} y f_Y(y)\,dy$$

Odotusarvon ominaisuudet

Satunnaismuuttujien X ja Y odotusarvojen muodostama järjestetty pari

$$(\mu_X, \mu_Y)$$

määrää satunnaismuuttujien X ja Y yhteisjakauman todennäköisyysmassan *painopisteen*.

Satunnaismuuttujien X ja Y *summan* $X + Y$ *odotusarvo*:

$$\mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y)$$

Satunnaismuuttujien X ja Y *erotuksen* $X - Y$ *odotusarvo*:

$$\mathrm{E}(X-Y) = \mathrm{E}(X) - \mathrm{E}(Y)$$

Jos satunnaismuuttujat X ja Y ovat *riippumattomia*, niin *tulon* XY *odotusarvo on odotusarvojen tulo*:

$$\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y) = \mu_X \mu_Y$$

Huomautus:

Käänteinen ei päde: Siitä, että

$$\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y) = \mu_X \mu_Y$$

ei seuraa, että satunnaismuuttujat X ja Y ovat riippumattomia.

Kaksiulotteisen jakauman varianssit ja standardipoikkeamat

Olkoot satunnaismuuttujien X ja Y *odotusarvot*

$$\mathrm{E}(X) = \mu_X \qquad \mathrm{E}(Y) = \mu_Y$$

Satunnaismuuttujien X ja Y **varianssit** yhtyvät vastaavien *reunajakaumien variansseihin*:

$$\mathrm{Var}(X) = \mathrm{D}^2(X) = \sigma_X^2 = \mathrm{E}[(X-\mu_X)^2]$$
$$\mathrm{Var}(Y) = \mathrm{D}^2(Y) = \sigma_Y^2 = \mathrm{E}[(Y-\mu_Y)^2]$$

Satunnaismuuttujien X ja Y varianssien kaavat voidaan kirjoittaa seuraaviin yhtäpitäviin muotoihin:

$$\mathrm{D}^2(X) = \mathrm{E}[(X-\mu_X)^2] = \mathrm{E}(X^2) - \mu_X^2 = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$$
$$\mathrm{D}^2(Y) = \mathrm{E}[(Y-\mu_Y)^2] = \mathrm{E}(Y^2) - \mu_Y^2 = \mathrm{E}(Y^2) - [\mathrm{E}(Y)]^2$$

Satunnaismuuttujien X ja Y **standardipoikkeamat** yhtyvät vastaavien *reunajakaumien standardipoikkeamiin*:

$$\mathrm{D}(X) = \sigma_X = \sqrt{\mathrm{E}[(X-\mu_X)^2]}$$
$$\mathrm{D}(Y) = \sigma_Y = \sqrt{\mathrm{E}[(Y-\mu_Y)^2]}$$

Diskreetin kaksiulotteisen jakauman varianssit

Olkoon satunnaismuuttujan X *reunajakauman pistetodennäköisyysfunktio* $f_X(x)$ ja satunnaismuuttujan Y *reunajakauman pistetodennäköisyysfunktio* $f_Y(y)$.

Tällöin satunnaismuuttujien X ja Y **varianssit** ovat *vakioita*

$$\mathrm{D}^2(X) = \sum_x (x-\mu_X)^2 f_X(x)$$
$$\mathrm{D}^2(Y) = \sum_y (y-\mu_Y)^2 f_Y(y)$$

Jatkuvan kaksiulotteisen jakauman varianssit

Olkoon satunnaismuuttujan X *reunajakauman tiheysfunktio* $f_X(x)$ ja satunnaismuuttujan Y *reunajakauman tiheysfunktio* $f_Y(y)$.

Tällöin satunnaismuuttujien X ja Y **varianssit** ovat *vakioita*

$$\mathrm{D}^2(X) = \int_{-\infty}^{+\infty} (x-\mu_X)^2 f_X(x)dx$$
$$\mathrm{D}^2(Y) = \int_{-\infty}^{+\infty} (y-\mu_Y)^2 f_Y(y)dy$$

Kovarianssi

Olkoot satunnaismuuttujien X ja Y *odotusarvot*

$$\mathrm{E}(X) = \mu_X \qquad \mathrm{E}(Y) = \mu_Y$$

Satunnaismuuttujien X ja Y **kovarianssi** on *vakio*

$$\mathrm{Cov}(X,Y) = \sigma_{XY} = \mathrm{E}[(X-\mu_X)(Y-\mu_Y)]$$

Erityisesti

$$\mathrm{Cov}(X,X) = \mathrm{Var}(X) = \sigma_X^2$$
$$\mathrm{Cov}(Y,Y) = \mathrm{Var}(Y) = \sigma_Y^2$$

Satunnaismuuttujien X ja Y *kovarianssin* kaava voidaan kirjoittaa seuraaviin yhtäpitäviin muotoihin:

$$\begin{aligned}\mathrm{Cov}(X,Y) &= \mathrm{E}[(X-\mu_X)(Y-\mu_Y)] \\ &= \mathrm{E}(XY) - \mu_X\mu_Y \\ &= \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)\end{aligned}$$

Jos satunnaismuuttujat X ja Y ovat *diskreettejä*, satunnaismuuttujien X ja Y **kovarianssi** on *vakio*

$$\mathrm{Cov}(X,Y) = \sum_x \sum_y (x-\mu_X)(y-\mu_Y) f_{XY}(x,y)$$

Jos satunnaismuuttujat X ja Y ovat *jatkuvia*, satunnaismuuttujien X ja Y **kovarianssi** on *vakio*

$$\mathrm{Cov}(X,Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x-\mu_X)(y-\mu_Y) f_{XY}(x,y)dxdy$$

Kovarianssin ominaisuudet

Jos

$$\mathrm{Cov}(X,Y) = \sigma_{XY} = 0$$

niin sanomme, että satunnaismuuttujat X ja Y ovat **korreloimattomia**.

Jos satunnaismuuttujat X ja Y ovat riippumattomia, niin ne ovat korreloimattomia. Sen sijaan *käänteinen ei päde*: Siitä, että satunnaismuuttujat X ja Y ovat korreloimattomia, *ei seuraa*, että satunnaismuuttujat X ja Y olisivat riippumattomia.

Olkoot satunnaismuuttujien X ja Y *varianssit*

$$\mathrm{Var}(X) = \sigma_X^2$$
$$\mathrm{Var}(Y) = \sigma_Y^2$$

ja kovarianssi

$$\mathrm{Cov}(X,Y) = \sigma_{XY}$$

Tällöin

$$\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2\,\mathrm{Cov}(X,Y) = \sigma_X^2 + \sigma_Y^2 \pm 2\sigma_{XY}$$

Jos

$$\mathrm{Cov}(X,Y) = \sigma_{XY} = 0$$

niin

$$\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) = \sigma_X^2 + \sigma_Y^2$$

Korrelaatiokerroin

Olkoon satunnaismuuttujilla X ja Y on seuraavat *odotusarvot*, *varianssit* ja *kovarianssi*:

$$\mathrm{E}(X) = \mu_X \qquad \mathrm{Var}(X) = \mathrm{D}^2(X) = \sigma_X^2$$

$$\mathrm{E}(Y) = \mu_Y \qquad \mathrm{Var}(Y) = \mathrm{D}^2(Y) = \sigma_Y^2$$

$$\mathrm{Cov}(X,Y) = \mathrm{E}[(X-\mu_X)(Y-\mu_Y)] = \sigma_{XY}$$

Satunnaismuuttujien X ja Y (*Pearsonin tulomomentti*-) **korrelaatiokerroin** on *vakio*

$$\begin{aligned}\mathrm{Cor}(X,Y) = \rho_{XY} &= \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}} \\ &= \frac{\mathrm{Cov}(X,Y)}{\mathrm{D}(X)\,\mathrm{D}(Y)} \\ &= \frac{\sigma_{XY}}{\sigma_X \sigma_Y}\end{aligned}$$

Korrelaatiokertoimen ominaisuudet

Huomaa, että

$$\mathrm{Cor}(X,Y) = \rho_{XY} = 0$$

täsmälleen silloin, kun

$$\mathrm{Cov}(X,Y) = \sigma_{XY} = 0$$

Jos siis

$$\mathrm{Cor}(X,Y) = \rho_{XY} = 0$$

niin satunnaismuuttujat X ja Y ovat *korreloimattomia*.

