Luku 3

Satunnaismuuttujat, ehdollistaminen ja riippumattomuus

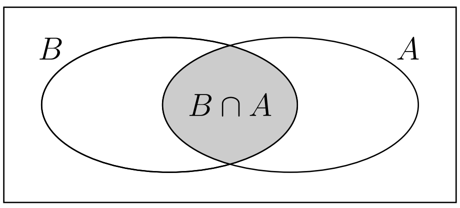
Tässä luvussa käsitellään satunnaismuuttujien ominaisuuksia ja täydennetään todennäköisyyslaskennan tietoja. Erityisesti satunnaismuuttujien odotusarvo on keskeinen käsite. Satunnaismuuttujien tarkastelussa rajoitutaan diskreettiin tapaukseen, mutta vastaavat tulokset pitävät paikkansa myös jatkuville satunnaismuuttujille. Tulosten todistaminen ja soveltaminen on huomattavasti helpompaa diskreettien satunnaismuuttujien yhteydessä.

3.1 Ehdollinen todennäköisyys

Määritelmä 3.1 (Ehdollinen todennäköisyys) Olkoot A ja B otosavaruuden Ω tapahtumia. Jos $P(A) > 0$, niin tapahtuman B ehdollinen todennäköisyys ehdolla A on

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)}. \tag{3.1.1}$$

Lauseke $P(B \mid A)$ luetaan ”B:n todennäköisyys ehdolla A”.



Voidaan ajatella, että $P(A)$ on alueen A pinta-ala ja $P(B \cap A)$ alueen $B \cap A$ pinta-ala. Ehdollinen todennäköisyys $P(B \mid A)$ on siis alueen $B \cap A$ pinta-alan suhteellinen osuus A:n pinta-alasta.

Esimerkki 3.1 Mikä on todennäköisyys, että saat pokerissa kuninkaallisen värisuoran K (samaa maata olevat kortit 10, 11, 12, 13 ja 14 = ässä)? Jos oletetaan, että kaikki 5 kortin kädet ovat yhtä todennäköisiä, niin

$$P(K) = \frac{4}{\binom{52}{5}} = \frac{1}{649740}.$$

Oletetaan, että jakaja jakaa 4 ensimmäistä korttia pöytään kuvapuoli alaspäin ja 5. kortin kuvapuoli ylöspäin. Viimenen korttisi on herttaässä (H_{14}). Millä todennäköisyydellä tämä käsi on kuninkaallinen värisuora? Ehdollisen todennäköisyyden (3.1.1) mukaan

$$P(K \mid H_{14}) = \frac{P(K \cap H_{14})}{P(H_{14})} = \frac{1/\binom{52}{5}}{\binom{51}{4}/\binom{52}{5}} = \frac{1}{\binom{51}{4}}.$$

Voimme nyt helposti todeta, että

$$P(K \mid H_{14}) = \frac{13}{5}\, P(K).$$

Kuninkaallisen värisuoran mahdollisuus siis yli kaksinkertaistuu, kun saat tietää, että viimeinen kortti on herttaässä. □

3.1.1 Todennäköisyyksien tulosääntö

Ehdollisen todennäköisyyden määritelmästä saadaan tulosääntö tapahtuman 'A ja B sattuvat' todennäköisyyden laskemiseksi. Jos tiedetään todennäköisyydet $P(A)$ ja $P(B \mid A)$, saadaan tulokaava

$$P(A \cap B) = P(A)\, P(B \mid A), \tag{3.1.2}$$

ja vastaavasti $P(A^c \cap B) = P(A^c)\, P(B \mid A^c)$. Lauseen 2.3 perusteella

$$P(B) = P(A \cap B) + P(A^c \cap B),$$

joten saamme *kokonaistodennäköisyyden* kaavan

$$P(B) = P(A)\, P(B \mid A) + P(A^c)\, P(B \mid A^c). \tag{3.1.3}$$

Ehdollisen todennäköisyyden määritelmän mukaan

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \qquad \text{kun} \quad P(B) > 0.$$

Kun tämän lausekkeen oikealle puolelle sijoitetaan $P(A \cap B)$:n paikalle (3.1.2) ja $P(B)$:n paikalle vastaavasti (3.1.3), saadaan *Bayesin kaava*

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

Jos siis tunnetaan todennäköisyydet $P(A)$, $P(B \mid A)$ ja $P(B \mid A^c)$, voidaan todennäköisyys $P(A \mid B)$ laskea Bayesin kaavan avulla.

Tulokaava (3.1.2) yleistyy myös useammalle kuin kahdelle tapahtumalle. Esimerkiksi

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

Tulokaavan, kokonaistodennäköisyyden ja Bayesin kaavan yleistykset käsitellään luvun loppupuolella.

Esimerkki 3.2 Suuri teollisuuskonserni valmistaa kännyköitä kolmessa eri maassa, jotka ovat nimeltään Fahru, Russo ja Swedla. Ostat kännykän, mut-

Taulukko 3.1. Kokonaistuotanto ja viallisten %-osuus eri maissa.

	Maa		
	Fahru	Russo	Swedla
Kokonaistuotanto	1000000	2000000	3000000
Viallisten %-osuus	20 %	10 %	5 %

ta et tiedä, missä se on valmistettu. Olkoon V tapahtuma, että tuote on viallinen. F on tapahtuma, että tuote on valmistettu Fahrussa. Vastaavasti R ja S viittaavat valmistusmaihin Russo ja Swedla. Lasketaan todennäköisyydet (a) $P(F \mid S^c)$, (b) $P(V \mid S^c)$, (c) $P(V)$, (d) $P(F \mid V)$. Oletetaan, että kaikki valmistetut 6000000 kännykkää ovat yhtä todennäköisiä.

Ratkaisu.

(a)
$$\begin{aligned} P(F \mid S^c) &= \frac{P(F \cap S^c)}{P(S^c)} \\ &= \frac{P(F)}{P(S^c)} \qquad (\text{koska } F \subseteq S^c) \\ &= \frac{1000000/6000000}{3000000/6000000} = \frac{1}{3}. \end{aligned}$$

(b)
$$\begin{aligned} P(V \mid S^c) &= \frac{V \cap S^c}{P(S^c)} \\ &= \frac{P[V \cap (F \cup R)]}{P(S^c)} \qquad (\text{koska } S^c = F \cup R) \\ &= \frac{P(V \cap F) + P(V \cap R)}{P(S^c)} \qquad (\text{koska } F \cap R = \emptyset) \\ &= \frac{P(V \mid F)\, P(F) + P(V \mid R)\, P(R)}{P(S^c)} \\ &= \frac{\frac{1}{5} \cdot \frac{1}{6} + \frac{1}{10} \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{2}{15}. \end{aligned}$$

Kohdat (c) ja (d) jätetään harjoitustehtäviksi. □

Esimerkki 3.3 (Väärä positiivinen) Oletetaan, että eräs verinäytteiden laboratoriotesti antaa kaksi ja vain kaksi tulosta: positiivisen ja negatiivisen. Tiedetään, että 95 % tautia A sairastavista saa testissä positiivisen tuloksen. Myös 2 % niistä, joilla ei ole tautia A, saa positiivisen tuloksen (väärän positiivisen!). Oletetaan, että 1 % populaatiosta sairastaa tautia A. Jos satunnaisesti valitun henkilön testitulos on positiivinen, mikä on todennäköisyys, että hän sairastaa tautia A?

Olkoon nyt $T = \{\text{sairastaa tautia}\}$ ja $+$ tarkoittaa positiivista testitulosta. Tiedämme, että

$$P(+ \mid T) = 0.95, \quad P(+ \mid T^c) = 0.02, \quad P(T) = 0.01 \quad \text{ja} \quad P(T^c) = 0.99.$$

Soveltamalla Bayesin kaavaa (3.7.4) saadaan

$$\begin{aligned} P(T \mid +) &= \frac{P(T)\,P(+ \mid T)}{P(T)\,P(+ \mid T) + P(T^c)\,P(+ \mid T^c)} \\ &= \frac{0.01 \cdot 0.95}{0.01 \cdot 0.95 + 0.99 \cdot 0.02} = \frac{95}{293} \approx 0.32. \end{aligned}$$

Todennäköisyys vaikuttaa ensi näkemältä kovin pieneltä. Alhainen todennäköisyys selittyy sillä, että positiiviset tulevat joukosta, joka on pieni verrattuna siihen joukkoon, josta väärät positiiviset tulevat. □

3.1.2 Riippumattomuus

Milloin käy niin, että ehdollinen todennäköisyys $P(B \mid A)$ on sama kuin ehdollistamaton todennäköisyys $P(B)$? Silloin on voimassa identiteetti

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

Tämä kysymys johtaa riippumattomuuden määritelmään.

Määritelmä 3.2 Tapahtumat A ja B ovat *riippumattomat*, jos

$$P(A \cap B) = P(A)\,P(B) \tag{3.1.4}$$

Jos tapahtumat A ja B ovat riippumattomat, niin silloin identiteetit

$$P(A \mid B) = P(A) \qquad \text{ja} \qquad P(B \mid A) = P(B)$$

pitävät paikkansa. Tapahtumien A ja B riippumattomuudesta seuraa, että myös niiden komplementit ovat riippumattomat.

Lause 3.1 *Jos tapahtumat A ja B ovat riippumattomat, niin myös*

1. A ja B^c,

2. A^c ja B,

3. A^c ja B^c

ovat riippumattomat.

Todistus. Todistetaan 1. kohta. On siis näytettävä, että A:n ja B:n riippumattomuudesta seuraa identiteetti $P(A \cap B^c) = P(A)\,P(B^c)$. Seurauslauseen 2.1 mukaan

$$\begin{aligned} P(A \cap B^c) &= P(A) - P(A \cap B) \\ &= P(A) - P(A)\,P(B) && [A \text{ ja } B \text{ riippumattomat}] \\ &= P(A)[1 - P(B)] \\ &= P(A)\,P(B^c) && [\text{Lause } 2.1(5)], \end{aligned}$$

joten A ja B^c ovat riippumattomat. Muut kohdat todistetaan vastaavalla tavalla. □

Esimerkki 3.4 Gynekologisen irtosolunäytteen eli Papa-kokeen avulla voidaan todeta kohdun kaulaosan syöpää edeltävät kudosmuutokset. Oletetaan, että 30–65-vuotiaista naisista $100p$ %:lla on epänormaaleja (muuntuneita) soluja (kohdunsuussa ja kohdunkaulassa). Papa-kokeen suorittamiseen liittyvät seuraavat virheet:

1. Tapahtuma B: Kohdunkaulassa on epänormaaleja soluja, mutta ne *eivät osu* otokseen. Olkoon $P(B) = b$.

2. Tapahtuma C: Otoksessa on poikkeavia soluja, mutta niitä *ei havaita.* Olkoon $P(C) = c$.

3. Tapahtuma D: Pelkästään normaaleja soluja sisältävä otos *luokitellaan väärin* poikkeavaksi. Olkoon $P(D) = d$.

Oletetaan, että kaikki mainitut otanta- ja määritysvirheet ovat toisistaan riippumattomat. Jos satunnaisesti valitulle 30–65-vuotiaalle naiselle tehdään Papa-koe, niin

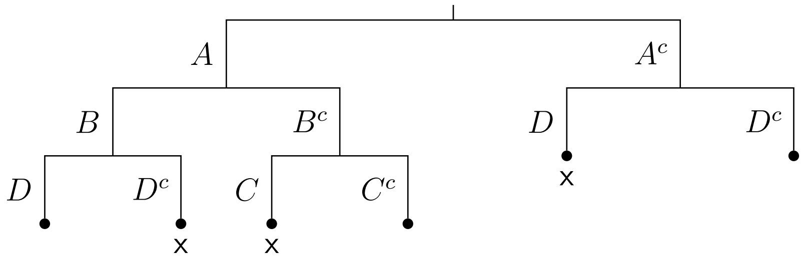
(a) millä todennäköisyydellä koe antaa väärän tuloksen?

(b) Jos testitulos osoittaa poikkeavia soluja löytyneen, millä todennäköisyydellä henkilöllä *ei ole* poikkeavia soluja?

Ratkaisu. (a) Tarkastellaan tapahtumia

V: Testi antaa virheellisen tuloksen,

A: Poikkeavia soluja on kohdunkaulassa



Kuvio 3.1. Kaaviokuva eri tulosvaihtoehdoista. Rastilla (x) merkityissä tilanteissa saadaan virheellinen testitulos.

ja tapahtumaa B (Poikkeavia soluja on, mutta ne eivät osu otokseen). Oletuksen mukaan $P(A) = p$, joten (Seurauslause 2.1)

$$\begin{aligned} P(V) &= P(A)\, P(V \mid A) + P(A^c)\, P(V \mid A^c) \\ &= p\, P(V \mid A) + (1-p)\, P(V \mid A^c). \end{aligned}$$

Virhetodennäköisyyden 3 mukaan $P(V \mid A^c) = d$. Toisaalta

$$P(V \mid A) = P(V \cap B \mid A) + P(V \cap B^c \mid A).$$

Virhetodennäköisyyksien 1 ja 3 mukaan

$$P(V \cap B \mid A) = (1-d)b$$

ja vastaavasti virheiden 1 ja 2 seurauksena

$$P(V \cap B^c \mid A) = c(1-b),$$

joten

$$P(V) = p[(1-d)b + c(1-b)] + (1-p)d.$$

(b) Jätetään harjoitustehtäväksi. □

Useamman kuin kahden tapahtuman riippumattomuuden määrittely vaatii hieman harkintaa. Milloin tapahtumat A, B ja C ovat riippumattomat? Ehdosta $P(A \cap B \cap C) = P(A)\, P(B)\, P(C)$ ei nimittäin seuraa, että tapahtumat ovat parittain riippumattomat.

Määritelmä 3.3 Tapahtumat A, B ja C ovat keskenään riippumattomat, jos

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

Esimerkki 3.5 Keskinäinen riippumattomuus ei seuraa parittaisesta riippumattomuudesta. Olkoon Ω otosavaruus, jonka alkeistapahtumia ovat tavallisen korttipakan kortit. Valitaan pakasta satunnaisesti yksi kortti. Olkoon $A = \{\spadesuit, \heartsuit\}$ tapahtuma, että saadaan pata tai hertta. Vastaavasti määritellään $B = \{\spadesuit, \clubsuit\}$ ja $C = \{\spadesuit, \diamondsuit\}$. Tapahtumien todennäköisyydet ovat $P(A) = P(B) = P(C) = \frac{26}{52} = \frac{1}{2}$. Mutta $A \cap B = A \cap C = B \cap C = \{\spadesuit\}$, joten

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = P(\{\spadesuit\}) = \frac{13}{52} = \frac{1}{4}.$$

Nyt A, B ja C ovat parittain riippumattomat, sillä $P(A \cap B) = P(A)\,P(B)$, $P(A \cap C) = P(A)\,P(C)$ ja $P(B \cap C) = P(B)\,P(C)$. Koska $A \cap B \cap C = \{\spadesuit\}$ ja

$$P(A \cap B \cap C) = P(\{\spadesuit\}) = \frac{1}{4} \neq P(A)\,P(B)\,P(C) = \left(\frac{1}{2}\right)^3 = \frac{1}{8},$$

niin A, B ja C eivät ole keskenään riippumattomat. $\square$

Esimerkki 3.6 Valitaan korttipakasta satunnaisesti yksi kortti. Määritellään tapahtumat $A = \{\text{ässä tai punainen kuningas tai punainen kuningatar}\}$, $M = \{\text{musta}\}$ ja $R = \{\text{risti}\}$. Silloin $P(A) = \frac{8}{52}$, $P(M) = \frac{1}{2}$ ja $P(R) = \frac{1}{4}$. Tapahtuma $A \cap M \cap R = \{\text{ristiässä}\}$ ja

$$P(A \cap M \cap R) = P(A)\,P(M)\,P(R) = \frac{8}{52} \cdot \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{52}.$$

Toisaalta

$$\begin{aligned} P(M \cap R) = P(R) = \frac{1}{4} &\neq P(M)\,P(R) = \frac{1}{8}, \\ P(A \cap M) = \frac{2}{52} &\neq P(A)\,P(M) = \frac{8}{52} \cdot \frac{1}{2} = \frac{4}{52}, \\ P(A \cap R) = \frac{1}{52} &\neq P(A)\,P(R) = \frac{8}{52} \cdot \frac{1}{4} = \frac{2}{52}, \end{aligned}$$

joten tapahtumat A, M ja R eivät ole parittain riippumattomia. Identiteetistä $P(A \cap M \cap R) = P(A)\,P(M)\,P(R)$ ei siis seuraa tapahtumien parittainen riippumattomuus. $\square$

Tapahtumien keskinäinen riippumattomuus vaatii toteutuakseen varsin voimakkaita ehtoja.

