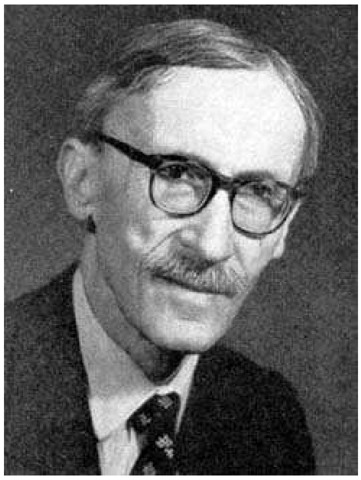
Todennäköisyysteoria

Teoria mitasta, mitallisuudesta, mitattomuudesta ja riippumattomuudesta

A. Kolmogorov



P. Lévy

Tommi Sottinen

tommi.sottinen@helsinki.fi
mathstat.helsinki.fi/~tsottine

1. joulukuuta 2006

Lukijalle

Nämä ovat luentomuistiinpanot Helsingin yliopiston Matematiikan ja tilastotieteen laitoksella syksyllä 2006 luennoimalleni 10 opintopisteen kurssille *Todennäköisyysteoria*.

Kurssi sisältää noin 48 tuntia luentoja ja 24 tuntia harjoituksia.

Luennot perustuvat E. Valkeilan luentoihin syksyltä 2001 ja H. Nyrhisen luentoihin keväältä 2006. Lisäksi olen käyttänyt lähteitä

1. Ya. Sinai *Probability Theory: An Introductory Course*, Springer-Verlag, Berliini, 1992.
2. G. Elfving ja P. Tuominen *Todennäköisyyslaskenta II*, toinen painos, Limes, Helsinki, 1990.
3. J. Jacod ja P. Protter *Probability Essentials*, toinen painos, Springer-Verlag, New York, 2004.
4. D. Williams *Probability with Martingales*, Cambridge University Press, Cambridge, 1991.
5. A. Shiryaev *Probability*, Springer-Verlag, New York, 1984.
6. O. Kallenberg *Foundations of modern probability*, toinen painos, Springer-Verlag, New York, 2002.

Lähteet on lueteltu vaikeusjärjestyksessä, mikä on myös laajuusjärjestys.

Helsingissä 1. joulukuuta 2006 T. S.

Sisältö

Luku 1

Johdanto

Aloitammme kertaamalla merkintöjä ja käsitteitä, jotka ovat luultavasti lukijalle varsin tuttuja (osio 1.1). Tämän jälkeen esitämme ylimalkaisesti todennäköisyyslaskennan "suuret raja-arvotulokset" (osio 1.2).

1.1 Merkintöjä

Luonnolliset luvut ovat $\mathbb{N} := \{1, 2, \ldots\}$, kokonaisluvut ovat $\mathbb{Z}$, rationaaliluvut ovat $\mathbb{Q}$, reaaliluvut ovat $\mathbb{R}$ ja $\mathbb{C}$ tarkoittaa kompleksilukuja. Lisäksi $\mathbb{R}_+ := \{x \in \mathbb{R}\,;\, x \geq 0\}$ ja $\mathbb{R}^d$ tarkoittaa d-ulotteista euklidista avaruutta.

Merkintä $|X|$ tarkoittaa joko X:n itseisarvoa, modulia, euklidista normia tai alkioiden lukumäärää riippuen siitä, onko X reaalinen, kompleksinen, vektori vai joukko.

Joukko N on *numeroituva*, jos on olemassa injektio $f : N \to \mathbb{N}$. äärelliset joukot sekä joukot $\mathbb{N}$, $\mathbb{Z}$ ja $\mathbb{Q}$ ovat numeroituvia. Joukot $\mathbb{R}$ ja $\mathbb{C}$ eivät ole numeroituvia. Jos joukko ei ole numeroituva, niin se on *ylinumeroituva*.

Perusjoukko tai -avaruus on Ω. Tästä joukosta kohtalon jumalatar valitsee satunnaiskokeen alkeistapauksen ω.

Osajoukon $A \subset \Omega$ *komplementti* on $A^c := \{\omega \in \Omega\,;\, \omega \notin A\}$.

Kaikkien joukon E osajoukkojen kokoelma on sen *potenssijoukko* $\mathrm{pot}\,(E)$: $A \in \mathrm{pot}\,(E)$ jos ja vain jos $A \subset E$.

Olkoon J jokin indeksijoukko ja $A_j \subset \Omega$ kaikilla $j \in J$. Tällöin

$$\begin{aligned}
\bigcup A_j \;&:=\; \bigcup_{j \in J} A_j \;:=\; \{\omega \in \Omega\,;\, \omega \in A_j \text{ jollakin } j \in J\}\,,\\
\bigcap A_j \;&:=\; \bigcap_{j \in J} A_j \;:=\; \{\omega \in \Omega\,;\, \omega \in A_j \text{ kaikilla } j \in J\}\,.
\end{aligned}$$

(Jätämme indeksijoukot merkitsemättä, jos sekaannuksen vaaraa ei ole.) Todennäköisyyslaskennan kielellä $\cap A_j$ on "kaikki A_j:t sattuvat" ja $\cup A_j$ on "vähintään yksi A_j sattuu". Tulkitsemme $\cup_{j\in\varnothing}A_j = \varnothing$ ja $\cap_{j\in\varnothing}A_j = \Omega$. Indeksijoukon J ei tarvitse olla numeroituva: esimerkiksi $\mathbb{R} = \cup_{x\in\mathbb{R}}\{x\}$.

Osajoukkokokoelmalle $A_j \subset \Omega$, $j \in J$, pätee *De Morganin lait*

$$\Big(\bigcup A_j\Big)^c = \bigcap A_j^c \quad \text{ja} \quad \Big(\bigcap A_j\Big)^c = \bigcup A_j^c.$$

Funktiolle $X : \Omega \to \mathbb{R}^d$ merkitsemme

$$\{X \in B\} := X^{-1}B := \{\omega \in \Omega\,;\, X(\omega) \in B\}\,.$$

Tässä siis $\{X \in B\} \subset \Omega$ ja $B \subset \mathbb{R}^d$.

Laajennetut reaaliluvut ovat $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$.

Reaalilukujonon (x_n) ala- ja yläraja-arvot ovat

$$\liminf x_n := \lim_n \inf_{k\ge n} x_k \quad \text{ja} \quad \limsup x_n := \lim_n \sup_{k\ge n} x_k.$$

$\liminf x_n$ ja $\limsup x_n$ ovat aina olemassa laajennettuina reaalilukuina (harjoitustehtävä 1.2). Lisäksi $x_n \to x$ jos ja vain jos $\liminf x_n = \limsup x_n = x$ (harjoitustehtävä 1.3). On myös selvää, että $\liminf x_n \le \limsup x_n$ kaikilla jonoilla (x_n). Nimitykset "ylä- ja alaraja-arvot" johtuvat tästä.

Joukkojonojen (A_n) *ala- ja yläraja-arvot* ovat

$$\liminf A_n := \bigcup_{n\ge 1}\bigcap_{k\ge n} A_k \quad \text{ja} \quad \limsup A_n := \bigcap_{n\ge 1}\bigcup_{k\ge n} A_k.$$

Todennäköisyyslaskennan kielellä $\limsup A_n$ on "A_n sattuu äärettömän usein" ja $\liminf A_n$ on "A_n sattuu jostakin lähtien". Usein käytetäänkin merkintöjä $\{A_n \text{ ä.u.}\} := \limsup A_n$ ja $\{A_n \text{ j.l.}\} := \liminf A_n$.

Jos $x_n \to x$ ja (x_n) on kasvava, niin merkitsemme $x_n \uparrow x$. Jos $x_n \to x$ ja (x_n) on vähenevä, niin merkitsemme $x_n \downarrow x$. Itse asiassa: jos (x_n) on kasvava, niin $x_n \uparrow \sup x_n \in \bar{\mathbb{R}}$, ja jos (x_n) on vähenevä, niin $x_n \downarrow \inf x_n \in \bar{\mathbb{R}}$.

Joukkojono (A_n) on kasvava, jos $A_n \subset A_{n+1}$ ja vähenevä, jos $A_n \supset A_{n+1}$. Merkintä $A_n \uparrow A$ tarkoittaa, että (A_n) on kasvava ja $A = \cup A_n$. Merkintä $A_n \downarrow A$ tarkoittaa, että (A_n) on vähenevä ja $A = \cap A_n$. Jos $\limsup A_n = \liminf A_n = A$, niin sanomme, että (A_n) suppenee kohti joukkoa A. Tällöin merkitsemme $A_n \to A$ tai $\lim A_n = A$.

Joukon A *indikaattori* on

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

Tällöin $A_n \to A$ jos ja vain jos $\mathbf{1}_{A_n}(\omega) \to \mathbf{1}_A(\omega)$ kaikilla $\omega \in \Omega$ (harjoitustehtävä 1.6).

1.2 Kolikonheittoa

Esitämme raja-arvotuloksia reilun kolikon heitoille ylimalkaisesti. Tarkemmin ja yleisemmin käsittelemme asiaa luvussa 8.

Olkoon

$$X_n \;:=\; \begin{cases} 1, & \text{jos } n. \text{ heitto on klaava,} \\ 0, & \text{jos } n. \text{ heitto on kruuna.} \end{cases}$$

Oletamme, että heitot ovat riippumattomia ja että kolikko on reilu. Toisin sanoen kaikille mahdollisille tuloksille $(x_1, \ldots, x_n) \in \{0,1\}^n$, $n \in \mathbb{N}$, pätee

$$\mathbf{P}[X_1 = x_1, \ldots, X_n = x_n] \;=\; \frac{1}{2^n}.$$

Olkoon

$$Y_n \;:=\; \frac{1}{n} \sum_{k \leq n} X_k$$

klaavojen *suhteellinen frekvenssi* n kolikonheiton sarjassa.

Kun $n \to \infty$, todennäköisyyksille $\mathbf{P}[Y_n = y]$, $y \in [0,1]$, pätee

$$\mathbf{P}[Y_n = y] \;\simeq\; \mathbf{1}_{\{y = \frac{1}{2}\}}, \tag{1.2.1}$$

$$\simeq\; e^{-n\left(y \ln y + (1-y)\ln(1-y) + \ln 2\right)}, \tag{1.2.2}$$

$$\simeq\; \frac{1}{\sqrt{\pi \frac{n}{2}}} e^{-n\left(2(y-\frac{1}{2})^2\right)}. \tag{1.2.3}$$

Arvio (1.2.1) on (heikko) *suurten lukujen raja*, arvio (1.2.2) on *suurten poikkeamien raja* ja arvio (1.2.3) on *keskeinen raja*. Näitä arvioita voitaneen pitää todennäköisyyslaskennan keskeisimpinä tuloksina

Perustelemme nyt kaavat (1.2.1)–(1.2.3) heuristisesti. Aluksi huomaamme, että kuvaus

$$y \mapsto I(y) \;:=\; y \ln y + (1-y)\ln(1-y) + \ln 2$$

on positiivinen ja saa pienimmän arvonsa (nolla) pisteessä $y = 1/2$. Siten arvio (1.2.1) seuraa arviosta (1.2.2). Perustelemme sitten arvion (1.2.2). Työkalumme on *Stirlingin kaava*:

$$n! \;\simeq\; \sqrt{2\pi n}\, n^n e^{-n}, \qquad \text{kun } n \to \infty.$$

Koska nY_n on binomijakautunut parametrein n ja $1/2$, niin (kun $ny \in \mathbb{N}$)

$$\mathbf{P}[Y_n = y] \;=\; \mathbf{P}[nY_n = ny] \;=\; \binom{n}{ny} 2^{-n} \;=\; \frac{n!\,2^{-n}}{(ny)!(n-ny)!}.$$

Soveltamalla tähän Stirlingin kaavaa saamme

$$\begin{aligned}
\mathbf{P}[Y_n = y] &\simeq \frac{\sqrt{2\pi n} n^n e^{-n}\, 2^{-n}}{\sqrt{2\pi ny}(ny)^{ny} e^{-ny} \sqrt{2\pi(n-ny)}(n-ny)^{n-ny} e^{-(n-ny)}} \\
&= \frac{1}{\sqrt{2\pi ny(1-y)}} y^{-ny}(1-y)^{-n(1-y)} 2^{-n} \\
&= \frac{1}{\sqrt{2\pi ny(1-y)}} e^{-n(y\ln y + (1-y)\ln(1-y) + \ln 2)}. \qquad (1.2.4)
\end{aligned}$$

Arvio (1.2.2) seuraa arviosta (1.2.4) hävittämällä hitaasti muuttuva termi:

$$\frac{1}{\sqrt{2\pi ny(1-y)}} e^{-n(y\ln y + (1-y)\ln(1-y) + \ln 2)} \simeq e^{-n(y\ln y + (1-y)\ln(1-y) + \ln 2)}.$$

Perustelemme sitten arvion (1.2.3). Juuri perustellun suurten lukujen arvion (1.2.1) nojalla voimme olettaa, että

$$y \simeq \frac{1}{2}.$$

Arvio (1.2.3) seuraa sijoittamalla $y(1-y) \simeq 1/4$ arvion (1.2.4) hitaasti muuttuvaan termiin ja toisen asteen Taylor-arvion

$$I(y) \simeq I\left(\frac{1}{2}\right) + I'\left(\frac{1}{2}\right)\left(y - \frac{1}{2}\right) + \frac{1}{2} I''\left(\frac{1}{2}\right)\left(y - \frac{1}{2}\right)^2$$

arvion (1.2.4) eksponenttiin. Nimittäin nyt

$$\begin{aligned}
I'(y) &= \ln y - \ln(1-y), \\
I''(y) &= \frac{1}{y} + \frac{1}{1-y},
\end{aligned}$$

ja

$$I\left(\frac{1}{2}\right) = 0, \qquad I'\left(\frac{1}{2}\right) = 0, \quad \text{sekä} \quad I''\left(\frac{1}{2}\right) = 4.$$

Siten Taylor-arvio on

$$I(y) \simeq 2\left(y - \frac{1}{2}\right)^2.$$

Suurten poikkeamien arvio (1.2.2) toimii paremmin, kun suhteellinen frekvenssi y on kaukana suurten lukujen arviostaan $y \simeq 1/2$. Keskeinen arvio (1.2.3) taas toimii paremmin, kun y on lähellä suurten lukujen arvion (1.2.1) ennustetta $y \simeq 1/2$. Arvioiden (1.2.2) ja (1.2.3) erilaiset “toiminta-alueet” ovat intuitiivisesti selviä, kun huomaa miten ne on johdettu: arvio (1.2.2) johdettiin hylkäämällä hitaasti muuttuva termi ja arvio (1.2.3) johdettiin käyttämällä Taylor-kehitelmää tyypillisen arvon $y \simeq 1/2$ ympärillä.

Heikon suurten lukujen arvion (1.2.1) taustalla on *heikko suurten lukujen laki*: $\lim \mathbf{P}[|Y_n - 1/2| > \varepsilon] = 0$ kaikilla $\varepsilon > 0$. *Vahva suurten lukujen* laki taas sanoo, että $\mathbf{P}[\lim Y_n = 1/2] = 1$. Tässä meidän on tulkittava $\{\lim Y_n = 1/2\}$ tapahtumaksi todennäköisyysavaruudella, joka koostuu äärettömän pitkistä kolikonheittojen sarjoista. Tätä varten tarvitsemme todennäköisyys*teoriaa*. Suurten poikkeamien arvion (1.2.2) takana on *suurten poikkeamien periaate* ja keskeisen arvion (1.2.3) takana on *keskeinen raja-arvolause*.

Harjoitustehtäviä

1.1. Olkoon $j \in J$, missä J on mielivaltainen indeksijoukko. Perustele De Morganin lait $(\cup A_j)^c = \cap A_j^c$ ja $(\cap A_j)^c = \cup A_j^c$.

1.2. Olkoon (x_n) lukujono. Osoita, että $\liminf x_n$ ja $\limsup x_n$ ovat aina olemassa laajennettuina reaalilukuina.

1.3. Olkoon (x_n) lukujono. Osoita, että (x_n) suppenee jos ja vain jos $\liminf x_n = \limsup x_n$, ja tällöin $\lim x_n = \liminf x_n$.

1.4. Olkoon A jokin joukko. Jokaiselle $n \in \mathbb{N}$ asetamme $A_{2n-1} := A$ ja $A_{2n} := A^c$. Mitä ovat joukot $\liminf A_n$ ja $\limsup A_n$? Suppeneeko jono (A_n)?

1.5. Olkoon (A_n) joukkojono. Osoita, että $\liminf A_n \subset \limsup A_n$.

1.6. Olkoon (A_n) jono perusjoukon Ω joukkoja. Osoita, että $A_n \to A$ jos ja vain jos $\mathbf{1}_{A_n}(\omega) \to \mathbf{1}_A(\omega)$ kaikilla $\omega \in \Omega$.

1.7. (a) Muotoile ja (b) perustele osion 1.2 tulokset (1.2.1)–(1.2.3) painotetun kolikon heitolle.

Oletamme siis, että kaikilla $(x_1, \ldots, x_n) \in \{0,1\}^n$, $n \in \mathbb{N}$,

$$\begin{aligned} \mathbf{P}[X_1 = x_1, \ldots, X_n = x_n] &= p^{|\{x_i\,;\, x_i=1, i\le n\}|}(1-p)^{|\{x_1\,;\, x_i=0, i\le n\}|} \\ &= p^{\sum_{i\le n} x_i}(1-p)^{n-\sum_{i\le n} x_i} \end{aligned}$$

jollakin $p \in [0,1]$. Kolikko on siis painotettu, mutta muistiton (eli kolikonheitto on riippumaton).

Luku 2

Todennäköisyysavaruus

Valitettavasti monissa mielenkiintoisissa tilanteissa emme voi käsitellä kaikkia perusjoukon Ω osajoukkoja tapahtumina eli "mitallisina" joukkoina. Emme nimittäin aina pysty liittämään jokaiseen osajoukkoon $A \subset \Omega$ todennäköisyyttä $\mathbf{P}[A]$ johdonmukaisella tavalla, vaan meidän on tyydyttävä johonkin $\operatorname{pot}(\Omega)$:n osajoukkoon. Konkreettisesti näemme tämän esimerkissä 2.4.1. Toisaalta, jos A on tapahtuma, niin A^c on tapahtuma. Nimittäin "terve järki" sanoo, että tällöin $\mathbf{P}[A^c] = 1 - \mathbf{P}[A]$. Lisäksi, jos A ja B ovat erillisiä tapahtumia, niin sama"terve järki" sanoo, että $A \cup B$ on tapahtuma ja $\mathbf{P}[A \cup B] = \mathbf{P}[A] + \mathbf{P}[B]$. Matemaattisesti on myöskin huomattu välttämättömäksi olettaa todennäköisyyden jatkuvuus: jos $(A_n) \downarrow A$, niin $\mathbf{P}[A_n] \downarrow \mathbf{P}[A]$. Erityisesti siis tällöin osajoukko $A \subset \Omega$ on tapahtuma, jos osajoukot $A_n \subset \Omega$, $n \in \mathbb{N}$, ovat tapahtumia. Matemaattinen välttämättömyys tarkoittaa tässä sitä, että ilman tätä jatkuvuutta vahva suurten lukujen laki ei pitäisi paikkaansa. Siten tämäkin tapahtumien ominaisuus seuraa "terveestä järjestä".

Edellä mainituista syistä seuraa, että tapahtumien (eli "mitallisten" joukkojen) luokan on oltava suljettu numeroituvan monien joukko-operaatioiden suhteen. Tällaista joukkoluokkaa kutsumme σ-algebraksi.

2.1 σ-algebrat

2.1.1 Määritelmä. Kokoelma $\mathcal{F}$ joukon Ω osajoukkoja on σ-*algebra*, jos

(i) $\Omega \in \mathcal{F}$,
(ii) jos $A \in \mathcal{F}$, niin $A^c \in \mathcal{F}$,
(iii) jos $A_n \in \mathcal{F}$ kaikilla $n \in \mathbb{N}$, niin $\cap A_n \in \mathcal{F}$.

Määritelmän 2.1.1 ehdon (i) voi korvata ehdolla

(i') $\varnothing \in \mathcal{F}$,

sillä ehdon (ii) nojalla tällöin $\varnothing^c = \Omega \in \mathcal{F}$. Ehdon (iii) voi korvata ehdolla

(iii') jos $A_n \in \mathcal{F}$ kaikilla $n \in \mathbb{N}$, niin $\cup A_n \in \mathcal{F}$,

sillä $\cup A_n = (\cap A_n^c)^c$ ja $(\cap A_n^c)^c \in \mathcal{F}$ ehdon (ii) nojalla.

2.1.2 Esimerkki. (i) Täysi σ-algebra $\operatorname{pot}(\Omega)$ on σ-algebra.
(ii) Triviaali σ-algebra $\{\Omega, \varnothing\}$ on σ-algebra.
(iii) Olkoon $\Omega = \{1,2,3\}$. Tällöin $\{\Omega, \{2,3\}, \{1\}, \varnothing\}$ on σ-algebra.
(iv) Olkoon $\Omega = \{1,2,3\}$. Kokoelma $\{\Omega, \{1\}, \{2\}, \{3\}, \varnothing\}$ ei ole σ-algebra.
(v) Olkoon $A \subset \Omega$. Tällöin $\{\Omega, A, A^c, \varnothing\}$ on σ-algebra.
(vi) Reaaliakselin avoimien välien muodostama kokoelma ei ole σ-algebra, sillä esimerkiksi $\{0\} = \cap(-\frac{1}{n}, \frac{1}{n})$ ei ole avoin väli.

2.1.3 Apulause. *Jos $\mathcal{F}_j$, $j \in J$, ovat σ-algebroja, niin $\cap\mathcal{F}_j$ on σ-algebra.*

Todistus. (i) $\Omega \in \cap\mathcal{F}_j$, sillä $\Omega \in \mathcal{F}_j$ kaikilla $j \in J$. (ii) jos $A \in \cap\mathcal{F}_j$, niin $A \in \mathcal{F}_j$ kaikilla $j \in J$. Siten $A^c \in \mathcal{F}_j$, sillä $\mathcal{F}_j$ on σ-algebra. Joten $A^c \in \cap\mathcal{F}_j$. (iii) Jos A_n, $n \in \mathbb{N}$, kuuluvat joukkoon $\cap\mathcal{F}_j$, niin ne kuuluvat joukkoon $\mathcal{F}_j$ kaikilla $j \in J$. Siten $\cap A_n \in \mathcal{F}_j$ kaikilla $j \in J$, sillä joukot $\mathcal{F}_j$, $j \in J$, ovat σ-algebroja. Siispä $\cap A_n \in \cap\mathcal{F}_j$. □

Apulause 2.1.3 mahdollistaa seuraavan määritelmän.

2.1.4 Määritelmä. Kokoelman $\mathcal{C}$ *virittämä σ-algebra* $\sigma(\mathcal{C})$ on pienin σ-algebra, joka sisältää kokoelman $\mathcal{C}$:

$$\sigma(\mathcal{C}) \;:=\; \bigcap \big\{\mathcal{F} : \mathcal{F} \text{ on } \sigma\text{-algebra ja } \mathcal{C} \subset \mathcal{F}\big\}.$$

Esimerkin 2.1.2(v) σ-algebra on kokoelman $\{A\}$, tai yhtä hyvin joukon A, virittämä σ-algebra. Samoin esimerkin 2.1.2(iii) σ-algebra on kokoelman $\{\{1\}\}$, eli joukon $\{1\}$, virittämä σ-algebra.

2.1.5 Apulause. *Jos $\mathcal{C}' \subset \mathcal{C}$, niin $\sigma(\mathcal{C}') \subset \sigma(\mathcal{C})$.*

Todistus. $\sigma(\mathcal{C})$ on σ-algebra, joka sisältää kokoelman $\mathcal{C}'$. Toisaalta $\sigma(\mathcal{C}')$ on pienin σ-algebra, joka sisältää kokoelman $\mathcal{C}'$. Siten $\sigma(\mathcal{C}') \subset \sigma(\mathcal{C})$. □

Osajoukko $U \subset \mathbb{R}^d$ on *avoin*, jos U:n jokainen piste x on posiitiivisen etäisyyden päässä joukon U komplementista. Toisin sanoen kaikille $x \in U$ pätee

$$\operatorname{dist}(x, U^c) \;:=\; \inf\{|x-y|\,;\, y \in U^c\} \;>\; 0.$$

2.1.6 Määritelmä. Euklidisen avaruuden $\mathbb{R}^d$ *Borelin σ-algebra* $\mathcal{B}_d$ on sen avointen joukkojen virittämä. Joukot $B \in \mathcal{B}_m$ ovat *Borel-joukkoja.* Lisäksi merkitsemme $\mathcal{B} := \mathcal{B}_1$.

2.1.7 Apulause. *Jokainen seuraavista kokoelmista virittää $\mathcal{B}$:n.*

(i) $[a,b]$, $a \le b$, $a,b \in \mathbb{R}$,
(ii) $[a,b]$, $a < b$, $a,b \in \mathbb{R}$,
(iii) $(a,b]$, $a \le b$, $a,b \in \mathbb{R}$,
(iv) $(a,b]$, $a < b$, $a,b \in \mathbb{R}$,
(v) (a,b), $a < b$, $a,b \in \mathbb{R}$,
(vi) $(-\infty,b]$, $b \in \mathbb{R}$,
(vii) $(-\infty,b)$, $b \in \mathbb{R}$,
(viii) $[a,\infty)$, $a \in \mathbb{R}$,
(ix) (a,∞), $a \in \mathbb{R}$,
(x) *mikä tahansa edellisistä kokoelmista, mutta niin että $a,b \in \mathbb{Q}$.*

Todistus. Näytämme kohdan (v) ja jätämme loput harjoitustehtäväksi 2.3. Osoitamme, että jokainen $\mathbb{R}$:n avoin joukko on numeroituva yhdiste avoimista väleistä (a,b). Väite seuraa tällöin määritelmän 2.1.1 ehdosta (iii'). Olkoon $B \subset \mathbb{R}$ avoin ja $B' = B \cap \mathbb{Q}$. Tällöin

$$B = \bigcup_{q \in B'} (q - \varepsilon_q, q + \varepsilon_q),$$

missä $\varepsilon_q = \mathrm{dist}(q, B^c) > 0$. □

Apulauseella 2.1.7 on myös luonnollinen d-ulotteinen vastine. Esimerkiksi joukot $\times_{k \le d}[a_k, \infty)$, $a_k \in \mathbb{Q}$, virittävät σ-algebran $\mathcal{B}_d$ (harjoitustehtävä 2.4).

2.2 Todennäköisyysmitta ja -avaruus

2.2.1 Määritelmä. Pari $(\Omega, \mathcal{F})$, missä $\mathcal{F}$ on perusjoukon Ω σ-algebra on *mitallinen avaruus.* Joukot $A \in \mathcal{F}$ ovat mitallisia joukkoja eli *tapahtumia.*

2.2.2 Määritelmä. *Todennäköisyysmitta* on funktio $\mathbf{P} : \mathcal{F} \to [0,1]$, jolle

(i) $\mathbf{P}[\Omega] = 1$,
(ii) jos tapahtumat $A_n \in \mathcal{F}$, $n \in \mathbb{N}$, ovat erillisiä, niin

$$\mathbf{P}\left[\bigcup A_n\right] = \sum \mathbf{P}[A_n].$$

Todennäkoisyysmitta on siis mitan erikoistapaus. Yleisesti nimittäin *mitta* on joukkokuvaus $\mu : \mathcal{F} \to \mathbb{R}_+ \cup \{\infty\}$, jolle $\mu[\varnothing] = 0$ ja

$$\mu\left[\bigcup A_n\right] \;=\; \sum \mu[A_n],$$

kun mitalliset joukot $A_n \in \mathcal{F}$, $n \in \mathbb{N}$, ovat erillisiä. Mikäli $\mu[\Omega] < \infty$, niin sanomme, että mitta μ on *äärellinen* tai *rajoitettu*.

Seuraava määritelmä yhdistettynä määritelmiin 2.1.1 ja 2.2.2 on *Kolmogorovin aksiomatisointi* todennäköisyyslaskennalle. Huomattavaa määritelmässä on, että se on puhtaasti tekninen eikä ollenkaan filosofinen: se ei siis sano, mitä todennäköisyys on “an sich” tai missään muussakaan filosofisessa mielessä. Se sanoo vain, miten todennäköisyys “käyttäytyy”.

2.2.3 Määritelmä. Kolmikko $(\Omega, \mathcal{F}, \mathbf{P})$ on *todennäköisyysavaruus.*

2.2.4 Esimerkki. (i) Tavallista nopanheittoa vastaa todennäköisyysavaruus $\Omega = \{1, \ldots, 6\}$, $\mathcal{F} = \mathrm{pot}\,(\Omega)$ ja $\mathbf{P}[A] = |A|/6$ kaikilla $A \in \mathcal{F}$.

(ii) n-kertaista kolikonheittoa vastaa todennäköisyysavaruus $\Omega = \{0,1\}^n$, $\mathcal{F} = \mathrm{pot}\,(\Omega)$ ja $\mathbf{P}[A] = |A|/2^n$ kaikilla $A \in \mathcal{F}$.

(iii) Loputonta kolikonheittoa vastaa todennäköisyysavaruus $\Omega = \{0,1\}^\infty$. Tällöin $\omega \in \Omega$ on siis äärettömän pitkä $\{0,1\}$-jono:

$$\omega \;=\; (\omega_1, \omega_2, \ldots, \omega_n, \omega_{n+1}, \ldots).$$

Todennäköisyys virittyy asettamalla $\mathbf{P}[\{\mathrm{proj}_n \omega\}] \;=\; 1/2^n$, missä $\mathrm{proj}_n(\omega_1, \omega_2, \ldots) \;=\; (\omega_1, \ldots, \omega_n)$. Vastaavasti σ-algebraksi valitaan $\mathcal{F} \;=\; \sigma(\mathrm{proj}_n A\,;\, A \in \mathrm{pot}\,(\Omega), n \in \mathbb{N})$. Joukkoja $\mathrm{proj}_n A$ kutsutaan *sylinterijoukoiksi.* Tässä tapauksessa on valitettavasti mahdotonta valita σ-algebraksi koko potenssijoukkoa $\mathrm{pot}\,(\{0,1\}^\infty)$. Tämän näkee myöhemmin esitettävästä esimerkistä 2.4.1 tulkitsemalla avaruuden $\{0,1\}^\infty$ avaruudeksi $[0,1]$ käyttämällä reaaliluvun desimaalikehitelmän binääriesitystä.

Todennäköisyyden ominaisuutta (ii) kutsutaan *σ-additiivisuudeksi.* Tavallinen äärellinen *additiivisuus* joukkofunktioille $\mathbf{P}$ tarkoittaa sitä, että $\mathbf{P}[A_1 \cup A_2] = \mathbf{P}[A_1] + \mathbf{P}[A_2]$, kun A_1 ja A_2 ovat erillisiä eli $A_1 \cap A_2 = \varnothing$. Itse asiassa hakasulkeiden käytöllä haluamme korostaa todennäköisyysmitan additiivisuutta.

Seuraavan apulauseen sanoma on, että σ-additiivisuus lisää äärelliseen additiivisuuteen monotonisen jatkuvuuden.

2.2.5 Apulause. *Olkoon* $\mathbf{P} : \mathcal{F} \to [0,1]$ *additiivinen funktio, jolle* $\mathbf{P}[\Omega] = 1$. *Tällöin seuraavat ehdot ovat yhtäpitäviä:*

(i) $\mathbf{P}$ *on* σ*-additiivinen,*
(ii) *jos* $A_n, A \in \mathcal{F}$, $n \in \mathbb{N}$, *ja* $A_n \downarrow A$, *niin* $\mathbf{P}[A_n] \downarrow \mathbf{P}[A]$,
(iii) *jos* $A_n, A \in \mathcal{F}$, $n \in \mathbb{N}$, *ja* $A_n \uparrow A$, *niin* $\mathbf{P}[A_n] \uparrow \mathbf{P}[A]$.

Todistus. (ii) $\Leftrightarrow$ (iii): $A_n \downarrow A$ jos ja vain jos $A_n^c \uparrow A^c$. Väite seuraa siten De Morganin laeista ja funktion $\mathbf{P}$ additiivisuudesta.

(iii) $\Leftrightarrow$ (i): Olkoon $A_n \uparrow A$. Asetamme $B_n := A_n \setminus A_{n-1}$. Tällöin B_n:t ovat erillisiä ja $\cup_{k\le n} B_k = A_n$. Siten

$$\mathbf{P}[A_n] \;=\; \mathbf{P}\left[\bigcup_{k\le n} B_k\right] \;=\; \sum_{k\le n} \mathbf{P}[B_k].$$

Koska $A = \cup A_n = \cup B_n$, näemme että (iii) $\Leftrightarrow$ (i). □

Merkinnällä $(A_n) \subset \mathcal{F}$ tarkoitamme, että $A_n \in \mathcal{F}$ kaikilla $n \in \mathbb{N}$.

Seuraava tulos sanoo, että todennäköisyys on jatkuva.

2.2.6 Lause. *Olkoon* $(A_n) \subset \mathcal{F}$. *Jos* $A_n \to A$, *niin* $A \in \mathcal{F}$ *ja* $\mathbf{P}[A_n] \to \mathbf{P}[A]$.

Todistus. Osoitamme aluksi, että $A \in \mathcal{F}$. Koska $A_n \to A$, niin erityisesti $A = \liminf A_n$. Mutta $\liminf A_n$ on tapahtuma, sillä se on saatu numeroituvalla määrällä joukko-operaatioita tapahtumista A_n, $n \in \mathbb{N}$.

Osoitamme sitten jatkuvuuden. Olkoot $A_n^+ = \cap_{k\ge n} A_k$ ja $A_n^- = \cup_{k\ge n} A_k$. Tällöin (A_n^+) on kasvava ja (A_n^-) on vähenevä. Lisäksi $\liminf A_n = \lim A_n^+$ ja $\limsup A_n = \lim A_n^-$. Koska $A_n \to A$, niin $A_n^+ \uparrow A$ ja $A_n^- \downarrow A$. Huomaamalla, että $A_n^+ \subset A_n \subset A_n^-$ ja käyttämällä apulausetta 2.2.5 saamme

$$\mathbf{P}[A] \;=\; \lim \mathbf{P}[A_n^+] \;\le\; \lim \mathbf{P}[A_n] \;\le\; \lim \mathbf{P}[A_n^-] \;\le\; \mathbf{P}[A].$$

Mutta tämä tarkoittaa, että $\mathbf{P}[A_n] \to \mathbf{P}[A]$. □

Esitämme lopuksi vielä niin sanotun *inkluusio/ekskluusio-kaavan.* Tämä kaava on periaatteessa tuttu todennäköisyyslaskennan johdantokursseilta. Nyt vain esitämme sen laajemmassa ääretöntä joukkoperhettä koskevassa muodossa.

2.2.7 Lause. *Olkoon* $A_n \in \mathcal{F}$, $n \in \mathbb{N}$. *Tällöin*

$$\mathbf{P}\left[\bigcup_{n\ge 1} A_n\right] \;=\; \lim_{n\to\infty} \sum_{m\le n} (-1)^{m-1} \sum_{k_1<\cdots<k_m\le n} \mathbf{P}\left[\bigcap_{l\le m} A_{k_l}\right].$$

Todistus. Harjoitustehtävä 2.6. □

2.3 Laajennuslauseita

Usein mallinnuksessa perusjoukko Ω ja todennäköisyydet $\mathbf{P}[C]$, $C \in \mathcal{C}$, on annettu jollekin joukkoluokalle $\mathcal{C} \subset \operatorname{pot}(\Omega)$ (mieti loputonta kolikonheittoa ja sylinterijoukkoja). Nyt keskeisiä kysymyksiä ovat:

1. *yksikäsitteisyys* eli määräytyykö $\mathbf{P}$ yksikäsitteisesti σ-algebralle $\sigma(\mathcal{C})$,
2. *olemassaolo* eli onko olemassa todennäköisyyttä $\mathbf{P}$ kokoelmalta $\sigma(\mathcal{C})$, kun todennäköisyydet $\mathbf{P}[C]$, $C \in \mathcal{C}$, on kiinnitetty.

σ-algebrat ovat vaikeita. Algebrat, π-luokat ja d-luokat ovat helppoja.

2.3.1 Määritelmä. Kokoelma $\mathcal{A}$ joukon Ω osajoukkoja on *algebra*, jos

(i) $\Omega \in \mathcal{A}$,
(ii) jos $A \in \mathcal{A}$, niin $A^c \in \mathcal{A}$,
(iii) jos $A_1, A_2 \in \mathcal{A}$, niin $A_1 \cap A_2 \in \mathcal{A}$.

2.3.2 Määritelmä. Kokoelma $\mathcal{I}$ joukon Ω osajoukkoja on π*-luokka*, jos $I_1 \cap I_2 \in \mathcal{I}$, kun $I_1, I_2 \in \mathcal{I}$.

2.3.3 Määritelmä. Kokoelma $\mathcal{D}$ joukon Ω osajoukkoja on d*-luokka* eli *monotoninen luokka*, jos

(i) $\Omega \in \mathcal{D}$,
(ii) jos $A, B \in \mathcal{D}$ ja $A \subset B$, niin $B \setminus A \in \mathcal{D}$,
(iii) jos $A_n \in \mathcal{D}$ kaikilla $n \in \mathbb{N}$ ja $A_n \uparrow A$, niin $A \in \mathcal{D}$.

Algebra on siis suljettu äärellisten monien joukko-operaatioiden suhteen, π-luokka on suljettu äärellisten monien leikkausten suhteen ja d-luokka on suljettu numeroituvien monotonisten operaatioiden suhteen.

Kuten σ-algebrojen tapauksessa, myös algebrojen, π-luokkaen ja d-luokkaen tapauksessa voimme puhua virittäjistä (apulause 2.1.3).

2.3.4 Apulause. *Jos $\mathcal{C}_j$, $j \in J$, ovat algebroja (π-luokkia, d-luokkia), niin $\cap \mathcal{C}_j$ on algebra (vast. π-luokka, d-luokka).*

Todistus. Harjoitustehtävä 2.7. □

2.3.5 Määritelmä. Kokoelman $\mathcal{C}$ virittämät algebra $a(\mathcal{C})$, π-luokka $\pi(\mathcal{C})$ ja d-luokka $d(\mathcal{C})$ ovat

$$
\begin{aligned}
a(\mathcal{C}) &:= \bigcap \{\mathcal{A}\,;\, \mathcal{A} \text{ on algebra ja } \mathcal{A} \subset \mathcal{C}\}\,,\\
\pi(\mathcal{C}) &:= \bigcap \{\mathcal{I}\,;\, \mathcal{I} \text{ on } \pi\text{-luokka ja } \mathcal{I} \subset \mathcal{C}\}\,,\\
d(\mathcal{C}) &:= \bigcap \{\mathcal{D}\,;\, \mathcal{D} \text{ on } d\text{-luokka ja } \mathcal{D} \subset \mathcal{C}\}\,.
\end{aligned}
$$

2.3.6 Apulause. *Kokoelma $\mathcal{F}$ on σ-algebra jos ja vain jos se on sekä π-luokka että d-luokka.*

Todistus. On selvää, että σ-algebrat ovat sekä π- että d-luokkia.

Olkoon sitten $\mathcal{F}$ sekä π- että d-luokka. Olkoon $A \in \mathcal{F}$. Tällöin $A^c = \Omega \setminus A \in \mathcal{F}$, sillä $\mathcal{F}$ on d-luokka. Olkoon lisäksi $B \in \mathcal{F}$. Tällöin $A \cup B \in \mathcal{F}$ De Morganin lain $A \cup B = (A^c \cap B^c)^c$ nojalla. Olemme nyt siis osoittaneet, että $\mathcal{F}$ on algebra. Olkoon sitten $A_n \in \mathcal{F}$ kaikilla $n \in \mathbb{N}$. Merkitsemme $B_n := \cup_{k \le n} A_k$. Tällöin $B_n \uparrow \cup A_n$, joten $\cup A_n \in \mathcal{F}$. Käyttämällä De Morganin lakia $\cap A_n = (\cup A_n^c)^c$ näemme, että $\mathcal{F}$ on σ-algebra. □

Seuraava tulos on *Dynkinin lemma* tai *monotonisen luokan lause.*

2.3.7 Apulause. *Jos $\mathcal{I}$ on π-luokka, niin $d(\mathcal{I}) = \sigma(\mathcal{I})$.*

Todistus. Apulauseen 2.3.6 nojalla riittää osoittaa, että $d(\mathcal{I})$ on π-luokka.

Askel 1: Merkitsemme

$$\mathcal{D}_1 \ := \ \{B \in d(\mathcal{I})\,;\, B \cap I \in d(\mathcal{I}) \text{ kaikilla } I \in \mathcal{I}\}\,.$$

Selvästi $\mathcal{I} \subset \mathcal{D}_1$. Osoitamme, että $\mathcal{D}_1$ on d-luokka. Selvästi $\Omega \in \mathcal{D}_1$. Olkoon sitten $A, B \in \mathcal{D}_1$, $A \subset B$ ja $I \in \mathcal{I}$. Tällöin

$$\begin{aligned}(B \setminus A) \cap I &= B \cap I \cap A^c \\ &= (B \cap I) \cap (A^c \cap I^c) \\ &= (B \cap I) \setminus (A \cap I).\end{aligned}$$

Koska $(B \cap I) \in d(\mathcal{I})$ ja $(A \cap I) \in d(\mathcal{I})$, niin $(B \cap I) \setminus (A \cap I) \in d(\mathcal{I})$. Siten $B \setminus A \in \mathcal{D}_1$. Olkoon lopuksi $(A_n) \subset \mathcal{D}_1$ ja $A_n \uparrow A$. Tällöin kaikille $I \in \mathcal{I}$ pätee $A_n \cap I \in d(\mathcal{I})$ ja $A_n \cap I \uparrow A \cap I$. Näemme, että $A \cap I \in d(\mathcal{I})$. Siten $A \in \mathcal{D}_1$. Olemme osoittaneet, että $\mathcal{D}_1$ on d-luokka. Toisaalta, koska $\mathcal{I} \subset \mathcal{D}_1$, niin $d(\mathcal{I}) \subset \mathcal{D}_1$. Siten $d(\mathcal{I}) = \mathcal{D}_1$.

Askel 2: Merkitsemme

$$\mathcal{D}_2 \ := \ \{A \in d(\mathcal{I})\,;\, A \cap B \in d(\mathcal{I}) \text{ kaikilla } B \in d(\mathcal{I})\}\,.$$

Askeleen 1 nojalla $\mathcal{D}_1 = d(\mathcal{I})$. Siten $\mathcal{I} \subset \mathcal{D}_2$. Toisaalta, kuten askeleessa 1, on helppo nähdä, että $\mathcal{D}_2$ on d-luokka. Lisäksi $\mathcal{I} \subset \mathcal{D}_2 \subset \mathcal{D}_1 \subset d(\mathcal{I})$. Siten $\mathcal{D}_2 = d(\mathcal{I})$. Lopulta väite seuraa huomaamalla, että $\mathcal{D}_2$ on selvästi π-luokka. □

Seuraava tulos on *Dynkinin laajennuslause.* Se takaa todennäköisyysmitan laajennuksen yksikäsitteisyyden.

2.3.8 Lause. *Olkoon* $\mathcal{I}$ π*-luokka ja* $\mathcal{F} = \sigma(\mathcal{I})$. *Olkoot* $\mathbf{P}$ *ja* $\mathbf{P}'$ *kaksi todennäköisyysmittaa mitallisella avaruudella* $(\Omega, \mathcal{F})$. *Jos* $\mathbf{P} = \mathbf{P}'$ π*-luokalla* $\mathcal{I}$, *niin* $\mathbf{P} = \mathbf{P}'$ σ*-algebralla* $\mathcal{F}$.

Todistus. Merkitsemme

$$\mathcal{D} \;:=\; \{A \in \mathcal{F};\, \mathbf{P}[A] = \mathbf{P}'[A]\}.$$

Osoitamme, että $\mathcal{D}$ on d-luokka. Selvästi $\Omega \in \mathcal{D}$. Olkoon sitten $A, B \in \mathcal{D}$ ja $A \subset B$. Tällöin

$$\mathbf{P}[B \setminus A] \;=\; \mathbf{P}[B] - \mathbf{P}[A] \;=\; \mathbf{P}'[B] - \mathbf{P}'[A] \;=\; \mathbf{P}'[B \setminus A].$$

Siten $B \setminus A \in \mathcal{D}$. Olkoon lopuksi $(A_n) \subset \mathcal{D}$, $A_n \uparrow A$. Todennäköisyyden jatkuvuuden nojalla

$$\mathbf{P}[A] \;=\; \lim \mathbf{P}[A_n] \;=\; \lim \mathbf{P}'[A_n] \;=\; \mathbf{P}'[A].$$

Siten $A \in \mathcal{D}$.