Jos satunnaismuuttujat X ja Y ovat riippumattomia, niin ne ovat korreloimattomia. Sen sijaan *käänteinen ei päde*: Siitä, että satunnaismuuttujat X ja Y ovat korreloimattomia, *ei seuraa*, että satunnaismuuttujat X ja Y olisivat riippumattomia.

Olkoon satunnaismuuttujien X ja Y *korrelaatiokerroin* Cor(X,Y). Tällöin

(i) $-1 \le \mathrm{Cor}(X,Y) \le +1$

(ii) Jos X ja Y ovat *riippumattomia*, niin $\mathrm{Cor}(X,Y) = 0$

(iii) $\mathrm{Cor}(X,Y) = \pm 1$, jos ja vain, jos

$$Y = \alpha + \beta X,$$

jossa α ja β ovat reaalisia *vakiota*, $\beta \neq 0$

Lineaarimuunnokset ja 2-ulotteisen jakauman tunnusluvut

Olkoot satunnaismuuttujilla X ja Y seuraavat odotusarvot, varianssit ja kovarianssi:

$$\mathrm{E}(X) = \mu_X \qquad \mathrm{Var}(X) = \sigma_X^2$$
$$\mathrm{E}(Y) = \mu_Y \qquad \mathrm{Var}(Y) = \sigma_Y^2$$
$$\mathrm{Cov}(X,Y) = \mathrm{E}[(X-\mu_X)(Y-\mu_Y)] = \sigma_{XY}$$

Olkoot

$$W = a + bX$$
$$Z = c + dY$$

jossa a, b, c, $d \in \mathbb{R}$ ovat reaalisia vakioita. Tällöin

$$\mathrm{E}(W) = a + b\,\mathrm{E}(X) = a + b\mu_X$$
$$\mathrm{E}(Z) = c + d\,\mathrm{E}(Y) = c + d\,\mu_Y$$

$$\mathrm{Var}(W) = b^2\,\mathrm{Var}(X) = b^2\sigma_X^2$$
$$\mathrm{Var}(Z) = d^2\,\mathrm{Var}(Y) = d^2\sigma_Y^2$$

$$\mathrm{Cov}(W,Z) = bd\,\mathrm{Cov}(X,Y) = bd\sigma_{XY}$$
$$\mathrm{Cor}(W,Z) = \mathrm{sgn}(bd)\,\mathrm{Cor}(X,Y) = \mathrm{sgn}(bd)\rho_{XY}$$

Ehdolliset jakaumat

Olkoon satunnaismuuttujien X ja Y yhteisjakauman *pistetodennäköisyys*- tai *tiheysfunktio* $f_{XY}(x,y)$ ja satunnaismuuttujien X ja Y *reunajakaumien* pistetodennäköisyys- tai tiheysfunktiot $f_X(x)$ ja $f_Y(x)$.

Satunnaismuuttujan X **ehdollinen jakauma** satunnaismuuttujan Y suhteen (ehdolla $Y = y$) on

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} \text{ , jos } f_Y(y) > 0$$

Satunnaismuuttujan Y **ehdollinen jakauma** satunnaismuuttujan X suhteen (ehdolla $X = x$) on

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} \text{ , jos } f_X(x) > 0$$

Ehdolliset jakaumat ja riippumattomuus

Jos satunnaismuuttujat X ja Y ovat *riippumattomia*, satunnaismuuttujan X *ehdollinen jakauma* satunnaismuuttujan Y suhteen yhtyy satunnaismuuttujan X *reunajakaumaan*:

$$f_{X|Y}(x|y) = f_X(x) \text{ , jos } f_Y(y) > 0$$

Jos satunnaismuuttujat X ja Y ovat *riippumattomia*, satunnaismuuttujan Y *ehdollinen jakauma* satunnaismuuttujan X suhteen yhtyy satunnaismuuttujan Y *reunajakaumaan*:

$$f_{Y|X}(y|x) = f_Y(y) \text{ , jos } f_X(x) > 0$$

Diskreetin kaksiulotteisen jakauman ehdolliset odotusarvot

Olkoot satunnaismuuttujat X ja Y *diskreettejä*.

Satunnaismuuttujan X **ehdollinen odotusarvo** satunnaismuuttujan Y suhteen on satunnaismuuttujan X ehdollisen jakauman odotusarvo:

$$\mathrm{E}(X \mid Y = y) = \sum_x x f_{X|Y}(x \mid y)$$

Satunnaismuuttujan Y **ehdollinen odotusarvo** satunnaismuuttujan X suhteen on satunnaismuuttujan Y ehdollisen jakauman odotusarvo:

$$\mathrm{E}(Y \mid X = x) = \sum_y y f_{Y|X}(y \mid x)$$

Jatkuvan kaksiulotteisen jakauman ehdolliset odotusarvot

Olkoot satunnaismuuttujat X ja Y *jatkuvia*.

Satunnaismuuttujan X **ehdollinen odotusarvo** satunnaismuuttujan Y suhteen on satunnaismuuttujan X ehdollisen jakauman odotusarvo:

$$\mathrm{E}(X \mid Y = y) = \int_{-\infty}^{+\infty} x f_{X|Y}(x \mid y)\,dx$$

Satunnaismuuttujan Y **ehdollinen odotusarvo** satunnaismuuttujan X suhteen on satunnaismuuttujan Y ehdollisen jakauman odotusarvo:

$$\mathrm{E}(Y \mid X = x) = \int_{-\infty}^{+\infty} y f_{Y|X}(y \mid x)\,dy$$

Ehdolliset odotusarvot ja riippumattomuus

Jos satunnaismuuttujat X ja Y ovat *riippumattomia*, ehdolliset odotusarvot *yhtyvät* niiden *reuna-jakaumien odotusarvoihin*.

Jos siis X ja Y ovat *riippumattomia*, niin

$$\mathrm{E}(X \mid Y) = \mathrm{E}(X)$$
$$\mathrm{E}(Y \mid X) = \mathrm{E}(Y)$$

Iteroidun odotusarvon laki

Ehdolliset odotusarvot voidaan tulkita *satunnaismuuttujiksi* ehtomuuttujan suhteen.

Siten satunnaismuuttujan *Y ehdollisen odotusarvon odotusarvo* (satunnaismuuttujan X suhteen) on

$$\mathrm{E}_X\left[\mathrm{E}(Y \mid X)\right] = \mathrm{E}(Y)$$

Siten satunnaismuuttujan *X ehdollisen odotusarvon odotusarvo* (satunnaismuuttujan Y suhteen) on

$$\mathrm{E}_Y\left[\mathrm{E}(X \mid Y)\right] = \mathrm{E}(X)$$

Regressiofunktiot

Tarkastellaan satunnaismuuttujan *X ehdollista odotusarvoa*

$$\mathrm{E}(X \mid Y = y)$$

ehtomuuttujan Y arvojen y *funktiona*. Tätä funktiota kutsutaan satunnaismuuttujan X **regressio-funktioksi** satunnaismuuttujan Y suhteen.

Satunnaismuuttujan X regressiofunktio muuttujan Y suhteen määrittelee **regressiokäyrän**

$$x = g_y(y) = \mathrm{E}(X \mid Y = y)$$

Tarkastellaan satunnaismuuttujan Y *ehdollista odotusarvoa*

$$\mathrm{E}(Y \mid X = x)$$

ehtomuuttujan X arvojen x *funktiona*. Tätä funktiota kutsutaan satunnaismuuttujan Y **regressio-funktioksi** satunnaismuuttujan X suhteen.

Satunnaismuuttujan Y regressiofunktio muuttujan X suhteen määrittelee **regressiokäyrän**

$$y = g_x(x) = \mathrm{E}(Y \mid X = x)$$

15. Moniulotteisia jakaumia

Multinomijakauma

Multinomijakauma on *binomijakauman yleistys* useamman toisensa poissulkevan tapahtuman tilanteeseen.