Määritelmä 3.4 Tapahtumat $A_1, \ldots, A_n$ ovat keskenään riippumattomat, jos jokainen tapahtumien osakokoelma $A_{i_1}, \ldots, A_{i_k}$ $(1 \leq k \leq n)$ toteuttaa ehdon

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P(A_{i_j}).$$

Ehdollinen riippumattomuus. Tapahtumat A ja B ovat riippumattomat ehdolla C, jos $P(A \cap B \mid C) = P(A \mid C)\, P(B \mid C)$.

3.1.3 Joukko-oppi ja todennäköisyys

Todennäköisyyslaskennan kannalta hyödylliset joukko-opin merkinnät esitettiin 1. luvussa. Tapahtumat A ja sen komplementti A^c eivät voi sattua samanaikaisesti, sillä $A \cap A^c = \emptyset$ ja $P(A \cap A^c) = P(\emptyset) = 0$. Toisaalta A, A^c on otosavaruuden Ω ositus, joten $A \cup A^c = \Omega$ ja $P(A \cup A^c) = P(\Omega) = 1$. Tapahtuma "$A$ tai A^c" sattuu varmasti. Lauseen 2.1 (kohta 4) perusteella tiedämme tuloksen $P(A \cup A^c) = P(A) + P(A^c)$, josta seuraa erittäin käyttökelpoinen sääntö (Lause 2.1, kohta 5)

$$P(A) = 1 - P(A^c).$$

De Morganin sääntö

$$(A \cap B)^c = A^c \cup B^c \tag{3.1.5}$$

on tärkeä apuväline todennäköisyyslaskennassa. Se pitää paikkansa myös mielivaltaisen monille tapahtumille. Tapahtuma-avaruuden kielellä luemme identiteetin (3.1.5) seuraavasti

Vasen puoli: Ei ole totta, että sekä A että B sattuvat.

Oikea puoli: Ainakin toinen tapahtumista A, B ei satu.

Soveltamalla kaksinkertaisen komplementin sääntöä $(A^c)^c = A$ saadaan De Morganin säännöstä (3.1.5) toinen vastaava sääntö

$$(A \cup B)^c = A^c \cap B^c.$$

3.2 Ehdolliset jakaumat

Olkoon X jossakin (numeroituvassa) otosavaruudessa Ω määritelty satunnaismuuttuja ja $P(\cdot)$ samassa otosavaruudessa määritelty todennäköisyys. Oletetaan, että tapahtuma $A \subset \Omega$, $P(A) > 0$, on sattunut. Määrittelemme nyt ehdollisen jakauman ehdollisen todennäköisyyden määritelmää mukaillen.

Jokaista X:n arvoa $x \in \mathbb{R}$ kohti voimme määritellä joukon

$$B_x = \{\, \omega \mid X(\omega) = x \,\}.$$

Ehdollisen todennäköisyyden määritelmän mukaan

$$P(X(\omega) = x \mid A) = P(B_x \mid A) = \frac{P(B_x \cap A)}{P(A)} \geq 0. \tag{3.2.1}$$

Koska $\bigcup_x B_x = \Omega$ ja $B_x \cap B_y = \emptyset$ kaikilla $x \neq y$, niin

$$\sum_x P(B_x \mid A) = \sum_x \frac{P(B_x \cap A)}{P(A)} = \frac{P(\Omega \cap A)}{P(A)} = 1. \tag{3.2.2}$$

Määritellään nyt funktio

$$f(x \mid A) = P(B_x \mid A) = P(X = x \mid A), \tag{3.2.3}$$

joka on (3.2.1):n ja (3.2.2):n perusteella todennäköisyysfunktio. Funktio (3.2.3) on X:n *ehdollinen todennäköisyysfunktio ehdolla* A.

Esimerkki 3.7 Oletetaan, että X noudattaa diskreettiä tasajakaumaa $\mathrm{Tasd}(1, N)$. Silloin X:n arvojoukko on $S_X = \{1, 2, \ldots, N\}$ ja $P(X = i) = 1/N$ kaikilla $i \in S_X$. Määritellään tapahtuma $A = \{\,\omega \mid a \leq X \leq b\,\}$, missä a, b ja N, $1 \leq a < b \leq N$, ovat kokonaislukuja. Silloin

$$P(A) = \sum_{i=a}^{b} \frac{1}{N} = \frac{b - a + 1}{N}$$

ja

$$P(\{X = k\} \cap A) = \begin{cases} 1/N; & a \leq k \leq b \\ 0; & \text{muutoin.} \end{cases}$$

Siksi X:n ehdollinen todennäköisyysfunktio ehdolla A on

$$f(x \mid A) = \begin{cases} \dfrac{1}{b - a + 1}; & a \leq x \leq b \\ 0; & \text{muutoin.} \end{cases}$$

□

3.3 Satunnaismuuttujien ominaisuuksia

3.3.1 Diskreetin satunnaismuuttujan odotusarvo

Numeroituvassa otosavaruudessa Ω määritellyn satunnaismuuttujan X *odotusarvo* on

$$E(X) = \sum_{\omega \in \Omega} X(\omega)\, P(\{\omega\}), \tag{3.3.1}$$

jos

$$\sum_{\omega \in \Omega} |X(\omega)\, P(\{\omega\})| < \infty. \tag{3.3.2}$$

Jos ehto (3.3.2) toteutuu, sarja (3.3.1) suppenee itseisesti. Tässä tapauksessa sanomme, että satunnaismuuttujalla X on odotusarvo. Muutoin satunnaismuuttujalla ei ole odotusarvoa. Jos $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ on äärellinen, niin

$$E(X) = \sum_{i=1}^{n} X(\omega_i)\, P(\{\omega_i\})$$

on aina olemassa.

Tarkastellaan nyt odotusarvon laskemista yleisemmin numeroituvassa otosavaruudessa. Olkoon A_1, A_2, … sellainen otosavaruuden jako

$$\Omega = \bigcup_i A_i,$$

että X saa saman arvon x_i koko joukossa A_i. Voimme kirjoittaa

$$X(\omega) = x_i, \quad \text{kun} \quad \omega \in A_i.$$

Merkitään nyt $P(A_i) = P(X = x_i) = p_i$, joten

$$E(X) = \sum_i P(A_i) x_i = \sum_i p_i x_i. \tag{3.3.3}$$

Tämä kaava saadaan ryhmittelemällä alkeistapaukset kaavassa (3.3.1) osajoukkoihin A_i ja summaamalla sitten yli indeksin i.

Kaavasta (3.3.1) saadaan myös minkä tahansa satunnaismuuttujan X funktion $h(X)$ odotusarvo. Koska $h(X)$ on satunnaismuuttuja, niin

$$E[h(X)] = \sum_{\omega \in \Omega} h[X(\omega)]\, P(\{\omega\}) = \sum_i p_i h(x_i).$$

Näin siis X:n jakauma määrittää $h(X)$:n odotusarvon. Jos erityisesti $h(X) = X^r$, saamme X:n *r. momentin*

$$E(X^r) = \sum_i p_i x_i^r. \tag{3.3.4}$$

Määrittelemme seuraavassa diskreetin satunnaismuuttujan *odotusarvon* todennäköisyysfunktion avulla. Jatkossa kutsumme satunnaismuuttujan odotusarvoa myös satunnaismuuttujan *keskiarvoksi.*

Määritelmä 3.5 (Odotusarvo) Olkoon X diskreetti satunnaismuuttuja, jonka arvojoukko on S ja todennäköisyysfunktio $f_X(x)$. Silloin X:n odotusarvo μ_X on

$$\mu_X = E(X) = \sum_{x \in S} x f_X(x) = \sum_{x \in S} x\, P(X = x), \tag{3.3.5}$$

jos summa suppenee itseisesti.

Odotusarvo μ_X on siis X:n arvojen todennäköisyyksillä painotettu keskiarvo. Jätämme usein merkinnästä satunnaismuuttujaan viittaavan alaindeksin X pois ja merkitsemme lyhyesti $f_X(x) = f(x)$ ja $\mu = E(X)$. Jos summan $\sum_{x \in S} x f_X(x)$ yhteenlaskettavien määrä on äärellinen, niin odotusarvo on aina olemassa. Mikäli yhteenlaskettavien määrä on ääretön, tulee summan supeta itseisesti.

Lause 3.2 *Oletetaan, että otosavaruudessa Ω määritellyllä diskreeteillä satunnaismuuttujilla X ja Y on odotusarvo ja $a \in \mathbb{R}$ on vakio. Silloin*

1. $E(aX) = a\,E(X)$ *ja* $E(X+Y) = E(X) + E(Y)$, *joten odotusarvo on lineaarinen operaattori.*

Olkoot $h(x)$, $h_1(x)$ ja $h_2(x)$ sellaisia funktioita, että satunnaismuuttujilla $h(X)$, $h_1(X)$ ja $h_2(X)$ on odotusarvo. Silloin seuraavat tulokset pitävät paikkansa:

2. $E[h(X)] = \sum_x h(x) f_X(x) = \sum_x h(x)\,P(X = x)$

3. *Jos* $h_1(x) \geq h_2(x)$ *kaikilla* x, *niin* $E[h_1(X)] \geq E[h_2(X)]$.

Todistus. 1. Todistetaan ensin $E(aX) = a\,E(X)$. Määritelmän mukaan

$$\begin{aligned} E(aX) &= \sum_x ax\,P(aX = ax) = a \sum_x x\,P(aX = ax) \\ &= a \sum_x x\,P(X = x) = a\,E(X). \end{aligned}$$

Identiteetti $P(aX = ax) = P(X = x)$ pitää paikkansa kaikilla $a \neq 0$, koska $\{\,\omega \mid aX(\omega) = ax\,\} = \{\,\omega \mid X(\omega) = x\,\}$. Jos $a = 0$, niin $aX = 0$ ja $E(aX) = 0 = 0 \cdot E(X)$. Odotusarvo $E(aX)$ on olemassa, koska $E(X)$ on olemassa (oletus). Huomaa, että X:n arvojoukko S_X on numeroituva ja merkintä $\sum_x$ tarkoittaa summaa yli arvojen S_X eli $\sum_x \equiv \sum_{x \in S_X}$.

Todistetaan $E(X+Y) = E(X) + E(Y)$:

$$\begin{aligned} E(X+Y) &= \sum_x \sum_y (x+y)\,P(X = x, Y = y) \\ &= \sum_x \sum_y [x\,P(X = x, Y = y) + y\,P(X = x, Y = y)] \\ &= \sum_x \sum_y x\,P(X = x, Y = y) + \sum_x \sum_y y\,P(X = x, Y = y) \\ &= \sum_x \sum_y x\,P(X = x)\,P(Y = y \mid X = x) \\ &\qquad + \sum_x \sum_y y\,P(Y = y)\,P(X = x \mid Y = y) \end{aligned}$$

$$
\begin{aligned}
&= \sum_x x\,P(X=x)\Big[\sum_y P(Y=y \mid X=x)\Big] \\
&\qquad + \sum_y y\,P(Y=y)\Big[\sum_x P(X=x \mid Y=y)\Big] \\
&= \sum_x x\,P(X=x) + \sum_y y\,P(Y=y) = E(X) + E(Y).
\end{aligned}
$$

Viimeistä edellinen yhtäsuuruus seuraa siitä, että $P(Y=y \mid X=x)$ on Y:n ehdollinen todennäköisyysfunktio ehdolla $X=x$ ja $P(X=x \mid Y=y)$ on X:n ehdollinen todennäköisyysfunktio ehdolla $Y=y$. Odotusarvon $E(X+Y)$ olemassaolo seuraa siitä, että $E(X)$ ja $E(Y)$ ovat olemassa ja $|x+y| \leq |x|+|y|$.

2. Seuraa suoraan odotusarvon määritelmästä.

3. Jos $h_1(x) \geq h_2(x)$ kaikilla $x \in \mathbb{R}$, niin

$$E[h_1(X)] - E[h_2(X)] = E[h_1(X) - h_2(X)]$$

1. kohdan mukaan. Nyt

$$E[h_1(X) - h_2(X)] = \sum_x [h_1(x) - h_2(x)]\,P(X=x) \geq 0,$$

koska $h_1(x) - h_2(x) \geq 0$ ja $P(X=x) \geq 0$ kaikilla $x \in \mathbb{R}$. Näin väite on todistettu. □

Olkoon I_A tapahtuman A indikaattorifunktio. Silloin

$$E(I_A) = P(A) \cdot 1 + [1 - P(A)] \cdot 0 = P(A).$$

Huomaa, että $1 - I_A = I_{A^c}$ on A:n komplementin indikaattorifunktio ja $I_\Omega = I_A + I_{A^c} = 1$ kaikilla $\omega \in \Omega$. Määritellään vastaavasti tapahtuman 'kruunu k. heitossa' indikaattorifunktio X_k:

$$X_k(\omega) = \begin{cases} 1, & \text{kun } \omega = \text{kruunu}; \\ 0, & \text{kun } \omega = \text{klaava}. \end{cases}$$

Oletetaan, että kruunun sattumisen todennäköisyys $P(X_k = 1) = p$, $k = 1, 2, \ldots, n$. Nyt satunnaismuuttuja

$$X = X_1 + X_2 + \cdots + X_n$$

on kruunujen lukumäärä, kun heitetään lanttia n kertaa. Silloin odotusarvon lineaarisuuden nojalla

$$E(X) = E(X_1) + E(X_2) + \cdots + E(X_n) = p + p + \cdots + p = np.$$

Kruunujen lukumäärän odotusarvo n:ssä heitossa on heittojen lukumäärä kertaa kruunun todennäköisyys. Jos lantti on harhaton, niin $E(X) = \frac{n}{2}$.

Esimerkki 3.8 Olkoon satunnaismuuttujan X arvoalue $S_X = \{-1, 0, 1\}$ ja arvojen todennäköisyydet

$$P(X = -1) = 0.2, \qquad P(X = 0) = 0.5 \qquad \text{ja} \qquad P(X = 1) = 0.3.$$

Lasketaan odotusarvo $E(X^2)$. Merkitään $Y = X^2$. Satunnaismuuttuja Y on siis X:n funktio. Y:n arvoalue on $S_Y = \{0, 1\}$, koska

$$Y(\omega) = \begin{cases} 1, & \text{kun } X(\omega) = 1 \text{ tai } X(\omega) = -1; \\ 0, & \text{kun } X(\omega) = 0. \end{cases}$$

Y:n arvojen 1 ja 0 todennäköisyydet ovat

$$\begin{aligned} P(Y = 1) &= P(X = -1) + P(X = 1) = 0.5, \\ P(Y = 0) &= P(X = 0) = 0.5. \end{aligned}$$

Siksi

$$E(X^2) = E(Y) = 1 \cdot 0.5 + 0 \cdot 0.5 = 0.5.$$

Olemme siis ensin määrittäneet X^2:n jakauman ja laskeneet siitä odotusarvon $E(X^2)$.

Voimme kuitenkin laskea $E(X^2)$:n määrittämättä ensin X^2:n jakaumaa. Soveltamalla Lausetta 3.9 (kohta 2) saadaan

$$\begin{aligned} E(X^2) &= (-1)^2 \cdot 0.2 + 0^2 \cdot 0.5 + 1^2 \cdot 0.3 \\ &= 1 \cdot (0.2 + 0.3) + 0 \cdot 0.5 = 0.5. \end{aligned}$$

Määritellään nyt satunnaismuuttuja

$$h(X) = [X - E(X)]^2 = (X - 0.5)^2 = X^2 - X + 0.25.$$

Satunnaismuuttuja $h(X)$ saa arvot $h(-1) = 2.25$, $h(0) = 0.25$ ja $h(1) = 0.25$. Odotusarvo on

$$\begin{aligned} E\big([X - E(X)]^2\big) &= 0.2 \cdot 2.25 + 0.5 \cdot 0.25 + 0.3 \cdot 0.25 \\ &= 0.2 \cdot 2.25 + 0.8 \cdot 0.25 = 0.65. \end{aligned}$$

Odotusarvo $E\big([X - E(X)]^2\big)$ on satunnaismuuttujan X varianssi. □

Esimerkki 3.9 Indikaattorifunktio (Määritelmä 2.3) on käyttökelpoinen myös todennäköisyyksien tarkastelussa. Jos A ja B ovat tapahtumia, niin silloin

$$I_{A^c} = 1 - I_A \qquad \text{ja} \qquad I_{A\cap B} = I_A I_B.$$

Koska $E(I_A) = P(A)$ ja $E(I_{A^c}) = P(A^c)$, niin odotusarvon lineaarisuuden nojalla (Lause 3.9, 1. kohta)

$$E(I_{A^c}) = 1 - E(I_A),$$

josta saamme tutun tuloksen $P(A^c) = 1 - P(A)$. De Morganin sääntöjen avulla saadaan myös identiteetti

$$I_{A\cup B} = I_A + I_B - I_A I_B.$$

□

Esimerkki 3.10 Satunnaismuuttuja X noudattaa diskreettiä tasajakaumaa Tasd$(1, N)$, kun $P(X = i) = \frac{1}{N}$, $i = 1, 2, \ldots, N$ (ks. alaluku 2.5.4). Silloin

$$\begin{aligned} E(X) &= \sum_{x=1}^{N} x \frac{1}{N} = \frac{1}{N} \sum_{x=1}^{N} x \\ &= \frac{1}{N} \cdot \frac{N(N+1)}{2} = \frac{N+1}{2}. \end{aligned}$$

Vastaavasti

$$\begin{aligned} E(X^2) &= \sum_{x=1}^{N} x^2 \frac{1}{N} = \frac{1}{N} \sum_{x=1}^{N} x^2 \\ &= \frac{1}{N} \cdot \frac{N(N+1)(2N+1)}{6} = \frac{(N+1)(2N+1)}{6}. \end{aligned}$$

□

Esimerkki 3.11 Hypergeometrinen jakauma esiteltiin tarkasteltaessa otantaa palauttamatta (alaluku 2.6.1). Esimerkiksi tarkistusotannassa tuotteet luokitellaan viallisiksi tai hyväksyttäviksi. Olkoon tuote-erässä N tuotetta, joista viallisia a ja hyväksyttäviä $N - a$ kappaletta. Tehdään n:n alkion satunnaisotos palauttamatta. Viallisten lukumäärä X otoksessa noudattaa hypergeometrista jakaumaa parametrein n, N ja p, missä $p = \frac{a}{N}$ on viallisten suhteellinen osuus tuote-erässä. Merkitään $X \sim \text{HGeo}(n, N, p)$. Hypergeometrisen jakauman todennäköisyysfunktio on

$$P(X = x; N, n, p) = \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}}, \qquad x = 0, 1, \ldots, n, \tag{3.3.6}$$

missä $a = pN$. Huomaa, että $x \leq \min(a, n)$ ja $x \geq \max(0, a + n - N)$, joten X:n todellinen arvoalue saattaa olla suppeampi kuin (3.3.6):ssä annettu.