Nyt $\mathcal{D}$ on d-luokka ja $\mathcal{I} \subset \mathcal{D}$ oletuksen nojalla. Siten, Dynkinin lemman 2.3.7 nojalla, $\mathcal{D} \supset d(\mathcal{I}) = \sigma(\mathcal{I}) = \mathcal{F}$. □

Dynkinin laajennuslause 2.3.8 takasi ainoastaan todennäköisyysmitan laajennuksen yksikäsitteisyyden. Seuraava laajennuksen olemassaolo- ja yksikäsitteisyystulos on *Carathéodoryn laajennuslause.*

2.3.9 Lause. *Olkoon* $\mathcal{A}$ *algebra ja* $\mathbf{P}_0 : \mathcal{A} \to [0,1]$ σ*-additiivinen joukkofunktio, jolle* $\mathbf{P}_0[\Omega] = 1$. *Tällöin on olemassa yksikäsitteinen todennäköisyysmitta* $\mathbf{P} : \sigma(\mathcal{A}) \to [0,1]$, *jolle* $\mathbf{P}[A] = \mathbf{P}_0[A]$, *kun* $A \in \mathcal{A}$.

Todistus. Harjoitustehtävä (esimerkiksi graduksi). □

2.3.10 Seuraus. *Mitalliselle avaruudelle* $(\mathbb{R}, \mathcal{B})$ *on olemassa* Lebesguen mitta ℓ, *joka määräytyy yksikäsitteisesti ehdosta* $\ell[(a,b)] = |b-a|$ *kaikilla* $a, b \in \mathbb{R}$.

Todistus. Tulos seuraisi välittömästi Carathéodoryn laajennuslauseesta, jos vain Lebesguen mitta olisi todennäköisyysmitta. Voimme kuitenkin määritellä Lebesguen mitan tasaisten todennäköisyysmittojen avulla seuraavasti. Olkoon $m \in \mathbb{Z}$. Jokaiselle välille $[m, m+1]$ asetamme

$$\ell_m[(a,b)] \;:=\; |b-a|,$$

kun $a, b \in [m, m+1]$. Välin $[m, m+1]$ Borelin σ-algebra on niiden joukkojen B_m luokka, joille $B_m \in \mathcal{B}$ ja $B_m \subset [m, m+1]$. Nyt, Carathéodoryn laajennuslauseen 2.3.9 nojalla, mitat ℓ_m laajenevat välien $[m, m+1]$ Borelin σ-algebroille. Toisaalta, jos $B \in \mathcal{B}$, niin $B = \cup B_m$, missä B_m on välin $[m, m+1]$ Borel-joukko. Siten voimme asettaa

$$\ell[B] \ := \ \sum \ell_m[B_m].$$

Tämä ℓ on Legesguen mitta. $\square$

Esitääme vielä lopuksi todennäköisyysavaruuden täydellistymän, vaikkakaan emme tarvitse sitä myöhemmin kurssilla.

Olkoon $(\Omega, \mathcal{F}, \mathbf{P})$ todennäköisyysavaruus. Olkoon $\mathcal{N}$ *nollaluokka*:

> $N \in \mathcal{N}$ jos ja vain jos on olemassa sellainen $A \in \mathcal{F}$, että $N \subset A$ ja $\mathbf{P}[A] = 0$.

On "filosofisesti tyydyttävää" voida sanoa, että $\mathbf{P}[N] = 0$, kun $N \in \mathcal{N}$. Matemaattisesti tämä tehdään tarkastelemalla joukkoluokkaa $\mathcal{F}^*$, joka koostuu alkioista A^*, joille on olemassa sellaiset $A^-, A^+ \in \mathcal{F}$, että $A^- \subset A^* \subset A^+$ ja $\mathbf{P}[A^+ \setminus A^-] = 0$. Tällöin $\mathcal{F}^*$ on σ-algebra. Itse asiassa voidaan osoittaa, että $\mathcal{F}^* = \sigma(\mathcal{F}, \mathcal{N})$. Asetamme sitten $\mathbf{P}^*[A^*] = \mathbf{P}[A^+] = \mathbf{P}[A^-]$. Todennäköisyysavaruus $(\Omega, \mathcal{F}^*, \mathbf{P}^*)$ on todennäköisyysavaruuden $(\Omega, \mathcal{F}, \mathbf{P})$ *täydellistymä*.

2.4 Mitallisuus ja mitattomuus

Tämä osio koostuu esimerkistä, joka osoittaa, ettemme voi aina valita kaikkia joukkoja mitallisiksi. Toisin sanoen joudumme joskus tyytymään siihen, että σ-algebra $\mathcal{F}$ on potenssijoukon $\mathrm{pot}\,(\Omega)$ aito osajoukko.

2.4.1 Esimerkki. Olkoon $\Omega = [0,1]$ ja $\mathbf{P}$ tasainen jakauma eli Lebesguen mitta. Tällöin siis

$$\mathbf{P}[(a,b)] \ = \ |b-a|,$$

kun $a, b \in \Omega$. Osoitamme, että kaikki Ω:n osajoukot eivät voi olla mitallisia.

Olkoot $x, y \in [0,1]$. Jatkossa $x+y$ tarkoittaa summaa $x+y$, jos $x+y \le 1$ ja summaa $x+y-1$ muulloin. Erotus $x-y$ määritellään vastaavasti

Olkoon $B \subset \Omega$ ja $\omega \in \Omega$. Merkitsemme

$$B + \omega \ := \ \{\omega' \in \Omega\,;\, \omega' = \omega + b, b \in B\}.$$

Huomaamme, että $\mathbf{P}[B] = \mathbf{P}[B+\omega]$ kaikilla $\omega \in \Omega$ eli $\mathbf{P}$ on siirtoinvariantti. Lisäksi B on mitallinen jos ja vain jos $B + \omega$ on mitallinen kaikilla $\omega \in \Omega$.

Määrittelemme sitten pisteille $\omega, \omega' \in \Omega$ relaation $\sim$ asettamalla $\omega \sim \omega'$ jos ja vain jos $\omega - \omega' \in \mathbb{Q}$. Selvästi $\sim$ on ekvivalenssirelaatio. Siten $\sim$

jakaa avaruuden Ω erillisiin ekvivalenssiluokkiin. Olkoon sitten A joukko, johon on valittu jokaisesta näistä ekvivalenssiluokista yksi alkio. Tämä joukko on olemassa valinta-aksiooman nojalla. Osoitamme, että tämä A ei voi olla mitallinen.

Olkoon $q_1, q_2, \ldots$ joukon Ω rationaaliluvut. Tarkastelemme joukkoja $A_n := A + q_n$. Nämä joukot muodostavat perusjoukon Ω *osituksen* eli A_n:t ovat erillisiä ja $\Omega = \cup A_n$ (harjoitustehtävä 2.8).

Nyt, edellisen perusteella, jos A on mitallinen, niin jokainen joukoista A_n, $n \in \mathbb{N}$ on mitallinen. Lisäksi, koska joukot A_n, $n \in \mathbb{N}$, muodostavat avaruuden Ω osituksen, niin

$$1 = \mathbf{P}[\Omega] = \mathbf{P}\Big[\bigcup A_n\Big] = \sum \mathbf{P}[A_n]. \tag{2.4.2}$$

Mutta mitta $\mathbf{P}$ on siirtoinvariantti. Siten $\mathbf{P}[A_n]$ on vakio. Siten yhtälö (2.4.2) on mahdoton. Olemme päätyneet ristiriitaan. Oletus, joka tulee hylätä on, että A on mitallinen.

Harjoitustehtäviä

2.1. Olkoon $\Omega = \{1, 2, \ldots, n\}$. Kuinka monta erilaista σ-algebraa joukolla Ω on?

2.2. Olkoon $\mathcal{F}$ jokin ääretön σ-algebra. Osoita, että $\mathcal{F}$ ei voi olla numeroituva.

2.3. Todista lemman 2.1.7 kohdat (i)–(iv) ja (vi)–(x).

2.4. Osoita, että joukot $\times_{k \le d}[a_k, \infty)$, $a_k \in \mathbb{Q}$, virittävät $\mathbb{R}^d$:n Borelin σ-algebran $\mathcal{B}_d$.

2.5. Olkoot $A_n \in \mathcal{F}$, $n \in \mathbb{N}$. Osoita, että

$$\mathbf{P}\Big[\bigcup A_n\Big] \le \sum \mathbf{P}[A_n].$$

2.6. Todista lause 2.2.7.

2.7. Todista apulause 2.3.4.

2.8. Osoita, että esimerkin 2.4.1 joukot A_n, $n \in \mathbb{N}$, muodostavat avaruuden Ω osituksen: A_n:t ovat erillisiä ja $\cup A_n = \Omega$.

Luku 3

Satunnaismuuttujat ja -vektorit

Satunnaiskuvauksen, tai -muuttujan, yleinen idea on seuraava. Todennäköisyysavaruus $(\Omega, \mathcal{F}, \mathbf{P})$ *on abstraktio, jonka ei tarvitse liittyä millään ilmeisellä tavalla havaittavaan todellisuuteen. Todelliset havainnot liittyvät satunnaiskokeeseen, jossa havaitsemme satunnaiskuvauksen* $X : \Omega \to$ "jokin avaruus". *Havaitsemme siis tuloksen* $X(\omega)$, *mutta alkeistapaus* $\omega \in \Omega$ *jää kohtalon jumalattaren salaisuudeksi. Näin ollen päätelmämme todennäköisyyksistä eivät perustu abstraktiin todennäköisyysmittaan* $\mathbf{P}$, *vaan havaittavaan todennäköisyysmittaan* $\mathbf{P}_X[\,\cdot\,] := \mathbf{P}[X \in \,\cdot\,] := \mathbf{P} \circ X^{-1}[\,\cdot\,]$. *Todennäköisyysmittaa* $\mathbf{P}_X$ *kutsumme satunnaiskuvauksen* X *jakaumaksi.*

Yleisesti satunnaiskuvaus X *on mitallinen kuvaus mitalliselta perusavaruudelta* $(\Omega, \mathcal{F})$ *jollekin mitaliselle avaruudelle* $(\mathbb{M}, \mathcal{M})$. *Mitallisuus tarkoittaa sitä, että* $\{X \in M\} \in \mathcal{F}$ *kaikilla* $M \in \mathcal{M}$. *Emme kuitenkaan käsittele satunnaiskuvauksia tässä yleisyydessä, sillä se vaatisi lukijalta huomattavaa topologian ja funktionaalianalyysin osaamista. Tyydymme käsittelemään tilanteita* $(\mathbb{M}, \mathcal{M}) = (\mathbb{R}^d, \mathcal{B}_d)$ *ja* $(\mathbb{M}, \mathcal{M}) = (\mathbb{R}, \mathcal{B})$. *Ensimmäisessä tapauksessa kutsumme satunnaiskuvauksia satunnaisvektoreiksi ja jälkimmäisessä tapauksessa satunnaismuuttujiksi.*

3.1 Määritelmä mitallisena kuvauksena

3.1.1 Määritelmä. Kuvaus $X : \Omega \to \mathbb{R}$ on *satunnaismuuttuja* eli *mitallinen kuvaus*, jos $\{X \in B\} \in \mathcal{F}$ kaikilla $B \in \mathcal{B}$. Vastaavasti vektoriarvoinen kuvaus $X = (X_1, \ldots, X_d) : \Omega \to \mathbb{R}^d$ on *satunnaisvektori*, jos $\{X \in B\} \in \mathcal{F}$ kaikilla $B \in \mathcal{B}_d$.

Itse asiassa $(X_1, \ldots, X_d)$ on satunnaisvektori jos ja vain jos sen komponentit $X_1, \ldots, X_d$ ovat satunnaismuuttujia (harjoitustehtävä 3.1).

Se, onko kuvaus X perusavaruudelta Ω satunnaismuuttuja, tai -vektori, riippuu σ-algebran $\mathcal{F}$ rikkaudesta: mitä rikkaampi $\mathcal{F}$ on sitä enemmän

$(\Omega, \mathcal{F})$:ltä on mitallisia kuvauksia. Seuraava esimerkki valaisee tätä ilmiötä.

3.1.2 Esimerkki. Tarkastelemme satunnaismuuttujia $X : \Omega \to \mathbb{R}$.

(i) Jos $\mathcal{F} = \operatorname{pot}(\Omega)$, niin kaikki kuvaukset ovat satunnaismuuttujia.
(ii) Jos $\mathcal{F} = \{\Omega, \varnothing\}$, niin vain vakiokuvaukset ovat satunnaismuuttujia.
(iii) Olkoon $\Omega = \{1, 2, 3, 4, 5, 6\}$ ja X kuusisivuisen nopan heiton tulos. Jos $\mathcal{F} = \operatorname{pot}(\Omega)$, niin X on satunnaismuuttuja avaruudelta $(\Omega, \mathcal{F})$. Jos taas $\mathcal{G} = \{\Omega, \{1, 3, 5\}, \{2, 4, 6\}, \varnothing\}$, niin X ei ole satunnaismuuttuja avaruudelta $(\Omega, \mathcal{G})$. Sen sijaan $Y :=$ "silmäluku on parillinen" on satunnaismuuttuja avaruudelta $(\Omega, \mathcal{G})$.

3.1.3 Lause. *Olkoon $\mathcal{B}_d = \sigma(\mathcal{C})$. Tällöin $X = (X_1, \ldots, X_d)$ on satunnaisvektori jos ja vain jos $\{X \in C\} \in \mathcal{F}$ kaikilla $C \in \mathcal{C}$.*

Todistus. Jos X on satunnaisvektori, niin $\{X \in C\} \in \mathcal{F}$ kaikilla $C \in \mathcal{C}$, sillä $\mathcal{C} \subset \mathcal{B}_d$.

Oletamme sitten kääntäen, että $\{X \in C\} \in \mathcal{F}$ kaikilla $C \in \mathcal{C}$ pätee. Olkoon $\mathcal{B}'_d$ niiden $\mathbb{R}^d$:n joukkojen B kokoelma, joille $\{X \in B\} \in \mathcal{F}$. Olkoon $(B_n) \subset \mathcal{B}'_d$. Koska $\{X \in \cap B_n\} = \cap\{X \in B_n\}$ ja $\{X \in B_n^c\} = \{X \in B_n\}^c$, niin näemme, että $\mathcal{B}'_d$ on σ-algebra. Nyt, oletuksen nojalla $\mathcal{C} \subset \mathcal{B}'_d$. Mutta koska $\mathcal{C}$ virittää σ-algebran $\mathcal{B}_d$, niin $\mathcal{B}_d \subset \mathcal{B}'_d$. $\square$

3.1.4 Seuraus. *$X = (X_1, \ldots, X_d)$ on satunnaisvektrori jos ja vain jos $\{X_1 \le q_1, \ldots, X_d \le q_d\} \in \mathcal{F}$ kaikilla $q_1, \ldots, q_d \in \mathbb{Q}$.*

Todistus. Tulos seuraa lauseesta 3.1.3, jos joukot $\times_{k \le d}(-\infty, q_k]$, $q_k \in \mathbb{Q}$, virittävät σ-algebran $\mathcal{B}_d$. Mutta tämä tulos taas seuraa harjoitustehtävästä 2.4. $\square$

3.2 Rajat ja muunnokset

3.2.1 Lause. *Olkoot X_n, $n \in \mathbb{N}$, satunnaismuuttujia. Tällöin myös* $\liminf X_n$ *ja* $\limsup X_n$ *ovat satunnaismuuttujia.*

Todistus. Koska $\{\sup X_n \le a\} = \cap\{X_n \le a\}$ (harjoitustehtävä 3.2) ja $\{X_n \le a\} \in \mathcal{F}$ kaikilla $a \in \mathbb{R}$, on $\sup X_n$ satunnaismuuttuja. Samalla tavalla näemme, että myös $\inf X_n$ on satunnaismuuttuja. Mutta $\limsup X_n = \inf_{n \ge 1} \sup_{k \ge n} X_k$. Siten $\limsup X_n$ on satunnaismuuttuja. Samalla tavalla näemme, että myös $\liminf X_n$ on satunnaismuuttuja. $\square$

3.2.2 Seuraus. *Olkoot X_n, $n \in \mathbb{N}$, satunnaismuuttujia. Jos $X_n \to X$, niin X on satunnaismuuttuja.*

Kuvaus $f : \mathbb{R}^d \to \mathbb{R}^{d'}$ on *Borel-mitallinen*, jos $\{f \in B\} \in \mathcal{B}_d$ kun $B \in \mathcal{B}_{d'}$.

3.2.3 Lause. *Olkoon $X = (X_1, \ldots, X_d)$ satunnaisvektori ja $f : \mathbb{R}^d \to \mathbb{R}^{d'}$ Borel-mitallinen. Tällöin $f(X)$ on satunnaisvektori.*

Todistus. Olkoon $B \in \mathcal{B}_{d'}$. Nyt $\{f(X) \in B\} = \{X \in f^{-1}B\}$. Koska f on Borel-mitallinen, niin $B_f := f^{-1}B = \{f \in B\} \in \mathcal{B}_d$ ja koska X on satunnaisvektori, niin $\{f(X) \in B\} = \{X \in B_f\} \in \mathcal{F}$. □

3.2.4 Seuraus. *Olkoon (X, Y) satunnaisvektori. Tällöin $X + Y$, XY, $\max(X, Y)$ ja $\min(X, Y)$ ovat satunnaismuuttujia. Lisäksi, jos $Y \neq 0$, niin X/Y on satunnaismuuttuja.*

Todistus. Jatkuvat kuvaukset ovat Borel-mitallisia (harjoitustehtävä 3.6). Mutta kuvaukset $(x, y) \mapsto x + y$, $(x, y) \mapsto xy$, $(x, y) \mapsto \max(x, y)$, $(x, y) \mapsto \min(x, y)$ ja $(x, y) \mapsto x/y$, $y \neq 0$ ovat jatkuvia. Tulos seuraa siten lauseesta 3.2.3. □

3.3 Jakauma

3.3.1 Määritelmä. Satunnaisvektorin $X = (X_1, \ldots, X_d)$ *jakauma* $\mathbf{P}_X$ on sen maalijoukkoonsa $\mathbb{R}^d$ indusoima todennäköisyysmitta

$$\mathbf{P}_X[B] \;:=\; \mathbf{P}[X \in B], \qquad B \in \mathcal{B}_d.$$

3.3.2 Apulause. *Olkoon $X = (X_1, \ldots, X_d)$ satunnaisvektori. Tällöin $\mathbf{P}_X$ on todennäköisyysmitta mitallisella avaruudella $(\mathbb{R}^d, \mathcal{B}_d)$.*

Todistus. Harjoitustehtävä 3.7. □

3.3.3 Esimerkki. Olkoon $B \in \mathcal{B}$.

(i) Jos X on deterministinen satunnaismuuttuja eli $\mathbf{P}[X = x_0] = 1$, niin sen jakauma on *Diracin pistemassa* pisteessä x_0:

$$\mathbf{P}_X[B] \;=\; \delta_{x_0}[B] \;:=\; \begin{cases} 1, & \text{jos } x_0 \in B, \\ 0, & \text{muulloin.} \end{cases}$$

(ii) Olkoon X symmetrisen N-sivuisen nopan heiton tulos. Sen jakauma voidaan esittää Diracin pistemassojen avulla:

$$\mathbf{P}_X[B] \;=\; \frac{1}{N} \sum_{k \le N} \delta_k[B].$$

Tämä tarkoittaa, että X on jakautunut tasaisesti joukkoon $\{1, \ldots, N\}$.

(iii) Satunnaismuuttuja on *diskreetti*, jos sen arvojoukko on numeroituva. Minkä tahansa diskreetin satunnaismuuttujan jakauma voidaan esittää Diracin pistemassojen : jos satunnaismuuttujan X arvojoukko on $\{a_1, a_2, \ldots\}$ ja $p_k := \mathbf{P}[X = a_k]$, niin

$$\mathbf{P}_X[B] \;=\; \sum p_k \delta_{a_k}[B]. \tag{3.3.4}$$

Itse asiassa satunnaismuuttuja X on diskreetti jos ja vain jos sen jakaumalla on esitys 3.3.4.

(iv) Satunnaismuuttuja X on *jatkuva*, jos sen jakaumalla on esitys

$$\mathbf{P}_X[B] \;=\; \int_B f_X(x)\,\mathrm{d}x$$

kaikilla $B \in \mathcal{B}$. Funktio $f : \mathbb{R} \to \mathbb{R}_+$ on satunnaismuuttujan X *tiheysfunktio*. Vastaavasti satunnaisvektori $X = (X_1, \ldots, X_d)$ on jatkuva, jos

$$\mathbf{P}_X[B] \;=\; \int \cdots \int_B f(x_1, \ldots, x_d)\,\mathrm{d}x_1 \cdots \mathrm{d}x_d$$

kaikilla $B \in \mathcal{B}_d$. Funktio $f : \mathbb{R} \to \mathbb{R}_+^d$ on satunnaisvektorin X *yhteistiheysfunktio*.

3.3.5 Lause. *Satunnaisvektorin $X = (X_1, \ldots, X_d)$ jakauman määrää sen* yhteiskertymäfunktio

$$F_X(x_1, \ldots, x_d) \;:=\; \mathbf{P}_X\left[\bigtimes_{k \le d} (-\infty, x_k]\right] \;=\; \mathbf{P}_X[X_1 \le x_1, \ldots, X_n \le x_d].$$

Todistus. Huomaamme, että

$$\mathcal{I} \;=\; \left\{\bigtimes_{k \le d} (-\infty, x_k]\,;\, x = (x_1, \ldots, x_d) \in \mathbb{R}^d\right\}$$

on π-luokka ja $\mathcal{B}_d = \sigma(\mathcal{I})$. Siten väite seuraa Dynkinin laajennuslauseesta 2.3.8. □

Seuraava lause karakterisoi ne funktiot $F : \mathbb{R} \to \mathbb{R}$, jotka ovat jonkin satunnaismuuttujan kertymäfunktioita. Lauseesta on olemassa myös satunnaisvektoreja koskeva version, jonka muotoilun ja todistamisen jätämme harjoitustehtäväksi 3.8.

3.3.6 Lause. *Funktio $F : \mathbb{R} \to \mathbb{R}$ on jonkin satunnaismuuttujan X* kertymäfunktio *eli $F(x) = \mathbf{P}[X \le x]$ jos ja vain jos*

(i) *F on kasvava: $F(x) \le F(y)$ kaikilla $x \le y$,*
(ii) *F on oikealta jatkuva: $F(x+) := \lim_{y \to x+} F(y) = F(x)$,*
(iii) *$\lim_{x\to\infty} F(x) = 1$ ja $\lim_{x\to -\infty} F(x) = 0$.*

Todistus. Tämä todistus on varsin ylimalkainen, toisin sanoen yksityiskohdissa olisi paljon täyttelemistä. Jätämme epämääräisyyksien huomaamisen ja korjaamisen ylimääräiseksi harjoitustehtäväksi.

Osoitamme aluksi, että kertymäfunktio toteuttaa ehdot (i)–(iii).

Olkoon X satunnaismuuttuja, jonka kertymäfunktio on F.

Ominaisuus (i) seuraa todennäköisyyden monotonisuudesta: jos $x \le y$, niin $\{X \le x\} \subset \{X \le y\}$ ja siten

$$F(x) \;=\; \mathbf{P}[X \le x] \;\le\; \mathbf{P}[X \le y] \;=\; F(y).$$

Ominaisuus (ii) seuraa todennäköisyyden jatkuvuudesta: jos $y \to x+$, niin $\{X \le y\} \to \{X \le x\}$ ja siten

$$F(x+) \;=\; \lim_{y \to x+} \mathbf{P}[X \le y] \;=\; \mathbf{P}[X \le x] \;=\; F(x).$$

(Edellä on huomattavaa, että yleisesti $F(x-) := \lim_{y \to x-} F(y) \ne F(x)$. Itse asiassa $\Delta F(x) := F(x) - F(x-) = \mathbf{P}[X = x]$.)

Myös ominaisuus (iii) seuraa todennäköisyyden jatkuvuudesta. Nimittäin jos $x \to \infty$, niin $\{X \le x\} \to \{X \le \infty\} = \Omega$ ja jos $x \to -\infty$, niin $\{X \le x\} \to \{X \le -\infty\} = \varnothing$.

Rakennamme sitten F-jakautuneen satunnaismuuttujan X tasaisesti jakautuneen satunnaismuuttujan U avulla.

Tarkastelemme "kanonista" todennäköisyysavaruutta $([0,1], \mathcal{B}_{[0,1]}, \ell)$, missä $\mathcal{B}_{[0,1]}$ on välin $[0,1]$ avointen joukkojen virittämä σ-algebra ja ℓ on Lebesguen mitta rajoitettuna välille $[0,1]$. Olkoon $U(x) = x$. Tällöin U on todennäköisyyden ℓ suhteen tasaisesti jakautunut. Koska F toteuttaa ehdot (i)–(iii), niin voimme määritellä $X := F^{\leftarrow}(U)$, missä

$$F^{\leftarrow}(x) \;:=\; \inf\{y \in [0,1]\,;\, F(y) > x\}.$$

Koska $F^{\leftarrow}$ on kasvava, se on Borel-funktio. Siten $F^{\leftarrow}(U)$ on satunnaismuuttuja. Lisäksi

$$\begin{aligned}
\mathbf{P}_X\big[(-\infty, x]\big] &= \ell\big[F^{\leftarrow}(U) \le x\big] \\
&= \ell[U \le F(x)] \\
&= F(x).
\end{aligned}$$

Siispä F on satunnaismuuttujan $X = F^{\leftarrow}(U)$ kertymäfunktio. □

Harjoitustehtäviä

3.1. Osoita, että $(X_1, \dots, X_d)$ on satunnaisvektori jos ja vain jos sen komponentit $X_1, \dots, X_d$ ovat satunnaismuuttujia.

3.2. Olkoon $a \in \mathbb{R}$ ja X_n, $n \in \mathbb{N}$, satunnaismuuttujia. Osoita, että $\{\sup X_n \leq a\} = \cap\{X_n \leq a\}$.

3.3. Olkoon $a \in \mathbb{R}$ ja X_n, $n \in \mathbb{N}$, satunnaismuuttujia. Mikä relaatioista $=$, $\subset$, $\supset$ pätee seuraavien tapahtumaparien välillä:

(a) $\{\inf X_n \leq a\}$ ja $\cup\{X_n \leq a\}$,
(b) $\{\limsup X_n \leq a\}$ ja $\{\liminf X_n \leq a\}$?

3.4. Olkoon X satunnaismuuttuja ja $A \in \mathcal{F}$. Osoita, että $X\mathbf{1}_A$ on satunnaismuuttuja.

3.5. Olkoon X satunnaismuuttuja ja $A \notin \mathcal{F}$. Onko $X\mathbf{1}_A$ satunnaismuuttuja?

3.6. Osoita, että jatkuvat kuvaukset ovat Borel-mitallisia.

3.7. Todista apulause 3.3.2.

3.8. Muotoile ja todista lauseen 3.3.6 satunnaisvektoreja vastaava versio.

Luku 4

Odotusarvo

Odotusarvon käsite voidaan heuristisesti johtaa monellakin eri tavalla.

Fyysikko ajattelee todennäköisyysjakaumaa todennäköisyysmassana reaaliakselin päällä. Hänelle odotusarvo on tämän jakauman massakeskipiste. Se on siis se piste, josta tuettuna todennäköisyysmassalla painotettu reaaliakseli pysyy tasapainossa.

Uhkapeluri taas ajattelee odotusarvoa hinnaksi, joka kannattaa maksaa uhkapeliin X osallistumisesta. Tämä näkemys perustuu suurten lukujen lakiin 8.1.9. Jos nimittäin perättäiset uhkapelit $X_1, X_2, \ldots$ ovat riippumattomia satunnaismuuttujia, jotka ovat jakautuneet kuten X, niin tällöin summa $(X_1 + \cdots + X_n)/n$ suppenee kohti satunnaismuuttujan X odotusarvoa. Uhkapelurin kielellä tämä tarkoittaa, että uhkapelissä X voittaa lopulta jos ja vain jos peliin osallistumisen hinta on korkeintaan pelin X odotusarvo.

Mittateoreetikolle satunnaismuuttujan X odotusarvo on mitallisen kuvauksen integraali mitan $\mathbf{P}$ suhteen: $\int_\Omega X \mathrm{d}\mathbf{P}$.

4.1 Määritelmä mittaintegraalina

4.1.1 Määritelmä. Satunnaismuuttuja X on *yksinkertainen*, jos

$$X = \sum_{k \le n} a_k \mathbf{1}_{A_k}, \tag{4.1.2}$$

joillakin $a_k \in \mathbb{R}$ ja $A_k \in \mathcal{F}$. Jos $A_k = \{X = a_k\}$, on esitys (4.1.2) *minimaalinen*. Yksinkertaisten satunnaismuuttujien luokkaa merkitsemme $\mathcal{E}$:llä.

Minimaalisessa esityksessä joukot $\{X = a_k\}$, $k = 1, \ldots, n$, muodostavat perusjoukon Ω *osituksen*: $\{X = a_k\}$:t ovat erillisiä ja $\cup\{X = a_k\} = \Omega$.

4.1.3 Esimerkki. Olkoon X klaavojen lukumäärä n heiton sarjassa. Tällöin X on yksinkertainen ja voimme esittää

$$X \;=\; \sum_{k\leq n} \mathbf{1}_{A_k},$$

missä A_k on tapahtuma "k. heitto on klaava". Tämä esitys ei ole minimaalinen. Minimaalinen esitys on

$$X \;=\; \sum_{0\leq k\leq n} k\mathbf{1}_{\{X=k\}}.$$

4.1.4 Apulause. *$\mathcal{E}$ on* vektoriavaruus*: jos $X, Y \in \mathcal{E}$ ja $a, b \in \mathbb{R}$, niin $aX + bY \in \mathcal{E}$. Lisäksi $\mathcal{E}$ on* vektorialgebra*: jos $X, Y \in \mathcal{E}$, niin $XY \in \mathcal{E}$.*

Todistus. Tulos seuraa huomaamalla, että $X \in \mathcal{E}$ jos ja vain jos sen arvojoukko $X\Omega$ on äärellinen. □

4.1.5 Määritelmä. Olkoon $X \in \mathcal{E}$ esityksenään (4.1.2). Sen *odotusarvo* on

$$\mathbf{E}[X] \;:=\; \sum_{k\leq n} a_k\mathbf{P}[A_k].$$

Odotusarvo on hyvin määritelty eli se ei riipu esityksestä (4.1.2). Tämä on helppo nähdä myöhemmin, kun huomaamme, että odotusarvo on lineaarinen operaattori (apulause 4.1.7). Erityisesti on hyvä huomata, että

$$\mathbf{E}[\mathbf{1}_A] \;=\; \mathbf{P}[A].$$

Lisäksi minimaalisen esityksen avulla huomaamme, että

$$\mathbf{E}[X] \;=\; \sum_{k\leq n} a_k\mathbf{P}[X=a_k].$$

4.1.6 Esimerkki. (i) Olkoon X tasaisesti jakautunut joukkoon $\{1,\ldots,N\}$ eli $\mathbf{P}[X=k] = \frac{1}{N}$, kun $k \in \{1,\ldots,N\}$ ja 0 muulloin. Tällöin

$$\mathbf{E}[X] \;=\; \sum_{k\leq N} k\mathbf{P}[X=k] \;=\; \frac{1}{N}\sum_{k\leq N} k \;=\; \frac{N+1}{2}.$$

(ii) Olkoon X binomijakautunut parametrein n ja p eli

$$\mathbf{P}[X=k] \;=\; \binom{n}{k}p^k q^{n-k}$$

($q := 1 - p$), kun $k \in \{0, \dots, n\}$ ja 0 muulloin. Tällöin

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{k \le n} k \binom{n}{k} p^k q^{n-k} \\
&= \sum_{k \le n} k \frac{n!}{k!(n-k)!} p^k q^{n-k} \\
&= np \sum_{k \le n} \frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} p^{k-1} q^{(n-1)-(k-1)} \\
&= np \sum_{k' \le n'} \binom{n'}{k'} p^{k'} q^{n'-k'} \\
&= np.
\end{aligned}$$

Yllä teimme muuttujanvaihdot $n' := n - 1$ ja $k' := k - 1$.

4.1.7 Apulause. $\mathbf{E} : \mathcal{E} \to \mathbb{R}$ *on positiivinen lineaarinen operaattori:*

(i) *jos* $X, Y \in \mathcal{E}$ *ja* $a, b \in \mathbb{R}$, *niin* $\mathbf{E}[aX + bY] = a\mathbf{E}[X] + b\mathbf{E}[Y]$,

(ii) *jos* $X \in \mathcal{E}$ *ja* $X \ge 0$, *niin* $\mathbf{E}[X] \ge 0$.

Todistus. (i) Olkoon

$$X = \sum_{k \le n} a_k \mathbf{1}_{A_k} \quad \text{ja} \quad Y = \sum_{l \le m} b_l \mathbf{1}_{B_l},$$

missä A_k:t ja B_l:t muodostavat Ω:n osituksen (voimme tarkastella esimerkiksi minimaalisia esityksiä). Tällöin

$$aX + bY = \sum_{k \le n} \sum_{l \le m} (aa_k + bb_l) \mathbf{1}_{A_k \cap B_l}$$

ja

$$\begin{aligned}
\mathbf{E}[aX + bY] &= \sum_{k \le n} \sum_{l \le m} (aa_k + bb_l) \mathbf{P}[A_k \cap B_l] \\
&= \sum_{k \le n} \sum_{l \le m} aa_k \mathbf{P}[A_k \cap B_l] + \sum_{k \le n} \sum_{l \le m} bb_l \mathbf{P}[A_k \cap B_l] \\
&= a \sum_{k \le n} a_k \sum_{l=1}^{n} \mathbf{P}[A_k \cap B_l] + b \sum_{l \le m} b_l \sum_{k \le n} \mathbf{P}[A_k \cap B_l] \\
&= a \sum_{k \le n} a_k \mathbf{P}[A_k] + b \sum_{l \le m} b_l \mathbf{P}[B_l] \\
&= a\mathbf{E}[X] + b\mathbf{E}[Y].
\end{aligned}$$

(ii) Tämä väite on triviaali. $\square$

4.1.8 Seuraus. *Olkoon* $X, Y \in \mathcal{E}$*. Jos* $X \leq Y$*, niin* $\mathbf{E}[X] \leq \mathbf{E}[Y]$*.*

Merkitsemme $X^+ := \max(X, 0)$ ja $X^- := -\min(X, 0)$. Tällöin X^+ ja X^- ovat positiivisia ja lisäksi $X = X^+ - X^-$ ja $|X| = X^+ + X^-$.

4.1.9 Määritelmä. Positiivisen satunnaismuuttujan X *odotusarvo* on

$$\mathbf{E}[X] \;\; := \;\; \sup\{\mathbf{E}[Y]\,;\, Y \in \mathcal{E}, Y \leq X\}$$

mikäli tämä supremum on äärellinen. Jos $\mathbf{E}[|X|] < \infty$, niin sanomme, että X on *integroituva* ja määrittelemme

$$\mathbf{E}[X] \;\; := \;\; \mathbf{E}[X^+] - \mathbf{E}[X^-].$$

Integroituvien satunnaismuuttujien luokka on $L^1 = L^1(\Omega, \mathcal{F}, \mathbf{P})$.

Odotusarvolle $\mathbf{E}[X]$ käytetään myös esimerkiksi merkintöjä

$$\int_\Omega X(\omega)\,\mathbf{P}[\mathrm{d}\omega], \quad \int_\Omega X(\omega)\,\mathrm{d}\mathbf{P}[\omega], \quad \int_\Omega X \mathrm{d}\mathbf{P},$$

$$\int_\mathbb{R} x\mathbf{P}_X[\mathrm{d}x] \quad \text{ja} \quad \int_\mathbb{R} x\,\mathrm{d}F_X(x).$$

Näistä kaksi viimeistä merkintää korostavat sitä seikkaa, että $\mathbf{E}[X]$ riippuu X:stä ainoastaan sen jakauman kautta: jos X:llä ja Y:llä on sama jakauma, niin $\mathbf{E}[X] = \mathbf{E}[Y]$.

4.1.10 Apulause. L^1 *on vektoriavaruus.*

Todistus. Harjoitustehtävä 4.2. □

4.1.11 Lause. *Olkoon* X *positiivinen.*

(i) *Tällöin on olemassa sellainen* $(X_n) \subset \mathcal{E}$*, että* $X_n \uparrow X$*.*
(ii) *Olkoon* $(X_n) \subset \mathcal{E}$ *sellainen, että* $X_n \uparrow X$*. Tällöin* $\mathbf{E}[X_n] \uparrow \mathbf{E}[X]$*, missä* $\mathbf{E}[X] < \infty$ *jos ja vain jos* $X \in L^1$*.*

Todistus. (i) Esimerkiksi valinta

$$X_n \;\; := \;\; \sum_{k \leq n2^n} (k-1)2^{-n}\mathbf{1}_{\{(k-1)2^{-n} \leq X < k2^{-n}\}} + n\mathbf{1}_{\{X \geq n\}}$$

antaa halutun jonon.

(ii) Koska $\mathbf{E}[X] = \sup\{\mathbf{E}[Y]; Y \in \mathcal{E}, Y \leq X\}$, niin $\mathbf{E}[X_n] \leq \mathbf{E}[X]$ kaikilla $n \in \mathbb{N}$. Siten välttämättä $\lim \mathbf{E}[X_n] \leq \mathbf{E}[X]$. Riittää siis osoittaa, että $\mathbf{E}[X] \leq \lim \mathbf{E}[X_n]$.

Osoitamme sitten, että jos $Y \in \mathcal{E}$ ja $Y \leq X$, niin $\mathbf{E}[Y] \leq \lim \mathbf{E}[X_n]$. Epäyhtälö $\mathbf{E}[X] \leq \lim \mathbf{E}[X_n]$ seuraa tästä, sillä tällöin

$$\begin{aligned}\mathbf{E}[X] &= \sup\{\mathbf{E}[Y]\,;\, Y \in \mathcal{E}, Y \leq X\} \\ &\leq \sup\{\lim \mathbf{E}[X_n]\,;\, Y \in \mathcal{E}, Y \leq X\} \\ &= \lim \mathbf{E}[X_n].\end{aligned}$$

Olkoon

$$Y = \sum_{k \leq m} a_k \mathbf{1}_{A_k},$$

missä $A_k = \{Y = a_k\}$. Olkoon $\varepsilon \in (0,1)$ mielivaltainen. Määrittelemme

$$Y_{n,\varepsilon} := (1-\varepsilon)Y\mathbf{1}_{\left\{(1-\varepsilon)Y \leq X_n\right\}}.$$

Olkoon

$$A_{n,k} := A_k \cap \{(1-\varepsilon)Y \leq X_n\}.$$

Koska $Y_{n,\varepsilon} \leq X_n$, niin $\mathbf{E}[Y_{n,\varepsilon}] \leq \mathbf{E}[X_n]$. Lisäksi

$$\mathbf{E}[Y_{n,\varepsilon}] = (1-\varepsilon)\sum_{k \leq m} a_k \mathbf{P}[A_{n,k}] \tag{4.1.12}$$

Nyt $A_{n,k} \uparrow A_k$, kun $n \to \infty$, sillä $X_n \uparrow X$ ja $Y \leq X$. Ottamalla raja-arvot yhtälössä (4.1.12) ja soveltamalla todennäköisyyden jatkuvuutta saamme

$$\begin{aligned}\lim_{n\to\infty} \mathbf{E}[Y_{n,\varepsilon}] &= (1-\varepsilon)\sum_{k \leq m} a_k \lim_{n\to\infty} \mathbf{P}[A_{n,k}] \\ &= (1-\varepsilon)\sum_{k \leq m} a_k \mathbf{P}[A_k] \\ &= (1-\varepsilon)\mathbf{E}[Y].\end{aligned}$$

Koska $\mathbf{E}[Y_{n,\varepsilon}] \leq \mathbf{E}[X_n]$, niin $\lim \mathbf{E}[Y_{n,\varepsilon}] \leq \lim \mathbf{E}[X_n]$. Siten, edellisen nojalla, $(1-\varepsilon)\mathbf{E}[Y] \leq \lim \mathbf{E}[X_n]$. Väite $\mathbf{E}[Y] \leq \lim \mathbf{E}[X_n]$ seuraa siitä, että $\varepsilon \in (0,1)$ oli mielivaltainen.

Lopuksi huomaamme, että todistuksen nojalla on selvää, että $\mathbf{E}[X] < \infty$ jos ja vain jos $\sup \mathbf{E}[X_n] < \infty$. □

4.1.13 Seuraus. *Olkoon $X \in L^1$. Tällöin on olemassa sellainen $(X_n) \subset \mathcal{E}$, että $X_n \to X$ ja $\mathbf{E}[X_n] \to \mathbf{E}[X]$.*

Todistus. Tämä väite seuraa soveltamalla lausetta 4.1.11 erikseen satunnaismuuttujan $X = X^+ - X^-$ positiiviseen ja negatiiviseen osaan. □

Lauseen 4.1.11 avulla voimme perustella vanhat tutut odotusarvon kaavat diskreeteille ja jatkuville satunnaismuuttujille.

4.1.14 Seuraus. (i) *Olkoon X diskreetti arvojoukkonaan $\{a_1, a_2, \ldots\}$. Tällöin $X \in L^1$ jos ja vain jos*

$$\sum |a_k| \mathbf{P}[X = a_k] \quad < \quad \infty$$

ja tällöin

$$\mathbf{E}[X] \quad = \quad \sum a_k \mathbf{P}[X = a_k].$$

(ii) *Olkoon X jatkuva. Tällöin $X \in L^1$ jos ja vain jos*

$$\int_{-\infty}^{\infty} |x| f_X(x) \, \mathrm{d}x \quad < \quad \infty$$

ja tällöin

$$\mathbf{E}[X] \quad = \quad \int_{-\infty}^{\infty} x f_X(x) \, \mathrm{d}x.$$

Todistuksen idea. (i) Väitteet näkee hajottamalla $X = X^+ - X^-$ ja aproksimoimalla positiivista osaa ja negatiivista osaa yksinkertaisilla satunnaismuuttujilla

$$\begin{aligned} X_n^+ &:= X^+ \mathbf{1}_{\{X^+ \le n\}}, \\ X_n^- &:= X^- \mathbf{1}_{\{X^- \le n\}}. \end{aligned}$$

(ii) Väitteet näkee hajottamalla $X = X^+ - X^-$ ja aproksimoimalla positiivista osaa ja negatiivista osaa yksinkertaisilla satunnaismuuttujilla

$$\begin{aligned} X_n^+ &:= \sum_{k \le n2^n} (k-1) 2^{-n} \mathbf{1}_{\{(k-1)2^{-n} \le X^+ < k2^{-n}\}} \\ X_n^- &:= \sum_{k \le n2^n} (k-1) 2^{-n} \mathbf{1}_{\{(k-1)2^{-n} \le X^- < k2^{-n}\}} \end{aligned}$$

ja käyttämällä epäoleellisen Riemann-integraalin määritelmää. □

4.1.15 Lause. $\mathbf{E} : L^1 \to \mathbb{R}$ *on positiivinen lineaarinen operaattori:*

(i) *jos $X, Y \in L^1$ ja $a, b \in \mathbb{R}$, niin $\mathbf{E}[aX + bY] = a\mathbf{E}[X] + b\mathbf{E}[Y]$,*
(ii) *jos $X \in L^1$ ja $X \ge 0$, niin $\mathbf{E}[X] \ge 0$.*

Todistus. (ii) Olkoot $X \in L^1$ positiivinen ja $Y \in \mathcal{E}$ sellainen, että $Y \leq X$. Tällöin $Y\mathbf{1}_{\{Y \geq 0\}} \in \mathcal{E}$ ja $Y\mathbf{1}_{\{Y \geq 0\}} \leq X$. Siten

$$\begin{aligned}\mathbf{E}[X] &= \sup\{\mathbf{E}[Y]\,;\, Y \in \mathcal{E}, Y \leq X\} \\ &\geq \sup\left\{\mathbf{E}\big[Y\mathbf{1}_{\{Y\geq 0\}}\big]\,;\, Y \in \mathcal{E}, Y \leq X\right\}.\end{aligned}$$

Mutta $\mathbf{E}[Y\mathbf{1}_{\{Y \geq 0\}}] \geq 0$ apulauseen 4.1.7 nojalla. Siten $\mathbf{E}[X] \geq 0$.

(i) Olkoot aluksi $X \in L^1$ positiivinen ja $a \geq 0$. Olkoon lisäksi $(X_n) \subset \mathcal{E}$ sellainen, että $X_n \uparrow X$. Tällöin $(aX_n) \subset \mathcal{E}$ ja $aX_n \uparrow aX$. Siten, apulauseen 4.1.7 ja lauseen 4.1.11 nojalla,

$$\mathbf{E}[aX] \;=\; \lim \mathbf{E}[aX_n] \;=\; a \lim \mathbf{E}[X_n] \;=\; a\mathbf{E}[X].$$

Olkoon sitten $X \in L^1$ ja $a \geq 0$. Koska $(aX)^{\pm} = aX^{\pm}$, niin edellisen nojalla

$$\begin{aligned}\mathbf{E}[aX] &= \mathbf{E}[(aX)^+] - \mathbf{E}[(aX)^-] \\ &= a\mathbf{E}[X^+] - a\mathbf{E}[X^-] \\ &= a\left(\mathbf{E}[X^+] - \mathbf{E}[X^-]\right) \\ &= a\mathbf{E}[X].\end{aligned}$$

Jos taas $a < 0$, niin väite $\mathbf{E}[aX] = a\mathbf{E}[X]$ seuraa kuten edellä huomaamalla, että nyt $(aX)^{\pm} = -aX^{\mp}$.

