4 Odotusarvo

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

4.1 Diskreetin satunnaismuuttujan odotusarvo

Määritelmä

Jos X on diskreetti sm, jonka arvojoukko on $S = \{x_1, x_2, \dots\}$ ja ptnf on f, niin sen odotusarvo on reaaliluku

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

mikäli sarja on itseisesti summautuva, eli mikäli

$$\sum_i |x_i| \, f(x_i) < \infty.$$

Muussa tapauksessa sanotaan että X:llä ei ole odotusarvoa. Se seikka, että X:llä on odotusarvo voidaan ilmaista myös sanomalla, että X on integroituva sm.

Vähän sarjaoppia

- Äärellinen summa suppenee tietenkin itseisesti; itseinen suppeneminen on relevantti ehto silloin, kun summattavia arvoja on äärettömän monta.
- Jos sarja summautuu itseisesti, niin sille saadaan sama äärellinen summa kaikilla summausjärjestyksillä.
- Toinen tapaus, jossa sarjan summausjärjestyksellä ei ole väliä on se, jossa kaikki summan termit ovat ei-negatiivisia; tällöin sarjan summa voi olla jokin reaaliluku tai ∞.
- Jos summa suppenee kohti jotakin reaalilukua, mutta ei suppene itseisesti, niin erilaisilla summausjärjestyksillä sarjalle voidaan saada erilaisia summia.
- Itseisen suppenemisen takia odotusarvon määritelmässä ei tarvitse kiinnittää tiettyä summausjärjestystä.

Huomautuksia odotusarvosta

- Jos satunnaismuuttujilla X ja Y on sama diskreetti jakauma, niin niillä on sama odotusarvo. Diskreetin jakauman odotusarvo määritellään lukuna EX, jossa X on mikä tahansa kyseistä jakaumaa noudattava sm.
- Odotusarvolla on fysikaalinen tulkinta massajakauman painopisteenä.
- Jos massattoman sauvan (=reaaliakseli) pisteisiin x_i asetetaan massat p_i, niin sauva pysyy tasapainossa, mikäli sitä tuetaan altapäin pisteestä EX. Tämä odotusarvon tulkinta säilyy myös muunlaisille jakaumille.

Odotusarvon kaavan motivointia

- Olkoon diskreetti sm X sinun voittamasi rahasumma tietyn uhkapelin yhdessä toistossa.
- Jos toistat peliä riippumattomasti N kertaa, niin voitat summan x_i osapuilleen $NP(X = x_i)$ kertaa.
- Keskimääräinen voittosi yhtä peliä kohti on

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

Odotusarvon frekvenssitulkinta ja suurten lukujen laki

- Odotusarvon frekvenssitulkinnan mukaan EX on osapuilleen sama kuin keskiarvo suuresta määrästä $X_1, \dots, X_N$ keskenään riippumattomia ja X:n kanssa samoin jakautuneita satunnaismuuttujia:

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

- Kaavan oikealla puolella esintyy satunnaismuuttuja, joka suppenee ns. vahvan suurten lukujen lain perusteella melkein kaikilla ω kohti odotusarvoa EX, kun N kasvaa rajatta. Vahva suurten lukujen laki on voimassa, mikäli X:n odotusarvo on olemassa (yksi todennäköisyysteorian kuuluisimpia tuloksia).
- Odotusarvoa arvioidaan tietokonesimuloinneissa usein näin, eli laskemalla keskiarvo satunnaismuuttujien simuloiduista arvoista.

Esimerkki: vakion odotusarvo

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

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

Esimerkki: tapahtuman indikaattorin odotusarvo

- Olkoon A tapahtuma, ja asetetaan $p = P(A)$.
- Jos $X = 1_A$, niin $X \sim \mathrm{Bernoulli}(p)$, ja

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

- Tapahtuman indikaattorin odotusarvo on sama kuin kyseisen tapahtuman todennäköisyys.

Esimerkki: binomijakauman odotusarvo

- Olkoon $X \sim \mathrm{Bin}(n, p)$, ja asetetaan $q = 1 - p$.

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

- Tässä laskussa käytettiin suoraan odotusarvon määritelmää. Johdamme saman tuloksen myöhemmin helpommalla tavalla.

Esimerkki: diskreetti jakauma, jolla ei ole odotusarvoa

- Tunnetusti sarja $\sum_{x=1}^{\infty} 1/x^2$ suppenee, ja sarja $\sum_{x=1}^{\infty} 1/x$ hajaantuu.
- Olkoon $c = \sum_{x=1}^{\infty} 1/x^2$.
- Määritellään sm X siten, että sen arvojoukko on kokonaisluvut poislukien nolla, ja ptnf on

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

- Tällöin X:llä ei ole odotusarvoa.