Olkoon $A_1, A_2, \dots , A_k$ otosavaruuden S *ositus*. Tällöin

$$A_i \cap A_j = \varnothing \text{ , } i \neq j$$

$$S = A_1 \cup A_2 \cup \cdots \cup A_k$$

Olkoot tapahtumien $A_1, A_2, \dots , A_k$ *todennäköisyydet*:

$$\Pr(A_i) = p_i \text{ , } i = 1, 2, \dots , k$$

$$p_1 + p_2 + \cdots + p_k = 1$$

Määritellään *satunnaismuuttujat* X_i , $i = 1, 2, \dots , k$:

X_i = Tapahtuman A_i esiintymisten lukumäärä n-kertaisessa toistokokeessa

Tällöin

$$X_i \sim \mathrm{Bin}(n,\ p_i) \text{ , } i = 1, 2, \ldots, k$$

jossa

$$p_i = \Pr(A_i) \text{ , } i = 1, 2, \dots , k$$

Lisäksi

$$X_1 + X_2 + \cdots + X_k = n$$

Multinomijakaumalla tarkoitetaan satunnaismuuttujien

$$X_1, X_2, \dots , X_k$$

yhteisjakaumaa.

Huomautus:

Satunnaismuuttuja X_i *eivät ole riippumattomia*, koska niitä sitoo toisiinsa ehto

$$X_1 + X_2 + \cdots + X_k = n$$

jossa toistokokeiden lukumäärä n on *kiinteä* luku.

Satunnaismuuttujat $X_1, X_2, \dots , X_k$ noudattavat $(k - 1)$-ulotteista **multinomijakaumaa**, jos niiden *yhteisjakauman pistetodennäköisyysfunktio* on muotoa

$$\Pr(X_1 = n_1 \text{ ja } X_2 = n_2 \text{ ja } \dots \text{ ja } X_k = n_k) = \frac{n!}{n_1! n_2! \cdots n_k!} p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$$

jossa

$$p_1 + p_2 + \cdots + p_k = 1$$

$$n_1 + n_2 + \cdots + n_k = n$$

Merkintä:

$$(X_1, X_2, \dots , X_k) \sim \mathrm{Multinom}(p_1, p_2 , \dots , p_k \,;\, n)$$

Jos $k = 2$, niin multinomijakauma yhtyy *binomijakaumaan*:

$$\Pr_{Multinom}(X_1 = n_1 \text{ ja } X_2 = n - n_1) = \Pr_{Bin}(X_1 = n_1)$$

Multinomijakauman *yksiulotteiset reunajakaumat* ovat *binomijakaumia*.

Multinomitodennäköisyydet saadaan korottamalla *multinomi* $(p_1 + p_2 + \cdots + p_k)$ potenssiin n:

$$(p_1 + p_2 + \cdots + p_k)^n = \sum \frac{n!}{n_1!n_2!\cdots n_k!} p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$$

jossa summa lasketaan yli kaikkien lukujen $n_1, n_2, \ldots, n_k$, joille pätee ehto

$$n_1 + n_2 + \cdots + n_k = n$$

2-ulotteinen normaalijakauma

Satunnaismuuttujien X ja Y muodostama pari (X,Y) noudattaa **2-ulotteista normaalijakaumaa**, jos sen *tiheysfunktio* on

$$(1) \qquad f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \exp\left\{-\frac{1}{2(1-\rho_{XY}^2)}Q(x,y)\right\}$$

jossa

$$Q(x,y) = \left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho_{XY}\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2$$

ja

$$\begin{aligned} &-\infty < \mu_X < +\infty && \sigma_X > 0 \\ &-\infty < \mu_Y < +\infty && \sigma_Y > 0 \\ &-1 \le \rho_{XY} \le +1 \end{aligned}$$

2-ulotteisen normaalijakauman tiheysfunktion määrittelevän yhtälön (1) hakasulkulauseke [·] määrää tiheysfunktion **tasa-arvokäyrät**. Kaikki tasa-arvokäyrät ovat **ellipsejä**, joiden yhtälöt voidaan ilmaista muodossa

$$Q(x,y) = \left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho_{XY}\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 = c^2$$

missä c on vakio.

2-ulotteinen normaalijakauman tiheysfunktio (1) on *parametroitu* niin, että sen parametreina ovat satunnaismuuttujien X ja Y *odotusarvot*, *varianssit* ja *korrelaatio*.

Satunnaismuuttujien X ja Y **odotusarvot** ovat

$$\begin{aligned} \mu_X &= \mathrm{E}(X) \\ \mu_Y &= \mathrm{E}(Y) \end{aligned}$$

Satunnaismuuttujien X ja Y **varianssit** ovat

$$\sigma_X^2 = \mathrm{Var}(X) = \mathrm{E}\left[(X-\mu_X)^2\right]$$
$$\sigma_Y^2 = \mathrm{Var}(Y) = \mathrm{E}\left[(Y-\mu_Y)^2\right]$$

Satunnaismuuttujien X ja Y **korrelaatio** on

$$\rho_{XY} = \mathrm{Cor}(X,Y) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

jossa

$$\sigma_{XY} = \mathrm{Cov}(X,Y) = \mathrm{E}\left[(X-\mu_X)(Y-\mu_Y)\right]$$

on satunnaismuuttujien X ja Y **kovarianssi**.

2-ulotteista normaalijakaumaa noudattavan satunnaismuuttujien X ja Y parin (X, Y) **odotusarvovektori** on

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix}$$

ja **kovarianssimatriisi** on

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{bmatrix} = \begin{bmatrix} \sigma_X^2 & \rho_{XY}\sigma_X\sigma_Y \\ \rho_{XY}\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$$

2-ulotteisen normaalijakauman **reunajakaumat** ovat 1-*ulotteisia normaalijakaumia*:

$$Y \sim \mathrm{N}(\mu_Y, \sigma_Y^2)$$
$$X \sim \mathrm{N}(\mu_X, \sigma_X^2)$$

2-*ulotteisen normaalijakauman tapauksessa* satunnaismuuttujien X ja Y *korreloimattomuudesta seuraa* niiden *riippumattomuus*. Muista, että *aina pätee* se, että *riippumattomuudesta seuraa korreloimattomuus*.

2-ulotteisen normaalijakauman **ehdolliset jakaumat** ovat 1-*ulotteisia normaalijakaumia*:

$$(Y \mid X) \sim \mathrm{N}(\mu_Y + \beta_{YX}(x-\mu_X)\,,\, \sigma_Y^2(1-\rho_{XY}^2))\,,\ \beta_{YX} = \rho_{XY}\frac{\sigma_Y}{\sigma_X}$$

$$(X \mid Y) \sim \mathrm{N}(\mu_X + \beta_{XY}(y-\mu_Y)\,,\, \sigma_X^2(1-\rho_{XY}^2))\,,\ \beta_{XY} = \rho_{XY}\frac{\sigma_X}{\sigma_Y}$$

Ehdollinen odotusarvo

$$\mathrm{E}(Y \mid X) = \mu_Y + \beta_{YX}(x-\mu_X)$$

on muuttujan Y **regressiofunktio** muuttujan X suhteen. Ehdollinen odotusarvo E(Y|X) määrää *suoran*

$$y = \mu_Y + \beta_{YX}(x-\mu_X)$$

Suoran *kulmakerroin* on $\beta_{YX} = \rho_{XY}\dfrac{\sigma_Y}{\sigma_X}$ ja se kulkee pisteen (μ_X,μ_Y) kautta.

Ehdollinen odotusarvo

$$\mathrm{E}(X\,|\,Y) = \mu_X + \beta_{XY}(y - \mu_Y)$$

on muuttujan X **regressiofunktio** muuttujan Y suhteen. Ehdollinen odotusarvo $\mathrm{E}(X|Y)$ määrää *suoran*

$$x = \mu_X + \beta_{XY}(y - \mu_Y)$$

Suoran *kulmakerroin* on $\beta_{XY} = \rho_{XY}\dfrac{\sigma_X}{\sigma_Y}$ ja se kulkee pisteen (μ_X,μ_Y) kautta.