Olemme osoittaneet, että $\mathbf{E}[aX] = a\mathbf{E}[X]$. Seuraavaksi osoitamme, että $\mathbf{E}[X+Y] = \mathbf{E}[X] + \mathbf{E}[Y]$.

Olkoot aluksi $X, Y \in L^1$ positiivisia ja $(X_n), (Y_n) \subset \mathcal{E}$ sellaisia, että $X_n \uparrow X$ ja $Y_n \uparrow Y$. Tällöin $(X_n + Y_n) \subset \mathcal{E}$ ja $X_n + Y_n \uparrow X + Y$. Koska odotusarvo on lineaarinen $\mathcal{E}$:ssä (apulause 4.1.7), niin lauseen 4.1.11 nojalla

$$\begin{aligned}\mathbf{E}[X+Y] &= \lim \mathbf{E}[X_n + Y_n] \\ &= \lim\left(\mathbf{E}[X_n] + \mathbf{E}[Y_n]\right) \\ &= \lim \mathbf{E}[X_n] + \lim \mathbf{E}[Y_n] \\ &= \mathbf{E}[X] + \mathbf{E}[Y].\end{aligned}$$

Olkoon sitten $X, Y \in L^1$. Määritelmän nojalla

$$\mathbf{E}[X+Y] \;=\; \mathbf{E}[(X+Y)^+] - \mathbf{E}[(X+Y)^-].$$

Toisaalta, määritelmän ja edellisen nojalla,

$$\begin{aligned}\mathbf{E}[X] + \mathbf{E}[Y] &= \mathbf{E}[X^+] - \mathbf{E}[X^-] + \mathbf{E}[Y^+] - \mathbf{E}[Y^-] \\ &= \mathbf{E}[X^+ + Y^+] - \mathbf{E}[X^- + Y^-].\end{aligned}$$

Näemme, että $\mathbf{E}[X+Y] = \mathbf{E}[X] + \mathbf{E}[Y]$ jos ja vain jos

$$\mathbf{E}[(X+Y)^+] - \mathbf{E}[(X+Y)^-] \;=\; \mathbf{E}[X^+ + Y^+] - \mathbf{E}[X^- + Y^-].$$

Järjestemällä termejä ja yhdistämällä positiiviset satunnaismuuttujat saman odotusarvon alle saamme ehdon

$$\mathbf{E}[(X+Y)^+ + X^- + Y^-] \;=\; \mathbf{E}[X^+ + Y^+ + (X+Y)^-].$$

Mutta tämä ehto seuraa identiteetistä

$$(X+Y)^+ + X^- + Y^- \;=\; X^+ + Y^+ + (X+Y)^-,$$

jonka näkee yksinkertaisella kirjanpitoargumentilla. □

4.1.16 Seuraus. *Olkoon $X, Y \in L^1$ ja $X \le Y$. Tällöin*

(i) $\mathbf{E}[X] \le \mathbf{E}[Y]$,
(ii) $|\mathbf{E}[X]| \le \mathbf{E}[|X|]$.

4.2 Rajankäynti ja odotusarvo

Seuraavan tuloksen kohtia (i) ja (ii) kutsutaan *monotonisen konvergenssin lauseeksi*, kohtia (iii) ja (iv) *Fatoun lemmaksi* ja kohtaa (v) Lebesguen *dominoidun konvergenssin lauseeksi.*

4.2.1 Lause. *Olkoon $X, Y \in L^1$ ja (X_n) jono satunnaismuuttujia.*

(i) *Jos $X_n \uparrow X$ ja $\mathbf{E}[X_1] > -\infty$, niin $\mathbf{E}[X_n] \uparrow \mathbf{E}[X]$.*
(ii) *Jos $X_n \downarrow X$ ja $\mathbf{E}[X_1] < \infty$, niin $\mathbf{E}[X_n] \downarrow \mathbf{E}[X]$.*
(iii) *Jos $X_n \ge Y$, niin $\mathbf{E}[\liminf X_n] \le \liminf \mathbf{E}[X_n]$.*
(iv) *Jos $X_n \le Y$, niin $\mathbf{E}[\limsup X_n] \ge \limsup \mathbf{E}[X_n]$.*
(v) *Jos $X_n \to X$ ja $|X_n| \le Y$ kaikilla $n \in \mathbb{N}$, niin $\lim \mathbf{E}[X_n] = \mathbf{E}[X]$.*

Monotonisen kovergenssin lauseessa, eli kohdissa (i) ja (ii), ei tarvitse olettaa, että $X \in L^1$. Tällöin tosin kohdassa (i) voi käydä niin, että $\mathbf{E}[X] = \infty$ ja kohdassa (ii) on mahdollista, että $\mathbf{E}[X] = -\infty$.

Todistus. (i) Lauseen 4.1.11(ii) nojalla voimme valita sellaisen jonon $(Y_{n,k}) \subset \mathcal{E}$, että $Y_{n,k} \uparrow X_n$ ja $\mathbf{E}[Y_{n,k}] \uparrow \mathbf{E}[X_n]$. Tarkastelemme sitten maksimia

$$Z_k \;:=\; \max_{n \le k} Y_{n,k}.$$

Nyt $Z_k \in \mathcal{E}$. Lisäksi jono (Z_k) on kasvava, sillä $Y_{n,k} \leq Y_{n,k+1}$ kaikilla $n \in \mathbb{N}$. Koska (Z_k) on kasvava, niin on olemassa raja-arvo $Z := \lim Z_k$.

Nyt $Z_k \leq X_k$, sillä

$$\begin{aligned} Z_k &= \max(Y_{1,k}, \ldots, Y_{k,k}) \\ &\leq \max(X_1, \ldots, X_k) \\ &= X_k. \end{aligned}$$

Siispä, kun $n \leq k$, niin

$$Y_{n,k} \quad \leq \quad Z_k \quad \leq \quad X_k \quad \leq \quad X.$$

Ottamalla tästä raja-arvot ensin k:n ja sitten n:n suhteen näemme, että $Z = X$. Lisäksi näemme, että

$$\mathbf{E}[Y_{n,k}] \quad \leq \quad \mathbf{E}[Z_k] \quad \leq \quad \mathbf{E}[X_k]\,,$$

kun (edelleen) $n \leq k$. Pitämällä n kiinnitettynä ja antamalla k:n kasvaa näemme tästä, että

$$\mathbf{E}[X_n] \quad \leq \quad \mathbf{E}[Z] \quad \leq \quad \lim \mathbf{E}[X_k]\,.$$

Antamalla sitten n:n kasvaa näemme, että

$$\lim \mathbf{E}[X_n] \quad \leq \quad \mathbf{E}[Z] \quad \leq \quad \lim \mathbf{E}[X_k].$$

Mutta väite seuraa nyt muistamalla, että $Z = X$.

Lopuksi pitää päästä eroon oletuksesta, että X ja X_n ovat positiivisia. Helpointa tämä on tehdä tarkastelemalla "siirrettyjä" satunnaismuuttujia $X + X_0^-$ ja $X_n + X_0^-$. Jätämme yksityiskohdat harjoitustehtäviksi.

(ii) Tämä kohta on kohdan (i) peilikuva.

(iii) Olkoon $Z_n := \inf_{k \geq n} X_k$ Tällöin $Z_n \uparrow \liminf X_n$. Soveltamalla kohtaa (i) saamme

$$\mathbf{E}[Z_n] \quad \uparrow \quad \mathbf{E}[\liminf X_n]. \tag{4.2.2}$$

Toisaalta $X_k \geq Z_n$, kun $k \geq n$. Siten $\mathbf{E}[X_k] \geq \mathbf{E}[Z_n]$, kun $k \geq n$. Tästä taas seuraa, että $\inf_{k \geq n} \mathbf{E}[X_k] \geq \mathbf{E}[Z_n]$. Ottamalla raja-arvot n:n suhteen saamme

$$\liminf_n \mathbf{E}[X_n] \quad = \quad \lim_n \inf_{k \geq n} \mathbf{E}[X_k] \quad \geq \quad \lim_n \mathbf{E}[Z_n]. \tag{4.2.3}$$

Väite seuraa yhdistämällä (4.2.2) ja (4.2.3).

(iv) Tämä kohta on kohdan (iii) peilikuva.

(v) Tämä kohta seuraa kohdista (iii) ja (iv). Nimittäin

$$\begin{aligned}\mathbf{E}[X] &= \mathbf{E}[\liminf X_n] \\ &\leq \liminf \mathbf{E}[X_n] \\ &\leq \limsup \mathbf{E}[X_n] \\ &\leq \mathbf{E}[\limsup X_n] \\ &= \mathbf{E}[X].\end{aligned}$$

Siispä väite $\liminf \mathbf{E}[X_n] = \limsup \mathbf{E}[X_n] = \mathbf{E}[X]$ seuraa nyt hampurilaisperiaatteesta. □

Mikäli jokin ominaisuus A pätee todennäköisyydellä yksi eli $\mathbf{P}[A] = 1$, niin sanomme että A *melkein varmasti.*

4.2.4 Seuraus. (i) *Olkoot X_n, $n \in \mathbb{N}$, positiivisia satunnaismuuttujia. Tällöin*

$$\mathbf{E}\Big[\sum X_n\Big] = \sum \mathbf{E}[X_n],$$

missä yhtälö pätee mahdollisesti muodossa $\infty = \infty$.

(ii) *Olkoot $X_n, n \in \mathbb{N}$, satunnaismuuttujia, joille*

$$\sum \mathbf{E}[|X_n|] < \infty.$$

Tällöin $\sum X_n$ suppenee melkein varmasti *eli*

$$\mathbf{P}\Big[\omega \in \Omega\,;\, \sum X_n(\omega) \text{ suppenee}\Big] = 1.$$

Lisäksi

$$\mathbf{E}[X_n] = \sum \mathbf{E}[X_n].$$

Todistus. (i) Tämän kohdan jätämme harjoitustehtäväksi 4.8.

(ii) Olkoon $S := \sum |X_n|$. Kohdan (i) nojalla $\mathbf{E}[S] = \sum \mathbf{E}[|X_n|] < \infty$. Siten $\mathbf{P}[S = \infty] < \infty$, joten sarja $\sum |X_n|$ suppenee melkein varmasti. Siten sarja $\sum X_n$ suppenee itseisesti melkein varmasti, joten se suppenee melkein varmasti. Lopuksi huomaamme, että $\sum_{k \leq n} X_k \leq S$. Siten väite $\mathbf{E}[\sum X_n] = \sum \mathbf{E}[X_n]$ seuraa dominoidun konvergenssin lauseesta 4.2.1(v). □

4.3 Satunnaisvektorin muunnoksen odotusarvo

Seuraava tulos 4.3.1 mahdollistaa satunnaisvektorin X muunnoksen $f(X)$ odotusarvon $\mathbf{E}[f(X)]$ laskemisen tietämättä itse muunnoksen $f(X)$ jakaumaa $\mathbf{P}_{f(X)}$. Lause 4.3.1 tarjoaa myös helpon tavan perustella seuraus 4.1.14.

4.3.1 Lause. *Olkoon $X = (X_1, \ldots, X_d)$ satunnaisvektori ja $f : \mathbb{R}^d \to \mathbb{R}$ Borel-mitallinen. Jos $f(X) \in L^1$, niin*

$$\begin{aligned} \mathbf{E}[f(X)] &= \int_{\mathbb{R}^d} f(x_1, \ldots, x_d)\, \mathbf{P}_X[\mathrm{d}x_1, \ldots, \mathrm{d}x_d] \\ &= \int_{\mathbb{R}^d} f(x_1, \ldots, x_d)\, \mathrm{d}F_X(x_1, \ldots, x_d). \end{aligned} \tag{4.3.2}$$

Lisäksi $f(X) \in L^1$ jos ja vain jos

$$\int_{\mathbb{R}^d} \left| f(x_1, \ldots, x_d) \right| \mathbf{P}_X[\mathrm{d}x_1, \ldots, \mathrm{d}x_d] \;<\; \infty. \tag{4.3.3}$$

Todistus. Todistamme ainoastaan muuttujanvaihtokaavan (4.3.2) ja senkin vain yksiulotteisessa tapauksessa $d = 1$. Loput jätämme harjoitustehtäviksi 4.9 ja 4.10.

Askel 1: Oletamme aluksi, että funktio f on muotoa $f = \mathbf{1}_B$ jollekin $B \in \mathcal{B}$. Tällöin väite on selvä. Nimittäin

$$\mathbf{E}[\mathbf{1}_B(X)] \;=\; \mathbf{P}[X \in B] \;=\; \mathbf{P}_X[B] \;=\; \int_{\mathbb{R}} \mathbf{1}_B(x)\, \mathbf{P}_X[\mathrm{d}x].$$

Askel 2: Jos funktio f on yksinkertainen, eli muotoa $f = \sum_{k \le n} a_k \mathbf{1}_{B_k}$, niin väite seuraa kuten edellä käyttämällä odotusarvon ja integraalin lineaarisuutta. Nimittäin

$$\begin{aligned} \mathbf{E}[f(X)] &= \mathbf{E}\left[\sum_{k \le n} a_k \mathbf{1}_{B_k}(X)\right] \\ &= \sum_{k \le n} a_k \mathbf{E}[\mathbf{1}_{B_k}(X)] \\ &= \sum_{k \le n} a_k \int_{\mathbb{R}} \mathbf{1}_{B_k}(x)\, \mathbf{P}_X[\mathrm{d}x] \\ &= \int_{\mathbb{R}} \sum_{k \le n} \mathbf{1}_{B_k}(x)\, \mathbf{P}_X[\mathrm{d}x] \\ &= \int_{\mathbb{R}} f(x)\, \mathbf{P}_X[\mathrm{d}x]. \end{aligned}$$

Askel 3: Lopuksi aproksimoimme yleistä funktiota f yksinkertaisilla funktioilla f_n siten, että $f_n \leq |f|$ ja väite seuraa dominoidun konvergenssin lauseesta 4.2.1(v). $\square$

4.3.4 Esimerkki. Jos X tasaisesti jakautunut välille $[0,1]$ ja $Y = e^X$, niin

$$\mathbf{E}[Y] \;=\; \mathbf{E}\big[e^X\big] \;=\; \int_0^1 e^x \,\mathrm{d}x \;=\; e-1.$$

Emme siis tarvinneet tietoa satunnaismuuttujan Y jakaumasta. Selvitämme sen nyt kuitenkin huvin vuoksi:

$$\mathbf{P}[Y \leq y] \;=\; \mathbf{P}[e^X \leq y] \;=\; \mathbf{P}[X \leq \ln y] \;=\; \ln y \mathbf{1}_{[1,e]}(y) + \mathbf{1}_{(e,\infty)}(y),$$

Näemme, että Y on jatkuvasti jakautunut kertymä- ja tiheysfunktioinaan

$$F_Y(y) \;=\; \ln y \mathbf{1}_{[1,e]}(y) + \mathbf{1}_{(e,\infty)}(y) \quad \text{ja} \quad f_Y(y) \;=\; \frac{1}{y}\mathbf{1}_{[1,e]}(y).$$

4.4 Epäyhtälöitä

Funktion $f : \mathbb{R}^d \to \mathbb{R}$ *tangentti* pisteessä $x \in \mathbb{R}^d$ on mikä tahansa pistettä $(x, f(x)) \in \mathbb{R}^{d+1}$ sivuava suora (tai oikeammin aliavaruus) $l_x(y) = a \cdot (y-x) + f(x)$, missä kulmakerroin $a = a(x) \in \mathbb{R}^d$. Sivuavuus pisteessä $(x, f(x)) \in \mathbb{R}^{d+1}$ tarkoittaa sitä, että $l_x(x) = f(x)$ ja graafit $\{(y, f(y)); y \in \mathbb{R}^d\}$ ja $\{(y, l_x(y)); y \in \mathbb{R}^d\}$ eivät leikkaa jossakin pisteen $(x, f(x))$ avoimessa ympäristössä.

Funktio $f : \mathbb{R}^d \to \mathbb{R}$ on *konveksi*, jos se on kaikkien tangenttienssa yläpuolella: $f \geq l_x$ kaikilla f:n tangenteilla l_x.

4.4.1 Lause. *Olkoon $f : \mathbb{R}^d \to \mathbb{R}$ konveksi funktio ja $X = (X_1, \ldots, X_d)$ satunnaisvektori, jonka komponentit ovat integroituvia. Tällöin pätee* Jensenin epäyhtälö

$$\mathbf{E}\big[f(X_1, \ldots, X_d)\big] \;\geq\; f\big(\mathbf{E}[X_1], \ldots, \mathbf{E}[X_d]\big).$$

Todistus. Huomaamme aluksi, että väite on hyvin määritelty: konveksi funktio f on jatkuva (harjoitustehtävä 4.11) ja jatkuvat funktiot ovat Borel-mitallisia, joten $f(X)$ on satunnaismuuttuja.

Tarkastelemme funktion f tangenttia pisteessä $(\mathbf{E}[X_1], \ldots, \mathbf{E}[X_d]) \in \mathbb{R}^d$. Tällöin funktion f konveksisuuden nojalla

$$\begin{aligned} f(X_1, \ldots, X_d) &\geq l_{(\mathbf{E}[X_1], \ldots, \mathbf{E}[X_d])}(X_1, \ldots, X_d) \\ &= \sum_{i \leq d} a_i(X_i - \mathbf{E}[X_i]) + f(X_1, \ldots, X_d). \end{aligned}$$

Siten

$$\begin{aligned}\mathbf{E}\big[f(X_1,\ldots,X_d)\big] &\geq \mathbf{E}\left[\sum_{i\leq d} a_i(X_i-\mathbf{E}[X_i]) + f(\mathbf{E}[X_1],\ldots,\mathbf{E}[X_d])\right] \\ &= \sum_{i\leq d} a_i\mathbf{E}\big[X_i-\mathbf{E}[X_i]\big] + f(\mathbf{E}[X_1],\ldots,\mathbf{E}[X_d]) \\ &= f(\mathbf{E}[X_1],\ldots,\mathbf{E}[X_d]).\end{aligned}$$

□

Jensenin epäytälön avulla on kohtalaisen helppo todistaa mittateorian "suuret epäyhtälöt": Hölder, Minkowski ja Schwarz.

4.4.2 Seuraus. *Olkoot X ja Y satunnaismuuttujia ja $p,q\geq 1$.*

(i) *Olkoon $1/p+1/q=1$ ja $\mathbf{E}[|X|^p],\mathbf{E}[|Y|^q]<\infty$. Tällöin pätee* Hölderin epäyhtälö

$$|\mathbf{E}[XY]| \leq \mathbf{E}[|XY|] \leq \mathbf{E}[|X|^p]^{\frac{1}{p}}\mathbf{E}[|Y|^q]^{\frac{1}{q}}$$

(ii) *Olkoon $\mathbf{E}[|X|^p],\mathbf{E}[|Y|^p]<\infty$. Tällöin pätee* Minkowskin epäyhtälö

$$\mathbf{E}[|X+Y|^p]^{\frac{1}{p}} \leq \mathbf{E}[|X|^p]^{\frac{1}{p}} + \mathbf{E}[|Y|^p]^{\frac{1}{p}}.$$

(iii) *Olkoon $\mathbf{E}[X^2],\mathbf{E}[Y^2]<\infty$. Tällöin pätee* Schwarzin epäyhtälö

$$|\mathbf{E}[XY]| \leq \mathbf{E}[|XY|] \leq \sqrt{\mathbf{E}[X^2]\mathbf{E}[Y^2]}.$$

Todistus. (i) Ensimmäinen epäyhtälö on triviaali. Siten voimme olettaa, että $X,Y\geq 0$. Lisäksi voimme olettaa, että $\mathbf{E}[X^p]>0$, sillä muuten $\mathbf{P}[X=0]=1$ ja epäyhtälö toteutuu muodossa $0\leq 0$. Todistus perustuu nyt todennäköisyysmitan vaihtoon ja Jensenin epäyhtälöön. Olkoon $A\in\mathcal{F}$. Määrittelemme

$$\mathbf{Q}[A] := \frac{\mathbf{E}[X^p\mathbf{1}_A]}{\mathbf{E}[X^p]}.$$

Tällöin $\mathbf{Q}$ on todennäköisyysmitta avaruudella $(\Omega,\mathcal{F})$ (harjoitustehtävä 4.13). Odotusarvoa tämän todennäköisyysmitan suhteen merkitsemme symbolilla $\mathbf{E}_{\mathbf{Q}}$. Olkoon sitten

$$Z := \frac{Y}{X^{p-1}}\mathbf{1}_{\{X>0\}}.$$

Tällöin

$$\begin{aligned}\mathbf{E}[|XY|] &= \mathbf{E}[XY] \\ &= \mathbf{E}\big[XY\mathbf{1}_{\{X>0\}}\big] \\ &= \mathbf{E}\left[\frac{Y}{X^{p-1}}\mathbf{1}_{\{X>0\}}X^p\right] \\ &= \mathbf{E}[ZX^p] \\ &= \mathbf{E}[X^p]\mathbf{E}_{\mathbf{Q}}\left[Z\right]\end{aligned}$$

Koska $p \geq 1$, niin kuvaus $x \mapsto |x|^p$ on koveksi. Siten, Jensenin epäyhtälön 4.4.1 nojalla, $\mathbf{E}_{\mathbf{Q}}[Z]^q \leq \mathbf{E}_{\mathbf{Q}}[Z^q]$. Siispä

$$\begin{aligned}\mathbf{E}[|XY|] &\leq \mathbf{E}[X^p]\mathbf{E}_{\mathbf{Q}}\left[Z^q\right]^{\frac{1}{q}} \\ &= \mathbf{E}[X^p]^{1-\frac{1}{q}}\mathbf{E}\left[\left(\frac{Y}{X^{p-1}}\mathbf{1}_{\{X>0\}}\right)^q X^p\right]^{\frac{1}{q}} \\ &= \mathbf{E}[X^p]^{1-\frac{1}{q}}\mathbf{E}\left[\frac{Y^q}{X^{(p-1)q}}X^p\right]^{\frac{1}{q}} \\ &= \mathbf{E}[X^p]^{\frac{1}{p}}\mathbf{E}[Y^q]^{\frac{1}{q}}.\end{aligned}$$

Viimeisessä yhtälössä käytimme hyväksi sitä, että $1-1/q = 1/p$ ja $(p-1)q = p$.

(ii) Jos $p = 1$, niin tämä tulos seuraa kolmioepäyhtälöstä. Oletamme sitten, että $p > 1$. Todistus perustuu nyt Hölderin epäyhtälöön. Olkoon $q = \frac{p}{p-1}$. Tällöin

$$\begin{aligned}\mathbf{E}[|X+Y|^p] &\leq \mathbf{E}[|X||X+Y|^{p-1}] + \mathbf{E}[|Y||X+Y|^{p-1}] \\ &\leq \mathbf{E}[|X|^p]^{\frac{1}{p}}\mathbf{E}[|X+Y|^{(p-1)q}]^{\frac{1}{q}} + \mathbf{E}[|Y|^p]^{\frac{1}{p}}\mathbf{E}[|X+Y|^{(p-1)q}]^{\frac{1}{q}} \\ &= \left(\mathbf{E}[|X|^p]^{\frac{1}{p}} + \mathbf{E}[|Y|^p]^{\frac{1}{p}}\right)\mathbf{E}[|X+Y|^p]^{1-\frac{1}{p}}.\end{aligned}$$

(iii) Schwarzin epäyhtälö seuraa välittömästi Hölderin epäyhtälöstä valitsemalla $p = q = 2$. □

4.4.3 Lause. *Olkoon $a > 0$ ja $X \in L^1$ positiivinen. Tällöin pätee* Tšebyševin epäyhtälö

$$\mathbf{P}[X \geq a] \;\leq\; \frac{\mathbf{E}[X]}{a}.$$

Todistus. Koska $a\mathbf{1}_{\{X\geq a\}} \leq X\mathbf{1}_{\{X\geq a\}} \leq X$, niin

$$a\mathbf{P}[X \geq a] = a\mathbf{E}\big[\mathbf{1}_{\{X\geq a\}}\big] = \mathbf{E}\big[a\mathbf{1}_{\{X\geq a\}}\big] \leq \mathbf{E}\big[X\mathbf{1}_{\{X\geq a\}}\big] \leq \mathbf{E}[X]$$

Väite seuraa jakamalla puolittain a:lla. □

Tšebyševin epäyhtälöä voidaan parantaa satunnaismuuttujan X "integroituvuuden asteen" mukaan. Seurauksessa 4.4.4 käsittelemme neliöintegroituvaa ja eksponentiaalisesti integroituvaa astetta.

4.4.4 Seuraus. (i) *Olkoon* $\mathbf{E}[X^2] < \infty$ *ja* $a > 0$. *Tällöin pätee* Markovin epäyhtälö

$$\mathbf{P}\big[|X - \mathbf{E}[X]| \geq a\big] \ \leq \ \frac{\mathbf{Var}[X]}{a^2},$$

missä

$$\mathbf{Var}[X] \ := \ \mathbf{E}[(X - \mathbf{E}[X])^2]$$

on satunnaismuuttujan X varianssi.

(ii) *Olkoon*

$$\Lambda(\theta) \ := \ \ln \mathbf{E}[e^{\theta X}]$$

satunnaismuuttujan X kumulantit generoiva funktio. *Oletamme, että pätee* Cramérin ehto: $\Lambda(\theta) < \infty$ *kaikilla* $\theta \in \mathbb{R}$. *Tällöin pätee* Tšernovin epäyhtälö

$$\mathbf{P}[X \geq a] \ \leq \ e^{-\Lambda^*(a)},$$

missä

$$\Lambda^*(a) \ := \ \sup_{\theta \in \mathbb{R}} \{a\theta - \Lambda(\theta)\}$$

on funktion Λ Fenchel–Legendre-muunnos.

Todistus. (i) Markovin epäyhtälö seuraa Tšebyševin epäyhtälöstä sijoittamalla X:n paikalle $(X - \mathbf{E}[X])^2$. Nimittäin nyt

$$\begin{aligned}
\mathbf{P}\big[\big|X - \mathbf{E}[X]\big| \geq a\big] &= \mathbf{P}\big[(X - \mathbf{E}[X])^2 \geq a^2\big] \\
&\leq \frac{\mathbf{E}\big[(X - \mathbf{E}[X])^2\big]}{a^2} \\
&= \frac{\mathbf{Var}[X]}{a^2}.
\end{aligned}$$

(ii) Olkoon $\theta \in \mathbb{R}$. Tällöin, Tšebyševin epäyhtälön nojalla,

$$\begin{aligned}
\mathbf{P}[X \geq a] &= \mathbf{P}\left[e^{\theta X} \geq e^{\theta a}\right] \\
&\leq \frac{\mathbf{E}\big[e^{\theta X}\big]}{e^{\theta a}} \\
&= e^{\Lambda(\theta) - \theta a} \\
&= e^{-(\theta a - \Lambda(\theta))}.
\end{aligned}$$

Tšernovin epäyhtälö seuraa optimoimalla yli vapaan parametrin $\theta \in \mathbb{R}$. $\square$

4.4.5 Esimerkki. Ylioppilaskokeen toisen kotimaisen kielen kokeessa on n kappaletta neljän kohdan monivalintatehtävää. Ylioppilaaksi halajavan on vastattava oikein vähintään 40% monivalintatehtävistä. Kokeen suunnitelija ei halua päästää ylioppilaiksi "puhtaita arpojia", jotka eivät osaa toista kotimaista kieltä. Kuinka suuri pitää n:n olla, jotta todennäköisyys "puhtaan arpojan" ylioppilaaksi pääsylle on pienempi kuin 0,1%?

Ongelma voidaan ratkaista tarkasti huomaamalla, että "puhtaan arpojan" oikeiden vastausten lukumäärä X on binomijakautunut parametrein n ja $1/4$. Kokeen suunnittelija ei kuitenkaan viitsi tai osaa tehdä tietokoneohjelmaa, joka suorittaa vaadittavat työläät laskut. Siksi hän käyttää Tšebyševin, Markovin tai Tšernovin arvioita. Arviointimenetelmien antamat tulokset ovat:

Tšebyšev: Esimerkin 4.1.6 kohdan (ii) nojalla $\mathbf{E}[X] = n/4$. Saamme siis arvion

$$\mathbf{P}[X \geq 0{,}4n] \;\leq\; \frac{\mathbf{E}[X]}{0{,}4n} \;\leq\; \frac{n/4}{0{,}4n} \;=\; \frac{1}{1{,}6} \;=\; 0{,}625.$$

Tämä arvio on n:stä riippumaton ja suurempi kuin $0,1\%$. Siten se ei hyödytä kokeen suunnittelijaa ollenkaan: tämän arvon mukaan mikään $n \in \mathbb{N}$ ei ole riittävän suuri.

Markov: Tämä epäyhtälö ei sovellu sellaisenaan ongelmaan, sillä se käsittelee todennäköisyyttä $\mathbf{P}[|X - n/4| \geq 0{,}4n]$, eikä todennäköisyyttä $\mathbf{P}[X \geq 0{,}4n]$. Voimme kuitenkin käyttää arviota

$$\mathbf{P}[X \geq a] \;\leq\; \frac{\mathbf{E}[X^2]}{a^2},$$

joka saadaan Tšebyševin epäyhtälöstä huomaamalla, että $\mathbf{P}[X \geq a] = \mathbf{P}[X^2 \geq a^2]$, kun X on positiivinen. Tehtäväksi jää siis laskea $\mathbf{E}[X^2]$. Jätämme tämän laskun harjoitustehtäväksi 4.14 ja toteamme vain, että vastaus on $\mathbf{E}[X^2] = 3n/16 + n^2/16$. Saamme siis arvion

$$\mathbf{P}[X \geq 0{,}4n] \;\leq\; \frac{\mathbf{E}[X^2]}{(0,4n)^2} \;=\; \frac{5(3+n)}{128n}.$$

Tämäkään arvio ei auta meitä: senkään mukaan mikään n ei ole riittävän iso.

Tšernov: Laskemme aluksi binomijakauman kumulantit generoivan funktion. Käyttämällä *Newtonin binomikaavaa*

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

näemme, että

$$\begin{aligned}\Lambda(\theta) &= \ln \sum_{k=0}^{n} e^{\theta k} \binom{n}{k} \left(\frac{1}{4}\right)^k \left(1 - \frac{1}{4}\right)^{n-k} \\ &= \ln \sum_{k=0}^{n} \binom{n}{k} \left(e^{\theta}\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{n-k} \\ &= \ln \left(e^{\theta}\frac{1}{4} + \frac{3}{4}\right)^n \\ &= n \ln \left(e^{\theta}\frac{1}{4} + \frac{3}{4}\right).\end{aligned}$$

Sitten ratkaisemme optimointitehtävän $\Lambda^*(a) = \sup_{\theta\in\mathbb{R}}\{a\theta - \Lambda(\theta)\}$ normaaliin tapaan tarkastelemalla derivaatan nollakohtia. Nyt

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\theta}\{a\theta - \Lambda(\theta)\} &= \frac{\mathrm{d}}{\mathrm{d}\theta}\left\{a\theta - n\ln\left(e^{\theta}\frac{1}{4} + \frac{3}{4}\right)\right\} \\ &= a - n\frac{e^{\theta}/4}{e^{\theta}/4 + 3/4}.\end{aligned}$$

Siten optimipiste $\theta = \theta^*$ saadaan kaavasta

$$a - n\frac{e^{\theta^*}}{e^{\theta^*} + 3} = 0.$$

Siispä

$$\theta^* = \ln\frac{3a}{n-a},$$

ja

$$\begin{aligned}\Lambda^*(a) &= a\theta^* - \Lambda(\theta^*) \\ &= a\ln 3 + a\ln\frac{a}{n-a} - n\ln 3 + 2n\ln 2 - n\ln\frac{n}{n-a}.\end{aligned}$$

Saamme siis arvion

$$\begin{aligned}\mathbf{P}[X \geq 0{,}4n] &\leq e^{-\Lambda^*(0{,}4n)} \\ &\leq 0{,}948^n\end{aligned}$$

Siispä ehto on $0{,}948^n \leq 0{,}001$. Tästä kokeilemalla näemme, että 130 tehtävää riittää. Tämä on melkoinen parannus Markovin arvioon.

Esitämme lopuksi tarkan ratkaisun. Tarkka kaava on

$$\mathbf{P}[X \geq 0{,}4n] \quad = \quad \sum_{k=0{,}4n}^{n} \binom{n}{k} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{n-k}.$$

Kokeilemalla tietokoneella näemme, että 91 tehtävää ei riitä, mutta 92 tehtävää riittää.

Harjoitustehtäviä

4.1. Olkoon satunnaismuuttujan X kertymäfunktio

$$F_X(x) \quad = \quad \frac{1}{\pi} \arctan x + \frac{1}{2}.$$

Onko X integroituva?

4.2. Todista apulause 4.1.10.

4.3. Todista seuraus 4.1.14.

4.4. Osoita, että $X \in L^1$ jos ja vain jos

$$\sum \mathbf{P}[|X| \geq n] \quad < \quad \infty.$$

4.5. Olkoon X Poisson-jakautunut parametrilla λ, eli X on diskreetti arvojoukkonaan $\mathbb{N} \cup \{0\}$ ja

$$\mathbf{P}[X = k] \quad = \quad e^{-\lambda} \frac{\lambda^k}{k!}.$$

Laske $\mathbf{E}[X]$ ja $\mathbf{Var}[X]$.

4.6. Olkoon X *normaalisti jakautut* parametrein μ ja σ^2 eli X on jatkuva arvojoukkonaan $\mathbb{R}$ ja

$$\mathbf{P}[X \in A] \quad = \quad \frac{1}{\sqrt{2\pi}\sigma} \int_A e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \,\mathrm{d}x.$$

Laske $\mathbf{E}[X]$ ja $\mathbf{Var}[X]$.

4.7. Olkoon X Cauchy-jakautunut, eli X on jatkuva arvojoukkonaan $\mathbb{R}$ ja

$$\mathbf{P}[X \in A] \quad = \quad \frac{1}{\pi} \int_A \frac{1}{1+x^2} \,\mathrm{d}x.$$

Laske $\mathbf{E}[X]$ ja $\mathbf{Var}[X]$.

4.8. Todista seurauksen 4.2.4 kohta (i).

4.9. Todista lauseen 4.3.1 muuttujanvaihtokaava (4.3.2) d-ulotteisessa tapauksessa.

4.10. Todista muuttujanvaihtolauseen 4.3.1 integroituvuusehto (4.3.3).

4.11. Osoita, että konveksit funktiot ovat jatkuvia.

4.12. Osoita, että kaksi kertaa derivoituva funktio $f : \mathbb{R} \to \mathbb{R}$ on konveksi jos ja vain jos f'' on positiivinen.

4.13. Olkoon X posiivinen satunnaismuuttuja, jolle $\mathbf{E}[X] > 0$. Olkoon $A \in \mathcal{F}$. Asetamme

$$\mathbf{Q}[A] \quad := \quad \frac{\mathbf{E}[X\mathbf{1}_A]}{\mathbf{E}[X]}.$$

Osoita, että $\mathbf{Q}$ on todennäköisyysmitta avaruudella $(\Omega, \mathcal{F})$.

4.14. Olkoon X binomijakautunut parametrein n ja p. Laske $\mathbf{E}[X^2]$.

Luku 5

Riippumattomuus

Riippumattomuuden, tai tarkemmin stokastisen riippumattomuuden, idea on ennustamattomuus. Esimerkiksi tapahtumat A ja B ovat riippumattomia, jos tapahtuman A sattuminen ei muuta näkemystämme tapahtuman B sattumisen todennäköisyydestä: $\mathbf{P}[B|A] = \mathbf{P}[B]$. *Siis A:n sattuminen ei anna meille mitään informaatiota siitä, miten B:lle käy. On huomattavaa, että stokastinen riippumattomuus on aivan eri käsite kuin kausaalinen riippumattomuus. Siten vanhasta kansanviisaudesta, että kaikki riippuu kaikesta ei seuraa, että kaikki asiat ovat stokastisesti riippuvia.*

Tämän luvun tarkoitus on esittää riippumattomuus yleisemmin σ-algebroja koskevana ominaisuutena. Kun σ-algebrat tulkitaan "informaatioiksi", niin tämä yleistys on varsin luonnollinen.

5.1 Riippumattomat σ-algebrat, satunnaisvektorit ja tapahtumat

5.1.1 Määritelmä. Olkoon J jokin indeksijoukko.

(a) Ali-σ-algebrat $\mathcal{G}_j \subset \mathcal{F}$, $j \in J$, ovat *riippumattomia*, jos kaikilla erillisillä indekseillä $j_1, \ldots, j_n \in J$ ja kaikilla $A_k \in \mathcal{G}_{j_k}$ pätee

$$\mathbf{P}\left[\bigcap_{k \le n} A_k\right] \;=\; \prod_{k \le n} \mathbf{P}[A_k]\,.$$

Ali-σ-algebrat on *pareittain riippumattomat*, jos

$$\mathbf{P}[A_k \cap A_l] \;=\; \mathbf{P}[A_k]\mathbf{P}[A_l]$$

kaikilla kaikilla erillisillä indekseillä $k, l \in J$ ja $A_k \in \mathcal{G}_k$, $A_l \in \mathcal{G}_l$.

(b) Satunnaismuuttujat tai -vektorit X_j, $j \in J$, ovat (pareittain) riippumattomia, jos niiden virittämät σ-algebrat $\sigma(X_j)$, $j \in J$, ovat (pareittain) riippumattomia.
Satunnaisvektorin X virittämä σ-algebra $\sigma(X)$ on pienin σ-algebra, jonka suhteen X on mitallinen. Itse asiassa $\sigma(X)$ koostuu joukoista $\{X \in B\}$, $B \in \mathcal{B}_d$. Tämä seuraa siitä, että kokoelma $\{X \in B\}$, $B \in \mathcal{B}_d$, on jo valmiiksi σ-algebra, ja toisaalta sen on kuitenkin sisällytettävä σ-algebraan $\sigma(X)$.

(c) Tapahtumat $A_j \in \mathcal{F}$, $j \in J$, ovat (pareittain) riippumattomia, jos indikaattorit $\mathbf{1}_{A_j}$, $j \in J$, ovat (pareittain) riippumattomia.

5.1.2 Esimerkki. Koska $\sigma(\mathbf{1}_{A_i}) = \{\varnothing, \Omega, A_i, A_i^c\}$, niin näemme että tapahtumat A_1 ja A_2 ovat riippumattomia jos ja vain jos $\mathbf{P}[A_1 \cap A_2] = \mathbf{P}[A_1]\mathbf{P}[A_2]$. Nimittäin tästä ehdosta seuraa välittömästi, että $\mathbf{P}[A_1 \cap A_2^c] = \mathbf{P}[A_1]\mathbf{P}[A_2^c]$, $\mathbf{P}[A_1^c \cap A_2] = \mathbf{P}[A_1^c]\mathbf{P}[A_2]$ ja $\mathbf{P}[A_1^c \cap A_2^c] = \mathbf{P}[A_1^c]\mathbf{P}[A_2^c]$ (siis jos tapahtumat A_1 ja A_2 ovat riippumattomia, niin myös tapahtumat A_1 ja A_2^c ovat riippumattomia, samoin tapahtumat A_1^c ja A_2 ovat riippumattomia, kuten myös tapahtumat A_1^c ja A_2^c). Siten määritelmä 5.1.1(c) on vanhan tutun riippumattomuuden määritelmän yleistys.

5.1.3 Lause. *Seuraavat ovat yhtäpitäviä:*

(i) *satunnaisvektorit $X = (X_1, \ldots, X_d)$ ja $Y = (Y_1, \ldots, Y_{d'})$ ovat riippumattomia,*

(ii) *kaikille $A \in \mathcal{B}_d$ ja $B \in B_{d'}$ pätee*

$$\mathbf{P}[X \in A, Y \in B] \quad = \quad \mathbf{P}[X \in A]\,\mathbf{P}[Y \in B],$$

(iii) *kaikille $A \in \mathcal{I}$ ja $B \in \mathcal{I}'$, missä $\mathcal{I}$ ja $\mathcal{I}'$ ovat sellaisia π-luokkia, että $\mathcal{B}_d = \sigma(\mathcal{I})$ ja $\mathcal{B}_{d'} = \sigma(\mathcal{I}')$ pätee*

$$\mathbf{P}[X \in A, Y \in B] \quad = \quad \mathbf{P}[X \in A]\,\mathbf{P}[Y \in B],$$

(iv) *yhteiskertymäfunktiolle pätee hajotelma*

$$F_{X,Y}(x_1, \ldots, x_d, y_1, \ldots, y_{d'}) \quad = \quad F_X(x_1, \ldots, x_d)\,F_Y(y_1, \ldots, y_{d'}),$$

(v) *satunnaismuuttujat $f(X)$ ja $g(Y)$ ovat riippumattomia kaikilla Borel-funktioilla $f : \mathbb{R}^d \to \mathbb{R}$ ja $g : \mathbb{R}^{d'} \to \mathbb{R}$,*

(vi)

$$\mathbf{E}[f(X)g(Y)] \quad = \quad \mathbf{E}[f(X)]\mathbf{E}[g(Y)]$$

kaikilla rajoitetuilla Borel-funktioilla $f : \mathbb{R}^d \to \mathbb{R}$ ja $g : \mathbb{R}^{d'} \to \mathbb{R}$, *Todistus.* (i) $\Leftrightarrow$ (ii): Kohta (ii) vain muotoilee määritelmän uudella tavalla. Nimittäin $\sigma(X)$ koostuu joukoista $\{X \in A\}$, $A \in \mathcal{B}_d$, ja $\sigma(Y)$ koostuu joukoista $\{Y \in B\}$, $B \in \mathcal{B}_{d'}$.

(ii) $\Rightarrow$ (iii): Tämä implikaatio on selvä.

(ii) $\Leftarrow$ (iii): Tämä implikaatio seuraa Dynkinin laajennuslauseesta 2.3.8.

(ii) $\Rightarrow$ (iv): Tämä implikaatio on selvä.

(ii) $\Leftarrow$ (iv): Tämä implikaatio seuraa Dynkinin laajennuslauseesta 2.3.8, sillä joukkoluokka $\times_{k\leq d}(-\infty, x_k]$, $x_1, \ldots, x_k \in \mathbb{R}$, on π-luokka, joka virittää σ-algebran $\mathcal{B}_d$.

(ii) $\Leftarrow$ (v): Tämän implikaation jätämme harjoitustehtäväksi 5.4.

(ii) $\Rightarrow$ (v): Huomaamme, että $\sigma(f(X)) \subset \sigma(X)$ ja $\sigma(g(Y)) \subset \sigma(Y)$. Nimittäin $\sigma(f(X))$ koostuu joukoista $\{f(X) \in B\}$, $B \in \mathcal{B}$. Mutta

$$\{f(X) \in B\} \quad = \quad \{X \in f^{-1}B\},$$

ja $f^{-1}B \in \mathcal{B}_d$, sillä f on Borel-mitallinen. Siten $\{f(X) \in B\} \in \sigma(X)$. Samalla tavalla näemme, että $\{g(Y) \in B\} \in \sigma(Y)$.

Väite seuraa nyt siitä, että σ-algebrojen riippumattomuus periytyy ali-σ-algebroille (harjoitustehtävä 5.5)

(ii) $\Leftarrow$ (vi): Tämän kohdan jätämme harjoitustehtäväksi 5.6.

(ii) $\Rightarrow$ (vi): Kohdan (ii) nojalla

$$\mathbf{E}[f(X)g(Y)] \quad = \quad \mathbf{E}[f(X)]\mathbf{E}[g(Y)] \tag{5.1.4}$$

pätee, kun f ja g ovat indikaattorifunktiota: $f = \mathbf{1}_A$ ja $g = \mathbf{1}_B$ joillekin $A \in \mathcal{B}_d$ ja $B \in \mathcal{B}_{d'}$. Mutta odotusarvon lineaarisuudesta seuraa sitten, että yhtälö (5.1.4) pätee yksinkertaisille funktioille

$$\begin{aligned} f(x) &= \sum_{k\leq n} a_k \mathbf{1}_{A_k}(y), \quad A_k \in \mathcal{B}_d, \\ g(y) &= \sum_{l\leq n'} b_l \mathbf{1}_{B_l}(y), \quad B_l \in \mathcal{B}_{d'}. \end{aligned}$$

Lopuksi huomaamme, että väite seuraa aproksimoimalla satunnaismuuttujia $f(X)$ ja $g(Y)$ yksinkertaisilla satunnaismuuttujilla ja käyttämällä dominoidun konvergenssin lausetta 4.2.1(v). Dominoidun konvergenssin lauseen käyttö on perusteltua, koska funktiot f ja g ovat rajoitettuja. $\square$

Lauseelle 5.1.3 on olemassa luonnolinen vastine, joka koskee perhettä $X_j = (X_{j,1}, \ldots, X_{j,d_j})$, $j \in J$, satunnaisvektoreita. Jätämme tämän lauseen muotoilun ja todistamisen harjoitustehtäväksi 5.7.