Tarkistamme ensin, että kyseessä on todennäköisyysjakauma. Selvästikin $P(X = x) \geq 0$, kun $x = 0, 1, \ldots, n$. Mutta identiteetin

$$\sum_{x=0}^{n} P(X = x) = \frac{1}{\binom{N}{n}} \sum_{x=0}^{n} \binom{a}{x}\binom{N-a}{n-x} = 1$$

oikeellisuuden tarkistaminen ei ole täysin vaivaton tehtävä. Voimme kuitenkin tässä nojautua hypergeometriseen identiteettiin (2.4.10), jonka mukaan

$$\sum_{x=0}^{n} \binom{a}{x}\binom{N-a}{n-x} = \binom{N}{n}.$$

Lasketaan nyt hypergeometrisen jakauman odotusarvo

$$E(X) = \sum_{x=0}^{n} x \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}} = \sum_{x=1}^{n} x \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}}.$$

Identiteetin (2.4.5) nojalla saadaan

$$x\binom{a}{x} = a\binom{a-1}{x-1}$$

ja

$$\binom{N}{n} = \frac{N}{n}\binom{N-1}{n-1},$$

joten

$$E(X) = \sum_{x=1}^{n} \frac{a\binom{a-1}{x-1}\binom{N-a}{n-x}}{\frac{N}{n}\binom{N-1}{n-1}} = \frac{na}{N}\sum_{x=1}^{n} \frac{\binom{a-1}{x-1}\binom{N-a}{n-x}}{\binom{N-1}{n-1}}.$$

Kun merkitään $y = n - 1$, voidaan kirjoittaa

$$\begin{aligned}\sum_{x=1}^{n} \frac{\binom{a-1}{x-1}\binom{N-a}{n-x}}{\binom{N-1}{n-1}} &= \sum_{y=0}^{n-1} \frac{\binom{a-1}{y}\binom{N-a}{n-1-y}}{\binom{N-1}{n-1}} \\ &= \sum_{y=0}^{n-1} P(Y = y; N-1, n-1, p_1) = 1,\end{aligned}$$

missä $p_1 = \frac{a-1}{N-1}$. Satunnaismuuttuja Y noudattaa siis jakaumaa HGeo$(n - 1, N - 1, p_1)$. Siksi hypergeometrisen jakauman HGeo(n, N, p) odotusarvo on

$$E(X) = n\frac{a}{N} = np.$$

Summa laskettiin muuntamalla alkuperäinen jakauma hypergeometriseksi jakaumaksi, jonka parametrit ovat $n - 1$, $N - 1$ ja $p_1 = \frac{a-1}{N-1}$. Vastaavilla laskelmilla voidaan osoittaa, että

$$\mathrm{Var}(X) = \frac{na}{N} \cdot \frac{(N-a)(N-n)}{N(N-1)} = np(1-p)\frac{N-n}{N-1}.$$

□

3.3.2 Ehdollisen jakauman odotusarvo

Koska $f(x \mid A)$ on todennäköisyysfunktio (ks. identiteetti (3.2.3)), niin sen avulla voidaan määritellä odotusarvo. Jos $\sum_x |x| f(x \mid A) < \infty$, niin X:n *ehdollinen odotusarvo ehdolla* A on

(3.3.7) $$E(X \mid A) = \sum_x x f(x \mid A).$$

Esimerkki 3.12 Oletetaan, että $X \sim \mathrm{Tasd}(1, N)$ ja $A = \{\, \omega \mid a \le X(\omega) \le b \,\}$, $1 \le a < b \le N$, kuten Esimerkissä 3.7. Nyt X:n ehdollinen odotusarvo ehdolla A on

$$E(X \mid A) = \sum_x x f(x \mid A) = \sum_{x=a}^{b} x\frac{1}{b-a+1} = \frac{a+b}{2}.$$

□

Ehdollisen odotusarvon ja odotusarvon välillä on olemassa seuraavassa lauseessa esitetty erittäin tärkeä yhteys.

Lause 3.3 *Olkoon satunnaismuuttujan X odotusarvo $E(X)$ ja olkoon A sellainen tapahtuma, että $P(A)\,P(A^c) > 0$. Silloin*

$$E(X) = P(A)\,E(X \mid A) + P(A^c)\,E(X \mid A^c).$$

Todistus. Seurauslauseen 2.1 mukaan

$$P(X = x) = P(\{X = x\} \cap A) + P(\{X = x\} \cap A^c)$$

ja ehdollisen todennäköisyyden määritelmän nojalla

$$P(\{X = x\} \cap A) = P(A)\,P(X = x \mid A)$$

ja

$$P(\{X = x\} \cap A^c) = P(A^c)\,P(X = x \mid A^c).$$

Tästä seuraa, että

$$f(x) = P(X = x) = P(A)f(x \mid A) + P(A^c)f(x \mid A^c).$$

Siksi

$$\begin{aligned} E(X) &= \sum_x xf(x) = P(A)\sum_x xf(x \mid A) + P(A^c)\sum_x xf(x \mid A^c) \\ &= P(A)\,E(x \mid A) + P(A^c)\,E(x \mid A^c), \end{aligned}$$

niinkuin väitettiin. □

Jos joukkokokoelma $\{A_i; i \geq 1\}$ muodostaa otosavaruuden Ω osituksen (ks. alaluku 1.3.2), niin voidaan todistaa seuraava yleinen tulos:

$$E(X) = \sum_i P(A_i)\,E(X \mid A_i).$$

Alaluvussa 1.3.2 tarkasteltiin vain äärellisiä osituksia. On syytä huomata, että joukkokokoelma $\{A_i; i \geq 1\}$ voi olla numeroituvasti ääretön. Koska $\{A_i; i \geq 1\}$ on Ω:n ositus, niin

(i) $\bigcup_{i=1}^{\infty} A_i = \Omega$,

(ii) $A_i \cap A_j = \emptyset$, kun $i \neq j$, ja

(iii) $P(A_i) > 0$, $i \geq 1$.

3.3.3 Satunnaismuuttujan varianssi

Varianssin laskemiseksi tarvitaan funktion $h(X) = X^2$ odotusarvo (Vertaa Lauseen 3.9 kohta 2). Odotusarvoa $E(X^2)$ sanotaan satunnaismuuttujan X 2. *momentiksi.* Vastaavasti odotusarvo $E(X)$ on X:n 1. *momentti.* Ennen varianssin määrittelyä esitetään muutamia jatkossa tärkeitä aputuloksia.

Apulause 3.1 *Oletetaan, että satunnaismuuttujilla X ja Y on 2. momentti ja $c \in \mathbb{R}$ on vakio. Silloin odotusarvot*

$$E[(cX)^2], \quad E[(X+Y)^2], \quad E(X), \quad E(Y) \quad ja \quad E(XY) \tag{3.3.8}$$

ovat olemassa.

Todistus.

1. Koska $E[(cX)^2] = c^2\, E(X^2)$ ja $E(X^2)$ on oletuksen mukaan olemassa, niin $E[(cX)^2]$ on olemassa.

2. Koska $0 \leq (X+Y)^2 = 2(X^2+Y^2)-(X-Y)^2 \leq 2(X^2+Y^2)$ ja oletuksen mukaan $E(X^2+Y^2) = E(X^2)+E(X^2)$ on olemassa, niin Lauseen 3.9 (kohta 3) mukaan $E[(X+Y)^2]$ on olemassa.

3. Koska $0 \leq (|X|-|Y|)^2 = |X|^2 + |Y|^2 - 2|X||Y|$, niin niin Lauseen 3.9 (kohta 3) mukaan

   $$E(|XY|) \leq \frac{1}{2}\, E(X^2+Y^2),$$

   joten $E(XY)$ on olemassa.

$\square$

Lause 3.4 (Cauchyn ja Schwarzin epäyhtälö) *Jos satunnaismuuttujilla X ja Y on 2. momentti, niin*

$$[E(XY)]^2 \leq E(X^2)\, E(Y^2). \tag{3.3.9}$$

Yhtäsuuruus on voimassa jos ja vain jos $P(aX + bY = 0) = 1$, joillain $a, b \in \mathbb{R}$, joista ainakin toinen poikkeaa nollasta.

Todistus. (1) Oletetaan, että $E(X^2) \neq 0$. Koska oletuksen mukaan $E(X^2)$ ja $E(Y^2)$ ovat olemassa, niin Apulauseen 3.1 mukaan myös $E(XY)$ on olemassa. Merkitään nyt $c = E(XY)/E(X^2)$. Silloin

$$0 \leq E[(Y-cX)^2] = E(Y^2) - \frac{[E(XY)]^2}{E(X^2)},$$

mistä väite seuraa. Yhtäsuuruus on voimassa silloin ja vain silloin kun

$$P(Y - cX = 0) = 1.$$

(2) Jos $E(X^2) = 0$, niin $P(X = 0) = 1$. Silloin $P(XY = 0) = 0$ ja $E(XY) = 0$, joten epäyhtälö (3.3.9) pitää triviaalisti paikkansa. $\square$

Yhtäsuuruus (3.3.9):ssä vallitsee silloin, kun $aX = -bY$ (todennäköisyydellä 1). Silloin $Y = -\frac{a}{b}X$, jos $b \neq 0$. Epäyhtälössä (3.3.9) pätee siis yhtäsuuruus, kun X ja Y ovat lineaarisesti riippuvia. Epäyhtälö (3.3.9) voidaan lausua myös muodossa

$$|E(XY)| \leq E(|XY|) \leq \sqrt{E(X^2)}\sqrt{E(Y^2)}.$$

Määritelmä 3.6 (Varianssi) Jos satunnaismuuttujalla X on 2. momentti $E(X^2)$, niin sillä on odotusarvo μ_X ja X:n varianssi on

$$\sigma_X^2 = \mathrm{Var}(X) = E[(X - \mu_X)^2]. \tag{3.3.10}$$

Merkintöjen μ_X ja σ_X^2 sijasta käytämme tavallisesti lyhyempiä versioita μ ja σ^2, jos sekaannuksen vaaraa ei ole. Odotusarvon lineaarisuutta soveltaen voidaan todeta, että

$$\begin{aligned} E[(X-\mu)^2] &= E(X^2 - 2\mu X + \mu^2) \\ &= E(X^2) - 2\mu\, E(X) + \mu^2 \\ &= E(X^2) - 2\mu^2 + \mu^2, \end{aligned}$$

joten

$$\sigma^2 = \mathrm{Var}(X) = E(X^2) - \mu^2 = E(X^2) - [E(X)]^2. \tag{3.3.11}$$

satunnaismuuttujan X hajonta $\sigma_X = \sqrt{\mathrm{Var}(X)}$. Odotusarvon määritelmästä ja identiteetistä (3.3.11) saamme erittäin käyttökelpoisen tuloksen:

$$\mathrm{Var}(cX) = c^2\,\mathrm{Var}(X), \qquad E(X^2) = \mu^2 + \mathrm{Var}(X). \tag{3.3.12}$$

Esimerkki 3.13 Lasketaan diskreettiä tasajakaumaa $\mathrm{Tasd}(1, N)$ noudattavan satunnaismuuttujan varianssi. Esimerkin 3.10 mukaan

$$E(X) = \frac{N+1}{2} \qquad \text{ja} \qquad E(X^2) = \frac{(N+1)(2N+1)}{6}.$$

Soveltamalla kaavaa (3.3.11) saadaan

$$\begin{aligned} \mathrm{Var}(X) &= E(X^2) - [E(X)]^2 \\ &= \frac{(N+1)(2N+1)}{6} - \left(\frac{N+1}{2}\right)^2 = \frac{N^2-1}{12}. \end{aligned}$$

□

3.3.4 Kovarianssi ja korrelaatio

Oletetaan, että satunnaismuuttujilla X ja Y on 2. momentti. Silloin odotusarvot $E(XY)$ ja $E[(X-\mu_X)(Y-\mu_Y)]$ ovat olemassa Apulauseen 3.1 nojalla.

Määritelmä 3.7 (Kovarianssi) Satunnaismuuttujien X ja Y *kovarianssi* σ_{XY} määritellään odotusarvona

$$\begin{aligned}\sigma_{XY} &= \operatorname{Cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] \\ &= E(XY) - \mu_X\mu_Y.\end{aligned} \tag{3.3.13}$$

Kovarianssin avulla voidaan sitten määritellä korrelaatiokerroin.

Määritelmä 3.8 (Korrelaatiokerroin) Satunnaismuuttujien X ja Y *korrelaatiokerroin*

$$\rho_{XY} = \operatorname{Cor}(X,Y) = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}. \tag{3.3.14}$$

Sanomme, että X ja Y ovat positiivisesti (negatiivisesti) korrelotuneita, jos $\rho_{XY} > 0$ (< 0). X ja Y eivät korreloi (korreloimattomia), jos $\rho_{XY} = 0$.

Apulause 3.2 (Summan varianssi) *Oletetaan, että satunnaismuuttujilla X ja Y on varianssi. Silloin*

1. $\operatorname{Var}(X+Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) + 2\operatorname{Cov}(X,Y)$.
2. *Jos satunnaismuuttujalla $X_1, X_2, \ldots, X_n$ on varianssi, niin*

$$\begin{aligned}\operatorname{Var}\left(\sum_{i=1}^n X_i\right) &= \sum_{i=1}^n\sum_{j=1}^n \operatorname{Cov}(X_i, X_j) \\ &= \sum_{i=1}^n \operatorname{Var}(X_i) + \sum_{i=1}^n\sum_{j\neq i}^n \operatorname{Cov}(X_i, X_j).\end{aligned} \tag{3.3.15}$$

Todistus. Todistetaan 1. kohta. Määritelmän mukaan

$$\operatorname{Var}(X+Y) = E[X+Y-(\mu_X+\mu_Y)]^2$$

ja

$$\begin{aligned}[X+Y-(\mu_X+\mu_Y)]^2 &= [(X-\mu_X)+(Y-\mu_Y)]^2 \\ &= (X-\mu_X)^2 + (Y-\mu_Y)^2 + 2(X-\mu_X)(Y-\mu_Y),\end{aligned}$$

missä $\mu_X = E(X)$ ja $\mu_Y = E(Y)$. Odotusarvon lineaarisuuden nojalla

$$\begin{aligned}E[X+Y-(\mu_X+\mu_Y)]^2 &= E(X-\mu_X)^2 + E(Y-\mu_Y)^2 \\ &\quad + 2\,E[(X-\mu_X)(Y-\mu_Y)] \\ &= \operatorname{Var}(X) + \operatorname{Var}(Y) + 2\operatorname{Cov}(X,Y).\end{aligned}$$

Kaava (3.3.15) voidaan todistaa induktiolla. □

3.3.5 Satunnaismuuttujan funktion jakauma

Lauseen 3.9 kohdassa 2 esitetään satunnaismuuttujan X funktion odotusarvo X:n jakauman avulla. Jos Y on X:n funktio, voidaan Y:n todennäköisyysjakauma johtaa X:n jakaumasta. Olkoon $Y = h(X)$ satunnaismuuttujan X funktio ja S_Y satunnaismuuttujan Y arvoalue. Jos $A \subset S_Y$, niin

$$P(Y \in A) = P(h(X) \in A).$$

Esimerkki 3.14 Olkoon X diskreetti satunnaismuuttuja, jonka arvoalue on $S = \{-1, 0, 1, 2\}$ ja todennäköisyysfunktio määritellään seuraavasti:

x:	-1	0	1	2
$f_X(x)$:	0.2	0.3	0.4	0.1

Jos $Y = X^2$, niin Y:n todennäköisyysfunktio on

y:	0	1	4
$f_Y(y)$:	0.3	0.6	0.1

Nyt siis esimerkiksi $P(Y = 1) = P(X = -1) + P(X = 1) = 0.2 + 0.4 = 0.6$. Y:n todennäköisyysfunktion määrittäminen X:n todennäköisyysfunktion avulla on suoraviivainen, vaikkakin joskus työläs prosessi.