Huomaa, että regressiosuorien kulmakertoimet β_{YX} ja β_{XY} toteuttavat yhtälön

$$\beta_{YX}\beta_{XY} = \rho_{XY}^2$$

2-ulotteisen normaalijakauman tapauksessa satunnaismuuttujien Y ja X **yhteiskorrelaatiokerroin** on

$$R_{Y|X} = \frac{\sigma_{YX}}{\sigma_Y\sigma_X} = \rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} = R_{X|Y}$$

ja **ehdolliset varianssit** ovat

$$\sigma_{Y|X}^2 = \sigma_Y^2(1 - \rho_{XY}^2)$$

$$\sigma_{X|Y}^2 = \sigma_X^2(1 - \rho_{XY}^2)$$

2-ulotteisen normaalijakauman tiheysfunktiota muodon määräävien **tasa-arvoellipsien**

$$Q(x,y) = \left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 = c^2$$

pääakseleiden pituudet (oik. pituuksien suhteet) ja **suunnat** saadaan määräämällä *kovarianssi-matriisin*

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{bmatrix} = \begin{bmatrix} \sigma_X^2 & \rho_{XY}\sigma_X\sigma_Y \\ \rho_{XY}\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$$

ominaisarvot ja **ominaisvektorit**.

Matriisin $\mathbf{\Sigma}$ *ominaisarvot* saadaan määräämällä determinanttiyhtälön

$$|\mathbf{\Sigma} - \lambda\mathbf{I}| = \begin{vmatrix} \sigma_X^2 - \lambda & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 - \lambda \end{vmatrix} = \lambda^2 - (\sigma_X^2 + \sigma_Y^2)\lambda + \sigma_X^2\sigma_Y^2 - \sigma_{XY}^2 = 0$$

nollakohdat muuttujan λ suhteen. Yhtälöllä on (aina) kaksi reaalista ja ei-negatiivista nollakohtaa, jotka ovat siis matriisin $\mathbf{\Sigma}$ ominaisarvot.

Matriisin $\mathbf{\Sigma}$ ominaisarvoja λ_1 ja λ_2 vastaavat *ominaisvektorit*

$$\mathbf{q}_1 = (q_{11}, q_{21})$$

$$\mathbf{q}_2 = (q_{12}, q_{22})$$

saadaan yhtälöistä

$$\mathbf{\Sigma q}_i = \lambda_i \mathbf{q}_i \text{ , } i = 1, 2$$

ottamalla huomioon ehdot

$$\mathbf{q}_i'\mathbf{q}_i = q_{1i}^2 + q_{2i}^2 = 1 \text{ , } i = 1, 2$$

Tasa-arvoellipsien

$$Q(x,y) = c^2$$

pääakseleiden *pituudet* suhtautuvat toisiinsa kuten *ominaisarvojen* λ_1 ja λ_2 *neliöjuuret* ja pääakseleiden *suunnat* yhtyvät vastaavien *ominaisvektoreiden suuntiin*.

16. Momenttiemäfunktio ja karakteristinen funktio

Momenttiemäfunktio

Olkoon X satunnaismuuttuja. Jos on olemassa $h > 0$ siten, että odotusarvo

$$m_X(t) = \mathrm{E}(e^{tX})$$

on olemassa kaikille

$$t \in (-h, +h)$$

niin sanomme, että $m_X(t)$ on satunnaismuuttujan X **momenttiemäfunktio**.

Aina pätee

$$m_X(0) = \mathrm{E}(e^0) = \mathrm{E}(1) = 1$$

Momenttiemäfunktion yksikäsitteisyys

Jos satunnaismuuttujan X momenttiemäfunktio

$$m_X(t) = \mathrm{E}(e^{tX})$$

on olemassa jossakin pisteen $t = 0$ ympäristössä, *se on yksikäsitteinen ja määrää täysin satunnaismuuttujan X todennäköisyysjakauman*.

Momenttiemäfunktion yksikäsitteisyydestä seuraa, että jos satunnaismuuttujilla U ja V on *sama* momenttiemäfunktio jossakin pisteen $t = 0$ ympäristössä, niin ne noudattavat *samaa* jakaumaa.

Momenttiemäfunktio ja satunnaismuuttujan momentit

Satunnaismuuttujan X ***k*. origomomentti**

$$\alpha_k = \mathrm{E}(X^k)\,,\ k = 1, 2, 3, \ldots$$

saadaan määräämällä satunnaismuuttujan X momenttiemäfunktion $m_X(t)$ k. derivaatta pisteessä $t = 0$:

$$\alpha_k = \mathrm{E}(X^k) = \left.\frac{d^k m_X(t)}{dt^k}\right|_{t=0}\,,\ k = 1, 2, 3, \ldots$$

Satunnaismuuttujan X *odotusarvo* ja *varianssi* saadaan kaavoilla

$$\mu = \mathrm{E}(X) = \alpha_1$$
$$\sigma^2 = \mathrm{Var}(X) = \mathrm{E}[(X-\mu)^2] = \mathrm{E}(X^2) - \mu^2 = \alpha_2 - \alpha_1^2$$

Momenttiemäfunktion Taylorin sarja

Satunnaismuuttujan X momenttiemäfunktio $m_X(t)$ voidaan kehittää **Taylorin sarjaksi**

$$m_X(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathrm{E}(X^k) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \alpha_k$$

jossa

$$\alpha_k = \mathrm{E}(X^k)\,,\ k = 1, 2, 3, \ldots$$

on satunnaismuuttujan X *k. momentti*.

Diskreetin jakauman momenttiemäfunktio

Olkoon satunnaismuuttuja X *diskreetti* ja sen *pistetodennäköisyysfunktio*

$$f(x) = \Pr(X = x)$$

Tällöin satunnaismuuttujan X *momenttiemäfunktio* saadaan kaavalla

$$m_X(t) = \mathrm{E}(e^{tX}) = \sum_x e^{tx} f(x)$$

Jatkuvan jakauman momenttiemäfunktio

Olkoon satunnaismuuttuja X *jatkuva* ja sen *tiheysfunktio*

$$f(x)$$

Tällöin satunnaismuuttujan X *momenttiemäfunktio* saadaan kaavalla

$$m_X(t) = \mathrm{E}(e^{tX}) = \int_{-\infty}^{+\infty} e^{tx} f(x)dx$$

Riippumattomien satunnaismuuttujien summan momenttiemäfunktio

Olkoot

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

riippumattomia satunnaismuuttujia, joiden momenttiemäfunktiot ovat

$$m_1(t), m_2(t), \ldots, m_n(t)$$

Tällöin **summan**

$$X = X_1 + X_2 + \cdots + X_n$$

momenttiemäfunktio on satunnaismuuttujien $X_1, X_2, \ldots, X_n$ momenttiemäfunktioiden tulo:

$$m_X(t) = m_1(t)m_2(t) \cdots m_n(t)$$

Riippumattomien samoin jakautuneiden satunnaismuuttujien summan momenttiemäfunktio

Olkoot

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

riippumattomia ja *samoin jakautuneita* satunnaismuuttujia, joiden momenttiemäfunktio on

$$m(t)$$

Tällöin satunnaismuuttujien $X_1, X_2, \ldots, X_n$ **summan**

$$X = X_1 + X_2 + \cdots + X_n$$

momenttiemäfunktio on muotoa

$$m_X(t) = [m(t)]^n$$

Riippumattomien samoin jakautuneiden satunnaismuuttujien aritmeettisen keskiarvon momenttiemäfunktio

Olkoot

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

riippumattomia ja *samoin jakautuneita* satunnaismuuttujia, joiden momenttiemäfunktio on

$$m(t)$$

Tällöin satunnaismuuttujien $X_1, X_2, \ldots, X_n$ **aritmeettisen keskiarvon**

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

momenttiemäfunktio on muotoa

$$m_{\overline{X}}(t) = [m(t/n)]^n$$

Diskreetin tasaisen jakauman momenttiemäfunktio

Oletetaan, että satunnaismuuttuja X noudattaa **diskreettiä tasaista jakaumaa**.

Tällöin sen *pistetodennäköisyysfunktio* on

$$f(x) = \Pr(X = x) = \frac{1}{n}\,,\; x = x_k\,,\; k = 1, 2, \ldots, n$$

jossa $\{x_1, x_2, \ldots, x_n\}$ on reaaliakselin *erillisten* pisteiden muodostama joukko.

