Todennäköisyyslaskenta, syksy 2013

Petri Koistinen
Matematiikan ja tilastotieteen laitos
Helsingin yliopisto

30. elokuuta 2013

2

Esipuhe

Tämä luentomoniste on tarkoitettu Helsingin yliopiston matematiikan ja tilastotieteen laitoksella pidettävälle 10 opintopisteen laajuiselle aineopintotasoiselle todennäköisyyslaskennan kurssille. Kurssilla käsitellään sellaisia puolia todennäköisyyslaskennasta, joita jokaisen tilastotieteilijän (tai muun todennäköisyyslaskennan soveltajan) tulisi tuntea. Stokastisiin prosesseihin liittyvät käsitteet ovat kuitenkin rajattu kurssin ulkopuolelle. Lukijalla oletetaan ennestään olevan jonkin verran esitietoja aiheesta, esimerkiksi P. Tuomisen teokseen Todennäköisyyslaskenta I [11] perustuvalta kurssilta Johdatus todennäköisyyslaskentaan (5 op). Tästä syystä kurssin alkuvaiheessa tiettyjä asioita vain kerrataan nopeasti. Joitakin monisteen kohtia joudutaan jättämään väliin ajanpuutteen takia.

Lähestymistapa pyrkii olemaan käytännönläheinen: tavoitteena on opetella satunnaismuuttujiin ja todennäköisyysjakaumiin liittyvää laskentoa pikemminkin kuin käsitellä todennäköisyyslaskentaa matemaattisena struktuurina. Eräs tämän kurssin tärkeimpiä tavoitteita on oppia käsittelemään kaksi- ja useampiulotteisia jakaumia. Tällaiset tiedot ja taidot ovat välttämättömiä esim. sellaiselle opiskelijalle, joka tahtoo ymmärtää tilastotieteen menetelmien perusteita ja lukea alan kirjallisuutta.

Matemaattisesti aukoton esitys vaatisi mittateoriaa ja Lebesguen integraalia, mutta näitä työkaluja ei tällä kurssilla käytetä. Valtaosan todennäköisyyslaskennan sovelluksista pystyy ymmärtämään ilman mitta- ja integrointiteoriaa (mutta tiettyjä sovelluksia ei). Tietyissä teorian kohdissa tässä monisteessa käytetään Lebesguen integraalin ominaisuuksia, koska Riemannin integraalin käyttö johtaisi vaikeuksiin. Joissakin kohdissa luentomonistetta esitetään yhteyksiä mittateoreettisen todennäköisyyslaskennan käsitteisiin. Nämä huomautukset on tarkoitettu niille opiskelijoille, jotka ovat kiinnostuneet matemaattisesta täsmällisyydestä ja kenties opiskelevat myöhemmin todennäköisyysteoriaa; muut opiskelijat voivat sivuuttaa ne kaikessa rauhassa.

Kaikki tässä monisteessa esitettävät asiat löytyvät myös monesta teoreettisen tilastotieteen ja todennäköisyyslaskennan oppikirjasta, jollaisia ovat esimerkiksi teokset [5, 1, 13, 8, 2, 10].

Todennäköisyyslaskennan matemaattisesti täsmällistä mittateoreettista muotoilua voi opiskella paitsi todennäköisyysteorian kurssilla myös lukuisista teoksista, kuten esimerkiksi [6, 4, 12, 7, 3]. Esimerkiksi Schervish [9] näyttää, kuinka tilastotieteen teoria voidaan esittää matemaattisesti täsmällisesti todennäköisyysteoriaan pohjautuen.

Kirjallisuutta

[1] S. F. Arnold. *Mathematical Statistics.* Prentice-Hall, Inc., 1990.

[2] Robert B. Ash. *Basic Probability Theory.* Dover, 2008.

[3] Krishna B. Athreya and Soumendra N. Lahiri. *Measure Theory and Probability Theory.* Springer Texts in Statistics. Springer, 2006.

[4] P. Billingsley. *Probability and Measure.* John Wiley & Sons, Inc., 2nd edition, 1986.

[5] George Casella and Roger L. Berger. *Statistical Inference.* Duxbury Press, 2nd edition, 2002.

[6] G. Elfving and P. Tuominen. *Todennäköisyyslaskenta II.* Limes ry, Helsinki, 1985.

[7] Jean Jacod and Philip Protter. *Probability Essentials.* Springer, 2nd edition, 2002.

[8] Sheldon M. Ross. *Introduction to Probability Models.* Academic Press, 7th edition, 2000.

[9] Mark J. Schervish. *Theory of Statistics*. Springer series in statistics. Springer-Verlag, 1995.

[10] Henk Tijms. *Understanding Probability*. Cambridge University Press, 3rd edition, 2012.

[11] P. Tuominen. *Todennäköisyyslaskenta I*. Limes ry, Helsinki, 1993.

[12] D. Williams. *Probability with Martingales*. Cambridge University Press, 1991.

[13] David Williams. *Weighing the Odds: A Course in Probabiliby and Statistics*. Cambridge University Press, 2001.

Sisältö

1 Tapahtumat ja niiden todennäköisyydet **1**
1.1 Todennäköisyyden tulkintoja . . . 1
1.2 Joukko-oppia . . . 2
1.3 Matemaattinen todennäköisyyden käsite . . . 5
1.4 Kombinatoriikkaa . . . 7
1.5 Multinomikertoimet . . . 9
1.6 Ehdollinen todennäköisyys . . . 10
1.7 Tapahtumien riippumattomuus . . . 12
1.8 Kokonaistodennäköisyys ja Bayesin kaava . . . 14
1.9 Monotoninen jatkuvuus . . . 14

2 Satunnaismuuttuja **16**
2.1 Satunnaismuuttuja ja sen jakauma . . . 16
2.2 Kertymäfunktio . . . 18
2.3 Diskreetti jakauma . . . 20
2.4 Esimerkkejä diskreeteistä jakaumista . . . 22
2.5 Integraalilaskennan kertausta . . . 23
2.6 Jatkuva jakauma . . . 25
2.7 Esimerkkejä jatkuvista jakaumista . . . 29
2.8 Satunnaismuuttujan muunnos . . . 30
2.9 Kvantiilifunktio ja sen käyttö simuloinnissa . . . 31
2.10 Muunnetun satunnaismuuttujan tiheys . . . 34

3 Yhteisjakauma **42**
3.1 Kaksiulotteinen satunnaisvektori ja sen jakauma . . . 42
3.2 Diskreetti kaksiulotteinen jakauma . . . 44
3.3 Useampiulotteinen satunnaisvektori . . . 46
3.4 Satunnaismuuttujien riippumattomuus . . . 47
3.5 Trinomijakauma ja multinomijakauma . . . 50
3.6 Satunnaisvektoreiden riippumattomuus . . . 52

4 Odotusarvo **53**
4.1 Diskreetin satunnaismuuttujan odotusarvo . . . 53
4.2 Jatkuvasti jakautuneen satunnaismuuttujan odotusarvo . . . 57
4.3 Odotusarvon ominaisuuksia . . . 58
4.4 Muunnoksen odotusarvo . . . 59
4.5 Momentit . . . 60
4.6 Varianssi, keskihajonta ja kovarianssi . . . 61

4.7 Momenttiemäfunktio ja kumulanttiemäfunktio 64
4.8 Karakteristinen funktio 66

5 Tärkeitä yksiulotteisia jakaumia **68**
5.1 Diskreettejä jakaumia 68
5.1.1 Binomijakauma 68
5.1.2 Hypergeometrinen jakauma 69
5.1.3 Geometrinen jakauma 69
5.1.4 Negatiivinen binomijakauma 70
5.1.5 Poissonin jakauma 72
5.2 Gamma- ja beetafunktio 73
5.3 Jatkuvia jakaumia 75
5.3.1 Skaalaus ja siirto 75
5.3.2 Tasajakauma 76
5.3.3 Eksponenttijakauma 76
5.3.4 Gammajakauma 76
5.3.5 Beetajakauma 78
5.3.6 Normaalijakauma 78
5.3.7 Normaalijakaumasta johdettuja jakaumia 79

6 Epäyhtälöitä **77**
6.1 Markovin ja Tšebyševin epäyhtälöt 77
6.2 Konveksit funktiot ja Jensenin epäyhtälö 78
6.3 Hölderin epäyhtälö 80
6.4 Cauchyn–Schwarzin epäyhtälö ja korrelaatio 81
6.5 Momenttiemäfunktiota koskevia epäyhtälöitä 82

7 Kaksiulotteinen jakauma **83**
7.1 Jatkuva kaksiulotteinen jakauma 83
7.2 Tasajakauma alueessa 87
7.3 Riippumattomuus 88
7.4 Satunnaisvektorin muunnoksen odotusarvo 90
7.5 Kovarianssi ja muita yhteismomentteja 92
7.6 Paras lineaarinen ennuste 92
7.7 Odotusarvovektori ja kovarianssimatriisi 94
7.8 Tiheysfunktion muuntokaava 95
7.9 t-jakauman ominaisuuksia 100

8 Ehdollinen jakauma **102**
8.1 Ehdolliset jakaumat 102
8.2 Kertolaskusääntö eli ketjusääntö 104
8.3 Diskreetin ja jatkuvan muuttujan yhteisjakauma 105
8.4 Ehdollinen odotusarvo 106
8.5 Yhteisjakauman määrittely hierarkkisesti 109

9 Moniulotteinen jakauma **113**
9.1 Satunnaisvektori 113
9.2 Odotusarvovektori ja kovarianssimatriisi 115
9.3 Ehdolliset jakaumat, kertolaskusääntö ja ehdollinen odotusarvo 118
9.4 Ehdollinen riippumattomuus 120

9.5 Tilastollisia malleja 120
9.6 Tiheysfunktion muuntokaava 123
9.7 Satunnaisvektorin momenttiemäfunktio 125

10 Moniulotteinen normaalijakauma **128**
10.1 Standardinormaalijakauma $N_n(\mathbf{0}, \mathbf{I})$ 128
10.2 Yleinen multinormaalijakauma 129
10.3 Affiinin muunnoksen jakauma 130
10.4 Tiheysfunktio 131
10.5 Tiheysfunktion tasa-arvopinnat 132
10.6 Korreloimattomuus ja riippumattomuus 134
10.7 Ehdolliset jakaumat 136
10.8 Kaksiulotteinen normaalijakauma 137
10.9 Normaalijakauman otoskeskiarvon ja otosvarianssin yhteisjakauma 138

11 Raja-arvolauseita ja approksimaatioita **141**
11.1 Suurten lukujen laki 141
11.2 Jakaumasuppeneminen 142
11.3 Keskeinen raja-arvolause 143
11.4 Normaaliapproksimaatio 144
11.5 Deltamenetelmä 146
11.6 Deltamenetelmän todistus 147

Luku 1

Tapahtumat ja niiden todennäköisyydet

1.1 Todennäköisyyden tulkintoja

Tilastotiedettä, ja tätä kautta todennäköisyyslaskentaa sovelletaan kaikilla tieteen ja useilla elinkeinoelämän aloilla. Jotkin ilmiöt ovat luonteeltaan deterministisiä, mikä tarkoittaa sitä, että niiden kehitys voidaan ennustaa (esim. jonkin luonnonlain avulla), kun tunnetaan ilmiöön liittyvät alkuehdot ja muut relevantit tekijät. Toisaalta on olemassa ilmiöitä, joiden lopputulosta ei osata ennustaa täysin varmasti. Tällöin lopputulokseen liittyvää epävarmuutta voidaan yrittää käsitellä stokastisen mallin (todennäköisyysmallin, tilastollisen mallin) avulla. Todennäköisyyslaskenta on stokastisiin malleihin liittyvää matematiikkaa.

Tarkastelemme esimerkkeinä kahta erilaista ilmiötä, joiden lopputuloksia voidaan kuvailla todennäköisyyksien avulla:

a) nopanheitto,

b) jonkin tietyn vaalin tulokset.

Nopanheiton yhteydessä tavallisesti sovelletaan todennäköisyyden ns. *klassista tulkintaa*, jonka mukaan eri silmäluvut ovat yhtä todennäköisiä nopan fysikaalisen *symmetrian* nojalla. Jos mahdollisten lopputulosten joukko Ω on äärellinen, ja $A \subset \Omega$ on kiinnostuksen kohteena olevan tapahtuman lopputulosten joukko ja todennäköisyydet määritellään symmetrian perusteella, niin tapauksen A todennäköisyydeksi valitaan A:lle suotuisten tapausten lukumäärä jaettuna kaikkien mahdollisten tapausten lukumäärällä, eli

$$P(A) = \frac{\#A}{\#\Omega} = \frac{A\text{:n alkioiden lkm}}{\Omega\text{:n alkioiden lkm}}$$

Nopanheiton tapauksessa mahdolliset lopputulokset (simäluvut) ovat $\Omega = \{1, 2, 3, 4, 5, 6\}$, ja silmäluvun 6 todennäköisyydeksi saadaan ylläolevalla määritelmällä (valitsemalla $A = \{6\}$)

$$P(\text{“kuutonen”}) = \frac{1}{6}.$$

Nopanheittoa on mahdollista toistaa useita kertoja. Yksittäisen nopanheiton lopputulosta on mahdotonta ennustaa, mutta pitkissä toistosarjoissa eri silmälukujen suhteelliset esiintymisfrekvenssit käyttäytyvät säännöllisesti. Nopanheitossa toistot ovat *riippumattomia* siinä mielessä, että edeltävien heittojen tulokset eivät vaikuta seuraavien heittojen tuloksiin. Nopanheittoon voidaan klassisen tulkinnan sijasta soveltaa myös todennäköisyyden *frekvenssitulkintaa*, jonka mukaan tietyn tapahtuman tai lopputuloksen A (esim. “yhdessä nopan heitossa saadaan silmäluku kuusi”) todennäköisyys on tämän tapahtuman esiintymisten suhteellisen frekvenssin raja-arvo, kun riippumattomien toistojen lukumäärä kasvaa rajatta. Ts. tapahtuman A todennäköisyys voidaan yrittää määritellä kaavalla

$$P(A) = \lim_{N \to \infty} \frac{N(A)}{N}, \tag{1.1}$$

jossa $N(A)$ on niiden kokeiden frekvenssi (eli lukumäärä), joissa saatiin lopputulos A, ja N on toistojen lukumäärä. Suure $N(A)/N$ on tapahtuman A suhteellinen esiintymisfrekvenssi.

Frekvenssitulkinnan oikeutus perustuu pohjimmiltaan ns. suurten lukujen lakiin, joka on eräs todennäköisyyslaskennan kuuluisimmista tuloksista. Tämän takia todennäköisyyden käsitettä ei voida määritellä frekvenssitulkinnan avulla. Sitä paitsi tällaista raja-arvoa ei voida koskaan saada reaalimaailman kokeilujen avulla tarkasti selville.

Vaikka frekvenssitulkinta ei kelpaa todennäköisyyden määritelmäksi, silti sitä kannatta kuitenkin yrittää mielessään soveltaa todennäköisyyslaskennan käsitteiden ja erilaisten todennäköisyysmallien ominaisuuksien ymmärtämiseksi. Erityisen hyödyllistä on yrittää simuloida todennäköisyysmallin käyttäytymistä tietokoneella. Tällöin puhtaan stokastisesta simuloinnista tai Monte Carlo -menetelmästä. Tietyn tapahtuman todennäköisyyttä voidaan tällöin arvioida toistamalla simulointia useita kertoja ja laskemalla ko. tapahtuman esiintymien suhteellinen frekvenssi.

Toisin kuin nopanheitto, yksittäiset vaalit ovat ainutkertainen tapahtuma, joten tässä yhteydessä ei voida puhua kokeen toistamisesta. Silti monet tahot ovat valmiita arvioimaan vaalien lopputulosta ennen kuin se on tiedossa. Esimerkiksi vaalituloksista on tavallisesti mahdollista lyödä vetoa jossakin vedonlyöntitoimistossa. Vedonlyöntitoimisto asettaa kertoimensa arvioimalla kyseessä olevien tapahtumien todennäköisyyksiä. Todennäköisyydelle on ainutkertaisten tapahtumien yhteydessä annettava toisenlainen, ns. *subjektiivinen* eli henkilökohtainen tulkinta. Tällöin todennäköisyys kuvaa henkilön (tai muun tahon) uskomuksen astetta väitteen totuuteen hänen käytettävissä olevan tiedon pohjalta. Eri tahot voivat tällöin saada erilaisia numeerisia vastauksia saman tapahtuman todennäköisyydelle. Saman henkilön todennäköisyydet tietylle tapahtumalle voivat olla eri aikoina erilaisia, jos henkilön käytössä oleva informaatio muuttuu.

Joskus tarvitaan vielä edellisistä poikkeavia tapoja tulkita todennäköisyyden käsitettä. Esimerkiksi kvanttimekaniikan yhteydessä tiettävästi edelleen kiistellään siitä, miten kvanttimaailman ilmiöiden todennäköisyydet pitäisi tulkita. Lisäksi kvanttimekaniikan todennäköisyyslaskenta poikkeaa joiltain osin siitä, mitä tällä kurssilla käsitellään.

1.2 Joukko-oppia

Koska todennäköisyyslaskenta perustuu joukko-oppiin, kertaamme varmuuden vuoksi joukko-opin merkintöjä.

Määritelmä 1.1. Olkoon Ω tarkasteltavan satunnaiskokeen kaikkien mahdollisten lopputulosten joukko. Kutsumme joukkoa Ω *perusjoukoksi*, ja sen alkioita $\omega \in \Omega$ *alkeistapauksiksi* (engl. *elementary event*). *Tapahtumiksi* (engl. *event*) kutsumme perusjoukon Ω osajoukkoja.

Huomautus. Usein tilastotieteessä perusjoukolle käytetään nimitystä otosavaruus (engl. *sample space*) sekä esim. merkintää S.

Olkoot nyt A ja B perusjoukon Ω osajoukkoja. Seuraavat merkinnät lienevät tuttuja:

- Merkintä $x \in A$ tarkoittaa, että x on joukon A alkio, eli että x kuuluu joukkoon A. Merkintä $x \notin A$ tarkoittaa, että x ei ole joukon A alkio, eli että x ei kuuluu joukkoon A.
- A^c tai $\bar{A}$ tarkoittaa joukon A komplementtia, eli niitä $x \in \Omega$, jotka eivät ole A:n alkioita.
- $A \subset B$ tarkoittaa, että joukko A on joukon B osajoukko eli että jokainen A:n alkio on myös B:n alkio; tämä asia voidaan merkitä myös $B \supset A$.
- $A = B$ tarkoittaa, että joukot A ja B ovat samoja, eli että niillä on samat alkiot (ts. sitä, että $A \subset B$ ja $B \subset A$).
- $A \cup B = \{x : x \in A \text{ tai } x \in B\}$ on joukkojen A ja B yhdiste.
- $A \cap B = \{x : x \in A \text{ ja } x \in B\}$ on joukkojen A ja B leikkaus.
- $A \setminus B = A \cap B^c = \{x : x \in A \text{ ja } x \notin B\}$ on joukkojen A ja B erotus.
- $\emptyset$ on tyhjä joukko.

Joukko-operaatiota ja niiden ominaisuuksia kannattaa havainnollistaa Vennin diagrammien avulla.

Joukko-operaatioilla on lukuisia tärkeitä algebrallisia ominaisuuksia. Yhdisteellä ja leikkauksella on mm. seuraavat ominaisuudet.

- Kommutatiivisuus (eli vaihdantalait): kaikilla A ja B pätee

$$A \cup B = B \cup A, \qquad A \cap B = B \cap A.$$

- Assosiatiivisuus (eli liitäntälait): kaikilla A, B ja C pätee

$$\begin{aligned}(A \cup B) \cup C &= A \cup (B \cup C) = A \cup B \cup C \\ (A \cap B) \cap C &= A \cap (B \cap C) = A \cap B \cap C.\end{aligned}$$

- Distributiivisuus (eli osittelulait): kaikilla A, B ja C pätee

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C), \qquad (A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

  Osittelulain muistamista auttaa vastaavan säännön palauttaminen mieleen reaalilukujen laskutoimituksille: $(a+b)c = ac+bc$. Joukko-opin osittelulaissa kumpi tahansa operaatioista $\cup$ tai $\cap$ voi olla yhteenlaskun roolissa, jolloin niistä toiselle jää kertolaskun rooli.

De Morganin lait yhdistävät yhdisteen, leikkauksen ja komplementoinnin,

$$(A \cup B)^c = A^c \cap B^c \quad \text{ja} \quad (A \cap B)^c = A^c \cup B^c \qquad \text{kaikille } A \text{ ja } B.$$

De Morganin lait yleistyvät myös usemman kuin kahden joukon yli lasketulle yhdisteelle ja leikkaukselle.

Yhdisteen ja leikkauksen voi laskea myös äärettömän monen joukon kokoelmalle. Jos $A_1, A_2, \ldots$ ovat perusjoukon osajoukkoja, niin niiden yhdistettä merkitään

$$\bigcup_{j=1}^{\infty} A_j = \{x : x \in A_j \text{ jollakin } j\}$$

ja leikkausta merkitään

$$\bigcap_{j=1}^{\infty} A_j = \{x : x \in A_j \text{ kaikilla } j\}.$$

Todennäköisyyslaskennan yhteydessä joukko-operaatioille voidaan antaa havainnollisempia tulkintoja. Tuomisen Todennäköisyys I -teoksesta mukailtu taulukko 1.1 antaa tästä esimerkkejä.

Taulukko 1.1 Eräitä joukko-opin merkintöjen tulkintoja.

Kaava	Tulkinta
Ω	varma tapahtuma
$\emptyset$	mahdoton tapahtuma
$x \in A$	x on A:lle suotuisa alkeistapaus
A	A sattuu
A^c	A ei satu
$A \cup B$	A tai B (tai molemmat) sattuu
$A \cap B$	sekä A ja B sattuvat
$A \cap B = \emptyset$	A ja B ovat erillisiä eli toisensa poissulkevia
$A \subset B$	jos A sattuu, niin myös B sattuu
$A \setminus B$	A sattuu, mutta B ei satu
$\bigcup_{j=1}^{\infty} A_j$	ainakin yksi tapahtumista $A_j, j \geq 1$ sattuu
$\bigcap_{j=1}^{\infty} A_j$	kaikki tapahtumat $A_j, j \geq 1$ sattuvat

Esimerkki 1.1. Perusjoukko eräissä sovelluksissa. Reaalimaailman satunnaisilmiön mallintamiseksi perusjoukko voidaan valita muuten vapaasti, mutta sen pitää olla riittävän rikas, jotta kaikki kiinnostuksen kohteena olevat tapahtumat voidaan esittää sen osajoukkoina. Tietty sovellus voidaan käsitellä käyttämällä useita erilaisia valintoja.

- Nopanheitossa tulos on jokin silmäluvuista 1, 2, . . . , 6. Perusjoukoksi voidaan valita suoraviivaisesti

  $$\Omega = \{1, 2, 3, 4, 5, 6\}$$

  siten, että alkeistapaus kertoo suoraan saadun silmäluvun.

- Kahden nopan heitossa voidaan alkeistapaukset esittää kokonaislukupareilla (i, j), jossa ensimmäinen koordinaatti i kertoo ensimmäisen nopan tuloksen ja toinen koordinaatti j toisen nopan tuloksen. Siis voidaan valita

  $$\Omega = \{(i, j) : 1 \leq i \leq 6, 1 \leq j \leq 6\}.$$

  Tämä valinta kelpaisi myös yhden nopan heiton kuvailuun; meidän tarvitsisi vain jättää laskuissa jälkimmäinen koordinaatti huomioimatta.

- Matemaattisissa ohjelmistoissa on yleensä ns. satunnaislukugeneraattori, joka poimii satunnaisesti reaaliluvun väliltä $(0, 1)$. Tällöin luonteva perusjoukko on kyseinen lukuväli, $\Omega = (0, 1)$. Toisin kuin edellisissä sovelluksissa, tällä kertaa perusjoukko ei ole äärellinen, vaan on ääretön ja sisältää ylinumeroituvan monta alkiota.
- Eräs tärkeä stokastisten prosessien tyyppi on sellainen, jonka jokainen realisaatio on välillä $(0, \infty)$ määritelty jatkuva funktio. Tällöin luonteva perusjoukko on kaikkien välillä $(0, \infty)$ määriteltyjen jatkuvien funktioiden muodostama joukko. Tällaisen perusjoukon käsittely on huomattavasti monimutkaisempaa kuin minkään aikaisemmin esitetyn perusjoukon käsittely, mutta tähän asiaan emme tällä kurssilla puutu.

$\triangle$

1.3 Matemaattinen todennäköisyyden käsite

Johdantoluvussa tapasimme erilaisia todennäköisyyden tulkintoja. Matemaattinen todennäköisyyden käsite ei ota kantaa siihen, miten sitä tulkitaan eli miten sitä sovelletaan reaalimaailman ilmiöiden kuvailuun. Sen sijaan ajatuksena on määritellä aksioomien avulla, minkälaisia ominaisuuksia todennäköisyydellä on. Nämä todennäköisyyden aksioomat esitti A. N. Kolmogorov 1930-luvulla

Ennen aksioomien läpikäyntiä teemme seuraavan määritelmän. Tapahtumat $A_1, A_2, \ldots$ ovat *erillisiä* (engl. *disjoint*) eli *toisensa poissulkevia* (engl. *mutually exclusive*), mikäli

$$A_i \cap A_j = \emptyset \quad \text{kun} \quad i \neq j.$$

Todennäköisyys(mitta) (engl. *probability measure*) P liittää perusjoukon Ω tapahtumiin $A \subset \Omega$ reaaliluvun $P(A)$, jota kutsutaan tapahtuman A todennäköisyydeksi. Todennäköisyydeltä vaaditaan seuraavat ominaisuudet.

Määritelmä 1.2. P on todennäköisyys (lyh. tn) eli todennäköisyysmitta (lyh. tn-mitta), mikäli se toteuttaa seuraavat kolme ominaisuutta.

(1) $P(A) \geq 0$ kaikille tapahtumille $A \subset \Omega$,

(2) $P(\emptyset) = 0$ ja $P(\Omega) = 1$,

(3) (täysadditiivusuus) $P(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} P(A_j)$, kun $A_1, A_2, \ldots$ ovat erillisiä tapahtumia.

Jos A ja B ovat erillisiä tapahtumia (ts. $A \cap B = \emptyset$), niin täysadditiivisuuden aksioomasta seuraa (valinnoilla $A_1 = A$, $A_2 = B$ ja $A_j = \emptyset$, kun $j \geq 3$) se, että

$$P(A \cup B) = P(A) + P(B) \quad \text{kun} \quad A \cap B = \emptyset. \tag{1.2}$$

Tätä ominaisuutta kutsutaan tn-mitan (äärelliseksi) *additiivisuudeksi*. Vastaavasti nähdään, että mikäli $A_1, \ldots, A_n$ ovat erillisiä tapahtumia, niin

$$P(A_1 \cup \cdots \cup A_n) = P(A_1) + \cdots + P(A_n).$$

Tämän tosiseikan voi johtaa paitsi täysadditiivisuuden aksioomasta myös induktiolla additiivisuudesta kahdelle erilliselle tapahtumalle (1.2). Sen sijaan täysadditiivisuutta ei ole mahdollista johtaa induktiolla äärellisestä additiivisuudesta (1.2). Pysähdymme nyt pohtimaan todennäköisyyden frekvenssitulkinnan valossa, miksi kaavassa (1.2) formalisoitu (äärellinen) additiivisuus on luonnollinen todennäköisyyden ominaisuus.

Olkoot A ja B tietyn satunnaiskokeen erillisiä tapahtumia. Oletetaan, että kyseistä satunnaiskoetta toistetaan riippumattomasti N kertaa. (Emme tässä vaiheessa määrittele, mitä tällainen riippumaton toistaminen oikeastaan tarkoittaa, vaan pohdinta pitää yrittää ymmärtää intuitioon nojautuen. Riippumattomissa toistoissa edeltävien kokeiden tulokset eivät vaikuta seuraavan kokeen tulokseen.) Sovitaan, että $N(E)$ tarkoittaa niiden kokeiden lukumäärää eli frekvenssiä, joissa tapahtuma E sattuu. Koska A ja B ovat erillisiä, niin $N(A \cup B) = N(A) + N(B)$, sillä kussakin toistossa korkeintaan yksi tapahtumista A tai B voi sattua. Tämän takia

$$\frac{N(A \cup B)}{N} = \frac{N(A)}{N} + \frac{N(B)}{N},$$

joten additiivisuus pätee suhteellisille frekvensseille. Rajankäynnin jälkeen additiivisuuden pitää päteä myös todennäköisyyksille.

Sen sijaan täysadditiiviisuus (aksiooma (3)) ei ole intuitiivisesti selvä ominaisuus. Se oletetaan matemaattisen mukavuudenhalun takia. Tällä tavalla menetellen saadaan matemaattisesti kaunis teoria, jonka puitteissa voidaan formuloida ja todistaa esim. äärettömiä pitkiä tapahtumajonoja koskevia lauseita (kuten vaikkapa vahva suurten lukujen lause).

Todennäköisyyden aksioomeista seuraa yksinkertaisilla laskuilla monia sen ominaisuuksia, kuten esimerkiksi seuraavassa lauseessa todetut ominaisuudet. Näistä osa jätetään harjoitustehtäviksi. Ideana todistuksissa on jakaa sopivasti valittu tapahtuma erillisiin tapahtumiin, ja sitten soveltaa (äärellistä) additiivisuutta (1.2) sekä muita todennäköisyyden ominaisuuksia.

Lause 1.1. *Todennäköisyydellä on seuraavat ominaisuudet.*

a) (Äärellinen additiivisuus) Jos $A \cap B = \emptyset$, *niin* $P(A \cup B) = P(A) + P(B)$.

b) $P(A^c) = 1 - P(A)$ *kaikille* A.

c) $0 \le P(A) \le 1$ *kaikille* A.

d) (Tn-mitan monotonisuus) Jos $A \subset B$, *niin* $P(A) \le P(B)$.

e) $P(A \setminus B) = P(A) - P(A \cap B)$ *kaikille* A *ja* B.

f) (Yhteenlaskukaava) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ *kaikille* A *ja* B.

g) (Boolen epäyhtälö) $P(A \cup B) \le P(A) + P(B)$ *kaikille* A *ja* B.

h) (Bonferronin epäyhtälö) $P(A \cap B) \ge P(A) + P(B) - 1$ *kaikille* A *ja* B.

Todistus. Äärellinen additiivisuus (a-kohta) todistettiin kaavassa (1.2). Todistetaan esimerkin vuoksi kohdat b) ja c). Jos A on tapahtuma, niin $\Omega = A \cup A^c$, jossa osat ovat erillisiä. Additiivisuuden ja aksiooman (2) nojalla

$$1 = P(A) + P(A^c),$$

joten b-kohta on todistettu. Koska A^c on tapahtuma, niin $P(A^c) \ge 0$ aksiooman (1) nojalla, joten

$$P(A) = 1 - P(A^c) \le 1,$$

mikä yhdessä aksiooman (1) kanssa todistaa c-kohdan. □

Huomautus. $0 \le P(A) \le 1$ kaikille tapahtumille A. Jos saat jonkin tapahtuman todennäköisyydeksi luvun, joka ei ole välillä $[0, 1]$, olet tehnyt laskuvirheen.

Yhdessä perusjoukkoa Ω, sen tapahtumia eli sitä Ω:n osajoukkojen kokoelmaa, jolla tn P on määritelty, sekä tn-mittaa P kutsutaan *todennäköisyysavaruudeksi* (tai todennäköisyyskentäksi).

Täydentävä huomautus

Jos perusjoukko Ω on äärellinen tai korkeintaan numeroituvasti ääretön, niin tavallisesti sallitaan tapahtumaksi mikä tahansa perusjoukon osajoukko. Jos kuitenkin Ω on suurempi kuin numeroituvasti ääretön, niin todennäköisyyden aksioomia ei saada toteutumaan kaikille sen osajoukoille, vaan joukkofunktion P määrittelyjoukkoa pitää rajoittaa. Kaikkien tapahtumien joukolla eli todennäköisyyden P määrittelyjoukolla kuuluu olla se ominaisuus, että se on ns. σ-algebra. Tästä seuraa se, että jos lähdetään likkeelle korkeintaan numeroituvasta määrästä tapahtumia, ja sovelletaan niihin korkeintaan numeroituva määrä joukko-operaatiota, niin tulokseksi saadaan tapahtuma. Tällä kurssilla sivuutamme σ-algebroihin liittyvät tarkastelut, ja **oletamme, että todennäköisyys on määritelty kaikille meitä kiinostaville perusjoukon osajoukoille**.

1.4 Kombinatoriikkaa

Lantinheiton, nopanheiton, korttipelien, loton ja muiden tällaisten uhkapelin stokastisena mallina pidetään — mikäli ei kerrota tarkempia tietoja — klassista todennäköisyyden tulkintaa, jonka mukaan

$$P(A) = \frac{\#A}{\#\Omega}. \tag{1.3}$$

Tapahtuman A todennäköisyys on tällöin tapahtumalle suotuisten alkeistapauksien lukumäärän $\#A$ (ts. joukon A alkioiden lukumäärän) ja kaikkien mahdollisten alkeistapauksien lukumäärän $\#\Omega$ (ts. perusjoukon Ω alkioiden lukumäärän) osamäärä. Tällainen valinta on mahdollista tietenkin vain silloin, kun perusjoukko on äärellinen. Alkeistapaukset $\omega \in \Omega$ ovat tällöin siinä mielessä symmetrisiä, että niillä on kaikilla sama todennäköisyys, eli

$$P(\{\omega\}) = \frac{1}{\#\Omega}, \qquad \text{kaikilla } \omega \in \Omega.$$

Tällaiseen malliin viitataan usein termillä *klassinen todennäköisyys*. Vaihtoehtoisesti voidaan puhua *tasaisesta todennäköisyysmallista* tai esim. *symmetrisestä todennäköisyysavaruudesta*. Kun käsitellään klassista todennäköisyyttä, niin stokastinen malli on määritelty heti, kun perusjoukko on valittu. Perusjoukko Ω pitää tietenkin valita järkevästi, jotta sen alkiot ja reaalimaailman symmetriat vastaavat toisiaan. Jotta osaisimme laskea todennäköisyyksiä, meidän pitää osata laskea eräitä kombinatorisia laskuja, jotka lienevät lukijalle pääosin jo ennestään tuttuja.

