Sisältö

Luku 4

Satunnaismuuttujien tunnusluvut

Tässä luvussa käsitellään satunnaismuuttujien ominaisuuksia. Erityisesti satunnaismuuttujien odotusarvo on keskeinen käsite. Tarkastelujen painopiste on diskreetteissä satunnaismuuttujissa ja kaikkia vastaavia tuloksia ei toisteta jatkuvien satunnaismuuttujien tapauksessa. Tulosten todistaminen ja soveltaminen on yleensä huomattavasti yksinkertaisempaa diskreettien satunnaismuuttujien yhteydessä.

4.1 Odotusarvo, varianssi ja kovarianssi

4.1.1 Odotusarvo

Tarkastellaan ensin diskreettejä satunnaismuuttujia, joiden arvojoukko on äärellinen.

Esimerkki 4.1 Heitetään lanttia 3 kertaa. Olkoon satunnaismuuttuja X on 'klaavojen lukumäärä'. Merkitään R = 'kruuna' ja L = 'klaava'. Silloin otosavaruus $\Omega = \{\mathrm{RRR}, \mathrm{RRL}, \mathrm{RLR}, \mathrm{RLL}, \mathrm{LRR}, \mathrm{LRL}, \mathrm{LLR}, \mathrm{LLL}\}$ ja X:n arvojoukko on $S_X = \{0, 1, 2, 3\}$. Nyt esimerkiksi $X(\mathrm{RRL}) = X(\mathrm{RLR}) = 1$. Olkoon A_r tapahtuma "klaavojen lukumäärä r". Merkintä $(X = 2)$ tarkoittaa tapahtumaa $A_2 = \{\mathrm{LLR}, \mathrm{LRL}, \ \mathrm{RLL}\}$ ja $P(X = 2) = P(A_2) = 3/8$. Tapahtuman indikaattori (Määritelmä 2.5) määriteltiin 2. luvussa. Esimerkiksi tapahtuman A_2 indikaattori $I_{A_2}(\omega) = 1$, jos $\omega \in A_2$ eli silloin kun klaavojen lukumäärä on 2. Vastaavasti määritellään jokainen I_{A_r}, $r = 0, 1, 2, 3$. Satunnaismuuttuja "klaavojen lukumäärä kolmessa heitossa"voidaan kirjoittaa indikaattorien avulla seuraavasti:

$$X = 0\, I_{A_0} + 1\, I_{A_1} + 2\, I_{A_2} + 3\, I_{A_3} = 1\, I_{A_1} + 2\, I_{A_2} + 3\, I_{A_3}.$$

Nyt $A_r = \{X(\omega) = r\}$ ja $P(A_r) = P(X = r)$, $r = 0, 1, 2, 3$. Yhteys otosavaruuteen häipyy näkyvistä, kun on merkitty yksinkertaisesti $\{\omega |\, X(\omega) = r\} = \{X(\omega) = r\}$. □

Olkoon $X : \Omega \to S_X$ diskreetti satunnaismuuttuj, jonka arvojoukko $S_X = \{x_1, \ldots, x_n\}$. Merkitään $A_i = (X = x_i)$ tapahtumaa, että X saa arvon

$x_i,\ i = 1, \dots, n$. Satunnaismuuttuja X voidaan lausua indikaattorien I_{A_i} avulla muodossa

$$X = x_1 \, I_{A_1} + \cdots + x_n \, I_{A_n}. \tag{4.1.1}$$

Todennäköisyyslaskennan pioneerit tarkastelivat usein odotusarvoa

$$E(X) = (\text{panos } x) \times \, P(A).$$

Tässä satunnaismuuttuja $X = x$, kun A sattuu ja muutoin $X = 0$. Panoksen x voittaa siis todennäköisyydellä $P(A)$. Odotusarvo on panos $\times$ todennäköisyys. Huomattakoon, että indikaattorin I_A odotusarvo on

$$E(I_A) = 1 \times P(A) + 01 \times [1 - P(A)] = P(A)$$

on A:n todennäköisyys.