Tarkastellaan vielä satunnaismuuttujaa $V = g(X) = (X - \mu_X)^2 = (X - 0.4)^2$, missä $\mu_X = 0.4$. V:n todennäköisyysfunktio on

v:	1.96	0.16	0.36	2.56
$f_Y(v)$:	0.2	0.3	0.4	0.1

ja $E(V) = E[(X - 0.4)^2] = \mathrm{Var}(X)$. □

Olkoot S_X ja S_Y satunnaismuuttujien X ja Y otosavaruudet (arvoalueet). Silloin funktio $h(x)$ määrittelee kuvauksen

$$h\colon S_X \to S_Y.$$

Määritellään *joukon A alkukuva* kuvauksessa h seuraavasti:

$$h^{-1}(A) = \{\, x \in S_X \mid h(x) \in A \,\}. \tag{3.3.16}$$

Joukko A voi olla myös yhden pisteen muodostama joukko eli $A = \{y\}$. Silloin

$$h^{-1}(\{y\}) = \{\, x \in S_X \mid h(x) = y \,\}.$$

Tässä tapauksessa merkitsemme $h^{-1}(y)$ merkinnän $h^{-1}(\{y\})$ sijasta. Huomaa, että $h^{-1}(y)$ on edelleen monen pisteen joukko, jos on useita sellaisia X:n arvoja x, että $h(x) = y$. Jos on vain yksi sellainen x, että $h(x) = y$, niin $h^{-1}(y)$ on yhden pisteen muodostama joukko $\{x\}$ ja kirjoitamme silloin $h^{-1}(y) = x$.

3.3.6 Identtisesti jakautuneet satunnaismuuttujat

Määritelmä 3.9 satunnaismuuttujat X ja Y ovat *identtisesti jakautuneet* eli noudattavat samaa jakaumaa, jos jokaiselle tapahtumalle $A \subset \Omega$ pätee $P(X \in A) = P(Y \in A)$.

Kun X ja Y noudattavat samaa jakaumaa, merkitään $X \sim Y$. Jos $X \sim Y$, niin siitä ei seuraa, että X ja Y ovat sama satunnaismuuttuja. Satunnaismuuttujat X ja Y ovat *identtiset* ($X \equiv Y$) eli samat, jos ne on määritelty samassa otosavaruudessa Ω ja $X(\omega) = Y(\omega)$ kaikilla $\omega \in \Omega$.

Esimerkki 3.15 Esimerkissä 2.6 heitettiin harhatonta lanttia 3 kertaa ja määriteltiin satunnaismuuttuja X = 'kruunujen lukumäärä'. Määritellään myös satunnaismuuttuja Y = 'klaavojen lukumäärä'. Merkitään R = 'kruunu' ja L = 'klaava'. Satunnaismuuttujilla X ja Y on sama jakauma, mutta $X \neq Y$, sillä esimerkiksi $X(\mathrm{RRL}) = 2 \neq Y(\mathrm{RRL}) = 1$. Satunnaismuuttujien X ja Y määritelmistä seuraa, että $X + Y \equiv 3$. $X + Y$ on vakio todennäköisyydellä 1: $P(X + Y = 3) = 1$. □

Satunnaismuuttujan jakauma voidaan luonnehtia kertymäfunktion avulla.

Lause 3.5 *Seuraavat kaksi väitettä ovat yhtäpitävät:*

1. *Satunnaismuuttujat X ja Y noudattavat samaa jakaumaa.*
2. *$F_X(x) = F_Y(x)$ kaikilla $x \in \mathbb{R}$, missä F_X on X:n ja F_Y on Y:n kertymäfunktio.*

Kun X ja Y ovat diskreettejä, niin $X \sim Y$, jos $f_X(x) = f_Y(x)$ kaikilla $x \in \mathbb{R}$.

Esimerkki 3.16 Heitetään harhatonta lanttia 4 kertaa. Olkoon kruunun todennäköisyys p. X ja Y on määritelty samoin kuin Esimerkissä 3.15. Mikä on tapahtuman $\{X = Y\}$ todennäköisyys? Tapahtuma $\{X = Y\}$ on

$$\{\,\omega \mid X(\omega) = Y(\omega)\,\} = \{\mathrm{RRLL}, \mathrm{LRRL}, \mathrm{LLRR}, \mathrm{LRLR}, \mathrm{RLLR}, \mathrm{RLRL}\}.$$

Jokaisen yksittäisen alkeistapahtuman (jonon) todennäköisyys on $p^2(1-p)^2$ ja jonoja on $\binom{4}{2} = 6$ kappaletta, joten

$$P(X = Y) = \binom{4}{2} p^2 (1-p)^2.$$

Milloin $X \sim Y$? Koska

$$f_X(x) = \binom{4}{x} p^x (1-p)^{4-x}, \qquad x = 0, 1, 2, 3, 4$$

ja

$$f_Y(y) = \binom{4}{y} (1-p)^y p^{4-y}, \qquad y = 0, 1, 2, 3, 4,$$

niin $f_X(x) = f_Y(x)$ kaikilla $x = 0, 1, 2, 3, 4$ jos ja vain jos $p = \frac{1}{2}$. Siis $X \sim Y$, kun $p = \frac{1}{2}$. □

3.3.7 Satunnaismuuttujien riippumattomuus

Määrittelimme tapahtumien riippumattomuuden alaluvussa 3.1.2. Tarkastelemme nyt satunnaismuuttujien riippumattomuutta.

Määritelmä 3.10 (Satunnaismuuttujien riippumattomuus) Satunnaismuuttujat X ja Y ovat riippumattomat jos

$$P(X \in A, Y \in B) = P(X \in A)\,P(Y \in B) \tag{3.3.17}$$

kaikilla joukoilla $A \subset \mathbb{R}$ ja $B \subset \mathbb{R}$.

Merkintä $P(X \in A,\, Y \in B)$ on lyhennys merkinnästä $P(\{X \in A\} \cap \{Y \in B\})$. Satunnaismuuttujat X ja Y ovat siis riippumattomat, jos tapahtumat $\{X \in A\}$ ja $\{X \in B\}$ ovat riippumattomat kaikilla $A \subset \mathbb{R}$ ja $B \subset \mathbb{R}$. Riippumattomuuden määritelmästä seuraa esimerkiksi, että kaikilla $x, y \in \mathbb{R}$

$$P(X = x, Y = y) = P(X = x)\,P(Y = y) = f_X(x) f_Y(y), \tag{3.3.18}$$

missä $f_X(x)$ on X:n ja $f_Y(y)$ on Y:n todennäköisyysfunktio.

Lause 3.6 *Jos X ja Y ovat riippumattomat, niin $U = g(X)$ ja $V = h(Y)$ ovat riippumattomat, missä $g(x)$ on pelkästään x:n (ts. X:n arvojen) funktio ja $h(y)$ pelkästään y:n funktio.*

Todistus. Määritellään $A_u = \{\, x \mid g(x) = u \,\}$ ja $A_v = \{\, y \mid h(y) = v \,\}$. Silloin kaikilla u ja v

$$\begin{aligned}
P(U = u, V = v) &= P[g(X) = u, h(Y) = v] \\
&= P(X \in A_u, Y \in A_v) \\
&= P(X \in A_u)\,P(Y \in A_v) \qquad (X \text{ ja } Y \text{ riippumattomat}) \\
&= P(U = u)\,P(V = v),
\end{aligned}$$

joten U ja V ovat riippumattomat. □

Määritelmä 3.10 pitää täsmälleen paikkansa vain diskreeteille satunnaismuuttujille. Koska yleisessä tapauksessa kaikki Ω:n osajoukot eivät ole tapahtumia, niin silloin on rajoituttava sopivasti määriteltyyn Ω:n osajoukkokokoelmaan. Yhtälö (3.3.17) pitää myös paikkansa, jos toinen oikean puolen tekijöistä on nolla. Huomaa, että $P(X \in A) = 0$ tarkoittaa, että $\{\, \omega \mid X(\omega) \in A \,\} = \emptyset$. Silloin

$$\{X \in A,\, Y \in B\} = \{\, \omega \mid X(\omega) \in A \,\} \cap \{\, \omega \mid Y(\omega) \in B \,\} = \emptyset,$$

joten $P(X \in A,\, Y \in B) = 0$.

Identiteettiä (3.3.18) voidaan myös pitää diskreettien satunnaismuuttujien X ja Y riippumattomuuden määritelmänä, sillä siitä seuraa identiteetti (3.3.17). Jos valitaan kaksi mielivaltaista numeroituvaa joukkoa $A \subset \mathbb{R}$ ja $B \subset \mathbb{R}$ sekä oletetaan (3.3.18), saadaan

$$\begin{aligned} P(X \in A, Y \in B) &= \sum_{x_i \in A} \sum_{y_j \in B} P(X = x_i, Y = y_j) \\ &= \sum_{x_i \in A} \sum_{y_j \in B} P(X = x_i)\, P(Y = y_j) \qquad [(3.3.18)] \\ &= \sum_{x_i \in A} P(X = x_i) \sum_{y_j \in B} P(Y = y_j) \\ &= P(X \in A)\, P(Y \in B). \end{aligned}$$

Näin olemme todenneet, että ehdot (3.3.17) ja (3.3.18) ovat yhtäpitävät.

Tämän luvun alussa määritelty tapahtumien riippumattomuus on itse asiassa satunnaismuuttujien riippumattomuuden erikoistapaus. Olkoon I_A tapahtuman A ja I_B tapahtuman B indikaattorifunktio. Huomaa, että I_A ja I_B ovat satunnaismuuttujia. Koska indikaattorifunktio saa vain arvot 1 tai 0, niin esimerkiksi

$$\{I_A = 1\} = A \qquad \text{ja} \qquad \{I_A = 0\} = A^c.$$

Jos I_A ja I_B ovat riippumattomat, niin

$$P(I_A = x, I_B = y) = P(I_A = x)\, P(I_B = y) \tag{3.3.19}$$

kaikilla $x, y \in \mathbb{R}$. Nyt siis $\{I_A = x\}$ on joko A, A^c tai $\emptyset$ ja $\{I_B = y\}$ on joko B, B^c tai $\emptyset$. Tästä seuraa mm. tapahtumien A ja B riippumattomuuden määritelmä

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

Lisäksi saadaan identiteetit

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

Lauseen 3.1 nojalla jokainen näistä identiteeteistä kelpaa A:n ja B:n riippumattomuuden määritelmäksi.

3.3.8 Useiden satunnaismuuttujien riippumattomuus

Satunnaismuuttujat $X_1, \ldots, X_n$ ovat riippumattomat, jos

$$\begin{aligned} P(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n) \\ = P(X_1 \in A_1)\, P(X_2 \in A_2) \cdots P(X_n \in A_n) \end{aligned} \tag{3.3.20}$$

kaikilla (sopivasti valituilla) joukoilla $A_i \subset \mathbb{R}$, $1 \le i \le n$. Jos $X_1, \ldots, X_n$ ovat diskreettejä, niin (3.3.20) pitää paikkansa kaikille joukoille $A_i \subset \mathbb{R}$, $1 \le i \le n$. Yleisessä tapauksessa on A_i:t ($1 \le i \le n$) valittava niin, että joukot $\{X_i \in A_i\} = \{\,\omega \mid X_i(\omega) \in A_i\,\}$ ovat tapahtumia. Huomaa, että riippumattomien satunnaismuuttujien $X_1, \ldots, X_n$ jokainen osajono $X_{i_1}, \ldots, X_{i_k}$ on riippumaton [$1 \le k \le n$ ja $\{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$]. Jos esimerkiksi X_1, X_2 ja X_3 ovat riippumattomat, niin myös X_1 ja X_2 ovat riippumattomat. Tämä nähdään, kun valitaan $A_3 = \mathbb{R}$. Silloin $\{X_3 \in \mathbb{R}\} = \Omega$ ja

$$\begin{aligned}\{X_1 \in A_1, X_2 \in A_2, X_3 \in \mathbb{R}\} &= \{X_1 \in A_1\} \cap \{X_2 \in A_2\} \cap \Omega \\ &= \{X_1 \in A_1, X_2 \in A_2\},\end{aligned}$$

joten identiteetin (3.3.20) mukaan

$$\begin{aligned}P(X_1 \in A_1, X_2 \in A_2) &= P(X_1 \in A_1)\,P(X_2 \in A_2)\,P(\Omega) \\ &= P(X_1 \in A_1)\,P(X_2 \in A_2).\end{aligned}$$

3.4 Suurten lukujen laki

Riippumattomat, samoin jakautuneet satunnaismuuttujat (rsj). Riippumattomien satunnaismuuttujien jono $X_1, X_2, \ldots$ (äärellinen tai ääretön) on samoin jakautunut, jos jokaisella jonon satunnaismuuttujalla on sama jakauma. Sanomme lyhyesti, että jono $X_1, X_2, \ldots$ on *rsj*. Silloin jonon satunnaismuuttujilla on sama kertymäfunktio F, joten

$$P(X_k \le x) = F(x) \qquad \text{kaikilla } x \in \mathbb{R}.$$

Jos siis yhden satunnaismuuttujan X_k odotusarvo on μ ja varianssi σ^2, silloin niiden kaikkien kaikkien odotusarvo on μ ja varianssi σ^2.

Lause 3.7 (Markovin epäyhtälö) *Olkoon $X \ge 0$ epänegatiivinen satunnaismuuttuja. Silloin*

$$P(X \ge a) \le \frac{E(X)}{a}, \qquad \textit{kun } a > 0.$$

Todistus. Olkoon I_A joukon $A = \{\,\omega \mid X(\omega) \ge a\,\}$ indikaattorifunktio [ks. (2.3)]. Koska sekä indikaattorifunktio että X ovat epänegatiiviset ja $I_A + I_{A^c} = 1$, niin

$$X = I_A X + I_{A^c} X \ge I_A X \ge a I_A.$$

Viimeinen epäyhtälö seuraa siitä, että $X(\omega) \ge a$ ja $I_A(\omega) = 1$, kun $\omega \in A$. Jos taas $\omega \notin A$, niin $I_A(\omega) = 0$, joten $I_A(\omega)X(\omega) = I_A(\omega)a = 0$. Keskiarvon monotoonisuuden (Lause 3.9, 3. kohta) ja lineaarisuuden (1. kohta) nojalla saadaan

$$E(X) \ge E(aI_A) = a\,E(I_A) = a\,P(X \in A) = a\,P(X \ge a),$$

koska tapahtumat $\{X \in A\}$ ja $\{X \ge a\}$ ovat määritelmän mukaan ekvivalentteja. □

Markovin epäyhtälön avulla on helppo todistaa erittäin käyttökelpoinen *Tšebyševin epäyhtälö.*

Lause 3.8 (Tšebyševin epäyhtälö) *Olkoon X satunnaismuuttuja, jonka keskiarvo on μ ja varianssi σ^2. Silloin*

$$P(|X-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}, \qquad \textit{kaikilla } \varepsilon > 0. \tag{3.4.1}$$

Todistus. Määritellään satunnaismuuttuja $Y = h(X) = (X-\mu)^2$ ja valitaan $a = \varepsilon^2 > 0$. Koska $Y \geq 0$ ja $E(Y) = \sigma^2$, seuraa Tšebyševin epäyhtälö (3.4.1) suoraan Markovin epäyhtälöstä. □

Lause 3.9 *Oletetaan, että otosavaruudessa Ω määritellyllä diskreeteillä satunnaismuuttujilla X ja Y on odotusarvo ja $a \in \mathbb{R}$ on vakio. Silloin*

1. *$E(aX) = a\,E(X)$ ja $E(X+Y) = E(X) + E(Y)$, joten odotusarvo on lineaarinen operaattori.*

Olkoot $h(x)$, $h_1(x)$ ja $h_2(x)$ sellaisia funktioita, että satunnaismuuttujilla $h(X)$, $h_1(X)$ ja $h_2(X)$ on odotusarvo. Silloin seuraavat tulokset pitävät paikkansa:

2. $E[h(X)] = \sum_x h(x) f_X(x) = \sum_x h(x)\,P(X=x)$
3. *Jos $h_1(x) \geq h_2(x)$ kaikilla x, niin $E[h_1(X)] \geq E[h_2(X)]$.*

Lause 3.10 (Tulon odotusarvo, riippumattomat SM:t) *Olkoot satunnaismuuttujat X ja Y riippumattomat.*

1. *Jos $E(X)$ ja $E(Y)$ ovat olemassa, niin $E(XY) = E(X)\,E(Y)$.*

Olkoot satunnaismuuttujat $X_1, X_2, \ldots, X_n$ riippumattomat.