Tarkastellaan n-alkioista joukkoa E,

$$E = \{e_1, e_2, \ldots, e_n\}.$$

Kuinka monella eri tavalla voidaan poimia k alkiota n-alkoisesta joukosta E?
Saamme tähän kysymykseen erilaisia vastauksia sen mukaan, valitaanko alkiot

- *takaisinpanolla* (eli palauttaen, engl. *with replacement*), jolloin sama alkio voidaan valita useaan kertaan, ja k voi olla suurempi kuin n;
- *ilman takaisinpanoa* (eli takaisinpanotta eli palauttamatta, engl. *without replacement*), jolloin valitaan $k \le n$ eri alkiota.

Termi "takaisinpano" viittaa tilanteeseen, jossa esim. arpalippuja poimitaan arvontauurnasta umpimähkään yksi kerrallaan. Otannassa takaisinpanolla kukin arpalippu palautetaan, eli pannaan takaisin uurnaan ennen seuraavan arvan poimintaa. Otannassa ilman takaisinpanoa arpalippua ei laiteta takaisin ennen seuravan arpalipun poimintaa.

Lisäksi pitää kertoa, onko poimintajärjestyksellä väliä vai ei.

- Jos *poimintajärjestyksellä on väliä*, niin valinta esitetään (järjestettynä) jonona.
- Jos *poimintajärjestyksellä ei ole väliä*, niin sellaiset kaksi valintaa ovat ekvivalentteja, joissa on poimittu samat alkiot (ja kukin alkio on molemmissa valinnoissa poimittu yhtä monta kertaa).

Pysytymme laskemaan lukumäärät eri tilanteissa ns. kombinatoriikan tuloperiaatteen avulla.

Määritelmä 1.3 (Kombinatoriikan tuloperiaate)**.** Tarkastellaan tehtävää, joka voidaan jakaa k:hon osatehtävään, joilla on seuraavat ominaisuudet. Ensimmäinen osatehtävä voidaan suorittaa n_1:llä eri tavalla. Kun ensimmäinen osatehtävä on suoritettu, toinen osatehtävä voidaan suorittaa n_2:lla eri tavalla (riippumatta siitä, mikä valinta tehtiin ensimmäisessä vaiheessa) jne. ja viimein, kun $k-1$ ensimmäistä osatehtävää on suoritettu, k:s osatehtävä voidaan suorittaa n_k eri tavalla (riippumatta siitä, mitkä valinnat tehtiin aikaisemmissa vaiheissa). Mikäli koko tehtävän ratkaisu saadaan koottua yksikäsitteisellä tavalla osatehtävien ratkaisuista, niin tällöin koko tehtävä voidaan suorittaa

$$n_1 n_2 \cdots n_k = \prod_{j=1}^{k} n_j$$

eri tavalla.

Poiminta takaisinpanolla, kun poimintajärjestyksellä on väliä: n-alkioisesta joukosta E voidaan muodostaa

$$n^k$$

erilaista k-alkioista jonoa $(x_1, \ldots, x_k)$ takaisinpanolla. Tämä lukumäärä on sama kuin joukon E k-kertaisen karteesisen tulon

$$E \times \cdots \times E \qquad (k \text{ tekijää})$$

kardinaliteetti. Kunkin koordinaatin x_i arvolle on n eri mahdollisuutta riippumatta muiden koordinaattien arvoista, joten lukumäärä n^k saadaan laskettua tuloperiaatteella. Tällaisten jonojen muodostama joukko ja vastaava tasainen todennäköisyysmalli on luonteva malli esim. k:lle lantin tai nopan heitolle: lantinheitossa joukolla E on $n=2$ alkiota, ja nopanheitossa $n=6$.

Poiminta ilman takaisinpanoa, kun poimintajärjestyksellä on väliä: n-alkioisesta joukosta E voidaan muodostaa

$$n(n-1)\cdots(n-k+1).$$

erilaista k-jonoa ($k \le n$), kun kaikkien jonon alkioiden pitää olla erisuuria. Tällaisiin jonoihin viitataan termillä E:n *k-permutaatio* (tai termillä k-variaatio, kuten Tuomisen kirjassa). Ensimmäinen jonon alkio voidaan valita n erilaisella tavalla. Tämän jälkeen toinen alkio voidaan valita $(n-1)$:llä eri tavalla, sillä toisen alkion pitää olla eri suuri kuin ensimmäisen. Tekemällä tällä tavoin peräkkäin k valintaa, joissa aina edelliset valinnat on suljettu pois päädytään yllä olevaan kaavaan.

Muistathan n-kertoman $n!$ määritelmän kokonaisluvulle $n \ge 0$:

$$n! = n \cdot (n-1) \cdots 2 \cdot 1, \quad \text{kun } n \ge 1, \text{ ja} \qquad 0! = 1. \tag{1.4}$$

Kertoma $n!$ ilmaisee n-alkoisen joukon n-permutaatioiden, eli lyhyesti permutaatioiden lukumäärän: n alkiota voidaan järjestää $n!$ erilaisella tavalla. Kertoman avulla kirjoitettuna n-alkioisen joukon k-permutaatioita on

$$\frac{n!}{(n-k)!} = n(n-1)\cdots(n-k+1).$$

Edellä tehty sopimus $0! = 1$ on tärkeä monien kaavojen yhteydessä. Esimerkiksi, jos mikäli laskemme n-alkioisen joukon n-permutaatioiden lukumäärän sijoittamalla edelliseen kaavaan $k = n$, niin kaavan vasemmalle puolelle saamme $n!/(n-n)! = n!/0! = n!$.

Poiminta ilman takaisinpanoa, kun poimintajärjestyksellä ei ole väliä: kun n-alkioisesta joukosta E valitaan k alkiota ilman takaisinpanoa, niin tämä on sama asia kuin että valitaan joukon E k-alkioinen osajoukko. Näitä k-osajoukkoja kutsutaan myös joukon E *k-kombinaatioiksi*. Niiden lukumäärän ilmaisee binomikerroin n yli k:n,

$$\begin{aligned}\binom{n}{k} &= \begin{cases} \dfrac{n!}{k!(n-k)!}, & \text{kun } 0 \le k \le n, \\ 0, & \text{muuten.} \end{cases} \\ &= \begin{cases} \dfrac{n(n-1)\cdots(n-k+1)}{k!}, & \text{kun } 0 \le k \le n, \\ 0, & \text{muuten.} \end{cases}\end{aligned} \tag{1.5}$$

Tämä johtuu siitä, että kukin tällainen osajoukko voidaan järjestää $k!$ eri tavalla k-jonoksi, jolloin saadaan generoitua kertaalleen kaikki joukon E k-permutaatiot.

Poiminta takaisinpanolla, kun poimintajärjestyksellä ei ole väliä: sivuutetaan tällä kurssilla.

Binomikertoimet esiintyvät myös ns. *binomikaavassa*, jonka mukaan

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k\, b^{n-k}, \tag{1.6}$$

kun a ja b ovat reaalilukuja ja $n \ge 0$ on kokonaisluku. Binomikaava seuraa edellä selostetusta binomikertoimien kombinatorisesta luonnehdinnasta sekä reaalilukujen laskutoimitusten ominaisuuksista.

1.5 Multinomikertoimet

Eräs tulkinta binomikertoimelle $\binom{n}{k}$ on, että se kertoo, kuinka monella tavalla n-alkioinen joukko voidaan osittaa kahteen osajoukkoon siten, että ensimmäiseen osaan tulee k alkiota ja toiseen osaan $n-k$ alkiota.

Kuinka monta vaihtoehtoa on, jos n-alkioinen joukko ositetaan kolmeen osajoukkoon siten, että ensimmäiseen osaan tulee k_1 alkioita, toiseen osaan k_2 alkioita ja kolmanteen osaan k_3 alkiota? Jotta kysymys olisi mielekäs, pitää tietenkin olettaa, että kukin $0 \le k_i \le n$, ja että

$$k_1 + k_2 + k_3 = n.$$

Samme vastattua kysymykseen helposti tuloperiatteen avulla. Voimme valita ensimmäisen osajoukon alkiot $\binom{n}{k_1}$ tavalla, ja sen jälkeen toisen osan alkiot $\binom{n-k_1}{k_2}$ tavalla, jonka jälkeen kaikki jäljelle jääneet alkiot kuuluvat kolmanteen osaan. Näin ollen eri mahdollisuuksia on yhteensä

$$\binom{n}{k_1}\binom{n-k_1}{k_2} = \frac{n!}{k_1!(n-k_1)!}\frac{(n-k_1)!}{k_2!(n-k_1-k_2)!} = \frac{n!}{k_1!k_2!k_3!}.$$

Tälle luvulle käytetään seuraavia merkintöjä,

$$\binom{n}{k_1 k_2 k_3} = \binom{n}{k_1, k_2, k_3} = \frac{n!}{k_1!k_2!k_3!},$$

eli toisinaan alakerrassa olevat luvut erotetaan toisistaan selvyyden vuoksi pilkuilla, ja toisinaan taas pilkkuja ei käytetä. Tälle luvulle käytetään nimitystä trinomikerroin (tai multinomikerroin).

Yleistetään edellinen kysymys m:lle osalle. Olkoon annettuna luvut $k_1, \dots, k_m$ siten, että kukin $0 \le k_i \le n$, ja

$$k_1 + \cdots + k_m = n.$$

Miten monella eri tavalla n-alkioinen joukko voidaan osittaa m osaan $A_1, \dots, A_m$ siten, että kussakin osassa A_i on k_i alkiota? Vastaavalla tavalla päättelemällä kuin edellä tapauksessa $m = 3$ nähdään, että eri mahdollisuuksia on

$$\frac{n!}{k_1!\, k_2! \cdots k_m!} = \binom{n}{k_1, k_2, \dots, k_m} \tag{1.7}$$

kappaletta. Näitä lukuja kutsutaan multinomikertoimiksi. Binomikerroin on tietenkin multinomikertoimen erikoistapaus, sillä

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

Multinomikertoimet esiintyvät ns. *multinomikaavassa*, jonka mukaan

$$(a_1 + a_2 + \cdots + a_m)^n = \sum \binom{n}{k_1, k_2, \dots, k_m} a_1^{k_1} a_2^{k_2} \dots a_m^{k_m}. \tag{1.8}$$

Tässä kaavassa summataan yli kaikkien sellaisten ei-negatiivisten kokonaislukujen $k_1, k_2, \dots, k_m$, joiden summa on n. Binomikaava (1.6) on multinomikaavan erikoistapaus.

1.6 Ehdollinen todennäköisyys

Toisinaan saamme tapahtuneesta osittaista informaatiota, ja tämä informaatio voidaan esittää siinä muodossa, että tapahtuma B on sattunut, jossa $P(B) > 0$. Kun tämä tieto otetaan huomioon, niin jonkin muun tapahtuman A todennäköisyytenä ei enää pidetä sen alkuperäistä todennäköisyyttä $P(A)$ vaan sen sijaan lasketaan *tapahtuman A ehdollinen todennäköisyys ehdolla B*, joka määritellään kaavalla

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \qquad \text{jossa } P(B) > 0. \tag{1.9}$$

Tämä on tapahtuman A todennäköisyys, kun otetaan huomioon (täsmälleen se) että B on sattunut.

Esimerkki 1.2. *Kysymys:* Heitetään kerran noppaa, ja silmäluku on parillinen. Millä todennäköisyydellä silmäluku on a) yksi b) kaksi?

Vastaus maalaisjärkeä soveltamalla: Koska silmäluku on parillinen, niin se ei tietenkään voi olla yksi, joten vastaus a-kohtaan on nolla; koska silmäluku on parillinen, niin vaihtoehdot 2, 4 ja 6 ovat yhtä todennäköisiä, joten vastaus b-kohtaan on 1/3.

Vastaus ehdollisen todennäköisyyden kaavaa soveltamalla: Olkoon perusjoukkona $\{1, \dots, 6\}$ siten, että perusjoukon alkio kertoo nopan silmäluvun. Tällöin a-kohdan todennäköisyys on

$$P(\{1\} \mid \{2,4,6\}) = \frac{P(\{1\} \cap \{2,4,6\})}{P(\{2,4,6\})} = \frac{P(\emptyset)}{3/6} = 0$$

ja b-kohdan todennäköisyys on

$$P(\{2\} \mid \{2,4,6\}) = \frac{P(\{2\} \cap \{2,4,6\})}{P(\{2,4,6\})} = \frac{1/6}{3/6} = \frac{1}{3}.$$

△

Ehdollisen todennäköisyyden $P(A \mid B)$ kaavassa (1.9) otetaan uudeksi perusjoukoksi B, jonka ehdolliseksi todennäköisyydeksi $P(B \mid B)$ pitää saada yksi. Tämä saadaan aikaan normittamalla alkuperäinen todennäköisyys: se jaetaan ehdon B todennäköisyydellä. Koska B on sattunut, niin kaavassa lasketaan ei-ehdollisia todennäköisyyksiä vain joukon B osajoukoille $A \cap B$.

Ehdollisen todennäköisyyden kaavan (1.9) voi motivoida myös sen frekvenssitulkinnan avulla. Tarkastellaan N-kertaista toistokoetta, jossa voi sattua A tai B tai molemmat. Tapahtuman A suhteellinen esiintymisfrekvenssi niissä kokeissa, joissa on sattunut tapahtuma B on

$$\frac{N(A \cap B)}{N(B)} = \frac{N(A \cap B)/N}{N(B)/N}.$$

Frekvenssitulkinnan mukaan jälkimmäisessä muodossa osoittaja lähestyy lukua $P(A \cap B)$ ja nimittäjä lukua $P(B)$, kun N kasvaa rajatta. Siis ehdollisen todennäköisyyden $P(A \mid B)$ määritelmä osamääränä $P(A \cap B)/P(B)$ on järkeenkäypä.

"Alkuperäiset"eli ei-ehdolliset todennäköisyydet $P(A)$ voidaan ymmärtää ehdollisina todennäköisyyksinä ehdolla Ω,

$$P(A \mid \Omega) = \frac{P(A \cap \Omega)}{P(\Omega)} = \frac{P(A)}{1}.$$

On lisäksi mahdollista tarkistaa, että ehdollinen todennäköisyys $P(A \mid B)$ on argumentin A funktiona todennäköisyys, eli että se toteuttaa todennäköisyyden kolme aksioomaa. Tästä seuraa, että **ehdolliselle todennäköisyydelle saadaan käyttää samoja laskusääntöjä (pystyviivan vasemmalla puolella esiintyvän argumentin suhteen), kuin tavalliselle todennäköisyydelle**.

Ehdollisen todennäköisyyden määritelmä (1.9) voidaan kirjoittaa myös tulomuodossa,

$$P(A \cap B) = P(A \mid B)\, P(B). \tag{1.10}$$

Tämä erittäin tärkeä kaava on nimeltään todennäköisyyksien *kertolaskusääntö* eli *kertolaskukaava*. Siitä käytetään myös nimeä todennäköisyyslaskennan *ketjusääntö*.

Mikäli $P(A) > 0$ ja $P(B) > 0$, on

$$P(A \cap B) = P(A \mid B)\, P(B) = P(A)\, P(B \mid A),$$

josta saadaan ratkaistua toinen ehdollisista todennäköisyyksistä, nimittäin

$$P(B \mid A) = \frac{P(A \mid B)\, P(B)}{P(A)}.$$

Kertolaskusääntö (1.10) voidaan yleistää myös usemmalle tapahtumalle. Tarkastellaan esimerkiksi kolmea tapahtumaa A, B ja C, ja oletetaan, että $P(A \cap B) > 0$. Tällöin myös $P(A) > 0$, ja

$$\begin{aligned} P(A)\, P(B \mid A)\, P(C \mid A \cap B) &= P(A)\, \frac{P(A \cap B)}{P(A)}\, \frac{P(A \cap B \cap C)}{P(A \cap B)} \\ &= P(A \cap B \cap C). \end{aligned}$$

Jos tapahtumia on n kappaletta $A_1, \dots, A_n$ ja $P(A_1 \cap \dots \cap A_{n-1}) > 0$, niin niiden leikkauksen todennäköisyys saadaan kertolaskusäännöllä

$$\begin{aligned} &P(A_1 \cap \dots \cap A_n) \\ &\quad = P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 \cap A_2) \cdots P(A_n \mid A_1 \cap \dots \cap A_{n-1}). \end{aligned} \tag{1.11}$$

Esimerkki 1.3. Perusteellisesti sekoitetusta 52 kortin pakasta jaetaan kaksi korttia. Tarkastellaan tapahtumia

$$A_1 = \{\text{“1. kortti on ässä”}\}, \quad A_2 = \{\text{“2. kortti on ässä”}\}.$$

Laske todennäköisyydet $P(A_1)$, $P(A_2)$, $P(A_1 \cap A_2)$ ja $P(A_2 \mid A_1)$.

Ratkaisu 1. Valitaan perusjoukoksi 52 kortin 2-permutaatiot, joita on $52 \cdot 51$ kappaletta. Tuloperiaatteen nojalla

$$\#A_1 = \#A_2 = 4 \cdot 51, \qquad \#(A_1 \cap A_2) = 4 \cdot 3.$$

Siis

$$P(A_1) = P(A_2) = \frac{4 \cdot 51}{52 \cdot 51} = \frac{1}{13}, \qquad P(A_1 \cap A_2) = \frac{4 \cdot 3}{52 \cdot 51},$$

ja näistä saadaan

$$P(A_2 \mid A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)} = \frac{4 \cdot 3}{4 \cdot 51} = \frac{1}{17}.$$

Ratkaisu 2. Ajatellaan tilannetta kortti kerrallaan. Todennäköisyys, että ensimmäinen kortti on ässä on tietenkin $4/52$, koska ässiä on pakassa neljä. Kun tarkastellaan vain yhden kortin arvoa kerrallaan, on samantekevää, monentenako se jaetaan, joten $4/52$ on myös tn sille, että toinen kortti on ässä. Siis

$$P(A_1) = P(A_2) = \frac{4}{52} = \frac{1}{13}.$$

Kun ensimmäinen kortti on jaettu ja osoittautunut ässäksi, niin jäljellä on perusteellisesti sekoitettu 51 kortin pakka, joka sisältää 3 ässää. Tämän takia

$$P(A_2 \mid A_1) = \frac{3}{51}.$$

Kertolaskusäännön nojalla

$$P(A_1 \cap A_2) = P(A_1)\, P(A_2 \mid A_1) = \frac{4}{52}\frac{3}{51}$$

Huomaa, että ratkaisussa 2 emme lainkaan ottaneet kantaa siihen, mitä pidämme perusjoukkona Ω. Sen sijaan laskimme käyttämällä (ehdollisia) todennäköisyyksiä, joiden lukuarvot ovat tässä tehtävässä intuitiivisesti selviä. $\triangle$

1.7 Tapahtumien riippumattomuus

Määritelmä 1.4 (Kahden tapahtuman riippumattomuus)**.** Tapahtumat A ja B ovat riippumattomia, jos

$$P(A \cap B) = P(A)\, P(B).$$

Tämä asia voidaan merkitä $A \perp\!\!\!\perp B$.

Huomautus. Tarkista, että ymmärrät, mitä eroa on seuraavilla käsitteillä.

a) A ja B ovat erillisiä tapahtumia,

b) A ja B ovat riippumattomia tapahtumia.

Jos A ja B ovat riippumattomia, niin tapauksen $A \cap B$ todennäköisyys saadaan kertomalla keskenään kyseisten tapauksien ei-ehdolliset todennäköisyydet. (Yleisemmässä tapauksessa tarvitaan kertolaskusääntöä (1.10).)

Jos $P(B) > 0$ ja $A \perp\!\!\!\perp B$, niin

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)\,P(B)}{P(B)} = P(A).$$

Siis mikäli A ja B ovat riippumattomia, niin tieto B:n sattumisesta ei muuta A:n todennäköisyyttä. Toisaalta, jos $P(B) > 0$ ja $P(A \mid B) = P(A)$, niin $A \perp\!\!\!\perp B$, sillä tällöin

$$P(A \cap B) = P(B \cap A) = P(B)\,P(A \mid B) = P(A)\,P(B).$$

Yhteenveto edellisistä tarkasteluista: jos $P(B) > 0$, on

$$A \perp\!\!\!\perp B \quad \Longleftrightarrow \quad P(A \mid B) = P(A). \tag{1.12}$$

Esimerkki 1.4. Kahden nopan heitto. Perusjoukko on $\Omega = E \times E$, jossa $E = \{1, 2, 3, 4, 5, 6\}$. Sen alkeistapaukset ovat symmetrisiä, ja alkeistapauksessa $(x, y) \in E \times E$ ensimmäinen koordinaatti x kertoo ensimmäisen nopan silmäluvun ja toinen koordinaatti y toisen nopan silmäluvun. Olkoot $1 \le i, j \le 6$ annettuja kokonaislukuja, ja tarkastellaan tapahtumia

$$A = \{\text{“nopan yksi tulos on } i\text{”}\}, \quad B = \{\text{“nopan kaksi tulos on } j\text{”}\}$$

Tällöin

$$P(A) = \frac{6}{36}, \quad P(B) = \frac{6}{36},$$

ja

$$P(A \cap B) = \frac{1}{36} = \frac{1}{6}\,\frac{1}{6} = P(A)\,P(B),$$

joten $A \perp\!\!\!\perp B$. △

Määritelmä 1.5 (Useamman tapahtuman riippumattomuus)**.** Tapahtumat $A_1, \ldots, A_n$ ovat riippumattomia, mikä asia voidaan merkitä $A_1, \ldots, A_n \perp\!\!\!\perp$, jos kaikilla $k \ge 2$ on voimassa

$$P(A_{i_1} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k})$$

aina kun $i_1, \ldots, i_k$ ovat keskenään erisuuria kokonaislukuja, jotka kaikki ovat välillä $1 \le i_j \le n$.

Esimerkiksi kolme tapahtumaa A, B ja C ovat riippumattomia, jos kaikki seuraavat neljä ehtoa pätevät,

$$\begin{aligned}
P(A \cap B) &= P(A)\,P(B) \\
P(A \cap C) &= P(A)\,P(C) \\
P(B \cap C) &= P(B)\,P(C) \\
P(A \cap B \cap C) &= P(A)\,P(B)\,P(C).
\end{aligned}$$

Tässä neljäs ehto ei seuraa kolmesta ensimmäisestä, eli tapahtumien A, B ja C parittaisesta riippumattomuudesta, mikä voidaan osoittaa yksinkertaisella esimerkillä.

Määritelmä 1.6 (Ehdollinen riippumattomuus)**.** Olkoon C tapahtuma, jolle $P(C) > 0$. Tapahtumat A ja B ovat ehdollisesti riippumattomia ehdolla C, jos

$$P(A \cap B \mid C) = P(A \mid C)\, P(B \mid C).$$

Tämä voidaan merkitä $(A \perp\!\!\!\perp B) \mid C$.

Toisin sanoen kaksi tapahtumaa ovat ehdollisesti riippumattomia ehdolla C, jos ne ovat riippumattomia ehdollisen todennäköisyyden $P(\cdot \mid C)$ mielessä. Ehdollisen riippumattomuuden käsite voidaan tällä periaatteella tietenkin yleistää useammalle kuin kahdelle tapahtumalle.

1.8 Kokonaistodennäköisyys ja Bayesin kaava

Määritelmä 1.7 (Ositus)**.** Tapahtumat $B_1, \dots, B_n$ muodostavat perusjoukon Ω *osituksen*, jos

- ne ovat erillisiä: $B_i \cap B_j = \emptyset$, kun $i \neq j$
- ne peittävät (eli tyhjentävät) Ω:n, eli $\Omega = B_1 \cup \dots \cup B_n$.

Olkoon A tapahtuma ja $B_1, \dots, B_n$ perusjoukon Ω ositus. Joukko A voidaan osittaa leikkamalla se kullakin joukoista B_i, eli

$$A = (A \cap B_1) \cup \dots \cup (A \cap B_n),$$

jossa osat $(A \cap B_1), \dots, (A \cap B_n)$ ovat erillisiä, joten (additiivisuus ja kertolaskusääntö)

$$P(A) = \sum_{i=1}^{n} P(B_i)\, P(A \mid B_i). \tag{1.13}$$

Tämä on *kokonaistodennäköisyyden kaava*.

Olkoon A tapahtuma siten, että $P(A) > 0$, ja olkoon $B_1, \dots, B_n$ perusjoukon Ω ositus. Jos tapahtuma A on sattunut, niin tapahtuman B_i (ehdollinen) tn on

$$P(B_i \mid A) = \frac{P(A \cap B_i)}{P(A)} = \frac{P(B_i)\, P(A \mid B_i)}{P(A)}$$

Jos tässä $P(A)$ vielä esitetään kokonaistodennäköisyyden kaavalla (1.13), niin saadaan *Bayesin kaava*

$$P(B_i \mid A) = \frac{P(B_i)\, P(A \mid B_i)}{\sum_{j=1}^{n} P(B_j)\, P(A \mid B_j)}, \qquad 1 \le i \le n. \tag{1.14}$$

1.9 Monotoninen jatkuvuus

Tarkastellaan jonoa erillisiä tapahtumia $A_1, A_2, \dots$ Määritellään kaikilla $n \ge 1$ tapahtuma B_n yhdisteenä n ensimmäisestä A_i-tapahtumasta,

$$B_n = \bigcup_{i=1}^{n} A_i.$$

Tapahtumat $B_1, B_2, \dots$ muodostavat nyt kasvavan jonon, ts.

$$B_1 \subset B_2 \subset \dots$$

Lisäksi

$$\bigcup_{i=1}^{\infty} A_i = \bigcup_{n=1}^{\infty} B_n.$$

Kehitetään seuraavaksi esitys ylläolevan numeroituvasti äärettömän yhdisteen tn:lle.

Todennäköisyyden täysadditiivisuuden ja tapahtumien $A_1, A_2, \ldots$ erillisyyden nojalla on voimassa

$$P(B_n) = \sum_{i=1}^{n} P(A_i) \xrightarrow[n\to\infty]{} \sum_{i=1}^{\infty} P(A_i) = P(\bigcup_{i=1}^{\infty} A_i) = P(\bigcup_{m=1}^{\infty} B_m).$$

Toisaalta, jos lähdetään liikkeelle kasvavasta jonosta tapahtuma $B_1, B_2, \ldots,$ jossa siis

$$B_1 \subset B_2 \subset \ldots,$$

niin on helppoa konstruoida erilliset tapahtumat $A_1, A_2, \ldots$ siten, että kullakin n joukko B_n on yhdiste n ensimmäisestä joukosta A_i. Meidän tarvitsee vain valita

$$A_1 = B_1, \quad A_2 = B_2 \setminus A_1, \quad A_3 = B_3 \setminus (A_1 \cup A_2)$$

ja niin edelleen, ts. asetetaan $A_i = B_i \setminus (\cup_{j=1}^{i-1} A_j)$, kun $i \geq 2$. Edellisen päättelyn nojalla on voimassa

$$P(\bigcup_{i=1}^{\infty} B_i) = \lim_{n\to\infty} P(B_n).$$

Komplementteihin siirtymällä nähdään että samanlainen raja-arvotulos pätee, kun tarkastellaan laskevaa jonoa tapahtumia $B_1, B_2, \ldots,$ jossa siis

$$B_1 \supset B_2 \supset \ldots.$$

Tällöin on voimassa

$$P(\bigcap_{i=1}^{\infty} B_i) = \lim_{n\to\infty} P(B_n).$$

Kirjataan nämä huomiot lauseeksi.

Lause 1.2 (Todennäköisyyden monotoninen jatkuvuus)**.** *Olkoon $B_1, B_2, \ldots$ jono tapahtumia.*

(a) Jos jono on kasvava, eli $B_1 \subset B_2 \subset B_3 \subset \ldots$, niin

$$P(\bigcup_{i=1}^{\infty} B_i) = \lim_{n\to\infty} P(B_n).$$

(b) Jos jono on laskeva, eli $B_1 \supset B_2 \supset B_3 \supset \ldots$, niin

$$P(\bigcap_{i=1}^{\infty} B_i) = \lim_{n\to\infty} P(B_n).$$

Luku 2

Satunnaismuuttuja

2.1 Satunnaismuuttuja ja sen jakauma

Satunnaismuuttuja (lyhenne sm, engl. *random variable, rv*; myös *variate*) on satunnaiskokeeseen liittyvä numeerinen muuttuja, jonka arvo määräytyy kokeen lopputuloksesta. Satunnaismuuttuja on siis perusjoukolla määritelty reaaliarvoinen funktio.

Määritelmä 2.1. Olkoon Ω perusjoukko. Kuvaus $X : \Omega \to \mathbb{R}$ on *satunnaismuuttuja.*

Huomautus. Tarkasti ottaen edellinen määritelmä ei ole riittävä. Lisäksi pitäisi vaatia, että jokaisen riittävän säännöllisen joukon (kuten esim. jokaisen välin) $B \subset \mathbb{R}$ alkukuva kuvauksessa X pitää olla tapahtuma, eli sellainen Ω:n osajoukko, jonka tn on määritelty. Tämä vaatimus ilmaistaan sanomalla, että X on Borelin funktio (eli Borel-mitallinen funktio). Tästä lähtien sivuutamme tällaiset mitallisuustarkastelut.

Esimerkki 2.1. Esimerkkejä satunnaismuuttujista.

- Silmäluku nopanheitossa. Jos alkeistapaus $\omega \in \{1, \dots, 6\}$ kertoo nopan silmäluvun, niin satunnaismuuttuja X, jolle
$$X(\omega) = \omega, \qquad \text{kaikilla } \omega$$
on nopan silmäluku.
- Silmälukujen summa kahdessa nopanheitossa. Jos $E = \{1, \dots, 6\}$, perusjoukko $\Omega = E \times E$, ja alkeistapaukselle $(i, j) \in E \times E$ ensimmäinen koordinaatti i kertoo nopan yksi ja toinen koordinaatti j nopan kaksi silmäluvun, niin noppien silmälukujen summan ilmoittaa satunnaismuuttuja
$$X((i, j)) = i + j.$$

△

Määritelmä 2.2 (Tapahtuman indikaattori)**.** Olkoon $A \subset \Omega$ tapahtuma. Sen indikaattori (eli ilmaisin eli osoitinmuuttuja) 1_A on satunnaismuuttuja, jonka määrittelee lauseke

$$1_A(\omega) = \begin{cases} 1, & \text{jos } \omega \in A, \\ 0, & \text{muuten.} \end{cases}$$

Huomautus. Tapahtuman A indikaattorille käytetään yleisesti myös merkintää I_A. Jos tapahtuma A esitetään monimutkaisella lausekkeella, niin usein käytetään edellisten sijasta merkintää $1(A)$ tai $I(A)$. Jos on lisäksi tarpeen merkitä argumentti ω näkyviin (mutta harvoin on), niin voidaan käyttää merkintöjä $1(A)(\omega)$ tai $I(A)(\omega)$.

Usein samalla perusjoukolla on määritelty useampi kuin yksi satunnaismuuttuja, esim. X ja Y. Tällöin voidaan tarkastella myös *satunnaismuuttujien välisiä laskutoimituksia*, kuten aX (vakiolle $a \in \mathbb{R}$), $X+Y$, XY jne. Myös ne ovat satunnaismuuttujia. Esim. $X+Y$ tarkoittaa funktioiden yhteenlaskua, ts. $X+Y$ on se funktio $\Omega \to \mathbb{R}$, jolle

$$(X+Y)(\omega) = X(\omega) + Y(\omega).$$

Toisin sanoen satunnaismuuttujien (ja muiden funktioiden) laskutoimitukset tulkitaan pisteittäin. Jos Y ei saa arvoa nolla, niin myös X/Y on sm.

Jos X on sm, niin voidaan kysyä, millä todennäköisyydellä se saa arvon joukosta B, jossa $B \subset \mathbb{R}$. Kyseistä tapahtumaa voidaan merkitä

$$\{\omega \in \Omega : X(\omega) \in B\},$$

mutta sitä merkitään tavallisesti lyhyemmin seuraavasti,

$$\{X \in B\} = \{\omega \in \Omega : X(\omega) \in B\}.$$

Sen todennäköisyyttä voidaan merkitä jollakin seuraavista tavoista,

$$P(X \in B) = P\{X \in B\} = P(\{X \in B\}) = P(\{\omega \in \Omega : X(\omega) \in B\}). \tag{2.1}$$

Tavallisesti käytetään lyhyitä merkintöjä, joissa ei esiinny perusjoukkoa Ω eikä sisäkkäisiä sulkuja.

Erityisesti voidaan olla kiinnostuneita seuraavista valinnoista joukolle B. Huomaa, minkälaisia merkintöjä niiden yhteydessä käytetään.

- B on yksiö $\{x\}$, jossa $x \in \mathbb{R}$. Tällöin kysytään pistetodennäköisyyttä

$$P(X = x) = P(\{\omega \in \Omega : X(\omega) = x\}).$$

- B on väli kuten esim. suljettu väli $[a, b]$, jossa $a < b$,

$$P(a \le X \le b) = P(\{\omega \in \Omega : a \le X(\omega) \le b\}).$$

  Vastaavasti määritellään $P(a < X \le b)$, $P(a \le X < b)$ ja $P(a < X < b)$.

- Voidaan myös tarkastella tapausta $B = (-\infty, x]$. Tällöin merkitään

$$P(X \le x) = P(\{\omega \in \Omega : X(\omega) \le x\}).$$

Määritelmä 2.3 (Satunnaismuuttujan jakauma)**.** Jos X on satunnaismuuttuja, niin sen *jakauma* on

$$P(X \in B)$$

ymmärrettynä argumentin $B \subset \mathbb{R}$ funktiona.

Huomautus. Itse asiassa tn $P(X \in B)$ on määritelty vain silloin, kun $B \subset \mathbb{R}$ on ns. Borelin joukko, jollaisia ovat esim. kaikki välit ja kaikki joukot, jotka voidaan muodostaa numeroituvasta määrästä välejä soveltamalla numeroituva määrä joukko-operaatiota. Tästä lähtien jätämme tämän seikan vaille huomiota. (Oletamme, että kaavoissa tarkastelemme vain sellaisia $A \subset \Omega$, jotka ovat tapahtumia; vain sellaisia $B \subset \mathbb{R}$, jotka ovat Borelin joukkoja; vain sellaisia funktiota $X : \Omega \to \mathbb{R}$, jotka ovat Borelin funktiota eli satunnaismuuttujia; vain sellaisia funktioita $f : \mathbb{R} \to \mathbb{R}$ jotka ovat Borelin funktiota. Näitä seikkoja ei tarvitse pelästyä; muunlaisia joukkoja $B \subset \mathbb{R}$ tai funktioita $f : \mathbb{R} \to \mathbb{R}$ on huomattavan vaikea konstruoida.)