Johdattelua seuraavaan lauseeseen

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

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

- Mikäli X:llä on odotusarvo, niin se saadaan summasta

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

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

Apulause

Lause

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

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

niin

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

Odotusarvon ominaisuuksia

Lause

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

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

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

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

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

(e) (Lineaarisuus) Jos $a, b \in \mathbb{R}$, niin $E(aX + bY) = a\,EX + b\,EY$.

(f) Jos $X \perp\!\!\!\perp Y$, niin satunnaismuuttujalla XY on odotusarvo, ja

$$E(XY) = E(X)\,E(Y).$$

(g) EX riippuu vain X:n jakaumasta: jos X:llä ja Y:llä on sama jakauma, niin $EX = EY$.

Lineaarisuus n:lle satunnaismuuttujalle

- Jos X_1, X_2 ja X_3 ovat integroituvia sm:ia, niin

$$\begin{aligned} E(X_1 + X_2 + X_3) &= E[(X_1 + X_2) + X_3] = E(X_1 + X_2) + EX_3 \\ &= EX_1 + EX_2 + EX_3. \end{aligned}$$

- Tämä ominaisuus yleistyy mille tahansa äärelliselle lukumäärälle integroituvia satunnaismuuttujia, ts.

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

mikäli kaikki odotusarvot EX_i ovat reaalilukuja.

Esimerkki: binomijakauman odotusarvo helpommin

- Jos $X \sim \mathrm{Bin}(n, p)$, niin se voidaan esittää summana

$$X = \sum_{i=1}^{n} Y_i,$$

jossa $Y_i \sim \mathrm{Bernoulli}(p)$ ja Y_i:t ovat riippumattomia.

- Odotusarvon lineaarisuuden nojalla

$$EX = E\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} EY_i = np.$$

- Tässä laskussa ei tarvittu satunnaismuuttujien Y_i riippumattomuutta.

4.2 Jatkuvasti jakautuneen satunnaismuuttujan odotusarvo

Määritelmä

Jos X on jatkuvasti jakautunut sm, ja sen tiheysfunktio on f, niin sen odotusarvo on luku

$$EX = E(X) = \int_{-\infty}^{\infty} x\, f(x)\,\mathrm{d}x,$$

mikäli kyseinen integraali suppenee itseisesti, eli mikäli

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

Muussa tapauksessa sanotaan että X:llä ei ole odotusarvoa. Se seikka, että X:llä on odotusarvo voidaan ilmaista myös sanomalla, että X on integroituva sm.

Esimerkkejä

- *Tasajakauman odotusarvo*: jos $X \sim \mathrm{U}(a, b)$, niin (helppo lasku) $EX = \frac{1}{2}(a+b)$.
- *Cauchyn jakaumalla ei ole odotusarvoa.* Sm X noudattaa Cauchyn jakaumaa, jos sillä on tiheys

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}, \qquad x \in \mathbb{R}\,.$$

Tämä on tf, sillä $f \geq 0$ ja

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \frac{2}{\pi}\int_0^{\infty}\frac{1}{1+x^2}\,\mathrm{d}x = \frac{2}{\pi}\ \Big/_{0}^{\infty} \arctan x = \frac{2}{\pi}\frac{\pi}{2} = 1$$

X:llä ei ole odotusarvoa, sillä

$$\int_{-\infty}^{\infty} |x|\, f(x)\,\mathrm{d}x = \frac{1}{\pi}\int_0^{\infty}\frac{2x}{1+x^2}\,\mathrm{d}x = \frac{1}{\pi}\lim_{M\to\infty}\ \Big/_{0}^{M} \ln(1+x^2) = \infty.$$

4.3 Odotusarvon ominaisuuksia

- Diskreettien ja jatkuvien jakaumien lisäksi on olemassa myös muunlaisia jakaumia.
- Todennäköisyysteoriassa odotusarvo EX määritellään perusjoukon Ω yli laskettuna abstraktina Lebesguen integraalina

$$EX = \int X \, dP = \int_{\Omega} X(\omega) \, P(\mathrm{d}\omega)$$

mikäli integrandi on itseisesti integroituva.
- Emme varsinaisesti käytä tätä määritelmää, mutta teemme joitakin sitä koskevia huomautuksia.