2. *Jos satunnaismuuttujilla $X_1, X_2, \ldots, X_n$ on odotusarvo, niin*

$$E(X_1 X_2 \cdots X_n) = E(X_1)\,E(X_2) \cdots E(X_n).$$

Todistus. 1. Odotusarvon määritelmän mukaan

$$\begin{aligned}
E(XY) &= \sum_x \sum_y xy\,P(X=x, Y=y) \\
&= \sum_x \sum_y xy\,P(X=x)\,P(Y=y) \qquad [X \text{ ja } Y \text{ riippumattomat}] \\
&= \Big[\sum_x x\,P(X=x)\Big]\Big[\sum_y y\,P(Y=y)\Big] \\
&= E(X)\,E(Y).
\end{aligned}$$

Koska $\sum_x x\,P(X=x)$ ja $\sum_y y\,P(Y=y)$ suppenevat itseisesti odotusarvojen olemassaolon nojalla, pitää 3. yhtäsuuruus paikkansa ja myös odotusarvon $E(XY)$ olemassaolo seuraa odotusarvojen $E(X)$ ja $E(Y)$ olemassaolosta.

Kohta 2. voidaan todistaa soveltamalla toistuvasti 1. kohdan tulosta. □

Apulause 3.3 (Summan varianssi, riippumattomat SM:t) *Oletetaan, että X_1, X_2, ..., X_n ovat riippumattomat ja niillä on varianssi. Silloin*

$$\operatorname{Cov}(X_i, X_j) = 0, \qquad i \neq j,$$

ja

$$\operatorname{Var}(X_1 + X_2 + \cdots + X_n) = \operatorname{Var}(X_1) + \operatorname{Var}(X_2) + \cdots + \operatorname{Var}(X_n).$$

Todistus. Jos $i \neq j$, niin

$$\begin{aligned}\operatorname{Cov}(X_i, X_j) &= E(X_i X_j) - E(X_i)\,E(X_j) \\ &= E(X_i)\,E(X_j) - E(X_i)\,E(X_j) = 0,\end{aligned}$$

koska X_i:n ja X_j:n riippumattomuuden nojalla $E(X_i X_j) = E(X_i)\,E(X_j) = 0$. Summan varianssin $\operatorname{Var}(\sum_{i=1}^n X_i)$ lauseke seuraa nyt suoraan Apulauseesta 3.2. □

Apulause 3.4 (Otoskeskiarvon odotusarvo ja varianssi) *Olkoot X_1, X_2, ..., X_n RSJ satunnaismuuttujat, joiden keskiarvo on μ ja varianssi σ^2. Määritellään satunnaismuuttujat*

$$S_n = X_1 + X_2 + \cdots + X_n, \qquad \overline{X}_n = \frac{S_n}{n}.$$

Silloin

$$E(S_n) = n\mu, \qquad \operatorname{Var}(S_n) = n\sigma^2, \qquad E(\overline{X}_n) = \mu, \qquad \operatorname{Var}(\overline{X}_n) = \frac{\sigma^2}{n}.$$

Voimme nyt todistaa Tšebyševin epäyhtälön avulla ns. *heikon suurten lukujen lain* (HSLL).

Lause 3.11 (Heikko suurten lukujen laki (HSLL)) *Olkoon X_1, X_2, ..., X_n ääretön RSJ satunnaismuuttujien jono, jossa jokaisen satunnaismuuttujan keskiarvo on μ ja varianssi σ^2. Olkoon $S_n = X_1 + X_2 + \cdots + X_n$ ja*

$$\overline{X}_n = \frac{S_n}{n}.$$

Silloin jokaisella $\varepsilon > 0$,

$$P(|\overline{X}_n - \mu| \geq \varepsilon) \to 0, \qquad \textit{kun } n \to \infty.$$

Todistus. Apulauseen 3.4 ja Tšebyševin epäyhtälön mukaan

$$P(|\overline{X}_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}.$$

Kun $n \to \infty$, niin $\sigma^2/(n\varepsilon^2) \to 0$, joten

$$P(|\overline{X}_n - \mu| \geq \varepsilon) \to 0.$$

Näin on lause todistettu. □

Heikko suurten lukujen laki sanoo, että otoskeskiarvo lähenee todennäköisyyden mielessä todellista keskiarvoa, kun otoskoko kasvaa.

3.5 Generoivat funktiot ja momentit

3.5.1 Momentit

Eräs tapa luonnehtia satunnaismuuttujan jakaumaa, on laskea jakauman momentit. Ne määritellään odotusarvon avulla.

Määritelmä 3.11 Olkoon r positiivinen kokonaisluku. Jos odotusarvo

$$\alpha_r = E(X^r)$$

on olemassa, se on satunnaismuuttujan X (tai X:n jakauman) r. momentti. Vastaavasti X:n r. keskusmomentti on

$$\mu_r = E[(X - \mu)^r],$$

missä $\mu = E(X) = \alpha_1$.

Momenttia α_r kutsutaan joskus myös *origomomentiksi.* Jakauman keskiarvo on siis 1. origomomentti ja varianssi 2. keskusmomentti. Satunnaismuuttujan X *tekijämomentit* g_r, $r = 1, 2, \ldots$ määritellään seuraavasti:

$$g_r = E[X^{(r)}] = E[X(X-1)\cdots(X-r+1)].$$

Ensimmäiset kaksi tekijämomenttia ovat

$$\begin{aligned} g_1 &= E(X) = \alpha_1 = \mu, \\ g_2 &= E[X(X-1)] = E(X^2 - X) = E(X^2) - E(X) = \alpha_2 - \mu. \end{aligned}$$

Koska $\sigma^2 = \alpha_2 - \mu^2$, niin

$$\sigma^2 = g_2 + \mu - \mu^2.$$

3.5.2 Momenttifunktio

Esittelemme nyt uuden todennäköisyysjakaumaan liittyvän funktion, *momentteja generoivan funtion,* jota kutsutaan lyhyesti *momenttifunktioksi* (mf). Momenttifunktio tarjoaa erään yleisen menetelmän momenttien laskemiseksi, vaikka se ei aina ole siihen tarkoitukseen helpoin tai tehokkain menetelmä. Momenttien laskemista tärkeämpää on se, että jakaumat voidaan luonnehtia kätevästi momenttifunktion avulla (mikäili se on olemassa).

Määritelmä 3.12 Olkoon X diskreetti satunnaismuuttuja, jonka todennäköisyysfunktio on $f(x)$ ja arvoavaruus S. Silloin reaalimuuttujan t funktio

$$M(t) = E(\mathrm{e}^{tX})$$

on satunnaismuuttujan X (tai X:n jakauman) momenttifunktio (mf), jos odotusarvo

$$E(\mathrm{e}^{tX}) = \sum_{x \in S} \mathrm{e}^{tx} f(x)$$

on olemassa jollain avoimella välillä $-a < t < a$, missä $a > 0$.

Määritelmän perusteella on selvää, että

$$M(0) = E(\mathrm{e}^{0\cdot X}) = \sum_{x\in S} f(x) = 1.$$

Olkoon $S = \{x_1, x_2, \ldots\}$. Silloin

$$M_X(t) = \mathrm{e}^{tx_1} f(x_1) + \mathrm{e}^{tx_2} f(x_2) + \cdots ,$$

missä e^{tx_k}:n kertoimet

$$f(x_k) = P(X = x_k), \qquad k = 1, 2, \ldots$$

ovat todennäköisyyksiä. Olkoon $f(x)$ satunnaismuuttujan X todennäköisyysfunktio, $g(y)$ satunnaismuuttujan Y todennäköisyysfunktio ja $S = \{a_1, a_2, \ldots\}$ X:n ja Y:n yhteinen arvoavaruus. Jos

$$M_X(t) = M_Y(t), \qquad \text{kaikilla } t,\ -h < t < h,$$

niin matemaattisen analyysin teorian nojalla

$$f(a_k) = g(a_k), \qquad k = 1, 2, \ldots$$

Jos siis kahdella satunnaismuuttujalla on sama momenttifunktio, niin niillä täytyy olla sama jakauma. Olkoon $F_X(u)$ X:n ja $F_Y(u)$ Y:n kertymäfunktio. Esitetään nyt momenttifunktion yksikäsitteisyyttä koskeva tulos lauseen muodossa.

Lause 3.12 *Olkoot satunnaismuuttujien X ja Y momenttifunktiot $M_X(t)$ ja $M_Y(t)$. Jos $M_X(t) = M_Y(t)$ kaikilla t jossain nollan ympäristössä, niin $F_X(u) = F_Y(u)$ kaikilla u:n arvoilla eli X:llä ja Y:llä on sama jakauma.*

Esimerkki 3.17 Jos $X \sim \mathrm{Ber}(p)$, niin

$$M(t) = E(\mathrm{e}^{tX}) = \mathrm{e}^{t\cdot 1}p + \mathrm{e}^{t\cdot 0}q = \mathrm{e}^{t}p + q,$$

missä $q = 1 - p$. □

Lause 3.13 *Olkoot X ja Y riippumattomat satunnaismuuttujat, joiden momenttifunktiot ovat $M_X(t)$ ja $M_Y(t)$. Silloin satunnaismuuttujan $Z = X + Y$ momenttifunktio on*

$$M_Z(t) = M_X(t) M_Y(t). \tag{3.5.1}$$

Todistus. Koska e^{tX} on pelkästään x:n (X:n arvojen) funktio ja e^{tY} pelkästään y:n funktio, niin Lauseen 3.6 mukaan e^{tX} ja e^{tY} ovat riippumattomat. Väite

$$E(\mathrm{e}^{tZ}) = E[\mathrm{e}^{t(X+Y)}] = E[\mathrm{e}^{tX}\mathrm{e}^{tY}] = E(\mathrm{e}^{tX})\, E(\mathrm{e}^{tY})$$

seuraa sitten suoraan Lauseesta 3.10. □

Usean satunnaismuuttujan tapauksessa on voimassa vastaava tulos.

Seuraus 3.1 *Olkoot X_1, X_2, ..., X_n riippumattomat satunnaismuuttujat, joiden momenttifunktiot ovat $M_{X_i}(t)$, $i = 1, 2, \ldots, n$. Silloin summan*

$$S_n = X_1 + X_2 + \cdots + X_n$$

momenttifunktio on

$$M_{S_n}(t) = M_{X_1}(t)M_{X_2}(t)\cdots M_{X_n}(t).$$

Jos momenttifunktio $M(t)$ on olemassa välillä $(-h, h)$, niin momenttifunktiolla on kaikkien kertalukujen derivaatat pisteessä $t = 0$. Kun identiteetti

$$M(t) = \sum_{x \in S} \mathrm{e}^{tx} f(x) \tag{3.5.2}$$

derivoidaan puolittain, voidaan oikea puoli derivoida termeittäin ja yhtäsuuruus säilyy. Derivoimalla lauseke (3.5.2) puolittain muuttujan t suhteen saadaan

$$M(t)' = \sum_{x \in S} x\mathrm{e}^{tx} f(x),$$
$$M(t)'' = \sum_{x \in S} x^2\mathrm{e}^{tx} f(x)$$

ja jokaisella positiivisella kokonaisluvulla r

$$M(t)^{(r)} = \sum_{x \in S} x^r\mathrm{e}^{tx} f(x).$$

Sijoittamalla $t = 0$ saadaan

$$M(0)' = \sum_{x \in S} x f(x) = E(X),$$
$$M(0)'' = \sum_{x \in S} x^2 f(x) = E(X^2)$$

ja yleisesti

$$M(0)^{(r)} = \sum_{x \in S} x^r f(x) = E(X^r).$$

Erityisesti

$$\mu = M(0)' \qquad \text{ja} \qquad \sigma^2 = M(0)'' - [M(0)']^2.$$

Lause 3.14 *Olkoon $M_X(t)$ satunnaismuuttujan X momenttifunktio ja $Y = aX + b$, missä a ja b ovat annettuja reaaliarvoisia vakioita. Silloin $M_Y(t) = \mathrm{e}^{bt}M_X(at)$.*

Lause 3.15 (Momenttifunktioiden suppeneminen) *Olkoon* $X_1, X_2, X_3, \ldots$ *satunnaismuuttujien jono, jossa jokaisella* X_n*:llä on momenttifunktio* $M_{X_n}(t)$, $n = 1, 2, 3, \ldots$ *Oletetaan lisäksi, että*

$$M_{X_n}(t) \to M_X(t)$$

kaikilla t:n arvoilla jossain nollan ympäristössä $(-h, h)$*, kun* $n \to \infty$*. Jos* $M_X(t)$ *on momenttifunktio, niin silloin on olemassa yksikäsitteinen kertymäfunktio* $F_X(x)$*, jonka momenttifunktio on* $M_X(t)$ *ja*

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

kaikissa pisteissä x*, joissa* $F_X(x)$ *on jatkuva.*

Satunnaismuuttujien momenttifunktioiden suppenemisesta seuraa siis satunnaismuuttujien kertymäfunktioiden suppeneminen.

3.5.3 Todennäköisyydet generoiva funktio (tgf)

Diskreetin satunnaismuuttujan X todennäköisyydet generoiva funktio (tgf) $G(t)$ määritellään seuraavasti:

$$G(t) = E(t^X) = \sum_{i=1}^{\infty} f(x_i) t^{x_i}.$$

Nähdään helposti, että $G(1) = \sum_{i=1}^{\infty} f(x_i) = 1$. Sarja suppenee ainakin silloin, kun $|t| < 1$. Kun sarja derivoidaan termeittäin, saadaan

$$G'(t) = \sum_{i=1}^{\infty} x_i f(x_i) t^{x_i - 1}.$$

Jos $G(t)$ on olemassa jollain välillä $(-h-1, h+1)$, $h > 0$, niin

$$G'(1) = E(X)$$

ja yleisesti

$$G^{(r)}(1) = E(X^{(r)}) = E[X(X-1)\cdots(X-r+1)]$$

kaikilla positiivisilla kokonaisluvuilla r. Todennäköisyydet generoiva funktio liittyy läheisesti momenttifunktioon, sillä

$$G(\mathrm{e}^t) = E(\mathrm{e}^{tX}) = M(t).$$

3.6 Kokeiden yhdistäminen ja tulomallit

Tarkastellaan nyt satunnaiskokeita $\mathcal{E}_1$ ja $\mathcal{E}_2$, joiden otosavaruudet ovat vastaavasti Ω_1 ja Ω_2. Olkoot satunnaiskokeisiin liittyvät todennäköisyysjakaumat $\{p_i\}$ ja $\{q_i\}$ $i = 1, 2, \ldots$ Tarkastelemme seuraavassa vain numeroituvia otosavaruuksia. Yhdistetään kokeet siten, että tehdään kokeet $\mathcal{E}_1$ ja $\mathcal{E}_2$. Merkitään yhdistettyä koetta $\mathcal{E}_1 \times \mathcal{E}_1$. Yhdistetyn kokeen tulos esitetään järjestettynä parina (ω_i, ω_j), missä $\omega_i \in \Omega_1$ on kokeen $\mathcal{E}_1$ tulos ja $\omega_j \in \Omega_2$ on kokeen $\mathcal{E}_2$ tulos. Yhdistetyn kokeen otosavaruus on siis otosavaruuksien Ω_1 ja Ω_2 *karteesinen tulo* $\Omega_1 \times \Omega_2 = \{\, (\omega_i, \omega_j) \mid \omega_i \in \Omega_1 \text{ ja } \omega_j \in \Omega_2 \,\}$. Vastaavalla tavalla voidaan yhdistää useampiakin kokeita.