5.2 Borelin ja Cantellin lemmat

5.2.1 Lause. *Olkoot* A_n, $n \in \mathbb{N}$, *tapahtumia. Tällöin pätee* Borelin ja Cantellin lemmat*:*

(i) *Jos*

$$\sum \mathbf{P}[A_n] \;<\; \infty,$$

niin

$$\mathbf{P}[\limsup A_n] \;=\; 0.$$

(ii) *Jos tapahtumat* A_n, $n \in \mathbb{N}$, *ovat riippumattomia ja*

$$\sum \mathbf{P}[A_n] \;=\; \infty,$$

niin

$$\mathbf{P}[\limsup A_n] \;=\; 1.$$

Todistus. (i) Olkoon $B_n := \cup_{k \geq n} A_k$. Tällöin $\limsup A_n = \cap B_n$. Koska $\cap B_n \subset B_m$ kaikilla $m \in \mathbb{N}$, niin

$$\mathbf{P}[\cap B_n] \;\leq\; \mathbf{P}[B_m] \;\leq\; \sum_{k \geq m} \mathbf{P}[A_k].$$

Mutta suppenevan sarjan jäännösterminä $\sum_{k \geq m} \mathbf{P}[A_k] \to 0$, kun $m \to \infty$. Siten $\mathbf{P}[\limsup A_n] = \mathbf{P}[\cap B_n] = 0$.

(ii) Koska

$$(\limsup A_n)^c \;=\; \bigcup_{m \geq 1} \bigcap_{n \geq m} A_n^c \;=\; \liminf A_n^c,$$

niin

$$\mathbf{P}[(\limsup A_n)^c] \;=\; \lim_{m \to \infty} \mathbf{P}\left[\bigcap_{n \geq m} A_n^c\right].$$

Olkoon sitten $N > m$ kiinteä. Tällöin, riippumattomuuden nojalla,

$$\begin{aligned}
\mathbf{P}\left[\bigcap_{n \geq m} A_n^c\right] \;&\leq\; \mathbf{P}\left[\bigcap_{N \geq n \geq m} A_n^c\right] \\
&=\; \prod_{N \geq n \geq m} \mathbf{P}[A_n^c] \\
&=\; \prod_{N \geq n \geq m} \big(1 - \mathbf{P}[A_n]\big).
\end{aligned}$$

Käytämme sitten arviota $1 - x \leq e^{-x}$, kun $x \geq 0$. Tällöin

$$\mathbf{P}\Bigg[\bigcap_{n\geq m} A_n^c\Bigg] \;\leq\; \exp\Bigg\{-\sum_{N\geq n\geq m} \mathbf{P}[A_n]\Bigg\} \;\to\; 0,$$

kun $N \to \infty$, sillä eksponenttiin tulee tällöin hajaantuvan sarjan jäännöstermi. Siten jokaisella $m \in \mathbb{N}$ pätee

$$\mathbf{P}\Bigg[\bigcap_{n\geq m} A_n^c\Bigg] \;=\; 0,$$

joten

$$\mathbf{P}[(\limsup A_n)^c] \;=\; \lim_{m\to\infty} \mathbf{P}\Bigg[\bigcap_{n\geq m} A_n^c\Bigg] \;=\; 0$$

eli $\mathbf{P}[\limsup A_n] = 1 - \mathbf{P}[\limsup A_n)^c] = 1 - 0 = 1$. □

5.2.2 Esimerkki. (i) Tarkastelemme loputonta nopanheittoa. Olkoon

$$A \;:=\; \text{“heittojen sarjassa on äärettömän monta kuutosta”}.$$

Jos A_n := “n. heitto on 6”, niin $A = \limsup A_n$. Selvästi tapahtumat A_n, $n \in \mathbb{N}$, ovat riippumattomia. Toisaalta $\mathbf{P}[A_n] = 1/6$, joten $\sum \mathbf{P}[A_n] = \infty$. Siten $\mathbf{P}[A] = 1$ Borelin Cantellin lemman 5.2.1(ii) nojalla.

(ii) Olkoon apina istutettu kirjoituskoneen ääreen. Apina painelee näppäimiä “täysin umpimähkään”. Shakespearen kootuissa on N kappaletta kirjaimia. Olkoon A_n tapahtuma “apina kirjoittaa Shakespearen kootut näppäinpainalluksilla $(n-1)N+1, \ldots, nN$. Todennäköisyys

$$p \;:=\; \mathbf{P}[A_n] \;=\; \left(\frac{1}{\text{näppäinten lukumäärä}}\right)^N$$

on häviävän pieni, mutta positiivinen. Nyt tapahtuma “apina kirjoittaa Shakespearen kootut äärettömän monta kertaa” sisältää tapahtuman $\limsup A_n$. Koska tapahtumat A_n, $n \in \mathbb{N}$, ovat riippumattomia, ja $\sum \mathbf{P}[A_n] = \sum p = \infty$, niin Borelin Cantellin lemman 5.2.1(ii) nojalla apina kirjoittaa melkein varmasti Shakespearen kootut äärettömän monta kertaa.

Seuraavaksi esitettävä *Kolmogorovin* 0–1*-laki* (lause 5.2.4) tarjoaa hauskemman tavan perustella esimerkin 5.2.2 kohdan (ii). Ennen itse tuloksen esittämistä tarvitsemme yhden käsitteen.

5.2.3 Määritelmä. Olkoon (X_n) jono satunnaismuuttujia. Merkitsemme $\mathcal{G}_n := \sigma(X_k; k > n)$, $n \in \mathbb{N}$. Tällöin σ-algebra

$$\mathcal{G}_\infty \;:=\; \cap \mathcal{G}_n$$

on jonon (X_n) virittämä *häntä-σ-algebra.*

5.2.4 Lause. *Olkoon (X_n) jono riippumattomia satunnaismuuttujia. Tällöin niiden virittämä hänta-σ-algebra $\mathcal{G}_\infty$ on triviaali: $\mathbf{P}[A] = 0$ tai $\mathbf{P}[A] = 1$ kaikilla $A \in \mathcal{G}_\infty$.*

Todistus. Olkoon $\mathcal{F}_n := \sigma(X_k; k \leq n)$ ja $\mathcal{G}_n := \sigma(X_k; k > n)$, $n \in \mathbb{N}$. Oletuksen nojalla σ-algebrat $\mathcal{F}_n$ ja $\mathcal{G}_n$ ovat riippumattomia. Siten, jos $F \in \mathcal{F}_n$ ja $G \in \mathcal{G}_n$, niin

$$\mathbf{P}[F \cap G] \;=\; \mathbf{P}[F]\mathbf{P}[G]. \tag{5.2.5}$$

Mutta jos $G \in \mathcal{G}_\infty$, niin (5.2.5) pätee kaikille $A \in \cup \mathcal{F}_n$. Siten, Dynkinin laajennuslauseen 2.3.8 nojalla, (5.2.5) pätee kaikilla $A \in \sigma(X_1, X_2, \ldots) =: \mathcal{F}_\infty$. Mutta $\mathcal{G}_\infty \subset \mathcal{F}_\infty$. Siten yhtälöstä (5.2.5) seuraa, että $\mathbf{P}[A] = \mathbf{P}[A]\mathbf{P}[A]$ kaikille $A \in \mathcal{G}_\infty$. Mutta tämä on mahdollista ainoastaa, jos $\mathbf{P}[A] = 0$ tai $\mathbf{P}[A] = 1$. □

5.2.6 Seuraus. *Olkoon (X_n) jono riippumattomia satunnaismuuttujia. Tällöin*

(i) *jono (X_n) joko suppenee tai hajaantuu melkein varmasti eli*

$$\mathbf{P}\big[\omega \in \Omega\,;\, \lim X_n(\omega) \text{ on olemassa}\big] \;=\; 0 \textit{ tai } 1,$$

(ii) *seuraavat satunnaismuuttujat ovat melkein varmasti vakioita*

$$\limsup X_n, \qquad \liminf X_n,$$
$$\limsup \frac{1}{n}\sum_{k\leq n} X_k, \qquad \liminf \frac{1}{n}\sum_{k\leq n} X_k.$$

5.3 Tuloavaruudet ja Fubinin lause

Tuloavaruudet vastaavat toistokokeita ja tulomitta vastaa toistokokeiden riippumattomuutta.

Olkoot $(\Omega_1, \mathcal{F}_1)$ ja $(\Omega_2, \mathcal{F}_2)$ kaksi mitallista avaruutta. Olkoon

$$\mathcal{F}_1 \times \mathcal{F}_2 \;:=\; \{A_1 \times A_2\,;\, A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\}$$

joukkoluokka tulojoukossa $\Omega_1 \times \Omega_2$. Ongelma on, että $\mathcal{F}_1 \times \mathcal{F}_2$ ei ole σ-algebra. Nimittäin joukon $A_1 \times A_2$ komplementtia ei voi kirjoittaa muodossa $B_1 \times B_2$, joillakin $B_1 \in \mathcal{F}_1$, $B_2 \in \mathcal{F}_2$. Itse asiassa

$$(A_1 \times A_2)^c = (A_1 \times A_2^c) \cup (A_1^c \times A_2) \cup (A_1^c \times A_2^c).$$

5.3.1 Määritelmä. *Tulo-σ-algebra* joukolla $\Omega_1 \times \Omega_2$ on

$$\mathcal{F}_1 \otimes \mathcal{F}_2 := \sigma\left(A_1 \times A_2\,;\, A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\right).$$

Avaruutta $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ kutsumme *tuloavaruudeksi.* Joukkoja $A_1 \times A_2$, $A_1 \in \mathcal{F}_1$, $A_2 \in \mathcal{F}_2$ kutsumme *mitallisiksi suorakaiteiksi.*

5.3.2 Esimerkki. Heitämme noppaa kaksi kertaa. Yhdelle nopanheitolle $\Omega_1 = \{1, \ldots, 6\}$ ja $\mathcal{F}_1 = \mathrm{pot}\,(\Omega)$. Kahdelle nopanheitolle $\Omega = \Omega_1^2 = \{(i,j); i,j = 1, \ldots, 6\}$ ja $\mathcal{F} = \mathcal{F}_1^{\otimes 2} = \mathrm{pot}\,(\Omega_1 \times \Omega_1)$. Tapahtuma “ensimmäinen heitto on 6 ja toinen heitto on parillinen” on mitallinen suorakaide, mutta tapahtuma “heittojen summa on parillinen” ei ole mitallinen suorakaide.

5.3.3 Apulause. *Olkoon X satunnaisvektori tuloavaruudelta $(\Omega_1 \times \Omega_2, \mathcal{F}_2 \otimes \mathcal{F}_2)$. Olkoon $\omega_1 \in \Omega_1$ kiinteä. Tällöin leikekuvaus $X(\omega_1, \cdot)$ on satunnaisvektori avaruudelta $(\Omega_2, \mathcal{F}_2)$. Vastaavasti leikekuvaus $X(\cdot, \omega_2)$ on satunnaisvektori avaruudelta $(\Omega_1, \mathcal{F}_1)$ kaikilla $\omega_2 \in \Omega_2$.*

Todistus. Todistamme väitteen vain satunnaismuuttujille. Vektoritapaus voidaan todistaa samalla tavalla tarkastelemalla satunnaisvektorin $X = (X_1, \ldots, X_d)$ komponentteja erikseen. Lisäksi on selvää, että riittää osoittaa leikekuvausta $X(\omega_1, \cdot)$ koskeva väite. Nimittäin leikettä $X(\cdot, \omega_2)$ koskeva väite on täysin symmetrinen.

Askel 1: Olkoon aluksi $X = \mathbf{1}_A$, missä $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$. Tarkastelemme joukkoperhettä

$$\mathcal{H} := \{A \in \mathcal{F}_1 \otimes \mathcal{F}_2\,;\, \mathbf{1}_A(\omega_1, \cdot) \text{ on } \mathcal{F}_2\text{-mitallinen kaikilla } \omega_1 \in \Omega_1\}.$$

Jos A on mitallinen suorakaide $A_1 \times A_2$, niin $\mathbf{1}_A(\omega_1, \omega_2) = \mathbf{1}_{A_1}(\omega_1)\mathbf{1}_{A_2}(\omega_2)$. Selvästi $\mathbf{1}_{A_1}(\omega_1)\mathbf{1}_{A_2}(\cdot)$ on $\mathcal{F}_2$-mitallinen. Siten

$$\mathcal{F}_1 \times \mathcal{F}_2 \subset \mathcal{H}.$$

Toisaalta joukkoperhe $\mathcal{H}$ on σ-algebra (harjoitustehtävä 5.10). Mutta, koska $\mathcal{H}$ sisältää mitalliset suorakaiteet, niin $\mathcal{F}_1 \otimes \mathcal{F}_2 \subset \mathcal{H}$. Toisaalta määritelmän perusteella on selvää, että $\mathcal{H} \subset \mathcal{F}_1 \otimes \mathcal{F}_2$. Siten $\mathcal{H} = \mathcal{F}_1 \otimes \mathcal{F}_2$.

Päättelemme, että lemman väite pätee satunnaismuuttujille, jotka ovat muotoa $\mathbf{1}_A$, $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$.

Askel 2: Olkoon sitten X yksinkertainen eli muotoa

$$X(\omega_1, \omega_2) \;=\; \sum_{k \le n} c_k \mathbf{1}_{A_k}(\omega_1, \omega_2),$$

missä $c_k \in \mathbb{R}$ ja $A_k \in \mathcal{F}_1 \otimes \mathcal{F}_2$. Askeleen 1 nojalla $X(\omega_1, \cdot)$ on $\mathcal{F}_2$-mitallisten kuvausten lineaarikombinaatio. Siten se on $\mathcal{F}_2$-mitallinen.

Askel 3: Olkoon sitten X positiivinen. Tällöin on olemassa sellainen jono (X_n) yksinkertaisia satunnaismuuttujia, että $X = \lim X_n$. Kun $\omega_1 \in \Omega$, niin $X(\omega_1, \cdot) = \lim X_n(\omega_1, \cdot)$. Siten, askeleen 2 nojalla, $X(\omega_1, \cdot)$ on $\mathcal{F}_2$-mitallisten kuvausten pisteittäisenä rajana $\mathcal{F}_2$-mitallinen.

Askel 4: Olkoon sitten lopuksi X yleinen. Väite seuraa tällöin askeleesta 3, sillä nyt $X(\omega_1, \cdot) = X^+(\omega_1, \cdot) - X^-(\omega_1, \cdot)$. □

5.3.4 Määritelmä. Olkoot $(\Omega_1, \mathcal{F}_1, \mathbf{P}_1)$ ja $(\Omega_2, \mathcal{F}_2, \mathbf{P}_2)$ todennäköisyysavaruuksia. Tuloavaruuden $(\Omega_1 \times \Omega_1, \mathcal{F}_1 \otimes \mathcal{F}_2)$ *tulomitta* $\mathbf{P}_1 \otimes \mathbf{P}_2$ määräytyy ehdosta

$$(\mathbf{P}_1 \otimes \mathbf{P}_2)[A_1 \times A_2] \;=\; \mathbf{P}_1[A_1]\, \mathbf{P}_2[A_2].$$

Laajennuslauseiden 2.3.8 ja 2.3.9 nojalla tulo-σ-algebralla $\mathcal{F}_1 \times \mathcal{F}_1$ on olemassa täsmälleen yksi tulomitta. Tulomitan suhteen σ-algebrat $\mathcal{F}_1$ ja $\mathcal{F}_2$, tulkittuna tulo-σ-algebran $\mathcal{F}_1 \otimes \mathcal{F}_2$ ali-σ-algebroina, ovat riippumattomia.

5.3.5 Esimerkki.
(i) Esimerkin 5.3.2 kahden nopanheiton tapauksessa tulomitta on $\mathbf{P}^{\otimes 2}[A] = |A|/36$. Tämä on nimittäin ainoa tapa saada heitot riippumattomiksi, kun $\mathbf{P}[A_i] = |A_i|/6$, missä A_i on heittoon $i = 1, 2$ liittyvä tapahtuma.
(ii) Olkoon $\Omega_1 = \Omega_2 = [0, 1]$, $\mathcal{F}_1 = \mathcal{F}_2 = \mathcal{B}_{[0,1]}$ välin $[0, 1]$ Borel-joukot sekä $\mathbf{P}_1 = \mathbf{P}_2 = \ell$ välin $[0, 1]$ Lebesguen mitta. Tällöin tulomitta $\ell^{\otimes 2}$ on tavallinen kaksiulotteinen Lebesguen mitta neliössä $[0, 1]^2$.

Seuraaava tulos 5.3.6 on kuuluisa *Fubinin lause*. Muuttujanvaihtokaavan (lause 4.3.1) ohella se on integraalilaskennan keskeisin kaava. Itse asiassa kaikki integrointikaavat voidaan johtaa kohtalaisen helposti käyttämällä ainoastaan muuttujanvaihtoa ja Fubinin lausetta.

5.3.6 Lause. *Olkoon X satunnaismuuttuja tuloavaruudelta $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ ja $\mathbf{P}_1 \otimes \mathbf{P}_2$ tulomitta. Tällöin: jos $X \ge 0$ tai $X \in L^1$, niin kuvaus*

$$\int_{\Omega_2} X(\cdot, \omega_2)\, \mathbf{P}_2[\mathrm{d}\omega_2]$$

on $\mathcal{F}_1$-mitallinen. Vastaavasti kuvaus

$$\int_{\Omega_1} X(\omega_1, \cdot)\, \mathbf{P}_1[\mathrm{d}\omega_1]$$

on $\mathcal{F}_2$-mitallinen. Lisäksi pätee Fubinin kaava

$$\begin{aligned}
&\int_{\Omega_1\times\Omega_2} X(\omega_1,\omega_2)\,(\mathbf{P}_1\otimes\mathbf{P}_2)[\mathrm{d}\omega_1,\mathrm{d}\omega_2] \\
&= \int_{\Omega_1}\left(\int_{\Omega_2} X(\omega_1,\omega_2)\,\mathbf{P}_2[\mathrm{d}\omega_2]\right)\mathbf{P}_1[\mathrm{d}\omega_1] \\
&= \int_{\Omega_2}\left(\int_{\Omega_1} X(\omega_1,\omega_2)\,\mathbf{P}_1[\mathrm{d}\omega_1]\right)\mathbf{P}_2[\mathrm{d}\omega_2],
\end{aligned}$$

missä jokainen rivi on olemassa (äärellisenä) täsmälleen samaan aikaan.

Todistus. Lauseen voi perustella apulauseella 5.3.3 ja aproksimoimalla satunnaismuuttujaa X yksinkertaisilla satunnaismuuttujilla sekä käyttämällä lopuksi monotonisen konvergenssin lausetta 4.2.1. Jätämme yksityiskohdat harjoitustehtäväksi 5.11. □

Itse asiassa Fubinin lause pätee myös niin sanotuille *σ-äärellisille mitoille.* Mitta $\mu : \mathcal{F} \to \mathbb{R}_+ \cup \{\infty\}$ on σ-äärellinen, jos on olemassa sellainen $(\Omega_n) \subset \mathcal{F}$, että $\mu[\Omega_n] < \infty$ kaikilla $n \in \mathbb{N}$ ja $\cup\Omega_n = \Omega$.

5.3.7 Seuraus. *Olkoot $X, Y \in L^1$ riippumattomia satunnaismuuttujia. Tällöin $XY \in L^1$ ja*

$$\mathbf{E}[XY] \;=\; \mathbf{E}[X]\mathbf{E}[Y]$$

Todistus. Tämä väite seuraa välittömästi Fubinin lauseesta 5.3.6 ja lauseesta 5.1.3(iv). Nimittäin (kertymäfunktioiden avulla kirjoitettuna)

$$\begin{aligned}
\mathbf{E}[XY] &= \int_{\mathbb{R}^2} xy\,\mathrm{d}F_{X,Y}(x,y) \\
&= \int_{\mathbb{R}^2} xy\,\mathrm{d}F_X(x)\mathrm{d}F_Y(y) \\
&= \int_{\mathbb{R}}\left(\int_{\mathbb{R}} xy\,\mathrm{d}F_X(x)\right)\mathrm{d}F_Y(y) \\
&= \int_{\mathbb{R}} y\left(\int_{\mathbb{R}} x\,\mathrm{d}F_X(x)\right)\mathrm{d}F_Y(y) \\
&= \int_{\mathbb{R}} x\,\mathrm{d}F_X(x)\int_{\mathbb{R}} y\,\mathrm{d}F_Y(y) \\
&= \mathbf{E}[X]\mathbf{E}[Y].
\end{aligned}$$

□

5.4 Äärettömät tuloavaruudet

Olkoot $(\Omega_n, \mathcal{F}_n, \mathbf{P}_n)$, $n \in \mathbb{N}$, todennäköisyysavaruuksia. Tarkastelemme nyt loputtomiin toistokokeisiin liittyvää tuloavaruutta $\Omega = \Omega_1 \times \Omega_2 \times \cdots$. Osion 2.4 esimerkin 2.4.1 perusteella on selvää, ettemme voi toivoa tulomittaa σ-algebralle, jonka virittää joukot $A_1 \times A_2 \times \cdots$, $A_n \in \mathcal{F}_n$. Joudumme siis tyytymään hieman karkeampaan σ-algebraan.

Esitämme nyt yleisen tuloavaruuden määritelmän.

Olkoon J joukko ja $(\Omega_j, \mathcal{F}_j, \mathbf{P}_j)$, $j \in J$, todennäköisyysavaruuksia.

Perusjoukkojen Ω_j, $j \in J$, karteesinen tuloavaruus $\times\Omega_j$ koostuu perheistä, tai kuvauksista, $\omega = (\omega_j)_{j \in J}$, missä $\omega_j \in \Omega_j$. Jos siis esimerkiksi $\Omega_j = \Omega_0$ on sama perusvaruus kaikilla $j \in J$, niin $\times\Omega_j = \Omega_0^J$. Siis tällöin $\omega \in \Omega_0^J$ on funktio $\omega : J \to \Omega_0$.

Ääretön tulo-σ-algebra $\otimes\mathcal{F}_j$ perusjoukolle $\times\Omega_j$ rakennetaan sylinterijoukkojen avulla. Joukko C on *sylinterijoukko*, jos se on muotoa

$$C \;=\; A_{j_1} \times A_{j_2} \times \cdots \times A_{j_k} \times \bigtimes_{j \neq j_1, \ldots, j_k} \Omega_j \tag{5.4.1}$$

joillakin $A_{j_l} \in \mathcal{F}_{j_l}$, $l \leq k$. Olkoon $\mathcal{C}$ sylinterijoukkojen luokka. Tällöin

$$\bigotimes \mathcal{F}_j \;:=\; \sigma(\mathcal{C}).$$

Ääretön tulomitta $\otimes\mathbf{P}_j$ määritellään myös sylinterijoukkojen kautta. Olkoon C, kuten edellä kaavassa (5.4.1). Tällöin asetamme

$$\otimes\mathbf{P}_j[C] \;:=\; \mathbf{P}_{j_1}[A_{j_1}] \cdots \mathbf{P}_{j_k}[A_{j_k}].$$

Laajennuslauseiden 2.3.8 ja 2.3.9 nojalla tämä todennäköisyys laajenee yksikäsitteisesti tulo-σ-algebralle $\otimes\mathcal{F}_j$. Lisäksi on huomattavaa, että todennäköisyyden $\otimes\mathbf{P}_j$ suhteen σ-algebrat $\mathcal{F}_j$, $j \in J$, ovat riippumattomia, kun ne tulkitaan tulo-σ-algebran $\otimes\mathcal{F}_j$ ali-σ-algebroiksi.

5.4.2 Määritelmä. Olkoot $(\Omega_j, \mathcal{F}_j, \mathbf{P}_j)$, $j \in J$. Todennäköisyysavaruuksia. Niiden *tuloavaruus* edellä määritelty kolmikko $(\times\Omega_j, \otimes\mathcal{F}_j, \otimes\mathbf{P}_j)$.

5.4.3 Esimerkki. Olkoon $I = \mathbb{N}$, $\Omega_n = \{0,1\}$, $\mathcal{F}_n = \mathrm{pot}\,(\{0,1\})$ ja $\mathbf{P}_n[\{0\}] = 1/2 = \mathbf{P}_n[\{1\}]$. Tällöin tuloavaruus

$$(\times\Omega_n, \otimes\mathcal{F}_n, \otimes\mathbf{P}_n) \;=\; (\Omega_1^\infty, \mathcal{F}_1^{\otimes\infty}, \mathbf{P}_1^{\otimes\infty})$$

vastaa loputonta reilun kolikon heittoa. Itse asiassa, jos merkitsemme

$$X_n \;:=\; \begin{cases} 1, & \text{jos } n. \text{ heitto on klaava,} \\ 0, & \text{jos } n. \text{ heitto on kruuna,} \end{cases}$$

niin
$$\bigotimes \mathcal{F}_n \;=\; \sigma(X_1, X_2, \ldots).$$
Ääretön tulo-σ-algebra on siis pienin σ-algebra, jonka suhteen koordinaattikuvaukset $\omega \mapsto X_n(\omega) = \omega_n$, $n \in \mathbb{N}$, ovat mitallisia.

Tulkitsemalla sopivasti esimerkkiä 2.4.1 näemme, että tulo-σ-algebra $\mathcal{F}_1^{\otimes\infty}$ on aidosti pienempi kuin täysi σ-algebra pot$\,(\Omega_1^\infty)$. Samoin näemme, että tulomittaa $\mathbf{P}_1^{\otimes\infty}$ ei voi laajentaa täydelle σ-algebralle pot$\,(\Omega_1^\infty)$. Jätämme näiden väitteiden tarkan perustelun harjoitustehtäväksi 5.13.

5.5 Riippumattomien satunnaismuuttujien summat

Tässä osiossa tarkastelemme ainoastaan satunnaismuuttujia eli $\mathbb{R}$-arvoisia satunnaisvektoreita. Vastaavat tulokset toki pätevät, luonnollisin muutoksin, myös satunnaisvektoreille.

Olkoot $X_1, \ldots, X_n$ riippumattomia satunnaismuuttujia. Tarkastelemme summan $X_1 + \cdots + X_n$ jakaumaa $\mathbf{P}_{X_1+\ldots+X_n}$. Vaihtelun vuoksi ilmoitamme tulokset kertymäfunktion
$$\begin{aligned} F_{X_1+\cdots+X_n}(y) &:= \mathbf{P}_{X_1+\cdots+X_n}\big[(-\infty, y]\big] \\ &= \mathbf{P}[X_1 + \cdots + X_n \le y] \end{aligned}$$
avulla.

5.5.1 Määritelmä. Olkoot $F_1, \ldots, F_n$ kertymäfunktioita. Niiden *konvoluutio* $F_1 * \cdots * F_n$ määräytyy rekursiivisesti kaavoista
$$\begin{aligned} F_1 * F_2(y) &:= \int_{\mathbb{R}}\int_{\mathbb{R}} \mathbf{1}_{(-\infty,y]}(x_1 + x_2)\, \mathrm{d}F_1(x_1)\mathrm{d}F_2(x_2), \\ F_1 * \cdots * F_n(y) &:= (F_1 * \cdots * F_{n-1}) * F_n(y) \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} \mathbf{1}_{(-\infty,y]}(x_1 + x_2)\, \mathrm{d}F_1 * \cdots * F_{n-1}(x_1)\mathrm{d}F_n(x_2). \end{aligned}$$

Konvoluution määritelmässä on huomattavaa, että se on symmetrinen: $F_1 * \cdots * F_n = F_{\pi(1)} * \cdots * F_{\pi(n)}$, missä $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$ on mikä tahansa joukon $\{1, \ldots, n\}$ permutaatio. Lauseen 5.5.3 jälkeen tämä väite on täysin selvä.

5.5.2 Esimerkki. Olkoon
$$F_1(x) \;=\; F_2(x) \;=\; x\mathbf{1}_{[0,1]}(x) + \mathbf{1}_{(1,\infty)}(x).$$

Tällöin

$$\mathrm{d}F(x) \;=\; \mathbf{1}_{[0,1]}(x)\mathrm{d}x$$

ja pienellä geometrisella päättelyllä näemme, että

$$\begin{aligned} F^{*2}(y) &:= F * F(y) \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} \mathbf{1}_{(-\infty,y]}(x_1+x_2)\,\mathbf{1}_{[0,1]}(x_1)\mathrm{d}x_1\mathbf{1}_{[0,1]}(x_2)\mathrm{d}x_2 \\ &= \ell\Big[(x,y)\in\mathbb{R}^2\,;\,0\le x_1,x_2\le 1, x_1+x_2\le y\Big] \\ &= \frac{y^2}{2}\mathbf{1}_{[0,1]}(y) + \left(1-\frac{(2-y)^2}{2}\right)\mathbf{1}_{[1,2]}(y). \end{aligned}$$

Kertymäfunktioiden konvoluutio on aina kertymäfunktio. Tämä seuraa välittömästi seuraavasta lauseesta.

5.5.3 Lause. *Olkoot $X_1,\ldots,X_n$ riippumattomia satunnaismuuttujia kertymäfunktioinaan $F_1,\ldots,F_n$. Tällöin summan $X_1+\cdots+X_n$ kertymäfunktio on $F_1*\cdots*F_n$.*

Todistus. Todistamme vain tapauksen $n=2$. Yleinen tapaus seuraa helposti induktiolla käyttämällä konvoluution rekursiivista määritelmää.

Koska X_1 ja X_2 ovat riippumattomia, niin lauseen 5.1.3(iv) nojalla $F_{X_1,X_2}(x_1,x_2) = F_1(x_1)F_2(x_2)$. Siten kaikille rajoitetuille Borel-funktioille $g:\mathbb{R}^2\to\mathbb{R}$ pätee

$$\mathbf{E}[g(X_1,X_2)] \;=\; \int_{\mathbb{R}}\int_{\mathbb{R}} g(x_1,x_2)\,\mathrm{d}F_1(x_1)\mathrm{d}F_2(x_2).$$

Erityisesti valitsemalla $g(x_1,x_2)=f(x_1+x_2)$ saamme

$$\mathbf{E}[f(X_1+X_2)] \;=\; \int_{\mathbb{R}}\int_{\mathbb{R}} f(x_1+x_2)\,\mathrm{d}F_1(x_1)\mathrm{d}F_2(x_2).$$

Väite seuraa nyt valitsemalla $f(x)=\mathbf{1}_{(-\infty,y]}(x)$, sillä tällöin

$$\begin{aligned} F_{X_1+X_2}(y) &= \mathbf{P}[X_1+X_2\le y] \\ &= \mathbf{E}\big[\mathbf{1}_{\{X_1+X_2\le y\}}\big] \\ &= \mathbf{E}\big[f(X_1+X_2)\big]. \end{aligned}$$

$\square$

Esimerkissä 5.5.2 laskimme siis kahden riippumattoman tasaisesti välille $[0,1]$ jakautuneen satunnaismuuttujan summan kertymäfunktion.

Määritelmä 5.5.1 koskee kertymäfunktioita eli mittoja. Funktioiden $f_1,\ldots,f_n : \mathbb{R} \to \mathbb{R}$ konvoluutio määritellään asettamalla

$$
\begin{aligned}
f_1 * f_2(y) &:= \int_{-\infty}^{\infty} f_1(y-x_2) f_2(x_2)\,\mathrm{d}x_2,\\
f_1 * \cdots * f_n(y) &:= (f_1 * \cdots f_{n-1}) * f_n(y)\\
&= \int_{-\infty}^{\infty} f_1 * \cdots * f_{n-1}(y-x_n) f_n(x_n)\,\mathrm{d}x_n.
\end{aligned}
$$

Samoin lukujonojen $p_{1,\cdot},\ldots,p_{n,\cdot}$ konvoluutio määritellään asettamalla

$$
\begin{aligned}
(p_{1,\cdot} * p_{2,\cdot})_k &:= \sum_l p_{1,k-l} p_{2,l},\\
(p_{1,\cdot} * \cdots * p_{n,\cdot})_k &:= (p_{1,\cdot} * \cdots * p_{n-1,\cdot}) * p_{n,\cdot}\\
&= \sum_l (p_{1,\cdot} * \cdots * p_{n-1,\cdot})_{k-l} p_{n,l}.
\end{aligned}
$$

Seuraava lauseen 5.5.3 seuraus valaisee näiden eri konvoluutiokäsitteiden yhteyttä.

5.5.4 Seuraus. *Olkoot $X_1,\ldots,X_n$ riippumattomia satunnaismuuttujia.*

(i) *Jos $X_1,\ldots,X_n$ ovat diskreettejä ja $p_{l,k} := \mathbf{P}[X_l = k]$, niin summa $X_1+\cdots+X_n$ on diskreetti ja $\mathbf{P}[X_1+\cdots+X_n = k] = p_{1,\cdot} * \cdots * p_{n,\cdot}$.*

(ii) *Jos $X_1,\ldots,X_n$ ovat jatkuvia tiheysfunktioinaan $f_1,\ldots,f_n$, niin summa $X_1+\cdots+X_n$ on jatkuva ja sen tiheysfunktio on $f_1 * \cdots * f_n$.*

Todistus. Harjoitustehtävä 5.14 □

5.5.5 Esimerkki. Olkoot X_1 ja X_2 riippumattomia Poisson-jakautuneita satunnaismuuttujia parametrein λ_1 ja λ_2. Toisin sanoen ne ovat diskreetisti jakautuneita satunnaismuuttujia pistetodennäköisyysfunktionaan

$$
\begin{aligned}
p_{1,k} &:= \mathbf{P}[X_1 = k] &= e^{-\lambda_1}\frac{\lambda_1^k}{k!},\\
p_{1,k} &:= \mathbf{P}[X_2 = k] &= e^{-\lambda_2}\frac{\lambda_2^k}{k!},
\end{aligned}
$$

missä $k = 0, 1, \ldots$. Tällöin, käyttämällä Newtonin binomikaavaa, saamme

$$\begin{aligned}
\mathbf{P}[X_1 + X_2 = k] &= \sum_l p_{1,k-l} p_{2,l} \\
&= \sum_l e^{-\lambda_1} \frac{\lambda_l^{k-l}}{(k-l)!} e^{-\lambda_2} \frac{\lambda_2^l}{l!} \\
&= e^{-(\lambda_1+\lambda_2)} \sum_l \frac{1}{(k-l)!l!} \lambda_1^{k-l} \lambda_2^l \\
&= e^{-(\lambda_1+\lambda_2)} \frac{1}{k!} \sum_l \binom{n}{k} \lambda_1^{k-1} \lambda_2^k \\
&= e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^k}{k!}.
\end{aligned}$$

Näemme, että $X_1 + X_2$ on Poisson-jakautunut parametrilla $\lambda_1 + \lambda_2$.

Yleisemmin näemme induktiolla, että jos $X_1, \ldots, X_n$ ovat riippumattomia Poisson-jakautuneita satunnaismuuttujia parametrein $\lambda_1, \ldots, \lambda_n$, niin summa $X_1 + \cdots + X_n$ on myös Poisson-jakautunut parametrilla $\lambda_1 + \cdots + \lambda_n$.

Harjoitustehtäviä

5.1. Anna esimerkki tapahtumista, jotka ovat kausaalisesti riippuvia, mutta tilastollisesti riippumattomia.

5.2. Anna esimerkki kokoelmasta $\mathcal{G}_j$, $j \in J$, σ-algebroja, joka on pareittain riippumaton, muttei riippumaton.

5.3. Osoita, että satunnaisvektori X on riippumaton itsestää jos ja vain jos se on melkein varmasti vakio eli on olemassa sellainen $x \in \mathbb{R}^d$, että $\mathbf{P}[X = x] = 1$.

5.4. Todista lauseen 5.1.3 implikaatio (v) $\Rightarrow$ (ii).

5.5. Olkoot $\mathcal{F}_1$ ja $\mathcal{F}_2$ riippumattomia σ-algebroja. Olkoon lisäksi $\mathcal{F}'_1 \subset \mathcal{F}_1$ ali-σ-algebra ja $\mathcal{F}'_2 \subset \mathcal{F}_2$ ali-σ-algebra. Osoita, että σ-algebrat $\mathcal{F}'_1$ ja $\mathcal{F}'_2$ ovat riippumattomia

5.6. Todista lauseen 5.1.3 implikaatio (vi) $\Rightarrow$ (ii).

5.7. Muotoile ja todista lauseen 5.1.3 vastine, joka koskee perhettä $X_j = (X_{j,1}, \ldots, X_{j,d_j})$, $j \in J$, satunnaisvektoreita.

5.8. Perustele esimerkki 5.2.2(ii) käyttämällä Kolmogorovin 0–1-lakia 5.2.4.

5.9. Osoita, että $\mathcal{B}_d = \mathcal{B}^{\otimes d}$.

5.10. Osoita, että lemman 5.3.3 todistuksessa esiintyvä joukkoluokka

$$\mathcal{H} \;:=\; \{A \in \mathcal{F}_1 \otimes \mathcal{F}_2\,;\, \mathbf{1}_A(\omega_1, \cdot) \text{ on } \mathcal{F}_2\text{-mitallinen kaikilla } \omega_1 \in \Omega_1\}$$

on σ-algebra.

5.11. Todista Fubinin lause 5.3.6.

5.12. Olkoot $X, Y, XY \in L^1$. Osoita, että ehdosta

$$\mathbf{E}[XY] \;=\; \mathbf{E}[X]\mathbf{E}[Y]$$

ei seuraa, että satunnaismuuttujat X ja Y ovat riippumattomia.

5.13. Tarkastelemme esimerkkiä 5.4.3.

(a) Osoita, että tulo-σ-algebra on aidosti täyttä σ-algebraa $\operatorname{pot}(\{0,1\}^\infty)$ pienempi.
(b) Osoita, että tulomittaa $\mathbf{P}_1^{\otimes\infty}$ ei voida laajentaa täydelle σ-algebralle $\operatorname{pot}(\{0,1\}^\infty)$.

5.14. Todista seuraus 5.5.4.

5.15. Olkoot X ja Y riippumattomia $\mathbb{N}$-arvoisia satunnaismuuttujia ja

$$\mathbf{P}[X = n] \;=\; \mathbf{P}[Y = n] \;=\; \frac{1}{2^n}.$$

Laske seuraavat todennäköisyydet:

(a) $\mathbf{P}[\min(X, Y) \le n]$,
(b) $\mathbf{P}[X = Y]$,
(c) $\mathbf{P}[X > Y]$,
(d) $\mathbf{P}[X \text{ jakaa } Y\!:\!\text{n}]$.

5.16. Olkoot $X_1, \ldots, X_n$ riippumattomia satunnaismuuttujia. Oletamme, että niillä on samat odotusarvot ja varianssit. Merkitsemme $\mu := \mathbf{E}[X_k]$ ja $\sigma^2 := \mathbf{Var}[X_k]$, $k \le n$. Olkoon

$$\bar{X} \;:=\; \frac{1}{n}\sum_{k\le n} X_k \quad \text{ja} \quad S^2 \;:=\; \frac{1}{n}\sum_{k\le n}\big(X_k - \bar{X}\big)^2.$$

Osoita, että

(a) $\mathbf{E}[\bar{X}] = \mu$,
(b) $\mathbf{Var}[\bar{X}] = \sigma^2/n$,
(c) $\mathbf{E}[S^2] = \sigma^2\,(n-1)/n$.

Miten tämä tulos liittyy tilastolliseen estimointiin?

5.17. Olkoot X_1 ja X_2 riippumattomia normaalisti jakautuneita satunnaismuuttujia odotusarvoinaan μ_1 ja μ_2 ja variansseinaan σ_1 ja σ_2. Toisin sanoen satunnaismuuttujan X_i, $i = 1, 2$, tiheysfunktio on

$$f_i(x) \quad = \quad \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{1}{2}\left(\frac{x-\mu_i}{\sigma_i}\right)^2}.$$

Laske satunnaismuuttujan $X_1 + X_2$ tiheysfunktio.

5.18. Olkoot $X_1, \ldots, X_n$ satunnaismuuttujia ja $g : \mathbb{R}^n \to \mathbb{R}$ Borel-kuvaus. Esitä satunnaismuuttujan $g(X_1, \ldots, X_n)$ jakauma integraalina yhteiskertymäfunktion $F_{X_1,\ldots,X_n}$ avulla.

5.19. Olkoot X_1 ja X_2 riippumattomia satunnaismuuttujia kertymäfunktioinaan F_{X_1} ja F_{X_2}. Esitä seuraavien satunnaismuuttujien kertymäfunktiot kertymäfuntioiden F_{X_1} ja F_{X_2} avulla:

(a) $\max(X_1, X_2)$,
(b) $\min(X_1, X_2)$,
(c) $X_1 X_2$,
(d) X_1/X_2.

5.20. Laske harjoitustehtävän 5.19 satunnaismuuttujien muunnosten tiheysfunktiot, kun oletamme että X_1 ja X_2 ovat jatkuvia tiheysfunktioinaan f_{X_1} ja f_{X_2}.

Luku 6

Karakteristiset funktiot

Karakteristiset funktiot ovat teknisiä työkaluja, joiden avulla voimme tarkastella jakaumia. Erityisen hyödyllisiä ne ovat riippumattomien satunnaismuuttujien summan jakauman tarkastelussa. Tällä kurssilla käytämme karakteristisia funktioita lähinnä keskeisen raja-arvolauseen todistamiseen.

Tämä luku on varsin tekninen, mutta karakteristisiin funktioihin perustuvat tekniikat ovat niin tehokkaita, että niiden opiskeluun kannattaa kuluttaa aikaa.

6.1 Kompleksiluvut ja -funktiot

Esitämme varsin lyhyen johdannon kompleksianalyysiin. Tarkemmin asiasta kiinnostuneet lukekoot funktioteoriaa.

Vektoriavaruutena (eli summan suhteen) ja topologisena avaruutena (eli normin suhteen) kompleksitaso $\mathbb{C}$ vastaa euklidista tasoa $\mathbb{R}^2$. Lukualgebrana (eli otetaan tulot summien lisäksi) $\mathbb{C}$ on jotain aivan muuta kuin $\mathbb{R}^2$; itse asiassa euklidisessa avaruudessa $\mathbb{R}^2$ ei edes ole mitään luonnollista tuloa.

Kompleksiluku $z \in \mathbb{C}$ on muotoa $z = x + iy$, missä $i := \sqrt{-1}$; $x \in \mathbb{R}$ on kompleksiluvun $z \in \mathbb{C}$ reaaliosa ja $y \in \mathbb{R}$ on sen imaginääriosa. Kompleksiluvun normi eli *moduli* on

$$|z| \quad := \quad \sqrt{x^2 + y^2}$$

ja sen vaihekulma eli *argumentti* $\arg z \in (-\pi, \pi]$ määräytyy ehdoista

$$\cos(\arg z) \quad = \quad \frac{x}{\sqrt{x^2+y^2}} \qquad \text{ja} \qquad \sin(\arg z) \quad = \quad \frac{y}{\sqrt{x^2+y^2}}.$$

Tällöin parin $(x, y) \in \mathbb{R}^2$ napakoordinaattiesitys on $(|z|, \arg z) \in \mathbb{R}_+ \times (-\pi, \pi]$ eli

$$x \quad = \quad |z| \cos(\arg z) \qquad \text{ja} \qquad y \quad = \quad |z| \sin(\arg z).$$

Kompleksilukujen $z_1 = (x_1, y_1)$ ja $z_2 = (x_2, y_2)$ summa määritellään komponenteittain:

$$z_1 + z_2 \;:=\; (x_1 + x_2) + i(y_1 + y_2).$$

Kompleksilukujen tulo määritellään kertomalla parit $z_1 = (x_1, y_1)$ ja $z_2 = (x_2, y_2)$ formaalisti komponenteittain sekä käyttämällä kaavaa $i^2 = -1$:

$$\begin{aligned} z_1 z_2 &= (x_1 + iy_1)(x_2 + iy_2) \\ &= x_1x_2 + ix_1y_2 + iy_1x_2 + i^2y_1y_2 \\ &= (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1). \end{aligned}$$

Kompleksiluvun z potenssi z^n, $n \in \mathbb{N} \cup \{0\}$, määritellään luonnollisella tavalla:

$$\begin{aligned} z^0 &:= 1, \\ z^n &:= z^{n-1}\, z. \end{aligned}$$

Tällöin kaikille $n, m \in \mathbb{N} \cup \{0\}$ pätee $z^n z^m = z^{n+m}$.

Jos $z \neq 0$, on sen käänteisluku

$$\frac{1}{z} \;:=\; \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2}.$$

Tällöin nimittäin pätee $z\frac{1}{z} = 1$ (harjoitustehtävä 6.1). Jakolasku määritellään sitten asettamalla

$$\frac{z_1}{z_2} \;:=\; z_1 \frac{1}{z_2},$$

kun $z_2 \neq 0$. Jos vielä asetamme $z^{-n} := 1/z^n$, niin kaava $z^n z^m = z^{n+m}$ pätee kaikilla $n, m \in \mathbb{Z}$, kun $z \neq 0$.