On mahdollista osoittaa, että sm:n X jakauma

$$P_X(B) = P(X \in B), \qquad B \subset \mathbb{R}$$

toteuttaa todennäköisyyden aksioomat, joten jakauma on perusjoukossa $\mathbb{R}$ määritelty tn-mitta. Käyttöön otetut lyhyet merkinnät johtavat siihen, että alkuperäinen perusjoukko Ω häviää tavallisesti kokonaan merkinnöistä. Tilanne voidaan mieltää myös siten, että perusjoukkona onkin $\mathbb{R}$ ja että tapahtumat ovat sen osajoukkoja.

Jakauma on reaaliakselin osajoukkojen kokoelmalla määritelty funktio, joten se on hyvin abstrakti olio. Seuraavaksi tarkastelemme konkreettisempia välineitä, joiden avulla voimme kuvailla ja hallita jakaumia: kertymäfunktio, diskreetin jakauman pistetodennäköisyysfunktio ja jatkuvan jakauman tiheysfunktio.

2.2 Kertymäfunktio

Määritelmä 2.4. Jos X on satunnaismuuttuja, niin funktio $F : \mathbb{R} \to \mathbb{R}$

$$F(x) = P(X \le x), \qquad x \in \mathbb{R}$$

on X:n kertymäfunktio (lyhenne kf).

Huomautus. Englannin kielellä jakauma on *(probability) distribution* (joskus *law*), ja kertymäfunktio on *distribution function* tai *cumulative distribution function, cdf.*

Esimerkki 2.2. Tapahtuman indikaattorin kf. Olkoon A tapahtuma, ja olkoon $X = 1_A$, eli X on tapahtuman A indikaattori. Tällöin X:n kf on

$$F(x) = \begin{cases} 0, & \text{jos } x < 0, \\ 1 - P(A), & \text{jos } 0 \le x < 1, \\ 1, & \text{jos } x \ge 1. \end{cases}$$

Huomaa, että kf:lla voi olla hyppyjä ja että se voi olla vakio jollakin välillä. △

Palautetaan mieleen ennen seuraavaa määritelmää, että monotonisella funktiolla $G : \mathbb{R} \to \mathbb{R}$ on jokaisessa pisteessä $x \in \mathbb{R}$ olemassa sekä vasemmanpuoleinen että oikeanpuoleinen raja-arvo. Lisäksi sillä on raja-arvo pisteissä $-\infty$ ja ∞ (mutta äärettömässä raja-arvo voi olla joko jokin reaaliluku, tai jompikumpi äärettömistä arvoista $-\infty$ tai ∞). Näitä raja-arvoja merkitään seuraavasti

$$G(x-) = \lim_{y \to x-} G(y), \qquad G(x+) = \lim_{y \to x+} G(y)$$
$$G(-\infty) = \lim_{y \to -\infty} G(y), \qquad G(\infty) = \lim_{y \to \infty} G(y).$$

Lause 2.1 (Kertymäfunktion ominaisuudet)**.** *Satunnaismuuttujan X kertymäfunktiolla $F : \mathbb{R} \to \mathbb{R}$ on seuraavat ominaisuudet.*

(a) F on kasvava funktio.

(b) F on oikealta jatkuva funktio, eli

$$F(x+) = F(x), \qquad \text{kaikilla } x \in \mathbb{R}.$$

(c) $F(-\infty) = 0$ ja $F(\infty) = 1$.

Kääntäen, jos funktiolla $F : \mathbb{R} \to \mathbb{R}$ on nämä ominaisuudet, niin on olemassa sm X siten, että F on X:n kf.

Todistus. Jos F on X:n kf ja $x < y$, niin $\{X \le x\} \subset \{X \le y\}$, joten

$$F(x) = P(X \le x) \le P(X \le y) = F(y),$$

mikä todistaa ominaisuuden (a).

Koska F on kasvava funktio, niin raja-arvo $F(x+)$ on olemassa missä tahansa pisteessä $x \in \mathbb{R}$. Koska oikeanpuoleinen raja-arvo on olemassa, niin se voidaan laskea minkä tahansa pistettä x oikealta lähestyvän jonon avulla, esim. seuraavasti,

$$\begin{aligned} F(x+) &= \lim_{n\to\infty} F(x + \frac{1}{n}) = \lim_{n\to\infty} P(X \le x + \frac{1}{n}) = P(\bigcap_{n=1}^{\infty} \{X \le x + \frac{1}{n}\}) \\ &= P(X \le x) = F(x). \end{aligned}$$

Päättely perustui siihen, että tapahtumat $\{X \le x + \frac{1}{n}\}$ muodostavat laskevan jonon, kun $n = 1, 2, \ldots$. Sen takia voitiin käyttää tn-mitan monotonista jatkuvuutta (ks. lause 1.2).

Samaan tapaan c-kohta saadaan perusteltua monotonisen jatkuvuuden avulla, ts.

$$\begin{aligned} F(-\infty) &= \lim_{n\to\infty} F(-n) = P(\bigcap_n \{X \le -n\}) = P(\emptyset) = 0. \\ F(\infty) &= \lim_{n\to\infty} F(n) = P(\bigcup_n \{X \le n\}) = P(\Omega) = 1. \end{aligned}$$

Sivuutamme perustelun sille, miksi ominaisuuksista (a), (b) ja (c) seuraa se, että F on jonkin satunnaismuuttujan kf. □

Lasketaan tn $P(a < X \le b)$, jossa $a < b$. Koska

$$\{X \le b\} = \{X \le a\} \cup \{a < X \le b\},$$

jossa osat ovat erillisiä, on

$$P(a < X \le b) = F(b) - F(a).$$

Seuraavan lauseen jälkeen pystymme toistamaan vastaavan laskun myös väleille (a, b), $[a, b)$ ja $[a, b]$.

Lause 2.2. *Jos F on sm:n X kf, niin kaikilla x*

$$P(X < x) = F(x-).$$

Todistus. Todistus onnistuu helposti käyttämällä tn-mitan monotonista jatkuvuutta. □

Huomaa, että kertymäfunktiolla voi olla hyppyjä. Koska

$$\{X \le x\} = \{X < x\} \cup \{X = x\},$$

jossa osat ovat erillisiä, niin

$$P(X = x) = P(X \le x) - P(X < x) = F(x) - F(x-).$$

Toisin sanoen pistetodennäköisyys $P(X = x)$ on yhtä suuri kuin kertymäfunktion hyppy pisteessä x.

Merkintäsopimus (satunnaismuuttuja voidaan ilmoittaa funktion alaindeksillä). Usein tarkastellaan useampaa kuin yhtä satunnaismuuttujaa, vaikkapa X ja Y. Tällöin käytetään sellaista merkintätapaa, jossa satunnaismuuttuja kirjoitetaan alaindeksiksi: esim. F_X on X:n kf ja F_Y on Y:n kf. Tätä merkintätapaa voidaan käyttää muidenkin jakauman esitysten kuin kertymäfunktioiden kohdalla.

Määritelmä 2.5. Satunnaismuuttujilla X ja Y on *sama jakauma* eli ne ovat *samoin jakautuneet*, jos

$$P(X \in B) = P(Y \in B), \qquad \text{kaikilla } B \subset \mathbb{R}.$$

Tämä seikka voidaan ilmaista merkinnällä

$$X \overset{\mathrm{d}}{=} Y.$$

Seuraava lause toteaa, että kertymäfunktiot ja jakaumat vastaavat toisiaan yksikäsitteisesti. Tämän takia voidaan puhua tietyn jakauman kertymäfunktiosta, eikä aina tarvitse puhua sm:n kertymäfunktiosta.

Lause 2.3 (Kertymäfunktio määrää jakauman yksikäsitteisesti)**.** *Seuraavat asiat ovat yhtäpitäviä.*

(a) X ja Y ovat samoin jakautuneet.

(b) $F_X = F_Y$ (ts. X:n ja Y:n kertymäfunktiot ovat samat).

Todistus. **(a)** $\Rightarrow$ **(b)**: Koska jakaumat ovat samat, on kaikilla $t \in \mathbb{R}$ voimassa

$$F_X(t) = P(X \le t) = P(Y \le t) = F_Y(t),$$

joten kertymäfunktiot ovat samat.

Implikaation (b) $\Rightarrow$ (a) todistaminen sivuutetaan, sillä se ei onnistu ilman mittateoriaa. □

2.3 Diskreetti jakauma

Määritelmä 2.6. Satunnaismuuttujan X jakauma on *diskreetti*, jos sen arvojoukko $\{x_1, x_2, \dots\}$ on äärellinen tai korkeintaan numeroituvasti ääretön.

Se, että sm:n X jakauma on diskreetti voidaan vaihtoehtoisesti ilmaista myös jollakin seuraavista tavoista

- X on diskreetti sm,

- sm X on diskreetisti jakautunut
- X:llä on diskreetti jakauma.

Joskus määritelmää laajennetaan siten, että vaaditaan vain, että X saa todennäköisyydellä yksi arvon korkeintaan numeroituvasti äärettömästä joukosta S. Tällöin diskreetti sm X voi saada arvoja joukon S ulkopuolelta, mutta tämän poikkeustapahtuman todennäköisyys on nolla. Yksinkertaisuuden vuoksi kutsumme myös tässä tapauksessa X:n arvojoukoksi jotakin sellaista korkeintaan numeroituvaa joukkoa S, jolle $P(X \in S) = 1$.

Määritelmä 2.7. Satunnaismuuttujan X pistetodennäköisyysfunktio (lyh. ptnf) on funktio $f : \mathbb{R} \to \mathbb{R}$, jolle

$$f(x) = P(X = x), \qquad x \in \mathbb{R}\,.$$

Huomautus. Pistetodennäköisyysfunktio $x \mapsto P(X = x)$ (engl. *probability (mass) function, pmf*) on määritelty koko reaaliakselilla kaikenlaisille satunnaismuuttujille X. Tämä käsite on kuitenkin käyttökelpoinen lähinnä vain diskreeteille satunnaismuuttujille.

Merkintäsopimuksia.

- Tyypillisesti pistetodennäköisyydet ilmoitetaan vain niille x_i, jotka kuuluvat diskreetin sm:n määritelmässä esiintyvään, korkeintaan numeroituvasti äärettömään joukkoon, joka voi esimerkiksi olla jokin kokonaislukujen osajoukko. Tällöin asiayhteydestä pitää ymmärtää, että muilla $x \in \mathbb{R}$ on $P(X = x) = 0$.
- Diskreetin satunnaismuuttujan ptnf:lle käytetään muuten samaa symbolia kuin sen kf:lle, mutta ptnf:n symboli kirjoitetaan pienellä kirjaimella (esim. f, g tai f_X) ja kf:n symboli vastaavalla isolla kirjaimella (esim. F, G tai F_X).

Lause 2.4 (Ptnf:n ominaisuuksia)**.** *Olkoon X diskreetti sm, f sen ptnf, ja olkoon S sen arvojoukko. Tällöin*

(a) $0 \le f(x) \le 1$ *kaikilla x, ja $f(x) = 0$, kun $x \notin S$.*

(b) f määrää X:n jakauman kaavalla

$$P(X \in B) = \sum_{x \in B} f(x)$$

Todistus. Olkoon $S = \{x_1, x_2, \dots\}$. Kohta (a) on ilmeinen. Kohdan (b) pohjustamiseksi pitää huomauttaa, että siinä lasketaan yhteen korkeintaan numeroituva määrä ei-negatiivisia termejä $f(x_i)$, jossa $x_i \in S \cap B$, minkä takia kyseinen summamerkintä on mielekäs. Kohdan (b) kaava seuraa kokonaistodennäköisyyden kaavasta (tai mahdollisesti sen versiosta numeroituvasti äärettömällä ositukselle), kun perusjoukolle Ω käytetään ositusta

$$\{X \in S^c\}, \{X = x_1\}, \{X = x_2\}, \dots.$$

□

Lause 2.5. *Olkoon $f : \mathbb{R} \to \mathbb{R}$ ei-negatiivinen funktio, joka saa positiivisia arvoja korkeintaan numeroituvassa $\mathbb{R}$:n osajoukossa siten, että*

$$\sum_x f(x) = 1$$

Tällöin se on jonkin diskreetin satunnaismuuttujan ptnf.

Todistus. Ideana on se, että ensin konstruoidaan pistetodennäköisyyksistä vastaava kertymäfunktioehdokas, ja sitten tarkistetaan, että sillä on kertymäfunktion ominaisuudet. □

2.4 Esimerkkejä diskreeteistä jakaumista

Diskreetti tasajakauma. Satunnaismuuttujalla X on diskreetti tasajakauma joukossa $\{1, \dots, N\}$, mikäli sen ptnf on

$$f(x) = \frac{1}{N}, \qquad \text{kun } x = 1, 2, \dots, N.$$

Esimerkiksi, jos X on silmäluku nopanheitossa, niin X:llä on diskreetti tasajakauma joukossa $1, \dots, 6$.

Bernoullin jakauma. Olkoon A on tapahtuma, ja merkitään $p = P(A)$, jolloin $0 \le p \le 1$. Tällöin indikaattorilla $X = 1_A$ on Bernoullin jakauma parametrilla p (eli X noudattaa Bernoullin jakaumaa parametrillä p). Tämä asia voidaan merkitä

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

X:n arvojoukko on $\{0, 1\}$, ja sen ptnf on

$$f(x) = f(x \mid p) = \begin{cases} 1 - p, & \text{kun } x = 0, \\ p, & \text{kun } x = 1, \\ 0, & \text{muuten.} \end{cases}$$

Tämä ptnf voidaan kirjoittaa lyhyemmin

$$f(x \mid p) = p^x (1 - p)^{1-x}, \qquad \text{kun } x = 0, 1,$$

kun käytetään edellä tehtyä merkintäsopimusta, jonka mukaan ptnf:n arvot tarvitsee kertoa vain kyseisen satunnaismuuttujan arvojoukossa. Tässä tulkitaan $0^0 = 1$, mikäli $p = 0$ tai $p = 1$.

Tavallisesti sanotaan, että satunnaiskoe onnistuu silloin, kun $X = 1$ ja epännoistuu muuten. Tällöin p on yhtä kuin kokeen onnistumistodennäköisyys.

Huomautus. Useimmat mielenkiintoiset jakaumat riippuvat yhden tai useamman parametrin arvosta. Pitääksemme lukua parametrin arvosta, merkitsemme sen tarvittaessa pystyviivan oikealle puolelle. Tätä merkintää voidaan käyttää kertymäfunktioille, pistetodennäköisyysfunktiolle ynnä muille funktioille. Jos taas sekaantumisen vaaraa ei ole, parametri voidaan jättää pois merkinnöistä.

Binomijakauma. Olkoon $n \ge 1$ kokonaisluku ja $0 \le p \le 1$. Sm X noudattaa binomijakaumaa parametreillä n ja p, jos sen arvojoukko on $\{0, 1, \dots, n\}$ ja ptnf on

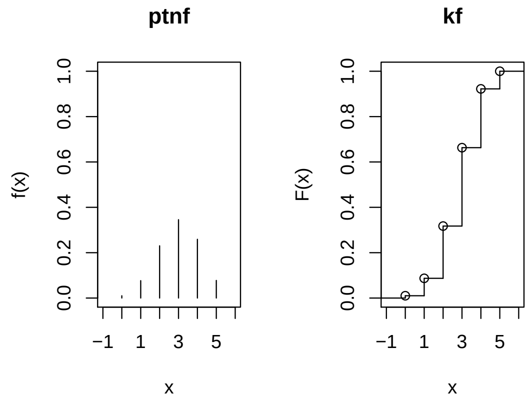
$$f(x) = f(x \mid n, p) = \binom{n}{x} p^x (1 - p)^{n-x}, \qquad \text{kun } x = 0, 1, \dots, n.$$

Tämä voidaan merkitä

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

Binomijakaumaa $\text{Bin}(n, p)$ noudattava sm syntyy toistokokeessa, jossa toistetaan riippumattomasti n kertaa sellaista (Bernoullin) koetta, jossa yhdessä kokeessa onnistumistodennäköisyys on p. Jos X on onnistumisten lukumäärä n toistossa, niin $X \sim \text{Bin}(n, p)$. Tämä tulos perustellaan esimerkissä 3.3. Bernoullin jakauma $\text{Bernoulli}(p)$ on tietenkin sama kuin binomijakauma $\text{Bin}(1, p)$.

Kuva 2.1 Binomijakauman $\mathrm{Bin}(n, p)$ ptfn ja kf, kun $n = 5$ ja $p = 0.6$.



2.5 Integraalilaskennan kertausta

Ennen kuin määrittelemme jatkuvan jakauman, kertaamme joitakin asioita integroinnista. Matemaattisessa analyysissä on määritelty erilaisia integrointiprosesseja, joista lukija on tässä vaiheessa tutustunut Riemannin integraaliin. Todennäköisyyslaskennan yhteydessä Riemannin integraalin käyttö johtaisi tiettyihin asian kannalta epäoleellisiin hankaluuksiin, minkä takia matemaatikot käyttävät sen sijasta ns. Lebesguen integraalia. Tämän kurssin puitteissa näiden kahden integraalikäsitteen eroilla ei ole merkitystä, vaan kurssin tehtävät saadaan ratkaistua käyttämällä tuttua Riemannin integraalia, mutta usein tarvitaan ns. epäolellista Riemannin integraalia.

Olkoon f välillä I määritelty reaalifunktio. Tällöin sanomme, että F on funktion f *antiderivaatta* eli *integraalifunktio*, jos

$$F'(x) = f(x), \qquad \text{kaikilla } x \in I. \tag{2.2}$$

(Välin päätepisteissä derivaatta tässä tarvittaessa määritellään toispuoleisena derivaattaa.) Jos F on f:n antiderivaatta ja C on vakio, niin myös $F+C$ on f:n antiderivaatta. Jokaisella suljetulla äärellisellä välillä $I = [a, b]$ määritellyllä jatkuvalla funktiolla f on olemassa antiderivatta välillä I.

Funktion f Riemannin interaali välin $[a, b]$ yli määritellään eräänlaisena summien raja-arvona, mikäli ko. raja-arvo on olemassa. Tällöin sanotaan, että f on Riemann-integroituva välin $[a, b]$ yli. Määritelmä toimii ainoastaan äärellisille integrointiväleille $[a, b]$, ja lisäksi integroitavan funktion eli integrandin f tulee on rajoitettu integrointivälillä. Integraalien arvoja lasketaan hyvin harvoin suoraan määritelmän perusteella, vaan tämän sijasta integraali tyypillisesti lasketaan antiderivaatan avulla.

Jos f on äärellisellä välillä $[a, b]$ määritelty jatkuva funktio, niin sen määrätty integraali välin $[a, b]$ yli saadaan laskettua seuraavalla kaavalla, jos tunnetaan jokin f:n antiderivaatta F:

$$\int_a^b f(x)\,\mathrm{d}x = \Big/_{a}^{b} F(x) = F(b) - F(a). \tag{2.3}$$

Edellisen lisäksi on hyödyllistä tietää, että integrandin muuttaminen äärellisessä määrässä pisteitä ei muuta integraalin arvoa. Nimittäin, jos f on Riemann-integroituva välin $[a, b]$ yli ja

$g = f$ kaikkialla muualla paitsi äärellisessä määrässä pisteitä, niin g on Riemann-integroituva, ja

$$\int_a^b g(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x. \tag{2.4}$$

Kaavasta (2.3) sekä tutuista derivointikaavoista saadaan johdettua monia integrointikaavoja, joista kertaamme osittaisintegroinnin sekä sijoituskeinon.

Olkoot u ja v välillä $[a, b]$ määriteltyjä jatkuvasti derivoituvia funktioita. Kun otetaan huomioon tulon derivointikaava

$$\frac{\mathrm{d}}{\mathrm{d}x}[u(x)\,v(x)] = u'(x)\,v(x) + u(x)\,v'(x),$$

saadaan tulos

$$\Big/_a^b u(x)\,v(x) = \int_a^b (u'(x)\,v(x) + u(x)\,v'(x))\,\mathrm{d}x,$$

josta järjestelmällä saadaan *osoittaisintegrointikaava* (engl. *integration by parts*)

$$\int_a^b u(x)\,v'(x)\,\mathrm{d}x = \Big/_a^b u(x)\,v(x) - \int_a^b u'(x)\,v(x)\,\mathrm{d}x. \tag{2.5}$$

Sijoituskeino (engl. *integration by substitution*) perustuu yhdistetyn funktion derivointikaavaan. Olkoon f jatkuva funktio välillä I, joka sisältää integrointivälin $[a, b]$, ja olkoon $g : [c, d] \to I$ jatkuvasti derivoituva funktio siten, että $g(c) = a$ ja $g(d) = b$. Olkoon F funktion f integraalifunktio. Tällöin

$$\int_a^b f(x)\,\mathrm{d}x = \Big/_a^b F(x) = F(b) - F(a) = F(g(d)) - F(g(c)) = \Big/_c^d F(g(t)).$$

Yhdistetyn funktion derivointikaavan nojalla

$$\frac{\mathrm{d}}{\mathrm{d}t}F(g(t)) = f(g(t))\,g'(t), \qquad \text{kaikilla } t \in [c, d]$$

jota käyttämällä saadaan sijoituskeinon kaava

$$\int_a^b f(x)\,\mathrm{d}x = \int_c^d f(g(t))\,g'(t)\,\mathrm{d}t. \tag{2.6}$$

Formaalisti tässä ainoastaan sijoitetaan $x = g(t)$ ja ja $\mathrm{d}x = g'(t)\,\mathrm{d}t$ integraaliin sekä muutetaan alkuperäiset x-rajat (a, b) vastaaviksi t-rajoiksi (c, d), joille $a = g(c)$ ja $b = g(d)$.

Huomaa, mitä ehtoja kaavassa (2.3) asetettiin: integrandin f tulee olla jatkuva funktio kyseessä olevalla äärellisellä suljetulla välillä ennen kuin integraalin saa laskettua antiderivaatan F avulla. Jos f on muuten jatkuva, mutta sillä on hyppy pisteessä $a < c < b$, niin tällöin integrointialue voidaan pilkkoa jakopisteellä c väleihin $[a, c]$ ja $[c, b]$ ja yrittää laskea integraali summana

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x,$$

jossa suotuisassa tapauksessa kumpikin integraaleista saadaan laskettua antiderivaatan avulla. (Pisteessä c joudutaan ainakin toisessa integraaleista integrandin määritelmää muuttamaan siten, että saadaan ko. suljetulla välillä jatkuva funktio, mutta tämä on vaaratonta, koska menettely ei muuta integraalin arvoa.)

Jos integrointivälin päätepiste on äärettömyydessä tai jos funktio f kasvaa rajatta integrointivälin päätepisteessä, niin Riemannin integraalin perusmääritelmä ei tuota mitään tulosta, vaan integraali joudutaan laskemaan ns. epäoleellisena integraalina. Tämä tarkoittaa käytännössä raja-arvon laskemista sellaisesta integraalista, jossa ongelmallinen päätepiste on korvattu hyvin käyttäytyvällä pisteellä, jonka sitten annetaan lähestyä alkuperäistä ongelmallista päätepistettä. Palautamme tämän menettelyn mieleen esimerkkien avulla.

Esimerkki 2.3. Laskemme integraalin $\int_0^1 x^{-a}\,\mathrm{d}x$, jossa $0 < a < 1$. Tämä integraali on epäoleellinen, sillä funktio x^{-a} kasvaa rajatta, kun $x \to 0+$. Jos $\epsilon > 0$, niin integraali välin $[\epsilon, 1]$ yli saadaan laskettua antiderivaatan avulla, sillä

$$\int_\epsilon^1 x^{-a}\,\mathrm{d}x = \Big/_{\!\!\epsilon}^{1} \frac{x^{-a+1}}{-a+1} = \frac{1}{1-a} - \frac{\epsilon^{1-a}}{1-a}.$$

Tässä $\epsilon^{1-a} \to 0$, kun $\epsilon \to 0+$ sen takia, että $1 - a > 0$. Koska raja-arvo on olemassa, niin alkuperäinen integraali saadaan laskettua epäoleellisena integraalina, ja

$$\int_0^1 x^{-a}\,\mathrm{d}x = \lim_{\epsilon \to 0+} \int_\epsilon^1 x^{-a}\,\mathrm{d}x = \frac{1}{1-a}.$$

△

Esimerkki 2.4. Laskemme integraalin $\int_1^\infty x^{-a}\,\mathrm{d}x$, kun $a > 1$. Nyt integraali on epäoleellinen sen takia, että ylempi päätepiste on ääretön. Integraali on kuitenkin äärellinen, sillä

$$\int_1^\infty x^{-a}\,\mathrm{d}x = \lim_{M\to\infty} \int_1^M x^{-a}\,\mathrm{d}x = \lim_{M\to\infty} \Big/_{\!\!1}^{M} -\frac{x^{1-a}}{a-1} = \lim_{M\to\infty} \left(-\frac{M^{1-a}}{a-1} + \frac{1}{a-1} \right) = \frac{1}{a-1}.$$

Lausekkeessa M^{1-a} potenssi on negatiivinen, joten tämä lauseke lähestyy nollaa, kun $M \to \infty$.
△

2.6 Jatkuva jakauma

Määritelmä 2.8. Satunnaismuuttujalla X on jatkuva jakauma tiheysfunktiolla f (lyh. tf), jos f on ei-negatiivinen ja

$$P(a \le X \le b) = \int_a^b f(x)\,\mathrm{d}x$$

kaikilla $a, b \in \mathbb{R}$, $a < b$.

Jatkuvalle jakaumalle todennäköisyys $P(a \le X \le b)$ saadaan käyrän $y = f(x)$, x-akselin sekä suorien $x = a$ ja $x = b$ rajaaman tason osajoukon pinta-alana.

Tiheysfunktio on englanniksi *(probability) density function, pdf.* Toisinaan puhutaan tiheysfunktion sijasta lyhyesti tiheydestä.

Se, että sm:lla X on jatkuva jakauma voidaan vaihtoehtoisesti ilmaista myös jollakin seuraavista tavoista.

- X:n jakauma on jatkuva,

- X on jatkuvasti jakautunut,
- X on jatkuva satunnaismuuttuja.

Sanontapa "X on jatkuva satunnaismuuttuja" on yleisesti käytössä, vaikka se on harhaanjohtava. Asiaa tuntematon henkilö voi nimittäin sekoittaa sen siihen väitteeseen, että funktio $X : \Omega \to \mathbb{R}$ olisi jatkuva. (Funktion X jatkuvuus tai epäjatkuvuus on todennäköisyyslaskennan näkökulmasta tavallisesti epärelevantti kysymys.)

Merkintäsopimus

- Jatkuvan jakauman tiheysfunktiolle käytetään muuten samaa symbolia kuin sen kf:lle, mutta tiheysfunktion symboli kirjoitetaan pienellä kirjaimella (esim. f, g tai f_X) ja kf:n symboli vastaavalla isolla kirjaimella (esim. F, G tai F_X).
- Joskus tiheysfunktio häviää jonkin joukon $B \subset \mathbb{R}$ ulkopuolella. Tällöin tiheysfunktion kaava voidaan ilmoittaa vain joukossa B, ja asiayhteydestä pitää ymmärtää, että $f(x) = 0$, kun $x \notin B$.

Jatkuvan jakauman määritelmässä esiintyvä integraali tarkoittaa itseasiassa Lebesguen integraalia, mutta sovelluksissa esiintyvät tiheysfunktiot ovat jatkuvia funktioita kaikkialla muualla paitsi äärellisessä määrässä pisteitä ja niiden integraalit voidaan käytännössä laskea Riemannin integraalien avulla. Todennäköisyyslaskennassa integrointialue ei välttämättä ole äärellinen väli. Lisäksi tiheysfunktio voi olla yhdessä tai useammassa pisteessä singulaarinen siten, että ainakin toinen sen toispuoleisista raja-arvoista on ääretön. Näiden asioiden takia integrointialue joudutaan usein ensin pilkkomaan paloihin, minkä jälkeen tiheysfunktion integraali saadaan palautettua summaksi (mahdollisesti epäoleellisia) Riemannin integraaleja.

Jakauman tiheysfunktio ei ole yksikäsitteinen, mutta melkein yksikäsitteinen. Esimerkiksi, jos f on tietyn jakauman tf ja g on yhtä kuin f muualla paitsi äärellisessä määrässä pisteitä, niin

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b g(x)\,\mathrm{d}x$$

kaikilla $a < b$. Tämän takia myös g on kyseisen jakauman tf. Kaikkein yleisin tulos on seuraava. Ei-negatiiviset funktiot f ja g ovat saman jakauman tiheysfunktioita, jos ja vain jos ne yhtyvät melkein kaikkialla, eli joukko jossa $f \neq g$ on nollamittainen (mikä käsite määritellään reaalianalyysin kursseilla).

Jos X:n jakauma on jatkuva ja sen tf on f, niin

$$P(X = x) = P(x \le X \le x) = \int_x^x f(u)\,\mathrm{d}u = 0$$

kaikilla x, joten pistetodennäköisyys on aina nolla. Tämän takia ptnf ei ole hyödyllinen käsite jatkuville jakaumille.

Koska kaikki pistetodennäköisyydet ovat nollia, niin

$$P(a < X < b) = P(a \le X < b) = P(a < X \le b) = P(a \le X \le b) = \int_a^b f(x)\,\mathrm{d}x,$$

kun $a < b$. Rajankäynnin jälkeen huomataan, että puoliäärettömien välien todennäköisyydet saadaan myöskin tf:a integroimalla. Kaikille x pätee

$$P(X \le x) = P(X < x) = \int_{-\infty}^x f(u)\,\mathrm{d}u, \qquad P(X > x) = P(X \ge x) = \int_x^\infty f(u)\,\mathrm{d}u.$$

Itse asiassa pätee vielä yleisempi tulos

$$P(X \in B) = \int_B f(x)\,\mathrm{d}x, \qquad \text{kaikille } B \subset \mathbb{R}, \tag{2.7}$$

jossa integraalin alaindeksi B tarkoittaa joukon B yli laskettua integraalia, ts.

$$\int_B f(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} 1_B(x)\,f(x)\,\mathrm{d}x.$$

Erityisesti

$$1 = P(\Omega) = P(X \in \mathbb{R}) = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x,$$

joten f:n integraali koko reaaliakselin yli on yksi.

Lause 2.6 (Kertymäfunktion ja tiheysfunktion yhteys)**.** *Olkoon X:llä jatkuva jakauma.*

(a) Jos X:n tiheysfunktio on f, niin sen kertymäfunktio F on

$$F(x) = \int_{-\infty}^{x} f(u)\,\mathrm{d}u.$$

(b) Jos X:n kertymäfunktio on F, niin F on derivoituva melkein kaikkialla, ja X:n tiheysfunktioksi voidaan valita F'.

Todistus. Kertymäfunktion a-kohdassa annettu lauseke tuli jo perusteltua. Kohdan (b) väite (tässä yleisyydessä) seuraa mitta- ja integrointiteoriasta. □

Huomautus. Kun sanotaan, että jatkuvan jakauman tiheysfunktioksi voidaan valita $f = F'$, niin tällä tarkoitetaan sitä, että tiheysfunktioksi valitaan jokin funktio f, joka yhtyy funktion F derivaattaan niissä pisteissä, joissa derivaatta on määritelty, ja muissa pisteissä f saadaan määritellä mielivaltaisesti.

Lause 2.7 (Riittävä ehto sille, että annettu kf on jatkuvan jakauman kf)**.** *Olkoon F kertymäfunktio siten, että*

(1) F on jatkuva koko reaaliakselilla,

(2) F on derivoituva kaikkialla muualla paitsi mahdollisesti äärellisessä määrässä pisteitä,

(3) ja lisäksi derivaatta $f = F'$ on jatkuva kaikkialla muualla paitsi mahdollisesti äärellisessä määrässä pisteita.

Tällöin F on jatkuvan jakauman kertymäfunktio, ja jakauman tiheysfunktioksi voidaan valita sen derivaatta $f = F'$.

Todistus. (Hahmotelma) Jos $[a, b]$ on sellainen äärellinen väli, joka ei sisällä yhtään pistettä, jossa F:n derivaatta ei ole jatkuva, niin

$$F(b) - F(a) = \int_a^b f(u)\,\mathrm{d}u,$$

jossa f on kf:n F derivaatta. Tämä yhtälö todistetaan mielivaltaisille $a \le b$ ottamalla raja-arvoja sekä käyttämällä oletuksia, erityisesti sitä oletusta, että F:llä ei ole hyppyä niissä pisteissä, joissa sen derivaatta ei ole jatkuva. □

Täydentäviä huomautuksia

- Jos X:llä on jatkuva jakauma, niin sen kertymäfunktio on jatkuva. Sen sijaan tiheysfunktio voi olla hyvin epäsäännöllinen funktio. Tosin tilastollisissa sovelluksissa esiintyvät tiheysfunktiot ovat varsin säännöllisiä funktioita, sillä ne ovat jatkuvia kaikkialla paitsi äärellisessä määrässä pisteitä. Tiheysfunktion ei tarvitse olla rajoitettu funktio.
- Täydellinen karakterisointi niille kertymäfunktioille, jotka ovat jonkin jatkuvan jakauman kertymäfunktioita on se, että kf F on jatkuvan jakauman kf täsmälleen silloin, kuin F on *absoluuttisesti jatkuva*, mutta tämän asian määritteleminen jätetään reaalianalyysin kurssien huoleksi.
- Siitä, että jakauman kf on jatkuva ei vielä seuraa, että jakauma olisi jatkuva. Tätä varten kf:n pitää siis olla absoluuttisesti jatkuva. Täsmällisempi nimitys jatkuvalle jakaumalle olisi *absoluuttisesti jatkuva jakauma*, mutta tätä nimitystä käytetään harvoin.