Määrittelemme nyt yhdistettyyn kokeeseen $\mathcal{E}_1 \times \mathcal{E}_2$ liittyvän todennäköisyysjakauman $\Omega_1 \times \Omega_2$:ssa. *Kokeet ovat riippumattomat* jos ja vain jos

$$P(\omega_i, \omega_j) = p_i q_j \tag{3.6.1}$$

kaikilla $\omega_i \in \Omega_1$ ja $\omega_j \in \Omega_2$, missä $p_i = p(\omega_i)$ on ω_i:n todennäköisyys Ω_1:ssä ja $q_j = p(\omega_j)$ on ω_j:n todennäköisyys Ω_2:ssä. Selvästikin $P(\omega_i, \omega_j) \geq 0$ kaikilla $(\omega_i, \omega_j) \in \Omega_1 \times \Omega_2$. Koska $\sum\limits_{\omega_i \in \Omega_1} p_i = \sum\limits_{\omega_j \in \Omega_2} q_j = 1$, niin

$$\sum_{(\omega_i,\omega_j)\in\Omega_1\times\Omega_2} P(\omega_i, \omega_j) = \sum_{\omega_i\in\Omega_1} \sum_{\omega_j\in\Omega_2} p_i q_j = \Big(\sum_{\omega_i\in\Omega_1} p_i\Big)\Big(\sum_{\omega_j\in\Omega_2} q_j\Big) = 1.$$

Identiteetti (3.6.1) siis määrittelee todennäköisyysjakauman $\Omega_1 \times \Omega_2$:ssa. Sitä kutsutaan yhdistetyn kokeen $\mathcal{E}_1 \times \mathcal{E}_1$ *tulomalliksi.*

Riippumattomat toistot

Tulomallin tärkeä erikoistapaus saadaan toistamalla n kertaa koe $\mathcal{E}$, jonka otosavaruus on Ω. Tällaista koetta sanotaan *toistokokeeksi* ja sitä merkitään $\mathcal{E}^n$. Yhdistetyn kokeen otosavaruus on $\Omega \times \Omega \times \cdots \times \Omega$, jonka alkeistapaukset ovat muotoa $\boldsymbol{\omega} = (\omega_1, \omega_2, \ldots, \omega_n)$, missä ω_i on i. toiston tulos. Olkoon $p(\omega)$ satunnaiskokeeseen $\mathcal{E}$ liittyvässä otosavaruudessa Ω määritelty jakaumafunktio. Toistokokeeseen $\mathcal{E}^n$ liittyvä jakaumafunktio määritellään seuraavasti:

$$p(\boldsymbol{\omega}) = p(\omega_1)p(\omega_2)\cdots p(\omega_n).$$

Bernoullin koe

Bernoullin koe (nimetty James Bernoullin mukaan) on koe, jossa on täsmälleen kaksi tulosvaihtoehtoa. Usein toista tulosvaihtoehtoa kutsutaan onnistumiseksi (O) ja toista epäonnistumiseksi (E), joten Bernoullin kokeen otosavaruus $\Omega = \{\mathrm{O}, \mathrm{E}\}$. Satunnaismuuttuja X noudattaa *Bernoullin jakaumaa,* kun

$$X = \begin{cases} 1, & \text{todennäköisyydellä } P(\mathrm{O}) = p; \\ 0, & \text{todennäköisyydellä } 1 - p, \end{cases} \tag{3.6.2}$$

missä $0 \leq p \leq 1$. Myös satunnaismuuttujan arvoa $X = 1$ kutsutaan onnistumiseksi ja p:tä onnistumistodennäköisyydeksi. Vastaavasti arvoa $X = 0$ kutsutaan epäonnistumiseksi. Huomaa, että X on 'onnistumisen' indikaattorifunktio. Bernoullin kokeen riippumattomat toistot muodostavat *Bernoullin toistokokeen.*

Esimerkki 3.18 Esimerkissä 2.6 heitetään harhatonta lanttia 3 kertaa. Yhdessä lantin heitossa otosavaruus $\Omega = \{\mathrm{R}, \mathrm{L}\}$. Voidaan sopia esimerkiksi, että kruunu (R) on onnistuminen ja klaava (L) on epäonnistuminen. Vastaavan Bernoullin jakaumaa noudattavaan satunnaismuuttujaan liittyvä otosavaruus $S = \{1, 0\}$. Lantin heitto on Bernoullin koe. Tehdään kolme riippumatonta Bernoullin koetta. Tähän yhdistettyyn kokeeseen liittyvä otosavaruus on $S \times S \times S = \{\,(s_1, s_2, s_3) \mid s_i \in S\,\} = \{111, 110, 101, 100, 011, 010, 001, 000\}$. $\square$

Kun toistetaan Bernoullin koe n kertaa (riippumattomat toistot), ovat kokeen mahdolliset tulokset n:n pituisia 1:n ja 0:n muodostamia jonoja. Tyypillinen jono on muotoa $111011000\ldots110$, jonka todennäköisyys on

$$ppp(1-p)p(1-p)(1-p)ppp\cdots pp(1-p) = p^k(1-p)^{n-k},$$

missä k on onnistumisten lukumäärä ja $n-k$ epäonnistumisten lukumäärä. Erilaisten mahdollisten jonojen lukumäärä on 2^n.

Binomijakauma voidaan määritellä Bernoullin toistokokeen avulla. Olkoon $X_1, X_2, \ldots, X_n$ samaa Bernoullin jakaumaa noudattavien riippumattomien satunnaismuuttujien jono, missä $P(X_i = 1) = p$ ja $P(X_i = 0) = 1 - p = q$, $i = 1, 2, \ldots, n$. Silloin $E(X_i) = p$ ja $\mathrm{Var}(X_i) = pq$. Onnistumisten lukumäärä n:ssä riippumattomassa Bernoullin kokeessa on

$$X = X_1 + X_2 + \cdots + X_n.$$

Mikä on todennäköisyys, että onnistumisia on x $(0 \leq x \leq n)$ kappaletta? Jos jonossa on täsmälleen x ykköstä, niin jonon todennäköisyys on $p^x(1-p)^{n-x}$. Tällaisia jonoja on yhteensä $\binom{n}{x}$ kappaletta. Onnistumisten lukumäärä n:ssä Bernoullin kokeessa noudattaa *binomijakaumaa*

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x},$$

missä siis $f(x) = P(X = x)$.

Onnistumisten lukumäärän otoskeskiarvo on

$$\overline{X}_n = \frac{S_n}{n}.$$

Se on onnistumisten suhteellinen frekvenssi n:ssä riippumattomassa Bernoullin kokeessa, esimerkiksi kruunujen suhteellinen frekvenssi lantin heitossa. Apulauseen 3.4 mukaan $E(\overline{X}_n) = p$ ja $\mathrm{Var}(\overline{X}_n) = pq/n$. HSSL:n mukaan

$$P(|\overline{X}_n - p| > \varepsilon) \to 0$$

kaikilla $\varepsilon > 0$, kun n kasvaa. Kruunujen suhteellinen frekvenssi lähenee p:tä todennäköisyyden mielessä, kun heittojen määrä kasvaa. Bernoulli todisti tämän tuloksen 1713. Tulosta kutsutaan hänen mukaansa *Bernoullin suurten lukujen laiksi.* Ensimmäisessä luvussa tarkasteltiin suhteellisen frekvenssin raja-arvoa todennäköisyyden tulkintana ja eräänlaisena perusteluna todennäköisyydelle. Nyt näemme, että tämä suhteellisen frekvenssin raja-arvotulos on yksi todennäköisyyslaskennan perustuloksista.

3.6.1 Yleinen tulokaava

Yleensä todennäköisyysongelmat koskevat useita tapahtumia tai satunnaismuuttujia, joiden keskinäisiä riippuvuuksia tarkastellaan. Tietyssä mielessä kaikki todennäköisyydet ovat ehdollisia, mutta tavallisesti selvänä pidetyt ehdot jätetään mainitsematta. Rahanheitossa mainitsemme vain vaihtoehdot 'kruunu' ja 'klaava', vaikka lantti voi jäädä myös reunalleen. Presidenttiehdokkaasta tulee presidentti vain sillä ehdolla, että säilyy hengissä vaalikampanjan ajan. Valitsemistodennäköisyyttä laskettaessa ei hengissäpysymisen todennäköisyyttä tavallisesti oteta huomioon.

Seuraavassa esitetään yleinen tulokaava. Huomaa, että jatkossa leikkausta $A_1 \cap A_2$ merkitään kaavojen yksinkertaistamiseksi lyhyesti $A_1 A_2$.

Väittämä 3.1 (Tulokaava) *Olkoot A_1, A_2, …, A_n mitä tahansa tapahtumia. Silloin*

$$P(A_1 A_2 \cdots A_n) = P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 A_2) \cdots \cdot P(A_n \mid A_1 A_2 \cdots A_{n-1}), \tag{3.6.3}$$

jos $P(A_1 A_2 \cdots A_{n-1}) > 0$.

Todistus. Jos $P(A_1 A_2 \cdots A_{n-1}) > 0$, niin kaavassa (3.6.3) esitetyt ehdolliset todennäköisyydet ovat hyvin määritellyt, koska

$$P(A_1) \geq P(A_1 A_2) \geq \cdots \geq P(A_1 A_2 \cdots A_{n-1}) > 0.$$

Kun yhtälön (3.6.3) oikea puoli kirjoitetaan auki ehdollisen todennäköisyyden kaavaa (3.1.1) soveltaen, saadaan

$$\frac{P(A_1)}{P(\Omega)} \cdot \frac{P(A_1 A_2)}{P(A_1)} \cdot \frac{P(A_1 A_2 A_3)}{P(A_1 A_2)} \cdots \frac{P(A_1 A_2 \cdots A_n)}{P(A_1 A_2 \cdots A_{n-1})},$$

joka supistuu todennäköisyydeksi $P(A_1 A_2 \cdots A_n)$. □

Kutsumme kaavaa (3.6.3) tapahtumien yhdisteen *yleiseksi tulokaavaksi.* Jos A_1, A_2, …, A_n ovat keskenään riippumattomat, niin saadaan

$$P(A_1 A_2 \cdots A_n) = P(A_1)\, P(A_2) \cdots P(A_n).$$

Oletetaan, että satunnaismuuttujien X_1, X_2, …, X_n, … arvoalueet S_i ovat numeroituvia. Määritellään tapahtumat

$$A_i = \{X_i = x_i\}, \qquad i = 1, 2, \ldots,$$

missä $x_i \in S_i$. Silloin voimme kirjoittaa kertolaskukaavan (3.6.3) avulla

$$\begin{aligned}(3.6.4)\quad & P(X_1 = x_1,\, X_2 = x_2,\, \ldots,\, X_n = x_n) \\ &= P(X_1 = x_1)\, P(X_2 = x_2 \mid X_1 = x_1)\, P(X_3 = x_3 \mid X_1 = x_1, X_2 = x_2) \cdots \\ &\qquad \cdot P(X_n = x_n \mid X_1 = x_1,\, \ldots,\, X_{n-1} = x_{n-1}).\end{aligned}$$

$P(X_1 = x_1,\, X_2 = x_2,\, \ldots,\, X_n = x_n)$ on satunnaismuuttujien X_1, X_2, …, X_n *yhteistodennäköisyys,* joka on lausuttu peräkkäisten ehdollisten todennäköisyyksien avulla.

Esimerkki 3.19 (Syntymäpäiväongelma uudelleen) Olemme jo aikaisemmin implisiittisesti soveltaneet yleistä tulokaavaa (3.6.3). Tarkastellaan uudelleen Esimerkin 2.3 syntymäpäiväongelmaa. Kutsuilla on r henkilöä. Millä todennäköisyydellä ainakin kahdella henkilöllä on sama syntymäpäivä? Käytössämme on osanottajalista, johon syntymäpäivät on merkitty (karkausvuotta ei oteta huomioon). Käydään listaa läpi alusta lähtien järjestyksessä niin pitkälle, kunnes löydetään syntymäpäivä, joka jo oli listalla aikaisemmin. Silloin etsintä lopetetaan siihen ja todetaan, että ainakin kahdella vieraalla on sama syntymäpäivä. Jos lista päästään läpi löytämättä toistoa, kellään ei ole samaa syntymäpäivää.

Olkoon B_j tapahtuma, että tarkistus lopetetaan j. vieraaseen, koska hänen kohdallaan huomataan 1. toistuva syntymäpäivä. Olkoon A_j tapahtuma, että j:llä ensimmäisellä on eri syntymäpäivä. Silloin

$$A_r^c = B_2 \cup B_3 \cup \cdots \cup B_r$$

on tapahtuma, että ainakin kahdella on sama syntymäpäivä. Koska tapahtumat B_2, B_3, …, B_r ovat toisensa poissulkevat, niin

$$P(A_r^c) = P(B_2) + P(B_3) + \cdots + P(B_r),$$

missä

$$P(A_r^c) = 1 - P(A_r).$$

Lasketaan kysytty todennäköisyys $P(A_r^c)$ todennäköisyyden $P(A_r)$ avulla.

Kuvataan tarkistusprosessi toistokokeena:

$$\begin{tikzcd}
B_2 & B_3 & B_4 & B_5 & & B_r & \\
A_1 \arrow[u, "\frac{1}{365}"] \arrow[r, "\frac{364}{365}"] & A_2 \arrow[u, "\frac{2}{365}"] \arrow[r, "\frac{363}{365}"] & A_3 \arrow[u, "\frac{3}{365}"] \arrow[r, "\frac{362}{365}"] & A_4 \arrow[u, "\frac{4}{365}"] & \cdots & A_{r-1} \arrow[u, "\frac{r-1}{365}"] \arrow[r, "\frac{365-(r-1)}{365}"] & A_r
\end{tikzcd}$$

Jotta tarkistusprosessi menee koko listan läpi, sattuu tapahtuma A_r, eli kaikilla vierailla on eri syntymäpäivä. Sitä ennen ovat sattuneet A_2, A_3, ..., A_{r-1}. Esimerkiksi A_2 on tapahtuma, että tarkistusprosessi ei pysähdy 2. vieraaseen, vaan hänellä on eri syntymäpäivä kuin 1. vieraalla. Todennäköisyys

$$P(A_2) = \frac{364}{365} = 1 - \frac{1}{365} = 1 - P(B_2).$$

koska valittavana on 364 päivää, jotka poikkeavat 1. vieraan syntymäpäiväpäivästä. Jos j:n ensimmäisen syntymäpäivän joukossa ei ole samoja, niin ei myöskään i:n ensimmäisen, jos $i < j$, jolloin $A_i \subset A_j$. Tästä seuraa, että $A_2 A_3 \cdots A_j = A_j$ ja

$$P(A_{j+1} \mid A_2 A_3 \cdots A_j) = P(A_{j+1} \mid A_j) = \frac{365 - j}{365} = 1 - \frac{j}{365}.$$

Soveltamalla tapahtumien yhdisteen tulokaavaa saadaan

$$\begin{aligned} P(A_r) &= P(A_2 A_3 A_4 \cdots A_r) \\ &= P(A_2)\, P(A_3 \mid A_2)\, P(A_4 \mid A_2 A_3) \cdots P(A_r \mid A_2 \cdots A_{r-1}) \\ &= P(A_2)\, P(A_3 \mid A_2)\, P(A_4 \mid A_3) \cdots P(A_r \mid A_{r-1}) \\ &= \frac{364}{365} \cdot \frac{362}{365} \cdot \frac{362}{365} \cdots \frac{365 - r + 1}{365} = \frac{365^{(r)}}{365^r}. \end{aligned}$$

□

3.7 Bayesin lause

Pastori Thomas Bayesin (1763) mukaan nimetty lause seuraa suoraan ehdollisen todennäköisyyden määritelmästä. Bayesilainen lähestymistapa tilastotieteeseen perustuu tähän lauseeseen. Olkoot H_1, H_2, ..., H_k sellaiset tapahtumat, että

$$H_i H_j = \emptyset \quad (i \neq j) \qquad \text{ja} \qquad \sum_{i=1}^{k} H_i = \Omega.$$

Nyt siis tapahtumajoukko H_1, H_2, ..., H_k muodostaa otosavaruuden Ω osituksen. Tämä tarkoittaa sitä, että yksi ja vain yksi tapahtumista H_1, H_2, ..., H_k sattuu, kun tehdään satunnaiskoe $\mathcal{E}$, jonka otosavaruus on Ω. Oletamme lisäksi, että $P(H_i) > 0$ kaikilla $i = 1, 2, \ldots, k$.