Tunnetusti reaalinen eksponenttifunktio e^x, $x \in \mathbb{R}$, voidaan esittää sarjana $e^x = \sum_{n \geq 0} x^n/n!$. Funktioteorian avulla voidaan osoittaa, että vastaava sarja suppenee (itseisesti eli modulin suhteen), kun reaaliluku $x \in \mathbb{R}$ korvataan kompleksiluvulla $z \in \mathbb{C}$. Tämä johtaa *kompleksisen eksponenttifunktion* määritelmään

$$e^z \;:=\; \sum_{n \geq 0} \frac{z^n}{n!}. \tag{6.1.1}$$

Käyttämällä *Cauchyn kertokaavaa* ja *Newtonin binomikaavaa* saamme

$$\begin{aligned} e^{z_1}e^{z_2} &= \sum_{n\ge 0}\frac{z_1^n}{n!}\sum_{m\ge 0}\frac{z_2^m}{m!} \\ &= \sum_{n\ge 0}\sum_{0\le m\le n}\frac{z_1^{n-m}z_2^m}{(n-m)!m!} \\ &= \sum_{n\ge 0}\frac{1}{n!}\sum_{0\le m\le n}\binom{n}{m} z_1^{n-m}z_2^m \\ &= \sum_{n\ge 0}\frac{(z_1+z_2)^n}{n!}. \end{aligned}$$

Tästä näemme, että

(6.1.2) $$e^{z_1+z_2} = e^{z_1}e^{z_2}.$$

Kaavasta (6.1.2) seuraa välittömästi, että $e^z = e^{x+iy} = e^x e^{iy}$. Mutta

$$\begin{aligned} e^{iy} &= \sum_{n\ge 0}\frac{(iy)^n}{n!} \\ &= \sum_{n\ge 0}(-1)^n\frac{y^{2n}}{(2n)!} + i\sum_{n\ge 0}(-1)^n\frac{y^{2n+1}}{(2n+1)!}. \end{aligned}$$

Tunnistamalla tästä kosinin ja sinin sarjakehitelmät

$$\begin{aligned} \cos y &= \sum_{n\ge 0}(-1)^n\frac{y^{2n}}{(2n)!}, \\ \sin y &= \sum_{n\ge 0}(-1)^n\frac{y^{2n+1}}{(2n+1)!} \end{aligned}$$

ja käyttämällä kaavaa (6.1.2) saamme *Eulerin kaavan*

$$e^z := e^x\big(\cos x + i\sin y\big).$$

Eulerin kaavasta puolestaan seuraa, että kompleksiluvulla $z = x+iy$ on napakoordinaattiesitys $z = re^{i\theta}$, missä $r = |z|$ ja $\theta = \arg z$. Lisäksi näemme, että $|e^z| = e^x$, $e^{z+2n\pi i} = e^z$ kaikilla $n\in\mathbb{Z}$ ja että kuvaus $z\mapsto e^z$ on jatkuva (modulin suhteen). Lisäksi Eulerin kaavasta näemme, että kompleksisen eksponenttifunktion $z\mapsto e^z$ kuvajoukko on $\mathbb{C}\setminus\{(0,0)\}$. Kompleksinen eksponenttifunktio ei siis ole "positiivinen" toisin kuin sen reaalinen vastine. Esimerkiksi -1 on kompleksisen eksponenttifunktion kuvajoukossa. Nimittäin "matematiikan kaunein kaava" sanoo, että $e^{i\pi}+1=0$.

Seuraavat kaavat "imaginääriselle" eksponenttifunktiolle $x\mapsto e^{ix}$, $x\in\mathbb{R}$, ovat jatkossa hyödyllisiä.

6.1.3 Apulause. *Olkoon* $x, y, u \in \mathbb{R}$. *Tällöin*

(i) $|e^{ix} - 1| \leq |x|$,
(ii) $\frac{\partial}{\partial u} e^{iux} = ixe^{iux}$.

Todistus. (i) Osoitamme, että $|e^{ix} - 1|^2 \leq x^2$. Nyt

$$\begin{aligned}\left|e^{ix} - 1\right|^2 &= \left|(\cos x - 1) + i \sin x\right|^2 \\ &= (\cos x - 1)^2 + \sin^2 x \\ &= \cos^2 x - 2\cos x + 1 + \sin^2 x \\ &= 2 - 2\cos x.\end{aligned}$$

Pitää siis osoittaa, että

$$1 - \cos x \leq \frac{x^2}{2}.$$

Mutta tämän epäyhtälön näkee esimerkiksi derivoimalla puolittain ja käyttämällä sitten epäyhtälöä $\sin x \leq x$, jonka puolestaan näkee niinikään derivoimalla puolittain ja käyttämällä sitten epäyhtälöä $\cos x \leq 1$.

(ii) Tämä on suora lasku:

$$\begin{aligned}\frac{\partial}{\partial u} e^{iux} &= \frac{\partial}{\partial u}\big(\cos(ux) + i \sin(ux)\big) \\ &= \frac{\partial}{\partial u}\cos(ux) + i\frac{\partial}{\partial u}\sin(ux) \\ &= -x\sin(ux) + ix\cos(ux) \\ &= ix\, i\sin(ux) + ix\cos(ux) \\ &= ix\big(\cos(ux) + i\sin(ux)\big) \\ &= ixe^{iux}.\end{aligned}$$

□

Lopuksi määrittelemme kompleksisen logaritmifunktion kompleksisen eksponenttifunktion käänteisfunktiona. Tulee siis ratkaista yhtälö $e^w = z$, $z \in \mathbb{C} \setminus \{(0,0)\}$, muuttujan $w \in \mathbb{C}$ suhteen. Koska kompleksinen eksponenttifunktio on jaksollinen, on ratkaisuja ääretön määrä. Siten meidän on valittava jokin ratkaisuhaara. Ratkaisun "päähaara"

$$z \mapsto w(z) = \ln|z| + i \arg z =: \ln z, \qquad z \in \mathbb{C} \setminus \{(0,0)\},$$

on *kompleksinen logaritmifunktio*.

6.2 Kompleksiarvoiset satunnaisvektorit

Kompleksiarvoinen satunnaisvektori on muotoa $Z = X + iY$, missä X ja Y ovat reaalisia satunnaisvektoreita.

Kompleksiarvoiset satunnaisvektorit $Z_j = X_j + iY_j$, $j \in J$, ovat riippumattomia, jos satunnaisvektorit X_j, Y_j, $j \in J$, ovat riippumattomia.

Kompleksiarvoinen satunnaismuuttuja $Z = X + iY$ on integroituva, jos sen reaali- ja imaginääriosat X ja Y ovat erikseen integroituvia satunnaismuuttujia. Tällöin määrittelemme

$$\mathbf{E}[Z] \ := \ \mathbf{E}[X] + i\mathbf{E}[Y].$$

6.2.1 Apulause. *Olkoot Z_1 ja Z_2 integroituvia kompleksiarvoisia satunnaismuuttujia ja $z \in \mathbb{C}$. Tällöin*

(i) $\mathbf{E}[zZ_1] = z\mathbf{E}[Z_1]$,
(ii) $\mathbf{E}[Z_1 + Z_2] = \mathbf{E}[Z_1] + \mathbf{E}[Z_2]$,
(iii) *jos Z_1 ja Z_2 ovat riippumattomia, niin Z_1Z_2 on integroituva ja* $\mathbf{E}[Z_1Z_2] = \mathbf{E}[Z_1]\mathbf{E}[Z_2]$,
(iv) $|\mathbf{E}[Z_1]| \leq \mathbf{E}[|Z_1|]$.

Todistus. Lauseen 4.1.15 valossa kohdat (i) ja (ii) ovat triviaaleja. Todistamme ne kuitenkin valaistaksemme kompleksiluvuilla operointia.

(i) Väite seuraa hajottamalla Z_1 ja z reaali- ja imaginääriosiin ja käyttämällä vastaavaa reaaliarvoisten satunnaismuuttujien tulosta 4.1.15(i). Nimittäin

$$\begin{aligned}
\mathbf{E}[zZ_1] &= \mathbf{E}\big[(x+iy)(X_1+iY_1)\big] \\
&= \mathbf{E}\big[xX_1 + ixY_1 + iyX_1 - yY_1\big] \\
&= \mathbf{E}\big[(xX_1 - yY_1) + i(xY_1 + yX_1)\big] \\
&= \mathbf{E}\big[xX_1 - yY_1\big] + i\mathbf{E}\big[xY_1 + yX_1\big] \\
&= x\mathbf{E}[X_1] - y\mathbf{E}[Y_1] + ix\mathbf{E}[Y_1] + y\mathbf{E}[X_1] \\
&= z\mathbf{E}[X_1] + ix\mathbf{E}[Y_1] + iy\mathbf{E}[X_1] - y\mathbf{E}[Y_1] \\
&= (x+iy)\left(\mathbf{E}[X_1] + i\mathbf{E}[Y_1]\right) \\
&= z\mathbf{E}[Z_1].
\end{aligned}$$

(ii) Tämä väite seuraa täsmälleen samalla tekniikalla kuin kohta (i). Nimittäin

$$
\begin{aligned}
\mathbf{E}[Z_1 + Z_2] &= \mathbf{E}[(X_1 + iY_1) + (X_2 + iY_2)] \\
&= \mathbf{E}[(X_1 + X_2) + i(Y_1 + Y_2)] \\
&= \mathbf{E}[X_1 + X_2] + i\mathbf{E}[Y_1 + Y_2] \\
&= \big(\mathbf{E}[X_1] + \mathbf{E}[X_2]\big) + i\big(\mathbf{E}[Y_1] + \mathbf{E}[Y_2]\big) \\
&= \big(\mathbf{E}[X_1] + i\mathbf{E}[Y_1]\big) + \big(\mathbf{E}[X_2] + i\mathbf{E}[Y_2]\big) \\
&= \mathbf{E}[Z_1] + \mathbf{E}[Z_2]
\end{aligned}
$$

(iii) tämän kohdan jätämme harjoitustehtäväksi 6.2.

(iv) Huomaamme aluksi, että väite on itse asiassa muotoa

$$
\sqrt{\mathbf{E}[X_1]^2 + \mathbf{E}[Y_1]^2} \;\leq\; \mathbf{E}\left[\sqrt{X_1^2 + Y_1^2}\right].
$$

Mutta tämä epäyhtälö seuraa välittömästi Jensenin epäyhtälöstä 4.4.1, sillä kuvaus $(x_1, x_2) \mapsto f(x_1, x_2) := \sqrt{x_1^2 + x_2^2}$ on konveksi (konveksisuuden näkee joko suoralla visualisoinnilla tai laskemalla Hessin matriisin

$$
H(x_1, x_2) \;:=\; H[f](x_1, x_2) \;:=\; \left[\frac{\partial^2}{\partial x_k \partial x_l} f(x_1, x_2)\right]_{k,l \leq 2}
$$

ja toteamalla, että se on positiivisesti semidefiniitti eli

$$
\sum_{k,l \leq 2} y_k \, H_{kl}(x_1, x_2) \, y_l \;\geq\; 0
$$

kaikilla $(x_1, x_2), (y_1, y_2) \in \mathbb{R}^2$). □

6.3 Karakteristinen funktio ja sen ominaisuudet

Karakteristinen funktio on (integrandin eksponentin etumerkkiä ja vakiota $(2\pi)^{-d/2}$ vaille) se kuvaus, jota analyysin puolella kutsutaan tyypillisesti Fourier–Stieltjes-muunnokseksi. Analyytikkojen karakteristinen funktio puolestaan on stokastikkojen indikaattori, jolle analyytikot käyttävät tyypillisesti merkintää χ_A stokastikkojen merkinnän $\mathbf{1}_A$ sijaan.

Karakteristisen funktion hyödyllisyys (tämän kurssin näkökulmasta) perustuu lähinnä kahteen (myöhemmin perusteltavaan) seikkaan:

1. satunnaisvektorin karakteristinen funktio määrää täysin (eli karakterisoi) sen jakauman,

2. riippumattomien satunnaisvektorien summan karakteristinen funktio on summattavien satunnaisvektorien karakteristisen funktioiden tulo.

6.3.1 Määritelmä. Satunnaisvektorin $X = (X_1, \ldots, X_d)$ *karakteristinen funktio* $\phi_X : \mathbb{R}^d \to \mathbb{C}$ on

$$\begin{aligned}\phi_X(u) &:= \mathbf{E}\left[e^{i\,u\cdot X}\right] \\ &= \mathbf{E}[\cos(u\cdot X)] + i\,\mathbf{E}[\sin(u\cdot X)]\,,\end{aligned}$$

missä $u{\cdot}X := \sum_{k\le d} u_k X_k$ on euklidisen avaruuden $\mathbb{R}^d$ tavanomainen sisätulo.

Karakteristinen funktio ϕ_X riippuu siis satunnaisvektorista X ainoastaan sen jakauman $\mathbf{P}_X$ kautta: jos satunnaisvektoreilla X ja Y on sama jakauma, eli $\mathbf{P}_X = \mathbf{P}_Y$ tai yhtä hyvin $F_X = F_Y$, niin $\phi_X = \phi_Y$. Tämä on selvää suoraan määritelmän perusteella. Nimittäin yhteiskerymäfunktion F_X avulla kirjoitettuna karakteristinen funktio ϕ_X on Riemann–Stieltjes-integraali

$$\begin{aligned}\phi_X(u) &= \int_{\mathbb{R}^d} e^{i\,u\cdot x}\,\mathrm{d}F_X(x) \\ &= \int_{\mathbb{R}^d} e^{i\sum_{k\le d} u_k x_k}\,\mathrm{d}F_X(x_1,\ldots,x_d).\end{aligned}$$

Erityisesti siis jatkuvasti jakautuneille satunnaisvektoreille X pätee

$$\begin{aligned}\phi_X(u) &= \int_{\mathbb{R}^d} e^{i\,u\cdot x} f_X(x)\,\mathrm{d}x \\ &= \int_{\mathbb{R}^d} e^{i\sum_{k\le d} u_k x_k} f_X(x_1,\ldots,x_d)\,\mathrm{d}x_1\cdots\mathrm{d}x_d,\end{aligned}$$

missä f_X on satunnaisvektorin X yhteistiheysfunktio. Fourier-analyysiä tuntevat tunnistanevat, että tällöin karakteristinen funktio ϕ_X on (vakiota $(2\pi)^{-d/2}$ vaille) yhteistiheysfuntion f_X Fourier-muunnos.

Myöhemmin näemme (seuraus 6.3.9), että myös käänteinen väite pätee: jos $\phi_X = \phi_Y$, niin $\mathbf{P}_X = \mathbf{P}_Y$.

6.3.2 Esimerkki. (i) Olkoon X Bernoulli-jakautunut parametrilla p. Toisin sanoen $\mathbf{P}[X = 1] = p$ ja $\mathbf{P}[X = 0] = (1 - p)$. Tällöin

$$\begin{aligned}\phi_X(u) &= e^{iu0}(1-p) + e^{iu}p \\ &= pe^{iu} + 1 - p.\end{aligned}$$

(ii) Olkoon X Poisson-jakautunut parametrilla λ. Tällöin

$$\begin{aligned}\phi_X(u) &= e^{-\lambda}\sum_{k\geq 0} e^{iuk}\frac{\lambda^k}{k!}\\ &= e^{-\lambda}\sum_{k\geq 0}\frac{(\lambda e^{iu})^k}{k!}\\ &= e^{-\lambda}e^{\lambda e^{iu}}\\ &= e^{\lambda(e^{iu}-1)}.\end{aligned}$$

(iii) Olkoon X normaalisti jakautunut parametrein $\mu = 0$ ja $\sigma^2 = 1$ eli

$$\mathbf{P}[X\in B] = \frac{1}{\sqrt{2\pi}}\int_B e^{-\frac{1}{2}x^2}\,\mathrm{d}x.$$

Tällöin, käyttämällä eksponenttifunktion sarjakehitelmää ja osittaisintegroimalla saatua kaavaa

$$\begin{aligned}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} x^{2n}e^{-\frac{1}{2}x^2}\,\mathrm{d}x &= (2n-1)!!\\ &:= (2n-1)(2n-3)\cdots 3\cdot 1,\end{aligned}$$

saamme

$$\begin{aligned}\phi_X(u) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{iux}\,e^{-\frac{1}{2}x^2}\,\mathrm{d}x\\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\sum_{n\geq 0}\frac{(iux)^n}{n!}e^{-\frac{1}{2}x^2}\,\mathrm{d}x\\ &= \sum_{n\geq 0}\frac{(iu)^n}{n!}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} x^n e^{-\frac{1}{2}x^2}\,\mathrm{d}x\\ &= \sum_{n\geq 0}\frac{(iu)^{2n}}{(2n)!}(2n-1)!!\\ &= \sum_{n\geq 0}\frac{(iu)^{2n}}{(2n)!}\frac{(2n)!}{2^n n!}\\ &= \sum_{n\geq 0}\left(-\frac{1}{2}u^2\right)^n\frac{1}{n!}\\ &= e^{-\frac{1}{2}u^2}.\end{aligned}$$

(iv) Olkoon $X = (X_1, \ldots, X_d)$ tasaisesti jakautunut hyperkuutioon $[0,1]^d$ eli

$$\mathbf{P}_X[B] \;=\; \int_B \mathbf{1}_{[0,1]^d}(x_1, \ldots, x_d)\, \mathrm{d}x_1 \cdots \mathrm{d}x_d, \qquad B \in \mathcal{B}_d.$$

Tällöin

$$\begin{aligned}
\phi_X(u_1, \ldots, u_d) &= \int_{[0,1]^d} e^{i \sum_{k \le d} u_k x_k} \mathrm{d}x_1 \cdots \mathrm{d}x_d \\
&= \int_0^1 \cdots \int_0^1 \prod_{k \le d} e^{i u_k x_k}\, \mathrm{d}x_1 \cdots \mathrm{d}x_d \\
&= \prod_{k \le d} \int_0^1 e^{i u_k x_k}\, \mathrm{d}x_k \\
&= \prod_{k \le d} \left(\int_0^1 \cos(u_k x_k)\, \mathrm{d}x_k + i \sin(u_k x_k)\, \mathrm{d}x_k \right) \\
&= \prod_{k \le d} \left(\frac{\sin u_k}{u_k} + i \frac{1 - \cos u_k}{u_k} \right).
\end{aligned}$$

6.3.3 Apulause. *Olkoon* $X = (X_1, \ldots, X_d)$ *satunnaisvektori,* $A \in \mathbb{R}^{d \times d}$ *ja* $b \in \mathbb{R}^d$*. Tällöin*

$$\phi_{AX+b}(u) \;=\; e^{iu \cdot b} \phi_X(A^\mathsf{T} u),$$

missä A^T *on matriisin* A transpoosi*:* $A^\mathsf{T}_{i,j} = A_{j,i}$.

Todistus. Osoitamme väitteen suoralla laskulla:

$$\begin{aligned}
\phi_{aX+b}(u) &= \mathbf{E}\left[e^{iu \cdot (AX+b)}\right] \\
&= \mathbf{E}\left[e^{iu \cdot AX} e^{iu \cdot b}\right] \\
&= e^{iu \cdot b} \mathbf{E}\left[e^{i(A^\mathsf{T} u) \cdot X}\right] \\
&= e^{iu \cdot b} \phi_X(A^\mathsf{T} u).
\end{aligned}$$

□

6.3.4 Esimerkki. Olkoon X normaalisti jakautunut parametrein μ ja σ^2. Tällöin $Z := (X - \mu)/\sigma$ on normaalisti jakautunut parametrein 0 ja 1. Esimerkin 6.3.2(iii) nojalla

$$\phi_Z(u) \;=\; e^{-\frac{1}{2} u^2}.$$

Siten, apulauseen 6.3.3 nojalla,

$$\begin{aligned}\phi_X(u) &= \phi_{\sigma Z+\mu}(u) \\ &= e^{iu\mu}\phi_Z(\sigma u) \\ &= e^{ui\mu-\frac{1}{2}\sigma^2u^2}.\end{aligned}$$

6.3.5 Apulause. *Karakteristinen funktio $\phi_X : \mathbb{R}^d \to \mathbb{C}$ on*

(i) *rajoitettu,*
(ii) *tasaisesti jatkuva*
(iii) *ja $\phi_X(\bar{0}) = 1$.*

Todistus. (i) Huomaamme aluksi, että $|e^{iu\cdot X}| \le 1$. Siten, epäyhtälön 6.2.1(iv) nojalla, $|\phi_X(u)| \le \mathbf{E}\big[|e^{iu\cdot X}|\big] \le 1$ eli ϕ_X on rajoitettu.

(ii) Osoitamme sitten, että ϕ_X on tasaisesti jatkuva. Työkalumme ovat arvio $|e^{iv\cdot X}| \le 1$, epäyhtälö 6.2.1(iv) ja arvio 6.1.3(i). Olkoot $u, v \in \mathbb{R}^d$. Tällöin

$$\begin{aligned}|\phi_X(u)-\phi_X(v)| &= \big|\mathbf{E}\big[e^{iu\cdot X}\big]-\mathbf{E}\big[e^{iv\cdot X}\big]\big| \\ &= \big|\mathbf{E}\big[e^{iu\cdot X}-e^{iv\cdot X}\big]\big| \\ &\le \mathbf{E}\big[\big|e^{iu\cdot X}-e^{iv\cdot X}\big|\big] \\ &= \mathbf{E}\big[\big|e^{iv\cdot X}\big(e^{i(u-v)\cdot X}-1\big)\big|\big] \\ &\le \mathbf{E}\big[\big|e^{iv\cdot X}\big|\big|e^{i(u-v)\cdot X}-1\big|\big] \\ &\le \mathbf{E}\big[\big|e^{i(u-v)\cdot X}-1\big|\big].\end{aligned}$$

Nyt, koska

$$\big|e^{i(u-v)\cdot X}-1\big| \le 2,$$

niin voimme käyttää dominoidun konvergenssin lausetta. Toisaalta

$$\lim_{h\to 0}\big|e^{ih\cdot X}-1\big| = 0,$$

joten

$$\lim_{h\to 0}\sup_{u\in\mathbb{R}^d}|\phi_X(u+h)-\phi_X(u)| = 0$$

eli ϕ_X on tasaisesti jatkuva.

(iii) Lopuksi huomaamme, että $e^{i\bar{0}\cdot X} = 1$. Siten $\phi_X(\bar{0}) = 1$. □

6.3.6 Lause. *Olkoon $X = (X_1, \ldots, X_d)$ satunnaisvektori. Olkoon lisäksi $\mathbf{E}[|X|^n] < \infty$ jollakin $n \in \mathbb{N}$. Tällöin*

(i) *karakteristisella funktiolla* ϕ_X *on jatkuvat* n*. kertaluvun osittaisderivaatat ja*

$$\frac{\partial^n}{\partial u_{j_1}\cdots\partial u_{j_n}}\phi_X(u) \;=\; i^n\mathbf{E}\big[X_{j_1}\cdots X_{j_n}e^{iu\cdot X}\big]\,,$$

(ii) ϕ_X *:llä on* n*. asteen Taylorin kehitelmä nollassa:*

$$\phi_X(u) \;=\; 1+\sum_{k\le n}\frac{1}{k!}\sum_{j_1,\dots,j_k\le d}\mathbf{E}[X_{j_1}\cdots X_{j_k}]i^k u_{j_1}\cdots u_{j_k}+|u|^n\varepsilon(u),$$

missä $\lim_{u\to 0}\varepsilon(u)=0$*.*

Todistus. (i) Haluamme vaihtaa derivoinnin ja odotusarvon järjestystä lausekkeessa

$$\frac{\partial^n}{\partial u_{j_1}\cdots\partial u_{j_n}}\mathbf{E}\big[e^{iu\cdot X}\big]\,.$$

Perustelemme järjestyksen vaihdon ainoastaan tapauksessa $d=1$ ja $n=1$. Yleinen tapaus seuraa samankaltaisista arvioista, mutta sivuutamme ne.

Järjestyksen vaihtaminen perustellaan dominoidun konvergenssin lauseella 4.2.1(v). Jotta voimme käyttää lausetta 4.2.1(v) tarkastelemme (skalaarista) arviota 6.1.3(i):

$$\big|e^{ix}-1\big| \;\le\; x, \qquad x\in\mathbb{R}.$$

Tästä arviosta näemme välittömästi, että

$$\left|\frac{e^{i(u+h)X}-e^{iuX}}{h}\right| \;\le\; |X|.$$

Antamalla $h\to 0$ jotakin jonoa pitkin saamme dominoidun konvergenssin lauseesta 4.2.1(v), että

$$\begin{aligned}
\phi_X'(u) &= \lim_{h\to 0}\frac{\phi_X(u+h)-\phi_X(u)}{h}\\
&= \lim_{h\to 0}\frac{\mathbf{E}\big[e^{i(u+h)X}\big]-\mathbf{E}\big[e^{iuX}\big]}{h}\\
&= \lim_{h\to 0}\mathbf{E}\left[\frac{e^{i(u+h)X}-e^{iuX}}{h}\right]\\
&= \mathbf{E}\left[\lim_{h\to 0}\frac{e^{i(u+h)X}-e^{iuX}}{h}\right]\\
&= \mathbf{E}\big[iuXe^{iuX}\big]\,.
\end{aligned}$$

Derivoinnin ja odotusarvon vaihtaminen on siis sallittua. Saamme

$$\begin{aligned}\frac{\partial^n}{\partial u_{j_1}\cdots\partial u_{j_n}}\phi_X(u) &= \frac{\partial^n}{\partial u_{j_1}\cdots\partial u_{j_n}}\mathbf{E}\big[e^{iu\cdot X}\big] \\ &= \mathbf{E}\left[\frac{\partial^n}{\partial u_{j_1}\cdots\partial u_{j_n}}e^{(iX)\cdot u}\right] \\ &= i^n\mathbf{E}\big[X_{j_1}\cdots X_{j_n}e^{iu\cdot X}\big].\end{aligned}$$

Kohta (i) on näin todistettu.

(ii) Tämä kohta seuraa välittömästi edellisetä kohdasta (i) soveltamalla tavallista d-ulotteista Taylorin kehitelmää funktioon ϕ_X origossa. □

Seuraava niin sanottu Lévyn kääntökaava kertoo, miten kertymäfunktio voidaan laskea karakteristisesta funktiosta.

6.3.7 Lause. *Olkoot ϕ_X ja F_X satunnaisvektorin $X = (X_1, \ldots, X_d)$ karakteristinen funktio ja yhteiskertymäfunktio. Olkoot $a = (a_1, \ldots, a_d), b = (b_1, \ldots, b_d) \in \mathbb{R}^d$ yhteiskertymäfunktion F_X jatkuvuuspisteitä joille $a_k \leq b_k$ kaikilla $k = 1, \ldots, d$. Tällöin pätee* Lévyn kääntökaava

$$\begin{aligned}&F_X(b_1, \ldots, b_d) - F_X(a_1, \ldots, a_d) \\ &= \frac{1}{(2\pi)^d}\lim_{T\to\infty}\int_{[-T,T]^d}\prod\frac{e^{-ia_ku_k} - e^{-ib_ku_k}}{iu_k}\,\phi_X(u_1, \ldots, u_d)\,\mathrm{d}u_1\cdots\mathrm{d}u_d.\end{aligned}$$

Ennen käätökaavan 6.3.7 tarkkaa todistusta esitämme heuristisesti, mistä on kyse.

Tarkastelemme vain skalaaritapausta $d = 1$, vaikka yleinen tapaus onkin täsmälleen samankaltainen. Olkoon $\hat{\mu}$ todennäkoisyysjakauman eli -mitan μ karakteristinen funktio:

$$\hat{\mu}(u) := \int_{\mathbb{R}} e^{iux}\,\mu[\mathrm{d}x].$$

Asetamme sitten bilineaarisen suhteen karaketeristisen funktioiden ja todennäköisyysmittojen välille seuraavasti:

$$\langle\hat{\nu}, \mu\rangle := \int_{\mathbb{R}}\hat{\nu}(x)\,\mu[\mathrm{d}x].$$

Huomaamme nyt, että relaatiolle $\langle\cdot,\cdot\rangle$ pätee kaava

$$\langle\hat{\nu}, \mu\rangle = \langle\hat{\mu}, \nu\rangle.$$

Nimittäin Fubinin kaavan 5.3.6 nojalla

$$
\begin{aligned}
\langle \hat{\nu}, \mu \rangle &= \int_{\mathbb{R}} \hat{\nu}(x)\, \mu[\mathrm{d}x] \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{iux}\, \nu[\mathrm{d}u] \right) \mu[\mathrm{d}x] \\
&= \int_{\mathbb{R}^2} e^{iux}\, \mu[\mathrm{d}x]\nu[\mathrm{d}u] \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{iux}\, \mu[\mathrm{d}x] \right) \nu[\mathrm{d}u] \\
&= \int_{\mathbb{R}} \hat{\mu}(u)\, \nu[\mathrm{d}u] \\
&= \langle \hat{\mu}, \nu \rangle\, .
\end{aligned}
$$

Olkoon sitten $B \in \mathcal{B}$ ja $\check{\mathbf{1}}_B$ se todennäköisyysmitta, jonka karakteristinen funktio on indikaattori $\mathbf{1}_B$ (oletamme tylysti tällaisen mitan olemassaolon). Tällöin, jos μ on satunnaismuuttujan X jakauma, niin

$$
\begin{aligned}
\mathbf{P}[X \in B] &= \int_{\mathbb{R}} \mathbf{1}_B(x)\mu[\mathrm{d}x] \\
&= \langle \mathbf{1}_B, \mu \rangle \\
&= \langle \hat{\mu}, \check{\mathbf{1}}_B \rangle \\
&= \int_{\mathbb{R}} \hat{\mu}(u)\, \check{\mathbf{1}}_B[\mathrm{d}u].
\end{aligned}
$$

Lévyn kääntökaava 6.3.7 siis sanoo heuristisesti, että

$$
\check{\mathbf{1}}_{[a,b]}[\mathrm{d}u] = \frac{1}{2\pi} \frac{e^{-iau} - e^{-ibu}}{iu}\, \mathrm{d}u.
$$

Siirrymme sitten kääntökaavan 6.3.7 rigoröösiin todistukseen.

Lauseen 6.3.7 todistuksessa tarvitsemme seuraavaa aputulosta, jonka todistuksen jätämme harjoitustehtäväksi 6.6.

6.3.8 Apulause. *Olkoon*

$$
\mathrm{sgn}(x) := \begin{cases} -1, & \textit{kun } x < 0, \\ 0, & \textit{kun } x = 0, \\ 1, & \textit{kun } x > 0. \end{cases}
$$

Tällöin

$$
\lim_{T \to \infty} \int_{-T}^{T} \frac{\sin(xu)}{u}\, \mathrm{d}u = \mathrm{sgn}(x)\pi
$$

ja

$$\left|\int_{-T}^{T} \frac{\sin(ux)}{u}\,\mathrm{d}u\right| \;\le\; 2\pi.$$

Lauseen 6.3.7 todistus. Todistamme vain tapauksen $d = 1$. Yleinen tapaus voidaan todistaa samankaltaisilla argumenteilla.

Olkoon $a \le b$. Merkitsemme

$$I_T \;:=\; \frac{1}{2\pi}\int_{-T}^{T} \frac{e^{-iua}-e^{-iub}}{iu}\phi_X(u)\,\mathrm{d}u.$$

Fubinin lauseesta 5.3.6 seuraa nyt, että

$$\begin{aligned}
I_T &= \frac{1}{2\pi}\int_{-T}^{T} \frac{e^{-iua}-e^{-iub}}{iu}\left(\int_{\mathbb{R}} e^{iux}\,\mathrm{d}F_X(x)\right)\mathrm{d}u \\
&= \frac{1}{2\pi}\int_{\mathbb{R}}\left(\int_{-T}^{T} \frac{e^{-iua}-e^{-iub}}{iu}e^{iux}\mathrm{d}u\right)\mathrm{d}F_X(x) \\
&=: \int_{\mathbb{R}} \Theta_T(x)\,\mathrm{d}F_X(x),
\end{aligned}$$

missä merkitsimme

$$\Theta_T(x) \;:=\; \frac{1}{2\pi}\int_{-T}^{T} \frac{e^{-iua}-e^{-iub}}{iu}e^{iux}\mathrm{d}u.$$

Fubinin käyttö oli muuten sallittua, koska

$$\left|\frac{e^{-iua}-e^{-iub}}{iu}e^{iux}\right| \;\le\; |b-a|.$$

Tarkastelemme sitten sisäintegraalia $\Theta_T(x)$. Pienen “kompleksitrigonometrisen” pyörityksen jälkeen näemme, että

$$\Theta_T(x) \;=\; \frac{1}{2\pi}\int_{-T}^{T} \frac{\sin(u(x-a))-\sin(u(x-b))}{u}\,\mathrm{d}u.$$

Nyt apulause 6.3.8 tulee käyttöön. Nimittäin sen nojalla

$$\begin{aligned}
\lim_{T\to\infty} I_T &= \lim_{T\to\infty}\int_{\mathbb{R}} \Theta_T(x)\,\mathrm{d}F_X(x) \\
&= \int_{\mathbb{R}} \lim_{T\to\infty} \Theta_T(x)\,\mathrm{d}F_X(x) \\
&= \frac{1}{2}\int_{\mathbb{R}} (\mathrm{sgn}(x-a)-\mathrm{sgn}(y-b))\,\mathrm{d}F_X(x) \\
&= \frac{1}{2}\left(-\mathbf{P}[X<a]+\mathbf{P}[X>a]+\mathbf{P}[X<b]-\mathbf{P}[X>b]\right) \\
&= \mathbf{P}[a<X\le b] \\
&= F_X(b)-F_X(a).
\end{aligned}$$

□

6.3.9 Seuraus. (i) *Karakteristinen funktio ϕ_X määrää jakauman $\mathbf{P}_X$ yksikäsitteisesti.*

(ii) *Satunnaisvektorit X ja Y ovat riippumattomia jos ja vain jos*

$$\phi_{X,Y}(u,v) \;=\; \phi_X(u)\,\phi_Y(v).$$

(iii) *Jos*

$$\int_{\mathbb{R}^d} |\phi_X(u)|\;\mathrm{d}u \;<\; \infty,$$

niin X on jatkuva ja sen tiheysfunktio saadaan kääntökaavasta

$$f_X(x) \;=\; \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-iu\cdot x}\phi_X(u)\,\mathrm{d}u.$$

(iv) *Jos X on diskreetti satunnaismuuttuja, niin*

$$\mathbf{P}[X=x_n] \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-iux_n}\phi_X(u)\,\mathrm{d}u.$$

Todistus. (i) Tämä tulos seuraa Lévyn kääntölauseesta 6.3.7 ja yhteiskertymäfunktion oikealta jatkuvuudesta. Jätämme perustelun yksityiskohdat harjoitustehtäväksi 6.7.

(ii) Harjoitustehtävä 6.8.

(iii) Harjoitustehtävä 6.9.

(iv) Harjoitustehtävä 6.10. □

Seuraava lause 6.3.10 kertoo, miten riippumattomien satunnaismuuttujien summan karakteristinen funktio esitetään summattavien satunnaismuuttujien karakterististen funktioiden avulla. Tämä esitys on keskeinen myöhemmin esitettävän keskeisen raja-arvolauseen 8.3.1 todistuksessa.

6.3.10 Lause. *Jos X ja Y riippumattomia satunnaismuuttujia, niin*

$$\phi_{X+Y}(u) \;=\; \phi_X(u)\,\phi_Y(u).$$

Todistus. Harjoitustehtävä 6.11. □

Lause 6.3.10 sanoo itse asiassa, että karakteristinen funktio muuttaa konvoluutiot tuloiksi: kertymäfunktiota $F_X * F_Y$ vastaava karakteristinen funktio on $\phi_X \cdot \phi_Y$.

6.3.11 Esimerkki. *Multinormaalijakauma* määritellään tyypillisesti sen karakteristisen funktion avulla: satunnaisvektori $X = (X_1, \ldots, X_d)$ on multinormaalisti jakautunut parametrein $\mu \in \mathbb{R}^d$ ja $\Sigma \in \mathbb{R}^{d\times d}$, jos sen karakteristinen funktio on muotoa

$$\phi_X(u) \;=\; e^{iu\cdot\mu - \frac{1}{2}u\cdot\Sigma u}.$$

Matriisista Σ oletetaan, että se on symmetrinen ja positiivisesti semidefiniitti:

$$\Sigma_{k,l} = \Sigma_{l,k} \qquad \text{ja} \qquad \sum_{k,l} c_k \Sigma_{k,l} c_l \geq 0$$

kaikilla $c_k, c_l \in \mathbb{R}$ ja $k, l \leq d$.

Tällöin μ on satunnaisvektorin X odotusarvovektori

$$\mu \;=\; (\mu_1, \ldots, \mu_d) \;=\; (\mathbf{E}[X_k])_{k\leq d}$$

ja Σ on satunnaisvektorin *kovarianssimatriisi*

$$\begin{aligned}\Sigma &= (\Sigma_{k,l})_{k,l\leq d} \\ &= (\mathbf{Cov}[X_k, X_l])_{k,l\leq d} \\ &:= \big(\mathbf{E}\big[(X_k-\mu_k)(X_l-\mu_l)\big]\big)_{k,l\leq d}.\end{aligned}$$

(harjoitustehtävä 6.12).

Lisäksi jokaista vektoria μ ja positiivisesti semidefiniittiä matriisia Σ vastaa jokin multinormaalisti jakautunut satunnaisvektori X. Itse asiassa X voidaan rakentaa riippumattomien normaalistijakautuneiden satunnaismuuttujien avulla. Nimittäin olkoot $\xi = (\xi_1, \ldots, \xi_d)$ riippumattomia samoin normaalisti jakautuneita satunnaismuuttujia odotusarvolla 0 ja varianssilla 1 ja $A \in \mathbb{R}^{d\times d}$. Tällöin

$$X \;:=\; \mu + A\xi \tag{6.3.12}$$

on multinormaalisti jakautunut parametrein μ ja $A^{\mathsf{T}}A$ (harjoitustehtävä 6.13). Lisäksi, jos X on multinormaalisti jakautunut parametrein μ ja Σ, niin on olemassa sellainen $A \in \mathbb{R}^{d\times d}$, että $\Sigma = A^{\mathsf{T}}A$ ja X voidaan esittää muodossa (6.3.12), missä vektorin $\xi = (\xi_1, \ldots, \xi_d)$ komponentit ovat riippumattomia samoin normaalisti jakautuneita satunnaismuuttujia odotusarvolla 0 ja varianssilla 1 (harjoitustehtävä 6.14).

Lopuksi huomautamme, että jos kaavassa (6.3.12) olevan matriisi A on täyttä astetta eli sen sarakevektorit virittävät koko avaruuden $\mathbb{R}^d$, niin X on jatkuvasti jakautunut ja sen tiheysfunktio on

$$f_X(x) \;=\; \frac{1}{(2\pi)^{d/2}\|\Sigma\|^{1/2}}\exp\left\{-\frac{1}{2}(x-\mu)\cdot\Sigma^{-1}(x-\mu)\right\},$$

missä $\|\Sigma\|$ on matriisin Σ determinantti ja Σ^{-1} on sen käänteismatriisi (harjoitustehtävä 6.15).

Harjoitustehtäviä

6.1. Olkoon $z = (x, y) \in \mathbb{C} \setminus \{(0,0)\}$. Asetamme

$$\frac{1}{z} \;:=\; \frac{x}{x^2+y^2} - i\frac{y}{x^2+y^2}.$$

Osoita, että tällöin

$$z \cdot \frac{1}{z} \;=\; 1.$$

6.2. Todista apulauseen 6.2.1 kohta (iii).

6.3. Olkoot X degeneroitunut pisteeseen $x_0 \in \mathbb{R}^d$. Toisin sanoen

$$\mathbf{P}[X \in B] \;=\; \delta_{x_0}[B], \quad B \in \mathcal{B}_d.$$

Laske ϕ_X.

6.4. Olkoot X yksinkertainen satunnaismuuttuja, joka saa arvot $x_1, \ldots, x_n$ todennäköisyyksin $p_1, \ldots, p_n$, $\sum p_n = 1$. Laske ϕ_X.

6.5. Olkoon X eksponentiaalisesti jakautunut parametrilla $\mu > 0$. Toisin sanoen sen tiheysfunktio on

$$f_X(x) \;=\; \mu e^{-\mu x}\mathbf{1}_{[0,\infty)}(x).$$

Laske ϕ_X.

6.6. Todista apulause 6.3.8.

6.7. Todista seuraus 6.3.9(i).

6.8. Todista seuraus 6.3.9(ii).

6.9. Todista seuraus 6.3.9(iii).

6.10. Todista seuraus 6.3.9(iv).

6.11. Todista lause 6.3.10.

6.12. Olkoot satunnaisvektorin X karakteristinen funktio muotoa

$$\phi_X(u) \;=\; e^{i\mu\cdot u - \frac{1}{2}u\cdot\Sigma u}.$$

Osoita, että tällöin

$$\begin{aligned} \mu_k &= \mathbf{E}[X_k], \\ \Sigma_{k,l} &= \mathbf{Cov}[X_k, X_l]. \end{aligned}$$

6.13. Olkoon $\xi = (\xi_1, \ldots, \xi_d)$ vektori, jonka komponentit ovat riippumattomia normaalistijakautuneita satunnaismuuttujia parametrein $\mu = 0$ ja $\sigma^2 = 1$. Olkoon $\mu \in \mathbb{R}^d$ ja $A \in \mathbb{R}^{d\times d}$. Asetamme

$$X \;:=\; \mu + A\xi.$$

Merkitään $\Sigma := A^{\mathsf{T}}A$. Osoita, että nyt

$$\begin{aligned} \mu_k &= \mathbf{E}[X_k], \\ \Sigma_{k,l} &= \mathbf{Cov}[X_k, X_l]. \end{aligned}$$

6.14. Olkoon X multinormaalisti jakautunut parametrein μ ja Σ. Osoita, että on olemassa sellainen matriisi K ja vektori ξ riippumattomia normaalisti jakautuneita satunnaismuuttujia, että

$$X \;=\; \mu + A\xi.$$

6.15. Olkoon X multinormaalisti jakautunut parametrein μ ja Σ. Osoita, että jos Σ on täyttä astetta, niin X on jatkuva ja sen tiheysfunktio on

$$f_X(x) \;=\; \frac{1}{(2\pi)^{d/2}\|\Sigma\|^{1/2}} \exp\left\{-\frac{1}{2}(x-\mu)\cdot\Sigma^{-1}(x-\mu)\right\},$$

missä $\|\Sigma\|$ on matriisin Σ determinantti ja Σ^{-1} on sen käänteismatriisi.

Luku 7

Suppenemiset

Kaikista todennäköisyysteoriassa käsiteltävistä suppenemisen muodoista melkein varma suppeneminen on luonnollisin, tai ainakin helpoin ymmärtää. Muita suppenemismuotoja tarvitsemme lähinnä kahdesta syystä:

1. *Melkein varma suppeneminen on varsin vahva suppenemisen muoto: monissa mielenkiintoisissa tilanteissa joudumme tyytymään heikompiin suppenemismuotoihin. Keskeinen raja-arvolause on esimerkki tällaisesta tilanteesta.*
2. *Integroituvan satunnaismuuttujajonon (X_n) melkein varmasta suppenemisesta kohti integroituvaa satunnaismuuttujaa X emme voi päätellä, että vastaavat odotusarvot $\mathbf{E}[X_n]$ suppenevat kohti odotusarvoa $\mathbf{E}[X]$. Toisin sanoen odotusarvo-operaattori $\mathbf{E}$ ei ole jatkuva melkein varman suppenemisen suhteen.*

Syy 1. johtaa stokastisen ja heikon suppenemisen käsitteisiin ja syy 2. johtaa L^p-suppenemisen käsitteeseen.

Esitämme suppenemiset enemmän tai vähemmän vahvuusjärjestyksessä vahvimmasta alkaen.

7.1 Melkein varma suppeneminen

Kaikkein yksinkertaisin satunnaismuuttujia tai -vektoreita koskeva suppenemisen muoto olisi tietysti pisteittäinen suppeneminen: $X_n(\omega) \to X(\omega)$ kaikilla $\omega \in \Omega$. Tällaista suppenemista on kuitenkin käytännössä turha odottaa. Jos esimerkiksi Y_n on klaavojen suhteellinen frekvenssi n reilun kolikon heitossa, niin vahva suurten lukujen laki pisteittäiselle suppenemiselle olisi: $Y_n(\omega) \to 1/2$ kaikilla $\omega \in \Omega$. Tällainen tulos ei kuitenkaan päde. Nimittäin heittosarja $\omega = (\mathsf{klaava}, \mathsf{klaava}, \ldots)$ on täysin mahdollinen ja tällöin $Y_n = 1$ kaikilla $n \in \mathbb{N}$. Samoin heittosarja $\omega = (\mathsf{kruuna}, \mathsf{kruuna}, \ldots)$ on myöskin mahdollinen ja tällöin $Y_n = 0$ kaikilla $n \in \mathbb{N}$. Huomattavaa on kuitenkin, että edellämainittujen yksittäistapausten todennäköisyys on nolla (kuten on tietysti minkä tahansa muunkin yksittäisen heittosarjan todennäköisyys). Siten, "heittämällä pois" nollajoukot, päädymme melkein varman suppenemisen käsitteeseen.

7.1.1 Määritelmä. Jono satunnaisvektoreita (X_n) *suppenee melkein varmasti* kohti satunnaisvektoria X, jos

$$\mathbf{P}\big[\omega \in \Omega\,;\, \lim X_n(\omega) = X(\omega)\big] \;=\; 1.$$

Tällöin merkitsemme $X_n \overset{m.v.}{\to} X$.

Tarkastelkaamme nyt vielä hieman käsitettä "melkein varmasti". Yleisesti sanomme, että jokin satunnaisvektoreita koskeva ominaisuus pätee "melkein varmasti", kun se pätee todennäköisyydellä yksi. Esimerkiksi merkintä $X \overset{m.v.}{=} Y$ tarkoittaa, että $\mathbf{P}[X = Y] = 1$. On helppo nähdä että $\overset{m.v.}{=}$ on ekvivalenssirelaatio:

- $X \overset{m.v.}{=} X$,
- jos $X \overset{m.v.}{=} Y$, niin $Y \overset{m.v.}{=} X$,
- jos $X \overset{m.v.}{=} Y$ ja $Y \overset{m.v.}{=} Z$, niin $X \overset{m.v.}{=} Z$.