Lause 2.8. *Jos $f : \mathbb{R} \to \mathbb{R}$ on ei-negatiivinen, ja sen integraali koko reaaliakselin yli on yksi, niin se on jonkin satunnaismuuttujan tiheysfunktio.*

Todistus. Ideana on konstruoida kertymäfunktioehdokas integroimalla funktiota f, ja tarkistaa, että konstruoidulla funktiolla on kertymäfunktion ominaisuudet. □

Lause 2.9. *Jos $g \geq 0$ on sellainen reaaliakselilla määritelty funktio, jonka integraali $\int_{-\infty}^{\infty} g(x)\,\mathrm{d}x$ on äärellinen ja aidosti positiivinen, niin tällöin on olemassa vakio k siten, että $f = k\,g$ on tiheysfunktio.*

Todistus. Valitaan $1/k = \int_{-\infty}^{\infty} g(x)\,\mathrm{d}x$, jonka jälkeen

$$\int f(x)\,\mathrm{d}x = k \int g(x)\,\mathrm{d}x = \frac{k}{k} = 1.$$

□

Huomautus. Tällä tavalla jokainen *normalisoimaton tiheysfunktio g* määrää yksikäsitteisellä tavalla jatkuvan jakauman, jonka tiheysfunktio on edellä konstruoitu f.

Esimerkki 2.5. Tiheysfunktion frekvenssitulkinta. Olkoon sm:lla X tiheysfunktio f. Oletetaan, että meillä on käytettävissä suuri aineisto $x_1, x_2, \ldots, x_N$, jota voidaan pitää N riippumattoman sm:n X:n jakaumaa noudattavan sm:n havaittuina arvoina. Toisin sanoen

$$x_1 = X_1(\omega^{\mathrm{act}}), x_2 = X_2(\omega^{\mathrm{act}}), \ldots x_N = X_N(\omega^{\mathrm{act}}),$$

jossa ω^{act} on aktualisoitunut alkeistapaus, kullakin X_i on sama jakauma kuin X:llä ja lisäksi $X_1, \ldots, X_N$ ovat riippumattomia (mikä määritellään myöhemmin). Kuinka tällaisessa tilanteessa voidaan arvioida tiheysfunktion arvoa pisteessä u? Oletetaan, että u on f:n jatkuvuuspiste.

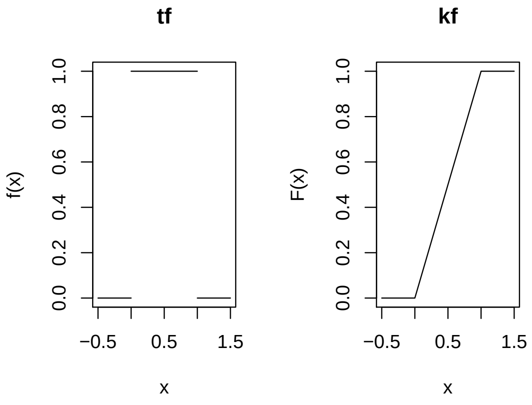
Tällaisessa tilanteessa voidaan menetellä seuraavasti. Valitaan jokin pieni luku $h > 0$ ja lasketaan, kuinka moni otospisteistä x_i osuu h:n pituiselle välille

$$[u - \frac{h}{2}, u + \frac{h}{2}].$$

Olkoon tämä lukumäärä $N_{u,h}$. Tiheysfunktion määritelmän nojalla ja todennäköisyyden frekvenssitulkinnan nojalla

$$P(u - \frac{h}{2} \leq X \leq u + \frac{h}{2}) = \int_{u-h/2}^{u+h/2} f(x)\,\mathrm{d}x \approx \frac{N_{u,h}}{N}.$$

Kuva 2.2 Tasajakauman $\mathrm{U}(0,1)$ tiheys- ja kertymäfunktio.



Toisaalta, integraalilaskennan väliarvolauseen nojalla

$$\int_{u-h/2}^{u+h/2} f(x)\,\mathrm{d}x \approx h\,f(u).$$

Kun nämä arviot yhdistetään, saadaan arvio

$$f(u) \approx \frac{N_{u,h}}{h\,N}.$$

Tietenkin tässä voitaisiin käyttää myös muita h:n pituisia välejä, jotka sisältävät pisteen u kuin tätä, jonka keskipiste on u. Sivuutamme nyt sen ongelman, kuinka arvo h kannattaisi valita, jotta arviosta saataisiin mahdollisimman tarkka.

Tällaista ns. parametritonta tiheysfunktion estimointia varten löytyy kirjallisuudesta monia erilaisia menetelmiä, joista eräät ovat perusajatukseltaan näin yksinkertaisia. $\triangle$

2.7 Esimerkkejä jatkuvista jakaumista

Tasajakauma (engl. *uniform distribution*). Olkoon $a, b \in \mathbb{R}$ ja $a < b$. Satunnaismuuttujalla X on välin (a, b) tasajakauma, $X \sim \mathrm{U}(a, b)$, jos sillä on jatkuva jakauma tiheysfunktiona f, jossa

$$f(x) = f(x \mid a, b) = \begin{cases} \dfrac{1}{b-a}, & \text{kun } a < x < b, \\ 0, & \text{muuten.} \end{cases}$$

Huomaa, että tasajakauman tiheysfunktio on vain paloittain jatkuva; se on epäjatkuva pisteissä $x = a$ ja $x = b$ ja jatkuva muualla. Aiemmin mainitun merkintäsopimuksen mukaan tämä tiheysfunktio voidaan ilmaista myös kaavalla

$$f(x) = \frac{1}{b-a}, \qquad \text{kun } a < x < b.$$

Tällöin pitää ymmärtää asiayhteydestä, että $f(x) = 0$, kun $x \le a$ tai $x \ge b$.

Tasajakauman kertymäfunktio on

$$F(x) = \int_{-\infty}^{x} f(u)\,\mathrm{d}u = \begin{cases} 0, & \text{kun } x \le a, \\ \dfrac{x-a}{b-a}, & \text{kun } a < x < b, \\ 1, & \text{kun } x \ge b. \end{cases}$$

Eksponenttijakauma (engl. *exponential distribution*). Satunnaismuuttujalla X on eksponenttijakauma parametrilla $\lambda > 0$, $X \sim \mathrm{Exp}(\lambda)$, jos X:llä on tf

$$f(x) = f(x \mid \lambda) = \begin{cases} \lambda\,\mathrm{e}^{-\lambda x}, & \text{kun } x > 0, \\ 0, & \text{kun } x \le 0. \end{cases}$$

Tiheysfunktio voidaan ilmoittaa lyhyemmin kaavalla

$$f(x) = \lambda\,\mathrm{e}^{-\lambda x}, \qquad \text{kun } x > 0.$$

Tarkistetaan varmuuden vuoksi, että tämä funktio f todellakin on tiheysfunktio. Ensiksi $f \ge 0$, ja toiseksi

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \int_{0}^{\infty} \lambda\,\mathrm{e}^{-\lambda x}\,\mathrm{d}x = \lim_{M\to\infty} \Big/_{0}^{M} (-\mathrm{e}^{-\lambda x}) = -\mathrm{e}^{-\lambda M} + 1 = 1,$$

joten f on tf.

Kun $x > 0$, niin eksponenttijakauman kf:lla on arvo

$$F(x) = \Big/_{0}^{x} (-\mathrm{e}^{-\lambda u}) = 1 - \mathrm{e}^{-\lambda x},$$

ja $F(x) = 0$, kun $x \le 0$.

2.8 Satunnaismuuttujan muunnos

Jos X on sm, ja $g : \mathbb{R} \to \mathbb{R}$ on funktio, niin myös $Y = g(X)$ on sm. Merkintä $Y = g(X)$ tarkoittaa sitä, että

$$Y(\omega) = g(X(\omega)), \qquad \text{kaikilla } \omega \in \Omega,$$

ts. Y on yhdistetty funktio $g \circ X$. Mitä voimme sanoa Y:n jakaumasta? Diskreetille jakaumalle muunnoksen jakauma on helppo karakterisoida, mutta jatkuvan tapauksen käsittely on monimutkaisempaa.

Jos $g : \mathbb{R} \to \mathbb{R}$ on funktio ja $A \subset \mathbb{R}$, niin merkintä $g^{-1}(A)$ tarkoittaa joukon A alkukuvaa kuvauksessa g, eli

$$g^{-1}(A) = \{x \in \mathbb{R} : g(x) \in A\}.$$

Huomaa, että

$$g(x) \in A \quad \Longleftrightarrow \quad x \in g^{-1}(A).$$

Lause 2.10. *Olkoon X diskreetti sm ptnf:lla f_X, ja olkoon $g : \mathbb{R} \to \mathbb{R}$ funktio. Tällöin sm $Y = g(X)$ diskreetti, ja sen ptnf f_Y on*

$$f_Y(y) = \sum_{x \in g^{-1}(\{y\})} f_X(x).$$

Todistus. Koska X:n arvojoukko on korkeintaan numeroituvasti ääretön, voi Y saada korkeintaan numeroituvasti äärettömän määrän eri arvoja. Siksi Y on diskreetti. Mille tahansa $y \in \mathbb{R}$ on

$$\begin{aligned} P(Y = y) &= P(g(X) = y) = P(g(X) \in \{y\}) = P(X \in g^{-1}(\{y\})) \\ &= \sum_{x \in g^{-1}(\{y\})} f_X(x), \end{aligned}$$

jossa viimeinen vaihe seuraa täysadditiivisuudesta. □

Huomautus. Jos yhtälöllä $y = g(x)$ on yksikäsitteinen ratkaisu $x = h(y)$ kaikilla y, jotka kuuluvat sm:n Y arvojoukkoon S_Y, niin

$$f_Y(y) = f_X(h(y)), \qquad y \in S_Y. \tag{2.8}$$

Jos taas ratkaisuja on useita, niin jokainen erisuuri ratkaisu antaa ptnf:n arvoon pisteessä y yhden tällaisen termin lisää.

Jos X:llä on jatkuva jakauma, niin sen muunoksen $Y = g(X)$ jakauma voi funktion g luonteesta riippuen joko diskreetti, jatkuva tai ei kumpaakaan näistä. Esimerkiksi, jos g on porrasfunktio, niin $Y = g(X)$ on diskreetti satunnaismuuttuja. Myöhemmin tarkastelemme tapausta, jossa satunnaismuuttujalla $Y = g(X)$ on jatkuva jakauma. On myös helppo konstruoida esimerkkejä, joissa muunnoksella $Y = g(X)$ ei ole diskreetti eikä jatkuva jakauma, vaan ns. sekatyypin jakauma.

Esimerkki 2.6. Sensuroinnin kautta syntyvä sekatyypin jakauma. Olkoon X:llä välin $[0, 4]$ tasajakauma $U(0, 4)$. Määritellään, että $Y = g(X)$, jossa

$$g(x) = \min(\max(1, x), 2).$$

Tällöin voidaan sanoa, että Y saadaan satunnaismuuttujasta X sensuroimalla (tai rajaamalla) sitä vasemmalta pisteessä 1 ja oikelta pisteessä 2. Helpolla laskulla nähdään, että Y:n kf voidaan esittää konveksina kombinaationa tietyn diskreetin jakauman kf:sta sekä tietyn jatkuvan jakauman kf:sta, nimittäin

$$F_Y = \frac{3}{4} F_d + \frac{1}{4} F_c.$$

Tässä F_d on seuraava diskreetin jakauman kf,

$$F_d(x) = \begin{cases} 0 & \text{kun } x < 1, \\ \frac{1}{3} & \text{kun } 1 \le x < 2, \\ \frac{2}{3} & \text{kun } x \ge 2, \end{cases}$$

ja F_c on välin $[1, 2]$ tasajakauman kf. (Konveksi kombinaatio on sellainen lineaarikombinaatio, jossa painot [edellä $\frac{3}{4}$ ja $\frac{1}{4}$] ovat ei-negatiivisia ja niiden summa on yksi.) Tällöin sanotaan, että Y:llä on sekatyypin jakauma.

Sensuroituja jakaumia esiintyy käytännössä esim. elinaika-analyysissa. △

2.9 Kvantiilifunktio ja sen käyttö simuloinnissa

Tässä jaksossa tarkastelemme sellaista jatkuvaa jakaumaa, jonka tiheysfunktio f on aidosti positiivinen välillä (a, b) ja nolla muualla. Päätepisteille sallitaan reaalisten arvojen lisäksi arvot $a = -\infty$ tai $b = \infty$.

Tällaisen jakauman kertymäfunktio F, on aidosti kasvava välillä (a,b), ja lisäksi $F(a)=0$ ja $F(b)=1$. Tämän takia kertymäfunktiolla F on olemassa käänteisfunktio $F^{-1} : (0,1) \to (a,b)$. (Tarkemmin sanoen F:n rajoittumalla joukkon (a,b) on käänteisfunktio F^{-1}.) Tässä yhteydessä merkintä F^{-1} ei siis tarkoita funktiota $1/F$, vaan kertymäfunktion käänteisfunktiota. Käytännössä $F^{-1}(u)$ määritetään ratkaisemalla x yhtälöstä

$$F(x) = u, \qquad 0 < u < 1.$$

Funktio F^{-1} toteuttaa identiteetit

$$F^{-1}(F(x)) = x, \quad \text{kaikilla} \quad a < x < b,$$

ja

$$F(F^{-1}(u)) = u, \quad \text{kaikilla} \quad 0 < u < 1.$$

Tällaista (avoimella välillä $(0,1)$ määriteltyä) kertymäfunktion käänteisfunktiota F^{-1} kutsutaan kyseisen jakauman *kvantiilifunktioksi* (engl. *quantile function*). (Myös nimitys fraktiilifunktio on käytössä.)

Kvantiilifunktion määritelmästä seuraa, että mielivaltaiselle $0 < u < 1$ osuus u jakauman todennäköisyysmassasta jää pisteen $F^{-1}(u)$ vasemmalle puolelle eli vasemmanpuoleiseen häntään. Toisin sanoen, jos sm:lla X on kf F, niin

$$P(X \le F^{-1}(u)) = F(F^{-1}(u)) = u, \qquad \text{mielivaltaiselle } 0 < u < 1.$$

Ts. tiheysfunktion f alle jää pisteen $F^{-1}(u)$ vasemmalle puolelle alue, jonka pinta-ala on u, sillä tiheysfunktion avulla ilmaistuna

$$u = P(X \le F^{-1}(u)) = \int_{-\infty}^{F^{-1}(u)} f(x)\,\mathrm{d}x.$$

Koska jakauma on jatkuva, niin pisteen $F^{-1}(u)$ oikealle puolelle jää jakauman todennäköisyysmassasta osuus $1-u$, eli tiheysfunktion alle jäävän oikeanpuoleisen häntäalueen pinta-ala on $1-u$. Tilannetta havainnollistaa kuva 2.3.

Kvantiilifunktio määritellään tarkastelemalla jakauman vasemmanpuoleista häntää. Tilastollisessa päättelyssä määritetään testien ns. kriittisiä pisteitä usein jakauman oikeanpuoleisen hännän avulla. Selvyyden vuoksi voimme tällöin puhua jakauman *alakvantiileista* ja *yläkvantiileista*. Jos $0 < u < 1$, niin jakauman u-alakvantiili (eli u-kvantiili tai u-fraktiili) on sellainen piste, josta vasemmalle jää jakaumasta osuus u. Ts. u-alakvantiili on $F^{-1}(u)$. Jakauman u-yläkvantiili taas on sellainen piste, josta oikealle jää jakaumasta osuus u, joten u-yläkvantiili on $F^{-1}(1-u)$, sillä

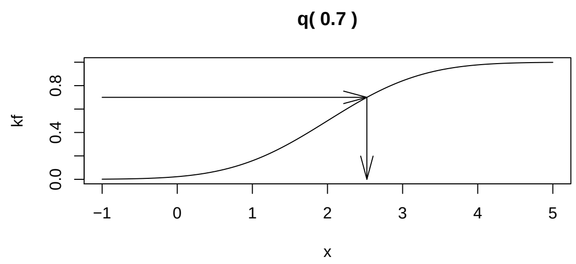
$$P(X \ge F^{-1}(1-u)) = 1 - P(X < F^{-1}(1-u)) = 1 - (1-u) = u.$$

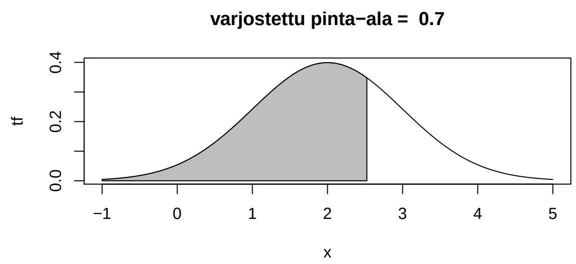
Jakauman tietyille (ala-)kvantiileille käytetään erikoisnimityksiä:

- mediaani on $\frac{1}{2}$-kvantiili,
- alakvartiili on $\frac{1}{4}$-kvantiili,
- yläkvartiili on $\frac{3}{4}$-kvantiili.

Olkoon X sm, jolla on kertymäfunktio F, ja tarkastellaan satunnaismuuttujaa $F(X)$. Koska F on aidosti kasvava funktio, niin myös F^{-1} on aidosti kasvava funktio. Jos sitä sovelletaan

Kuva 2.3 Erään jatkuvan jakauman kvantiilifunktion $q = F^{-1}$ arvo pisteessä $u = 0.7$. Ylemmässä kuvassa on piirretty jakauman kertymäfunktio ja alemmassa kuvassa sen tiheysfunktio.





epäyhtälön molemmille puolille, niin epäyhtälön ratkaisujoukko säilyy muuttumattomana. Jos $0 < u < 1$, niin

$$\begin{aligned} P[F(X) \le u] = P[F^{-1}(F(X)) \le F^{-1}(u)] &= P[X \le F^{-1}(u)] \\ &= F(F^{-1}(u)) = u. \end{aligned}$$

Tämä tarkoittaa sitä, että satunnaismuuttujalla $F(X)$ on tuttu jakauma,

$$F(X) \sim \mathrm{U}(0, 1). \tag{2.9}$$

Satunnaismuuttujaan $F(X)$, ja tulokseen $F(X) \sim \mathrm{U}(0, 1)$ viitataan toisinaan nimellä *kertymäfunktiomuunnos* (engl. *probability integral transform*).

Edellinen tulos voidaan kääntää. Olkoon $U \sim \mathrm{U}(0, 1)$, olkoon F sellaisen jakauman kf, jolla on tämän jakson alussa mainitut ominaisuudet, ja tarkastellaan satunnaismuuttujaa $Z = F^{-1}(U)$. Jos $a < x < b$, niin

$$\begin{aligned} P(Z \le x) = P[F^{-1}(U) \le x] &= P[F(F^{-1}(U)) \le F(x)] \\ &= P[U \le F(x)] = F(x) \end{aligned}$$

Siis

$$U \sim \mathrm{U}(0, 1) \quad \Rightarrow \quad \text{sm:n } F^{-1}(U) \text{ kertymäfunktio on } F.$$

Tällöin sanotaan, että sm $Z = F^{-1}(U)$ on saatu *käänteisfunktiomenetelmällä* (engl. (esim.) *inverse transform*).

Matemaattisissa ohjelmistoissa on yleensä satunnaislukugeneraattori, jonka tuottamia arvoja voidaan pitää riippumattomina jakaumaa $\mathrm{U}(0, 1)$ noudattavien satunnaismuuttujien arvoina. Jos jakauman kvantiilifunktiolla on yksinkertainen lauseke, niin käänteisfunktiomenetelmä antaa kätevän keinon simuloida kyseistä jakaumaa.

Esimerkki 2.7. Jakauman Exp(1) simulointi käänteisfunktiomenetelmällä. Jakauman kf on

$$F(x) = \begin{cases} 1 - \mathrm{e}^{-x}, & \text{jos } x \geq 0 \\ 0, & \text{muuten}, \end{cases}$$

Kvantiilifunktion arvo $F^{-1}(u)$ pisteessä $0 < u < 1$ saadaan ratkaisemalla x yhtälöstä

$$F(x) = u \quad \Leftrightarrow \quad x = -\ln(1-u).$$

Tämän takia jakaumaa Exp(1) voidaan simuloida asettamalla $X = -\ln(1-U)$, kun U on ensin generoitu jakaumasta $\mathrm{U}(0,1)$. △

Täydentäviä huomautuksia

Johdimme edellä käänteisfunktiomenetelmän asettamalla F:lle oletuksia. Itse asiassa nämä oletukset ovat tarpeettomia, ja menetelmä yleistyy mille tahansa kertymäfunktioille ts. lauseen 2.1 kolme ominaisuutta (a), (b) ja (c) toteuttavalle funktioille. Tätä varten kvantiilifunktion määritelmää pitää ensin yleistää.

Määritelmä 2.9 (Kertymäfunktion yleistetty käänteisfunktio, kvantiilifunktio)**.** Olkoon F mielivaltainen kf. Sen *yleistetty käänteisfunktio* määritellään kaavalla

$$F^{-1}(u) = \inf\{x : F(x) \geq u\}, \qquad 0 < u < 1. \tag{2.10}$$

Sanomme, että yleistetty käänteisfunktio F^{-1} on kertymäfunktiota F vastaava *kvantiilifunktio*, tai että F^{-1} on kyseisen *jakauman kvantiilifunktio*.

Edellä merkintä $\inf B$ tarkoittaa joukon $B \subset \mathbb{R}$ suurinta alarajaa. Koska F on kasvava ja oikealta jatkuva, tässä määritelmässä esiintyvä joukko $\{x : F(x) \geq u\}$ on muotoa $[t, \infty)$ oleva rajoittamaton väli, jonka suurin alaraja on välin vasen päätepiste t.

Lause 2.11. *Jos F on mielivaltainen kf, ja $U \sim \mathrm{U}(0,1)$, niin satunnaismuuttujalla $F^{-1}(U)$ on jakauma, jonka kertymäfunktio on F.*

Todistus. Lauseen erikoistapaus todistettiin jo. Yleisen tapauksen todistus sivuutetaan. □

Edellä määrittelimme (ala-)kvantiilin ainoastaan tämän jakson alussa kerrotussa jatkuvan jakauman tilanteessa. Yleisemmässä tapauksessa (ala-)kvantiili määritellään siten, että mikä tahansa luku x_p, joka toteuttaa ominaisuudet

$$P(X \leq x_p) \geq p \quad \text{ja} \quad P(X < x_p) \leq p$$

on X:n jakauman p-kvantiili. On helppo nähdä, että $F^{-1}(p)$ (jossa F^{-1} on nyt X:n kf:n yleistetty käänteisfunktio) on jakauman p-kvantiileista pienin.

2.10 Muunnetun satunnaismuuttujan tiheys

Tarkastelemaan tapausta, jossa X:llä on jatkuva jakauma, ja $g : \mathbb{R} \to \mathbb{R}$. Satunnaismuuttujalla $Y = g(X)$ ei tällöin välttämättä ole jatkuva jakauma, kuten olemme edellä nähneet.

Jos kuitenkin g on riittävän sileä ja aidosti monotoninen, tai jos sen määrittelyjoukko $\mathbb{R}$ voidaan osittaa paloihin, joissa g on riittävän sileä ja aidosti monotoninen, niin osoittautuu, että tällöin Y:llä on jatkuva jakauma. Jakauman tiheysfunktion lauseke on mahdollista johtaa laskemalla ensin sen kertymäfunktio ja sitten kertymäfunktion derivaatta. Tämän lisäksi pitää jollakin tavalla varmistaa, että Y:n jakauma todellakin on jatkuva, koska ainoastaan tässä tapauksessa tiheysfunktio saadaan kertymäfunktion derivaattana (vrt. lause 2.6). Paitsi derivoimalla kertymäfunktiota, tiheysfunktio on mahdollista johtaa käyttämällä muuttujanvaihtoa integraalissa.

Esimerkki 2.8. Olkoon X:llä jatkuva jakauma, jolla on jatkuva tiheysfunktio f_X. Johdamme satunnaismuuttujan $Y = \mathrm{e}^X$ tiheysfunktion kahdella erilaisella tavalla.

Tapa 1. (Lasketaan kertymäfunktio ja derivoidaan se.) Selvästi $Y > 0$, joten $F_Y(y) = 0$, kun $y < 0$. Kun $y > 0$, on

$$F_Y(y) = P(Y \le y) = P(\mathrm{e}^X \le y) = P(X \le \ln(y)) = F_X(\ln(y)),$$

joten arvaamme, että tiheysfunktioksi kelpaa

$$f_Y(y) = F_Y'(y) = f_X(\ln(y))/y, \qquad \text{kun } y > 0,$$

ja $f_Y(y) = 0$, kun $y \le 0$. Tarkistamme lopuksi lauseen 2.7 avulla, että Y:n jakauma on jatkuva. Tämän jälkeen edellä johdettu kaava kelpaa jakauman tiheysfunktioksi. Koska X:llä on jatkuva jakauma, sen kertymäfunktio F_X on jatkuva funktio. Edellä johdettujen kaavojen perusteella on selvää, että Y:n kertymäfunktio $F_Y(y)$ on jatkuva ja jatkuvasti derivoituva kaikkialla muualla paitsi mahdollisesti pisteessä $y = 0$. Y:n jakauma voidaan lauseen 2.7 nojalla todeta jatkuvaksi, jos osoitetaan, että $F_Y(y)$ on jatkuva pisteessä $y = 0$. Tämä on ilmeistä, sillä

$$F_Y(0+) = \lim_{y \to 0+} F_X(\ln(y)) = F_X(-\infty) = 0 = F_Y(0) = F_Y(0-).$$

Tapa 2. (Tehdään muuttujanvaihto integraalissa.) Tarkistetaan, että tn $P(a \le Y \le b)$ saadaan laskettua integroimalla tiettyä funktiota f_Y välin (a, b) yli, kun $a < b$ ovat mielivaltaisia pisteitä. Koska $Y > 0$, voidaan lisäksi rajoittua tilanteeseen $0 < a < b$.

$$\begin{aligned}
P(a \le Y \le b) &= P(a \le \mathrm{e}^X \le b) \\
&= P(\ln a \le X \le \ln b) \\
&= \int_{\ln a}^{\ln b} f_X(x)\,\mathrm{d}x && (X\text{:n tf on } f_X) \\
&= \int_a^b f_X(\ln(y)) \frac{1}{y}\,\mathrm{d}y && (\text{Sij. } y = \mathrm{e}^x \Leftrightarrow x = \ln(y).)
\end{aligned}$$

$\triangle$

Tämän esimerkin integrointiin nojautuva päättelytapa voidaan yleistää seuraavassa lauseessa väitettyyn muotoon. Tämä tapa yleistyy helposti useampaan ulottuvuuteen toisin kuin kertymäfunktiota käyttävä tapa. Kerrataan kuitenkin ensin yksi käsite.

Määritelmä 2.10 (Diffeomorfismi)**.** Kuvaus $g : A \to B$, jossa A ja B ovat d-ulotteisen avaruuden $\mathbb{R}^d$ avoimia osajoukkoja, on diffeomorfismi, jos

(a) g on bijektio joukkojen A ja B välillä

(b) sekä g että sen käänteisfunktio $g^{-1} : B \to A$ ovat jatkuvasti derivoituvia.

Lause 2.12 (Muuttujanvaihtokaava tiheysfunktioille)**.** *Olkoon sm:lla X jatkuva jakauma tf:lla f_X. Olkoon $g : A \to B$ diffeomorfismi, jossa $A, B \subset \mathbb{R}$ ovat avoimia välejä, ja olkoon*

$$P(X \in A) = 1.$$

Määritellään $h(y) = g^{-1}(y)$, kun $y \in B$. Tällöin sm:lla $Y = g(X)$ on jatkuva jakauma tiheysfunktiolla

$$f_Y(y) = \begin{cases} f_X(h(y))\,|h'(y)|, & \text{kun } y \in B \\ 0, & \text{muuten.} \end{cases} \tag{2.11}$$

Huomautus. Tulos (2.11) voidaan ilmaista lyhyemmin kaavalla

$$f_Y(y) = f_X(h(y))\,|h'(y)|, \qquad \text{kun } y \in B.$$

Vertaa jatkuvan jakauman tapauksen tulosta (2.11) vastaavan diskreetin tapauksen tulokseen (2.8). Vain jatkuvan jakauman tapauksessa tarvitaan käänteisfunktion derivaatan itseisarvoa.

Todistus. Koska g on bijektio avoimelta väliltä avoimelle välille, se on joko aidosti kasvava tai aidosti vähenevä funktio. Oletetaan ensin, että g on aidosti kasvava. Tällöin myös sen käänteisfunktio h on aidosti kasvava. Jos $a < b$ ja $a, b \in B$, niin

$$P(a \le Y \le b) = P(a \le g(X) \le b) = P(h(a) \le X \le h(b)) = \int_{h(a)}^{h(b)} f_X(x)\,\mathrm{d}x.$$

Tekemällä integraalissa muuttujanvaihdos

$$y = g(x) \qquad \Longleftrightarrow \qquad x = h(y)$$

saadaan

$$P(a \le Y \le b) = \int_a^b f_X(h(y))\,h'(y)\,\mathrm{d}y.$$

(Tässä kohdassa sovelletaan muuttujanvaihtoa Lebesguen integraaliin; lauseen oletuksilla käsiteltävät funktiot eivät välttämättä olisi Riemann-integroituvia.) Koska tämä pätee kaikille väleille $[a, b] \in B$ ja koska $P(Y \in B) = 1$, ja koska $h' > 0$, niin väite on todistettu siinä tapauksessa, että g on aidosti kasvava.

Jos taas g on aidosti vähenevä, niin $h = g^{-1}$ on myös aidosti vähenevä, joten siis $h' < 0$. Edellinen lasku pätee pienten muutosten jälkeen: epäyhtälön suunta vaihtuu, mikä vaihtaa integraalin merkin, mutta tämä kompensoidaan ottamalla h:n derivaatasta itseisarvo. $\square$

Kuten jakson avaavasta esimerkistä nähdään, muunnetun satunnaismuuttujan tiheysfunktio voidaan johtaa kahdella erilaisella tekniikalla:

- johdetaan kertymäfunktio ja lasketaan tiheysfunktio kertymäfunktion derivaattana,
- käytetään muuttujanvaihtokaavaa.

Kertymäfunktiotekniikassa pitää erikseen tarkistaa, että saatu kertymäfunktio on jatkuvan jakauman kertymäfunktio. Ehdoton edellytys tälle asialle on se, että kertymäfunktion pitää olla jatkuva funktio koko reaaliakselilla. Lauseen 2.7 mukaan riittävä ehto jakauman jatkuvuudelle on se, jos kertymäfunktio on lisäksi jatkuvasti derivoituva kaikkialla muualla paitsi äärellisessä määrässä poikkeuspisteitä. (Näissä poikkeuspisteissä derivaatan ei tarvitse olla olemassa.)

Muuttujanvaihtokaavaa käytettäessä pitää miettiä, mitkä ovat avoimet joukot A ja B, ja lisäksi käänteisfunktio h pitää pystyä ratkaisemaan (ja tämä ei onnistu ellei kyseessä ole bijektio). Tämän jälkeen on tyypillisesti helppo tarkistaa, onko lauseke $g(x)$ jatkuvasti derivoituva joukossa A ja onko lauseke $h(y)$ jatkuvasti derivoituva joukossa B. Jos nämä ehdot täyttyvät, niin kyseessä on diffeomorfismi.

Huomautus. Kun muunnoksen $Y = g(X)$ tiheysfunktio ilmoitetaan, on tärkeää paitsi kertoa kaava $f_X(h(y))\,|h'(y)|$ myös ilmoittaa selkeästi kaavan pätevyysalue, eli se joukko B, jossa kyseinen kaava pätee.

Jos g on diffeomorfismi, niin muuttujanvaihtokaavan muistamista helpottaa seuraava muistisääntö. Yhtälön

$$f_X(x)\,|\mathrm{d}x| = f_Y(y)\,|\mathrm{d}y| \tag{2.12}$$

pitää säilyä bijektiivisessä muuttujanvaihdossa

$$y = g(x) \iff x = h(y).$$

Kun tiedosta (2.12) ratkaistaan $f_Y(y)$, saadaan

$$f_Y(y) = f_X(x)\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right| = f_X(h(y))\,|h'(y)|,$$

mutta tietenkin tämä kaava pätee vain g:n kuvajoukossa B.

Muistisääntöä

$$f_X(x)\,|\mathrm{d}x| = f_Y(y)\,|\mathrm{d}y|$$

voi yrittää ymmärtää differentiaaligeometrisen (tai infinitesimaalisen) järkeilyn avulla. Ks. kuvaa 2.4, joka havainnollistaa tilannetta kahdessa eri y-pisteessä. Jos y ja $y + \Delta y$ kuuluvat g:n kuvajoukkoon B, jossa $\Delta y > 0$ on pieni luku, niin

$$P(y \le Y \le y + \Delta y) = \int_y^{y+\Delta y} f_Y(y)\,\mathrm{d}y \approx f_Y(y)\,\Delta y.$$

Toisaalta tämä tn on sama kuin

$$P(y \le g(X) \le y + \Delta y).$$

Jos g on aidosti kasvava funktio, niin tässä esiintyvä kaksoisepäyhtälö voidaan ratkaista g:n käänteisfunktion h avulla muotoon

$$P(y \le g(X) \le y + \Delta y) = P\left(h(y) \le X \le h(y + \Delta y)\right) = \int_x^{x+\Delta x} f_X(x)\,\mathrm{d}x \approx f_X(x)\,\Delta x,$$

missä $x = h(y)$ eli $y = g(x)$, ja $\Delta x = h(y + \Delta y) - h(y)$. Kun nämä kaksi arviota yhdistetään, saadaan

$$f_Y(y)\,\Delta y \approx f_X(x)\,\Delta x.$$

Edeltävät arviot paranevat, kun Δy pienenee. Kun Δy:n annetaan supeta kohti nollaa, voidaan ajatella että päädytään kasvavalle muunnokselle g pätevään differentiaaliseen muistisääntöön

$$f_X(x)\,\mathrm{d}x = f_Y(y)\,\mathrm{d}y.$$

Edellä suuretta Δx voidaan vielä arvioida derivaatan avulla, sillä

$$\Delta x = h(y + \Delta y) - h(y) \approx h'(y)\,\Delta y.$$

Jos edelliset arviot yhdistetään, ja luku Δy supistetaan pois, päädytään kaavan (2.11) mukaiseen tulokseen

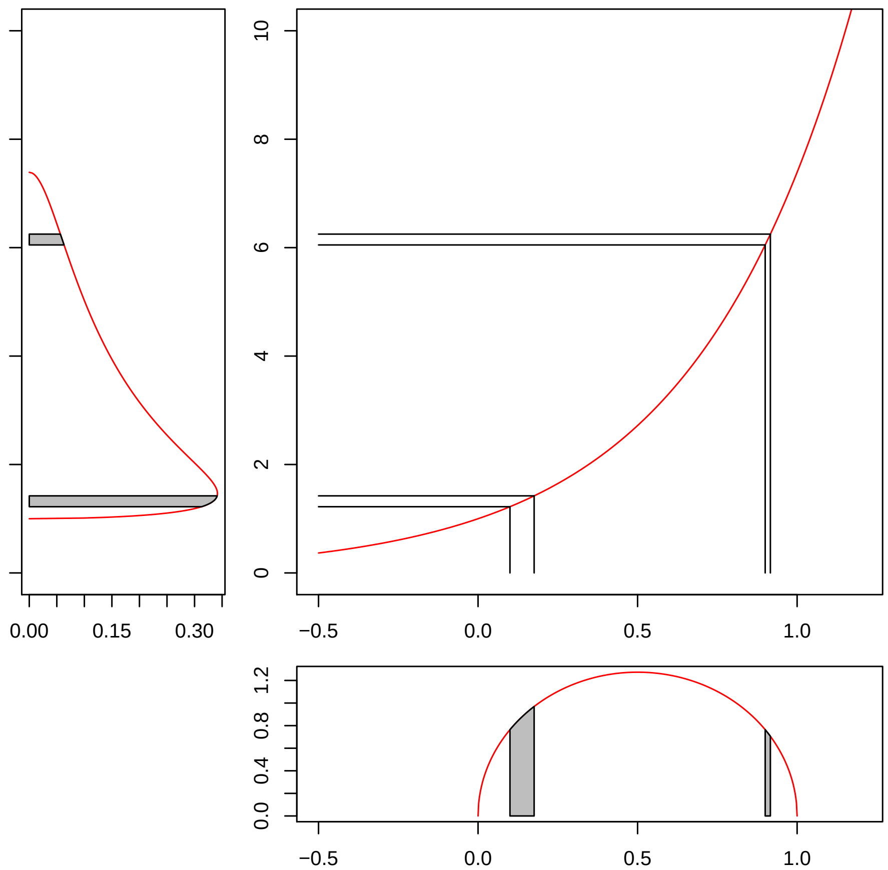
$$f_Y(y) = f_X(x)\,h'(y) = f_X(h(y))\,h'(y)$$

Tässä laskussa g oli kasvava, minkä takia ei tarvittu itseisarvoja lisäysten ympärillä tai derivaatan ympärillä. Jos sen sijaan g on vähenevä, itseisarvot ovat tarpeen.