Määritelmä 4.1 (Odotusarvo) Olkoon X *diskreetti satunnaismuuttuja*, jonka arvojoukko on $S_X = \{x_1, \dots, x_n\}$ ja $p_i = P(X = x_i),\ i = 1, \dots, n$. Silloin X:n odotusarvo on

$$E(X) = \sum_{i=1}^{n} x_i p_i.$$

Jatkossa saatamme kutsua satunnaismuuttujan odotusarvoa myös satunnaismuuttujan keskiarvoksi. Huomattakoon, että satunnaismuuttujan indikaattoriesityksen (4.1.1) nojalla odotusarvo voidaan kirjoittaa muodossa

$$E(X) = \sum_{i=1}^{n} x_i \, E(I_{A_i}) = \sum_{i=1}^{n} x_i \, P(X = x_i). \tag{4.1.2}$$

Jos otosavaruus Ω on numeroituva, niin X:n odotusarvo voidaan kirjoittaa

$$\begin{aligned} E(X) &= \sum_{\omega \in \Omega} X(\omega) \, P(\{\omega\}) = \sum_{i=1}^{n} \sum_{\omega \in \{X = x_i\}} X(\omega) \, P(\{\omega\}) \\ &= \sum_{i=1}^{n} x_i p_i. \end{aligned} \tag{4.1.3}$$

Odotusarvon (4.1.2) esitystapa satunnaismuuttujan arvojen todennäköisyyksillä painotettuna keskiarvona on käyttökelpoinen. Jos X:n arvojoukko $S_X\{x_1, x_2, \dots\}$ on numeroituvasti ääretön, niin niin (4.1.3) voi olla ääretön tai odotusarvoa ei ole olemassa (sarja ei suppene).

Kaavasta (4.1.3) saadaan myös minkä tahansa satunnaismuuttujan X funktion $h(X)$ odotusarvo. Koska $h(X)$ on satunnaismuuttuja, niin

$$E[h(X)] = \sum_{i=1}^{n} h(x_i) p_i.$$

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

$$E(X^r) = \sum_i p_i x_i^r. \tag{4.1.4}$$

Jos X on *jatkuva satunnaismuuttuja*, jonka tiheysfunktio on f, niin X:n odotusarvo on

$$E(X) = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x, \text{ mikäli integraali on olemassa.}$$

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 4.1 *Oletetaan, että otosavaruudessa Ω määritellyillä diskreeteillä satunnaismuuttujilla X ja Y on odotusarvo, X:n arvojoukko on $S_X = \{x_1, \ldots, x_n\}$ ja Y:n arvojoukko $S_Y = \{y_1, \ldots, y_m\}$, $P(X = x_i) = p_i, P(Y = y_j) = q_j$ ja $P(X = x_i, Y = y_j) = p_{ij}$ sekä $a \in \mathbb{R}$ 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. *$E[h(X)] = \sum_{i=1}^{n} h(x_i) p_i = \sum_{i=1}^{n} h(x_i) p_i$*

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

$$\begin{aligned} E(aX) &= \sum_{i=1}^{n} a x_i\, P(aX = a x_i) = a \sum_{i=1}^{n} x_i\, P(aX = a x_i) \\ &= a \sum_{i=1}^{n} x_i\, P(X = x_i) = a\, E(X). \end{aligned}$$

Identiteetti $P(aX = a x_i) = P(X = x_i)$ pitää paikkansa kaikilla $a \neq 0$, koska $\{\, \omega \mid aX(\omega) = a x_i \,\} = \{\, \omega \mid X(\omega) = x_i \,\}$. 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).

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

$$\begin{aligned}
E(X+Y) &= \sum_i \sum_j (x_i + y_j)\, P(X = x_i,\, Y = y_j) = \sum_i \sum_j (x_i + y_j) p_{ij} \\
&= \sum_i \sum_j (x_i p_{ij} + y_j p_{ij}) \\
&= \sum_i \sum_j x_i p_{ij} + \sum_i \sum_j y_j p_{ij} \\
&= \sum_i x_i \sum_j p_{ij} + \sum_j y_j \sum_i p_{ij} \\
&= \sum_i x_i p_i + \sum_j y_j q_j \\
&= E(X) + E(Y).
\end{aligned}$$