Lisäksi esimerkiksi $\mathbf{E}[X] = \mathbf{E}[Y]$, jos $X \overset{m.v.}{=} Y$. Itse asiassa $\mathbf{E}[|X - Y|] = 0$ jos ja vain jos $X \overset{m.v.}{=} Y$.

On siis luonnollista, että melkein varmasti samat satunnaisvektorit tyypillisesti samaistetaan.

Melkein varma suppeneminen vastaa käytännössä monella tavoin täysin varmaa pisteittäistä suppenemista $X_n(\omega) \to X(\omega)$ kaikilla $\omega \in \Omega$. Itse asiassa on helppo nähdä, että aikaisemmissa luvuissa todistamamme lauseet, joissa oletamme pisteittäisen suppenemisen, pätevät myös melkein varman suppenemisen vallitessa. Näytämme esimerkin vuoksi, että monotonisen konvergenssin lause 4.2.1(i) pätee melkein varmalle suppenemiselle. Jätämme harjoitustehtäväksi 7.1 tarkistaa, että myös dominoidun konvergenssin lause 4.2.1(v) pätee melkein varman suppenemisen vallitessa.

7.1.2 Lause. *Olkoon (X_n) melkein varmasti kasvava jono satunnaismuuttujia, jolle $X_n \overset{m.v}{\to} X$. Jos $\mathbf{E}[X_1] > -\infty$, niin $\mathbf{E}[X_n] \uparrow \mathbf{E}[X]$.*

Todistus. Olkoon $\Omega_n \subset \Omega$, $n \in \mathbb{N}$, se joukko jolle $X_n \le X_{n+1}$ pätee. Olkoon samoin $\Omega_0 \subset \Omega$ se joukko, jolle $X_n \to X$ pätee. Voimme olettaa, että

$\Omega_0, \Omega_n \in \mathcal{F}$. Nimittäin näiden joukkojen todennäköisyys on yksi, ja voimme aina tarvittaessa täydellistää σ-algebran $\mathcal{F}$ sisältämään melkein varmat joukot. Mutta nyt joukossa

$$\tilde{\Omega} \ := \ \Omega_0 \cap \bigcap \Omega_n$$

pätee: (X_n) on kasvava ja $X_n \to X$. Siten $(X_n \mathbf{1}_{\tilde{\Omega}})$ on kasvava ja $\mathbf{E}[X_n \mathbf{1}_{\tilde{\Omega}}] \to \mathbf{E}[X \mathbf{1}_{\tilde{\Omega}}]$. Mutta toisaalta

$$\begin{aligned}
\mathbf{P}[\tilde{\Omega}^c] &= \mathbf{P}\Big[\Big(\bigcap (\Omega_n \cap \Omega_0)\Big)^c\Big] \\
&= \mathbf{P}\Big[\bigcup (\Omega_n \cap \Omega_0)^c\Big] \\
&\leq \sum \mathbf{P}[\Omega_n^c \cup \Omega_0^c] \\
&\leq \sum \left(\mathbf{P}[\Omega_n^c] + \mathbf{P}[\Omega_0^c]\right) \\
&= 0.
\end{aligned}$$

Siten $\mathbf{P}[\tilde{\Omega}] = 1$. Mutta huomaamalla, että $\mathbf{E}[X] = \mathbf{E}[Y]$, kun $X \overset{m.v.}{=} Y$, tästä seuraa, että $\mathbf{E}[X_n \mathbf{1}_{\tilde{\Omega}}] = \mathbf{E}[X_n]$ ja $\mathbf{E}[X \mathbf{1}_{\tilde{\Omega}}] = \mathbf{E}[X]$. □

Edellisen todistuksen nojalla on selvää, että voimme käytännössä samaistaa satunnaisvektorijonot $X = (X_n)_{n \in \mathbb{N}}$ ja $\tilde{X} = (\tilde{X}_n)_{n \in \mathbb{N}}$, jos

$$\mathbf{P}\Big[X_n = \tilde{X}_n\Big] \ = \ 1 \qquad \text{kaikilla } n \in \mathbb{N}.$$

Nimittäin tällöin

$$\begin{aligned}
\mathbf{P}\Big[X \neq \tilde{X}\Big] &= \mathbf{P}\Big[X_n \neq \tilde{X}_n \text{ jollakin } n \in \mathbb{N}\Big] \\
&\leq \sum \mathbf{P}\Big[X_n \neq \tilde{X}_n\Big] \\
&= 0.
\end{aligned}$$

Mainittakoon kuitenkin, että tätä samaistusta ei voi tehdä ilman merkittäviä lisäoletuksia satunnaisvektoriperheille $X = (X_j)_{j \in J}$, ja $\tilde{X} = (\tilde{X}_j)_{j \in J}$, jos indeksijoukko J on ylinumeroituva. Nimittäin ylinumeroituvassa tapauksessa ehto

$$\mathbf{P}\Big[X_j = \tilde{X}_j\Big] \ = \ 1 \qquad \text{kaikilla } j \in J$$

ei takaa, että

$$\mathbf{P}\Big[X \neq \tilde{X}\Big] \ = \ \mathbf{P}\Big[X_j \neq \tilde{X}_j \text{ jollakin } j \in J\Big] \ = \ 0.$$

Itse asiassa joukko $\{X \neq \tilde{X}\} = \{X_j \neq \tilde{X}_j \text{ jollakin } j \in J\}$ voi olla jopa ei-mitallinen. Tämä samaistuksen mahdottomuus aiheuttaa ongelmia esimerkiksi jatkuva-aikaisten stokastisten prosessien teoriassa.

Borelin ja Cantellin lemman 5.2.1 avulla voidaan todistaa seuraava ehto satunnaisvektorijonon (X_n) melkein varmalle suppenemiselle. Ehto on hyödyllinen silloin, kun rajavektori X voidaan "arvata" ja todennäköisyyksiä $\mathbf{P}[|X_n - X| > \varepsilon]$ osataan arvioida.

7.1.3 Apulause. *Olkoon X satunnaisvektori ja (X_n) jono satunnaisvektoreita. Jos kaikille $\varepsilon > 0$ pätee*

$$\sum \mathbf{P}[|X_n - X| > \varepsilon] \;<\; \infty,$$

niin $X_n \stackrel{m.v.}{\to} X$.

Todistus. Harjoitustehtävä 7.5. □

Osion lopuksi annamme ehdon satunnaisvektorisarjan $\sum X_n$ melkein varmalle suppenemiselle. (Tarkkaavainen luentojen seuraaja on jo saattanut huomata yhden tähän aihepiiriin liittyvän tuloksen: seuraus 4.2.4.)

7.1.4 Apulause. *Riittävä ehto sarjan $\sum X_n$ melkein varmalle suppenemiselle on sellaisen positiivitermisen sarjan $\sum \varepsilon_n$ olemassaolo, että*

(i) $\sum \varepsilon_n \;<\; \infty$,
(ii) $\sum \mathbf{P}[|X_n| > \varepsilon_n] \;<\; \infty$.

Todistus. Olkoon

$$A \;:=\; \{|X_n| > \varepsilon_n \text{ ä.u.}\}\,.$$

Oletuksesta (ii) seuraa Borelin ja Cantellin lemman 5.2.1(i) nojalla, että $\mathbf{P}[A] = 0$. Olkoon sitten $\omega \in A^c = \{|X_n| \leq \varepsilon_n \text{ j.l.}\}$. Tällöin on olemassa sellainen $n_0(\omega)$, että $|X_n(\omega)| < \varepsilon_n$, kun $n \geq n_0(\omega)$. Tästä seuraa majoranttiperiaatteen nojalla, että

$$\begin{aligned}
\sum_{n\geq 1} |X_n(\omega)| \;&<\; \sum_{n\leq n_0(\omega)} |X_n(\omega)| + \sum_{n>n_0(\omega)} \varepsilon_n \\
&\leq\; \sum_{n\leq n_0(\omega)} |X_n(\omega)| + \sum_{n\geq 1} \varepsilon_n \\
&<\; \infty.
\end{aligned}$$

Siispä sarja $\sum X_n(\omega)$ suppenee itseisesti kaikilla $\omega \in A^c$. Koska $\mathbf{P}[A^c] = 1$, niin sarja $\sum X_n$ suppenee (itseisesti) melkein varmasti. □

7.2 L^p-suppeneminen

Ilman esimerkiksi monotonista tai dominoitua lisäoletusta suppenemisesta $X_n \overset{m.v.}{\to} X$ ei seuraa, että $\mathbf{E}[X_n] \to \mathbf{E}[X]$. Esitämme nyt suppenemismuodon, joka takaa odotusarvojen suppenemisen $\mathbf{E}[X_n] \to \mathbf{E}[X]$ ilman mitään lisäoletuksia.

7.2.1 Määritelmä. Olkoon $p \geq 1$. Jono satunnaisvektoreita (X_n) *suppenee L^p:ssä* kohti satunnaisvektoria X, jos

$$\mathbf{E}\big[|X_n - X|^p\big] \;\to\; 0.$$

Tällöin merkitsemme $X_n \overset{L^p}{\to} X$.

Avaruus $L^p := L^p(\Omega, \mathcal{F}, \mathbf{P})$ on niiden satunnaisvektorien avaruus, joille $|X|^p \in L^1$. Jos samaistamme melkein varmasti samat satunnaisvektorit, niin

$$\|X\|_p \;:=\; \mathbf{E}[|X|^p]^{1/p}$$

on normi avaruudella L^p. Nimittäin normiin liittyvä ehto $\|cX\|_p = |c|\|X\|_p$ on ilmiselvä ja kolmioepäyhtälöehto $\|X+Y\|_p \leq \|X\|_p + \|Y\|_p$ on Minkowskin epäyhtälö 4.4.2(ii). Ehto $\|X\|_p = 0$, kun $X = 0$ on selvä. Ehtoon $X = 0$, kun $\|X\|_p = 0$ tarvitsemme melkein varmasti samojen satunnaisvektorien samaistusta.

Määritelmä 7.2.1 olettaa implisiittisesti, että $X \in L^p$ ja $(X_n) \subset L^p$.

Myöhemmin, osiossa 9.2, osoitamme, että normiavaruus $(L^p, \|\cdot\|_p)$ on *täydellinen*, toisin sanoen sen Cauchy-jonot suppenevat. Siten L^p on Banachin avaruus eli täydellinen normiavaruus.

L^p-suppenemiset suhtautuvat toisiinsa kauniisti eri p:n arvoilla. Tämä seuraa todennäköisyysmitan rajoittuneisuudesta: $\mathbf{P}[\Omega] = 1$. Yleisille rajoittamattomille mitoille seuraava tulos ei nimittäin päde.

7.2.2 Apulause. *Olkoon $p \geq q \geq 1$.*

(i) *Jos $X \in L^p$, niin $X \in L^q$.*

(ii) *Jos $X_n \overset{L^p}{\to} X$, niin $X_n \overset{L^q}{\to} X$.*

Todistus. Jätämme todistuksen harjoitustehtäväksi 7.7. Vihjeenä todettakoon, että kannattaa tarkastella joukkoja $\{|X_n - X| < 1\}$ ja $\{|X_n - X| \geq 1\}$ sekä käyttää Tšebyševin epäyhtälöa 4.4.3. □

Seuraava tulos näyttää, että L^p-suppeneminen soveltuu loistavasti odotusarvojen suppenemisen tarkasteluun.

7.2.3 Lause. *Jos* $X_n \xrightarrow{L^p} X$ *jollakin* $p \geq 1$*, niin* $\mathbf{E}[X_n] \to \mathbf{E}[X]$*.*

Todistus. Apulauseen 7.2.2 nojalla $X_n \xrightarrow{L^1} X$. Mutta nyt

$$\begin{aligned} \big|\mathbf{E}[X_n] - \mathbf{E}[X]\big| &= \big|\mathbf{E}[X_n - X]\big| \\ &\leq \mathbf{E}\big[|X_n - X|\big] \\ &\to 0. \end{aligned}$$

□

Seuraava esimerkki näyttää, että L^p-suppeneminen ei seuraa melkein varmasta suppenemisesta (kohta (i)), eikä myöskään toisin päin (kohta (ii)).

7.2.4 Esimerkki. Olkoon $(\Omega, \mathcal{F}, \mathbf{P}) := ([0,1], \mathcal{B}, \ell)$. Tarkastelemme siis tasaista todennäköisyysmittaa tai -jakaumaa välillä $[0,1]$.

(i) Olkoon

$$X_n(\omega) := n\mathbf{1}_{(0,\frac{1}{n})}(\omega).$$

Selvästi tällöin $X_n \xrightarrow{m.v.} 0$ (itse asiassa suppeneminen on jopa täysin varmaa). Mutta toisaalta

$$\begin{aligned} \mathbf{E}\big[|X_n - 0|^p\big] &= \mathbf{E}[|X_n|^p] \\ &= \int_0^1 |X(\omega)|^p \,\mathrm{d}\omega \\ &= \int_0^1 \Big|n\mathbf{1}_{(0,\frac{1}{n})}(\omega)\Big|^p \,\mathrm{d}\omega \\ &= n^p \int_0^1 \mathbf{1}_{(0,\frac{1}{n})}(\omega)\,\mathrm{d}\omega \\ &= n^{p-1} \\ &\not\to 0, \end{aligned}$$

sillä $p \geq 1$.

(ii) Olkoon $m \in \mathbb{N}$ ja $\Delta_{m,k}$, $k \leq 2^m$ välin $[0,1]$ ositus tasapituisiin osaväleihin. Tällöin siis $|\Delta_{m,k}| = 2^{-m}$. Määrittelemme sitten jonon (X_n) asettamalla

$$X_1 := \mathbf{1}_{\Delta_{1,1}}, \quad X_2 := \mathbf{1}_{\Delta_{1,2}}, \quad \ldots \quad , X_{2^m} := \mathbf{1}_{\Delta_{1,2^m}},$$

$$X_{2^m+1} := \mathbf{1}_{\Delta_{2,1}}, \quad \ldots, \quad X_{k2^m+1} := \mathbf{1}_{\Delta_{k,1}}, \ldots$$

Tällöin jono (X_n) ei voi supeta melkein varmasti, sillä $\limsup X_n = 1$ ja $\liminf X_n = 0$. Mutta toisaalta $X_n \xrightarrow{L^p} 0$, sillä

$$\lim \mathbf{E}[|X_n|^p] = \lim \mathbf{P}[\Delta_{n,1}] = 0.$$

7.3 Stokastinen suppeneminen

Stokastinen suppeneminen on sekä melkein varmaa suppenemista että L^p-suppenemista heikompi suppenemisen muoto. Stokastisen suppenemisen käyttökelpoisuus on esimerkiksi siinä, että suurten lukujen laki saadaan pätemään heikompien riippuvuusoletusten vallitessa kuin jos vaadittaisiin melkein varmaa suppenemista.

7.3.1 Määritelmä. Jono satunnaisvektoreita (X_n) *suppenee stokastisesti* eli *mitan suhteen* kohti satunnaisvektoria X, jos kaikilla $\varepsilon > 0$ pätee

$$\mathbf{P}\big[|X_n - X| > \varepsilon\big] \quad \to \quad 0.$$

Tällöin merkitsemme $X_n \xrightarrow{\mathbf{P}} X$.

Seuraava tulos sanoo, että sekä melkein varma että L^p-suppeneminen ovat molemmat stokastista suppenemista vahvempia suppenemisen muotoja. L^p-suppenemisen kohdalla joudumme tietysti olettamaan a priori, että jono $(X_n) \subset L^p$. Yleisesti tietenkin satunnaisvektorijono (X_n) voi supeta kohti satunnaisvektoria X stokastisesti, vaikka $X_n \notin L^p$ millään $p \geq 1$.

7.3.2 Lause. *Olkoon (X_n) jono satunnaisvektoreita.*

(i) *Jos $X_n \xrightarrow{m.v.} X$, niin $X_n \xrightarrow{\mathbf{P}} X$.*
(ii) *Jos $X_n \xrightarrow{L^p} X$ jollakin $p \geq 1$, niin $X_n \xrightarrow{\mathbf{P}} X$.*

Todistus. (i) Tämän kohdan jätämme harjoitustehtäväksi 7.11.

(ii) Apulauseen 7.2.2 nojalla riittää tarkastella L^1-suppenemista. Mutta nyt väite seuraa Tšebyševin epäyhtälöstä 4.4.3. Nimittäin

$$\begin{aligned} \mathbf{P}[|X_n - X| > \varepsilon] &\leq \frac{1}{\varepsilon}\mathbf{E}[|X_n - X|] \\ &\to 0 \end{aligned}$$

kaikilla $\varepsilon > 0$. □

Seuraava esimerkki osoittaa, että stokastinen suppeneminen on aidosti sekä melkein varmaa että L^p-suppenemista heikompi suppenemisen muoto.

7.3.3 Esimerkki. Olkoon $(\Omega, \mathcal{F}, \mathbf{P})$ ja $\Delta_{m,k}$ kuten esimerkissä 7.2.4. Olkoon samoin $X_n = \mathbf{1}_{\Delta_{k,l}}$, $n = m2^k + l$, kuten esimerkissä 7.2.4.

(i) Nyt jono (X_n) ei suppene melkein varmasti (edelleenkään), mutta kaikilla $\varepsilon > 0$ pätee

$$\lim \mathbf{P}[|X_n| > \varepsilon] \;=\; \lim \mathbf{P}[\Delta_{n,1}] \;=\; 0.$$

Siten $X_n \xrightarrow{\mathbf{P}} 0$.

(ii) Olkoon sitten $Y_n := a_n X_n$, missä

$$a_n \;=\; \frac{n}{\mathbf{P}[\Delta_{k,l}]},$$

kun $n = m2^k + l$. Tällöin (Y_n) ei suppene L^p:ssä, sillä

$$\mathbf{E}[|Y_n|^p] \;=\; n^p \;\to\; \infty.$$

Mutta

$$\lim \mathbf{P}[|Y_n| > \varepsilon] \;=\; \lim \mathbf{P}[\Delta_{n,1}] \;=\; 0.$$

Siten $X_n \xrightarrow{\mathbf{P}} 0$.

Kun seuraavan tuloksen yhdistää lauseeseen 7.3.2(ii) ja esimerkkiin 7.3.3(ii) on kuvamme melkein varman ja stokastisen suppenemisen välisestä yhteydestä täydellinen.

7.3.4 Lause. *Jos $X_n \xrightarrow{\mathbf{P}} X$, niin on olemassa sellainen osajono (n_k), että $X_{n_k} \xrightarrow{m.v.} X$.*

Todistus. Rakennamme osajonon (n_k). Olkoon $n_1 = 1$. Etsimme sitten indeksin n_k induktiivisesti: n_k on pienin indeksi, joka on aidosti suurempi kuin indeksi n_{k-1} ja jolle pätee

$$\mathbf{P}\left[|X_{n_k} - X| > 2^{-k}\right] \;<\; 2^{-k}.$$

Koska $X_n \xrightarrow{\mathbf{P}} X$, niin tällainen indeksi n_k löytyy kaikille k. Olemme siis saaneet osajonon (n_k).

Nyt

$$\sum \mathbf{P}\left[|X_{n_k} - X| > 2^{-k}\right] \;<\; \sum 2^{-k} \;<\; \infty.$$

Siten, Borelin ja Cantellin lemman 5.2.1(i) nojalla,

$$\mathbf{P}\left[|X_{n_k} - X| > 2^{-k} \text{ ä.u.}\right] \;=\; 0.$$

Mutta tämä tarkoittaa, että $X_{n_k} \xrightarrow{m.v.} X$. Nimittäin nyt kaikille $\varepsilon > 0$ pätee

$$\begin{aligned} 0 \;&=\; \mathbf{P}\left[|X_{n_k} - X| > 2^{-k} \text{ ä.u.}\right] \\ &\geq\; \mathbf{P}[|X_{n_k} - X| > \varepsilon \text{ ä.u.}]. \end{aligned}$$

Koska $\varepsilon > 0$ on mielivaltainen, niin $\mathbf{P}[X_{n_k} \not\to X] = 0$. □

7.4 Jakaumasuppeneminen

Edes stokastinen suppeneminen ei ole riittävän heikko keskeistä raja-arvolausetta varten (se on riittävän heikko heikkoa suurten lukujen lakia varten). Joudumme esittämään vielä stokastista suppenemistakin heikomman suppenemiskäsitteen: jakauma- eli heikon suppenemisen.

7.4.1 Määritelmä. Jono satunaisvektoreita (X_n) *suppenee heikosti* eli *jakaumaltaan* kohti satunnaisvektoria X, jos

$$F_{X_n} \to F_X$$

kaikissa F_X:n jatkuvuuspisteissä. Tällöin merkitsemme $X_n \xrightarrow{d} X$.

Yksinkertaisempi määritelmä jakaumasuppenemiselle olisi tietysti vaatia $F_{X_n}(x) \to F_X(x)$ kaikilla $x \in \mathbb{R}^d$. Seuraava esimerkki kuitenkin osoittaa, että tämä on aivan liikaa vaadittu.

7.4.2 Esimerkki. Olkoon $(\Omega, \mathcal{F}, \mathbf{P}) := ([0,1], \mathcal{B}, \ell)$. Tarkastelemme siis jälleen kerran tasaista jakaumaa välillä $[0,1]$. Asetamme

$$X_n(\omega) \ := \ \omega^n.$$

Tällöin on varsin helppo nähdä, että $X_n \to 0$ melkein varmasti, L^p:ssä ja stokastisesti. Osoitamme myöhemmin lauseessa 7.4.4, että tällöin $X_n \xrightarrow{d} 0$. Nyt “satunnaismuuttujan” $X = 0$ kertymäfunktio on $F_X(x) = \mathbf{1}_{[0,\infty)}(x)$. Mutta

$$\begin{aligned}
F_{X_n}(x) &= \mathbf{P}[X_n \le x] \\
&= \ell[\omega \in [0,1]\,;\, \omega^n \le x] \\
&= \ell\big[\omega \in [0,1]\,;\, \omega \le x^{1/n}\big] \\
&= x^{1/n}\mathbf{1}_{[0,1]}(x) + \mathbf{1}_{(1,\infty)}(x) \\
&\to \mathbf{1}_{(0,\infty)}(x).
\end{aligned}$$

Siten $0 = F_{X_n}(0) \not\to F_X(0) = 1$. Siispä kertymäfunktiot F_n, $n \in \mathbb{N}$, eivät suppene kohti kertymäfunktiota F_X sen epäjatkuvuuspisteessä $x = 0$.

Huomattavaa jakaumasuppenemisessa on, että se on jakaumia eli todennäköisyysmittoja koskeva käsite: se tarkastelee satunnaisvektoreita X_n, $n \in \mathbb{N}$, ja X ainoastaan niiden jakaumien F_{X_n}, $n \in \mathbb{N}$, ja F_X tasolla. Erityisesti siis voimme puhua satunnaisvektorijonon (X_n) suppenemisesta, vaikka satunnaisvektorit olisi määritelty eri todennäköisyysavaruuksilta. Tämä ei ole, ainakaan helposti, mahdollista muiden edellä esitettyjen suppenemiskäsitteiden tapauksessa. Lisäksi jakaumasuppenemisessa on huomattavaa, että jos $(\tilde{X}_n)$ on jokin sellainen jono satunnaisvektoreita, että X_n ja $\tilde{X}_n$ ovat samoin jakautuneita kaikilla $n \in \mathbb{N}$, niin jonon (X_n) ja $(\tilde{X}_n)$ jakaumasuppenemista koskevat ominaisuudet ovat samoja. Tämä seuraa siitä, että funktiot F_{X_n}, $n \in \mathbb{N}$, eivät "näe" satunnaisvektorien X_n, $n \in \mathbb{N}$, riippuvuussuhteita. Tästä seuraa esimerkiksi se, että jakaumasuppenemista käsiteltäessä voimme halutessamme olettaa jonon (X_n) koostuvan riippumattomista satunnaisvektoreista.

On vielä syytä mainita, että yleisessä mittateoriassa ja funktionaalianalyysissä on myös heikon suppenemisen käsitteet, jotka erovat hieman meidän stokastikkojen käsitteestä. Yleisessä mittateoriassa mittajono (μ_n) suppenee heikosti kohti mittaa μ, jos $\int g \mathrm{d}\mu_n \to \int g \mathrm{d}\mu$ kaikilla jatkuvilla funktioilla g, joilla $g(x) = 0$, kun $|x|$ on riittävän iso. Funktionaalianalyysissä puolestaan Banachin avaruuden B jono (b_n) suppenee heikosti kohti Banachin avaruuden alkiota b, jos $f(b_n) \to f(b)$ kaikilla rajoitetuilla lineaarisilla funktioilla, eli *funktionaaleilla*, $f : B \to \mathbb{R}$.

Heikko suppeneminen on epäilemättä hieman hämyinen käsite. Siksi onkin syytä tarkastella lähemmin, mitä se tarkoittaa. Seuraavaksi esitettävä ns. *Portmanteau-lause* kokoaa yhteen joukon yhtäpäteviä ehtoja heikolle suppenemiselle. Itse asiassa usein heikko suppeneminen määritelläänkin Portmanteau-lauseen 7.4.3 kohdan (vi) avulla. Tämä määritelmä jakauma- eli heikolle suppenemiselle on jo varsin lähellä mittateoreetikkojen vastaavaa määritelmää.

7.4.3 Lause. *Seuraavat ovat yhtäpitäviä:*

(i) $X_n \xrightarrow{d} X$,

(ii) $\lim F_{X_n} = F_X$ *kaikissa* F_X*:n jatkuvuuspisteissä,*

(iii) $\lim \mathbf{P}[X_n \in B] = \mathbf{P}[X \in B]$ *kaikilla* $B \in \mathcal{B}_d$, *joille* $\mathbf{P}[X \in \partial B] = 0$. *Merkintä* ∂B *tarkoittaa joukon* B reunaa.

(iv) $\limsup \mathbf{P}[X_n \in F] \leq \mathbf{P}[X \in F]$ *kaikilla suljetuilla* $F \subset \mathbb{R}^d$,

(v) $\liminf \mathbf{P}[X_n \in G] \geq \mathbf{P}[X \in G]$ *kaikilla avoimilla* $G \subset \mathbb{R}^d$,

(vi) $\lim \mathbf{E}[g(X_n)] = \mathbf{E}[g(X)]$ *kaikilla rajoitetuilla ja jatkuvilla funktioilla* $g : \mathbb{R}^d \to \mathbb{R}$,

(vii) $\lim \phi_{X_n} = \phi_X$.

Todistus. (i) $\Leftrightarrow$ (ii): Tämä on määritelmä.

(ii) $\Leftrightarrow$ (iii): Huomaamme aluksi, että ominaisuus "$\mathbf{P}[X \in \partial B] > 0$ jos ja vain jos ∂B sisältää kertymäfunktion F_X epäjatkuvuuspisteen" säilyy joukkojen monotonisissa rajoissa. Siten, monotonisen luokan lauseen eli Dynkinin laajennuslauseen 2.3.8 nojalla, riittää tarkastella suorakaiteita

$$B_x \;:=\; \bigtimes_{k\le d} \big(-\infty, x_k\big], \qquad x = (x_1,\ldots,x_d) \in \mathbb{R}^d.$$

Mutta näille joukoille väite on selvä. Nimittäin $x \in \mathbb{R}^d$ on kertymäfunktion F_X epäjatkuvuuspiste jos ja vain jos $\mathbf{P}[X \in \partial B_x] > 0$, sillä

$$\mathbf{P}[X \in \partial B_x] \;=\; F_X(x_1,\ldots,x_d) - \lim_{y_1\to x_1-,\ldots,y_d\to x_d-} F_X(y_1,\ldots,y_d).$$

(iii) $\Leftrightarrow$ (iv) & (v): Olkoon $A \in \mathcal{B}_d$ mielivaltainen joukko. Tällöin joukko $F := A \cup \partial A$ on aina suljettu. Samoin joukko $G := A \setminus \partial A$ on aina avoin. Nyt huomaamme, että $\mathbf{P}[X \in \partial A] = 0$ jos ja vain jos

$$\begin{aligned} \mathbf{P}[X \in G] &= \mathbf{P}[X \in A \setminus \partial A] \\ &= \mathbf{P}[X \in A] \\ &= \mathbf{P}[X \in A \cup \partial A] \\ &= \mathbf{P}[X \in F]. \end{aligned}$$

Mutta tällöin (ja vain tällöin) kohtien (iv) ja (v) ala- ja yläraja-arvot ovat samoja, joten kyseessä on aito raja.

(iv) $\Leftrightarrow$ (v): Jos G on avoin, niin G^c on suljettu, ja päinvastoin. Olkoon sitten F on suljettu ja $G = F^c$. Tällöin

$$\begin{aligned} \limsup \mathbf{P}[X_n \in F] &= \limsup \left(1 - \mathbf{P}[X_n \in G]\right) \\ &= 1 + \limsup \left(-\mathbf{P}[X_n \in G]\right) \\ &= 1 - \liminf \mathbf{P}[X_n \in G]. \end{aligned}$$

Siten

$$\begin{aligned} & \limsup \mathbf{P}[X_n \in F] \le \mathbf{P}[X \in F] \\ \Leftrightarrow\quad & 1 - \liminf \mathbf{P}[X_n \in G] \le 1 - \mathbf{P}[X \in G] \\ \Leftrightarrow\quad & \liminf \mathbf{P}[X_n \in G] \ge \mathbf{P}[X \in G]. \end{aligned}$$

(ii) $\Leftarrow$ (vi): Valitsemme

$$g(x_1,\ldots,x_d) \;:=\; \mathbf{1}_{(-\infty,y_1]\times\cdots\times(-\infty,y_d]}(x_1,\ldots,x_d).$$

Tällöin

$$\begin{aligned} \mathbf{E}[g(X_n)] &= \mathbf{E}\big[\mathbf{1}_{\{X_{n,1}\le y_1,\ldots,X_{n,d}\le y_d\}}\big] \\ &= \mathbf{P}[X_{n,1} \le y_1,\ldots,X_{n,d} \le y_d] \\ &= F_{X_n}(y_1,\ldots,y_d) \end{aligned}$$

Väite seuraisi tästä, jos indikaattorifunktiomme g olisi jatkuva (mitä se ei ole). Voimme kuitenkin silottaa sen epäjatkuvuuspisteen $(y_1,\ldots,y_d)$ ympärillä jatkuvaksi siten, että $g(x_1,\ldots,x_d)=1$, kun $x_k\le y_k$ ja $g(x_1,\ldots,x_d)=0$, kun $x_k+\varepsilon>y_k$. Antamalla $\varepsilon\to 0$ ja käyttämällä dominoidun konvergenssin lausetta 4.2.1(v) näemme, että väite pätee kaikissa yhteiskertymäfunktion F_X jatkuvuuspisteissä $(y_1,\ldots,y_d)$.

(ii) $\Rightarrow$ (vi): Harjoitustehtävä 7.15.

(i) $\Leftrightarrow$ (vii): Tämä seuraa myöhemmin esitettävästä lauseesta 7.4.7, jonka todistuksen vaikean puolen (i) $\Leftarrow$ (vii) kuitenkin sivuutamme. □

Seuraava lause sanoo, että jakaumasuppeneminen on heikoin esitetyistä suppenemisen muodoista eli että se seuraa stokastisesta suppenemisesta.

7.4.4 Lause. *Jos $X_n \xrightarrow{\mathbf{P}} X$, niin $X_n \xrightarrow{d} X$.*

Todistus. Harjoitustehtävä 7.16. □

Seuraava esimerkki osoittaa, että jakaumasuppeneminen on aidosti heikompi suppenemisen muoto kuin skokastinen suppeneminen.

7.4.5 Esimerkki. Olkoon X ja $\tilde{X}$ riippumattomia normaalisti jakautuneita satunnaismuuttujia parametrein $\mu=0$ ja $\sigma^2=1$. Olkoon $X_n=X$, kun n on pariton ja $X_n=\tilde{X}$, kun n on parillinen. Nyt

$$F_{X_n}(x) \;=\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}y^2}\,\mathrm{d}y$$

kaikilla $n\in\mathbb{N}$, joten mitä ilmeisimmin (X_n) suppenee heikosti. Toisaalta $X-\tilde{X}$ on “aidosti satunnainen”. Itse asiassa voidaan osoittaa, että $X-\tilde{X}$ on normaalisti jakautunut parametrein $\mu=0$ ja $\sigma^2=2$. Siten todennäköisyys $\mathbf{P}[|X_n-X_{n+1}|>\varepsilon]$ on ε:sta riippuva, mutta n:stä riippumaton aidosti positiivinen vakio. Niinpä (X_n) ei voi supeta stokastisesti.

Esimerkistä 7.4.5 huolimatta seuraava käänteinen tulos pätee.

7.4.6 Lause. *Jos $X_n \xrightarrow{d} X$ ja X on deterministinen, niin $X_n \xrightarrow{\mathbf{P}} X$.*

Todistus. Olkoon $c := X\in\mathbb{R}^d$ (vakio) ja $\varepsilon>0$ mielivaltainen. Todistus perustuu siihen havaintoon, että joukko $\{x\in\mathbb{R}^d; |x-c|<\varepsilon\}$ on avoin. Nimittäin nyt Portmateau-lauseen 7.4.3 kohdan (v) nojalla

$$\liminf \mathbf{P}[|X_n-c|<\varepsilon] \;\ge\; \mathbf{P}[|X-c|<\varepsilon] \;=\; 1.$$

Siten $\lim \mathbf{P}[|X_n-c|<\varepsilon]=1$ eli $X_n \xrightarrow{\mathbf{P}} c$. □

Seuraava tulos, tai pikemminkin sen vaikea kohta (ii), on *Lévyn jatkuvuuslause*. Se karakterisoi jakaumasuppenemisen karakterististen funktioiden avulla.

7.4.7 Lause. *Olkoon (X_n) jono satunnaisvektoreita.*

(i) *Jos $X_n \xrightarrow{d} X$, niin $\phi_{X_n} \to \phi_X$.*
(ii) *Jos $\phi_{X_n} \to \phi$ ja ϕ on jatkuva origossa, niin on olemassa sellainen satunnaisvektori X, että ϕ on sen karakteristinen funktio ja $X_n \xrightarrow{d} X$.*

Todistus. (i) Tämä väite seuraa Portmanteau-lauseen 7.4.3 kohdasta (vi). Nimittäin kuvaus $u \mapsto e^{iu\cdot x}$ on jatkuva ja rajoitettu. Se, että kuvaus $u \mapsto e^{iu\cdot x}$ on kompleksiarvoinen ei tuota ongelmia: meidän pitää vain katsoa reaali- ja imaginääriosia erikseen.

(ii) Tämä osa on se vaikea osa. Jätämme todistuksen harjoitustehtäväksi (graduksi). □

7.5 Suppenemisten väliset suhteet

Seuraava tulos on kokoomalause. Siinä ei itse asiassa todisteta paljoakaan mitään uutta, vaan lähinnä ainoastaan kootaan tässä luvussa aikaisemmin esitettyjä tuloksia yhdeksi "isoksi kuvaksi" (Piirrä toki se kuva!). Jätämme harjoitustehtäväksi 7.18 koota lauseen perustelu. Itse asiassa lähinnä vain käänteinen väite vaatii tarkempaa perustelua: muu on lähestulkoon pelkkää keräilyä.

7.5.1 Lause. *Seuraavat implikaatiot pätevät ja mikään seuraavista implikaatioista ei päde kääntäen.*

(i) *Jos $X_n \xrightarrow{m.v.} X$, niin $X_n \xrightarrow{\mathbf{P}} X$.*
(ii) *Jos $X_n \xrightarrow{L^p} X$ jollakin $p \geq 1$, niin $X_n \xrightarrow{\mathbf{P}} X$.*
(iii) *Olkoon $p \geq q \geq 1$. Jos $X_n \xrightarrow{L^p} X$, niin $X_n \xrightarrow{L^q} X$.*
(iv) *Jos $X_n \xrightarrow{\mathbf{P}} X$, niin on olemassa osajono (n_k), jolle $X_{n_k} \xrightarrow{m.v.} X$.*
(v) *Jos $X_n \xrightarrow{\mathbf{P}} X$, niin $X_n \xrightarrow{d} X$.*

Harjoitustehtäviä

7.1. Osoita dominoidun konvergenssin lause 4.2.1(v), kun oletamme vain, että $X_n \xrightarrow{m.v} X$ ja $|X_n| \leq Y$ melkein varmasti kaikilla $n \in \mathbb{N}$.

7.2. Olkoon (X_n) jono satunnaismuuttujia. Osoita, että $X_n \to X$ täsmälleen joukossa

$$\bigcap_{m\geq 1}\bigcup_{n\geq 1}\bigcap_{k\geq n}\left\{|X_k - X| \leq \frac{1}{m}\right\}.$$

7.3. Olkoon $X_n \overset{m.v.}{\to} X$ ja $f : \mathbb{R}^d \to \mathbb{R}^d$ jatkuva. Osoita, että tällöin $f(X_n) \overset{m.v.}{\to} f(X)$.

7.4. Olkoon $X_n \overset{m.v.}{\to} X$, $Y_n \overset{m.v.}{\to} Y$ ja $a, b \in \mathbb{R}^d$. Osoita, että tällöin

$$aX_n + bY_n \quad \overset{m.v.}{\to} \quad aX + bY.$$

7.5. Todista apulause 7.1.3.

7.6. Osoita, että apulauseen 7.1.3 ehto $\sum \mathbf{P}[|X_n - X| > \varepsilon] < \infty$ ei ole välttämätön heikolle suppenemiselle.

7.7. Todista apulause 7.2.2.

7.8. Olkoon $X_n \overset{L^p}{\to} X$, $Y_n \overset{L^p}{\to} Y$ ja $a, b \in \mathbb{R}^d$. Osoita, että tällöin

$$aX_n + bY_n \quad \overset{L^p}{\to} \quad aX + bY.$$

7.9. Olkoon $X_n \overset{L^p}{\to} X$ ja $f : \mathbb{R}^d \to \mathbb{R}^d$ jatkuva. Päteekö tällöin $f(X_n) \overset{L^p}{\to} f(X)$?

7.10. Osoita, että apulause 7.2.2 ei päde, jos todennäköisyysmitan sijasta tarkastelemme rajoittamatonta mittaa.

7.11. Todista lause 7.3.2(i).

7.12. Osoita, että $X_n \overset{\mathbf{P}}{\to} X$ jos ja vain jos

$$\mathbf{E}\left[\frac{|X_n - X|}{1 + |X_n - X|}\right] \quad \to \quad 0.$$

7.13. Olkoon $X_n \overset{\mathbf{P}}{\to} X$ ja $f : \mathbb{R}^d \to \mathbb{R}^d$ jatkuva. Päteekö tällöin $f(X_n) \overset{\mathbf{P}}{\to} f(X)$?

7.14. Olkoon $X_n \overset{\mathbf{P}}{\to} X$, $Y_n \overset{\mathbf{P}}{\to} Y$ ja $a, b \in \mathbb{R}^d$. Päteekö tällöin, että

$$aX_n + bY_n \quad \overset{\mathbf{P}}{\to} \quad aX + bY?$$

7.15. Osoita ekvivalenssi (ii) $\Rightarrow$ (vi) Portmanteau-lauseessa 7.4.3.

7.16. Todista lause 7.4.4.

7.17. Olkoon $X_n \xrightarrow{d} X$ ja $f : \mathbb{R}^d \to \mathbb{R}^d$ jatkuva kertymäfunktion F_X jatkuvuuspisteissä. Osoita, että tällöin $f(X_n) \xrightarrow{d} f(X)$.

7.18. Kerää todistus lauseelle 7.5.1.

7.19. Ovatko melkein varma, stokastinen, L^p- ja jakaumasuppeneminen konsistentteja? Toisin sanoen jos $X_n \to X$ jossakin mielessä ja $X_n \to Y$ jossakin toisessa mielessä, niin onko välttämättä $X = Y$?

Luku 8

Raja-arvolauseita

Todennäköisyys liittyy ennustamattomiin ilmiöihin. Esimerkiksi kolikonheitto on ennustamaton siinä mielessä, ettemme voi täydellä varmuudella sanoa tuleeko kruuna vai klaava. Kolikonheitto ei ole kuitenkaan kaoottinen. Toisin sanoen siinä esiintyy jotakin säännönmukaisuutta. Vaikka yksittäisen heiton tulos onkin täysin ennustamaton eli kaoottinen, niin tiedämme esimerkiksi, että kuitenkin pitkässä heittosarjassa klaavojen suhteellisella osuudella on tapana supeta kohti jotakin tiettyä lukua, jota sitten meillä (jos olemme frekventistejä) on tapana kutsua klaavan todennäköisyydeksi. Tämä erottaa kolikonheiton puhtaasti kaoottisesta ilmiöstä, jota ei voi ennustaa mitenkään. Yleisemmin ja tarkemmin sanottuna raja-arvolauseet tuovat todennäköisyyslaskentaan järjestystä kaaoksen tilalle.

Tässä luvussa käsittelemme ainoastaan satunnaismuuttujia, vaikka vastaavat tulokset pätevät myös satunnaisvektoreille.

8.1 Suurten lukujen lait

Kertaamme aluksi korrelaation käsitteen.

Neliöintegroituvat satunnaismuuttujat X_n, $n \in \mathbb{N}$, ovat *korreloimattomia*, jos kaikille $n \neq m$ pätee

$$\mathbf{E}[X_n X_m] \;=\; \mathbf{E}[X_n]\mathbf{E}[X_m].$$

Nimitys “korreloimattomuus” johtuu siitä että tällöin

$$\mathbf{Cor}[X_n, X_m] \;:=\; \frac{\mathbf{Cov}[X_n, X_m]}{\sqrt{\mathbf{Var}[X_n]\mathbf{Var}[X_m]}} \;=\; 0.$$

Korreloimattomuudesta seuraa välittömästi, että

$$\mathbf{Var}\left[\sum_{k\le n} X_k\right] \;=\; \sum_{k\le n}\mathbf{Var}[X_k]$$

(harjoitustehtävä 8.1).

Huomattavaa on, että korrelaatio käsitteenä, samoin kuin kovarianssi, vaatii satunnaismuuttujilta integroituvuutta. Siis odotusarvojen $\mathbf{E}[XY]$, $\mathbf{E}[X]$ ja $\mathbf{E}[Y]$ on oltava olemassa. Riittävä ehto näiden odotusarvojen olemassaololle on $X, Y \in L^2$. Nimittäin tällöin, Schwarzin epäyhtälön nojalla,

$$\mathbf{E}[|XY|] \;\leq\; \mathbf{E}[X^2]^{1/2}\mathbf{E}[Y^2]^{1/2} \;<\; \infty,$$

mistä seuraa myös, valitsemalla vuoron perään $X = 1$ ja $Y = 1$, että myös $\mathbf{E}[|X|]$ ja $\mathbf{E}[|Y|]$ ovat äärellisiä.

Seuraava tulos on *heikko suurten lukujan laki* korreloimattomille satunnaismuuttujille.