Kuva 2.4 Kuvassa on esitetty tiheysfunktiot f_X ja f_Y sekä muunnoksen g kuvaaja. Kahdelle pisteelle y_1 ja y_2 on väritetty integraalia

$$\int_{y_i}^{y_i+\Delta y} f_Y(y)\,\mathrm{d}y$$

vastaava tiheysfunktion f_Y alle jäävä pinta-ala. Kuvassa on näytetty, kuinka nämä integraalit saadaan laskettua tiheysfunktion f_X integraalien avulla värittämällä vastinalueiden pinta-alat. Pisteet on valittu siten, että $f_X(h(y_1)) = f_X(h(y_2))$ ja luku Δy on sama pisteille y_1 ja y_2, joten ero arvoissa $f_Y(y_1)$ ja $f_Y(y_2)$ johtuu yksinomaan siitä, että $h'(y_1) \neq h'(y_2)$. Huomaa, kuinka tämä ero näkyy vastinalueiden leveyksissä.



Yhtälöstä (2.12) voidaan ratkaista $f_Y(y)$ myös toisella tavalla, nimittäin seuraavasti,

$$f_Y(y) = f_X(x)\frac{1}{\left|\frac{\mathrm{d}y}{\mathrm{d}x}\right|} = \frac{f_X(x)}{|g'(x)|} = \frac{f_X(h(y))}{|g'(h(y))|}.$$

Myös tämä kaava pitää paikkansa joukossa B, sillä kaava

$$\frac{\mathrm{d}x}{\mathrm{d}y} = \frac{1}{\frac{\mathrm{d}y}{\mathrm{d}x}}$$

kertoo oikean yhteyden funktion ja sen käänteisfunktion derivaattojen välillä, ja aidosta monotonisuudesta seuraa se, että $|g'| > 0$ joukossa A ja $|h'| > 0$ joukossa B.

Esimerkki 2.9. Käsittelemme uudestaan esimerkin, jossa $Y = \mathrm{e}^X$ ja X:llä on tf f_X. Tällä kertaa ratkaisemme tehtävän soveltamalla muistisääntöä

$$f_X(x)\,|\mathrm{d}x| = f_Y(y)\,|\mathrm{d}y|.$$

Tarkastelemme bijektiivistä vastaavuutta

$$y = \mathrm{e}^x \quad \Leftrightarrow \quad x = \ln(y)$$

pisteiden $x \in \mathbb{R}$ ja $y > 0$ välillä. Koska molemmat lausekkeet ovat jatkuvasti derivoituvia, on kyseessä diffeomorfismi. Ratkaisemme muistisäännöstä seuraavan kaavan $f_Y(y)$:lle,

$$f_Y(y) = f_X(x)\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right|$$

Koska tahdomme laskea f_Y:n arvon pisteessä y, pitää kaavan oikealla puolella x esittää y:n avulla kaavalla $x = \ln(y)$. Kaavassa tarvitaan myös tämän lausekkeen derivaattaa $1/y$. Kun teemme nämä sijoitukset, saamme tuloksen

$$f_Y(y) = f_X(\ln(y))\,\frac{1}{y}, \qquad y > 0.$$

Muistisäännöstä oltaisiin voitu johtaa myös kaava

$$f_Y(y) = f_X(x)\,\frac{1}{\left|\frac{\mathrm{d}y}{\mathrm{d}x}\right|}.$$

Tässä tarvitaan lausekkeen e^x derivaattaa, jonka avulla saadaan.

$$f_Y(y) = \frac{f_X(x)}{\mathrm{e}^x} = \frac{f_X(\ln(y))}{\mathrm{e}^{\ln(y)}} = \frac{f_X(\ln(y))}{y}, \qquad y > 0.$$

△

Joskus tiheysfunktion muuttujanvaihtokaavaa tarvitaan myös sellaisessa tilanteessa, jossa g ei ole monotoninen, mutta jossa sen määrittelyjoukko voidaan pilkkoa väleiksi, joille rajoitettuna g on aidosti monotoninen. Tällaiset tilanteet saadaan hoidettua samaan tapaan kuin seuraavassa esimerkissä.

Esimerkki 2.10. Olkoon X:llä jatkuva jakauma, jolla on jatkuva tiheysfunktio f_X, ja tarkastellaan satunnaismuuttujaa $Y = X^2$.

Tapa 1 (kertymäfunktiotekniikka) Lasketaan ensin Y:n kf. Selvästi $Y \geq 0$, joten $F_Y(y) = 0$, kun $y < 0$. Kun $y > 0$, on

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}).$$

Derivaatan laskemisen jälkeen arvaamme, että Y:n tf on

$$f_Y(y) = f_X(\sqrt{y})\,\frac{1}{2\sqrt{y}} + f_X(-\sqrt{y})\,\frac{1}{2\sqrt{y}}, \qquad \text{kun } y > 0,$$

ja $f_Y(y) = 0$, kun $y \leq 0$. Se seikan, että Y:n jakauma on jatkuva, voidaan tarkistaa lauseen 2.7 avulla. Funktio F_X on jatkuvan jakauman kertymäfunktiona jatkuva, joten ainoa ongelmallinen piste on $y = 0$. Tässä pisteessä pitää tarkistaa F_Y:n jatkuvuus, mutta tarkistus sujuu helposti, koska $F_X(x)$ on jatkuva pisteessä $x = 0$.

Tapa 2 (muuttujanvaihtokaava integraalissa) Olkoot $0 < a < b$ mielivaltaisia. Tällöin

$$\begin{aligned}
P(a \leq Y \leq b) &= P(a \leq X^2 \leq b) = P(\sqrt{a} \leq |X| \leq \sqrt{b}) \\
&= P(-\sqrt{b} \leq X \leq -\sqrt{a}) + P(\sqrt{a} \leq X \leq \sqrt{b}) \\
&= \int_{-\sqrt{b}}^{-\sqrt{a}} f_X(x)\,\mathrm{d}x + \int_{\sqrt{a}}^{\sqrt{b}} f_X(x)\,\mathrm{d}x.
\end{aligned}$$

Nyt ensimmäisessä integraalissa tehdään muuttujanvaihto

$$y = x^2 \qquad \Longleftrightarrow \qquad x = -\sqrt{y} \qquad (\text{jossa } x < 0 \text{ ja } y > 0)$$

ja toisessa

$$y = x^2 \qquad \Longleftrightarrow \qquad x = \sqrt{y} \qquad (\text{jossa } x, y > 0).$$

Tällöin saadaan

$$P(a \leq Y \leq b) = \int_a^b \left(f_X(-\sqrt{y})\,\frac{1}{2\sqrt{y}} + f_X(\sqrt{y})\,\frac{1}{2\sqrt{y}} \right) \mathrm{d}y.$$

Koska tulos on muotoa $\int_a^b f_Y(y)\,\mathrm{d}y$, tiheysfunktion f_Y kaava voidaan lukea tästä tuloksesta. △

Yleistämällä tämän esimerkin jälkimmäisen tavan päättely päädytään seuraavaan tulokseen.

Lause 2.13. *Olkoon sm:lla X jatkuva jakauma tiheysfunktiolla f_X ja olkoon $Y = g(X)$. Oletetaan, että $\mathbb{R}$ voidaan osittaa joukkoihin A_0 ja $A_1, A_2, \ldots$ siten, että*

- $P(X \in A_0) = 0,$
- *kukin A_i on avoin väli, kun $i \geq 1$ ja välejä $A_i, i \geq 1$ on äärellinen tai numeroituvasti ääretön määrä*
- *kuvaus $g_{|A_i}$, eli g rajoitettuna joukkoon A_i, on diffeomorfisimi avoimelta väliltä A_i sen kuvajoukolle B_i kaikilla $i \geq 1$*

Koska kukin kuvaus $g_{|A_i}$ on kääntyvä kun $i \geq 1$, on sillä käänteisfunktio, jota merkitään

$$h_i = (g_{|A_i})^{-1}, \qquad h_i : B_i \to A_i.$$

Tällöin Y:llä on jatkuva jakauma, jonka tiheysfunktio on

$$\sum_{i \geq 1} 1_{B_i}(y) \, f_X(h_i(y)) \, |h_i'(y)|.$$

Tässä termi

$$1_{B_i}(y) \, f_X(h_i(y)) \, |h_i'(y)|$$

tulkitaan nollaksi, jos $y \notin B_i$ (jolloin indikaattorin jälkeinen termi ei ole hyvin määritelty).

Todistus. Ideana on pilkkoa tapahtuma $\{a \leq Y \leq b\}$ erillisiin tapahtumiin

$$\{a \leq g(X) \leq b, X \in A_i\}, \qquad i \geq 0.$$

Kun $i = 0$, tämän tapahtuman tn on nolla. Muilla i tapahtuman todennäköisyys esitetään integraalina, jossa tehdään muuttujanvaihdos. □

Huomautus. Jos edellisen lauseen oletukset ovat voimassa, niin muunnoksen $Y = g(X)$ tiheysfunktio saadaan laskea myös johtamalla ensin sm:n Y kertymäfunktio F_Y ja sitten derivoimalla tämä funktio. Tämän lauseen avulla voidaan käsitellä monenlaisia muunnoksia, kuten vaikkapa $Y = |X|$ tai $Y = \sin(X)$, kun sm:lla X on jatkuva jakauma. Tiheysfunktion $f_Y(y)$ kaavaan saadaan kutakin yhtälön $y = g(x)$ ratkaisua kohti yksi termi, joka on samaa muotoa (2.11) kuin diffeomorfisen muunnoksen tapauksessa.

Esimerkki 2.11. Esimerkissä 2.10 käsiteltiin tapausta $Y = X^2$. Siinä yhteydessä tehtiin implisiittisesti valinnat

$$A_0 = \{0\}, \quad A_1 = (-\infty, 0), \quad A_2 = (\infty, 0),$$

jolloin

$$B_1 = (0, \infty), \quad B_2 = (0, \infty)$$

ja kuvaukset h_1 ja h_2 ovat

$$\begin{aligned} h_1(y) &= -\sqrt{y}, \qquad y > 0 \\ h_2(y) &= \sqrt{y}, \qquad y > 0. \end{aligned}$$

△

Luku 3

Yhteisjakauma

Edellisessä kappaleessa tarkastelimme aina yhtä satunnaismuuttujaa kerrallaan. Tässä kappaleessa näemme, miten aikaisemmat käsitteet yleistyvät, kun samalla perusjoukolla on määriteltynä useampi kuin yksi satunnaismuuttuja. Yksi päätavoitteista on määritellä, milloin satunnaismuuttujat ovat riippumattomia. Syvennämme moniulotteisten jakaumien käsittelyä myöhemmissä luvuissa.

3.1 Kaksiulotteinen satunnaisvektori ja sen jakauma

Yleistämme aluksi jakson 2.1 määritelmät kahteen dimensioon.

Määritelmä 3.1 (Kaksiulotteinen satunnaisvektori)**.** Olkoot X ja Y samalla perusjoukolla Ω määriteltyjä satunnaismuuttujia, $X : \Omega \to \mathbb{R}$ ja $Y : \Omega \to \mathbb{R}$. Tällöin (X, Y) on kaksiulotteinen *satunnaisvektori* (lyh. sv) (engl. *two-dimensional random vector; bivariate random vector*).

Kaksiulotteinen sv (X, Y) on siis kuvaus (funktio)

$$(X, Y) : \Omega \to \mathbb{R}^2, \quad \text{siten, että} \quad (X, Y)(\omega) = (X(\omega), Y(\omega)).$$

Satunnaisvektorista (X, Y) käytetään myös nimitystä kaksiulotteinen satunnaismuuttuja (engl. *two-dimensional random variable*) tai satunnaismuuttujapari.

Määritelmä 3.2 (Satunnaisvektorin jakauma, yhteisjakauma)**.** Satunnaisvektorin (X, Y) jakauma eli satunnaismuuttujien X ja Y *yhteisjakauma* (engl. *joint distribution*) on $\mathbb{R}^2$:n osajoukoilla A määritelty funktio

$$P((X, Y) \in A) = P(\{\omega \in \Omega : (X(\omega), Y(\omega)) \in A\}).$$

Ennen seuraavaa määritelmää tehdään seuraava **merkintäsopimus, jonka mukaan pilkku tarkoittaa leikkausta**. Kun lasketaan tapahtumien leikkauksen todennäköisyyttä, niin tapahtumien väliin voidaan symbolin $\cap$ sijasta merkitä pilkku, siis seuraavaan tapaan

$$P(X \in A, Y \in B) = P(\{X \in A, X \in B\}) = P(\{X \in A\} \cap \{Y \in B\}).$$

Tässä yhteydessä voidaan käyttää jaksossa 2.1 esiteltyjä merkintöjä. Erityisesti merkintä $P(X \leq x, Y \leq y)$ tarkoittaa seuraavaa

$$P(X \leq x, Y \leq y) = P(\{X \leq x \text{ ja } Y \leq y\}).$$

Määritelmä 3.3 (Satunnaisvektorin kertymäfunktio, yhteiskertymäfunktio)**.** Satunnaisvektorin (X, Y) kertymäfunktio eli satunnaismuuttujien X ja Y *yhteiskertymäfunktio* (lyh. ykf) (engl. *joint (cumulative) distribution function, joint cdf*) on

$$F_{X,Y}(x, y) = F(x, y) = P(X \le x, Y \le y), \qquad x, y \in \mathbb{R}\,.$$

Kuten yksiulotteisessa tapauksessa, niin myös kaksi- ja useampiulotteisessa tapauksessa kyseessä olevat satunnaismuuttujat voidaan ilmoittaa kertymäfunktion (tai muun jakauman esityksen) alaindekseillä, ellei asiayhteyden perusteella muuten ole selvää, minkä muuttujien yhteisjakauman esityksestä on kyse.

Satunnaismuuttujien X ja Y yhteisjakauma tietenkin määrää niiden yhteiskertymäfunktion, sillä

$$P(X \le x, Y \le y) = P((X, Y) \in (-\infty, x] \times (-\infty, y]).$$

Kääntäen, mikäli tunnetaan ykf, niin on mahdollista osoittaa, että sen avulla voidaan laskea *kaikkien* joukkojen A todennäköisyydet kaavassa

$$P((X, Y) \in A), \qquad A \subset \mathbb{R}^2,$$

eli yhteisjakauma määräytyy yhteiskertymäfunktiosta.

Lause 3.1. *Yhteiskertymäfunktio määrää jakauman.*

Todistus. Todistus siivutetaan, sillä se vaatisi mittateoriaa. □

Esimerkki 3.1. Olkoot $x_1 < x_2$ ja $y_1 < y_2$. Lasketaan ykf:n F avulla

$$P(x_1 < X \le x_2,\ y_1 < Y \le y_2) = P((X, Y) \in (x_1, x_2] \times (y_1, y_2]).$$

Nyt (piirrä kuva!)

$$\{X \le x_2,\ Y \le y_2\} = \{x_1 < X \le x_2,\ y_1 < Y \le y_2\} \cup C,$$

jossa joukot ovat erillisiä, ja $C = A \cup B$, jossa

$$A = \{X \le x_1, Y \le y_2\}, \quad B = \{X \le x_2, Y \le y_1\}.$$

Niinpä

$$\begin{aligned}
P(x_1 < X \le x_2,\ y_1 < Y \le y_2) &= F(x_2, y_2) - P(C) \\
&= F(x_2, y_2) - (P(A) + P(B) - P(A \cap B)) \\
&= F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1).
\end{aligned}$$

△

Kun tarkastellaan kahta satunnaismuuttujaa X ja Y, niin yksittäisen sm:n X (tai Y) (yksiulotteista) jakaumaa kutsutaan sen *reunajakaumaksi* (engl. *marginal distribution*). Näiden jakaumien kertymäfunktiot eli X:n ja Y:n *reunakertymäfunktiot* (engl. *marginal cdf's*) saadaan yhteiskertymäfunktiosta $F_{X,Y}$ raja-arvoina,

$$F_X(x) = P(X \le x, Y \in \mathbb{R}) = \lim_{y\to\infty} F_{X,Y}(x, y) = F_{X,Y}(x, \infty),$$

ja vastaavasti

$$F_Y(y) = \lim_{x\to\infty} F_{X,Y}(x, y) = F_{X,Y}(\infty, y).$$

Ykf:n sijasta kaksiulotteisia jakaumia käsitellään yleensä niiden yhteispistetodennäköisyysfunktioiden tai yhteistiheysfunktioiden avulla. Yhteisjakauma kuvaillaan usein kertolaskusäännön eli ketjusäännön avulla, mitä varten pitää ensin kuvailla yhden muuttujan reunajakauma sekä toisen muuttujan ehdollinen jakauma. Käymme seuraavaksi nämä asiat läpi diskreetin yhteisjakauman tapauksessa. Jatkuvan yhteisjakauman tapaus käsitellään myöhemmässä luvussa.

3.2 Diskreetti kaksiulotteinen jakauma

Määritelmä 3.4 (Diskreetti (yhteis)jakauma, yhteispistetodennäköisyysfunktio)**.** Satunnaisvektori (X, Y) on diskreetti, eli satunnaismuuttujilla X ja Y on *diskreetti yhteisjakauma*, jos sv:lla (X, Y) on enintään numeroituva arvojoukko, tai tarkemmin sanoen, on olemassa enintään numeroituva joukko $S \subset \mathbb{R}^2$ siten, että

$$P((X, Y) \in S) = 1.$$

Joukkoa S voidaan kutsua satunnaisvektorin (X, Y) arvojoukoksi tai sen jakauman kantajaksi. Tällöin sv:n (X, Y) pistetodennäköisyysfunktio eli satunnaismuuttujien X ja Y *yhteispistetodennäköisyysfunktio* (lyh. yptnf) (engl. *joint probability (mass) function, joint pmf*) on

$$f(x, y) = f_{X,Y}(x, y) = P(X = x, Y = y), \qquad x, y \in \mathbb{R}\,.$$

Yptnf:llä on analogiset ominaisuudet yksiulotteisen jakauman ptnf:n kanssa.

Lause 3.2. *Olkoon f ja S kuten määritelmässä. Tällöin*

a) $0 \le f(x, y) \le 1$, *ja* $f(x, y) = 0$, *kun* $(x, y) \notin S$.

b) f *määrää yhteisjakauman kaavalla*

$$P((X, Y) \in B) = \sum_{(x,y)\in B} f(x, y).$$

Todistus. Idea on sama kuin yksiulotteisessa tapauksessa. Kohdassa b summataan yhteen korkeintaan numeroituva määrä nollasta poikkeavia termejä, jotka ovat kaikki positiivisia, minkä takia summausjärjestyksestä ei tarvitse välittää, ja merkintä on mielekäs. □

Erityisesti X:n ja Y:n pistetodennäköisyydet saadaan summaamalla toinen muuttujista pois yptnf:stä,

$$f_X(x) = P(X = x) = P(X = x, Y \in \mathbb{R}) = \sum_y f(x, y), \tag{3.1}$$

$$f_Y(y) = P(Y = y) = P(X \in \mathbb{R}, Y = y) = \sum_x f(x, y). \tag{3.2}$$

Itse asiassa nämä tulokset sisältyvät kokonaistodennäköisyyden kaavaan (1.13) (tai sen yleistykseen numeroituvasti äärettömälle ositukselle).

Lisäksi yptnf summautuu ykköseksi, sillä

$$1 = P((X, Y) \in S) = \sum_{(x,y)\in S} f(x, y).$$

Seuraava lause on lauseen 2.5 kaksiulotteinen vastine. Sen todistus sivuutetaan.

Lause 3.3. *Olkoon $S \subset \mathbb{R}^2$ äärellinen tai numeroituvasti ääretön joukko, ja olkoon $f : S \to \mathbb{R}$ ei-negatiivinen funktio siten, että*

$$\sum_{(x,y)\in S} f(x, y) = 1.$$

Tällöin se on eräiden diskreettien satunnaismuuttujien X ja Y yptnf.

Seuraavassa lauseessa näytetään, että diskreetin yhteisjakauman tapauksessa sekä X ja Y ovat diskreettejä. Edellä on johdettu kaavat niiden (reuna)pistetodennäköisyysfunktioille.

Lause 3.4. *Satunnaismuuttujien X ja Y yhteisjakauma on diskreetti silloin ja vain silloin, kun reunajakaumat ovat diskreettejä.*

Todistus. Oletetaan ensin, että reunajakaumat ovat diskreettejä eli että sekä X että Y ovat diskreettejä satunnaismuuttujia. Olkoot niiden arvojoukot S_X ja S_Y, ts. $P(X \in S_X) = 1$ ja $P(Y \in S_Y) = 1$ ja $S_X \subset \mathbb{R}$ ja $S_Y \subset \mathbb{R}$ ovat korkeintaan numeroituvia joukkoja. Tällöin joukko $S = S_X \times S_Y \subset \mathbb{R}^2$ on myös korkeintaan numeroituva, ja helpolla laskulla nähdään, että

$$P((X, Y) \in S) = 1.$$

Siis yhteisjakauma on diskreetti.

Oletetaan seuraavaksi, että yhteisjakauma on diskreetti. Olkoon S korkeintaan numeroituva joukko siten, että $P((X, Y) \in S) = 1$. Tässä

$$S = \{(x'_1, y'_1), (x'_2, y'_2), \dots\}.$$

Olkoon nyt $S_X = \{x_1, x_2, \dots\}$, jossa $x_1, x_2, \dots$ on jonossa $x'_1, x'_2, \dots$ esiintyvät *eri suuret* arvot, ts. S_X on joukon S projektio x-akselille. Määritellään vastaavasti $S_Y = \{y_1, y_2, \dots\}$, jossa $y_1, y_2, \dots$ ovat jonossa $y'_1, y'_2, \dots$ esiintyvät eri suuret arvot. Nyt S_X ja S_Y ovat korkeintaan numeroituvia joukkoja, ja

$$P(X \notin S_X) \le P(X \notin S_X \text{ tai } Y \notin S_Y) = P((X, Y) \notin S) = 0.$$

Siis $P(X \in S_X) = 1$, ja vastaavalla laskulla $P(Y \in S_Y) = 1$, joten X ja Y ovat diskreettejä. □

Esimerkki 3.2. Reunajakaumat eivät määrää yhteisjakaumaa. Tarkastellaan diskreettiä yhteisjakaumaa, jonka yptnf:n kantaja S on

$$\{0, 1\} \times \{0, 1\}$$

ja yptnf on $f(x, y)$. Tällainen yptnf voidaan esittää taulukon muodossa seuraavasti,

$x \backslash y$	0	1	$\sum_y$
0	$f(0,0)$	$f(0,1)$	$f_X(0)$
1	$f(1,0)$	$f(1,1)$	$f_X(1)$
$\sum_x$	$f_Y(0)$	$f_Y(1)$	

Huomaa, että reunajakaumien ptnf:t ovat taulukon reunasummia ts.

$$f_X(x) = \sum_y f(x, y) = f(x, 0) + f(x, 1), \qquad x = 0, 1,$$
$$f_Y(y) = \sum_x f(x, y) = f(0, y) + f(1, y), \qquad y = 0, 1.$$

(Termi *reunajakauma* selittyy tällaisesta huomiosta.) Nyt tietenkin

$$X \sim \text{Bernoulli}(f_X(1)) \quad \text{ja} \quad Y \sim \text{Bernoulli}(f_Y(1)).$$

Moni erilainen yhteisjakauma tuottaa samat reunajakaumat. Esim. reunajakaumat $X \sim$ Bernoulli(1/2) ja $X \sim$ Bernoulli(1/2) saadaan valitsemalla $f(0,0) = f(1,1) = a$ ja $f(0,1) = f(1,0) = \frac{1}{2} - a$, jossa a on mikä tahansa luku väliltä $0 \le a \le \frac{1}{2}$. Taulukon muodossa esitettynä yptnf näyttää seuraavalta,

$x \backslash y$	0	1	$\sum$
0	a	$\frac{1}{2} - a$	$\frac{1}{2}$
1	$\frac{1}{2} - a$	a	$\frac{1}{2}$
$\sum$	$\frac{1}{2}$	$\frac{1}{2}$	

△

3.3 Useampiulotteinen satunnaisvektori

Avaruuden $\mathbb{R}^n$ vektoria $(x_1, \dots, x_n)$ merkitään $\mathbf{x} = (x_1, \dots, x_n)$. Silloin, kun se esiintyy matriisien kanssa samassa lausekkeessa, se ymmäretään pystyvektoriksi eli $n \times 1$-matriisiksi,

$$\mathbf{x} = (x_1, \dots, x_n) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Määritelmä 3.5. Jos $X_1, \dots, X_n$ ovat samalla perusjoukolla Ω määriteltyjä satunnaismuuttujia, niin $\mathbf{X} = (X_1, \dots, X_n)$ on n-ulotteinen satunnaisvektori (lyh. sv). Se on kuvaus $\Omega \to \mathbb{R}^n$ siten, että

$$\mathbf{X}(\omega) = (X_1(\omega), \dots, X_n(\omega)) = \begin{bmatrix} X_1(\omega) \\ \vdots \\ X_n(\omega) \end{bmatrix}$$

Määritelmä 3.6 (Yhteiskertymäfunktio)**.** Satunnaismuuttujien $X_1, \dots, X_n$ yhteiskertymäfunktio (ykf) eli satunnaisvektorin $\mathbf{X} = (X_1, \dots, X_n)$ kertymäfunktio on

$$F_{X_1,\dots,X_n}(x_1, \dots, x_n) = F_{\mathbf{X}}(\mathbf{x}) = P(X_1 \le x_1, \dots, X_n \le x_n),$$

jossa $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$.

Satunnaisvektorin jakauma määritellään samoin kuin aikaisemmin, eli se on funktio

$$B \mapsto P(\mathbf{X} \in B), \qquad B \subset \mathbb{R}^n .$$

Kuten kaksiulotteisessa tapauksessa, ykf määrää jakauman myös dimensiossa n. Korkeissa dimensioissa ykf:a käytetään harvoin konkreettisissa laskuissa, vaan se on lähinnä teoreettinen apuväline.

Jos kaikki satunnaisvektorin $\mathbf{X}$ komponentit X_i ovat diskreettejä sm:ia, niin sen ptnf eli sm:ien X_i yhteispistetodennäköisyysfunktio (yptnf) on

$$f_{\mathbf{X}}(\mathbf{x}) = P(\mathbf{X} = \mathbf{x}) = P(X_1 = x_1, \dots, X_n = x_n). \tag{3.3}$$

3.4 Satunnaismuuttujien riippumattomuus

Määritelmä 3.7 (Kahden sm:n riippumattomuus)**.** Satunnaismuuttujat X ja Y ovat riippumattomia, jos

$$P(X \in A \text{ ja } Y \in B) = P(X \in A)\, P(Y \in B), \qquad \text{kaikilla } A, B \subset \mathbb{R}.$$

Tämä voidaan merkitä $X \perp\!\!\!\perp Y$.

Jos $X \perp\!\!\!\perp Y$ ja $P(Y \in B) > 0$, niin mille tahansa A pätee

$$\begin{aligned} P(X \in A \mid Y \in B) &= \frac{P(X \in A, Y \in B)}{P(Y \in B)} \\ &= \frac{P(X \in A)\, P(Y \in B)}{P(Y \in B)} = P(X \in A). \end{aligned}$$

Lasku voidaan tietenkin toistaa vaihtamalla X:n ja Y:n roolit, jolloin nähdään että

$$P(X \in A \mid Y \in B) = P(X \in A), \qquad P(Y \in B \mid X \in A) = P(Y \in B)$$

kaikille joukoille A ja B, joille edellä merkityt ehdolliset todennäköisyydet ovat hyvin määriteltyjä. Tämä voidaan tulkita siten, että jos $X \perp\!\!\!\perp Y$, niin tieto Y:n arvosta ei anna tietoa X:n jakaumasta ja kääntäen tieto X:n arvosta ei anna tietoa Y:n jakaumasta.

Lause 3.5. *Seuraavat seikat ovat yhtäpitäviä,*

(a) $X \perp\!\!\!\perp Y$

(b) Yhteiskertymäfunktio faktoroituu reunakertymäfunktioiden tuloksi, eli

$$F_{X,Y}(x, y) = F_X(x)\, F_Y(y) \qquad \textit{kaikilla } x, y \in \mathbb{R}.$$

Todistus. **(a) ⇒ (b):** Valitaan $A = (-\infty, x]$ ja $B = (-\infty, y]$ ja sovelletaan riippumattomuuden määritelmää.

Implikaation (b) ⇒ (a) todistaminen sivuutetaan, sillä todistus vaatisi mittateoriaa. □

Lause 3.6. *Jos X ja Y ovat diskreettejä satunnaismuuttujia, niin seuraavat seikat ovat yhtäpitäviä,*

(a) $X \perp\!\!\!\perp Y$

(b) Yhteispistetodennäköisyysfunktio faktoroituu reunapistetodennäköisyysfunktioiden tuloksi, eli

$$f_{X,Y}(x, y) = f_X(x)\, f_Y(y) \quad \textit{kaikilla } x, y.$$

Todistus. Implikaatio (a) ⇒ (b) seuraa valitsemalla riippumattomuuden määritelmässä $A = \{x\}$ ja $B = \{y\}$.

Todistetaan implikaatio (b) ⇒ (a) lähtemällä liikkeelle oletuksesta, että ominaisuus (b) on voimassa. Olkoot $A, B \subset \mathbb{R}$ mielivaltaisia. Olkoot sm X:n mahdolliset eri arvot $x_1, x_2, \ldots$ ja sm

Y:n mahdolliset eri arvot $y_1, y_2, \dots$ Tällöin

$$\begin{aligned}
P(X \in A, Y \in B) &= P\left(\bigcup_{\substack{x_i \in A \\ y_j \in B}} \{X = x_i, Y = y_j\}\right) \\
&= \sum_{x_i \in A, y_j \in B} P(X = x_i, Y = y_j) && \text{(täysadditiivisuus)} \\
&= \sum_{x_i \in A, y_j \in B} f_{X,Y}(x_i, y_j) && \text{(määritelmä)} \\
&= \sum_{x_i \in A, y_j \in B} f_X(x_i)\, f_Y(y_j) && \text{(oletus)} \\
&= \sum_{x_i \in A, y_j \in B} P(X = x_i)\, P(Y = y_j) && \text{(määritelmä)} \\
&= \sum_{x_i \in A} P(X = x_i) \sum_{y_j \in B} P(Y = y_j) && \text{(osittelulaki)} \\
&= P(X \in A)\, P(Y \in B),
\end{aligned}$$

joten X ja Y ovat riippumattomia. □

Huomautus. Myöhemmässä luvussa todistetaan, että jatkuvan yhteisjakauman tapauksessa satunnaismuuttujat ovat riippumattomia silloin ja vain silloin, kun niiden yhteistiheysfunktio voidaan esittää samalla tavalla reunajakaumien tiheysfunktioiden tulona.

Huomautus. Jos A ja B ovat tapahtumia, niin edellistä lausetta soveltamalla voidaan tarkistaa, että tapahtumat ovat riippumattomia silloin ja vain silloin kuin niiden indikaattorit ovat riippumattomia sm:ia, eli

$$A \perp\!\!\!\perp B \quad \Longleftrightarrow \quad 1_A \perp\!\!\!\perp 1_B.$$

Seuraavaksi näytetään, että riippumattomien satunnaisfunktioiden muunnokset ovat riippumattomia.