Lause 3.16 *Olkoon*

$$\Omega = \sum_i H_i$$

jokin otosavaruuden ositus. Silloin minkä tahansa tapahtuman $T \subset \Omega$ todennäköisyys voidaan lausua muodossa

$$P(T) = \sum_i P(H_i)\, P(T \mid H_i). \tag{3.7.1}$$

Todistus. Joukko-opin sääntöjen nojalla saadaan

$$T = \Omega T = \Big(\sum_i H_i\Big) T = \sum_i H_i T,$$

josta todennäköisyyden P additiivisuuden (Määritelmä 2.5) perusteella seuraa kaava

$$P(T) = P\Big(\sum_i H_i T\Big) = \sum_i P(H_i T).$$

Kun kaavaan sijoitetaan

$$P(H_i T) = P(H_i)\, P(T \mid H_i),$$

saadaan (3.7.1). □

Jos kaavassa (3.7.1) jokin $P(H_i) = 0$, vastaava summan termi on 0, vaikka $P(T \mid H_i)$ ei olekaan määritelty. Kaavaa (3.7.1) kutsutaan *kokonaistodennäköisyyden kaavaksi.*

Olkoot X ja Y kokonaislukuarvoiset satunnaismuuttujat ja k jokin kokonaisluku. Soveltamalla kaavaa (3.7.1) tapahtumiin

$$H_i = \{X = i\}, \qquad T = \{Y = k\}$$

saadaan

$$P(Y = k) = \sum_i P(X = i)\, P(Y = k \mid X = i), \tag{3.7.2}$$

missä summa käy yli kaikkien kokonaislukujen. Jos $P(X = i) = 0$, niin vastaava yhteenlaskettava summassa on 0. Kaava on helppo yleistää mille tahansa satunnaismuuttujalle X, jonka arvojoukko S_X on numeroituva. Y voi olla jokin yleisempi satunnaismuuttuja, ei välttämättä kokonaislukuarvoinen, ja tapahtuma $T = \{Y = k\}$ voidaan korvata vaikkapa tapahtumalla $T = \{Y > a\}$, $a \in \mathbb{R}$.

Esimerkki 3.20 (Mikä laatikko?) Meillä on 3 samanlaista laatikkoa. Laatikossa i on i valkoista palloa ja yksi musta, $i = 1, 2, 3$. Tilanne on siis oheisen kuvion kaltainen

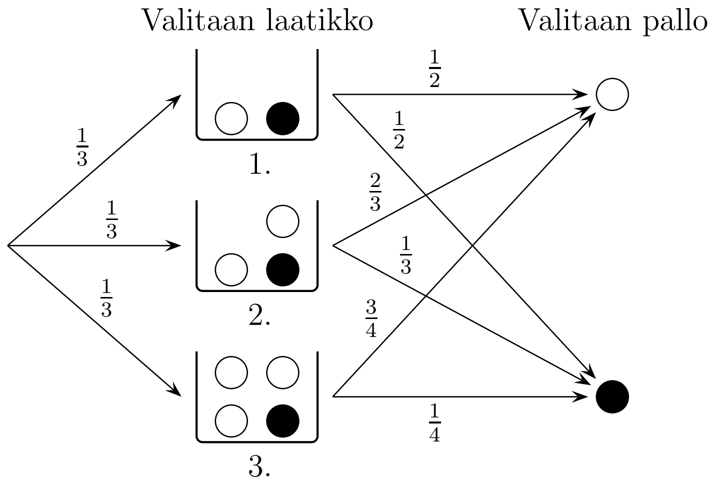
Laatikko valitaan harhattoman nopan heitolla. Jos silmäluku on k, valitaan laatikko, jonka numero $i = \lceil \frac{k}{2} \rceil$ on $\frac{k}{2}$ pyöristettynä lähimpään (suurempaan) kokonaislukuun. Jos esimerkiksi $k = 3$, niin $\lceil \frac{k}{2} \rceil = 2$ ja valitaan Laatikko

2. Kun valitun pallon väri on tiedossa, arvataan, mistä laatikosta pallo on valittu.

Mikä on arvauksesi, jos valittu pallo on valkoinen? Tuntuisi järkevältä arvata Laatikko 3, koska siellä on suhteellisesti eniten valkoisia. Olkoon $H_i = \{\text{Pallo Laatikosta } i\}$, $T = \{\text{Pallo valkoinen}\}$. Arvion varmentamiseksi lasketaan todennäköisyydet

$$P(H_i \mid T) = \frac{P(H_iT)}{P(T)}, \qquad i = 1, 2, 3. \tag{3.7.3}$$

Seuraavassa kuviossa on esitetty havainnollisesti tilanteeseen liittyvät todennäköisyydet.



Kaavassa (3.7.3) osoittaja on

$$P(H_iT) = P(H_i)\,P(T \mid H_i) = \frac{1}{3} \cdot \frac{i}{i+1}, \qquad i = 1, 2, 3.$$

Koska $\sum_{i=1}^{3} H_iT = T$ ja T_1, T_2 ja T_3 muodostavat T:n osituksen, niin yhteenlaskulauseen perusteella

$$P(T) = \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{2}{3} + \frac{1}{3} \cdot \frac{3}{4} = \frac{23}{36}.$$

Kaavasta (3.7.3) saadaan

$$P(H_i \mid T) = \frac{\frac{1}{3} \cdot \frac{i}{i+1}}{\frac{23}{36}} = \frac{12}{23} \cdot \frac{i}{i+1}, \qquad i = 1, 2, 3.$$

Jos veikkaat Laatikkoa 3, todennäköisyys osua oikeaan on $\frac{9}{23}$. Laatikolla 1 vastaava todennäköisyys on $\frac{6}{23}$ ja Laatikolla 2 se on $\frac{8}{23}$. Intuitiivisesti oikealta tuntunut Laatikon 3 valinta on siis paras arvaus. □

Väittämä 3.2 (Bayesin lause) *Olkoon $H_1, H_2, \ldots, H_k$ otosavaruuden Ω ositus ja T sellainen tapahtuma, että $P(T) > 0$. Silloin*

$$P(H_i \mid T) = \frac{P(H_i)\,P(T \mid H_i)}{\sum_{j=1}^{k} P(H_j)\,P(T \mid H_j)}. \tag{3.7.4}$$

Todistus. Todennäköisyyksien tulokaavan nojalla sadaan

$$P(H_i T) = P(H_i)\, P(T \mid H_i) = P(T)\, P(H_i \mid T),$$

mistä seuraa

$$P(H_i \mid T) = \frac{P(H_i)\, P(T \mid H_i)}{P(T)}.$$

Väittämän 3.16 mukaan $P(T) = \sum_j P(H_j)\, P(T \mid H_j)$, joten kaava (3.7.4) on todistettu. $\square$

Kaavaa (3.7.4) kutsutaan Bayesin säännöksi. Tapahtumat $H_1, H_2, \dots, H_k$ voidaan usein ajatella *hypoteeseiksi*, joista täsmälleen yksi on tosi. T on taas jokin tunnettu tieto satunnaiskokeen tuloksesta: tiedämme, että tapahtuma T on sattunut. Todennäköisyydet $P(H_i)$, $i = 1, 2, \dots, k$ ovat hypoteeseja koskevia ns. *prioritodennäköisyyksiä,* jotka voivat kuvastaa uskoa tai luottamusta kyseisiin hypoteeseihin. Ehdollista todennäköisyyttä $P(H_i \mid T)$ kutsutaan hypoteesin H_i *posterioritodennäköisyydeksi* tai posterioriluottamukseksi hypoteesiin H_i. Tapahtuman T todennäköisyys $P(T \mid H_i)$ ehdolla, että hypoteesi H_i on tosi, on tapahtuman T *uskottavuus* (likelihood) ehdolla H_i.

3.7.1 Peräkkäisotanta

Populaatiossa on N henkilöä, joista Np $(0 \le p \le 1)$ henkilöä kannattaa puoluetta B ja loput $N - Np$ eivät kannata B:tä (ts. kannattavat jotain muuta puoluetta, eivät kannata mitään puoluetta, eivät ota kantaa yms.). Haluamme estimoida kannattajien suhteellisen osuuden p, joka on tuntematon parametri. Haastattelija kysyy n:n satunnaisesti valitun henkilön mielipiteen (otanta palauttamatta). Määritellään

$$X_i = \begin{cases} 1, & \text{jos } i. \text{ haastateltava kannattaa } B\text{:tä;} \\ 0 & \text{muutoin,} \end{cases}$$

missä $1 \le i \le n$ ja $1 \le n \le N$. Tarkastellaan siis satunnaismuuttujien jonoa $\{X_1, X_2, \dots, X_n\}$ tai lyhyesti $\{\, X_i \mid 1 \le i \le n \,\}$. Tällaista jonoa kutsutaan *stokastiseksi prosessiksi,* mikä on satunnaismuuttujien perheestä käytetty nimitys.

Merkitään nyt $A_i = \{X_i = 1\}$ ja $A_i^c = \{X_i = 0\}$. Silloin kokonaistodennäköisyyden kaavan mukaan

$$P(A_2) = P(A_1)\, P(A_2 \mid A_1) + P(A_1^c)\, P(A_2 \mid A_1^c). \tag{3.7.5}$$

Helposti nähdään, että

$$P(A_1) = \frac{Np}{N} = p \qquad \text{ja} \qquad P(A_1^c) = \frac{N - Np}{N} = 1 - p.$$

Toisaalta

$$P(A_2 \mid A_1) = \frac{Np-1}{N-1} \qquad \text{ja} \qquad P(A_2 \mid A_1^c) = \frac{Np}{N-1},$$

koska 1. haastatellun jälkeen jäljellä on $N-1$ haastateltavaa, joiden joukossa on $Np-1$ B:n kannattajaa, jos 1. haastateltava oli B:n kannattaja. Jos 1. haastateltava ei ollut B:n kannattaja, niin jäljellä on vielä Np B:n kannattajaa. Kun nämä todennäköisyydet sijoitetaan kaavaan (3.7.5), saadaan

$$P(A_2) = p\frac{Np-1}{N-1} + (1-p)\frac{Np}{N-1} = p.$$

Näin olemme osoittaneet, että $P(A_1) = P(A_2)$. Mutta tämä tulos pitää paikkansa yleisesti:

(3.7.6) $$P(A_i) = p, \qquad i = 1, 2, \ldots, n; \quad 1 \le n \le N.$$

Näytämme nyt, että tämä yleinen tulos pitää paikkansa. Voimme ajatella, että B:n kannattajat on numeroitu $1, 2, \ldots, Np$ ja muut $Np+1, Np+2, \ldots, N$. Kysymyksessä on otanta palauttamatta, kun järjestys otetaan huomioon. Tarkastellaan tapahtumaa A_{i+1}, että $(i+1)$. haastateltava on B:n kannattaja. Kaikkien $(i+1)$:n kokoisten järjestettyjen jonojen (otosten) lukumäärä on $N^{(i+1)}$. Sellaisia jonoja, joissa $(i+1)$. alkio on 1 (B:n kannattaja) on $Np(N-1)^{(i)}$ kappaletta, koska B:n kannattaja voidaan valita Np tavalla ja loput i otosalkiota $(N-1)^{(i)}$ tavalla. Tuloperiaatteen mukaan suotuisia otoksia on siis $Np(N-1)^{(i)}$ kappaletta. Tästä seuraa, että

$$\begin{aligned} P(A_{i+1}) &= \frac{Np(N-1)^{(i)}}{N^{(i+1)}} = \frac{pN(N-1)\cdots(N-1-i+1)}{N^{(i+1)}} \\ &= \frac{pN^{(i+1)}}{N^{(i+1)}} = p. \end{aligned}$$

Olemme näin todistaneet tuloksen (3.7.6)

Määritellään nyt satunnaismuuttuja

$$X = X_1 + X_2 + \cdots + X_n,$$

joka on B:n kannattajien lukumäärä otoksessa. Tiedämme aikaisempien tarkastelujen perusteella, että X noudattaa hypergeometrista jakaumaa $\mathrm{HGeo}(n, N, p)$. Johdimme Esimerkissä 3.11 hypergeometrisen jakauman odotusarvon. Nyt tämä odotusarvo on helppo laskea satunnaismuuttujan X avulla, koska

$$\begin{aligned} E(X) &= E(X_1) + E(X_2) + \cdots + E(X_n) \\ &= p + p + \cdots + p = np, \end{aligned}$$

koska

$$E(X_i) = 1 \cdot p + 0 \cdot (1-p) = p, \qquad i = 1, 2, \ldots, n.$$

Jos satunnaismuuttuja X/n valitaan p:n estimaattoriksi, voimme todeta, että

$$E\left(\frac{X}{n}\right) = \frac{1}{n}\,E(X) = \frac{1}{n} \cdot np = p.$$

Sanomme, että X/n on *harhaton estimaattori*.

3.8 Usean tapahtuman unionin todennäköisyys

Lauseessa 2.5 esitettiin kolmen tapahtuman A_1, A_2 ja A_3 unionin todennäköisyyden $P(A_1 \cup A_2 \cup A_3)$ lauseke. Yleistetään nyt tämä tulos n:n tapahtuman $A_1, A_2, \ldots, A_n$ unionin tapaukseen.

Lause 3.17 *Samassa otosavaruudessa määriteltyjen tapahtumien A_1, A_2, $\ldots$, A_n unionin $\bigcup_{i=1}^n A_i$ todennäköisyys on*

$$P\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{i=1}^{n} P(A_i) - \sum_{j>i}^{n} P(A_iA_j) + \sum_{k>j>i}^{n} P(A_iA_jA_k) - + \cdots + (-1)^{n-1} P(A_1A_2\cdots A_n). \tag{3.8.1}$$

Todistus. Olkoon $\alpha_i = I_{A_i}$ tapahtuman A_i indikaattorifunktio, eli

$$\alpha_i(\omega) = \begin{cases} 1, & \text{kun } \omega \in A_i \\ 0, & \text{kun } \omega \in A_i^c. \end{cases}$$

Silloin tapahtuman $A_1^cA_2^c\cdots A_n^c$ indikaattorifunktio on $\prod_{i=1}^n (1-\alpha_i)$. Koska $\bigcup_{i=1}^n A_i = (A_1^cA_2^c\cdots A_n^c)^c$, niin sen indikaattorifunktio on

$$\begin{aligned} I_{\bigcup A_i} &= 1 - \prod_{i=1}^{n}(1-\alpha_i) \\ &= \sum_{i=1}^{n} \alpha_i - \sum_{j>i}^{n} \alpha_i\alpha_j + \sum_{k>j>i}^{n} \alpha_i\alpha_k\alpha_k \\ &\qquad - + \cdots + (-1)^{n-1}\alpha_1\alpha_2\cdots\alpha_n. \end{aligned} \tag{3.8.2}$$

Kun nyt yhtälössä (3.8.2) otetaan odotusarvot puolittain ja käytetään hyväksi odotusarvon lineaarisuutta, saadaan tulos (3.8.1). Huomaa, että indikaattorifunktion I_A odotusarvo $E(I_A) = P(A)$ on vastaavan tapahtuman todennäköisyys. Silloin $E(I_{\bigcup A_i}) = P\big(\bigcup_{i=1}^n A_i\big)$, $E(\alpha_i) = P(A_i)$, $E(\alpha_i\alpha_j) = P(A_iA_j)$, $\ldots$, $E(\alpha_1\alpha_2\cdots\alpha_n) = P(A_1A_2\cdots A_n)$. □

Esimerkki 3.21 (Yhteensopivuusongelma) Meillä on kaksi n:n kortin korttipakkaa, joiden kortit on numeroitu juoksevasti 1:stä n:ään. Asetetaan 1. pakan kortit pöydälle riviin numerojärjestyksessä $1, 2, \ldots, n$. Sekoitetaan 2. pakka ja asetetaan kortit riviin pöydälle saadussa satunnaisjärjestyksessä. Mikä on todennäköisyys, että i. kortin numero on i? Silloin molemmissa riveissä i. kortti on i eli on saatu i-pari. Mikä on todennäköisyys, että saadaan ainakin yksi pari?

Ratkaisu. Olkoon A_i tapahtuma, että saadaan i-pari. Pakan 2 kortit voidaan asettaa $n!$ erilaiseen järjestykseen. Jos numero i kiinnitetään i. paikalle, niin loput kortit voidaan asettaa $(n-1)!$ erilaiseen järjestykseen, joten

$$P(A_i) = \frac{(n-1)!}{n!} = \frac{1}{n}. \tag{3.8.3}$$

Jos kiinnitetään i-pari ja j-pari ($i \neq j$), niin loput $(n-2)$ korttia voidaan permutoida $(n-2)!$ tavalla. Silloin

$$P(A_i A_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}. \tag{3.8.4}$$

Vastaavalla tavalla voidaan laskea todennäköisyys, että saadaan i-pari, j-pari ja k-pari ($i \neq j \neq k$):

$$P(A_i A_j A_k) = \frac{(n-3)!}{n!} = \frac{1}{n(n-1)(n-2)} \tag{3.8.5}$$

ja yleisesti

$$P(A_{i_1} A_{i_2} \dots A_{i_m}) = \frac{(n-m)!}{n!} = \frac{1}{n(n-1)\cdots(n-m+1)}, \qquad 1 \leq m \leq n.$$

Todennäköisyys, että saadaan ainakin yksi pari on siis Lauseen 3.17 mukaan

$$\begin{aligned} P\Big(\bigcup_{i=1}^{n} A_i\Big) &= \binom{n}{1}\frac{1}{n} - \binom{n}{2}\frac{1}{n(n-1)} + \binom{n}{3}\frac{1}{n(n-1)(n-2)} \\ &\qquad - + \cdots + (-1)^{n-1}\frac{1}{n!} \\ &= 1 - \frac{1}{2!} + \frac{1}{3!} - + \cdots + (-1)^{n-1}\frac{1}{n!}. \end{aligned}$$

Huomaa, että

$$1 - \frac{1}{2!} + \frac{1}{3!} - + \cdots + (-1)^{n-1}\frac{1}{n!} + \cdots = \sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{i!} = 1 - \mathrm{e}^{-1} = 0.632\dots$$

Kun siis n on suuri, niin

$$P\Big(\bigcup_{i=1}^{n} A_i\Big) \approx 1 - \mathrm{e}^{-1} = 0.632\dots$$

Suurilla n:n arvoilla todennäköisyys saada ainakin yksi pari on hyvin lähellä lukua 0.632. □

Satunnaismuuttujat, ehdollistaminen ja riippumattomuus: Yhteenveto

Todennäköisyys

- Ehdollinen todennäköisyys

$$P(B \mid A) = \frac{P(AB)}{P(A)}, \qquad P(A) \neq 0.$$

- Tulosääntö $\quad P(AB) = P(A)\, P(B \mid A)$.
- Yleinen tulokaava

$$P(A_1A_2A_3 \cdots A_{n-1}A_n) = P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1A_2) \cdots \\ \cdot P(A_n \mid A_1A_2 \cdots A_{n-1}).$$

- Riippumattomuus. A ja B ovat riippumattomat, jos $P(AB) = P(A)\, P(B)$.
- $P(A_1$ tai A_2 tai $A_3)$

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1A_2) \\ - P(A_1A_3) - P(A_2A_3) + P(A_1A_2A_3).$$

Satunnaismuuttujat

- Odotusarvo

$$E(X) = \sum_{\omega \in \Omega} X(\omega)\, P(\{\omega\}),$$
$$E(X) = \sum_{x_i \in S} x_i\, P(X = x_i),$$

  missä S on X:n arvojoukko.