8.1.1 Lause. *Olkooon $(X_n) \subset L^2$ jono korreloimattomia satunnaismuuttujia joille*

$$\begin{aligned} \mathbf{E}[X_n] &= \mu, \\ \frac{1}{n^2}\sum_{k\leq n}\mathbf{Var}[X_k] &\to 0. \end{aligned}$$

Tällöin

$$\frac{1}{n}\sum_{k\leq n} X_k \;\overset{\mathbf{P}}{\to}\; \mu.$$

Todistus. Olkoon

$$S_n \;:=\; \sum_{k\leq n} X_k.$$

Tällöin, odotusarvon lineaarisuuden ja jonon korreloimattomuuden, nojalla

$$\begin{aligned} \mathbf{E}[S_n] &= \mathbf{E}\left[\sum_{k\leq n} X_k\right] \\ &= \sum_{k\leq n}\mathbf{E}[X_k] \\ &= \sum_{k\leq n}\mu \\ &= n\mu \end{aligned}$$

ja

$$\begin{aligned} \mathbf{Var}[S_n] &= \mathbf{Var}\left[\sum_{k\leq n} X_k\right] \\ &= \sum_{k\leq n}\mathbf{Var}[X_k]. \end{aligned}$$

Siten, Markovin epäyhtälön 4.4.4(i) nojalla,

$$\begin{aligned}
\mathbf{P}\left[\left|\frac{1}{n}\sum_{k\leq n} X_k - \mu\right| > \varepsilon\right] &= \mathbf{P}\Big[|S_n - n\mu| > n\varepsilon\Big] \\
&= \mathbf{P}\Big[\big|S_n - \mathbf{E}[S_n]\big| > n\varepsilon\Big] \\
&\leq \frac{\mathbf{Var}[S_n]}{n^2\varepsilon^2} \\
&\leq \frac{\sum_{k\leq n}\mathbf{Var}[X_k]}{n^2\varepsilon^2} \\
&\to 0.
\end{aligned}$$

□

8.1.2 Seuraus. *Olkoon* $(X_n) \subset L^2$ *jono korreloimattomia ja samoin jakautuneita satunnaismuuttujia. Tällöin*

$$\frac{1}{n}\sum_{k\leq n} X_k \overset{\mathbf{P}}{\to} \mathbf{E}[X_1].$$

8.1.3 Esimerkki. Heikon suurten lukujen lain sovelluksena osoitamme *Weierstraßin aproksimaatiolauseen*:

> Olkoon $f : [0,1] \to \mathbb{R}$ jatkuva funktio. Tällöin jokaiselle $\varepsilon > 0$ on olemassa sellainen polynomi P, että
>
> $$\sup_{x\in[0,1]} |f(x) - P(x)| < \varepsilon.$$

Käytämme aproksimaatiossa *Bernsteinin polynomeja*

$$B_n(x) := \sum_{0\leq k\leq n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k(1-x)^{n-k}.$$

Olkoon S_n binomijakautunut parametrein n ja x. Tällöin siis

$$S_n = \sum_{k\leq n} X_k,$$

missä X_k:t ovat riippumattomia binomijakautuneita satunnaismuuttujia parameterin 1 ja x. Nyt

$$\mathbf{P}[S_n = k] = \binom{n}{k} x^k(1-x)^{n-k},$$

joten

$$\mathbf{E}\left[f\left(\frac{1}{n}S_n\right)\right] \;=\; \sum_{0\le k\le n} f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k} \;=\; B_n(x).$$

Olkoon sitten $\varepsilon > 0$. Koska f on suljetun välin jatkuva funktio, niin se on *tasaisesti jatkuva*. Toisin sanoen jokaiselle $\varepsilon > 0$ on olemassa sellainen $\delta = \delta(\varepsilon)$, että

$$\sup_{|x-y|\le\delta} |f(x) - f(y)| \;<\; \frac{1}{2}\varepsilon.$$

Nyt jokaiselle $x \in [0,1]$ pätee

$$\Big|B_n(x) - f(x)\Big| \;=\; \left|\mathbf{E}\left[f\left(\frac{1}{n}S_n\right) - f(x)\right]\right|.$$

Merkitsemme sitten

$$\begin{aligned} Y_n &:= \left|f\left(\frac{1}{n}S_n\right) - f(x)\right|, \\ Z_n &:= \left|\frac{1}{n}S_n - x\right|. \end{aligned}$$

Nyt $Y_n < \frac{1}{2}\varepsilon$, kun $Z_n < \delta$. Siten

$$\begin{aligned} \Big|B_n(x) - f(x)\Big| &\le \mathbf{E}[Y_n] \\ &= \mathbf{E}\big[Y_n \mathbf{1}_{\{Z_n\le\delta\}}\big] + \mathbf{E}\big[Y_n \mathbf{1}_{\{Z_n>\delta\}}\big] \\ &\le \frac{1}{2}\varepsilon\mathbf{P}[Z_n \le \delta] + 2\max_{y\in[0,1]}|f(y)|\mathbf{P}[Z_n > \delta] \\ &\le \frac{1}{2}\varepsilon + 2\max_{y\in[0,1]}|f(y)|\,\frac{1}{4n\delta^2}. \end{aligned}$$

Valitsemalla nyt n niin isoksi, että

$$2\max_{y\in[0,1]}|f(y)|\,\frac{1}{4n\delta^2} \;<\; \frac{1}{2}\varepsilon$$

saamme $|B_n(x) - f(x)| < \varepsilon$. Koska δ:n valinta ei riippunut pisteestä x, olemme näyttäneet väitteen.

Siirrymme sitten tarkastelemaan vahvaa suurten lukujen lakia. Aloitamme aputuloksella, jota kutsutaan *Skorohodin lemmaksi*.

8.1.4 Apulause. *Olkoot $X_1, \ldots, X_n$ riippumattomia satunnaismuuttujia ja $S_n := \sum_{k \le n} X_k$. Olkoon*

$$c_n \;:=\; \max_{k \le n} \mathbf{P}\left[\left|S_n - S_k\right| > s\right] \;<\; 1.$$

Tällöin

$$\mathbf{P}\left[\max_{k \le n} \left|S_k\right| > 2s\right] \;\le\; \frac{1}{1 - c_n} \mathbf{P}\left[\left|S_n\right| > s\right].$$

Todistus. Harjoitustehtävä 8.5. □

Seuraava mielenkiintoinen tulos osoittaa, että riippumattomassa tapauksessa stokastinen suppeneminen onkin itse asiassa melkein varmaa suppenemista. Tulos seuraa olennaisesti siitä, että riippumattomien satunnaismuuttujien summa joko suppenee tai hajaantuu melkein varmasti. Toisin sanoen ei ole mahdollista, että se suppenisi positiivisella ykköstä pienemmällä todennäköisyydellä.

8.1.5 Lause. *Olkoon (X_n) jono riippumattomia satunnaismuuttujia. Tällöin sarja $\sum X_n$ suppenee stokastisesti jos ja vain jos se suppenee melkein varmasti.*

Todistus. On selvää, että riittää ainoastaan osoittaa, että stokastisesta suppenemisesta seuraa melkein varma suppeneminen.

Teemme ristiriitatodistuksen.

Jos reaalilukujono (s_n) ei suppene, niin on olemassa sellainen $\varepsilon > 0$ että kaikille $m \in \mathbb{N}$ pätee

$$\sup_{n; n > m} |s_n - s_m| \;>\; \varepsilon.$$

Nyt sarja $\sum X_n$ joko suppenee tai hajaantuu melkein varmasti, koska sen summattavat ovat riippumattomia (seuraus 5.2.6).

Oletamme siis että $\sum X_n$ hajaantuu positiivisella todennäköisyydellä. Tällöin siis on olemassa sellainen $\varepsilon > 0$ ja $\delta > 0$, että jokaiselle $m \in \mathbb{N}$

$$\mathbf{P}\left[\sup_{n > m} \left|\sum_{m < k \le n} X_k\right| > \varepsilon\right] \;\ge\; \delta. \tag{8.1.6}$$

Olkoon sitten $N \in \mathbb{N}$. Skorohodin lemmasta 8.1.4 seuraa nyt, että

$$\mathbf{P}\left[\max_{m < n \le N} \left|\sum_{m < k \le n} X_k\right| > \varepsilon\right] \;\le\; \frac{1}{1 - C_{m,N}} \mathbf{P}\left[\left|\sum_{m < k \le N} X_k\right| > \frac{\varepsilon}{2}\right],$$

missä

$$C_{m,N} \ := \ \max_{m<n\leq N} \mathbf{P}\left[\left|\sum_{n\leq k\leq N} X_k\right| > \frac{\varepsilon}{2}\right].$$

Oletamme sitten, että $\sum X_n$ suppenee stokastisesti. Tällöin

$$\sum_{m<k\leq N} X_k \ \stackrel{\mathbf{P}}{\to} \ 0,$$

kun $m, N \to \infty$. Toisin sanoen, kun $m, N \to \infty$,

$$\mathbf{P}\left[\left|\sum_{m<k\leq N} X_k\right| > \frac{\varepsilon}{2}\right] \ \to \ 0$$

ja lisäksi

$$C_{m,N} \ \to \ 0.$$

Mutta antamalla nyt ensin $N \to \infty$ näemme, että

$$\lim_{m\to\infty} \mathbf{P}\left[\sup_{n>m}\left|\sum_{m<k\leq n} X_k\right| > \varepsilon\right] \ = \ 0.$$

Mutta tämä on ristiriidassa positiivisen hajaantumisen (8.1.6) kanssa. □

Käyttämällä L^2-suppenemista saamme seuraavan ehdon sarjan melkein varmalle suppenemiselle. Jätämme tarkan perustelun harjoitustehtäväksi 8.6.

8.1.7 Seuraus. *Olkoon* $(X_n) \subset L^2$ *jono riippumattomia satunnaismuuttujia, joille* $\mathbf{E}[X_n] = 0$. *Jos sarja* $\sum \mathbf{E}[X_n^2]$ *suppenee, niin sarja* $\sum X_n$ *suppenee melkein varmasti.*

Todistamme seuraavaksi *vahvan suurten lukujen lain* keskitetyille neliöintegroituville satunnaismuuttujille. Sitä ennen tarvitsemme vielä yhden aputuloksen: niin sanotun *Kroneckerin lemman.*

8.1.8 Apulause. *Olkoon* $\sum x_n$ *suppeneva sarja. Olkoon* $b_n \uparrow \infty$. *Tällöin*

$$\frac{1}{b_n}\sum_{k\leq n} b_k x_k \ \to \ 0.$$

Todistus. Harjoitustehtävä 8.7. □

8.1.9 Lause. *Olkoon $(X_n) \subset L^2$ jono riippumattomia satunnaismuuttujia, joille $\mathbf{E}[X_n] = 0$. Olkoon $b_n \uparrow \infty$ sellainen jono reaalilukuja, että*

$$\sum \frac{\mathbf{E}[X_n^2]}{b_n^2} \;<\; \infty.$$

Tällöin

$$\frac{1}{b_n} \sum_{k \le n} X_k \;\overset{m.v.}{\longrightarrow}\; 0.$$

Todistus. Kroneckerin lemma 8.1.8 nojalla väite seuraa sarjan $\sum X_n/b_n$ melkein varmasta suppenemisesta. Mutta tämä taas seuraa välittömästi seurauksesta 8.1.7, sillä oletimme, että sarja

$$\sum \frac{\mathbf{E}[X_n^2]}{b_n^2} \;=\; \sum \mathbf{E}\left[\left(\frac{X_n}{b_n}\right)^2\right]$$

suppenee. □

Lauseissa 8.1.1 ja 8.1.9 jouduimme olettamaan, että summattavat satunnaismuuttuja olivat neliöintegroituvia. Riippumattomien satunnaismuuttujien tapauksessa voimme luopua tästä oletuksesta. Integroivuusletus $(X_n) \subset L^1$ on tietysti pakko tehdä. Nimittäin odotettu raja on $\mu = \mathbf{E}[X_n]$.

Seuraava tulos on *Kolmogorovin vahva suurten lukujen laki.*

8.1.10 Lause. *Olkoon (X_n) jono riippumattomia samoin jakautuneita satunnaismuuttujia. Olkoon $\mu := \mathbf{E}[X_1]$ olemassa. Tällöin*

$$\frac{1}{n} \sum_{k \le n} X_k \;\overset{m.v.}{\longrightarrow}\; \mu.$$

Jos $\mathbf{E}[|X_1|] = \infty$, niin yllä oleva sarja hajaantuu melkein varmasti.

Todistus. Osoitamme vain tilannetta $\mathbf{E}[|X_1|] < \infty$ koskevan väitteen. Tilanteen $\mathbf{E}[|X_1|] = \infty$ tarkastelun jätämme harjoitustehtäväksi 8.8.

Todistus perustuu aikaisemmin todistettuun vahvaan suurten lukujen lakiin 8.1.9 ja niin sanottuun katkaisutekniikkaan.

Asetamme katkaisun

$$\tilde{X}_n \;:=\; X_n \mathbf{1}_{\left\{|X_n| \le n\right\}}.$$

Nyt

$$\begin{aligned}\sum \mathbf{P}\Big[X_n \neq \tilde{X}_n\Big] &= \sum \mathbf{P}[|\,|X_n| > n] \\ &= \sum \mathbf{P}[|X_1| > n] \\ &\leq \mathbf{E}[|X_1|] \\ &< \infty.\end{aligned}$$

Kayttämällä Borelin ja Cantellin lemmaa 5.2.1(i) näemme tästä, että

$$\mathbf{P}\Big[X_n \neq \tilde{X}_n \text{ ä.u.}\Big] = 0.$$

Siten riittää osoittaa, että

$$\frac{1}{n}\sum_{k\leq n} \tilde{X}_k \overset{m.v.}{\to} \mu.$$

Nimittäin jos $n_0(\omega) \in \mathbb{N}$ on sellainen, että $\tilde{X}_n(\omega) = X_n(\omega)$, kun $n \geq n_0(\omega)$, niin

$$\begin{aligned}\lim \frac{1}{n}\sum_{k\leq n} X_n(\omega) &= \lim \frac{1}{n}\sum_{k\leq n_0(\omega)} X_n(\omega) + \lim \frac{1}{n}\sum_{n_0(\omega)<k\leq n} X_n(\omega) \\ &= \lim \frac{1}{n}\sum_{n_0(\omega)<k\leq n} \tilde{X}_n(\omega) \\ &= \lim \frac{1}{n}\sum_{k\leq n} \tilde{X}_n(\omega).\end{aligned}$$

Nyt $\mathbf{E}[\tilde{X}_n] \to \mathbf{E}[X_1] = \mu$. Siten väite seuraa, jos

$$\frac{1}{n}\sum_{k\leq n}\Big(\tilde{X}_k - \mathbf{E}[\tilde{X}_k]\Big) \overset{m.v.}{\to} 0.$$

Perustelemme tämän käyttämällä vahvaa suurten lukujen lakia 8.1.9 neliöintegroituville satunnaismuuttujille. Katkaisumme takia satunnaismuuttujamme $\tilde{X}_n$, $n \in \mathbb{N}$, ovat neliöintegroituvia.

Koska

$$\mathbf{E}\Big[\Big(\tilde{X}_n - \mathbf{E}[\tilde{X}_n]\Big)^2\Big] \leq \mathbf{E}[\tilde{X}_n^2],$$

niin riittää osoittaa sarjaehto

$$\sum \frac{\mathbf{E}[\tilde{X}_n^2]}{n^2} < \infty.$$

Mutta

$$\begin{aligned}
\sum \frac{\mathbf{E}[\tilde{X}_n^2]}{n^2} &= \sum \frac{1}{n^2}\mathbf{E}\big[X_1^2 \mathbf{1}_{\{|X_1|\le n\}}\big] \\
&= \sum_{n\ge 1}\sum_{k\le n} \frac{1}{n^2}\mathbf{E}\big[X^2 \mathbf{1}_{\{k-1<|X_1|\le k\}}\big] \\
&= \sum_{k\ge 1}\sum_{n>k} \frac{1}{n^2}\mathbf{E}\big[X_1^2 \mathbf{1}_{\{k-1<|X_1|\le k\}}\big] \\
&= \sum_{k\ge 1}\mathbf{E}\big[X_1^2 \mathbf{1}_{\{k-1<|X_1|\le k\}}\big]\sum_{n>k}\frac{1}{n^2} \\
&\le \sum_{k\ge 1}\mathbf{E}\big[X_1^2 \mathbf{1}_{\{k-1<|X_1|\le k\}}\big]\,\frac{2}{k} \\
&= 2\sum_{k\ge 1}\mathbf{E}\left[\frac{X_1^2}{k}\mathbf{1}_{\{k-1<|X_1|\le k\}}\right] \\
&\le 2\sum_{k\ge 1}\mathbf{E}\left[\frac{X_1^2}{|X_1|}\mathbf{1}_{\{k-1<|X_1|\le k\}}\right] \\
&= 2\mathbf{E}[|X_1|] \\
&< \infty.
\end{aligned}$$

□

8.1.11 Esimerkki. Vahvan suurten lukujen lain sovelluksena käsittelemme *Monte Carlo -integrointia.*

Olkoon $g : \mathbb{R} \to \mathbb{R}$ Borel-mitallinen funktio. Haluamme laskea numeerisesti määrätyn integraalin

$$I \;:=\; \int_a^b g(x)\,\mathrm{d}x.$$

Olkoot $U_1, U_2, \ldots$ riippumattomia tasaisesti välille $[a, b]$ jakautuneita satunnaismuuttujia. Olkoon

$$X_k \;:=\; g(U_k)$$

ja

$$S_n \;:=\; \sum_{k\le n} X_k.$$

Tällöin

$$\mu \;:=\; \mathbf{E}[X_k] \;=\; \int_a^b g(x)\frac{1}{b-a}\,\mathrm{d}x \;=\; \frac{1}{b-a}\,I.$$

Mutta vahva suurten lukujen laki sanoo, että

$$(b-a)\frac{1}{n}S_n \overset{m.v.}{\to} (b-a)\mu = I.$$

8.1.12 Esimerkki. Toisena vahvan suurten lukujen lain sovelluksena käsittelemme tuntemattoman jakauman estimointia. Tilanne on seuraava: saamme riippumattomia havaintoja $X_1, X_2, \ldots$ tuntemattomasta jakaumasta $F(x) = \mathbf{P}[X \le x]$. Tehtävänä on selvittää kertymäfunktio F. Määrittelemme *empiirisen kertymäfunktion*

$$\hat{F}_n(x) := \frac{\big|\{k \le n\,;\, X_k \le x\}\big|}{n} = \frac{1}{n}\sum_{k \le n} \mathbf{1}_{\{X_k \le x\}}.$$

Nyt indikaattorimuuttujat $\mathbf{1}_{\{X_k \le x\}}$, $k \in \mathbb{N}$, ovat riippumattomia ja samoin jakautuneita, joten vahva suurten lukujen laki sanoo, että

$$\hat{F}_n(x) \overset{m.v.}{\to} \mathbf{E}\big[\mathbf{1}_{\{X_k \le x\}}\big] = \mathbf{P}[X_k \le x] = F(x).$$

Lopetamme tämän osion *Kolmogorovin kolmen sarjan lauseeseen.* Se karakterisoi riippumattomista summattavista koostuvan sarjan melkein varman suppenemisen.

8.1.13 Lause. *Olkoon (X_n) jono riippumattomia satunnaismuuttujia. Olkoon $\varepsilon > 0$. Merkitsemme*

$$X_n^\varepsilon = X_n \mathbf{1}_{\{|X_n| \le \varepsilon\}}.$$

Tällöin $\sum X_n$ suppenee melkein varmasti jos ja vain jos sarjat

$$\sum \mathbf{E}[X_n^\varepsilon], \quad \sum \mathbf{Var}[X_n^\varepsilon], \quad \text{ja} \quad \sum \mathbf{P}[X_n \ge \varepsilon]$$

suppenevat jollakin (ja siten kaikilla) $\varepsilon > 0$.

Todistus. Jätetään harjoitustehtäväksi (graduksi). □

8.2 Suurten poikkeamien periaate

Suurten poikkeamien periaate tutkii suurten lukujen lain suppenemisvauhtia. Olkoon (X_n) jono satunnaismuuttujia jotka toteuttavat heikon suurten lukujen lain

$$\frac{1}{n}S_n \overset{\mathbf{P}}{\to} \mu,$$

missä $S_n := \sum_{k\leq n} X_k$. Nyt siis

$$\mathbf{P}\left[\left|\frac{1}{n}S_n - \mu\right| \geq \varepsilon\right] \quad\to\quad 0.$$

Suurten poikkeamien periaate tarkastelee, kuinka nopeaa tämä suppeneminen on. Se sanoo, silloin kun se pätee, että suppeneminen on eksponentiaalisen nopeaa. Karkeasti sanottuna se sanoo, että

$$\mathbf{P}\left[\left|\frac{1}{n}S_n - \mu\right| \geq \varepsilon\right] \quad\simeq\quad e^{-nI(\varepsilon)},$$

missä I on suppenemiseen liittyvä vauhtifunktio.

Suurten poikkeamien teoria sisältää useita eri tuloksia. Tällä kurssilla esitämme ainoastaan klassisen *Cramérin lauseen.* Siinä vauhtifunktio I on summattavien satunnaismuuttujien yhteisen kumulantit generoivan funktion Λ Fenchel–Legendre-muunnos Λ^*.

8.2.1 Lause. *Olkoot (X_n) jono riippumattomia samoin jakautuneita satunnaismuuttujia, jotka toteuttavat* Cramérin ehdon*:*

$$\Lambda(\theta) \quad:=\quad \ln \mathbf{E}\big[e^{i\theta X_n}\big] \quad<\quad \infty$$

kaikilla $\theta \in \mathbb{R}$. Olkoon $\mu := \mathbf{E}[X_n]$ ja $S_n := \sum_{k\leq n} X_k$. Tällöin kaikille $x > \mu$ pätee

$$\lim \frac{1}{n}\ln \mathbf{P}\left[\frac{1}{n}S_n \geq x\right] \quad=\quad -\Lambda^*(x).$$

ja kaikille $x < \mu$ pätee

$$\lim \frac{1}{n}\ln \mathbf{P}\left[\frac{1}{n}S_n \leq x\right] \quad=\quad -\Lambda^*(x),$$

missä

$$\Lambda^*(x) \quad:=\quad \sup_{\theta\in\mathbb{R}^d}\{\theta x - \Lambda(\theta)\}$$

on kumulantit generoivan funktion Λ Fenchel–Legendre-muunnos.

Todistus. Aloitamme muutamilla huomioilla.

Ensinnäkin voimme olettaa, että satunaismuuttujat X_n, $n \in \mathbb{N}$, eivät ole degeneroituneita, sillä muuten väite on triviaali. Toiseksikin voimme olettaa, että $x = 0$ ja $\mu < 0$. Asetamme sitten

$$\rho \quad:=\quad \inf_{\theta\in\mathbb{R}} e^{\Lambda(\theta)}.$$

Tällöin $\Lambda^*(0) = -\log\rho$ ja siten tehtävänämme on itse asiassa osoittaa, että

$$\lim \frac{1}{n}\ln \mathbf{P}[S_n \geq 0] \;=\; \log\rho. \tag{8.2.2}$$

Kolmanneksikin huomaamme, että funktio e^{Λ} on aidosti konveksi ja

$$\frac{\mathrm{d}}{\mathrm{d}\theta}e^{\Lambda(0)} \;=\; \mathbf{E}[X_n] \;<\; 0.$$

Nyt voimme tarkastella kolmea tapausta:
Tapaus 1: $\mathbf{P}[X_n < 0] = 1$.
Tapaus 2: $\mathbf{P}[X_n \leq 0] = 1$ ja $\mathbf{P}[X_n = 0] > 0$.
Tapaus 3: $\mathbf{P}[X_n < 0]\mathbf{P}[X_n > 0] > 0$.
Tapausten 1 ja 2 käsittelemisen jätämme harjoitustehtäväksi 8.9.
Tarkastelemme sitten tapausta 3.
Huomaamme aluksi, että

$$\lim_{\theta\to\pm\infty} e^{\Lambda(\theta)} \;=\; \infty.$$

Koska e^{Λ} on aidosti konveksi, niin on olemassa yksikäsitteinen minimi $\tau \in \mathbb{R}$, jolle $\Lambda(\tau) = \log\rho$ ja $\Lambda'(\tau) = 0$. Siten, Tšebyševin epäyhtälön 4.4.3 ja satunnaismuuttujien X_n, $n \in \mathbb{N}$, riippumattomuuden nojalla,

$$\begin{aligned}
\mathbf{P}\left[\frac{1}{n}S_n \geq 0\right] &= \mathbf{P}\left[e^{\tau S_n} \geq 1\right] \\
&\leq \mathbf{E}\left[e^{\tau S_n}\right] \\
&= \mathbf{E}\left[e^{\tau X_1}\right]^n \\
&= e^{n\Lambda(\tau)} \\
&= \rho^n.
\end{aligned}$$

Siten

$$\limsup \frac{1}{n}\ln \mathbf{P}\left[\frac{1}{n}S_n \geq 0\right] \;\leq\; \log\rho.$$

Olemme siis osoittaneet väitteen (8.2.2) ylärajan.

Tarkastelemme sitten väitteen (8.2.2) alarajaa. Toisin sanoen haluamme osoittaa, että

$$\liminf \frac{1}{n}\ln \mathbf{P}\left[\frac{1}{n}S_n \geq 0\right] \;\geq\; \log\rho. \tag{8.2.3}$$

Kun väite (8.2.3) on todistettu, olemme todistaneet koko lauseen 8.2.1.

Väitteen (8.2.3) todistus perustuu seuraavaan mitanvaihtoon: olkoon τ kuten edellä ja olkoon F satunnaismuuttujien X_n, $n \in \mathbb{N}$, yhteinen kertymäfunktio. Asetamme uuden kertymäfunktion F_τ kaavalla

$$F_\tau(x) \;:=\; \frac{1}{\rho}\int_{-\infty}^{x} e^{\tau y}\,\mathrm{d}F(y).$$

On suoraviivaista tarkistaa, että F_τ on todellakin kertymäfunktio. Itse asiassa kertymäfunktiota F_τ kutsutaan kertymäfunktion F *Esscherin muunnokseksi* pisteessä τ.

Olkoot sitten $\mathbf{E}_\tau$ ja $\mathbf{P}_\tau$ odotusarvo ja todennäköisyys, jotka vastaavat kertymäfunktiota F_τ. Tällöin on helppo nähdä, että satunnaismuuttujat X_n, $n \in \mathbb{N}$, ovat riippumattomia ja samoin jakautuneita mitan $\mathbf{P}_\tau$ suhteen ja

$$\begin{aligned} 0 &\;=\; \mathbf{E}_\tau[X_n], \\ \sigma_\tau^2 &\;:=\; \mathbf{E}_\tau[X_n^2] \;<\; \infty. \end{aligned}$$

Lisäksi

$$\mathbf{P}\left[\frac{1}{n}S_n \geq 0\right] \;=\; \rho^n \mathbf{E}_\tau\left[\mathbf{1}_{\{S_n\geq 0\}}e^{-\tau S_n}\right] \tag{8.2.4}$$

(harjoitustehtävä 8.10). Osoitamme sitten, että

$$\liminf \frac{1}{n}\ln \mathbf{E}_\tau\left[\mathbf{1}_{\{S_n\geq 0\}}e^{-\tau S_n}\right] \;\geq\; 0. \tag{8.2.5}$$

Jos väite (8.2.5) pätee, niin yhdistämällä se väitteeseen (8.2.4) saamme alarajan (8.2.3) ja lause 8.2.1 on todistettu.

Enää siis pitää osoittaa (8.2.5). Perustelemme sen käyttämällä keskeistä raja-arvolausetta 8.3.1 (joka todistetaan vasta seuraavassa osiossa).

Olkoon

$$\Phi(x) \;:=\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}y^2}\,\mathrm{d}y$$

standardinormaalijakauman kertymäfunktio. Olkoon c sellainen vakio, että $\Phi(c) - \Phi(0) > 1/4$ (selvästi tällainen vakio on olemassa). Nyt, käyttämällä keskeistä raja-arvolausetta 8.3.1 saamme

$$\begin{aligned}
&\mathbf{E}_\tau\left[\mathbf{1}_{\{S_n\geq 0\}}e^{-\tau S_n}\right] \\
&= \mathbf{E}_\tau\left[\mathbf{1}_{\left\{\frac{1}{\sqrt{n}\sigma_\tau}\sum_{k\leq n}X_k\geq 0\right\}}e^{-\tau\sqrt{n}\sigma_\tau\frac{1}{\sqrt{n}\sigma_\tau}\sum_{k\leq n}X_k}\right] \\
&\geq \mathbf{E}_\tau\left[\mathbf{1}_{\left\{0\leq\frac{1}{\sqrt{n}\sigma_\tau}\sum_{k\leq n}X_k\leq c\right\}}e^{-\tau\sqrt{n}\sigma_\tau\frac{1}{\sqrt{n}\sigma_\tau}\sum_{k\leq n}X_k}\right] \\
&\geq e^{-\tau\sqrt{n}\sigma_\tau c}\mathbf{P}_\tau\left[0\leq\frac{1}{\sqrt{n}\sigma_\tau}\sum_{k\leq n}X_k\leq c\right] \\
&> \frac{1}{4}e^{-\tau\sqrt{n}\sigma_\tau c},
\end{aligned}$$

kun n on riittävän iso. Väite (8.2.5) seuraa tästä. □

Cramér toimi ruotsalaisen vakuutusyhtiön konsulttina. On siis vain oikein ja kohtuullista, että seuraavaksi sovellamme hänen lausettaan vakuutusyhtiöitä koskevaan ongelmaan.

8.2.6 Esimerkki. Oletetaan, että vakuutusyhtiö saa päivässä kiinteän ennalta tunnetun määrän p vakuutusmaksuja (preemioita). Joka päivä t vakuutusyhtiö saa satunnaisen määrän X_t korvausvaateita. Oletamme, että vaateet X_t, $t\in\mathbb{N}$, ovat riippumattomia ja samoin jakautuneita, ja että

$$\Lambda(\theta) \;:=\; \ln\mathbf{E}\left[e^{\theta X_n}\right] \;<\; \infty$$

kaikilla $\theta\in\mathbb{R}$.

Vakuutusyhtiön ongelma on nyt määrätä vakuutusmaksu niin, että vararikon todennäköisyys (ennen päivää T) on mahdollisimman pieni. Cramérin lause sanoo nyt, että vararikon todennäköisyys on

$$\mathbf{P}\left[\sum_{t\leq T}X_t > pT\right] \;\simeq\; e^{-T\Lambda^*(p)}.$$

Tehtäväksi jää siis ratkaista p yhtälöstä

$$\Lambda^*(p) \;=\; -\frac{\ln\varepsilon}{T},$$

missä ε on kiinnitetty riskitaso. Yhtälön ratkaiseminen vaatii tietysti funktion Λ^* tuntemista. Onneksi tämä funktio voidaan kuitenkin estimoida saapuneista korvaustvaateista.

8.3 Keskeinen raja-arvolause

Seuraava tulos on *keskeinen raja-arvolause* sen klassisessa, kohtalaisen yksinkertaisessa muodossa.

8.3.1 Lause. *Olkoon* $(X_n) \subset L^2$ *jono riippumattomia samoin jakautuneita satunnaismuuttujia. Olkoon* $\mu := \mathbf{E}[X_n]$ *ja* $\sigma^2 := \mathbf{Var}[X_n]$*. Tälloin*

$$\frac{1}{\sigma\sqrt{n}} \sum_{k \le n} (X_k - \mu) \quad \xrightarrow{d} \quad G,$$

missä G *on normaalisti jakautunut odotusarvolla* 0 *ja varianssilla* 1*.*

Seuraavaksi esitettävä todistus perustuu karakteristisiin funktioihin ja niiden kvadraattiseen aproksimointiin. Keskeiset huomiot ovat:

- Neliöintegroituvan satunnaismuuttujan $X \in L^2$ karakteristiselle funktiolle ϕ_X pätee

$$\begin{aligned} \phi_X(u) &\simeq \phi_X(0) + \frac{\mathrm{d}}{\mathrm{d}u}\phi_X(0)u + \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}u^2}\phi_X(0)u^2 \\ &= 1 + i\mathbf{E}[X]u - \frac{1}{2}\mathbf{E}[X^2]u^2. \end{aligned}$$

- Jos $X_1, \ldots, X_n$ ovat riippumattomia samoin jakautuneita satunnaismuuttujia, niin

$$\begin{aligned} \phi_{X_1+\cdots+X_n}(u) &= \phi_{X_1}(u) \cdots \phi_{X_n}(u) \\ &= \phi_{X_1}(u)^n. \end{aligned}$$

Lauseen 8.3.1 todistus. Olkoon ϕ satunnaismuuttujien $X_n - \mu$ yhteinen karakteristinen funktio. Koska $\mathbf{E}[X_n - \mu] = 0$ ja $\mathbf{E}[(X_n - \mu)^2] = \sigma^2$, niin lauseen 6.3.6(ii) nojalla

$$\phi(u) \;=\; 1 - \frac{\sigma^2}{2}u^2 + u^2\varepsilon(u),$$

missä $\varepsilon(u) \to 0$, kun $u \to 0$. Osasumman

$$\sum_{k \le n} (X_k - \mu)$$

karakteristinen funktio on riippumattomuuden nojalla ϕ^n. Lisäksi normeeratun osasumman

$$\frac{1}{\sigma\sqrt{n}} \sum_{k \le n} (X_k - \mu)$$

karakteristinen funktio ϕ_n on

$$\begin{aligned}\phi_n(u) &= \phi\left(\frac{u}{\sqrt{n}\sigma}\right)^n \\ &= e^{n\ln\phi\left(\frac{u}{\sigma\sqrt{n}}\right)} \\ &= e^{n\ln\left(1-\frac{u^2}{2n}+\frac{u^2}{n\sigma^2}\varepsilon\left(\frac{u}{\sigma\sqrt{n}}\right)\right)}.\end{aligned}$$

Nyt, kun $u \in \mathbb{R}$ on kiinteä, niin jätämme harjoitustehtäväksi 8.11 osoittaa, että

$$\lim_{n\to\infty}\left(n\ln\left(1-\frac{u^2}{2n}+\frac{u^2}{n\sigma^2}\varepsilon\left(\frac{u}{\sigma\sqrt{n}}\right)\right)\right) = -\frac{1}{2}u^2.$$

Siten, koska eksponenttifunktio on jatkuva

$$\lim_{n\to\infty}\phi_n(u) = e^{-\frac{1}{2}u^2}.$$

Mutta tämä on standardinormaalisti jakautuneen satunnaismuuttujan karakteristinen funktio. Väite seuraa siten Lévyn jatkuvuuslauseesta 7.4.7(ii). □

Klassinen esimerkki keskeisen raja-arvolauseen soveltamisesta on tietysti niin sanottu *normaaliaproksimaatio*. Seuraava esimerkki käsittelee tätä.

8.3.2 Esimerkki. Oletamme, että elektronisen komponentin elinikä on satunnaismuuttuja, jonka odotusarvo on on a ja keskihajonta, eli varianssin neliöjuuri, on samoin a. Tehtävänä on ratkaista kuinka monta (riippumatonta) komponenttia tarvitaan, jotta niiden yhteenlaskettu elinikä S olisi alle $8a$ korkeintaan todennäköisyydellä 5%. Varma epäyhtälö saataisiin tietysti käyttämälle Markovin epäyhtälöä. Jos kuitenkin tyydymme "hieman epävarmaan" epäyhtälöön, niin voimme käyttää normaaliaproksimaatiota ja saada Markovin epäyhtälöä tarkemman tuloksen. Nimittäin keskeisen raja-arvolauseen nojalla normeerattu yhteenlaskettu elinikä

$$\frac{S-na}{a\sqrt{n}}$$

on likimain normaalijakautunut parametrein $\mu = 0$ ja $\sigma^2 = 1$. Siten

$$\mathbf{P}[S \le 8a] = \mathbf{P}\left[\frac{S-na}{a\sqrt{n}} \le \frac{8-n}{\sqrt{n}}\right] \simeq \Phi\left(\frac{8-n}{\sqrt{n}}\right),$$

missä

$$\Phi(x) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}y^2}\,\mathrm{d}y$$

on standardinormaalijakautuneen satunnaismuuttujan kertymäfunktio. Saamme siis aproksimatiivisen epäyhtälön

$$\Phi\left(\frac{8-n}{\sqrt{n}}\right) \;\leq\; 0{,}05$$

eli

$$\frac{8-n}{\sqrt{n}} \;\leq\; \Phi^{-1}(0{,}05),$$

mistä, katsomalla esimerkiksi taulukosta standardinormaalijakauman *fraktiilifunktion* Φ^{-1} arvoja, samme ehdon $n \geq 15$.

Osion lopuksi esitämme *Lindebergin ja Fellerin keskeisen raja-arvolauseen.* Se on tietyssä mielessä täydellinen vastaus keskeiseen raja-arvo-ongelmaan. Toisin sanoen se antaa riittävät ja välttämättömät ehdot sille, koska satunnaismuuttujajono (X_n) suppenee heikosti kohti normaalijakaumaa.

8.3.3 Lause. *Olkoot $\xi_{n,l}$, $n \in \mathbb{N}$, $l \leq k_n$ satunnaismuuttujia. Oletamme, että $\xi_{n,1}, \ldots, \xi_{n,k_n}$ ovat riippumattomia jokaisella $n \in \mathbb{N}$, $\lim k_n = \infty$, ja*

$$\begin{aligned} \mathbf{E}[\xi_{n,l}] &= 0, \\ \sum_{l \leq k_n} \mathbf{Var}[\xi_{n,l}] &= 1. \end{aligned}$$

Jos $\xi_{n,l}$, $n \in \mathbb{N}$, $l \leq k_n$ toteuttavat Lindebergin ehdon

$$\lim \sum_{l \leq k_n} \mathbf{E}\left[\xi_{n,l}^2 \mathbf{1}_{\{|\xi_{n,l}| \geq \varepsilon\}}\right] \;=\; 0 \tag{8.3.4}$$

kaikilla $\varepsilon > 0$, niin

$$\sum_{l \leq k_n} \xi_{n,l} \;\xrightarrow{d}\; G, \tag{8.3.5}$$

missä G on standardinormaalisti jakautunut.

Lisäksi, jos

$$\lim \max_{l \leq k_n} \mathbf{P}[\xi_{n,l} \geq \varepsilon] \;=\; 0$$

kaikilla $\varepsilon > 0$, niin ehto (8.3.4) on välttämätön keskeiselle raja-arvolauseelle (8.3.5).

Todistus. Sivuutetaan. □

Harjoitustehtäviä

8.1. Olkoot $X_1, \ldots, X_n$ satunnaismuuttujia, joille $X_k^2 \in L^1$ kaikille $k \leq n$. Osoita, että

(a)
$$\mathbf{Var}\left[\sum_{k \leq n} X_k\right] \;=\; \sum_{k \leq n} \mathbf{Var}[X_k] + 2 \sum_{k < l \leq n} \mathbf{Cov}[X_k, X_l].$$

(b) jos satunnaismuuttujat $X_1, \ldots, X_n$ ovat korreloimattomia, niin
$$\mathbf{Var}\left[\sum_{k \leq n} X_k\right] \;=\; \sum_{k \leq n} \mathbf{Var}[X_k].$$

8.2. Olkoot $X, Y \in L^2$. Osoita, että

(a) $|\mathbf{Cor}[X, Y]| \leq 1$,

(b) jos $\mathbf{Cor}[X, Y] = 1$, niin on olemassa sellaiset luvut $a, b \in \mathbb{R}$, että $Y \overset{m.v}{=} aX + b$.

8.3. Muotoile ja todista heikko suurten lukujen laki 8.1.1 satunnaisvektoreille.

8.4. Olkoon $(X_n) \subset L^2$ jono korreloimattomia ja samoin jakautuneita satunnaismuuttujia. Osoita, että tällöin
$$\frac{1}{n} \sum_{k \leq n} X_k \;\overset{m.v.}{\longrightarrow}\; \mathbf{E}[X_1].$$

8.5. Todista Skorohodin lemma 8.1.4.

8.6. Todista seuraus 8.1.7.

8.7. Todista Kroneckerin lemma 8.1.8.

8.8. Osoita, että lauseen 8.1.10 sarja hajaantuu melkein varmasti, jos $\mathbf{E}[|X_1|] = \infty$.

8.9. Käsittele Cramérin lauseen 8.2.1 todistuksen kohdat 1 ja 2.

8.10. Käsittele Cramérin lauseen 8.2.1 todistuksen väite (8.2.4).

8.11. Osoita keskeisen raja-arvolauseen 8.3.1 todistuksessa käytetty väite
$$\lim_{n \to \infty} \left(n \ln \left(1 - \frac{u^2}{2n} + \frac{u^2}{n\sigma^2} \varepsilon \left(\frac{u}{\sigma \sqrt{n}} \right) \right) \right) \;=\; -\frac{1}{2} u^2.$$

Luku 9

Ehdollinen odotusarvo

Yksi elämän keskeisimmistä ongelmista on päätöksenteko epävarmuuden vallitessa eli optimaalinen arvaaminen tai ennustaminen. Ehdollinen odotusarvo antaa yhden vastauksen tähän ongelmaan. Se kertoo, mitä "keskimäärin" tapahtuu, kun otamme huomioon, mitä tiedämme. Käsite "mitä tiedämme" on matematisoitu kaiken mahdollisen informaation $\mathcal{F}$ *ali-*σ*-algebraksi* $\mathcal{G}$ *ja "keskimäärin" tarkoittaa luonnollisesti odotusarvomielessä.*

Tarkastelemme tässä luvussa ainoastaan satunnaismuuttujia, vaikka vastaavat tulokset pätevät myös satunnaisvektoreille.

9.1 Johdattelua

Ehdollisen odotusarvon idea perustuu tietysti *ehdolliseen todennäköisyyteen.*

9.1.1 Määritelmä. Olkoon $E \in \mathcal{F}$ sellainen, että $\mathbf{P}[E] > 0$. Tapahtuman $A \in \mathcal{F}$ *ehdollinen todennäköisyys* ehdolla E on

$$\mathbf{P}[A|E] \;:=\; \frac{\mathbf{P}[A \cap E]}{\mathbf{P}[E]}.$$

Ehdollisen todennäköisyyden idea perustuu tietysti tulosääntöön

$$\mathbf{P}[A \cap E] \;=\; \mathbf{P}[E]\mathbf{P}[A|E].$$

Naiivi tapa määritellä ehdollinen odotusarvo on tulkita se integraalina, eli odotusarvona, ehdollisen todennäköisyyden suhteen. Tätä tulkintaa varten meidän on tulkittava ehdollinen todennäköisyys todenäköisyytenä. Seuraava apulause sanoo, että tämä on mahdollista.

9.1.2 Apulause. *Olkoon* $(\Omega, \mathcal{F}, \mathbf{P})$ *todennäköisyysavaruus ja* $E \in \mathcal{F}$ *sellainen, että* $\mathbf{P}[E] > 0$. *Tällöin kuvaus*

$$A \quad \mapsto \quad \mathbf{P}[A|E], \qquad A \in \mathcal{F},$$

on todennäkäköisyysmitta avaruudella $(\Omega, \mathcal{F})$.

Todistus. Harjoitustehtävä 9.1. □

Tarkastelemme sitten satunnaismuuttujia X ja Y. Oletamme, että Y on diskreetti. Toisin sanoen oletamme, että on olemassa sellainen numeroituva joukko $\{y_1, y_2, \ldots\} \subset \mathbb{R}$, että $\sum \mathbf{P}[Y = y_n] = 1$ ja $\mathbf{P}[Y = y_n] > 0$ kaikilla $n \in \mathbb{N}$. Huomaamme, että apulauseen 9.1.2 nojalla kuvaus

$$B \quad \mapsto \quad \mathbf{P}[X \in B | Y = y]$$

on jokaisella $y \in \{y_1, y_2, \ldots\}$ todennäköisyysmitta avaruudella $(\mathbb{R}, \mathcal{B})$. Siten voimme määritellä ehdollisen odotusarvon jokaiselle $y \in \{y_1, y_2, \ldots\}$ asettamalla muuttujanvaihtokaavan 4.3.1 ehdottamalla tavalla

$$\mathbf{E}[X|Y = y] \quad := \quad \int_{\mathbb{R}} x \, \mathbf{P}[X \in \mathrm{d}x | Y = y].$$

Tässä siis ehdollinen todennäköisyys

$$\mathbf{P}[X \in B | Y = y] \quad = \quad \frac{\mathbf{P}[X \in B, Y = y]}{\mathbf{P}[Y = y]}, \qquad B \in \mathcal{B},$$

on melkein varmasti hyvinmääritelty, sillä $\mathbf{P}[Y = y] > 0$, kun $y \in \{y_1, y_2, \ldots\}$ ja $\mathbf{P}[Y = y_n \text{ jollakin } n \in \mathbb{N}] = 1$. Jos Y ei ole diskreetti, niin yllä oleva määrittely ei toimi.

Käyttämättä muuttujanvaihtokaavaa 4.3.1 voimme määritellä ehdollisen odotusarvon myös satunnaismuuttujana:

$$\mathbf{E}[X|Y](\omega) \quad := \quad \int_{\Omega} X(\omega') \, \mathbf{P}[\mathrm{d}\omega' | Y](\omega), \tag{9.1.3}$$

missä $\mathbf{P}[\cdot|Y]$ on *satunnaismitta*:

$$\mathbf{P}[A|Y](\omega) \quad = \quad \sum \mathbf{P}[A|Y = y_n] \mathbf{1}_{\{Y = y_n\}}(\omega).$$

Satunnaismitta $\mathbf{P}[\cdot|Y]$ on siis todennäköisyysmitta-arvoinen satunnaismuuttuja, joka on todennäköisyysmitta, kun satunnaismuuttujan Y arvo on annettu. Ehdollinen odotusarvo $\mathbf{E}[X|Y]$ on siis satunnaismuuttuja, jonka arvo on melkein varmasti tunnettu, kun tiedämme satunnaismuuttujan Y arvon.

Kaava (9.1.3) on epäilemättä hieman vaikea lukea. Syy tähän ei ole niinkään notaatio-ongelma, vaan pikemminkin on kyse siitä, että lähestymme vaikeita aiheita. Seuraava esimerkki toivottavasti kuitenkin helpottaa hämmentynyttä lukijaa.

9.1.4 Esimerkki. Tarkastelemme n-kertaista perättäistä riippumatonta toistokoetta, jossa yksittäisen onnistumisen todennäköisyys on $p \in (0,1)$. Olkoon X onnistumisten lukumäärä ja Y tapahtuman n. toisto onnistui indikaattori. Tällöin siis X on binomijakautunut parametrein n ja p ja Y on binomijakautunut parametrein 1 ja p.

Määräämme ehdollisen odotusarvon $\mathbf{E}[X|Y]$.