Diskreetin tasaisen jakauman *momenttiemäfunktio* on

$$m_X(t) = \frac{1}{n}\sum_{k=1}^{n} e^{tx_k}$$

Bernoulli-jakauman momenttiemäfunktio

Oletetaan, että satunnaismuuttuja X noudattaa **Bernoulli-jakaumaa** Ber(p).

Tällöin sen *pistetodennäköisyysfunktio* on

$$f(x) = \Pr(X = x) = p^x q^{1-x}\,,\; 0 < p < 1\,,\; q = 1 - p\,,\; x = 0, 1$$

Bernoulli-jakauman *momenttiemäfunktio* on

$$m_X(t) = q + pe^t$$

Binomijakauman momenttiemäfunktio

Oletetaan, että satunnaismuuttuja X noudattaa **binomijakaumaa** Bin(n, p).

Tällöin sen *pistetodennäköisyysfunktio* on

$$f(x) = \Pr(X = x) = \binom{n}{x} p^x q^{n-x}\,,\; 0 < p < 1\,,\; q = 1 - p\,,\; x = 0, 1, 2, \ldots, n$$

Binomijakauman *momenttiemäfunktio* on

$$m_X(t) = (q + pe^t)^n$$

Geometrisen jakauman momenttiemäfunktio

Oletetaan, että satunnaismuuttuja X noudattaa **geometrista jakaumaa** Geom(p).

Tällöin sen *pistetodennäköisyysfunktio* on

$$f(x) = \Pr(X = x) = q^{x-1}p \text{ , } 0 < p < 1 \text{ , } q = 1 - p \text{ , } x = 1, 2, 3, \ldots$$

Geometrisen jakauman *momenttiemäfunktio* on

$$m_X(t) = \frac{pe^t}{1 - qe^t}$$

Negatiivisen binomijakauman momenttiemäfunktio

Oletetaan, että satunnaismuuttuja X noudattaa **negatiivista binomijakaumaa** NegBin(r, p).

Tällöin sen *pistetodennäköisyysfunktio* on

$$f(x) = \Pr(X = x) = \binom{x-1}{r-1} q^{x-r}p^r \text{ , } 0 < p < 1 \text{ , } q = 1 - p \text{ , } r = 1, 2, 3, \ldots \text{ ;}$$
$$x = r, r+1, r+2, \ldots$$

Negatiivisen binomijakauman *momenttiemäfunktio* on

$$m_X(t) = \frac{(pe^t)^r}{(1 - qe^t)^r}$$

Poisson-jakauman momenttiemäfunktio

Oletetaan, että satunnaismuuttuja X noudattaa **Poisson-jakaumaa** Poisson(λ).

Tällöin sen *pistetodennäköisyysfunktio* on

$$f(x) = \Pr(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} \text{ , } \lambda > 0$$
$$x = 0, 1, 2, \ldots$$

Poisson-jakauman *momenttiemäfunktio* on

$$m_X(t) = e^{\lambda(e^t - 1)}$$

Jatkuvan tasaisen jakauman momenttiemäfunktio

Oletetaan, että satunnaismuuttuja X noudattaa **jatkuvaa tasaista jakaumaa** Uniform(a, b).

Tällöin sen *tiheysfunktio* on

$$f(x) = \frac{1}{b - a} \text{ , } a \le x \le b$$

Jatkuvan tasaisen jakauman *momenttiemäfunktio* on

$$m_X(t) = \frac{e^{bt} - e^{at}}{t(b-a)}$$

Eksponenttijakauman momenttiemäfunktio

Oletetaan, että satunnaismuuttuja X noudattaa **eksponenttijakaumaa** Exp(λ).

Tällöin sen *tiheysfunktio* on

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

Eksponenttijakauman *momenttiemäfunktio* on

$$m_X(t) = \frac{\lambda}{\lambda - t}$$

Normaalijakauman momenttiemäfunktio

Oletetaan, että satunnaismuuttuja X noudattaa **normaalijakaumaa** N(μ,σ^2).

Tällöin sen *tiheysfunktio* on

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}, -\infty < \mu < +\infty \, , \, \sigma > 0 \, , \, -\infty < x < +\infty$$

Normaalijakauman *momenttiemäfunktio* on

$$m_X(t) = \exp(\mu t + \tfrac{1}{2}\sigma^2 t^2)$$

Karakteristinen funktio

Olkoon X *satunnaismuuttuja*. Tällöin odotusarvo

$$\varphi_X(t) = \mathrm{E}(e^{itX}) \, , \, i = \sqrt{-1}$$

on satunnaismuuttujan X ja sen jakauman **karakteristinen funktio**. Karakteristinen funktio *on* – toisin kuin sen momenttiemäfunktio – *aina olemassa*.

Jos satunnaismuuttujan X *momenttiemäfunktio*

$$m_X(t) = \mathrm{E}(e^{tX})$$

tunnetaan, saadaan sen karakteristinen funktio

$$\varphi_X(t) = \mathrm{E}(e^{itX}) \, , \, i = \sqrt{-1}$$

momenttiemäfunktiosta sijoituksella

$$t \to it \, , \, i = \sqrt{-1}$$

Karakteristisen funktion ominaisuudet

Aina pätee:

(i) $\varphi_X(0) = \mathrm{E}(e^0) = \mathrm{E}(1) = 1$

(ii) $|\varphi_X(t)| \leq 1$ kaikille $t \in (-\infty, +\infty)$.

(iii) $\varphi_X(-t) = \bar{\varphi}_X(t)$
jossa merkintä $\bar{z}$ tarkoittaa kompleksiluvun *z konjugaattia*.

(iv) $\varphi_X(t)$ on *tasaisesti jatkuva* kaikille $t \in (-\infty, +\infty)$.

Diskreetin satunnaismuuttujan karakteristinen funktio

Olkoon *X diskreetti satunnaismuuttuja*, jonka *pistetodennäköisyysfunktio* on

$$f_X(x) = \Pr(X = x)$$

Satunnaismuuttujan X karakteristinen funktio saadaan kaavalla

$$\varphi_X(t) = \mathrm{E}(e^{itX}) = \sum_x e^{itx} f_X(x) \text{ , } i = \sqrt{-1}$$

Jatkuvan satunnaismuuttujan karakteristinen funktio

Olkoon *X jatkuva satunnaismuuttuja*, jonka *tiheysfunktio* on

$$f_X(x)$$

Satunnaismuuttujan X karakteristinen funktio saadaan kaavalla

$$\varphi_X(t) = \mathrm{E}(e^{itX}) = \int_{-\infty}^{+\infty} e^{itx} f_X(x)dx \text{ , } i = \sqrt{-1}$$

Inversioteoreema

Olkoon

$$F_X(x) = \Pr(X \leq x)$$

satunnaismuuttujan *X kertymäfunktio* ja

$$\varphi_X(t) = \mathrm{E}(e^{itX}) \text{ , } i = \sqrt{-1}$$

sen *karakteristinen funktio*. Oletetaan, että

$$(a - h, a + h)$$

sellainen reaaliakselin väli, että kertymäfunktio $F_X(x)$ on *jatkuva välin päätepisteissä*. Tällöin

$$F_X(a+h) - F_X(a-h) = \lim_{T \to +\infty} \frac{1}{\pi} \int_{-T}^{+T} \frac{\sin(ht)}{t} e^{-ita} \varphi_X(t)dt$$

Jos jakauman *karakteristinen funktio tunnetaan*, *voidaan* jakauman *kertymäfunktio määrätä* inversioteoreemassa määritellyn rajaprosessin avulla. Myös *karakteristisen funktion yksikäsitteisyys* voidaan todistaa inversioteoreeman avulla.

Karakteristisen funktion yksikäsitteisyys

Satunnaismuuttujan X karakteristinen funktio on **yksikäsitteinen** ja **määrää täysin satunnaismuuttujan *X* todennäköisyysjakauman**.

Tämä merkitsee seuraavaa: Jos satunnaismuuttujilla U ja V on *sama karakteristinen funktio*, satunnaismuuttujat U ja V noudattavat *samaa todennäköisyysjakaumaa*.