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

$$E[h_1(X) - h_2(X)] = \sum_i [h_1(x_i) - h_2(x_i)] p_i \geq 0,$$

koska $h_1(x_i) - h_2(x_i) \geq 0$ ja $p_i \geq 0$ kaikilla $i = 1, \ldots, n$. 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 4.2 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 4.1 (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 4.3 Indikaattorifunktio (Määritelmä 2.5) 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 4.1, 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 4.4 Satunnaismuuttuja X noudattaa diskreettiä tasajakaumaa Tasd$(1, N)$, kun $P(X = i) = \frac{1}{N}$, $i = 1, 2, \ldots, N$ (ks. alaluku 2.17). 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 4.5 Hypergeometrinen jakauma esiteltiin tarkasteltaessa otantaa palauttamatta (alaluku 2.7.1). 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 $X \sim \mathrm{HGeo}(n, N, p)$. Hypergeometrisen jakauman todennäköisyysfunktio on

$$P(X = x; N, n, p) = \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}}, \qquad x = 0, 1, \ldots, n, \tag{4.1.5}$$

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 (4.1.5):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 (2.4.6) 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

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

□

Esimerkki 4.6 Alaluvussa 3.3.3 tarkasteltiin peräkkäisotantaa äärellisestä populaatiosta. 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.). 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 $1 \le n \le N$.

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 HGeo(n, N, p). Johdimme Esimerkissä 4.5 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.* □

4.1.2 Ehdollinen 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

$$E(X \mid A) = \sum_x x f(x \mid A). \tag{4.1.6}$$

Esimerkki 4.7 Oletetaan, että $X \sim \text{Tasd}(1, N)$ 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 4.2 *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) $P(A_i) > 0$, $i \geq 1$.

4.1.3 Varianssi

Varianssin laskemiseksi tarvitaan funktion $h(X) = X^2$ odotusarvo (Vertaa Lauseen 4.1 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 4.1 *Oletetaan, että satunnaismuuttujilla X ja Y on 2. momentti ja $c \in \mathbb{R}$ on vakio. Silloin odotusarvot*

$$E[(cX)^2], \quad E[(X+Y)^2], \quad E(X), \quad E(Y) \quad \textit{ja} \quad E(XY) \tag{4.1.7}$$

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 $0 \le (X+Y)^2 = 2(X^2+Y^2)-(X-Y)^2 \le 2(X^2+Y^2)$ ja oletuksen mukaan $E(X^2+Y^2) = E(X^2)+E(X^2)$ on olemassa, niin Lauseen 4.1 (kohta 3) mukaan $E[(X+Y)^2]$ on olemassa.

3. Koska $0 \le (|X|-|Y|)^2 = |X|^2 + |Y|^2 - 2|X||Y|$, niin niin Lauseen 4.1 (kohta 3) mukaan

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

joten $E(XY)$ on olemassa.

$\square$

Lause 4.3 (Cauchyn ja Schwarzin epäyhtälö) *Jos satunnaismuuttujilla X ja Y on 2. momentti, niin*

$$[E(XY)]^2 \le E(X^2)\, E(Y^2). \tag{4.1.8}$$

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 4.1 mukaan myös $E(XY)$ on olemassa. Merkitään nyt $c = E(XY)/E(X^2)$. Silloin

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

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 (4.1.8) pitää triviaalisti paikkansa. $\square$

Yhtäsuuruus (4.1.8):ssä vallitsee silloin, kun $aX = -bY$ (todennäköisyydellä 1). Silloin $Y = -\frac{a}{b}X$, jos $b \neq 0$. Epäyhtälössä (4.1.8) pätee siis yhtäsuuruus, kun X ja Y ovat lineaarisesti riippuvia. Epäyhtälö (4.1.8) voidaan lausua myös muodossa

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

Määritelmä 4.2 (Varianssi) Jos satunnaismuuttujalla X on 2. momentti $E(X^2)$, niin sillä on odotusarvo μ_X ja X:n varianssi on

$$\sigma_X^2 = \mathrm{Var}(X) = E[(X-\mu_X)^2]. \tag{4.1.9}$$

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

$$\sigma^2 = \mathrm{Var}(X) = E(X^2) - \mu^2 = E(X^2) - [E(X)]^2. \tag{4.1.10}$$

satunnaismuuttujan X hajonta $\sigma_X = \sqrt{\mathrm{Var}(X)}$. Odotusarvon määritelmästä ja identiteetistä (4.1.10) saamme erittäin käyttökelpoisen tuloksen:

$$\mathrm{Var}(cX) = c^2\,\mathrm{Var}(X), \qquad E(X^2) = \mu^2 + \mathrm{Var}(X). \tag{4.1.11}$$

Esimerkki 4.8 Lasketaan diskreettiä tasajakaumaa $\mathrm{Tasd}(1, N)$ noudattavan satunnaismuuttujan varianssi. Esimerkin 4.4 mukaan

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

Soveltamalla kaavaa (4.1.10) saadaan

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

□

4.1.4 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 4.1 nojalla.