  $E(X)$ on todennäköisyyksillä painotettu X:n arvojen keskiarvo.

- Odotusarvon lineaarisuus

$$E(X+Y) = E(X) + E(Y) \quad \text{ja} \quad E(cX) = c\,E(X), \text{ missä } c \text{ on vakio.}$$

- Varianssi

$$\mathrm{Var}(X) = E(X-\mu)^2 = E(X^2) - \mu^2, \qquad \mu = E(X).$$

- Lineaarinen muunnos $cX + b$

$$E(cX+b) = c\,E(X) + b, \quad \mathrm{Var}(cX+b) = c^2\,\mathrm{Var}(X), \quad b \text{ ja } c \text{ vakioita.}$$

- Cauchyn ja Schwarzin epäyhtälö

$$[E(XY)]^2 \le E(X^2)\,E(Y^2).$$

- Kovarianssi

$$\mathrm{Cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = E(XY) - \mu_X\mu_Y,$$

  missä $\mu_X = E(X)$ ja $\mu_Y = E(Y)$.

- Summat

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y),$$
$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum\sum_{i\neq j} \mathrm{Cov}(X_i, X_j).$$

- Identtiset jakaumat. Diskreeteillä satunnaismuuttujilla X ja Y on sama jakauma, jos niillä on sama arvoalue S ja kaikilla $v \in S$

$$P(X=v) = P(Y=v).$$

- Samat satunnaismuuttujat. X ja Y ovat identtiset, jos $X(\omega) = Y(\omega)$ kaikilla $\omega \in \Omega$. Jos $P(X=Y)=1$, niin $X=Y$ (X ja Y diskreettejä).

- Riippumattomuus. X ja Y ovat riippumattomat, jos

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

  kaikilla $A \subset S_X$ ja $B \subset S_Y$.

  Jos X ja Y ovat riippumattomat, niin

  1) $g(X)$ ja $h(Y)$ ovat riippumattomat,
  2) $E(XY) = E(X)\,E(Y)$,
  3) $\mathrm{Cov}(X,Y) = 0$,
  4) $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

- Markovin epäyhtälö

$$P(X \ge a) \le \frac{E(X)}{a}, \qquad \text{missä } X \ge 0 \text{ ja } a > 0.$$

- Tšebyševin epäyhtälö

$$P(|X-\mu| \ge \varepsilon) \le \frac{\sigma^2}{\varepsilon^2},$$

  missä $\varepsilon > 0$, $\mu = E(X)$, $\sigma^2 = \mathrm{Var}(X)$.

- Otoskeskiarvo

$$\overline{X}_n = \frac{1}{n}(X_1 + X_2 + \cdots + X_n),$$
$$E(\overline{X}_n) = \mu \qquad \text{ja} \qquad \mathrm{Var}(\overline{X}_n) = \frac{\sigma^2}{n},$$

jos $E(X_i) = \mu$ ja $\mathrm{Var}(X_i) = \sigma^2$, $i = 1, 2, \ldots, n$.

- Suurten lukujen laki (heikko): $P(|\overline{X}_n - \mu| \geq \varepsilon) \to 0$, kun $n \to \infty$ ja X_1, X_2, …, X_n ovat riippumattomat ja noudattavat samaa jakaumaa.

Generoivat funktiot ja momentit

- Satunnaismuuttujan momentit

X:n r. momentti	$\alpha_r = E(X^r)$,
r. keskusmomentti	$\mu_r = E(X - \mu)^r$,
r. tekijämomentti	$g_r = E[X^{(r)}] = E[X(X-1)\cdots(X-r+1)]$.

- Momenttifunktio

$$M_X(t) = E(\mathrm{e}^{tX}); \qquad t \in (-a, a),\ a > 0.$$

- Summan $Z = X + Y$ momenttifunktio $M_Z(t) = M_X(t)M_Y(t)$, jos X ja Y ovat riippumattomat.

- r. momentti. Momenttifunktion r. derivaatta pisteessä $t = 0$ on r. momentti: $M(0)^{(r)} = E(X^r)$.

- Todennäköisyydet generoiva funktio

$$G(t) = E(t^X), \qquad X \text{ on diskreetti.}$$

- r. tekijämomentti. G:n r. derivaatta pisteessä $t = 1$ on X:n r. tekijämomentti: $G^{(r)}(1) = E[X^{(r)}]$.

- $G(t)$ vs. $M(t)$: $G(\mathrm{e}^t) = E(\mathrm{e}^{tX}) = M(t)$.

Bayesin lause

- Kokonaistodennäköisyys

$$P(T) = \sum_{i=1}^{k} P(H_i)\, P(T \mid H_i),$$

missä $T \subset \Omega$ ja H_1, H_2, …, H_k on Ω:n ositus.

- Bayesin kaava
$$P(H_i \mid T) = \frac{P(H_i)\,P(T \mid H_i)}{\sum_{j=1}^{k} P(H_j)\,P(T \mid H_j)}.$$
- Prioritodennäköisyydet $P(H_i)$.
- Posterioritodennäköisyydet $P(H_i \mid T)$, $i = 1, 2, \ldots, n$.
- Uskottavuus. $P(T \mid H_i)$ on tapahtuman T uskottavuus ehdolla, että H_i on tosi.

Harjoituksia

1. Oletetaan, että $P(X = 0) = 1 - P(X = 1)$ ja $E(X) = 3\,\mathrm{Var}(X)$. Laske $P(X = 0)$.

2. Olkoon satunnaismuuttujan X todennäköisyysfunktio
$$f(x) = \frac{(|x| + 1)^2}{9}, \qquad x = -1, 0, 1.$$
Laske $E(X)$, $E(X^2)$ ja $E(3X^2 - 2X + 4)$.

3. Olkoon $h(x) = (x - b)^2$, missä b ei ole X:n funktio. Millä b:n arvolla odotusarvo $E[(X - b)^2]$ saavuttaa miniminsä, kun oletetaan, että odotusarvo on olemassa. (Vihje: Tarkastele funktiota $g(b) = E[(X - b)^2] = E(X^2) - 2b\,E(X) + b^2$.)

4. Olkoon $\Omega = \{\omega_1, \omega_2, \omega_3\}$ ja $P(\omega_1) = P(\omega_2) = P(\omega_3) = \frac{1}{3}$. Määritellään satunnaismuuttujat X, Y ja Z seuraavasti:
$$\begin{array}{lll} X(\omega_1) = 1, & X(\omega_2) = 2, & X(\omega_3) = 3, \\ Y(\omega_1) = 2, & Y(\omega_2) = 3, & Y(\omega_3) = 1, \\ Z(\omega_1) = 3, & Z(\omega_2) = 1, & Z(\omega_3) = 2. \end{array}$$

(a) Osoita, että satunnaismuuttujilla X, Y ja Z on sama todennäköisyysjakauma.

(b) Määritä satunnaismuuttujien $X + Y$, $Y + Z$, $X + Z$ ja

(c) satunnaismuuttujien $\sqrt{(X^2 + Y^2)Z}$ ja $Z/|X - Y|$ todennäköisyysjakauma.

5. Populaatiossa on M miestä ja N naista. Miehissä on m ja naisissa n tupakoitsijaa. Populaatiosta valitaan satunnaisesti yksi. A on tapahtuma, että valittu on mies ja B tapahtuma, että on valittu tupakoitsija. Mitkä ehdot lukumäärien M, N, m ja n on toteutettava, jotta A ja B ovat toisistaan riippumattomat?

6. Tarkastellaan Esimerkin 2.8 tilannetta, jossa Pekka ja Paavo pelaavat ”kruunua ja klaavaa” (satunnaiskävely, $n = 20$).

(a) Mikä on todennäköisyys, että Pekka on 5 heiton jälkeen voitolla yhden euron, 10 heiton jälkeen 2 euroa, 20 heiton jälkeen 2 euroa?

(b) Mikä on Pekan voiton odotusarvo 20 heiton sarjassa?

(c) Jos Pekka on 5. heiton jälkeen voitolla euron, mikä on Pekan voiton odotusarvo 20. heiton jälkeen?

7. Tarkastellaan Tehtävän 6 peliä simuloimalla ($n = 20$).

(a) Mikä on todennäköisin voittosumma? Epätodennäköisin voittosumma? Hahmottele voittosumman todennäköisyysjakauma.

(b) Kuinka usein Pekka on voitolla pelin aikana? Hahmottele tämän satunnaismuuttujan todennäköisyysjakauma.

8. Oletetaan, että $X \sim \mathrm{Tasd}(1, N)$ noudattaa diskreettiä tasajakaumaa.

(a) Jos $E(x) = 6$, niin mitä on $\mathrm{Var}(X)$?

(b) Olkoon $X \sim \mathrm{Tasd}(3, 8)$. Laske $E(X)$ ja $\mathrm{Var}(X)$.

9. Suuressa tehtaassa sattuu 5:n päivän jakson aikana 3 onnettomuutta. Oletetaan, että kaikki mahdolliset 5^3 erilaista 3:n onnettomuuden sijoittumista 5:n päivän jaksolle ovat yhtä todennäköisiä.

Olkoon $Y = \{\text{Onnettomuuspäivien lukumäärä jakson aikana}\}$ ja X niiden päivien lukumäärä, jolloin onnettomuuksia ei satu.

(a) Määritä satunnaismuuttujan $X = 5 - Y$ todennäköisyysfuntio.

(b) Laske $E(X)$ ja $\mathrm{Var}(X)$.

10. Lennolla Havannasta Helsinkiin laukkuni eivät olleet perillä Helsingissä samaan aikaan kuin minä. Laukkuja on reitillä siirretty koneesta toiseen 3 kertaa ja todennäköisyydet, että siirtoa ei ole tehty ajoissa tai oikein, ovat siirtojärjestyksessä 0.4, 0.2 ja 0.1. Mikä on todennäköisyys, että moka sattui jo ensimmäisessä siirrossa?

11. Olkoot X ja Y riippumattomat kokonaislukuarvoiset satunnaismuuttujat, joilla on sama todennäköisyysfunktio $f_X(n) = f_Y(n) = p_n$, $n \geq 1$. Laske todennäköisyydet $P(X = Y)$ ja $P(X \leq Y)$.

12. Tarkastellaan kaksilapsisia perheitä. Oletetaan pojat ja tytöt yhtä todennäköisiksi ja 2. lapsen sukupuoli on riippumaton 1. lapsen sukupuolesta.

Tarkastellaan neljää tapahtumaa:

$$\begin{aligned} A &= \text{1. lapsi on poika,} \\ B &= \text{lapset ovat eri sukupuolta,} \\ C &= \text{1. lapsi on tyttö,} \\ D &= \text{2. lapsi on poika.} \end{aligned}$$

(a) Mitkä tapahtumaparit $\{A, B\}$, $\{A, C\}$, $\{B, C\}$ ovat keskenään riippumattomat?

(b) Ovatko tapahtumat A, B ja D keskenään eli täydellisesti riippumattomat?

13. A, B ja C ampuvat maaliin 20 laukausta. Yhden laukauksen osumistodennäköisyys on A:lla 0.4, B:llä 0.3, C:llä 0.1 ja laukaukset ovat toisistaan riippumattomat. Olkoot X_A, X_B ja X_C vastaavasti A:n, B:n ja C:n osumien lukumäärät ja X osumien kokonaismäärä.

(a) Määrittele X riippumattomien satunnaismuuttujien summana ja laske sen avulla X:n odotusarvo ja varianssi.

(b) Määritä Tšebyševin epäyhtälön avulla väli, jolle osumien kokonaismäärä osuu vähintään todennäköisyydellä $\frac{8}{9}$.

14. Liukuhihnalta tulevat pullot ovat vikaantuneita, toisistaan riippumatta, todennäköisyydellä 0.2. Hihnalta tulevat pullot tarkistetaan, vikaantuneet poistetaan ja loput pakataan 12 pullon laatikoihin.

(a) Millä todennäköisyydellä on tutkittava täsmälleen 17 pulloa, kunnes laatikko saadaan täyteen?

(b) Ainakin 17 pulloa, kunnes laatikko saadaan täyteen?

15. Lääkärillä oli oheisessa taulukossa esitetty uuden hoidon vaikutusta koskeva potilasaineisto:

	Asuu kaupungissa		Asuu maaseudulla	
	Saanut hoidon	Ei hoitoa	Saanut hoidon	Ei hoitoa
Elossa	1000	50	95	5000
Kuollut	9000	950	5	5000

Tarkastellaan tapahtumia A = 'potilas elossa', B = 'saanut hoidon' ja C = 'asuu kaupungissa'. Estimoi tarvittavat todennäköisyydet taulukon frekvenssien avulla ja laske

(a) $P(A \mid B)$ ja $P(A \mid B^c)$ sekä

(b) $P(A \mid BC)$, $P(A \mid B^cC)$, $P(A \mid BC^c)$ ja $P(A \mid B^cC^c)$.

(c) Oliko hoidosta apua?

16. Olkoot X ja Y sellaiset satunnaismuuttujat, että $E(X) = \mu_X$, $E(Y) = \mu_Y$, $\mathrm{Var}(X) = \sigma_X^2$, $\mathrm{Var}(Y) = \sigma_Y^2$ ja $\rho = \mathrm{Cor}(X, Y)$. Käytetään satunnaismuuttujan Y arvioimiseen regressioennustetta $\hat{Y} = \alpha + \beta X$, missä α ja β ovat vakioita. Ennusteen keskineliövirhe määritellään

$$\mathrm{MSE}(\hat{Y}) = E\big([Y - (\alpha + \beta X)]^2\big).$$

(a) Osoita laskemalla, että

$$\mathrm{MSE}(\hat{Y}) = [\mu_Y - (\alpha + \beta\mu_X)]^2 + \mathrm{Var}(Y - \beta X).$$

(b) Valitse edellisessä $\mathrm{MSE}(\hat{Y})$:n lausekkeessa $\alpha = \mu_Y - \beta\mu_X$ ja näytä, että silloin

$$\mathrm{MSE}(\hat{Y}) = (\beta\sigma_X - \rho\sigma_Y)^2 + \sigma_Y^2(1 - \rho^2).$$

(c) Päättele nyt, että $\mathrm{MSE}(\hat{Y})$ saavuttaa miniminsä $\sigma_Y^2(1 - \rho^2)$, kun $\alpha = \mu_Y - \beta\mu_X$ ja $\beta = \rho\sigma_Y/\sigma_X$.

17. Olkoon X sellainen diskreetti satunnaimuuttuja, että sen todennäköisyysfunktio on $P(X = x_i) = p_i$, $i \geq 1$ ja 2. momentti $E(X^2) = \sum_i p_i x_i^2$ on olemassa. Olkoon $A = \{\, i \mid |x_i| \geq \varepsilon \,\}$, missä $\varepsilon > 0$.

(a) Osoita, että

$$P(|X| \geq \varepsilon) = \sum_{i \in A} p_i \text{ ja } E(X^2) \geq \sum_{i \in A} p_i x_i^2,$$

(b) $\sum_{i \in A} p_i x_i^2 \geq \sum_{i \in A} p_i \varepsilon^2$

(c) ja lopuksi $P(|X| \geq \varepsilon) \leq E(X^2)/\varepsilon^2$.