Jatkuvan jakauman karakteristinen funktio

Olkoon

$$\varphi_X(t) = \mathrm{E}(e^{itX}) \text{ , } i = \sqrt{-1}$$

satunnaismuuttujan X karakteristinen funktio.

Oletetaan, että

$$|\varphi_X(t)|$$

on *integroituva* kaikille $t \in (-\infty, +\infty)$. Tällöin satunnaismuuttuja X on *jatkuva* ja sen *tiheysfunktio* $f_X(x)$ saadaan kaavalla

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx}\varphi_X(t)dt \text{ , } i = \sqrt{-1}$$

Huomaa, että *jatkuvan satunnaismuuttujan X* karakteristinen funktio

$$\varphi_X(t) = \mathrm{E}(e^{itX}) = \int_{-\infty}^{+\infty} e^{itx} f_X(x)dx \text{ , } i = \sqrt{-1}$$

on satunnaismuuttujan X tiheysfunktion **Fourier-muunnos** ja

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx}\varphi_X(t)dt \text{ , } i = \sqrt{-1}$$

on sen **käänteinen Fourier-muunnos**.

Karakteristinen funktio ja satunnaismuuttujan momentit

Olkoon

$$\varphi_X(t) = \mathrm{E}(e^{itX}) \text{ , } i = \sqrt{-1}$$

satunnaismuuttujan X karakteristinen funktio.

Oletetaan, että satunnaismuuttujan X ***r*. (origo-) momentti**

$$\alpha_r = \mathrm{E}(X^r)$$

on olemassa.

Tällöin karakteristinen funktio $\varphi_X(t)$ on *differentioituva kertalukuun r* ja

$$\alpha_k = \mathrm{E}(X^k) = \frac{1}{i^k}\cdot\left.\frac{d^k\varphi_X(t)}{dt^k}\right|_{t=0} \text{ , } i = \sqrt{-1} \text{ , } k = 1, 2, \ldots, r$$

Oletetaan, että satunnaismuuttujan X karakteristinen funktio $\varphi_X(t)$ on *differentioituva kertalukuun r*. Tällöin *kaikki momentit*

$$\alpha_k = \mathrm{E}(X^k) \text{ , } k = 1, 2, \ldots, r$$

ovat olemassa, jos r on *parillinen* ja *kaikki momentit*

$$\alpha_k = \mathrm{E}(X^k)\,,\ k = 1, 2, \ldots, r-1$$

ovat olemassa, jos r on *pariton*.

Karateristisen funktion Taylorin sarjakehitelmä

Olkoon

$$\varphi_X(t) = \mathrm{E}(e^{itX})\,,\ i = \sqrt{-1}$$

satunnaismuuttujan X karakteristinen funktio. Oletetaan, että satunnaismuuttujan X r. (origo-) *momentti*

$$\alpha_r = \mathrm{E}(X^r)$$

on olemassa. Tällöin karakteristinen funktio $\varphi_X(t)$ voidaan kehittää **Taylorin sarjaksi**

$$\varphi_X(t) = \sum_{k=0}^{r} \frac{(it)^k}{k!} \mathrm{E}(X^k) + o(t^r) = \sum_{k=0}^{r} \frac{(it)^k}{k!} \alpha_k + o(t^r)$$

jossa

$$o(t^r)/t^r \to 0\,,\ \text{jos } t \to 0$$

Riippumattomien satunnaismuuttujien summan karakteristinen funktio

Olkoot

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

riippumattomia satunnaismuuttujia, joiden *karakteristiset funktiot* ovat

$$\varphi_1(t), \varphi_2(t), \ldots, \varphi_n(t)$$

Tällöin **summan**

$$X = X_1 + X_2 + \cdots + X_n$$

karakteristinen funktio on satunnaismuuttujien $X_1, X_2, \ldots, X_n$ karakterististen funktioiden *tulo*:

$$\varphi_X(t) = \varphi_1(t)\varphi_2(t) \cdots \varphi_n(t)$$

Riippumattomien samoin jakautuneiden satunnaismuuttujien summan karakteristinen funktio

Olkoot

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

riippumattomia ja *samoin jakautuneita* satunnaismuuttujia, joiden karakteristinen funktio on

$$\varphi(t)$$

Tällöin satunnaismuuttujien $X_1, X_2, \ldots, X_n$ **summan**

$$X = X_1 + X_2 + \cdots + X_n$$

karakteristinen funktio on muotoa

$$\varphi_X(t) = [\varphi(t)]^n$$

Riippumattomien samoin jakautuneiden satunnaismuuttujien aritmeettisen keskiarvon karakteristinen funktio

Olkoot

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

riippumattomia ja *samoin jakautuneita* satunnaismuuttujia, joiden karakteristinen funktio on

$$\varphi(t)$$

Tällöin satunnaismuuttujien $X_1, X_2, \ldots, X_n$ **aritmeettisen keskiarvon**

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

karakteristinen funktio on muotoa

$$\varphi_{\bar{X}}(t) = [\varphi(t/n)]^n$$

17. Satunnaismuuttujien muunnokset ja niiden jakaumat

Satunnaismuuttujan lineaarimuunnoksen jakauma

Olkoon X *jatkuva satunnaismuuttuja*, jonka *tiheysfunktio* on

$$f_X(x)$$

Muodostetaan satunnaismuuttujan X *lineaarimuunnos*

$$Y = a + bX$$

jossa a ja $b \neq 0$ ovat *vakioita*. Tällöin satunnaismuuttujan Y tiheysfunktio on

$$f_Y(y) = \frac{1}{|b|} f_X\left(\frac{y-a}{b}\right)$$

Satunnaismuuttujien muunnosten jakaumat

Olkoon satunnaismuuttujan X *jakauman* tiheysfunktio

$$f_X(x)$$

Määritellään satunnaismuuttuja

$$U = h(X)$$

jossa muunnos

$$u = h(x)$$

on *aidosti monotoninen* ja *jatkuvasti derivoituva*. Tällöin satunnaismuuttujan U tiheysfunktio saadaan kaavalla

$$f_U(u) = f_X(h^{-1}(u))\left|\frac{dh^{-1}(u)}{du}\right|$$

Kaksiulotteisten satunnaismuuttujien muunnosten jakaumat

Olkoon satunnaismuuttujien X ja Y *yhteisjakauman* tiheysfunktio

$$f_{XY}(x,y)$$

Määritellään satunnaismuuttujat

$$U = g(X,Y)$$

$$V = h(X,Y)$$

Oletetaan, että muunnoksella

$$(*) \qquad \begin{cases} u = g(x,y) \\ v = h(x,y) \end{cases}$$

on seuraavat ominaisuudet:

(i) Muuttujat x ja y *voidaan ratkaista yksikäsitteisesti* yhtälöryhmästä $(*)$.

(ii) Funktioilla g ja h on *jatkuvat osittaisderivaatat* muuttujien x ja y suhteen.

(iii) Muunnoksen $(*)$ *Jacobin determinantti*

$$\frac{\partial(u,v)}{\partial(x,y)} = \frac{\partial u}{\partial x}\cdot\frac{\partial v}{\partial y} - \frac{\partial u}{\partial y}\cdot\frac{\partial v}{\partial x} \neq 0$$

kaikille x ja y, joille

$$f_{XY}(x,y) \neq 0$$

Tällöin satunnaismuuttujien U ja V yhteisjakauman tiheysfunktio on

$$f_{UV}(u,v) = f_{XY}(x,y)\left|\frac{\partial(u,v)}{\partial(x,y)}\right|^{-1}$$

jossa x ja y ratkaistaan yhtälöryhmästä $(*)$.

Riippumattomien satunnaismuuttujien summan jakauma

Olkoot X ja Y *riippumattomia* satunnaismuuttujia, joiden *tiheysfunktiot* ovat

$$f_X(x), f_Y(y)$$

Tällöin *summan*

$$U = X + Y$$

tiheysfunktio on

$$f_U(u) = \int_{-\infty}^{+\infty} f_Y(u-x) f_X(x)dx$$

Riippumattomien satunnaismuuttujien osamäärän jakauma

Olkoot X ja Y *riippumattomia* satunnaismuuttujia, joiden *tiheysfunktiot* ovat

$$f_X(x), f_Y(y)$$

Tällöin *osamäärän*

$$U = X / Y$$

tiheysfunktio on

$$f_U(u) = \int_{-\infty}^{+\infty} y f_Y(uy) f_Y(y)dy$$

Riippumattomien satunnaismuuttujien minimin jakauma

Olkoot satunnaismuuttujat

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

riippumattomia ja *samaa jakaumaa* noudattavia satunnaismuuttujia.