Määritelmä 4.3 (Kovarianssi) Satunnaismuuttujien X ja Y *kovarianssi* σ_{XY} määritellään odotusarvona

$$\begin{aligned} \sigma_{XY} = \mathrm{Cov}(X,Y) &= E[(X-\mu_X)(Y-\mu_Y)] \\ &= E(XY) - \mu_X\mu_Y. \end{aligned} \tag{4.1.12}$$

Kovarianssin avulla voidaan sitten määritellä korrelaatiokerroin.

Määritelmä 4.4 (Korrelaatiokerroin) Satunnaismuuttujien X ja Y *korrelaatiokerroin*

$$\rho_{XY} = \mathrm{Cor}(X,Y) = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}. \tag{4.1.13}$$

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

1. $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)$.

2. Jos satunnaismuuttujalla $X_1, X_2, \ldots, X_n$ on varianssi, niin

$$\mathrm{Var}\Bigl(\sum_{i=1}^{n} X_i\Bigr) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}(X_i, X_j) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{i=1}^{n}\sum_{j\neq i}^{n} \mathrm{Cov}(X_i, X_j). \tag{4.1.14}$$

Todistus. Todistetaan 1. kohta. Määritelmän mukaan

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

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

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

Kaava (4.1.14) voidaan todistaa induktiolla. □

4.2 Satunnaismuuttujan funktio

Lauseen 4.1 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 4.9 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:

$$h^{-1}(A) = \{\, x \in S_X \mid h(x) \in A \,\}. \tag{4.2.1}$$

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 $h^{-1}(y) = x$.

4.3 Satunnaismuuttujien identtisyys

Määritelmä 4.5 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 4.10 Esimerkissä 2.10 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 4.4 *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 4.11 Heitetään harhatonta lanttia 4 kertaa. Olkoon kruunun todennäköisyys p. X ja Y on määritelty samoin kuin Esimerkissä 4.10. 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 $p = \frac{1}{2}$. □

4.4 Satunnaismuuttujien riippumattomuus

Määrittelimme tapahtumien riippumattomuuden alaluvussa 3.1.2. Tarkastelemme nyt satunnaismuuttujien riippumattomuutta.

4.4.1 Kaksi satunnaismuuttujaa

Määritelmä 4.6 (Satunnaismuuttujien riippumattomuus) Satunnaismuuttujat X ja Y ovat riippumattomat jos

$$P(X \in A, Y \in B) = P(X \in A)\, P(Y \in B) \tag{4.4.1}$$

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 $B \subset \mathbb{R}$.

Jos X ja Y ovat diskreettejä, niin riippumattomuuden määritelmän nojalla

$$P(X = x, Y = y) = P(X = x)\, P(Y = y) = f_X(x) f_Y(y) \tag{4.4.2}$$

kaikilla $x, y \in \mathbb{R}$, missä $f_X(x)$ on X:n ja $f_Y(y)$ on Y:n todennäköisyysfunktio. Diskreettien satunnaismuuttujien X ja Y yhteisjakauman todennäköisyysfunktio määritellään:

$$P(X = x, Y = y) = f_{X,Y}(x, y) \tag{4.4.3}$$

kaikilla $x, y \in \mathbb{R}$. Huomattakoon, että $f_{X,Y}(x, y) > 0$ täsmälleen silloin, kun $(x, y) \in S_X \times S_Y$ ja muutoin $f_{X,Y}(x, y) = 0$. Diskreetit satunnaismuuttujat X ja Y ovat riippumattomat silloin ja vain silloin kun

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) \tag{4.4.4}$$

kaikilla $x, y \in \mathbb{R}$.