Lause 3.7. *Jos $X \perp\!\!\!\perp Y$, ja $g : \mathbb{R} \to \mathbb{R}$ sekä $h : \mathbb{R} \to \mathbb{R}$ ovat funktioita, niin myös $g(X) \perp\!\!\!\perp h(Y)$.*

Todistus. Olkoot $A, B \subset \mathbb{R}$ mielivaltaisia. Tällöin

$$\{g(X) \in A, h(Y) \in B\} = \{X \in g^{-1}(A), Y \in h^{-1}(B)\}.$$

Tässä $g^{-1}(A)$ on joukon A alkukuva kuvauksessa g, ja $h^{-1}(B)$ on joukon B alkukuva kuvauksessa h. Koska $X \perp\!\!\!\perp Y$, on

$$\begin{aligned}
P(g(X) \in A, h(Y) \in B) &= P(X \in g^{-1}(A), Y \in h^{-1}(B)) \\
&= P(X \in g^{-1}(A))\, P(Y \in h^{-1}(B)) \\
&= P(g(X) \in A)\, P(h(Y) \in B).
\end{aligned}$$

□

Määritelmä 3.8 (Usean sm:n riippumattomuus)**.** Satunnaismuuttujat $X_1, \dots, X_n$ ovat riippumattomia, jos

$$P\Big(\bigcap_{i=1}^{n} \{X_i \in B_i\}\Big) = \prod_{i=1}^{n} P(X_i \in B_i), \qquad \text{kaikilla } B_1, \dots, B_n \subset \mathbb{R}.$$

Tämä voidaan merkitä $X_1, \dots, X_n \perp\!\!\!\perp$.

Esimerkki 3.3. Binomijakauman synty toistetussa Bernoullin kokeessa. Olkoot $Y_1, \ldots, Y_n$ riippumattomia Bernoullin jakaumaa noudattavia satunnaismuuttujia, joilla on yhteinen onnistumistn $0 \le p \le 1$. Ts. kullakin niistä on ptnf:na

$$f(y) = p^y \, (1-p)^{1-y}, \qquad \text{kun } y = 0, 1.$$

Jos $Y_i = 1$ niin sanotaan, että onnistuttiin toistossa i, ja muuten sanotaan, että epäonnistuttiin toistossa i. Olkoon X onnistumisten lukumäärä n toistossa, eli

$$X = Y_1 + \cdots + Y_n = \sum_{i=1}^{n} Y_i.$$

Mikä on X:n jakauma?

Jos $(y_1, \ldots, y_n)$ on jokin jono nollia ja ykkösiä, niin

$$P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_n = y_n) = \prod_{i=1}^{n} p^{y_i} \, (1-p)^{1-y_i} = p^s \, (1-p)^{n-s},$$

jossa $s = \sum_{i=1}^{n} y_i$, eli s on ykkösten ja $n - s$ on nollien lukumäärä jonossa $(y_1, \ldots, y_n)$.

Satunnaismuuttuja X saa arvon $0 \le x \le n$ jos ja vain jos satunnaismuuttujat $Y_1, \ldots, Y_n$ saavat sellaiset arvot $y_1, \ldots, y_n$, että jonossa $(y_1, \ldots, y_n)$ on x ykköstä (ja $n-x$ nollaa). Tällaisia jonoja on $\binom{n}{x}$ kappaletta, ja niistä kunkin tn on $p^x \, (1-p)^{n-x}$. Todennäköisyyden additiivisuuden nojalla

$$f_X(x) = P(X = x) = \binom{n}{x} p^x \, (1-p)^{n-x}, \qquad \text{kun } x = 0, 1, \ldots, n.$$

Ts. $X \sim \mathrm{Bin}(n, p)$, eli X noudattaa binomijakaumaa otoskokoparametrilla n ja onnistumistodennäköisyydellä p. △

Huomautus. Kun määrittelimme useamman tapahtuman riippumattomuuden, jouduimme tarkastelemaan kaikkia jonoja $i_1, \ldots, i_k$ jossa indeksit ovat keskenään erisuuria. Satunnaismuuttujien riippumattomuuden määritelmässä tätä ei tarvita. Tämä johtuu siitä, että voimme poistaa tarkastelusta sellaiset sm:t

$$X_j, \quad \text{joilla} \quad j \notin \{i_1, \ldots, i_k\}$$

valitsemalla niiden kohdalla $B_j = \mathbb{R}$.

Huomautus. Jos $A_1, \ldots, A_n$ ovat tapahtumia, niin niiden indikaattorit $1_{A_1}, \ldots, 1_{A_n}$ ovat satunnaismuuttujia. Seuraavat kaksi seikkaa ovat yhtäpitäviä.

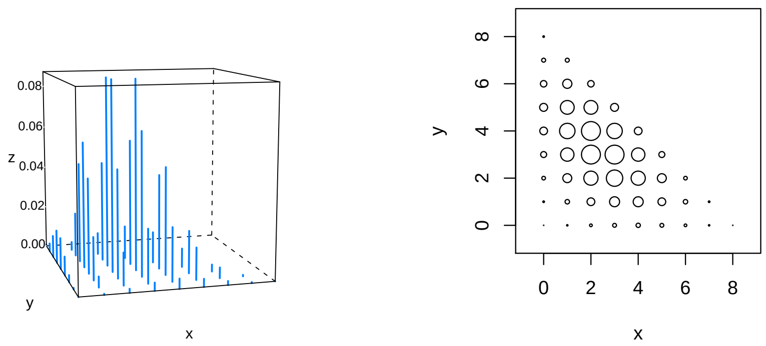
a) Tapahtumat $A_1, \ldots, A_n$ ovat riippumattomia,

b) satunnaismuuttujat $1_{A_1}, \ldots, 1_{A_n}$ ovat riippumattomia.

Numeroituvasti äärettömän monen satunnaismuuttujan $X_1, X_2, \ldots$ riippumattomuus määritellään seuraavasti.

Määritelmä 3.9 (Äärettömän monen sm:n riippumattomuus)**.** Satunnaismuuttujat $X_1, X_2, \ldots$ ovat riippumattomia, jos satunnaismuuttujat $X_1, \ldots, X_n$ ovat riippumattomia kaikilla $n \ge 2$.

Kuva 3.1 Trinomijakauman $\mathrm{Trin}(n,(p_1,p_2))$ yptfn, kun $n = 8$ ja $p_1 = 0.3$ ja $p_2 = 0.4$.



3.5 Trinomijakauma ja multinomijakauma

Trinomijakauma yleistää binomijakauman mutkattomasti kaksiulotteiseksi diskreetiksi jakaumaksi. Multinomijakauma on puolestaan trinomijakauman yleistys korkeampaan ulottuvuuteen.

Olkoot $p_1, p_2 > 0$ todennäköisyyksiä siten, että $p_1 + p_2 < 1$. Olkoon $n \ge 1$ positiivinen kokonaisluku. Tällöin

$$f(x,y) = \binom{n}{x, y, n-x-y} p_1^x \, p_2^y \, (1 - p_1 - p_2)^{n-x-y}, \qquad x, y \ge 0, \quad x + y \le n.$$

on *trinomijakauman* $\mathrm{Trin}(n,(p_1,p_2))$ yptnf. Kaavassa esiintyvä trinomikerroin (tai yleisemmin multinomikerroin) määriteltiin jaksossa 1.7.

Kuvassa 3.1 trinomijakauman yptnf on piirretty kahdella eri tavalla. Ensimmäisessä tavassa funktio $z = f(x,y)$ on esitetty perspektiivikuvana. Toisessa tavassa kolmas dimensio esitetään piirtosymbolin koon avulla: funktion $f(x,y)$ arvo on verrannollinen pisteeseen (x,y) piirretyn ympyrän pinta-alaan.

Se, että kyseessä on yptnf nähdään helposti multinomikaavan (1.8) avulla, sillä

$$\begin{aligned} 1 &= [p_1 + p_2 + (1 - p_1 - p_2)]^n \\ &= \sum_{x,y} \binom{n}{x, y, n-x-y} p_1^x \, p_2^y \, (1 - p_1 - p_2)^{n-x-y}, \end{aligned}$$

jossa summa otetaan niiden kokonaislukuparien (x,y) yli, joille $x, y \ge 0$ ja $x + y \le n$. Lisäksi jokainen summan termeistä on positiivinen.

Trinomijakaumalla on toistokokeen kautta syntyvä tulkinta. Tarkastellaan koetta, jonka yhdessä toistossa saadaan täsmälleen yksi tuloksista A, B tai C. Olkoon yhdessä toistossa

$$p_1 = P(A), \quad p_2 = P(B), \quad p_3 = P(C) = 1 - p_1 - p_2.$$

Toistetaan tätä koetta riippumattomasti n kertaa, ja annetaan satunnaismuuttujan Z_i kertoa lopputulos toistossa numero i siten, että

$$Z_i = \begin{cases} 1 & \text{jos tulos on } A, \\ 2 & \text{jos tulos on } B, \\ 3 & \text{jos tulos on } C. \end{cases}$$

Tällöin kaikilla i

$$P(Z_i = z) = p_z, \qquad \text{kun } z = 1, 2, 3.$$

Määritellään, että sm X on niiden i lkm, joilla $Z_i = 1$ (eli lopputuloksen A lkm) ja sm Y on niiden i lkm, joilla $Z_i = 2$ (eli lopputuloksen B lkm). Tällöin niitä i, joilla $Z_i = 3$ (eli lopputuloksen C lkm) on $n - X - Y$.

Jos $(z_1, z_2, \ldots, z_n)$ on arvoista 1, 2 tai 3 koostuva jono, niin

$$\begin{aligned} P(Z_1 = z_1, Z_2 = z_2, \ldots, Z_n = z_n) &= p_{z_1} p_{z_2} \ldots p_{z_n} \\ &= p_1^x p_2^y p_3^{n-x-y}. \end{aligned}$$

Tässä x on lopputuloksen A lkm annetussa jonossa eli niiden i lkm, joilla $z_i = 1$, ja y on lopputuloksen B lkm eli niiden i lkm, joilla $z_i = 2$. Tällöin lopputuloksen C lkm eli niiden i lkm, joille $z_i = 3$ on $n - x - y$. Toisaalta, jos kiinnitetään kokonaisluvut x ja y siten, että $x, y \ge 0$ ja $x + y \le n$, niin sellaisia jonoja $(z_1, z_2, \ldots, z_n)$ jotka koostuvat arvoista 1, 2 tai 3, joissa tulosta 1 on x kpl ja tulosta 2 on y kpl on multinomikertoimen

$$\binom{n}{x, y, n - x - y}$$

ilmaisema määrä, sillä kukin tälläinen jono vastaa täsmälleen yhtä tapaa osittaa n-alkionen joukko kolmeen osaan siten, että ensimmäiseen osaan tulee x kpl alkioita ja toiseen osaan y kpl alkioita (jolloin kolmanteen osaan jää $n - x - y$ kpl alkioita). Jokaiselle tällaiselle jonolle

$$P(Z_1 = z_1, Z_2 = z_2, \ldots, Z_n = z_n) = p_1^x p_2^y p_3^{n-x-y}.$$

Tällä tavalla nähdään, että

$$P(X = x, Y = y) = \binom{n}{x, y, n - x - y} p_1^x p_2^y p_3^{n-x-y},$$

jossa $p_3 = 1 - p_1 - p_2$. Toisin sanoen satunnaismuuttujien X ja Y yhteisjakauma on trinomijakauma otoskokoparametrilla n ja todennäköisyysparametreilla (p_1, p_2).

Toistokoetulkinnan valossa on ilmeistä, että sm:n X reunajakauma on $X \sim \text{Bin}(n, p_1)$. (Ajattele toistokoetta, jossa onnistutaan, kun saadaan lopputulos A ja muuten epäonnistutaan.) Vastaavasti, $Y \sim \text{Bin}(n, p_2)$.

Trinomijakauma saadaan suoraviivaisesti yleistettyä siihen tapaukseen, jossa tulosvaihtoehtoja voi olla vielä useampia. Olkoon $(p_1, \ldots, p_n)$ todennäköisyysvektori, eli kukin $p_i \ge 0$, ja

$$p_1 + \cdots + p_n = 1.$$

Multinomijakauman $\text{Mult}(k, (p_1, \ldots, p_n))$ ptnf on

$$f(x_1, \ldots, x_n) = \binom{k}{x_1, \ldots, x_n} p_1^{x_1} \ldots p_n^{x_n}, \tag{3.4}$$

kun $x_1 + \cdots + x_n = k$ ja nolla muuten.

Multinomijakauma syntyy, kun tarkastellaan k-kertaista toistokoetta, jossa kussakin toistossa toteutuu yksi n:sta vaihtoehdosta, jossa toistot ovat toisistaan riippumattomia. Muuttujat $x_1, \ldots, x_n$ ovat eri lopputulosten frekvenssit k:ssa toistossa. Kaava (3.4) perustellaan samalla tavalla kuin trinomijakauman tapauksessa; viimekädessä vedotaan jaksossa 1.5 esitettyyn multinomikertoimien kombinatoriseen luonnehdintaan.

Useimmiten multinomijakaumaa pidetään n-ulotteisena jakauma, mutta joissakin yhteyksissä on mielekkäämpää pitää sitä $(n - 1)$-ulotteisena jakauma (jolloin $X_n = k - X_1 - \cdots - X_{n-1}$). Tällä konventiolla binomijakauma ja trinomijakauma ovat multinomijakauman erikoistapauksia.

3.6 Satunnaisvektoreiden riippumattomuus

Kahden satunnaisvektorin $\mathbf{X}$ ja $\mathbf{Y}$ riippumattomuus määritellään täysin vastaavasti kuin satunnaismuuttujien kohdalla, ks. jakso 3.4.

Samalla päättelyllä kuin aikaisemmin,

$$\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \quad \Rightarrow \quad g(\mathbf{X}) \perp\!\!\!\perp h(\mathbf{Y}),$$

kun g ja h ovat mielivaltaisia vektoriargumentin reaaliarvoisia funktioita.

On hyvä huomata, että jos satunnaismuuttujat

$$X_1, X_2, \ldots, X_k; \quad Y_1, Y_2, \ldots, Y_m$$

ovat riippumattomia, niin tällöin niistä koostuvat satunnaisvektorit $\mathbf{X}$ ja $\mathbf{Y}$ ovat riippumattomia, jossa

$$\mathbf{X} = (X_1, \ldots, X_k), \qquad \mathbf{Y} = (Y_1, \ldots, Y_m).$$

Tällöin nimittäin satunnaisvektoreiden yhteiskertymäfunktio faktoroituu reunasatunnaisvektoreiden kertymäfunktioiden tuloksi, ts.

$$F_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y}) = F_{\mathbf{X}}(\mathbf{x})\, F_{\mathbf{Y}}(\mathbf{y}), \qquad \text{kaikilla } \mathbf{x}, \mathbf{y},$$

koska se faktoroituu peräti komponenttisatunnaismuuttujien kertymäfunktioiden tuloksi. Tämä riittää todistamaan, että $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}$.

Luku 4

Odotusarvo

Seuraavaksi kerrataan, miten satunnaismuuttujan odotusarvo määritellään diskreetissä ja jatkuvassa tapauksessa. Odotusarvolle käytetään englannikielisessä kirjallisuudessa lukuisia nimityksiä, kuten seuraavia: *mean (value)*, *(mathematical) expectation*, *expected value*. Satunnaisvektorin odotusarvo määritellään myöhemmässä kappaleessa.

4.1 Diskreetin satunnaismuuttujan odotusarvo

Määritelmä 4.1 (Diskreetin sm:n odotusarvo)**.** Jos X on diskreetti sm, jonka arvojoukko on $S = \{x_1, x_2, \dots\}$ ja ptnf on f, niin sen odotusarvo on reaaliluku

$$EX = E(X) = \sum_i x_i\, f(x_i) = \sum_x x\, f(x)$$

mikäli tämä sarja on itseisesti summautuva, eli mikäli $\sum_i |x_i|\, f(x_i) < \infty$. Muussa tapauksessa sanotaan että X:llä *ei ole odotusarvoa*. Se seikka, että X:llä on odotusarvo voidaan ilmaista myös sanomalla, että X on *integroituva* sm.

Edellä kysymys sarjan itseisestä summautumisesta on relevantti tietenkin vain silloin, kun summattavia arvoja on äärettömän monta. Palautetaan mieleen analyysistä se seikka, että jos sarja summautuu itseisesti, niin sille saadaan sama summa kaikilla summausjärjestyksillä. (Jos summa suppenee kohti jotakin reaalilukua, mutta ei suppene itseisesti, niin erilaisilla summausjärjestyksillä sarjalle voidaan saada erilaisia summia.) Itseisen suppenemisen takia edellisessä määritelmässä ei tarvitse kiinnittää tiettyä summausjärjestystä. Toinen tapaus, jossa sarjan summausjärjestyksellä ei ole väliä on se, jossa kaikki summan termit ovat ei-negatiivisia; tällöin sarjan summa voi olla jokin reaaliluku tai ∞.

Jos satunnaismuuttujilla X ja Y on sama diskreetti jakauma, niin niillä on sama odotusarvo. Sen takia diskreetin jakauman odotusarvo voidaan määritellä lukuna EX, jossa X on mikä tahansa kyseistä jakaumaa noudattava sm.

Odotusarvolla on *fysikaalinen tulkinta* massajakauman painopisteenä. Jos massattoman sauvan (=reaaliakseli) pisteisiin x_i asetetaan massat p_i, niin sauva pysyy tasapainossa, mikäli sitä tuetaan alta päin pisteestä EX. Tämä odotusarvon tulkinta säilyy myös muunlaisille jakaumille.

Edellä esitettyyn määritelmään voidaan päätyä myös todennäköisyyden *frekvenssitulkinnan* avulla. Ajatellaan diskreettiä satunnaismuuttujaa X, jonka mahdolliset arvot ovat $\{x_i\}$. Konkreettisuuden vuoksi olkoon X sinun voittamasi rahasumma tietyn uhkapelin yhdessä toistossa.

Jos toistat peliä riippumattomasti N kertaa (jossa N on hyvin suuri luku), niin voitat summan x_i osapuilleen $NP(X = x_i)$ kertaa, joten keskimääräinen voittosi yhtä peliä kohti on

$$\frac{1}{N} \sum_i x_i \, N \, P(X = x_i) = \sum_i x_i \, f(x_i) = EX.$$

Odotusarvon frekvenssitulkinnan mukaan satunnaismuuttujan X odotusarvo on osapuilleen sama kuin suuren määrän $X_1, \ldots, X_N$ keskenään riippumattomien ja X:n kanssa samoin jakautuneitten satunnaismuuttujien otoskeskiarvo, ts.

$$EX \approx \frac{1}{N} \sum_{i=1}^{N} X_i. \tag{4.1}$$

Kaavan oikealla puolella esintyy satunnaismuuttuja, jonka arvo suppenee ns. vahvan suurten lukujen lain perusteella melkein kaikilla ω kohti odotusarvoa EX, kun N kasvaa rajatta. Vahva suurten lukujen laki on puolestaan todennäköisyysteorian kuuluisimmista tuloksista. Odotusarvoa arvioidaan usein tietokonesimuloinneissa tällä periaatteella, eli laskemalla keskiarvo satunnaismuuttujien realisoituneista arvoista. Vahvan suurten lukujen lain tiedetään olevan voimassa, mikäli X:n odotusarvo on olemassa.

Esimerkki 4.1. Vakion odotusarvo. Vakiota $a \in \mathbb{R}$ voidaan pitää diskreettinä satunnaismuuttujana, joka saa aina arvon a. Vastaavaa jakaumaa sanotaan *degeneroituneeksi jakaumaksi.* Nyt

$$Ea = a \cdot 1 = a,$$

joten vakion odotusarvo on kyseinen vakio. △

Esimerkki 4.2. Tapahtuman A indikaattorin odotusarvo (eli Bernoullin jakauman odotusarvo). Jos $p = P(A)$ ja $X = 1_A$, niin $X \sim \text{Bernoulli}(p)$, ja

$$EX = 0 \cdot (1 - p) + 1 \cdot p = p.$$

Tapahtuman indikaattorin odotusarvo on sama kuin kyseisen tapahtuman todennäköisyys. △

Esimerkki 4.3. Binomijakauman odotusarvo. Olkoon $X \sim \text{Bin}(n, p)$, ja asetetaan $q = 1 - p$. Odotusarvo EX saadaan laskettua suoraan määritelmän nojalla seuraavasti.

$$\begin{aligned} EX &= \sum_{x=0}^{n} x f(x) = \sum_{x=1}^{n} x f(x) = \sum_{x=1}^{n} x \, \frac{n!}{x!(n-x)!} \, p^x q^{n-x} \\ &= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} \, p^{x-1} q^{n-x} = np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k q^{n-1-k} \\ &= np(p+q)^{n-1} = np \end{aligned}$$

Laskussa sovellettiin binomikaavaa (1.6). Laskemme binomijakauman odotusarvon myöhemmin helpommalla tavalla, jossa käytämme hyväksi odotusarvon lineaarisuutta. △

Esimerkki 4.4. Diskreetti jakauma, jolla ei ole odotusarvoa. Tunnetusti sarja $\sum_{x=1}^{\infty} 1/x^2$ suppenee, ja sarja $\sum_{x=1}^{\infty} 1/x$ hajaantuu. Olkoon $c = \sum_{x=1}^{\infty} 1/x^2$, ja määritellään sm X siten, että sen arvojoukko on kokonaisluvut poislukien nolla, ja ptnf on

$$f(x) = P(X = x) = \frac{1}{2c} \frac{1}{x^2}, \qquad x = \pm 1, \pm 2, \ldots$$

Tällöin X:llä ei ole odotusarvoa. △

Jos diskreetin sm:n X arvojoukko on $\{x_1, x_2, \dots\}$, niin X voidaan esittää erillisten tapahtumien $A_i = \{X = x_i\}$ indikaattorien avulla summana

$$X = \sum_{i \ge 1} x_i \, 1_{A_i}.$$

Mikäli X:llä on odotusarvo, niin se saadaan summasta

$$EX = \sum_{i \ge 1} x_i \, P(A_i).$$

Joskus on kätevää käyttää jotakin muuta perusjoukon Ω ositusta $B_1, B_2, \dots$ Jos X on vakio kussakin osituksen palassa B_i, niin seuraava lause totetaa, että edellinen kaava pitää yhä paikkansa.

Lause 4.1. *Jos $B_1, B_2, \dots$ on perusjoukon (äärellinen tai numeroituvasti ääretön) ositus, ja X on integroituva sm, joka saa vakioarvon y_j kullakin palalla B_j, eli*

$$X = \sum_{j \ge 1} y_j \, 1_{B_j},$$

niin

$$EX = \sum_{j \ge 1} y_j \, P(B_j).$$

Todistus. Selvästi X on diskreetti sm. Olkoot sen *eri* arvot $x_1, x_2, \dots,$ ja määritellään

$$A_i = \{X = x_i\}, \qquad i \ge 1.$$

Kun kukin A_i ositetaan tapahtumilla B_j, niin nähdään, että

$$\begin{aligned} EX &= \sum_{i \ge 1} x_i \, P(A_i) \\ &= \sum_{i \ge 1} x_i \sum_{j \ge 1} P(A_i \cap B_j) = \sum_{i \ge 1} \sum_{j \ge 1} x_i \, P(A_i \cap B_j) \end{aligned}$$

Mikäli $A_i \cap B_j \neq \emptyset$, niin edellä $x_i = y_j$, sillä X saa arvon x_i joukossa A_i ja arvon y_j joukossa B_j. Muussa tapauksessa summattava termi on nolla. Kaikissa tapauksissa

$$x_i \, P(A_i \cap B_j) = y_j \, P(A_i \cap B_j), \qquad \text{kaikilla } i, j \ge 1.$$

Näin ollen

$$\begin{aligned} EX &= \sum_{i \ge 1} \sum_{j \ge 1} y_j \, P(A_i \cap B_j) \\ &= \sum_{j \ge 1} y_j \sum_{i \ge 1} P(A_i \cap B_j) = \sum_{j \ge 1} y_j \, P(B_j), \end{aligned}$$

sillä tapahtumat $A_1, A_2, \dots$ osittavat kunkin tapahtuman B_j. □

Lause 4.2. *Olkoot X ja Y diskreettejä satunnaismuuttujia, joiden odotusarvot ovat olemassa. Odotusarvolla on seuraavat ominaisuudet.*

(a) (Positiivisuus) Jos $X \ge 0$, niin $EX \ge 0$.

(b) Jos $X \geq 0$ *ja* $EX = 0$*, niin* X *on vakio nolla, ts.* $P(X = 0) = 1$.

(c) Jos $X \leq Y$*, niin* $EX \leq EY$.

(d) (Vakion odotusarvo) Jos $a \in \mathbb{R}$*, niin* $Ea = a$.

(e) (Lineaarisuus) Jos $a, b \in \mathbb{R}$*, niin* $E(aX + bY) = a\,EX + b\,EY$.

(f) Jos $X \perp\!\!\!\perp Y$*, niin satunnaismuuttujalla* XY *on odotusarvo, ja se saadaan kaavalla*

$$E(XY) = E(X)\,E(Y).$$

(g) EX *riippuu vain* X*:n jakaumasta: jos* X*:llä ja* Y*:llä on sama jakauma, niin* $EX = EY$.

Todistus. Kohdat (a), (b), ja (g) ovat ilmeisiä. Vakion odotusarvo (kohta (d)) on jo laskettu.

Todistetaan lineaarisuus (kohta (e)) sekä riippumattomien satunnaismuuttujien tulon odotusarvoa koskeva väite (kohta (f)). Olkoon X:n eri arvot $x_1, x_2, \dots$ ja Y:n eri arvot $y_1, y_2, \dots$. Määritellään

$$A_i = \{X = x_i\}, \qquad B_j = \{Y = y_j\}.$$

Tällöin joukot $A_i \cap B_j, i, j \geq 1$ osittavat perusjoukon Ω, ja X:llä on arvo x_i ja Y:llä arvo y_j joukossa $A_i \cap B_j$, mikäli kyseinen joukko on epätyhjä. Muussa tapauksessa tämän joukon tn on nolla.

Jos $a, b \in \mathbb{R}$ ovat vakioita, niin lauseen 4.1 perusteella

$$\begin{aligned}
E(aX + bY) &= \sum_{i,j} (ax_i + by_j) P(A_i \cap B_j) \\
&= a \sum_{i,j} x_i\, P(A_i \cap B_j) + b \sum_{i,j} y_j\, P(A_i \cap B_j) \\
&= a \sum_i x_i\, P(A_i) + b \sum_j y_j P(B_j) = aEX + bEY.
\end{aligned}$$

Tämä todistaa lineaarisuuden. Jos $X \perp\!\!\!\perp Y$, ja X ja Y ovat integroituvia, niin

$$\begin{aligned}
E(XY) &= \sum_{i,j} x_i y_j\, P(A_i \cap B_j) \\
&= \sum_{i,j} x_i y_j\, P(A_i) P(B_j) = \sum_{i,j} x_i\, P(A_i)\, y_j\, P(B_j) \\
&= \left[\sum_i x_i\, P(A_i) \right] \left[\sum_j y_j\, P(B_j) \right] = E(X)E(Y)
\end{aligned}$$

Tämä todistaa kohdan (f).

Se seikka, että odotusarvo säilyttää järjestyksen (kohta (c)) seuraa lineaarisuudesta ja odotusarvon positiivisuudesta seuraavasti. Jos $X \leq Y$, niin $Y - X \geq 0$, joten

$$0 \leq E(Y - X) = EY - EX. \qquad \square$$

Huomautus. Jos X_1, X_2 ja X_3 ovat integroituvia sm:ia, niin

$$\begin{aligned}
E(X_1 + X_2 + X_3) &= E[(X_1 + X_2) + X_3] = E(X_1 + X_2) + EX_3 \\
&= EX_1 + EX_2 + EX_3.
\end{aligned}$$

Tämä ominaisuus yleistyy mille tahansa äärelliselle lukumäärälle integroituvia satunnaismuuttujia, ts.

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

mikäli kaikki odotusarvot EX_i ovat reaalilukuja.

Esimerkki 4.5. Binomijakauman odotusarvo uudestaan. Jos $X \sim \mathrm{Bin}(n,p)$, niin se voidaan esittää summana

$$X = \sum_{i=1}^{n} Y_i,$$

jossa $Y_i \sim \mathrm{Bernoulli}(p)$ ja Y_i:t ovat riippumattomia. Odotusarvon lineaarisuuden nojalla

$$EX = E\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} EY_i = np.$$

Tässä laskussa ei tarvittu satunnaismuuttujien Y_i riippumattomuutta. △

4.2 Jatkuvasti jakautuneen satunnaismuuttujan odotusarvo

Määritelmä 4.2 (Odotusarvo jatkuvassa tapauksessa)**.** Jos X on jatkuvasti jakautunut sm, ja sen tiheysfunktio on f, niin sen odotusarvo on luku

$$EX = E(X) = \int_{-\infty}^{\infty} x\, f(x)\,\mathrm{d}x,$$

mikäli kyseinen integraali suppenee itseisesti, eli mikäli $\int_{-\infty}^{\infty} |x|\, f(x)\,\mathrm{d}x < \infty$. Muussa tapauksessa sanotaan että X:llä ei ole odotusarvoa. Se seikka, että X:llä on odotusarvo voidaan ilmaista myös sanomalla, että X on *integroituva* sm.

Esimerkki 4.6. Tasajakauman odotusarvo. Jos $a < b$ ja $X \sim \mathrm{U}(a,b)$, niin helppo lasku osoittaa, että $EX = \frac{1}{2}(a+b)$. △

Esimerkki 4.7. Cauchyn jakauma esimerkkinä jatkuvasta jakaumasta, jolla ei ole odotusarvoa. Sm X noudattaa Cauchyn jakaumaa, jos sillä on tiheys

$$f(x) = \frac{1}{\pi}\,\frac{1}{1+x^2}, \qquad x \in \mathbb{R}.$$

Tämä on tf, sillä $f \geq 0$ ja f:n integraali reaaliakselin yli on yksi, sillä koska f on parillinen funktio, on

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \frac{2}{\pi}\int_{0}^{\infty} \frac{1}{1+x^2}\,\mathrm{d}x = \lim_{M\to\infty} \frac{2}{\pi} \Big/_{0}^{M} \arctan x = \frac{2}{\pi}\,\frac{\pi}{2} = 1.$$

X:llä ei ole odotusarvoa, sillä

$$\int_{-\infty}^{\infty} |x|\, f(x)\,\mathrm{d}x = \frac{1}{\pi}\int_{0}^{\infty} \frac{2x}{1+x^2}\,\mathrm{d}x = \frac{1}{\pi}\lim_{M\to\infty} \Big/_{0}^{M} \ln(1+x^2) = \infty.$$

△

4.3 Odotusarvon ominaisuuksia

Diskreettien ja jatkuvien jakaumien lisäksi on olemassa myös muunlaisia jakaumia. Yleisessä tapauksessa odotusarvo EX määritellään perusjoukon Ω yli laskettuna abstraktina Lebesguen integraalina

$$EX = \int X\,dP = \int_\Omega X(\omega)\,P(\mathrm{d}\omega)$$

mikäli integrandi on itseisesti integroituva. Vaikka emme varsinaisesti käytä tätä määritelmää, käsittelemme seuraavaksi kuitenkin hieman sitä, miten merkiltään rajoittamattoman satunnaismuuttujan odotusarvo määritellään todennäköisyysteoriassa.

Todennäköisyysteoriassa määritellään ensin ei-negatiivisen sm:n $Y \ge 0$ integraali EY tietyn rajankäynnin kautta. Tulokseksi saadaan joko jokin ei-negatiivinen reaaliluku (jolloin sanotaan, että Y on integroituva) tai ∞. Tämän jälkeen merkiltään rajoittamattoman sm:n X odotusarvo määritellään jakamalla se ensin *positiiviseen osaan* $X^+ \ge 0$ sekä *negatiiviseen osaan* $X^- \ge 0$ seuraavasti

$$X^+ = \max(X, 0), \quad X^- = \max(-X, 0). \tag{4.2}$$

Tällöin

$$X = X^+ - X^-, \quad \text{ja} \quad |X| = X^+ + X^-. \tag{4.3}$$

Mikäli sekä EX^+ ja EX^- ovat äärellisiä, niin EX määritellään kaavalla

$$EX = EX^+ - EX^-. \tag{4.4}$$

Tällöin sanotaan, että X on *integroituva*.

Joissakin tarkasteluissa (mutta harvoin) odotusarvolle sallitaan reaalilukujen lisäksi arvot $+\infty$ tai $-\infty$, ts. odotusarvo saa olla laajennettu reaaliluku. Jos vain toinen integraaleista EX^+ ja EX^- on äärellinen, niin EX voidaan määritellä käyttämällä kaavaa $EX = EX^+ - EX^-$ sekä laskusääntöjä

$$\infty - a = \infty, \quad a - \infty = -\infty, \qquad a \in \mathbb{R}\,.$$

Sen sijaan lausekkeelle $\infty - \infty$ ei voida määritellä arvoa. Mikäli X:n odotusarvo on määriteltävissä tähän tapaan, sanomme, että odotusarvo *on olemassa laajennettuna reaalilukuna* tai että se *on olemassa laajennetussa mielessä*.

Edellä nähtiin esimerkit sekä diskreetistä että jatkuvasti jakautuneesta satunnaismuuttujasta, joilla kummallakaan ei ole olemassa odotusarvoa edes tässä laajennetussa mielessä. Tämän kurssin puitteissa olemme pääasiassa kiinnostuneita vain äärellisistä odotusarvoista, ja sanomme muussa tapauksissa, että odotusarvo ei ole olemassa.

Yleisessä tapauksessa odotusarvolla on edelleen ne ominaisuudet, jotka todistimme diskreetissä tapauksessa. Seuraavaa lausetta emme todista, sillä se vaatisi mitta- ja integrointiteoriaa. Lauseen ominaisuuksia saadaan käyttää ikään kuin ne olisivat aksioomeja.

Lause 4.3. *Olkoot X ja Y integroituvia satunnaismuuttujia (siis: odotusarvot $EX, EY \in \mathbb{R}$). Odotusarvolla on seuraavat ominaisuudet.*

(a) (Positiivisuus) Jos $X \ge 0$, niin $EX \ge 0$.

(b) Jos $X \ge 0$ ja $EX = 0$, niin X on vakio nolla, ts. $P(X = 0) = 1$.

(c) Jos $X \le Y$, niin $EX \le EY$.

(d) (Vakion odotusarvo) Jos $a \in \mathbb{R}$, niin $Ea = a$.

(e) (Lineaarisuus) Jos $a, b \in \mathbb{R}$, niin $E(aX + bY) = a\,EX + b\,EY$.

(f) Jos $X \perp\!\!\!\perp Y$, niin satunnaismuuttujalla XY on odotusarvo, ja se saadaan kaavalla

$$E(XY) = E(X)\,E(Y).$$

(g) EX riippuu vain X:n jakaumasta: jos X:llä ja Y:llä on sama jakauma, niin $EX = EY$.