Olkoon satunnaismuuttujien $X_1, X_2, \ldots , X_n$ yhteinen *tiheysfunktio*

$$f(x)$$

ja yhteinen *kertymäfunktio*

$$F(x)$$

Tällöin satunnaismuuttujan

$$X_{(1)} = \min\{X_1, X_2, \ldots, X_n\}$$

tiheysfunktio on

$$f_{(1)}(x) = n[1 - F(x)]^{n-1} f(x)$$

Riippumattomien satunnaismuuttujien maksimin jakauma

Olkoot satunnaismuuttujat

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

riippumattomia ja *samaa jakaumaa* noudattavia satunnaismuuttujia.

Olkoon satunnaismuuttujien $X_1, X_2, \ldots, X_n$ yhteinen *tiheysfunktio*

$$f(x)$$

ja yhteinen *kertymäfunktio*

$$F(x)$$

Tällöin satunnaismuuttujan

$$X_{(n)} = \max\{X_1, X_2, \ldots, X_n\}$$

tiheysfunktio on

$$f_{(n)}(x) = n[F(x)]^{n-1} f(x)$$

18. Konvergenssikäsitteet ja raja-arvolauseet

Miksi todennäköisyyslaskennassa käytetään useita konvergenssikäsitteitä?

Olkoon

$$(S, \mathfrak{F}, \Pr)$$

todennäköisyyskenttä, jossa S on *otosavaruus*, $\mathfrak{F}$ otosavaruuden S osajoukoille määritelty σ-*algebra* ja Pr σ-algebran $\mathfrak{F}$ alkioille määritelty *todennäköisyysmitta*. Olkoon X (mitallinen) *funktio* otosavaruudesta S reaalilukujen joukkoon:

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

Tällöin sanomme, että funktio X on *satunnaismuuttuja*. Jos haluamme korostaa sitä, että *satunnaismuuttuja on otosavaruuden S kuvaus reaalilukujen joukkoon*, merkitsemme

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

Tarkastelemme seuraavassa satunnaismuuttujien

$$X_1, X_2, X_3, \ldots$$

muodostamia **jonoja** ja niiden *konvergenssia*. On syytä huomata, että satunnaismuuttujien X_1, X_2, X_3, … muodostamat jonot *eivät ole* lukujonoja tavanomaisessa mielessä, vaan jokainen satunnaismuuttujien X_1, X_2, X_3, … muodostama jono on itse asiassa *lukujonojen joukko*. Tämä nähdään siitä, että satunnaismuuttujien X_1, X_2, X_3, … muodostamassa jonossa *jokaiseen* otosavaruuden alkioon $s \in S$ liittyy lukujono

$$X_1(s), X_2(s), X_3(s), \ldots$$

Siten lukujono

$$X_1(s), X_2(s), X_3(s), \ldots$$

voi *konvergoida*, kun

$$s \in A \subset S$$

ja *hajaantua*, kun

$$s \in A^c \subset S$$

Tästä seuraa, että satunnaismuuttujien jonojen konvergenssin määritteleminen on paljon vaativampi tehtävä kuin tavanomaisten lukujonojen konvergenssin määritteleminen.

Varma konvergenssi

Satunnaismuuttujien

$$X_1, X_2, X_3, \ldots$$

muodostama jono **konvergoi varmasti** kohti satunnaismuuttujaa X, jos

$$\lim_{i\to\infty} X_i(s) = X(s)\text{, kaikille } s \in S$$

Varman konvergenssin käsite on tavallisesti liian rajoittava ollakseen hyödyllinen.

Melkein varma konvergenssi

Satunnaismuuttujien

$X_1, X_2, X_3, \ldots$

muodostama jono **konvergoi melkein varmasti** eli **todennäköisyydellä yksi** kohti satunnaismuuttujaa X, jos

$$\Pr(\lim_{n\to\infty} X_n = X) = 1$$

Merkitään tätä:

$$\lim_{n\to\infty} X_n = X \quad \text{(a.s.)}$$

tai

$$X_n \xrightarrow[n\to\infty]{\text{a.s.}} X$$

jossa lyhenne a.s. = *almost surely*.

Tällöin sen joukon, jossa konvergenssia ei tapahdu, on *nollamittainen*.

Kvadraattinen konvergenssi

Satunnaismuuttujien

$X_1, X_2, X_3, \ldots$

muodostama jono konvergoi **kvadraattisesti** kohti satunnaismuuttujaa X, jos

$$\lim_{n\to\infty} \mathrm{E}[(X_n - X)^2] = 0$$

Merkitään tätä:

$$\lim_{n\to\infty} X_n = X \quad \text{(q.m.)}$$

tai

$$X_n \xrightarrow[n\to\infty]{\text{q.m.}} X$$

jossa lyhenne q.m. = *in quadratic mean*.

Stokastinen konvergenssi

Satunnaismuuttujien

$X_1, X_2, X_3, \ldots$

muodostama jono konvergoi **stokastisesti** kohti satunnaismuuttujaa X, jos kaikille $\varepsilon > 0$ pätee

$$\lim_{n\to\infty} \Pr(|X_n - X| > \varepsilon) = 0$$

Merkitään tätä:

$$\lim_{n\to\infty} X_n = X \quad \text{(P)}$$

tai

$$X_n \xrightarrow[n\to\infty]{\text{P}} X$$

jossa lyhenne P = *in probability*.

Jakaumakonvergenssi

Olkoon

$$X_1, X_2, X_3, \ldots$$

jono satunnaismuuttujia, joiden kertymäfunktiot ovat

$$F_1(x), F_2(x), F_3(x), \ldots$$

Satunnaismuuttujien $X_1, X_2, X_3, \ldots$ muodostama jono konvergoi **jakaumaltaan** (tai **heikosti**) kohti satunnaismuuttujaa X, jonka kertymäfunktio on $F_X(x)$, jos

$$\lim_{n\to\infty} F_n(x) = F_X(x)$$

jokaisessa pisteessä x, jossa $F_X(x)$ on jatkuva.

Merkitään tätä:

$$\lim_{n\to\infty} X_n = X \quad \text{(L)}$$

tai

$$X_n \xrightarrow[n\to\infty]{\text{L}} X$$

jossa lyhenne L = *in law*.

Konvergenssikäsitteiden yhteydet

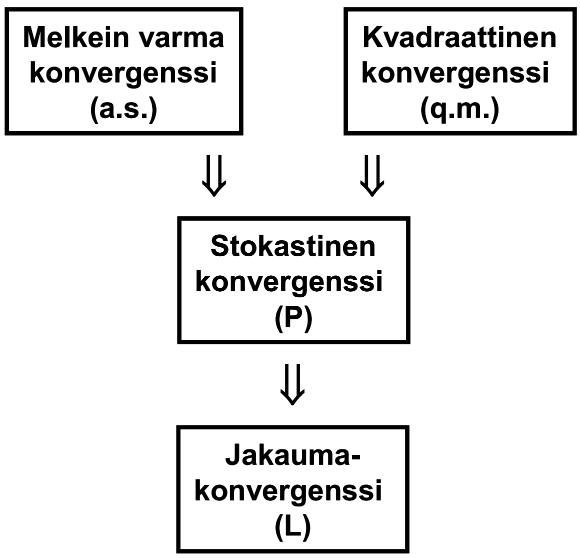
Voidaan osoittaa, että todennäköisyyslaskennan konvergenssikäsitteillä on seuraavat yhteydet:

(i) **Melkein varma konvergenssi (a.s.) implikoi stokastisen konvergenssin (P)**.

(ii) **Kvadraattinen konvergenssi (q.m.) implikoi stokastisen konvergenssin (P)**.

(iii) **Stokastinen konvergenssi (P) implikoi jakaumakonvergenssin eli heikon konvergenssin (L)**.