Satunnaismuuttujan odotusarvo todennäköisysteoriassa

- Todennäköisyysteoriassa määritellään ensin ei-negatiivisen sm:n $Y \geq 0$ integraali EY tietyn rajankäynnin kautta. Tulokseksi saadaan joko jokin ei-negatiivinen reaaliluku (jolloin sanotaan, että Y on integroituva) tai ∞.
- Merkiltään rajoittamattoman sm:n X odotusarvo määritellään jakamalla se ensin *positiiviseen osaan* $X^+ \geq 0$ sekä *negatiiviseen osaan* $X^- \geq 0$ seuraavasti

$$X^+ = \max(X, 0), \quad X^- = \max(-X, 0). \tag{2}$$

Tällöin

$$X = X^+ - X^-, \quad \text{ja} \quad |X| = X^+ + X^-. \tag{3}$$

Merkiltään rajoittamattoman sm:n odotusarvo

- Mikäli sekä EX^+ ja EX^- ovat äärellisiä, niin EX määritellään kaavalla
$$EX = EX^+ - EX^-. \tag{4}$$
Tällöin sanotaan, että X on *integroituva*.
- Joissakin tarkasteluissa odotusarvolle sallitaan reaalilukujen lisäksi arvot $+\infty$ tai $-\infty$, ts. odotusarvo saa olla laajennettu reaaliluku (engl. *extended real number*).
- Jos vain toinen integraaleista EX^+ ja EX^- on äärellinen, niin EX voidaan määritellä käyttämällä kaavaa $EX = EX^+ - EX^-$ sekä laskusääntöjä
$$\infty - a = \infty, \quad a - \infty = -\infty, \qquad a \in \mathbb{R}\,.$$
- Huomaa, että lausekkeelle $\infty - \infty$ ei voida määritellä arvoa.

Sanontatapoja

- Mikäli X:n odotusarvo on määriteltävissä laajennettujen reaalilukujen avulla kaavalla $EX = EX^+ - EX^-$, sanomme, että odotusarvo on olemassa laajennettuna reaalilukuna tai että se on olemassa laajennetussa mielessä.
- Jos sanomme, että odotusarvo on olemassa tai että satunnaismuuttuja on integroituva, niin tämä tarkoittaa sitä, että odotusarvo EX on reaaliluku. Tällöin $E|X| < \infty$.
- Edellä nähtiin esimerkit sekä diskreetistä että jatkuvasti jakautuneesta satunnaismuuttujasta, joilla kummallakaan ei ole olemassa odotusarvoa edes tässä laajennetussa mielessä.

Odotusarvon ominaisuuksia (yleinen tapaus)

Lause

Olkoot X ja Y integroituvia satunnaismuuttujia (siis: odotusarvot $EX, EY \in \mathbb{R}$).

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

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

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

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

(e) (Lineaarisuus) Jos $a, b \in \mathbb{R}$, niin $E(aX + bY) = a\, EX + b\, EY$.

(f) Jos $X \perp\!\!\!\perp Y$, niin satunnaismuuttujalla XY on odotusarvo, ja

$$E(XY) = E(X)\, E(Y).$$

(g) EX riippuu vain X:n jakaumasta: jos X:llä ja Y:llä on sama jakauma, niin $EX = EY$.

Tärkeä epäyhtälö

Lause

X on integroituva jos ja vain jos $|X|$ on integroituva. Tällöin $|EX| \leq E|X|$.

Todistus. Jos X on integroituva, niin $EX^+ \in \mathbb{R}$ ja $EX^- \in \mathbb{R}$, joten

$$E|X| = E(X^+ + X^-) = EX^+ + EX^- \in \mathbb{R}\,.$$

Toisaalta, jos $E|X| < \infty$, niin välttämättä

$$0 \leq EX^+ < \infty, \quad \text{ja} \quad 0 \leq EX^- < \infty,$$

joten $EX = EX^+ - EX^- \in \mathbb{R}$. Jos X on integroituva, niin kolmioepäyhtälön nojalla

$$|EX| = |EX^+ - EX^-| \leq EX^+ + EX^- = E|X|.$$

4.4 Muunnoksen odotusarvo

- Olkoon X sm, ja $g : \mathbb{R} \to \mathbb{R}$ funktio.
- Määritellään sm $Y = g(X) = g \circ X$.
- Yksi tapa laskea Y:n odotusarvo olisi ensin johtaa sm:n Y jakauma. Mikäli Y:n jakauma on joko diskreetti tai jatkuva, voidaan soveltaa tuttuja kaavoja.
- Seuraava lause tarjoaa erilaisen tavan.

Satunnaismuuttujan muunnoksen odotusarvo

Lause

Olkoon X sm, $g : \mathbb{R} \to \mathbb{R}$ ja $Y = g(X)$.