Lause 4.4. *X on integroituva eli sen odotusarvo on olemassa jos ja vain jos $|X|$ on integroituva. Tällöin*

$$|EX| \le E|X|$$

Todistus. Jos X on integroituva, niin $EX^+ \in \mathbb{R}$ ja $EX^- \in \mathbb{R}$, joten yhtälön (4.3) sekä lineaarisuuden nojalla

$$E|X| = EX^+ + EX^- \in \mathbb{R}\,.$$

Toisaalta, jos $E|X| < \infty$, niin välttämättä

$$0 \le EX^+ < \infty, \quad \text{ja} \quad 0 \le EX^- < \infty,$$

joten $EX = EX^+ - EX^- \in \mathbb{R}$. Jos X on integroituva, niin kolmioepäyhtälön nojalla

$$|EX| = |EX^+ - EX^-| \le EX^+ + EX^- = E|X|. \qquad \square$$

4.4 Muunnoksen odotusarvo

Jos X on sm, ja $g : \mathbb{R} \to \mathbb{R}$ on funktio, niin yhdistetty funktio $Y = g(X) = g \circ X$ on myös satunnaismuuttuja. Yksi tapa laskea Y:n odotusarvo olisi ensin johtaa sm:n Y jakauma. Tämä onnistuu periaatteessa helposti diskreetissä tapauksessa. Jatkuvassa tapauksessa johto onnistuu ainakin, jos g on riittävän säännöllinen. Mikäli Y:n jakauma on joko diskreetti tai jatkuva, voitaisiin sitten soveltaa tuttuja kaavoja. Toisen tavan tarjoaa seuraava lause.

Lause 4.5 (Satunnaismuuttujan muunnoksen odotusarvo)**.** *Olkoon X sm, $g : \mathbb{R} \to \mathbb{R}$ ja $Y = g(X)$.*

(a) Jos X:llä on diskreetti jakauma ptnf:lla f_X, niin

$$EY = Eg(X) = \sum_x g(x)\,f_X(x),$$

mikäli summa suppenee itseisesti, eli mikäli $\sum_x |g(x)| f_X(x) < \infty$

(b) Jos X:llä on jatkuva jakauma tf:lla f_X, niin

$$EY = Eg(X) = \int_{-\infty}^{\infty} g(x)\,f_X(x)\,\mathrm{d}x,$$

mikäli integraali suppenee itseisesti, eli mikäli $\int_{-\infty}^{\infty} |g(x)|\,f_X(x)\,\mathrm{d}x < \infty$.

Todistus. Esitetään todistus vain diskreetissä tapauksessa, jolloin myös Y on diskreetti sm. Olkoot X:n eri arvot $x_1, x_2, \ldots$, ja määritellään

$$A_i = \{X = x_i\}.$$

Tällöin $Y(\omega)$ saa vakioarvon $y_i = g(x_i)$, kun $\omega \in A_i$. Lauseen 4.1 perusteella

$$EY = \sum_{i \ge 1} y_i\,P(A_i) = \sum_{i \ge 1} g(x_i)\,f_X(x_i). \qquad \square$$

Yhteenvetona,

$$Eg(X) = \begin{cases} \sum_x g(x)\, f_X(x) & \text{jos } X\text{:llä on diskreetti jakauma} \\ \int_{-\infty}^{\infty} g(x)\, f_X(x)\,\mathrm{d}x & \text{jos } X\text{:llä on jatkuva jakauma} \end{cases} \tag{4.5}$$

mikäli kyseinen odotusarvo on hyvin määritelty. Tämä on erittäin tärkeä kaava käytännön laskuissa; myös odotusarvon EX laskeminen palautuu tähän kaavaan (valitsemalla funktioksi g identiteettifunktio, jolle $g(x) = x$). Todennäköisyyslaskennan englanninkielisissä oppikirjoissa kaavalle käytetään usein nimitystä *law of the unconscious statistician*, koska tilastotieteilijät toisinaan käyttävät sitä huomaamatta, että kyseessä ei ole määritelmä vaan että kyseinen kaava pitäisi jollakin tavalla johtaa.

Huomautus. Aloittelijat luulevat usein virheellisesti, että $Eg(X)$ on yhtäsuuri kuin $g(EX)$. Tämä pitää paikkansa affiinille funktiolle $g(x) = a + bx$, mutta ei yleisesti.

Esimerkki 4.8. Odotusarvo sensuroidulle jakaumalla. Esimerkissä 2.6 käsittelimme jakaumaa, joka saadaan muunnoksella $Y = g(X)$, kun $X \sim U(0, 4)$ ja $g(x) = \min(\max(1, x), 2)$. Osoittautui, että sm:n Y jakauma ei ole jatkuva eikä diskreetti vaan ns. sekatyypin jakauma, joten Y:n odotusarvon laskuun suoraan sen jakaumasta tarvittaisiin Lebesguen integraalia. Koska Y on muunnos jatkuvasti jakautuneesta satunnaismuuttujasta X, niin EY:n saa kuitenkin laskettua suoraviivaisesti lauseen 4.5 avulla, sillä

$$\begin{aligned} EY = Eg(X) &= \int_{-\infty}^{\infty} g(x)\, f_X(x)\,\mathrm{d}x \\ &= \int_0^1 \frac{1}{4}\,\mathrm{d}x + \int_1^2 \frac{x}{4}\,\mathrm{d}x + \int_2^4 \frac{2}{4}\,\mathrm{d}x = \frac{13}{8}. \end{aligned}$$

△

Lause 4.6. *(Riippumattomien sm:ien muunnoksien tulon odotusarvo.) Olkoot satunnaismuuttujat $X \perp\!\!\!\perp Y$, ja $g : \mathbb{R} \to \mathbb{R}$ ja $h : \mathbb{R} \to \mathbb{R}$ sellaisia funktioita, että $g(X)$ ja $h(Y)$ ovat integroituvia sm:ia. Tällöin*

$$E[g(X)\, h(Y)] = E[g(X)]\, E[h(Y)].$$

Todistus. Koska $g(X) \perp\!\!\!\perp h(Y)$ lauseen 3.7 nojalla, tämä lause seuraa suoraan lauseen 4.3 f-kohdasta. □

4.5 Momentit

Määritelmä 4.3. Satunnaismuuttujan X

- k:s origomomentti eli k:s momentti μ'_k on

$$\mu'_k = EX^k, \qquad \text{kun } k = 1, 2, \ldots$$

- k:s keskusmomentti (engl. *central moment*) μ_k on

$$\mu_k = E[(X - EX)^k], \qquad \text{kun } k = 1, 2, \ldots$$

- kertaluvun r absoluuttinen momentti (engl. *absolute moment*) on

$$E|X|^r, \qquad r > 0$$

mikäli kyseiset odotusarvot ovat olemassa. Muuten sanotaan, että kyseinen momentti ei ole olemassa.

Huomautuksia.

- Vaikka a^k on määritelty kaikille reaaliluvuille a ja kokonaisluvuille $k \geq 1$, niin a^r ei ole reaalinen, mikäli $a < 0$ ja r ei ole kokonaisluku. Tämän takia origo- ja keskusmomentit määritellään vain silloin, kun kertaluku on kokonaisluku.
- Odotusarvo EX on 1. momentti μ'_1, ja 1. keskusmomentti $\mu_1 = 0$.
- Keskusmomentti μ_k voidaan esittää origomomenttien $\mu'_j, j \leq k$ avulla kehittämällä lauseke $(X - EX)^k$ binomikaavalla. Kun merkitään $\mu = EX$, niin esim.

  $$\mu_2 = E[(X-\mu)^2] = E[X^2 - 2\mu X + \mu^2] = E[X^2] - \mu^2 = \mu'_2 - (\mu'_1)^2$$

  Tässä esiintyvä toinen keskusmomentti μ_2 on sovelluksissa niin tärkeä, että sille käytetään erityistä nimitystä: se on nimeltään *varianssi*.
- Sovelluksissa tärkeimpiä ovat ensimmäinen ja toinen momentti. Korkeammista momenteista esim. kolmas keskusmomentti kuvaa jakauman vinoutta ja neljäs keskusmomentti sen huipukkuutta. Tätä korkeampia momentteja käytetään harvoin.

Olkoon kertaluvun $s > 0$ absoluuttinen momentti äärellinen, eli $E|X|^s < \infty$. Jos $0 < r < s$, niin

$$\begin{aligned}|X|^r &= |X|^r \, 1_{\{|X| \leq 1\}} + |X|^r \, 1_{\{|X|>1\}} \\ &\leq 1 + |X|^s \, 1_{\{|X|>1\}} \leq 1 + |X|^s,\end{aligned}$$

minkä takia

$$E|X|^r \leq 1 + E|X|^s < \infty.$$

Siis jos $E|X|^s$ on äärellinen jollekin $s > 0$, niin kaikki sitä pienemmän kertaluvun absoluuttiset momentit

$$E|X|^r, \qquad 0 < r < s$$

ovat myös äärellisiä. Jos $s > 1$, niin tällöin kertalukujen $1 \leq k \leq s$ origomomentit sekä keskusmomentit kertalukua $1 \leq k \leq r$ ovat olemassa (binomikaavan ja odotusarvon lineaarisuuden nojalla).

4.6 Varianssi, keskihajonta ja kovarianssi

Määritelmä 4.4 (Varianssi ja keskihajonta)**.** Satunnaismuuttujan X *varianssi* $\operatorname{var} X$ on sen 2. keskusmomentti, eli

$$\operatorname{var} X = \operatorname{var}(X) = E[(X - EX)^2],$$

mikäli kyseinen odotusarvo on hyvin määritelty. Tällöin X:n *keskihajonta* (engl. *standard deviation*) on varianssin $\operatorname{var} X$ (positiivinen) neliöjuuri. Usein X:n varianssia merkitään σ^2 (tai σ_X^2). Tällöin varianssin positiivinen neliöjuuri $\sigma \geq 0$ (tai $\sigma_X \geq 0$) on X:n keskihajonta.

- Varianssi voidaan määritellä vain integroituvalle satunnaismuuttujalle, ts. oletetaan että $EX \in \mathbb{R}$. Jos EX on äärellinen, mutta $E[(X - EX)^2] = \infty$, niin voidaan sanoa, että $\operatorname{var} X = \infty$.

- Varianssille $\operatorname{var} X$ käytetään yleisesti myös merkintää $D^2(X)$.
- Edellä nähtiin, että mikäli EX^2 on äärellinen, niin myös varianssi on äärellinen, ja

$$\operatorname{var} X = EX^2 - (EX)^2. \tag{4.6}$$

Lause 4.7 (Varianssin ominaisuuksia)**.**

(i) Mikäli $\operatorname{var} X$ *on olemassa, niin* $\operatorname{var} X \geq 0$. *Jos* $\operatorname{var} X = 0$, *niin* X *on vakio.*

(ii) Jos $a, b \in \mathbb{R}$ *ja* $\operatorname{var} X$ *on olemassa, niin* $\operatorname{var}(aX + b) = a^2 \operatorname{var} X$.

Todistus. **(i):** Merkitään $\mu = EX$. Koska $(X - \mu)^2 \geq 0$, niin $\operatorname{var} X \geq 0$, mikäli se on reaalilukuna olemassa. Jos taas $\operatorname{var} X = \infty$, niin tietenkin $\operatorname{var} X \geq 0$.

Jos $\operatorname{var} X = 0$, niin tällöin

$$P[X = \mu] = P[(X - \mu)^2 = 0] = 1.$$

Todennäköisyys on yksi sen takia, että $(X - \mu)^2 \geq 0$ ja $E(X - \mu)^2 = 0$. Tästä syystä X saa arvon μ todennäköisyydellä yksi.

(ii): $E(aX + b) = aEX + b$, joten

$$\operatorname{var}(aX + b) = E[(a(X - EX))^2] = a^2 E[(X - EX)^2].$$

□

Toisinaan on tarpeen osata laskea kahden satunnaismuuttujan summan tai erotuksen varianssi. Tämä onnistuu niiden kovarianssin avulla.

Määritelmä 4.5. Satunnaismuuttujien X ja Y (välinen) *kovarianssi* $\operatorname{cov}(X, Y)$ on odotusarvo

$$\operatorname{cov}(X, Y) = E[(X - EX)(Y - EY)],$$

mikäli kyseinen odotusarvo on olemassa.

Huomautus: Cauchyn–Schwarzin epäyhtälön perusteella (joka tulee esille jakossa 6.4) kovarianssi on automaattisesti olemassa, jos satunnaismuuttujilla X ja Y on äärelliset toiset momentit (eli jos niillä on äärelliset varianssit).

Kovarianssin määritelmästä nähdään, että

- $\operatorname{cov}(Y, X) = \operatorname{cov}(X, Y)$,
- $\operatorname{var} X = \operatorname{cov}(X, X)$,

mikäli kyseiset suureet ovat äärellisiä. Kun kerrotaan sulut auki, ja lasketaan odotusarvo termi termiltä, saadaan

$$\begin{aligned}\operatorname{cov}(X, Y) &= E[(X - EX)(Y - EY)] \\ &= E[XY - X\,(EY) - (EX)\,Y + (EX)\,(EY)] \\ &= E(XY) + (-1 - 1 + 1)\,(EX)\,(EY),\end{aligned}$$

joten kovarianssi voidaan laskea myös kaavalla

$$\operatorname{cov}(X, Y) = E(XY) - (EX)(EY). \tag{4.7}$$

Lause 4.8. *Kovarianssi on bilineaarinen, eli molempien argumenttiensa suhteen lineaarinen operaattori. Ts. jos* $a_1, a_2 \in \mathbb{R}$ *ovat vakioita ja* X_1, X_2 *ja* Z *ovat satunnaismuuttujia, joilla on äärelliset toiset momentit, niin*

$$\begin{aligned} \operatorname{cov}(a_1X_1 + a_2X_2, Z) &= a_1 \operatorname{cov}(X_1, Z) + a_2 \operatorname{cov}(X_2, Z) \\ \operatorname{cov}(Z, a_1X_1 + a_2X_2) &= a_1 \operatorname{cov}(Z, X_1) + a_2 \operatorname{cov}(Z, X_2) \end{aligned}$$

Jos lisäksi $b_1, b_2 \in \mathbb{R}$ *ja* Y_1, Y_2 *ovat satunnaismuuttujia, joilla on äärelliset toiset momentit, niin*

$$\operatorname{cov}(a_1X_1 + a_2X_2, b_1Y_1 + b_2Y_2) = \sum_{i=1}^{2}\sum_{j=1}^{2} a_i\, b_j \operatorname{cov}(X_i, Y_j).$$

Todistus. Odotusarvon lineaarisuuden nojalla

$$E(a_1X_1 + a_2X_2) = a_1EX_1 + a_2EX_2,$$

joten

$$a_1X_1 + a_2X_2 - E(a_1X_1 + a_2X_2) = a_1(X_1 - EX_1) + a_2(X_2 - EX_2).$$

Siis odotusarvon lineaarisuuden nojalla,

$$\begin{aligned} \operatorname{cov}(a_1X_1 + a_2X_2, Z) &= E\left[(a_1(X_1 - EX_1) + a_2(X_2 - EX_2))\,(Z - EZ)\right] \\ &= a_1E[(X_1 - EX_1)(Z - EZ)] + a_2E[(X_2 - EX_2)(Z - EZ)] \\ &= a_1 \operatorname{cov}(X_1, Z) + a_2 \operatorname{cov}(X_2, Z). \end{aligned}$$

Kovarianssin symmetrisyyden nojalla

$$\operatorname{cov}(Z, \sum_i a_iX_i) = \operatorname{cov}(\sum_i a_iX_i, Z) = \sum_i a_i \operatorname{cov}(X_i, Z) = \sum_i a_i \operatorname{cov}(Z, X_i).$$

Kovarianssi on nyt todistettu molempien argumenttiensa suhteen lineaariseksi. Näin ollen

$$\begin{aligned} \operatorname{cov}(\sum_i a_iX_i, \sum_j b_jY_j) &= \sum_i a_i \operatorname{cov}(X_i, \sum_j b_jY_j) = \sum_i a_i \sum_j b_j \operatorname{cov}(X_i, Y_j) \\ &= \sum_i \sum_j a_ib_j \operatorname{cov}(X_i, Y_j). \end{aligned}$$

□

Laskemme seuraavaksi lineaarikombinaation $aX + bY$ varianssin, kun a ja b ovat vakioita. Tällöin

$$\begin{aligned} \operatorname{var}(aX + bY) &= \operatorname{cov}(aX + bY, aX + bY) \\ &= a \operatorname{cov}(X, aX + bY) + b \operatorname{cov}(Y, aX + bY) \\ &= a^2 \operatorname{cov}(X, X) + ab \operatorname{cov}(X, Y) + ab \operatorname{cov}(Y, X) + b^2 \operatorname{cov}(Y, Y) \\ &= a^2 \operatorname{var}(X) + 2ab \operatorname{cov}(X, Y) + b^2 \operatorname{var}(Y). \end{aligned}$$

Huomaa erityisesti, että

$$\operatorname{var}(X + Y) = \operatorname{var} X + \operatorname{var} Y + 2 \operatorname{cov}(X, Y) \tag{4.8}$$

ja

$$\operatorname{var}(X - Y) = \operatorname{var} X + \operatorname{var} Y - 2 \operatorname{cov}(X, Y). \tag{4.9}$$

Jos $\operatorname{cov}(X, Y) = 0$ eli muuttujat X ja Y eivät korreloi, niin

$$\operatorname{var}(X + Y) = \operatorname{var}(X - Y) = \operatorname{var} X + \operatorname{var} Y. \tag{4.10}$$

Jos $X \perp\!\!\!\perp Y$, niin

$$\operatorname{cov}(X, Y) = E[(X - EX)(Y - EY)] = E(X - EX)\, E(Y - EY) = 0 \cdot 0 = 0,$$

joten tällöin $\operatorname{var}(X \pm Y) = \operatorname{var} X + \operatorname{var} Y$.

Huomaa, että summan $X+Y$ varianssi on yhtä kuin muuttujien varianssien summa täsmälleen silloin, kun $\operatorname{cov}(X, Y) = 0$. Tässä tapauksessa sanotaan, että X ja Y eivät korreloi. Edellä nähtiin, että riippumattomuudesta seuraa korreloimattomuus.

Esimerkki 4.9. Korreloimattomuudesta ei seuraa riippumattomuus. Olkoon sm:lla X jatkuva jakauma tf:lla f, ja olkoon f parillinen funktio. Oletetaan lisäksi, että EX^4 on äärellinen. Tällöin $EX = E(X^3) = 0$, joten kaavan (4.7) avulla nähdään helposti, että

$$\operatorname{cov}(X, X^2) = E(X \cdot X^2) - (EX)(EX^2) = 0.$$

Siis X ja X^2 eivät korreloi, mutta ne eivät tietenkään ole riippumattomia satunnaismuuttujia. $\triangle$

Korreloimattomien satunnaismuuttujien summan varianssin kaava (4.10) yleistyy helposti myös äärellisen monen keskenään korreloimattoman satunnaismuuttujan summan varianssille (HT). Usein tarvitaan tämän tuloksen seuraavaa erikoistapausta. Jos $X_1, \dots, X_n$ ovat riippumattomia ja niillä on kaikilla äärellinen varianssi, niin tällöin

$$\operatorname{var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \operatorname{var}(X_i)$$

Esimerkki 4.10. Binomijakauman varianssi. Jos $X \sim \mathrm{Bin}(n, p)$, niin se voidaan esittää summana $\sum_{i=1}^n Y_i$, jossa $Y_i \sim \mathrm{Bernoulli}(p)$ ja Y_i:t ovat riippumattomia. Nyt $EY_i = p$, joten

$$\operatorname{var} Y_i = E[(Y_i - p)^2] = (0 - p)^2(1 - p) + (1 - p)^2 p = p(1 - p).$$

Riippumattomuuden nojalla $\operatorname{var} X = \sum_{i=1}^n \operatorname{var} Y_i = np(1 - p)$. $\triangle$

4.7 Momenttiemäfunktio ja kumulanttiemäfunktio

Määritelmä 4.6 (Momenttiemäfunktio)**.** Satunnaismuuttujan X (jakauman) momenttiemäfunktio eli momenttigeneroivafunktio eli momentit generoiva funktio (engl. *moment generating function, mgf*) on funktio

$$M(t) = M_X(t) = E e^{tX},$$

joka on (reaalifunktiona) olemassa niissä pisteissä t, joissa kyseinen odotusarvo on (äärellisenä) olemassa. Erityisesti $M(0) = 1$.

Huomautus. Momenttiemäfunktio on (argumentin merkkiä vaille) sama asia kuin Laplacen muunnos (engl. *Laplace transform*).

Jos seuraavassa laskussa derivoinnin ja odotusarvon järjestyksen vaihto pystytään jotenkin perustelemaan, niin helposti nähdään, että

$$M'_X(t) = \frac{\mathrm{d}}{\mathrm{d}t} M_X(t) = \frac{\mathrm{d}}{\mathrm{d}t} E\mathrm{e}^{tX} = E\frac{\partial}{\partial t}\mathrm{e}^{tX} = EX\mathrm{e}^{tX},$$

joten $EX = M'_X(0)$. Laskemalla toinen derivaatta nähdään, että

$$M''_X(t) = \frac{\mathrm{d}}{\mathrm{d}t} EX\mathrm{e}^{tX} = E\frac{\partial}{\partial t}X\mathrm{e}^{tX} = EX^2\mathrm{e}^{tX},$$

joten $EX^2 = M''_X(0)$. Samalla tavalla edeten nähdään, että aina $EX^k = M^{(k)}(0)$. Osoittautuu, että laskun läpiviennille riittävä ehto on se, että momenttiemäfunktio on määritelty jossakin origon ympäristössä.

Lause 4.9. *Jos momenttiemäfunktio $M_X(t)$ on reaalisena olemassa jossakin origon ympäristössä $|t| < h$, jossa $h > 0$, niin $M_X(t)$ on äärettömän monta kertaa derivoituva origossa, X:n kaikki momentit ovat olemassa, ja M_X voidaan esittää suppenevana Taylorin sarjana*

$$M_X(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} EX^k, \qquad |t| < h.$$

Erityisesti $\quad EX^k = M_X^{(k)}(0), \qquad k \geq 1.$

Todistus. Sarjaesityksen todistus löytyy esim. Billingsleyn teoksesta *Probability and Measure*. Momenttien esitys derivaattojen avulla seuraa derivoimalla Taylorin sarjaa. □

Määritelmä 4.7 (Kumulanttiemäfunktio)**.** Jos momenttiemäfunktio $M_X(t)$ on määritelty jossakin origon ympäristössä, niin X:n (jakauman) kumulanttiemäfunktio $K(t) = K_X(t)$ (engl. *cumulant generating function, cgf*) määritellään kyseisessä ympäristössä momenttiemäfunktion logaritmina,

$$K(t) = K_X(t) = \ln M_X(t).$$

Odotusarvo ja varianssi saadaan laskettua derivoimalla kumulanttiemäfunktiota origossa, sillä

$$K'_X(0) = \frac{M'_X(0)}{M_X(0)} = \frac{EX}{1} = EX,$$

ja

$$K''_X(0) = \frac{M''_X(0)M_X(0) - M'(0)M'(0)}{M_X(0)^2} = EX^2 - (EX)^2 = \operatorname{var} X.$$

Tämä tarjoaa näppärän tavan johtaa odotusarvo ja varianssi joillekin jakaumille, mutta toisille jakaumille sattaa olla kätevämpää johtaa odotusarvo ja varianssi laskemalla jakauman ensimmäinen ja toinen momentti.

Origossa laskettuja kumulanttiemäfunktion derivaattoja kutsutaan X:n (jakauman) *kumulanteiksi*: k:s kumulantti on $K_X^{(k)}(0)$. Ensimmäinen kumulantti on odotusarvo ja toinen kumulantti on varianssi.

Esimerkki 4.11. Binomijakauman odotusarvo ja varianssi kumulanttiemäfunktion avulla. Jos $X \sim \mathrm{Bin}(n, p)$, niin binomikaavan avulla nähdään helposti, että

$$M_X(t) = (p\mathrm{e}^t + 1 - p)^n,$$

joten
$$K_X(t) = n\ln(pe^t + 1 - p).$$
Laskemalla ensimmäinen ja toinen derivaatta origossa nähdään, että
$$EX = K'_X(0) = np, \qquad \operatorname{var} X = K''_X(0) = np(1-p).$$
△

Lause 4.10 (Momenttiemäfunktion ominaisuuksia)**.** *Seuraavassa kohtien (a) ja (b) kaavat pätevät niissä pisteissä, joissa kyseiset funktiot ovat määriteltyjä. (Ne pätevät muissa pisteissä muodossa $\infty = \infty$.)*

(a) Jos a, b ovat vakioita, niin $M_{aX+b}(t) = e^{bt} M_X(at)$.

(b) Jos $X \perp\!\!\!\perp Y$, niin $M_{X+Y}(t) = M_X(t)\, M_Y(t)$.

Todistus. Nämä ominaisuudet todistetaan yksinkertaisilla laskuilla, sillä
$$M_{aX+b}(t) = E\exp(t(aX+b)) = e^{bt}\, E\exp(atX) = e^{bt}\, M_X(at),$$
ja kun $X \perp\!\!\!\perp Y$, niin
$$M_{X+Y}(t) = E\exp(t(X+Y)) = E[\exp(tX)\, \exp(tY)] = Ee^{tX}\, Ee^{tY},$$
missä käytimme hyväksi lausetta 4.6. □

Lause 4.11 (Momenttiemäfunktio määrää jakauman)**.** *Jos X:n ja Y:n momenttiemäfunktiot ovat molemmat olemassa jossakin origon ympäristössä, ja ovat siellä yhtäsuuret, niin X:llä ja Y:llä on sama jakauma.*

Todistus. Ks. esim. Billingsleyn teos. □

Momenttiemäfunktiota ja kumulanttiemäfunktiota käytetään useissa asymptoottisissa tarkasteluissa (esim. suurten poikkeamien teoriassa sekä satulapistekehitelmässä). Näiden funktioiden käyttöön liittyy kuitenkin hankaluuksia. Jos jokin X:n momentti ei ole olemassa, niin momenttiemäfunktio on kyllä määritelty origossa, mutta ei missään origon ympäristössä. Lisäksi on olemassa jakaumia (esim. log-normaalinen jakauma), joilla on kaikkien kertalukujen momentit, mutta joiden momenttiemäfunktio ei ole olemassa missään origon ympäristössä. Näiden ongelmien takia edistyneemmissä todennäköisyyslaskennan kirjoissa käytetään momenttiemäfunktion sijasta jakauman karakteristista funktiota.

4.8 Karakteristinen funktio

Edellisen jakson lopussa mainituilta vaikeuksilta vältytään, mikäli momenttiemäfunktion sijasta käytetään jakauman karakteristista funktiota, sillä karakteristinen funktio on kaikkialla määritelty. Karakteristisen funktion määrittelyssä tarvitaan kompleksilukuja.

Määritelmä 4.8. Satunnaismuuttujan X karakteristinen funktio (engl. *characteristic function*) on koko reaaliakselilla määritelty kompleksiarvoinen funktio
$$\phi(t) = \phi_X(t) = Ee^{itX}, \quad \text{jossa } i = \sqrt{-1}.$$

Tässä kompleksinen eksponentti saadaan laskettua kaavalla

$$\mathrm{e}^{itX} = \cos(tX) + i \sin(tX),$$

ja sen odotusarvo määritellään laskemalla yhteen reaaliosan ja imaginääriosan odotusarvot,

$$E\mathrm{e}^{itX} = E \cos(tX) + i\, E \sin(tX).$$

Karakteristinen funktio $\phi_X(t)$ on määritelty kaikilla X ja kaikilla t sen takia, että se on rajoitetun funktion odotusarvo.

Lause 4.12. *Jos momenttiemäfunktio $M_X(t)$ on määritelty jossakin origon ympäristössä, niin X:n karakteristinen funktio on*

$$\phi_X(t) = M_X(it).$$

Huomautus. Tässä $M_X(it)$ tarkoittaa sitä, että Taylorin sarjaan sijoitetaan argumentin t sijalle it. Tämä on luvallista, koska kyseessä on suppeneva potenssisarja. Jos funktiolle $M_X(t)$ saadaan johdettua lauseke, niin tulos saadaan laskettua myös sijoittamalla ko. lausekkeessa argumentin t sijalle it.

Lause 4.13 (Karakteristinen funktio määrää jakauman)**.** *Jos X:n ja Y:n karakteristiset funktiot ovat samat, niin X:llä ja Y:llä on sama jakauma.*

Todistus. Ks. esim. Billingsleyn teos. □

Luku 5

Tärkeitä yksiulotteisia jakaumia

Tässä luvussa tarkastellaan tiettyjä sovelluksissa usein esiintyviä jakaumia. Näistä jakaumista löytyy lisätietoja ja kuvaajia Wikipediasta, ks.

`http://en.wikipedia.org/wiki/List_of_probability_distributions`

Varoitus. Kirjallisuudessa käytetään useille näistä jakaumista monia erilaisia parametrointeja. Kussakin lähteessä käytetty parametrointi paljastuu tavallisesti vasta tutkimalla lähemmin siinä käytettyjä kaavoja. Eri lähteet käyttävät jakaumille eri tunnuksia.

5.1 Diskreettejä jakaumia

Suurin osa sovelluksissa käytettävistä diskreeteistä satunnaismuuttujista on kokonaislukuarvoisia. Annamme ptnf:n f lausekkeen vain niissä pisteissä, joissa $f(x)$ voi olla nollaa suurempi.

5.1.1 Binomijakauma

Synty. Onnistumisten lkm n-kertaisessa toistetussa Bernoullin kokeessa, kun onnistumistn on p.

Tunnus $X \sim \mathrm{Bin}(n, p)$, jossa $n \geq 1$ kokonaisluku (otoskokoparametri) ja $0 \leq p \leq 1$ (onnistumistn; todennäköisyysparametri).

Ptnf

$$f(x) = f(x \mid n, p) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0, 1, \ldots, n.$$

Binomikaavan mukaan

$$1 = (p + (1-p))^n = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x}.$$

Koska lisäksi jokainen termi on ei-negatiivinen, muodostuu niistä ptnf kokonaisluvuilla $x = 0, 1, \ldots, n$.

Odotusarvo, varianssi, momenttiemäfunktio

$$EX = np, \quad \operatorname{var} X = np(1-p), \quad M(t) = (pe^t + 1 - p)^n.$$

Yhteyksiä muihin jakaumiin. $\mathrm{Bin}(1, p)$ on $\mathrm{Bernoulli}(p)$.

Yhteenlaskuominaisuus. Jos $X \sim \mathrm{Bin}(m,p)$ ja $Y \sim \mathrm{Bin}(n,p)$ (sama onnistumistn) ja $X \perp\!\!\!\perp Y$, niin

$$X + Y \sim \mathrm{Bin}(n+m,p).$$

Esimerkiksi, jos laatikossa on N palloa, joista $0 \le K \le N$ on valkoista ja loput mustia, ja laatikosta poimitaan umpimähkään n palloa takaisinpanolla, niin nostettujen valkoisten pallojen lukumäärällä X on jakauma $\mathrm{Bin}(n, K/N)$, jonka odotusarvo ja varianssi ovat

$$EX = n\,\frac{K}{N}, \qquad \operatorname{var} X = n\,\frac{K}{N}\,(1-\frac{K}{N}).$$

5.1.2 Hypergeometrinen jakauma

Synty. Laatikossa on N palloa, joista $0 \le K \le N$ on valkoista ja loput ovat mustia, ja laatikosta nostetaan umpimähkäisesti $n \le N$ palloa ilman takaisinpanoa. Jos X on sm, joka kertoo, kuinka moni nostetuista palloista on valkoinen, niin X:n **ptnf** on

$$f(x) = f(x \mid N, K, n) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}, \qquad x = 0, 1, \ldots, n.$$

Tämä on hypergeometrinen jakauma parametreilla N, K ja n. Edellä käytettiin binomikertoimille $\binom{m}{k}$ sopimusta (1.5), jonka mukaan $\binom{m}{k} = 0$ ellei $0 \le k \le m$. Hypergeometrisen jakauman pistetodennäköisyysfunktio poikkeaa nollasta vain, kun $0 \le x \le \min(n, K)$ ja $n - x \le N - K$.

Hypergeometrisen jakauman ptnf:n saa johdettua suoraan kombinatoriikan tuloperiaatteella. Nyt symmetrisiä alkeistapauksia ovat N pallon n-osajoukot, joiden lukumäärä on $\binom{N}{n}$. Sellaisen osajoukon valinta, jossa on x valkoista ja $n-x$ mustaa palloa, voidaan ajatella tehtävän kahdessa vaiheessa. Ensin valitaan x:n valkoisen pallon osajoukko K valkoisesta pallosta (eri mahdollisuuksia on $\binom{K}{x}$), ja sen jälkeen valitaan $n - x$ mustan pallon osajoukko $N - K$ mustasta pallosta (eri mahdollisuuksia on $\binom{N-K}{n-x}$). Jakauman ptnf saadaan jakamalla suotuisten alkeistapauksien lukumäärä kaikkien alkeistapausten lukumäärällä.

Hypergeometrisen jakauman odotusarvo ja varianssi ovat

$$EX = n\,\frac{K}{N}, \qquad \operatorname{var} X = n\,\frac{K}{N}\,(1-\frac{K}{N})\,\frac{N-n}{N-1}.$$

Nämä tulokset johdetaan esim. Tuomisen kirjassa. Hypergeometrisen jakauman käsittely on hankalahkoa, minkä takia suurten populaatioiden tapauksessa (N suuri ja $n \ll N$) sitä mielellään approksimoidaan vastaavalla binomijakaumalla $\mathrm{Bin}(n, K/N)$, joka syntyisi otannasta takaisinpanolla.

Binomijakauma ja hypergeometrinen jakauma ovat lähisukulaisia. Sen sijaan hypergeometrisella jakaumalla ja seuraavaksi esitettävällä geometrisella jakaumalla ei ole keskenään mitään ilmeistä yhteyttä. Hypergeometrisen jakauman nimen alkuperä on hämärän peitossa, mutta se saattaa johtua jakauman tietyistä löyhistä yhteyksistä ns. hypergeometriseen funktioon, joka on eräs klassinen erikoisfunktio.

5.1.3 Geometrinen jakauma

Synty. Toistetaan riippumattomasti Bernoullin koetta, jossa onnistumistn on $0 < p < 1$. Olkoon

$$X = \text{“epäonnistumisten lkm ennen ensimmäistä onnistumista”}.$$

Tällöin X:llä on geometrinen jakauma parametrilla p, eli $X \sim \mathrm{Geom}(p)$.