(iv) Melkein varman ja kvadraattisen konvergenssin yhteydestä ei voida sanoa mitään yleistä.



Vahva suurten lukujen laki

Olkoon

$$X_i\,,\ i = 1, 2, 3, \ldots$$

jono *riippumattomia* ja *samoin jakautuneita* satunnaismuuttujia, joilla on *sama* odotusarvo:

$$\mathrm{E}(X_i) = \mu\,,\ i = 1, 2, 3, \ldots$$

Määritellään satunnaismuuttujien X_i , $i = 1, 2, 3, \ldots$ *aritmeettinen keskiarvo*:

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Tällöin pätee **vahva suurten lukujen laki**: Satunnaismuuttujien $X_1, X_2, X_3, \ldots , X_n$ aritmeettisten keskiarvojen muodostama jono $\bar{X}_n = \Sigma X_i / n$ konvergoi **melkein varmasti** eli **todennäköisyydellä yksi** kohti satunnaismuuttujien yhteistä odotusarvoa μ:

$$\bar{X}_n \xrightarrow[n\to\infty]{\text{a.s.}} \mu$$

Vahva suurten lukujen laki voidaan ilmaista sanoin seuraavasti: Samoin jakautuneiden satunnaismuuttujien *aritmeettinen keskiarvo lähestyy muuttujien lukumäärän kasvaessa rajatta muuttujien yhteistä odotusarvoa melkein kaikkialla* eli se **otosavaruuden *S* osajoukko, jossa konvergenssia ei tapahdu on *nollamittainen***.

Heikko suurten lukujen laki

Olkoon X_i , $i = 1, 2, 3, \ldots$ jono *riippumattomia* satunnaismuuttujia, joilla on *sama* odotusarvo ja varianssi:

$$\mathrm{E}(X_i) = \mu\,,\ \mathrm{D}^2(X_i) = \sigma^2\,,\ i = 1, 2, 3, \ldots$$

Määritellään satunnaismuuttujien X_i , $i = 1, 2, \ldots , n$ *aritmeettinen keskiarvo*:

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Tällöin pätee **heikko suurten lukujen laki**: Satunnaismuuttujien $X_1, X_2, X_3, \ldots , X_n$ aritmeettisten keskiarvojen muodostama jono $\bar{X}_n = \Sigma X_i / n$ konvergoi **stokastisesti** kohti satunnaismuuttujien yhteistä odotusarvoa μ:

$$\bar{X}_n \xrightarrow[n\to\infty]{\mathrm{P}} \mu$$

Heikko suurten lukujen laki voidaan ilmaista sanoin seuraavasti: Samoin jakautuneiden satunnaismuuttujien *aritmeettinen keskiarvo lähestyy muuttujien lukumäärän kasvaessa muuttujien yhteistä odotusarvoa* sellaisella tavalla, että (**suurten**) **poikkeamien todennäköisyys satunnaismuuttujien yhteisestä odotusarvosta tulee yhä pienemmäksi eli** (**suuret**) **poikkeamat tulevat yhä *harvinaisemmiksi***.

Suurten lukujen lait

Suurten lukujen lakeja voidaan pitää matemaattisena formulointina **tilastollisen stabiliteetin** käsitteelle.

Koska melkein varma konvergenssi implikoi stokastisen konvergenssin, *vahva suurten lukujen laki implikoi heikon suurten lukujen lain*.

Suurten lukujen laeista on olemassa *yleisempiä muotoja*, joissa voidaan lieventää *samoin-jakautuneisuus*- ja *riippumattomuusoletuksia*.

Keskeinen raja-arvolause

Olkoon

$$X_i \,,\; i = 1, 2, 3, \ldots$$

jono *riippumattomia*, *samaa normaalijakaumaa* $\mathrm{N}(\mu,\sigma^2)$ *noudattavia satunnaismuuttujia*. Tällöin *satunnaismuuttujien* X_i *summa* Y_n on normaalinen:

$$Y_n = \sum_{i=1}^{n} X_i \sim \mathrm{N}(n\mu, n\sigma^2)$$

Kysymys: Mitä voidaan sanoa *riippumattomien*, *samaa jakaumaa noudattavien satunnais-muuttujien summan jakaumasta*, jos ko. satunnaismuuttujat *eivät noudata normaalijakaumaa*?

Vastaus: *Ei-normaalisten* satunnaismuuttujien summa *ei yleensä ole* normaalinen.

Mutta: Jos yhteenlaskettavia on ”tarpeeksi paljon”, *satunnaismuuttujien summa on kuitenkin* (hyvin yleisin ehdoin) **approksimatiivisesti normaalinen**. Tämä on **keskeisen raja-arvolauseen** olennainen sisältö.

Koska monia satunnaismuuttujia voidaan pitää *usean riippumattoman tekijän summana*, antaa keskeinen raja-arvolause selityksen *empiiriselle havainnolle* niiden normaalisuudesta.

Olkoon

$$X_i \,,\; i = 1, 2, 3, \ldots$$

jono *riippumattomia*, *samoin jakautuneita satunnaismuuttujia*, joiden odotusarvo ja varianssi ovat

$$\mathrm{E}(X_i) = \mu \,,\; i = 1,2,3,\ldots$$

$$\mathrm{D}^2(X_i) = \sigma^2 \,,\; i = 1,2,3,\ldots$$

Olkoon

$$Y_n = \sum_{i=1}^{n} X_i$$

satunnaismuuttujien X_i , $i = 1, 2, \ldots , n$ *summa*. Summan Y_n odotusarvo ja varianssi ovat

$$\mathrm{E}(Y_n) = n\mu$$

$$\mathrm{D}^2(Y_n) = n\sigma^2$$

Standardoidaan summa Y_n :

$$Z_n = \frac{Y_n - n\mu}{\sigma\sqrt{n}}$$

Annetaan $n \to \infty$. Tällöin satunnaismuuttujan Z_n *jakauma lähestyy standardoitua normaali-jakaumaa* $\mathrm{N}(0,1)$.

Siten **keskeinen raja-arvolause** sanoo, että

$$\lim_{n\to\infty} \Pr\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \le z \right) = \Phi(z)$$

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

Merkintä:

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \sim_a \mathrm{N}(0,1)$$

Keskeisen raja-arvolauseen mukaan usean satunnaismuuttujan *summa on* (*tietyin ehdoin*) *approksimatiivisesti normaalinen* (*lähes*) *riippumatta yhteenlaskettavien jakaumasta*.

Huomautus:

Yhteenlaskettavien ei tarvitse olla edes *jatkuvia*, vaan ne voivat olla jopa *diskreettejä*.

Approksimaation *hyvyys riippuu* yhteenlaskettavien satunnaismuuttujien lukumäärästä, niiden jakaumasta ja erityisesti niiden jakauman *vinoudesta*. Approksimaation *hyvyys paranee*, kun yhteenlaskettavien satunnaismuuttujien *lukumäärä kasvaa*.

Jos yhteenlaskettavien satunnaismuuttujien jakauma on *symmetrinen*, approksimaatio on hyvä jo suhteellisen pienillä yhteenlaskettavien lukumäärillä. Jos yhteenlaskettavien satunnaismuuttujien jakauma on *epäsymmetrinen*, hyvä approksimaatio vaatii enemmän yhteenlaskettavia.

Keskeisessä raja-arvolauseessa esiintyvä *rajakäyttäytymisen muoto* on esimerkki **jakauma-konvergenssista** eli **heikosta konvergenssista**.

Keskeisestä raja-arvolauseesta on olemassa *yleisempiä muotoja*, joissa lievennetään *samoin-jakautuneisuus*- ja *riippumattomuusoletuksia*.

Keskeisestä raja-arvolauseesta seuraa: Riippumattomien samoin jakautuneiden satunnaismuuttujien X_i , $i = 1, 2, \ldots , n$ **aritmeettinen keskiarvo**

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

on suurille (mutta äärellisille) *n approksimatiivisesti normaalinen parametrein* μ *ja* σ^2/n:

$$\bar{X}_n \sim_a \mathrm{N}\left(\mu, \frac{\sigma^2}{n} \right)$$