Lause 4.5 *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 4.6 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 (4.4.1) 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 (4.4.4) voidaan myös pitää diskreettien satunnaismuuttujien X ja Y riippumattomuuden määritelmänä, sillä siitä seuraa identiteetti (4.4.1). Jos valitaan kaksi mielivaltaista numeroituvaa joukkoa $A \subset \mathbb{R}$ ja $B \subset \mathbb{R}$ sekä oletetaan (4.4.4), saadaan

$$\begin{aligned} P(X \in A, Y \in B) &= \sum_{x_i \in A} \sum_{y_j \in B} P(X = x_i, Y = y_j) \\ &= \sum_{x_i \in A} \sum_{y_j \in B} P(X = x_i)\, P(Y = y_j) \qquad [(4.4.4)] \\ &= \sum_{x_i \in A} P(X = x_i) \sum_{y_j \in B} P(Y = y_j) \\ &= P(X \in A)\, P(Y \in B). \end{aligned}$$

Näin olemme todenneet, että ehdot (4.4.1) ja (4.4.4) 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

$$P(I_A = x, I_B = y) = P(I_A = x)\, P(I_B = y) \tag{4.4.5}$$

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

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

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.

4.4.2 Useita satunnaismuuttujia

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

$$(4.4.6)\quad \begin{aligned} &P(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n) \\ &\qquad = P(X_1 \in A_1)\, P(X_2 \in A_2) \cdots P(X_n \in A_n) \end{aligned}$$

kaikilla (sopivasti valituilla) joukoilla $A_i \subset \mathbb{R}$, $1 \le i \le n$. Jos $X_1, \ldots, X_n$ ovat diskreettejä, niin (4.4.6) 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 (4.4.6) mukaan

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

4.5 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 4.6 (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. (2.5)]. 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 $X(\omega) \geq a$ 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 4.1, 3. kohta) ja lineaarisuuden (1. kohta) nojalla saadaan

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

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

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

Lause 4.7 (Tšebyševin epäyhtälö) *Olkoon X satunnaismuuttuja, jonka keskiarvo on μ ja varianssi σ^2. Silloin*

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}, \qquad \textit{kaikilla } \varepsilon > 0. \tag{4.5.1}$$

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 (4.5.1) suoraan Markovin epäyhtälöstä. □

Lause 4.8 (Riippumattomat satunnaismuuttujat, tulon odotusarvo) *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*

$$E(X_1X_2\cdots X_n) = E(X_1)\,E(X_2)\cdots E(X_n).$$

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

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

ja

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

Todistus. Jos $i \neq j$, niin

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

koska X_i:n ja X_j:n riippumattomuuden nojalla $E(X_i X_j) = E(X_i)\, E(X_j) = 0$. Summan varianssin $\mathrm{Var}(\sum_{i=1}^{n} X_i)$ lauseke seuraa nyt suoraan Apulauseesta 4.2. □

Apulause 4.4 (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

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

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

Lause 4.9 (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 $\varepsilon > 0$,

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

Todistus. Apulauseen 4.4 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.

4.6 Generoivat funktiot ja momentit

4.6.1 Momentit

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

Määritelmä 4.7 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

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

4.6.2 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 4.8 Olkoon X satunnaismuuttuja, jonka tiheysfunktio on $f(x)$. Reaalimuuttujan t funktio

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

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

$$E(\mathrm{e}^{tX}) = \begin{cases} \sum_i \mathrm{e}^{tx_i} f(x_i) & \text{diskreetti satunnaismuuttuja} \\ \int_{-\infty}^{\infty} \mathrm{e}^{tx} f(x)\,\mathrm{d}x, & \text{jatkuva satunnaismuuttuja} \end{cases}$$

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

Määritelmän perusteella on selvää, että

$$M(0) = \sum_i f(x_i) = 1, \text{kun } X \text{ diskreetti ja } M(0) = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1,$$

kun X on jatkuva. 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 4.10 *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 4.12 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 4.11 *Olkoot X ja Y riippumattomat satunnaismuuttujat, joiden momenttifunktiot ovat $M_X(t)$ ja $M_Y(t)$. Silloin satunnaismuuttujan $Z = X + Y$ momenttifunktio on*

$$M_Z(t) = M_X(t) M_Y(t). \tag{4.6.1}$$

Todistus. Koska e^{tX} on pelkästään x:n (X:n arvojen) funktio ja e^{tY} pelkästään y:n funktio, niin Lauseen 4.5 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 4.8. □

Usean satunnaismuuttujan tapauksessa on voimassa vastaava tulos.