On selvää, että jos $Y = 0$, niin X on binomijakautunut parametrein $n-1$ ja p. Samoin on selvää, että jos $Y = 1$, niin $X - 1$ on binomijakautunut parametrein $n-1$ ja p. Siten

$$\begin{aligned}
&\mathbf{P}[X = k|Y] \\
&\quad = \binom{n-1}{k} p^k (1-p)^{n-1-k} \mathbf{1}_{\{Y=0\}} + \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} \mathbf{1}_{\{Y=1\}}
\end{aligned}$$

ja

$$\mathbf{E}[X|Y] \;=\; (n-1)p\,\mathbf{1}_{\{Y=0\}} + \big((n-1)p+1\big)\,\mathbf{1}_{\{Y=1\}}.$$

Haluamme sitten määritellä ehdollisen odotusarvon $\mathbf{E}[X|Y]$, kun Y on yleinen, ei välttämättä diskreetti, satunnaismuuttuja. Luonnollinen lähestymistapa olisi määritellä $\mathbf{E}[X|Y = y]$ integraalina ehdollisen todennäköisyyden $\mathbf{P}[\cdot|Y = y]$ suhteen, missä ehdollinen todennäköisyys $\mathbf{P}[\cdot|Y = y]$ on määritelty raja-arvona

$$\mathbf{P}[A|Y = y] \;:=\; \lim_{h \to 0+} \frac{\mathbf{P}\big[A, |Y - y| \le h\big]}{\mathbf{P}\big[|Y - y| \le h\big]}. \tag{9.1.5}$$

Tässä lähestymistavassa on seuraava ongelma: jos $\mathbf{P}[Y = y] = 0$, niin mikään ei takaa, että raja-arvo (9.1.5) on olemassa.

9.1.6 Esimerkki. Lähetymistapa (9.1.5) ei ole kuitenkaan aivan toivoton. Jos esimerkiksi oletamme, että pari (X, Y) on jatkuvasti jakautunut yhteistiheyfunktiolla $f_{X,Y}$, niin voidaan osoittaa että

$$\mathbf{P}[X \in B|Y = y] \;=\; \frac{1}{f_Y(y)} \int_B f_{X,Y}(x, y)\,\mathrm{d}x$$

Nimittäin vaihtamalla integroinnin ja raja-arvon järjestystä saamme, käyttämällä l'Hospitalin kaavaa, että

$$\begin{aligned}
\mathbf{P}[X \in B | Y = y] &= \lim_{h \downarrow 0} \frac{\mathbf{P}[X \in B, |Y - y| \leq h]}{\mathbf{P}[|Y - y| \leq h]} \\
&= \lim_{h \downarrow 0} \frac{\int_B \int_{y-h}^{y+h} f_{X,Y}(x, y')\, \mathrm{d}y' \mathrm{d}x}{\int_{y-h}^{y+h} f_Y(y')\, \mathrm{d}y'} \\
&= \lim_{h \downarrow 0} \int_B \frac{\int_{y-h}^{y+h} f_{X,Y}(x, y') \mathrm{d}y'}{\int_{y-h}^{y+h} f_Y(y')\, \mathrm{d}y'}\, \mathrm{d}x \\
&= \int_B \lim_{h \downarrow 0} \frac{\int_{y-h}^{y+h} f_{X,Y}(x, y') \mathrm{d}y'}{\int_{y-h}^{y+h} f_Y(y')\, \mathrm{d}y'}\, \mathrm{d}x \\
&= \int_B \frac{f_{X,Y}(x, y)}{f_Y(y)}\, \mathrm{d}x.
\end{aligned}$$

Esimerkin 9.1.6 nojalla naiivi lähestymistapa (9.1.5) toimii siis sillä lisäoletuksella, että (X, Y) on jatkuvasti jakautunut. Voisimme sitten yrittää naiivilla lähestymistavalla heikentämällä lisäoletuksia. Tämä ei olisi kuitenkaan kovin eleganttia. Ongelmalle on kuitenkin olemassa elegantti ratkaisu. Ongelma kierretään Radonin ja Nikodymin derivaattojen avulla. Idea perustuu siihen, että voimme tulkita raja-arvoon (9.1.5) liittyvän ehdollisen odotusarvon tiettyjen (merkkisten) mittojen Radonin ja Nikodymin derivaataksi. Derivoinnin käyttö on itse asiassa varsin luonnollista. Nimittäin esimerkin 9.1.6 tulos on formaalissa muodossa

$$\mathbf{P}[X \in \mathrm{d}x | Y = y] = \frac{1}{f_Y(y)} f_{X,Y}(x, y) \mathrm{d}x.$$

Käyttämällä formaaleja kaavoja

$$\begin{aligned}
f_{X,Y}(x, y)\, \mathrm{d}x \mathrm{d}y &= \mathbf{P}[X \in \mathrm{d}x, Y \in \mathrm{d}y], \\
f_Y(y)\, \mathrm{d}y &= \mathbf{P}[Y \in \mathrm{d}y],
\end{aligned}$$

näemme, että esimerkin 9.1.6 tulos on itse asiassa mittoja koskeva derivointitulos:

$$\begin{aligned}
\mathbf{P}[X \in \mathrm{d}x | Y = y] &= \frac{f_{X,Y}(x, y) \mathrm{d}x}{f_Y(y)} \\
&= \frac{f_{X,Y}(x, y) \mathrm{d}x \mathrm{d}y}{f_Y(y) \mathrm{d}y} \\
&= \frac{\mathbf{P}[X \in \mathrm{d}x, Y \in \mathrm{d}y]}{\mathbf{P}[Y \in \mathrm{d}y]}.
\end{aligned}$$

Yllä oleva pyörittely oli tietysti täysin formaalia ja matemaattisesti epäpätevää. Matemaattisesti täsmällisesti se saadaan tehtyä käyttämällä Radonin ja Nikodymin derivaattoja.

Rakennamme ehdollisen odotusarvon Radonin ja Nikodymin derivaattojen avulla seuraavassa osiossa. Itse asiassa seuraavassa osiossa määritellään ehdollinen odotusarvo ali-σ-algebran suhteen: määrittelemme otuksen $\mathbf{E}[X|\mathcal{G}]$. Yhteys ehdolliseen odotusarvoon $\mathbf{E}[X|Y]$ on tällöin $\mathbf{E}[X|Y] = \mathbf{E}[X|\sigma(Y)]$. Yhteys ehdolliseen todennäköisyyteen $\mathbf{P}[A|Y]$ on periaatteessa

$$\mathbf{P}[A|Y] \;=\; \mathbf{E}\left[\mathbf{1}_A|\sigma(Y)\right]. \tag{9.1.7}$$

Selvennämme nyt hieman, mitä tarkoitamme kun sanomme, että kaava (9.1.7) pätee “periaatteessa”. Elegantista Radonin ja Nikodymin lähestymistavasta joudumme maksamaan hinnan: ehdollinen odotusarvo $\mathbf{E}[\mathbf{1}_A|\sigma(Y)]$ on aina olemassa, mutta jos $\mathbf{P}[\cdot|Y]$ määritellään kaavalla (9.1.7), niin se ei ole välttämättä satunnaismitta. Toisin sanoen, kun $\omega \in \Omega$ on kiinnitetty, niin kuvaus $\mathbf{P}[\cdot|Y](\omega)$ ei ole välttämättä mitta. Se on sitä vain melkein varmasti eli melkein kaikilla $\omega \in \Omega$. Valitettavasti nyt melkein varmuus ei riitä, sillä ω on satunnaismitassa $\mathbf{P}[\cdot|Y](\cdot)$ kahdessa roolissa: sekä satunnaismuuttujan argumenttina että mitan joukon alkiona. Tämä ongelma voidaan kuitenkin monissa mielenkiintoisissa tilanteissa ratkaista käyttämällä säännöllisiä ehdollisia jakaumia. Näistä puhumme hieman viimeisessä osiossa.

9.2 Määritelmä Radonin ja Nikodymin derivaattana

Ennen kuin pääsemme derivoimaan todennäköisyysmittoja tarvitsemme yhden määritelmäkokoelman.

9.2.1 Määritelmä. Olkoot μ ja ν mittoja mitallisella avaruudella $(\Omega, \mathcal{F})$.

(a) Mitta ν on *absoluuttisesti jatkuva* mitan μ suhteen, jos kaikille $A \in \mathcal{F}$ pätee: $\nu[A] = 0$ aina, kun $\mu[A] = 0$. Tällöin merkitsemme $\nu \ll \mu$.

(b) Mitta ν on *ekvivalentti* mitan μ suhteen, jos $\nu \ll \mu$ ja $\mu \ll \nu$. Tällöin merkitsemme $\nu \sim \mu$.

(c) Mitat ν ja μ ovat *singulaarisia*, jos on olemassa sellainen joukko $A \in \mathcal{F}$, että $\nu[A^c] = 0$ ja $\mu[A'] = 0$ kaikilla $A' \subset A$. Tällöin merkitsemme $\nu \perp \mu$.

Itse asiassa emme tarvitse jatkossa määritelmän 9.2.1 kohtia (b) tai (c). Esitimme ne tässä vain täydellisyyden vuoksi.

Seuraava tulos on *Radonin ja Nikodymin lause.*

9.2.2 Lause. *Olkoon ν ja μ mittoja mitallisella avaruudella $(\Omega, \mathcal{F})$. Tällöin $\nu \ll \mu$ jos ja vain jos*

$$\nu[A] \;=\; \int_A Z(\omega)\,\mu[\mathrm{d}\omega] \tag{9.2.3}$$

jollekin mitalliselle kuvaukselle $Z : \Omega \to \mathbb{R} \cup \{\infty\}$.

Todistus. Sivuutetaan ajanpuutteen vuoksi. □

Radonin ja Nikodymin lauseen 9.2.2 satunnaismuuttujaa Z kutsutaan mitan ν *Radonin ja Nikodymin derivaataksi* mitan μ suhteen. Käytämme myös merkintää

$$\frac{\mathrm{d}\nu}{\mathrm{d}\mu} \;:=\; Z.$$

Tämän merkinnän syy on seuraava formalismi: Radonin ja Nikodymin derivaatan määräävän kaavan (9.2.3) voi kirjoittaa formaalisti muodossa

$$\nu[\mathrm{d}\omega] \;=\; Z(\omega)\,\mu[\mathrm{d}\omega]$$

tai lyhyemmin

$$\mathrm{d}\nu \;=\; Z\,\mathrm{d}\mu.$$

Mutta tällöin formaali sievennys sanoo, että

$$Z \;=\; \frac{\mathrm{d}\nu}{\mathrm{d}\mu}.$$

Seuraava esimerkki tulkitsee pistetodennäköisyysfunktion ja tiheysfunktion Radonin ja Nikodymin derivaatoiksi.

9.2.4 Esimerkki. (i) Tarkastelemme mitallista avaruutta $(\mathbb{R}, \mathcal{B})$ varustettuna Lebesguen mitalla ℓ. Olkoon $f : \mathbb{R} \to \mathbb{R}_+$ muotoa

$$\begin{aligned} f(x) - f(x_0) \;&=\; \int_{x_0}^{x} f'(y)\,\ell(\mathrm{d}y) \\ &=\; \int_{x_0}^{x} f'(y)\,\mathrm{d}y \end{aligned}$$

kaikilla $x_0 \in \mathbb{R}$, missä $f' : \mathbb{R} \to \mathbb{R}^+$. Tällöin kuvaus

$$\mu_f[A] \;:=\; \int_A f'(y)\,\ell(\mathrm{d}y), \qquad A \in \mathcal{B}_+$$

on mitta ja sen Radonin ja Nikodymin derivaatta Lebesguen mitan suhteen on funktion f tavanomainen derivaatta:

$$\frac{\mathrm{d}\mu_f}{\mathrm{d}\ell}(x) \ = \ = \ \frac{\mathrm{d}f}{\mathrm{d}x}(x) \ = \ f'(x).$$

Mikäli

$$\int_{-\infty}^{\infty} f'(y)\,\ell(\mathrm{d}y) \ = \ 1,$$

niin ν_f on todennäköisyysmitta ja f' on siihen liittyvä tiheysfunktio. Todennäköisyysittaa ν_f vastaava kertymäfunktio on

$$F(x) \ = \ \nu_f\big[(-\infty, x]\big].$$

(ii) Tarkastelemme mitallista avaruutta $(\mathbb{N}, \mathrm{pot}\,(\mathbb{N}))$. Olkoon $\mathfrak{n}$ laskurimittaa:

$$\mathfrak{n}(A) \ = \ |A|, \qquad A \subset \mathbb{N}.$$

Olkoon $p = (p_1, p_2, \ldots)$ jono positiivisia reaalilukuja. Olkoon $s = (s_1, s_2, \ldots)$ jonon p osasummien jono:

$$s_n \ := \ \sum_{k \leq n} p_k.$$

Tällöin s on mitta avaruudella $(\mathbb{N}, \mathrm{pot}\,(\mathbb{N}))$, kun tulkitsemme

$$s[A] \ := \ \sum_{k \in A} p_k, \qquad A \subset \mathbb{N}.$$

Lisäksi s:n Radonin ja Nikodymin derivaatta laskurimitan $\mathfrak{n}$ suhteen on p:

$$\left(\frac{\mathrm{d}s}{\mathrm{d}\mathfrak{n}}\right)_n \ = \ p_n.$$

Mikäli $\sum p_n = 1$, niin s on todennäköisyysmitta ja p on siihen liittyvä pistetodennäköisyysfunktio. Todennäköisyysmittaa s vastaava kertymäfunktio on

$$F(x) \ = \ \sum s_n \mathbf{1}_{(n-1,n]}(x).$$

Käyttämällä Radonin ja Nikodymin lausetta voimme määritellä ehdollisen odotusarvon, kun ehtona on jokin ali-σ-algebra $\mathcal{G} \subset \mathcal{F}$. Nimittäin olkoon X integroituva satunnaismuuttuja ja $\mathcal{G} \subset \mathcal{F}$ ali-σ-algebra. Tällöin voimme määritellä äärelliset mitat $\mathbf{Q}_X^+$ ja $\mathbf{Q}_X^-$ mitalliselle avaruudelle $(\Omega, \mathcal{G})$ asettamalla

$$\mathbf{Q}_X^{\pm}[A] \ := \ \mathbf{E}\big[X^{\pm}\mathbf{1}_A\big], \qquad A \in \mathcal{G}.$$

Tällöin on helppo nähdä, että $\mathbf{Q}_X^{\pm} \ll \mathbf{P}|\mathcal{G}$, missä $\mathbf{P}|\mathcal{G}$ on mitan $\mathbf{P}$ rajoittuma ali-σ-algebralle $\mathcal{G}$:

$$(\mathbf{P}|\mathcal{G})[A] \;:=\; \mathbf{P}[A], \qquad A \in \mathcal{G}.$$

Radonin ja Nikodymin lauseen 9.2.2 nojalla on siten olemassa sellaiset $\mathbf{P}$-melkein varmasti yksikäsitteiset $\mathcal{G}$-mitalliset positiiviset satunnaismuuttujat (Radonin ja Nikodymin derivaatat) Z^+ ja Z^-, että

$$\begin{aligned}
\mathbf{E}[X\mathbf{1}_A] &= \mathbf{E}\big[X^+\mathbf{1}_A\big] - \mathbf{E}\big[X^-\mathbf{1}_A\big] \\
&= \mathbf{Q}_X^+[A] - \mathbf{Q}_X^-[A] \\
&= \mathbf{E}\big[Z^+\mathbf{1}_A\big] - \mathbf{E}\,\big[Z^-\mathbf{1}_A\big] \\
&= \mathbf{E}\big[(Z^+ - Z^-)\mathbf{1}_A\big] \\
&=: \mathbf{E}[Z\mathbf{1}_A]
\end{aligned}$$

kaikilla $A \in \mathcal{G}$. Nyt on luonnollista valita merkkinen Radonin ja Nikodymin derivaatta Z ehdolliseksi odotusarvoksi $\mathbf{E}[X|\mathcal{G}]$. Ehdollinen odotusarvo voidaan siis karakterisoida seuraavalla operationaalisella määritelmällä, joka on peräisin Kolmogorovilta.

9.2.5 Määritelmä. Olkoon $\mathcal{G} \subset \mathcal{F}$ ali-σ-algebra. Satunnaismuuttujan $X \in L^1$ *ehdollinen odotusarvo* on se melkein varmasti yksikäsitteinen satunnaismuuttuja $\mathbf{E}[X|\mathcal{G}]$, jolle

$$\mathbf{E}\Big[\mathbf{E}[X|\mathcal{G}]\mathbf{1}_A\Big] \;=\; \mathbf{E}\Big[X\mathbf{1}_A\Big]$$

kaikilla $A \in \mathcal{G}$.

Huomattavaa on, että ehdollinen odotusarvo on vain melkein varmasti määritelty. Siten sitä koskevat kaavat tulee ymmärtää melkein varmassa muodossa. Emme kuitenkaan jatkossa mainitse tätä eksplisiittisesti, vaan oletamme lukijan ymmärtävän, että esimerkiksi kaava $\mathbf{E}[X|\mathcal{G}] = Y$ tarkoittaa $\mathbf{E}[X|\mathcal{G}] \overset{m.v.}{=} Y$.

Ehdollinen odotusarvo $\mathbf{E}[X|Y]$, missä Y on satunnaismuuttuja määritellään nyt luonnolliseen tapaan:

$$\mathbf{E}[X|Y] \;:=\; \mathbf{E}[X|\sigma(Y)].$$

Ehdollinen odotusarvo $\mathbf{E}[X|Y]$ on siis $\sigma(Y)$-mitallinen satunnaismuuttuja. Yhteys funktioon $y \mapsto \mathbf{E}[X|Y=y]$ on tietysti

$$\mathbf{E}[X|Y](\omega) \;=\; \sum_{y\in\mathbb{R}} \mathbf{E}[X|Y=y]\mathbf{1}_{\{Y=y\}}(\omega).$$

Käyttämällä Kolmogorovin operationaalista määritelmää on helppo todistaa ehdollisen odotusarvon ominaisuuksia (muutenhan emme kutsuisi määritelmää “operationaaliseksi”).

Osoitamme aluksi, että ehdollinen odotusarvo on positiivinen lineaarinen projektio-operaattori.

9.2.6 Lause. *Olkoot* $a, b \in \mathbb{R}$, $X, Y \in L^1$ *ja* $\mathcal{H}, \mathcal{G} \subset \mathcal{F}$ *ali-*σ*-algebroja. Tällöin*

(i) $\mathbf{E}[X|\mathcal{G}] \geq 0$, *kun* $X \geq 0$,
(ii) $\mathbf{E}[aX + bY|\mathcal{G}] = a\mathbf{E}[X|\mathcal{G}] + b\mathbf{E}[Y|\mathcal{G}]$,
(iii) $\mathbf{E}\big[\,\mathbf{E}[X|\mathcal{H}]\,\big|\,\mathcal{G}\big] = \mathbf{E}[X|\mathcal{H} \cap \mathcal{G}]$.

Todistus. (i) Kolmogorovin määritelmän nojalla pitää osoittaa, että

$$\mathbf{E}[X\mathbf{1}_A] \quad \geq \quad 0$$

kaikilla $A \in \mathcal{G}$. Mutta tämä on selvää, sillä $X\mathbf{1}_A \geq 0$, kun $X \geq 0$.

(ii) On selvää, että $a\mathbf{E}[X|\mathcal{G}] + b\mathbf{E}[Y|\mathcal{G}]$ on $\mathcal{G}$-mitallinen. Siten pitää osoittaa vain, että

$$\mathbf{E}\Big[\,\big(a\mathbf{E}[X|\mathcal{G}] + b\mathbf{E}[Y|\mathcal{G}]\big)\mathbf{1}_A\,\Big] \quad = \quad \mathbf{E}\Big[\,\big(aX + bY\big)\mathbf{1}_A\,\Big]$$

kaikilla $A \in \mathcal{G}$. Mutta tämä seuraa ehdottoman odotusarvon lineaarisuudesta ja ehdollisen odotusarvon operationaalisesta määritelmästä. Nimittäin kaikille $A \in \mathcal{G}$ pätee

$$\begin{aligned}
\mathbf{E}\Big[\big(aX + bY\big)\mathbf{1}_A\Big] &= \mathbf{E}[aX\mathbf{1}_A + bY\mathbf{1}_A] \\
&= a\mathbf{E}[X\mathbf{1}_A] + b\mathbf{E}[Y\mathbf{1}_A] \\
&= a\mathbf{E}[\,\mathbf{E}[X|\mathcal{G}]\mathbf{1}_A\,] + b\mathbf{E}[\,\mathbf{E}[Y|\mathcal{G}]\mathbf{1}_A\,] \\
&= \mathbf{E}\Big[\,\big(a\mathbf{E}[X|\mathcal{G}] + b\mathbf{E}[Y|\mathcal{G}]\big)\mathbf{1}_A\,\Big].
\end{aligned}$$

(iii) Väitteen mitallisuuskysymykset ovat selviä, sillä yhtälön vasen puoli on sekä $\mathcal{H}$- että $\mathcal{G}$-mitallinen, ja koska $\mathcal{H} \cap \mathcal{G}$ on σ-algebra, niin $\mathcal{H}$- ja $\mathcal{G}$-mitallisuus tarkoittaa $\mathcal{H} \cap \mathcal{G}$-mitallisuuttaa. Pitää siis enää osoittaa, että

$$\mathbf{E}[\,\mathbf{E}[X|\mathcal{H}]\mathbf{1}_A] \quad = \quad \mathbf{E}[X\mathbf{1}_A] \tag{9.2.7}$$

kaikilla $A \in \mathcal{H} \cap \mathcal{G}$. Mutta jos $A \in \mathcal{H} \cap \mathcal{G}$, niin $A \in \mathcal{H}$ ja siten (9.2.7) pätee ehdollisen odotusarvon $\mathbf{E}[X|\mathcal{H}]$ määritelmän nojalla. □

Lauseen 9.2.6 projektio-ominaisuutta (iii) kuvataan usein sanonnalla "karkeampi aina voittaa" (eikä vain useasti niin kuin Leino sanoi). Nimittäin jos $\mathcal{H}$ on $\mathcal{G}$:tä karkeampi eli $\mathcal{H} \subset \mathcal{G}$, niin ominaisuudesta 9.2.6(iii) seuraa, että

$$\mathbf{E}\big[\,\mathbf{E}[X|\mathcal{H}]\,\big|\,\mathcal{G}\big] \;=\; \mathbf{E}\big[\,\mathbf{E}[X|\mathcal{G}]\,\big|\,\mathcal{H}\big] \;=\; \mathbf{E}[X|\mathcal{H}].$$

Erityisesti siis ehdoton odotusarvo on ehdollisen odotusarvon erikoistapaus. Ehdottomassa odotusarvossa on ehtona triviaali ali-σ-algebra:

$$\mathbf{E}[X] \;=\; \mathbf{E}[X\,|\,\{\varnothing,\Omega\}]$$

Ehdollinen odotusarvo toteuttaa olennaisesti kaikki ehdottoman odotusarvon ominaisuudet. Esitämme nyt, esimerkin vuoksi, ehdollisen monotonisen konvergenssin lauseen.

9.2.8 Lause. *Olkoon* (X_n) *kasvava jono positiivisia satunnaismuuttujia, joille* $X_n \uparrow X$ *ja* $X \in L^1$. *Olkoon* $\mathcal{G}$ *ali-*σ*-algebra. Tällöin* $\mathbf{E}[X_n|\mathcal{G}] \uparrow \mathbf{E}[X|\mathcal{G}]$.

Todistus. Olkoon $A \in \mathcal{G}$. Tällöin $X_n\mathbf{1}_A \uparrow X\mathbf{1}_A$. Siten tavallisen, ehdottoman, monotonisen konvergenssin lauseen 4.2.1(i) nojalla

$$\mathbf{E}[X_n\mathbf{1}_A] \;\uparrow\; \mathbf{E}[X\mathbf{1}_A]\,.$$

Mutta, suoraan ehdollisen odotusarvon määritelmän nojalla, tämä tarkoittaa, että

$$\mathbf{E}[\mathbf{E}[X_n|\mathcal{G}]\mathbf{1}_A] \;\uparrow\; \mathbf{E}[\mathbf{E}[X|\mathcal{G}]\mathbf{1}_A]\,. \tag{9.2.9}$$

Väite seuraa tästä. Nimittäin lineaarisuuden ja tavallisen monotonisen konvergenssin lauseen nojalla (9.2.9) sanoo, että

$$\mathbf{E}[(\lim \mathbf{E}[X_n|\mathcal{G}] - \mathbf{E}[X|\mathcal{G}])\,\mathbf{1}_A] \;=\; 0$$

kaikilla $A \in \mathcal{G}$. Koska $\lim \mathbf{E}[X_n|\mathcal{G}] - \mathbf{E}[X|\mathcal{G}]$ on $\mathcal{G}$-mitallinen, niin on oltava $\lim \mathbf{E}[X_n|\mathcal{G}] = \mathbf{E}[X|\mathcal{G}]$ melkein varmasti. □

Ehdollisessa odotusarvossa "tunnetun voi ottaa ulos" ehdollistamisesta ja riippumattomassa tapauksessa ehdollistaminen on turhaa:

9.2.10 Lause. *Olkoon* $X, Y, XY \in L^1$.

(i) *Jos* X *on* $\mathcal{G}$*-mitallinen, niin* $\mathbf{E}[XY|\mathcal{G}] = X\mathbf{E}[Y|\mathcal{G}]$.
(ii) *Jos* Y *on riippumaton* σ*-algebrasta* $\mathcal{G}$*, niin* $\mathbf{E}[Y|\mathcal{G}] = \mathbf{E}[Y]$.

Todistus. (i) Mitallisuuskysymykset ovat selviä: kaavan molemmat puolet ovat $\mathcal{G}$-mitallisia. Tämän jälkeen kaavan $\mathbf{E}[XY|\mathcal{G}] = X\mathbf{E}[Y|\mathcal{G}]$ voi todistaa helposti "rakenteellisella induktiolla" satunnaismuuttujan X suhteen: ensiksi todetaan että kaava pätee, kun $X = \mathbf{1}_A$ jollakin $A \in \mathcal{G}$, sitten käytetään lineaarisuuttaa ja lopuksi käytetään (ehdollista) dominoidun konvergenssin lausetta. Sivuutamme yksityiskohdat.

(ii) Mitallisuuskysymykset ovat jälleen kerran täysin selviä. Nimittäin deterministinen otus $\mathbf{E}[Y]$ on varmasti mitallinen minkä tahansa ali-σ-algebran suhteen. Pitää siis osoittaa vain, että

$$\mathbf{E}[\mathbf{E}[Y]\mathbf{1}_A] \;=\; \mathbf{E}[Y\mathbf{1}_A]$$

kaikilla $A \in \mathcal{G}$. Mutta koska Y on riippumaton ali-σ-algebrasta $\mathcal{G}$, niin satunnaismuuttujat Y ja $\mathbf{1}_A$ ovat riippumattomia kaikilla $A \in \mathcal{G}$. Siten siis

$$\mathbf{E}[Y\mathbf{1}_A] \;=\; \mathbf{E}[Y]\mathbf{E}[\mathbf{1}_A].$$

Väite seuraa tästä. □

9.3 Ehdollinen odotusarvo parhaana L^2-ennusteena

Tässä osiossa määrittelemme ehdollisen odotusarvon käyttämättä Radonin ja Nikodymin lausetta. Määrittelemme nimittäin sen Hilbertin avaruuden ortoprojektiona. Tämän määritelmän hyvä puoli on että se antaa välittömästi meille seuraavan tulkinnan: ehdollinen odotusarvo $\hat{X} = \mathbf{E}[X|\mathcal{G}]$ on se melkein varmasti yksikäsitteinen $\mathcal{G}$-mitallinen satunnaismuuttuja $\hat{X}$, jolle "ennustusvirhe" $\mathbf{Var}[X-Y]$ minimoituu kaikkien $\mathcal{G}$-mitallisten satunnaismuuttujien Y joukossa:

$$\mathbf{E}[X|\mathcal{G}] \;=\; \underset{Y \text{ on } \mathcal{G}\text{-mitallinen}}{\operatorname{argmin}} \mathbf{Var}[Y-X].$$

Huono puoli tässä lähestymistavassa on, että a priori sitä voidaan käyttää vain neliöintegroituville satunnaismuuttujille $X \in L^2$. Radonin ja Nikodymin lähestymistapa toimii kaikille integroituville satunnaismuuttujille $X \in L^1$.

Aloitamme määrittelemällä *Hilbertin avaruuden* ja *sisätulon.*

9.3.1 Määritelmä. Olkoon H vektoriavaruus.

(a) Kuvaus $\langle\cdot,\cdot\rangle : H^2 \to \mathbb{R}$ on *sisätulo*, jos

- (i) $\langle x,y\rangle = \langle y,x\rangle$ kaikilla $x,y \in H$,
- (ii) $\langle ax+by,z\rangle = a\langle x,z\rangle + b\langle y,z\rangle$ kaikilla $a,b \in \mathbb{R},\ x,y,z \in H$,
- (iii) $\langle x,x\rangle \geq 0$ kaikilla $x \in H$,

(iv) $\langle x, x\rangle = 0$, jos ja vain jos $x = 0$.

(b) Avaruus $H := (H, \langle\cdot,\cdot\rangle)$ on *Hilbertin avaruus*, jos se on täydellinen normiavaruus normin $x \mapsto \|x\| := \sqrt{\langle x, x\rangle}$ suhteen.

Hilbertin avaruuksille pätee seuraava projektio-ominaisuus.

9.3.2 Apulause. *Olkoon H Hilbertin avaruus ja $\hat{H}$ sen suljettu aliavaruus. Olkoon $x \in H$. Tällöin on olemassa yksikäsitteinen* ortoprojektio $\hat{x} \in \hat{H}$*, jolle etäisyys $\|x - \hat{x}\|$ minimoituu. Lisäksi ortoprojektio $\hat{x}$ määräytyy ehdosta*

$$\langle x - \hat{x}, y\rangle \;=\; 0$$

kaikilla $y \in \hat{H}$.

Todistus. Esitetään funktionaalianalyysin peruskurssilla. □

Seuraava apulause sanoo, että L^p-avaruudet ovat täydellisiä ja siten Banachin avaruuksia.

9.3.3 Apulause. *Avaruus L^p on täydellinen. Toisin sanoen sen Cauchy-jonot suppenevat. Toisin sanoen, jos kaikille $\varepsilon > 0$ on olemassa sellainen n_ε, että*

$$\|X_n - X_m\|_p \;<\; \varepsilon,$$

kun $n, m \geq n_\varepsilon$, niin $X_n \xrightarrow{L^p} X$.

Todistus. Jätämme tämän väitteen todistuksen harjoitustehtäväksi 9.5. Vihjeenä todettakoon, että klassinen todistus perustuu lauseeseen 7.3.4. Siis siihen, että stokastisesti suppenevalla jonolla on melkein varmasti suppeneva osajono. □

9.3.4 Apulause. *Avaruus L^2 on Hilbertin avaruus, kun melkein varmasti samat satunnaismuuttujat samaistetaan ja sisätuloksi otetaan $\langle X, Y\rangle := \mathbf{E}[XY]$.*

Todistus. On täysin suoraviivaista nähdä, että $XY \mapsto \mathbf{E}[XY]$ on sisätulo ja että $\langle X, X\rangle = \|X\|_2^2$. Siten pitää vain osoittaa, että L^2 on täydellinen. Mutta apulause 9.3.3 sanoo, että se on. □

Seuraava aputulos seuraa yksinkertaisesti siitä, että mitallisten funktioiden rajat ovat mitallisia.

9.3.5 Apulause. *Olkoon $\mathcal{G} \subset \mathcal{F}$ ali-σ-algebra. Tällöin $L^2(\Omega, \mathcal{G}, \mathbf{P})$ on avaruuden $L^2(\Omega, \mathcal{F}, \mathbf{P})$ suljettu aliavaruus.*

Edellä esitetyn nojalla voimme määritellä ehdollisen odotusarvon $\mathbf{E}[X|\mathcal{G}]$ olemaan satunnaismuuttujan $X \in L^2(\Omega, \mathcal{F}, \mathbf{P})$ ortoprojektio aliavaruudelle $X \in L^2(\Omega, \mathcal{G}, \mathbf{P})$. Eli se on se melkein varmasti yksikäsitteinen $\mathcal{G}$-mitallinen satunnaismuuttuja $\hat{X}$, jolle etäisyys $\mathbf{E}[(X-\hat{X})^2]$ minimoituu. Seuraava lause sanoo, että tämä määritelmä on konsistentti Kolmogorovin operationaalisen määritelmän 9.2.5 kanssa.

9.3.6 Lause. *Olkoon $X \in L^2(\Omega, \mathcal{F}, \mathbf{P})$ ja $\mathcal{G} \subset \mathcal{F}$ ali-σ-algebra. Olkoon $\hat{X}$ satunnaismuuttuja X ortoprojektio aliavaruudelle $L^2(\Omega, \mathcal{G}, \mathbf{P})$. Tällöin $\hat{X} = \mathbf{E}[X|\mathcal{G}]$.*

Todistus. $\mathcal{G}$-mitallisuus on jälleen kerran selvää. Pitää siis vain osoittaa, että

$$\mathbf{E}\left[\hat{X}\mathbf{1}_A\right] = \mathbf{E}[X\mathbf{1}_A]$$

kaikilla $A \in \mathcal{G}$. Koska $\hat{X}$ on ortoprojektio, niin

$$\mathbf{E}\left[(X - \hat{X})Y\right] = 0$$

kaikilla $\mathcal{G}$-mitallisilla neliöintegroituvilla satunnaismuuttujilla Y. Väite seuraa nyt valitsemalla $Y := \mathbf{1}_A$. □

9.4 Säännöllinen ehdollinen jakauma

Tässä osiossa tarkastelemme satunnaismuuttujan ehdollista odotusarvoa toisen satunnaismuuttujan suhteen. Tällöin voimme valita sellaisen ehdollisen todennäköisyyden, että se on mitta. Tätä mittaa kutsumme säännölliseksi ehdolliseksi jakaumaksi.

Aloitamme kuitenkin säännöllisen ehdollisen jakauman yleisestä, σ-algebroihin liittyvästä, määritelmästä.

9.4.1 Määritelmä. Olkoon $\mathcal{G}$ ali-σ-algebra. Kuvaus $\mathbf{P}[\cdot|\mathcal{G}]$ on *säännöllinen ehdollinen jakauma*, jos

(i) jokaiselle $A \in \mathcal{F}$, $\mathbf{P}[A|\mathcal{G}] = \mathbf{E}[\mathbf{1}_A|\mathcal{G}]$ melkein varmasti,
(ii) jokaisella $\omega \in \Omega$, $\mathbf{P}[\cdot|\mathcal{G}](\omega)$ on todennäköisyysmitta mitallisella avaruudella $(\Omega, \mathcal{F})$.

9.4.2 Lause. *Olkoon $\mathbf{P}[\cdot|\mathcal{G}]$ säännöllinen ehdollinen jakauma ja $X \in L^1$. Tällöin*

$$\mathbf{E}[X|\mathcal{G}] = \int_\Omega X(\omega)\,\mathbf{P}[\mathrm{d}\omega|\mathcal{G}].$$

Todistus. Tämä väite on helppo osoittaa standardimasiinalla eli rakenteellisella induktiolla satunnaismuuttujan X suhteen: jos $X = \mathbf{1}_A$ jollekin $A \in \mathcal{F}$, niin väite on selvä. Siten väite seuraa yksinkertaisille satunnaismuuttujille lineaarisuuden nojalla ja lopulta väite seuraa integroituville satunnaismuuttujille (ehdollisen) dominoidun konvergenssin lauseen nojalla. □

Lause 9.4.2 palauttaa ehdollisen odotusarvon säännölliseen ehdolliseen jakaumaan eli ehdolliseen todennäköisyyteen. Kysymys onkin nyt, koska säännöllinen ehdollinen jakauma on olemassa. Tiedämme, että ehdollinen odotusarvo on olemassa kaikille $X \in L^1$. Siten toivottavaa olisi, että säännöllinen ehdollinen jakauma olisi myös olemassa jokaiselle $X \in L^1$. Mutta tällöin säännöllisen ehdollisen jakauman tulisi olla olemassa aina, sillä $\mathbf{1}_A \in L^1$ kaikilla $A \in \mathcal{F}$. Tämä ei kuitenkaan valitettavasti pidä paikkaansa. Ongelma on se, että (epäsäännöllinen) ehdollinen todennäköisyys $\mathbf{P}^0[A] := \mathbf{E}[\mathbf{1}_A|\mathcal{G}]$ on määritelty vain melkein varmasti. Toisin sanoen voi olla olemassa sellainen nollajoukko N, että

$$\mathbf{P}^0\left[\bigcup A_n \middle| \mathcal{G}\right](\omega) \neq \sum \mathbf{P}^0[A_n|\mathcal{G}](\omega), \tag{9.4.3}$$

kun $\omega \in N$. Epäyhtälössä (9.4.3) siis joukot $A_n \in \mathcal{F}$, $n \in \mathbb{N}$, ovat erillisiä ja nollajoukko N liittyy niihin. Säännöllisen ehdollisen jakauman $\mathbf{P}[\cdot|\mathcal{G}]$ ideana on muuttaa ehdollista todennäköisyyttä $\mathbf{P}^0[\cdot|\mathcal{G}]$ nollajoukossa N siten, että epäyhtälö (9.4.3) toteutuukin yhtälönä. Ongelma on, että tämä täytyy tehdä samanaikaisesti kaikille poikkeusjoukoille N ja A_n, $n \in \mathbb{N}$, joille ongelma (9.4.3) esiintyy. On kuitenkin periaatteessa mahdollista, että σ-algebra $\mathcal{F}$ on niin iso, että se sisältää niin paljon nollajoukkoja, että ne kasaantuvat, eikä löydy sellaista yhtä nollajoukkoa, että $\mathbf{P}^0[\cdot|\mathcal{G}]$ olisi mitta sen ulkopuolella. Itse asiassa voidaan osoittaa, muttei kovin helposti, että yleisesti σ-algebrat $\mathcal{F}$ voivat todellakin olla niin isoja. Mikäli $\mathcal{F}$ on kuitenkin jonkin satunnaismuuttujan X virittämä, niin se on riittävän pieni. Toisin sanoen $\mathcal{B}$ on riittävän pieni (itse asiassa $\mathcal{B}_d$ on myös riittävän pieni, muttemme nyt viitsi tarkastella vektoritapausta). Seuraava lause sanoo tämän.

9.4.4 Lause. *Olkoon X satunnaismuuttuja ja $\mathcal{G} \subset \mathcal{F}$ ali-σ-algebra. Tällöin säännöllinen ehdollinen jakauma $\mathbf{P}_X[\cdot|\mathcal{G}]$ on aina olemassa.*

Todistus. Olkoon $\mathbf{P}^0_X[\cdot|\mathcal{G}]$ jokin ehdollinen todennäköisyys. Sana "jokin" viittaa siihen, että ehdollinen todennäköisyys on yleisesti vain melkein varmasti määritelty. Nyt olemme siis valinneet jonkin version eli ekvivalenssiluokan edustajan. Dynkinin laajennuslauseen 2.3.8 nojalla riittää osoittaa, että voimme valita sellaisen version $F(\cdot|\mathcal{G})$ ehdollisesta kertymäfunktiosta

$$F^0(x|\mathcal{G}) := \mathbf{P}^0_X\big[(-\infty, x]\big|\mathcal{G}\big], \qquad x \in \mathbb{R},$$

että se on kertymäfunktio jokaisella $\omega \in \Omega$.

Rakennamme säännöllisen version $F(\cdot|\mathcal{G})$. Konstruktion toimivuus perustuu siihen, että avaruus $\mathbb{R}$ on separoituva.

Olkoon $\{q_1, q_2, \ldots\} := \mathbb{Q}$ ja

$$\begin{aligned} M_{k,l} &:= \Big\{\omega \in \Omega\,;\, F^0(q_k|\mathcal{G}) < F^0(q_l|\mathcal{G})\Big\}, \\ M &:= \bigcup_{q_k > q_l} M_{k,l}. \end{aligned}$$

Nyt siis M on joukko, missä kertymäfunktioiden monotonisuusominaisuus ei päde. Mutta M on numeroituva yhdiste joukkoja $M_{k,l}$, joille $\mathbf{P}[M_{k,l}] = 0$ kaikilla $k, l \in \mathbb{N}$. Siten $\mathbf{P}[M] = 0$. Olkoon sitten

$$\begin{aligned} N_k &:= \Big\{\omega \in \Omega\,;\, \lim_{q_l \downarrow q_k} F^0(q_l|\mathcal{G}) \neq F^0(q_k|\mathcal{G})\Big\}, \\ N &:= \bigcup_k N_k. \end{aligned}$$

Joukko N on siis se joukko, jossa kertymäfunktioiden jatkuvuusominaisuus ei päde. Jälleen huomaamme, että $\mathbf{P}[N_k] = 0$ kaikilla $k \in \mathbb{N}$ ja siten $\mathbf{P}[N] = 0$. Lopuksi olkoon $L \in \mathcal{F}$ se joukko, jossa

$$\lim_{q_k \to -\infty} F^0(q_k|\mathcal{G}) \neq 0 \qquad \text{tai} \qquad \lim_{q_k \to \infty} F^0(q_k|\mathcal{G}) \neq 1$$

eli kertymäfunktioiden rajaominaisuudet eivät päde. Edelleen $\mathbf{P}[L] = 0$.

Nyt olemme valmiit määrittelemään säännöllisen ehdollisen kertymäfunktion $F(\cdot|\mathcal{G})$. Olkoon G mikä tahansa kertymäfunktio. Asetamme

$$F(x|\mathcal{G}) \;:=\; \begin{cases} G(x), & \text{kun } \omega \in M \cup N \cup L, \\ \lim_{q_k \downarrow x} F^0(q_k|\mathcal{G}), & \text{muulloin.} \end{cases}$$

Nyt on helppo nähdä, että $F(\cdot|\mathcal{G})$ on kertymäfunktio jokaisella $\omega \in \Omega$ ja että $F(\cdot|\mathcal{G}) = F^0(\cdot|\mathcal{G})$ melkein varmasti. Siten olemme löytäneet (erään) säännöllisen ehdollisen jakauman. Nimittäin $\mathbf{P}_X[\cdot|\mathcal{G}]$ saadaan laajentamalla jokaisella $\omega \in \Omega$ yhteys

$$\mathbf{P}_X\big[(-\infty, x]\big|\mathcal{G}\big] \;=\; F(x|\mathcal{G}), \qquad x \in \mathbb{R},$$

mitaksi σ-algebralle $\mathcal{B}$. □

Tulkitsemalla säännöllistä ehdollista jakaumaa hieman saamme seuraavan tuloksen, mikä helpottaa usein ehdollisten odotusarvojen $\mathbf{E}[X|Y = y]$ laskemista.

9.4.5 Lause. *Olkoot X ja Y satunnaismuuttujia ja $g : \mathbb{R}^2 \to \mathbb{R}$ sellainen mitallinen kuvaus, että $\mathbf{E}[|g(X,Y)|] < \infty$. Tällöin*

$$\mathbf{E}[g(X,Y)] = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} g(x,y)\, \mathbf{P}[X \in \mathrm{d}y | Y = y] \right) \mathbf{P}[Y \in \mathrm{d}y]$$

Todistus. Harjoitustehtävä 9.6. □

Lausetta 9.4.5 kutsutaan joskus *mittojen hajotelmalauseeksi*. Se nimittäin sanoo formaalisti, että

$$\mathbf{P}[X \in \mathrm{d}x, Y \in \mathrm{d}y] = \mathbf{P}[X \in \mathrm{d}x | Y = y]\mathbf{P}[Y \in \mathrm{d}y].$$

Seuraava tulos kertoo, miten ehdollisia odotusarvoja voi laskea säännöllisen ehdollisen jakauman avulla. Se seuraa varsin helposti lauseesta 9.4.5. Jätämme varsinaisen todistuksen harjoitustehtäväksi 9.7.

9.4.6 Seuraus. *Olkoot X, Y ja g kuten lauseessa 9.4.5.*

(i) *Tällöin*

$$\mathbf{E}[g(X,Y)|Y = y] = \int_{\mathbb{R}} g(x,y)\, \mathbf{P}[X \in \mathrm{d}x | Y = y],$$

(ii) *ja jos X ja Y ovat riippumattomia, niin*

$$\mathbf{E}[g(X,Y)|Y = y] = \mathbf{E}[g(X,y)]\,.$$

Harjoitustehtäviä

9.1. Osoita, että ehdollinen todennäköisyys on todennäköisyys. Toisin sanoen osoita, että kiinnitetyllä $E \in \mathfrak{F}$, jolle $\mathbf{P}[E] > 0$ kuvaus

$$A \mapsto \mathbf{P}[A|E]$$

on todennäköisyysmitta avaruudella $(\Omega, \mathfrak{F})$.

9.2. Olkoon Y diskreetti satunnaismuuttuja ja $\mathbf{P}[\cdot|Y]$ osiossa 9.1 määritelty satunnaismitta. Osoita, että kaikille $A \in \mathfrak{F}$ pätee

$$\mathbf{P}[A] = \mathbf{E}\big[\mathbf{P}[A|Y]\big].$$

9.3. Osoita, että määritelmä 9.2.5 on kaavan (9.1.3) antaman määritelmän yleistys.

9.4. Olkoon $X, Y \in L^1$ ja $g : \mathbb{R} \to \mathbb{R}$ mitallinen bijektio. Osoita, että

$$\mathbf{E}[X|Y] \quad = \quad \mathbf{E}[X|g(Y)],$$

mutta yleisesti

$$\mathbf{E}[X|Y = y] \quad \neq \quad \mathbf{E}[X|g(Y) = y].$$

9.5. Todista apulause 9.3.3

9.6. Todista lause 9.4.5.

9.7. Todista seuraus 9.4.6.

Hakemisto