Ptnf on

$$f(x) = f(x \mid p) = P(X = x) = p\,(1-p)^x, \quad x = 0, 1, 2, \ldots$$

Tämä johtuu siitä, että $X = x$ jos ja vain jos koetta vastaavat riippumattomat Bernoullin muuttujat $Y_1, \ldots, Y_x, Y_{x+1}$ saavat arvot

$$Y_1 = 0, Y_2 = 0, \ldots, Y_x = 0 \quad \text{ja} \quad Y_{x+1} = 1.$$

Riippumattomuuden nojalla tämän tapahtuman tn on $p\,(1-p)^x$.

Se että kyseessä on todella ptnf nähdään *geometrisen sarjan* summasta,

$$\sum_{j=0}^{\infty} r^j = 1 + r + r^2 + \cdots = \frac{1}{1-r}, \qquad |r| < 1,$$

jonka seurauksena

$$\sum_{x=0}^{\infty} f(x) = p\,\frac{1}{1-(1-p)} = 1.$$

Koska lisäksi jokainen termeistä $f(x) = p\,(1-p)^x$ on positiivinen, muodostuu niistä ptnf einegatiivisilla kokonaisluvuilla.

Momenttiemäfunktio saadaan helpolla laskulla,

$$M(t) = \frac{p}{1-(1-p)\mathrm{e}^t}, \qquad t < -\ln(1-p).$$

Tästä johdetaan helposti jakauman odotusarvo ja varianssi,

$$EX = \frac{1-p}{p}, \qquad \operatorname{var} X = \frac{1-p}{p^2}.$$

Huomautus. Useissa lähteissä geometrinen jakauma määritellään edellisesti poikkeavasti niin, että se on satunnaismuuttujan Y jakauma, jossa

$$Y = \text{"sen toiston järjestysnumero, jolla onnistutaan ensimmäisen kerran"}.$$

Tässä $Y = X + 1$.

5.1.4 Negatiivinen binomijakauma

Synty. Toistetaan riippumattomasti Bernoullin koetta, jossa onnistumistn on $0 < p < 1$. Olkoon $r > 0$ kokonaisluku, ja tarkastellaan sm:a

$$X = \text{"epäonnistumisten lkm, ennenkuin onnistutaan } r\text{:nnen kerran"}.$$

Tällöin

$$\begin{aligned}\{X = x\} = &\{r + x - 1 \text{ ensimmäisessä toistossa epäonnistutaan } x \text{ kertaa} \\ &\text{ja onnistutaan } r-1 \text{ kertaa}\} \cap \{\text{toistossa } r + x \text{ onnistutaan}\}.\end{aligned}$$

Koska toistot ovat riippumattomia, on

$$f(x) = f(x \mid r, p) = \binom{r+x-1}{x} p^r (1-p)^x, \qquad x = 0, 1, 2, \ldots$$

Tässä binomikertoimet ilmaistaan usein gammafunktion avulla, joka yleistää kertoman positiivisille reaaliluvuille kaavan $r! = \Gamma(r+1)$ kautta. Tällöin parametrille $r > 0$ voidaan sallia myös muut kuin kokonaislukuarvot, ja ptnf voidaan kirjoittaa muotoon

$$f(x) = f(x \mid r, p) = \frac{\Gamma(r+x)}{\Gamma(r)\, x!}\, p^r (1-p)^x, \qquad x = 0, 1, 2, \ldots$$

Jos $r > 0$ ei ole kokonaisluku, jakaumalla ei kuitenkaan enää ole toistokokeeseen liittyvää tulkintaa

Huomautus: useissa lähteissä (esim. Tuomisen TN I) negatiivinen binomijakauma määritellään siten, että se on sm:n Y jakauma, jossa

$$Y = \text{"sen toiston järjestysnumero, jolla onnistutaan } r\text{:nnen kerran"}.$$

Tällöin $Y = X + r$.

Huomaa, että geometrinen jakauma ja negatiivinen binomijakauma ovat lähisukulaisia. Sen sijaan binomijakaumalla ja negatiivisella binomijakaumalla ei ole mitään ilmeistä yhteyttä keskenään.

Lisätietoja negatiivisesta binomijakaumasta

Mikä negatiivisessa binomijakaumassa sitten on negatiivista? Tätä varten meidän pitää ensin tehdä alustavia tarkasteluja. Binomikertoimet määritellään seuraavalla tavalla, kun ylempi indeksi ei ole välttämättä kokonaisluku, mutta alempi indeksi on,

$$\binom{r}{k} = \begin{cases} \dfrac{r(r-1)\cdots(r-k+1)}{k!}, & k \ge 0 \text{ kokonaisluku,} \\ 0, & k < 0 \text{ kokonaisluku.} \end{cases} \tag{5.1}$$

Tätä merkintää tarvitaan esim. *binomisarjassa*, jonka mukaan kaikilla $r \in \mathbb{R}$

$$(1+z)^r = \sum_{j=0}^{\infty} \binom{r}{j} z^j, \qquad |z| < 1. \tag{5.2}$$

Binomikaava on binomisarjan erikoistapaus: jos $n \ge 0$ on kokonaisluku ja kokonaisluku $j \ge n$, niin $\binom{n}{j} = 0$, joten

$$(1+z)^n = \sum_{j=0}^{\infty} \binom{n}{j} z^j = \sum_{j=0}^{n} \binom{n}{j} z^j.$$

Olkoon $r > 0$ ja olkoon $j \ge 0$ kokonaisluku. Tällöin

$$\begin{aligned} \binom{-r}{j} &= \frac{(-r)(-r-1)\cdots(-r-j+1)}{j!} \\ &= (-1)^j \frac{(r+j-1)\cdots(r+1)r}{j!} = (-1)^j \binom{r+j-1}{j}. \end{aligned}$$

NegBin(r, p)-jakauman ptnf voidaan siis kirjoittaa vielä uuteen, binomijakaumaa muistuttavaan muotoon, kun merkitään $q = 1 - p$,

$$f(x) = \binom{-r}{x} p^r (-q)^x, \qquad x = 0, 1, 2, \ldots$$

Binomisarjan avulla on nyt helppo näyttää, että kyseessä on todellakin ptnf, sillä

$$\sum_{x=0}^{\infty} f(x) = p^r \sum_{x=0}^{\infty} \binom{-r}{x} (-q)^x = p^r (1-q)^{-r} = 1.$$

Ts. negatiivisen binomijakauman ptnf saadaan kehittämällä lausekkeessa $p^r(1-q)^{-r}$ binomin $1-q$ negatiivinen potenssi $(1-q)^{-r}$ binomisarjalla.

Jakauman **momenttiemäfunktio** saadaan laskettua helposti binomisarjan avulla,

$$M(t) = \left(\frac{p}{1-(1-p)\mathrm{e}^t} \right)^r, \qquad t < -\ln(1-p).$$

Odotusarvo ja varianssi löytyvät helposti derivoimalla momenttiemäfunktiota,

$$EX = \frac{r(1-p)}{p}, \qquad \operatorname{var} X = \frac{r(1-p)}{p^2}.$$

Yhteyksiä muihin jakaumiin. NegBin$(1, p)$ on Geom(p). Negatiiviselle binomijakaumalla käytetään myös muita parametrointeja. Usein parametreiksi otetaan jakauman odotusarvo μ sekä r (dispersioparametri). Tällöin $p = r/(\mu + r)$ ja jakauman varianssi on $\mu + \mu^2/r$. Jos tässä parametroinnissa annetaan $r \to \infty$, niin rajalla saadaan Poi(μ) (ts. pistetodennäköisyydet suppenevat kohti tämän Poissonin jakauman pistetodennäköisyyksiä). Tämän takia negatiivista binomijakaumaa käytetään mallina sellaisissa yhtyksissä, joissa tavallisesti tahdottaisiin käyttää Poissonin jakaumaa, mutta joissa aineiston empiirinen varianssi näyttää selvästi suuremmalta kuin Poissonin jakauman varianssi (*negative binomial as an overdispersed Poisson distribution*).

Huomautus. Kun $r > 0$ on kokonaisluku, niin NegBin(r, p)-jakaumaa kutsutaan toisinaan Pascalin jakaumaksi. Jos $r > 0$ on kokonaisluku, niin joissakin lähteissä sm:n $Y = X + r$ (eikä sm:n X) jakaumaa kutsutaan negatiiviseksi binomijakaumaksi.

5.1.5 Poissonin jakauma

Tunnus $X \sim \mathrm{Poi}(\theta)$, jossa $\theta > 0$.

Ptnf on

$$f(x) = f(x \mid \theta) = \mathrm{e}^{-\theta} \frac{\theta^x}{x!}, \quad x = 0, 1, 2, \ldots$$

Tämä on ptnf sillä perusteella, että kaikki arvot ovat ei-negatiivisia, ja *eksponenttifunktion sarjakehitelmän* mukaan

$$\mathrm{e}^u = \exp(u) = \sum_{j=0}^{\infty} \frac{u^j}{j!}, \qquad \forall u \in \mathbb{R}.$$

Mikäli $u > 0$, niin sarjan kaikki termit ovat positiivisia, joten luvut $\mathrm{e}^{-\theta} \theta^j / j!$ muodostavat pistetodennäköisyysfunktion, kun $\theta > 0$ ja $j = 0, 1, 2, \ldots$.

Momenttiemäfunktio saadaan laskettua eksponenttifunktion sarjakehitelmän avulla,

$$M(t) = \sum_{x=0}^{\infty} \mathrm{e}^{xt} \, \mathrm{e}^{-\theta} \frac{\theta^x}{x!} = \exp(\theta(\mathrm{e}^t - 1)).$$

Odotusarvo ja varianssi saadaan tästä helposti derivoimalla,

$$EX = \operatorname{var} X = \theta.$$

Yhteenlaskuominaisuus. Jos $X \sim \mathrm{Poi}(\theta_1)$ ja $Y \sim \mathrm{Poi}(\theta_2)$, ja $X \perp\!\!\!\perp Y$, niin

$$M_{X+Y}(t) = M_X(t)\, M_Y(t) = \exp((\theta_1 + \theta_2)(\mathrm{e}^t - 1)),$$

joten $X + Y \sim \mathrm{Poi}(\theta_1 + \theta_2)$.

Poissonin prosessi. Yksi tilanne, josta syntyy Poissonin jakauma on Poissonin prosessi. Siinä mallinnetaan diskreettejä tapahtumia jatkuvalla aika- tai pituusvälillä tai tasossa tai avaruudessa. Saateetaisiin mallintaa esimerkiksi sitä,

- kuinka monta asiakasta palvelupisteeseen saapuu tietyllä aikavälillä,
- kuinka monta fotonia osuu tiettyyn filmin alueeseen,
- kuinka monta valkosolua löytyy tietystä veripisarasta.

Mikäli välin pituus (tai tasoalueen pinta-ala, avaruuden osan tilavuus) on s yksikköä, niin Poissonin prosessissa siinä havaittavien tapahtumien lukumäärällä on Poissonin jakauma $\mathrm{Poi}(s\lambda)$. Lisäksi erillisillä väleillä (tasoalueilla, avaruuden osilla) havaittavat lukumäärät ovat keskenään riippumattomia. Keskimääräinen tapahtumien lukumäärä yhtä (pituus-, pinta-ala- tai tilavuus-)yksikköä kohti on $\lambda s/s = \lambda$. Ts. $\lambda > 0$ eli Poissonin prosessin *intensiteetti* on tapahtumien odotusarvo yhtä (pituus-, pinta-ala-, tilavuus-)yksikköä kohti.

5.2 Gamma- ja beetafunktio

Eulerin *gammafunktio* Γ voidaan määritellä positiivisilla argumenteilla integraalilla

$$\Gamma(t) = \int_0^\infty x^{t-1}\mathrm{e}^{-x}\,\mathrm{d}x, \qquad t > 0. \tag{5.3}$$

Jos $t \geq 1$, niin tämä integraali on epäoleellinen vain ylärjallaan, mutta jos $0 < t < 1$, niin integraali on epäoleellinen myös alarajalla, koska tällöin integrandi kasvaa rajatta, kun t lähestyy nollaa. Gammafunktion määrittelevän integraalin (5.3) voidaan osoittaa olevan äärellinen jos ja vain jos $t > 0$.

Selvästi $\Gamma(1) = 1$, sillä

$$\Gamma(1) = \int_0^\infty \mathrm{e}^{-x}\,\mathrm{d}x = \lim_{M\to\infty} \Big/_0^M -\mathrm{e}^{-x} = 1.$$

Osittaisintegroinnilla nähdään, että

$$\Gamma(t+1) = \int_0^\infty x^t\,\mathrm{e}^{-x}\,\mathrm{d}x = -\Big/_0^\infty x^t\,\mathrm{e}^{-x} + t\int_0^\infty x^{t-1}\,\mathrm{e}^{-x}\,\mathrm{d}x,$$

ja tässä sijoitustermi häviää. (Yllä integraalia käsiteltiin ikään kuin se ei olisi epäoleellinen, vaikka tässä oikeastaan pitäisi tarkastella raja-arvoja; tarkempi käsittely antaa kuitenkin saman johtopäätöksen.) Tämän takia gammafunktiolle pätee funktionaaliyhtälö

$$\Gamma(t+1) = t\,\Gamma(t), \qquad \text{kaikilla } t > 0. \tag{5.4}$$

Näiden ominaisuuksien takia kokonaislukuargumenteilla n pätee

$$\Gamma(n) = (n-1)!, \qquad \text{kun } n = 1, 2, 3, \ldots \tag{5.5}$$

Tässä mielessä gammafunktio on kertoman yleistys.

Tarkastellaan seuraavaksi integraalia

$$I = \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}x^2}\,\mathrm{d}x,$$

joka on standardinormaalijakauman normalisointivakio. Ottamalla huomioon, että integrandi on parillinen funktio ja tekemällä sitten muuttujanvaihto $t = \frac{1}{2}x^2$ integraali saadaan muotoon

$$I = 2\int_0^{\infty} \mathrm{e}^{-\frac{1}{2}x^2}\,\mathrm{d}x = 2\int_0^{\infty} \mathrm{e}^{-t}\,\frac{1}{\sqrt{2t}}\,\mathrm{d}t = \sqrt{2}\,\Gamma(\frac{1}{2}). \tag{5.6}$$

Pian näytetään vielä, että $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Tarkastellaan tuloa $\Gamma(a)\,\Gamma(b)$, jossa $a, b > 0$,

$$\Gamma(a)\,\Gamma(b) = \int_0^{\infty} u^{a-1}\,\mathrm{e}^{-u}\,\mathrm{d}u \int_0^{\infty} v^{b-1}\,\mathrm{e}^{-v}\,\mathrm{d}v.$$

Tehdään muuttujanvaihdot $u = x^2$ ja $v = y^2$, jolloin saadaan

$$\Gamma(a)\,\Gamma(b) = 4\int_0^{\infty}\int_0^{\infty} x^{2a-1}\,y^{2b-1}\,\mathrm{e}^{-x^2-y^2}\,\mathrm{d}x\,\mathrm{d}y.$$

Tässä kahden integraalin tulo on tulkittu tasointegraaliksi. Seuraavaksi siirrytään napakoordinaatteihin r ja θ,

$$x = r\cos(\theta), \quad y = r\sin(\theta), \qquad \mathrm{d}x\,\mathrm{d}y = r\,\mathrm{d}r\,\mathrm{d}\theta.$$

Edellä tasointegraali laskettiin yli ensimmäisen kvadrantin, jota napakoordinaateissa vastaa alue $0 < \theta < \pi/2$ ja $r > 0$, joten tasointegraali muuntuu muotoon

$$\begin{aligned}\Gamma(a)\,\Gamma(b) &= 4\int_{r=0}^{\infty}\int_{\theta=0}^{\frac{\pi}{2}} (r\cos\theta)^{2a-1}\,(r\sin\theta)^{2b-1}\,\mathrm{e}^{-r^2}\,r\,\mathrm{d}r\,\mathrm{d}\theta \\ &= \left(2\int_0^{\infty} r^{2a+2b-1}\,\mathrm{e}^{-r^2}\,\mathrm{d}r\right)\left(2\int_0^{\frac{\pi}{2}} (\cos\theta)^{2a-1}\,(\sin\theta)^{2b-1}\,\mathrm{d}\theta\right).\end{aligned}$$

Muuttujanvaihdolla $t = r^2$ nähdään, että tässä

$$2\int_0^{\infty} r^{2a+2b-1}\,\mathrm{e}^{-r^2}\,\mathrm{d}r = \int_0^{\infty} t^{a+b-1}\,\mathrm{e}^{-t}\,\mathrm{d}t = \Gamma(a+b),$$

ja muuttujanvaihdolla $u = \cos^2\theta$ nähdään, että

$$\begin{aligned}&2\int_0^{\frac{\pi}{2}} (\cos\theta)^{2a-1}\,(\sin\theta)^{2b-1}\,\mathrm{d}\theta \\ &\qquad = \int_0^{\frac{\pi}{2}} (\cos^2\theta)^{a-1}\,(\sin^2\theta)^{b-1}\,2\cos\theta\,\sin\theta\,\mathrm{d}\theta \\ &\qquad = -\int_1^0 u^{a-1}\,(1-u)^{b-1}\,\mathrm{d}u = \int_0^1 u^{a-1}\,(1-u)^{b-1}\,\mathrm{d}u.\end{aligned}$$

Tässä funktiota

$$B(a,b) = \int_0^1 u^{a-1}\,(1-u)^{b-1}\,\mathrm{d}u, \qquad a, b > 0 \tag{5.7}$$

kutsutaan (Eulerin) *beetafunktioksi*. Beetafunktion määrittelevä integraali (5.7) on äärellinen jos ja vain jos $a > 0$ ja $b > 0$. Edellisen johdon kuluessa saimme beetafunktiolle myös toisen integraaliesityksen, nimittäin

$$B(a, b) = 2 \int_0^{\frac{\pi}{2}} (\cos\theta)^{2a-1} (\sin\theta)^{2b-1} \, \mathrm{d}\theta,$$

ja tätä esitystä sovellamme pian.

Edellä saatiin muuttujanvaihtojen avulla johdettua beetafunktiolle esitys gammafunktion avulla kaavalla

$$B(a, b) = \frac{\Gamma(a)\,\Gamma(b)}{\Gamma(a+b)}. \tag{5.8}$$

Erityisesti

$$\Gamma(\frac{1}{2})\,\Gamma(\frac{1}{2}) = B(\frac{1}{2}, \frac{1}{2})\,\Gamma(1) = B(\frac{1}{2}, \frac{1}{2}) = 2 \int_0^{\frac{\pi}{2}} (\cos\theta)^0 (\sin\theta)^0 \, \mathrm{d}\theta = \pi,$$

joten

$$\Gamma(\frac{1}{2}) = \sqrt{\pi}, \tag{5.9}$$

kuten edellä väitettiin.

Kun tulosta $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ käytetään yhtälössä (5.6), saadaan tulos

$$\int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}x^2} \, \mathrm{d}x = \sqrt{2\pi} \tag{5.10}$$

5.3 Jatkuvia jakaumia

5.3.1 Skaalaus ja siirto

Jos sm:lla Z on jatkuva jakauma tiheysfunktiolla f_0, ja sm X määritellään kaavalla

$$X = \mu + \sigma Z, \qquad \sigma > 0,$$

niin tiheysfunktion muuntokaavan (ks. kaava (2.11) tai muistisääntö (2.12)) nojalla sm:n X tf on

$$f_X(x \mid \mu, \sigma) = \frac{1}{\sigma}\, f_0\left(\frac{x-\mu}{\sigma}\right).$$

Jos erityisesti $\sigma = 1$ ja $\mu = 0$, niin X:n tf on f_0. Tällä tavoin saadaan perhe jakaumia, jossa parametria μ kutsutaan sijaintiparametriksi (engl. *location parameter*) ja parametria $\sigma > 0$ skaalaparametriksi (engl. *scale parameter*).

Erityisesti, jos $\mu = 0$, saadaan jakaumaperhe, jonka tiheysfunktiot ovat muotoa

$$\frac{1}{\sigma}\, f_0(x/\sigma), \qquad \sigma > 0.$$

Jotkin jakaumaperheet parametroidaan muodossa

$$\lambda\, f_0(\lambda x), \qquad \lambda > 0,$$

jossa f_0 on tf. Tällöin $\sigma = 1/\lambda$ on skaalaparametri, ja parametria λ voidaan kutsua *rate*-parametriksi. (Englannin kielen sana *rate* voidaan kääntää, asiayhteydestä riippuen, esim. sanoilla intensiteetti, vauhti, osuus, aste, taajuus jne.)

5.3.2 Tasajakauma

Jos $a < b$, niin välin (a, b) tasajakaumalla, $X \sim \mathrm{U}(a, b)$ on tiheysfunktio

$$f(x) = f(x \mid a, b) = \frac{1}{b-a} 1_{(a,b)}(x), \qquad x \in \mathbb{R}\,.$$

Sen kaksi ensimmäistä momenttia lasketaan helposti integroimalla,

$$EX = \frac{1}{2}(a + b), \qquad EX^2 = \frac{1}{3}(a^2 + ab + b^2),$$

ja tästä nähdään, että

$$\operatorname{var} X = \frac{(b-a)^2}{12}.$$

5.3.3 Eksponenttijakauma

Sm X noudattaa eksponenttijakaumaa parametrilla $\lambda > 0$, eli $X \sim \mathrm{Exp}(\lambda)$, jos sen tf on

$$f(x) = f(x \mid \lambda) = \lambda \mathrm{e}^{-\lambda x}, \qquad x > 0.$$

Tässä kaavassa λ on *rate*-parametri. Jakauman odotusarvo ja varianssi ovat

$$EX = \frac{1}{\lambda}, \qquad \operatorname{var} X = \frac{1}{\lambda^2} = (EX)^2,$$

ja sen momenttiemäfunktio on

$$M(t) = \frac{\lambda}{\lambda - t}, \qquad t < \lambda.$$

Nämä tulokset seuraavat gammajakauman vastaavista tuloksista, sillä eksponenttijakauma on gammajakauman erikoistapaus.

Eksponenttijakaumalle käytetään yleisesti myös sellaista parametrointia, jossa parametrina on $\theta = 1/\lambda$. Tällöin θ on eksponenttijakauman odotusarvo ja θ^2 sen varianssi.

Eksponenttijakaumaa käytetään usein komponentin eliniän mallina. Tällöin tehdään se oletus, että komponentti ei kulu käytössä. Eksponenttijakaumalla on nimittäin seuraava mielenkiintoinen ominaisuus, ns. muistinmenetysominaisuus. Oletetaan, että komponentin elinikä $X \sim \mathrm{Exp}(\lambda)$, ja $x, h > 0$ ja lasketaan todennäköisyys, että elinikä on suurempi kuin $x + h$, kun tiedetään, että se on suurempi kuin x,

$$\begin{aligned} P(X > x + h \mid X > x) &= \frac{P(X > x + h \text{ ja } X > x)}{P(X > x)} = \frac{P(X > x + h)}{P(X > x)} \\ &= \frac{1 - F(x + h)}{1 - F(x)} = \frac{\mathrm{e}^{-\lambda(x+h)}}{\mathrm{e}^{-\lambda x}} = \mathrm{e}^{-\lambda h} = P(X > h). \end{aligned}$$

Tämä todennäköisyys ei siis riipu lainkaan siitä, kuinka kauan komponenttia on jo käytetty.

5.3.4 Gammajakauma

Gammafunktion määritelmästä (5.3) nähdään, että

$$g(x) = \frac{1}{\Gamma(\alpha)} x^{\alpha-1} \mathrm{e}^{-x}, \qquad x > 0$$

on tiheysfunktio, kun parametri $\alpha > 0$. Kutsumme vastaavaa jakaumaa gammajakaumaksi parametreilla α ja 1, ja merkitsemme sitä tunnuksella $\mathrm{Gam}(\alpha, 1)$.

Jos $\alpha > 0$, $\lambda > 0$ ja

$$Z \sim \mathrm{Gam}(\alpha, 1),$$

niin tällöin määrittelemme, että satunnaismuuttuja

$$X = Z/\lambda$$

noudattaa gammajakaumaa parametrein $\alpha > 0$ (muotoparametri) ja $\lambda > 0$ (*rate*-parametri). Jakson 5.3.1 mukaan satunnaismuuttujalla X on tiheysfunktio

$$f(x \mid \alpha, \lambda) = \lambda\, g(\lambda x),$$

jossa g on jakauman $\mathrm{Gam}(\alpha, 1)$ tiheysfunktio. Helpolla laskulla nähdään, että gammajakauman $\mathrm{Gam}(\alpha, \lambda)$ tiheysfunktio on

$$f(x) = f(x \mid \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}\, \mathrm{e}^{-\lambda x}, \qquad x > 0 \tag{5.11}$$

Huomaa, että gammajakauma ja gammafunktio ovat eri asioita; gammajakauman tiheysfunktion lausekkeessa esiintyy gammafunktio.

Jakauman $\mathrm{Gam}(\alpha, \lambda)$ momenttiemäfunktio saadaan laskettua, kun huomataan, että siinä tarvittava integraali on erään toisen gammajakauman normalisointivakio,

$$\begin{aligned} M(t) = E\mathrm{e}^{Xt} &= \int_0^\infty \mathrm{e}^{xt}\, \frac{\lambda^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}\, \mathrm{e}^{-\lambda x}\, \mathrm{d}x \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)}\, \frac{\Gamma(\alpha)}{(\lambda - t)^\alpha} \int_0^\infty \frac{(\lambda - t)^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}\, \mathrm{e}^{-(\lambda - t)x}\, \mathrm{d}x \\ &= \left(\frac{\lambda}{\lambda - t}\right)^\alpha, \qquad t < \lambda. \end{aligned}$$

Tässä laskussa *integroimme kuten tilastotietielijä*: tunnistamme että integrandi on vakiota vaille tutun jakauman tiheysfunktio, jota integroidaan kantajansa yli. Tällöin osaamme suoraan kirjoittaa lausekkeen integraalin arvolle katsomalla ko. tiheysfunktion normalisointivakiota.

Momenttiemäfunktion lausekkeesta nähdään helposti, että

$$EX = \frac{\alpha}{\lambda}, \qquad \mathrm{var}\, X = \frac{\alpha}{\lambda^2},$$

kun $X \sim \mathrm{Gam}(\alpha, \lambda)$.

Yhteyksiä muihin jakaumiin.

- $\mathrm{Exp}(\lambda)$ on $\mathrm{Gam}(1, \lambda)$.
- χ^2_n (khiin neliön jakauma n:llä vapausasteella) on $\mathrm{Gam}(n/2, 1/2)$.

Yhteenlaskuominaisuus. Jos $X \sim \mathrm{Gam}(\alpha_1, \lambda)$ ja $Y \sim \mathrm{Gam}(\alpha_2, \lambda)$ (jälkimmäinen parametri sama) ja X ja Y ovat riippumattomia, niin

$$M_{X+Y}(t) = M_X(t)\, M_Y(t) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha_1 + \alpha_2}, \qquad t < \lambda.$$

Tästä nähdään, että

$$X + Y \sim \mathrm{Gam}(\alpha_1 + \alpha_2, \lambda).$$

Huomautus. Gammajakauma parametroidaan usein myös toisella tavalla käyttämällä parametreja α ja skaalaparametria $1/\lambda$. Kummassakin parametroinnissa jälkimmäistä parametria on tapana merkitä symbolilla β. Se, kummasta parametroinnista on kyse selviää, mikäli jakauman tiheysfunktion tai odotusarvon kaava annetaan.

5.3.5 Beetajakauma

Kun $\alpha, \beta > 0$, niin beetafunktion määritelmän nojalla funktio

$$f(x) = f(x \mid \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, \qquad 0 < x < 1,$$

on tiheysfunktio. Käytämme sille tunnusta $\mathrm{Be}(\alpha, \beta)$. Tämän jakauman momentit on helppo laskea, sillä ne saadaan suoraan erään toisen beetajakauman normalisointivakiosta (ts. integroimalla kuten tilastotieteilijä). Jos $r > 0$, on

$$EX^r = \frac{B(\alpha + r, \beta)}{B(\alpha, \beta)}.$$

Erityisesti

$$EX = \frac{\alpha}{\alpha + \beta}, \qquad EX^2 = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)},$$

josta

$$\operatorname{var} X = EX^2 - (EX)^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

Välin $(0,1)$ tasajakauma $\mathrm{U}(0,1)$ on sama kuin $\mathrm{Be}(1,1)$.

5.3.6 Normaalijakauma

Standardinormaalijakaumaa noudattavan sm:n $Z \sim N(0,1)$ tf on

$$f_Z(z) = \phi(z) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2} z^2), \qquad z \in \mathbb{R}\,. \tag{5.12}$$

Standardinormaalijakuman tiheysfunktiolle käytetään yleisesti merkintää ϕ ja sen kertymäfunktiolle merkintää Φ. (Asiayhteydestä riippuen ϕ voi tarkoittaa joko karakteristista funktiota tai standardinormaalijakauman tiheysfunktiota tai jotakin muuta.) Se seikka, että kyseessä on tiheysfunktio seuraa siitä, että funktio on aidosti positiivinen ja että sen integraali koko reaaliakselin yli on kaavan (5.10) nojalla yksi.

Lasketaan seuraavaksi $N(0,1)$:n momenttiemäfunktio.

$$M_Z(t) = E\mathrm{e}^{tZ} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2} z^2 + tz) \,\mathrm{d}z.$$

Täydennetään eksponenttifunktion argumentti neliöksi,

$$-\frac{1}{2} z^2 + tz = -\frac{1}{2}(z-t)^2 + \frac{1}{2} t^2,$$

minkä jälkeen (sijoituksella $u = z - t$) nähdään, että

$$M_Z(t) = \exp(\frac{1}{2} t^2) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} u^2\right) \mathrm{d}u = \exp(\frac{1}{2} t^2).$$

Laskemalla kumulanttiemäfunktion $K_Z(t) = \frac{1}{2}t^2$ derivaatat nähdään, että

$$EZ = 0, \qquad \operatorname{var} Z = 1.$$

Sm X noudattaa normaalijakaumaa parametrein μ, σ^2, eli $X \sim N(\mu, \sigma^2)$, jos se voidaan esittää muodossa

$$X = \mu + \sigma Z, \qquad Z \sim N(0, 1). \tag{5.13}$$

Tällöin

$$EX = \mu, \qquad \operatorname{var} X = \sigma^2 \operatorname{var} Z = \sigma^2,$$

joten normaalijakauman parametrit ovat sen odotusarvo ja varianssi. Jos $\sigma > 0$, niin jakauman tf on

$$f_X(x) = f_X(x \mid \mu, \sigma^2) = \frac{1}{\sigma}\,\phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right), \qquad x \in \mathbb{R}. \tag{5.14}$$

Jakauman momenttiemäfunktio on

$$\begin{aligned} M_X(t) &= M_{\mu+\sigma Z}(t) = E\exp((\mu + \sigma Z)t) = \exp(\mu t)\, M_Z(\sigma t) \\ &= \exp(\mu t + \frac{1}{2}\sigma^2 t^2). \end{aligned} \tag{5.15}$$

Yhteenlaskuominaisuus. Jos $X \sim N(\mu_1, \sigma_1^2)$ ja $Y \sim N(\mu_2, \sigma_2^2)$, ja ne ovat riippumattomia, niin

$$M_{X+Y}(t) = M_X(t)\, M_Y(t) = \exp((\mu_1 + \mu_2)t + \frac{1}{2}(\sigma_1^2 + \sigma_2^2)t^2),$$

joten $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

5.3.7 Normaalijakaumasta johdettuja jakaumia

Khiin neliön jakauma. Jos $X_1, \dots, X_n \sim N(0, 1)$ ja ne ovat riippumattomia, niin niiden neliöiden summalla sanotaan olevan χ^2_n-jakauma eli khiin neliön jakauma n:llä vapausasteella (tai vapausasteluvulla n).

Khiin neliön jakauma on erikoistapaus gammajakaumasta, mikä nähdään seuraavasti. Jos $X \sim N(0, 1)$, niin muuttujan $Y = X^2$ tiheysfunktio on (vrt. esim. 2.10)

$$f_Y(y) = (\phi(\sqrt{y}) + \phi(-\sqrt{y}))\frac{1}{2\sqrt{y}} = \frac{1}{\sqrt{2\pi}}\, y^{-1/2}\exp(-\frac{1}{2}y),$$

kun $y > 0$. Tästä nähdään, että $X^2 \sim \mathrm{Gam}(\frac{1}{2}, \frac{1}{2})$. Gammajakauman yhteenlaskuominaisuuden perusteella $X_1^2 + X_2^2 \sim \mathrm{Gam}(\frac{1}{2} + \frac{1}{2}, \frac{1}{2})$, ja samaa päättelyä jatkamalla

$$X_1^2 + \cdots + X_n^2 \sim \mathrm{Gam}(\frac{n}{2}, \frac{1}{2}).$$

Toisin sanoen χ^2_n on sama jakauma kuin $\mathrm{Gam}(\frac{n}{2}, \frac{1}{2})$. Jos $\nu > 0$ ei ole kokonaisluku, niin määrittelemme, että χ^2_ν jakauma on sama kuin gammajakauma $\mathrm{Gam}(\frac{\nu}{2}, \frac{1}{2})$.

t-jakauma. Jos $Z \sim N(0, 1)$ ja $V \sim \chi^2_\nu$ (jollekin $\nu > 0$) ja $Z \perp\!\!\!\perp V$, niin satunnaismuuttujalla

$$T = \frac{Z}{\sqrt{V/\nu}}$$

on jakauma, jota kutsutaan (Studentin) t-jakaumaksi ν:llä vapausasteella tai vapausasteluvulla ν (engl. *degrees of freedom, df*). Merkitsemme tämän asian $T \sim t_\nu$.

F-jakauma. Jos $U \sim \chi^2_k$ ja $V \sim \chi^2_m$ ja $U \perp\!\!\!\perp V$, niin sm:lla

$$Y = \frac{U/k}{V/m}$$

on jakauma, jota kutsutaan F-jakaumaksi parametreilla k (osoittajan vapausasteluku) ja m (nimittäjän vapausasteluku), eli $Y \sim F_{k,m}$.

Tarkastelemme myöhemmin lähemmin joitakin t-jakauman ominaisuuksia.