Seuraus 4.1 *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

$$M(t) = \sum_{x \in S} \mathrm{e}^{tx} f(x) \tag{4.6.2}$$

derivoidaan puolittain, voidaan oikea puoli derivoida termeittäin ja yhtäsuuruus säilyy. Derivoimalla lauseke (4.6.2) puolittain muuttujan t suhteen saadaan

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

ja jokaisella positiivisella kokonaisluvulla r

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

Sijoittamalla $t = 0$ saadaan

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

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 4.12 *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 4.13 (Lévyn jatkuvuuslause) *Olkoon $X_1, X_2, \ldots$ jono satunnaismuuttujia, joiden kertymäfunktiot ovat $F_{X_1}, F_{X_2}, \ldots$ ja vastaavasti momenttifunktiot $M_{X_1}(t), M_{X_2}(t), \ldots$. Olkoon X satunnaismuuttuja, jonka kertymäfunktio on F_X ja momenttifunktio $M_X(t)$. Jos n:n kasvaessa rajatta*

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

kaikilla t:n arvoilla jossain nollan ympäristössä $(-h, h)$, $h > 0$, niin silloin

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

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

Satunnaismuuttujien momenttifunktioiden suppenemisesta seuraa siis satunnaismuuttujien kertymäfunktioiden suppeneminen. Tällöin sanomme, että satunnaismuuttujat $X_1, X_2, \ldots$ suppenevat jakaumamielessä kohti satunnaismuuttujaa X.

4.6.3 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

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

4.7 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

$$P(\omega_i, \omega_j) = p_i q_j \tag{4.7.1}$$

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 $\sum_{\omega_i \in \Omega_1} p_i = \sum_{\omega_j \in \Omega_2} q_j = 1$, niin

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

Identiteetti (4.7.1) 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

$$X = \begin{cases} 1, & \text{todennäköisyydellä } P(\mathrm{O}) = p; \\ 0, & \text{todennäköisyydellä } 1 - p, \end{cases} \tag{4.7.2}$$

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 4.13 Esimerkissä 2.10 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 $S\times S\times S = \{\, (s_1, s_2, s_3) \mid s_i \in S \,\} = \{111, 110, 101, 100, 011, 010, 001, 000\}$.
□

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 4.4 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.

Satunnaismuuttujien tunnusluvut: Yhteenveto

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

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

- 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

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

- 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

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

- Tšebyševin epäyhtälö

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

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

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:

$$\begin{aligned} X(\omega_1) &= 1, & X(\omega_2) &= 2, & X(\omega_3) &= 3,\\ Y(\omega_1) &= 2, & Y(\omega_2) &= 3, & Y(\omega_3) &= 1,\\ Z(\omega_1) &= 3, & Z(\omega_2) &= 1, & Z(\omega_3) &= 2. \end{aligned}$$

(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. Tarkastellaan Esimerkin 2.13 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?

6. Tarkastellaan Tehtävän 5 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.

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

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

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

8. 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 $\text{Var}(X)$.

9. 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 $P(X \leq Y)$.

10. 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 $\frac{8}{9}$.

11. 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 $\beta = \rho\sigma_Y/\sigma_X$.

12. 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 $P(|X| \geq \varepsilon) \leq E(X^2)/\varepsilon^2$.

13. Osoita satunnaismuuttujille X, Y ja Z, että

(a) $\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$.

(b) $\mathrm{Cov}(X, Y + Z) = \mathrm{Cov}(X, Y) + \mathrm{Cov}(X, Z)$.

14. Jos $X \perp\!\!\!\perp Y$ (riippumattomat), niin osoita, että kaikilla Z

$$\mathrm{Cov}(X, Y + Z) = \mathrm{Cov}(X, Z).$$

15. Osoita, että

$$\mathrm{Var}(X - Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\,\mathrm{Cov}(X, Y).$$

16. Jos $X \perp\!\!\!\perp Y$, niin osoita että

$$\operatorname{Cov}(X, XY) = E(Y)\operatorname{Var}(X).$$

17. Jos $X \perp\!\!\!\perp Y$, niin osoita että

$$\operatorname{Var}(XY) = \sigma_1^2\sigma_2^2 + \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2,$$

Missä $E(X) = \mu_1,\ E(Y) = \mu_2,\ \operatorname{Var}(X) = \sigma_1^2$ ja $\operatorname{Var}(Y) = \sigma_2^2$.