(a) Jos X:llä on diskreetti jakauma ptnf:lla f_X, niin

$$EY = Eg(X) = \sum_x g(x)\, f_X(x),$$

mikäli summa suppenee itseisesti, eli mikäli $\sum_x |g(x)| f_X(x) < \infty$

(b) Jos X:llä on jatkuva jakauma tf:lla f_X, niin

$$EY = Eg(X) = \int_{-\infty}^{\infty} g(x)\, f_X(x)\, \mathrm{d}x,$$

mikäli integraali suppenee itseisesti, eli mikäli $\int_{-\infty}^{\infty} |g(x)|\, f_X(x)\, \mathrm{d}x < \infty$.

Tiedostamattoman tilastotieteilijän laki

- Mikäli kyseinen odotusarvo on olemassa, niin

$$Eg(X) = \begin{cases} \sum_x g(x)\, f_X(x) & \text{jos } X\text{:llä on diskreetti jakauma} \\ \int_{-\infty}^{\infty} g(x)\, f_X(x)\,\mathrm{d}x & \text{jos } X\text{:llä on jatkuva jakauma} \end{cases} \tag{5}$$

- Myös EX voidaan laskea tällä kaavalla (valitsemalla funktioksi g identiteettifunktio, jolle $g(x) = x$).
- Englanninkielisissä oppikirjoissa kaavalle käytetään usein nimitystä law of the unconscious statistician, koska tilastotieteilijät toisinaan käyttävät sitä huomaamatta, että kyseessä ei ole määritelmä vaan että kyseinen kaava pitäisi jollakin tavalla johtaa.

Onko $Eg(X)$ yhtä kuin $g(EX)$?

- Aloittelijat luulevat usein, että $Eg(X)$ on aina yhtäsuuri kuin $g(EX)$.
- Tämä pitää paikkansa affiinille funktiolle $g(x) = a + bx$, mutta ei (yleisesti) muille funktioille.

Riippumattomien sm:ien muunnoksien tulon odotusarvo

Lause

Olkoot satunnaismuuttujat $X \perp\!\!\!\perp Y$, ja $g : \mathbb{R} \to \mathbb{R}$ ja $h : \mathbb{R} \to \mathbb{R}$ sellaisia funktioita, että $g(X)$ ja $h(Y)$ ovat integroituvia sm:ia. Tällöin

$$E[g(X)\,h(Y)] = E[g(X)]\,E[h(Y)].$$

Todistus. Koska $g(X) \perp\!\!\!\perp h(Y)$ lauseen 3.7 nojalla, niin tämä lause seuraa riippumattomien satunnaismuuttujien tulon kaavasta.

4.5 Momentit

Määritelmä

Satunnaismuuttujan X

- k:s origomomentti eli k:s momentti μ'_k on

$$\mu'_k = EX^k, \qquad \text{kun } k = 1, 2, \ldots$$

- k:s keskusmomentti (engl. *central moment*) μ_k on

$$\mu_k = E[(X - EX)^k], \qquad \text{kun } k = 1, 2, \ldots$$

- kertaluvun r absoluuttinen momentti (engl. *absolute moment*) on

$$E|X|^r, \qquad r > 0$$

mikäli kyseiset odotusarvot ovat olemassa. Muuten sanotaan, että kyseinen momentti ei ole olemassa.

Huomautuksia momenteista

- Jos $a < 0$ ja $r \neq 0$, niin $|a|^r$ on hyvin määritelty, mutta a^r ei ole reaalinen, ellei r ole kokonaisluku. Tämän takia rajoitutaan kokonaislukukertalukuihin.
- Odotusarvo EX on 1. momentti μ_1', ja 1. keskusmomentti $\mu_1 = 0$.
- Keskusmomentti μ_k voidaan esittää origomomenttien $\mu_j', j \leq k$ avulla kehittämällä lauseke $(X - EX)^k$ binomikaavalla.
- Toinen keskusmomentti μ_2 eli varianssi on sovelluksissa tärkeä.
- Korkeammista momenteista esim. kolmas keskusmomentti kuvaa jakauman vinoutta ja neljäs keskusmomentti sen huipukkuutta. Tätä korkeampia momentteja käytetään harvoin.

Milloin momentit ovat olemassa?

- Jos $E|X|^s < \infty$ jollekin $s > 0$, niin

$$E|X|^r < \infty, \qquad \text{kaikilla } 0 < r < s.$$

  Tällöin myös kaikki origomomentit sekä keskusmomentit kertalukua $1 \leq k \leq s$ ovat olemassa.
- *Todistus*: Jos $0 < r < s$ ja $E|X|^s < \infty$, niin

$$\begin{aligned} |X|^r &= |X|^r 1_{\{|X| \leq 1\}} + |X|^r 1_{\{|X| > 1\}} \\ &\leq 1 + |X|^s 1_{\{|X| > 1\}} \leq 1 + |X|^s, \end{aligned}$$

  minkä takia

$$E|X|^r \leq 1 + E|X|^s < \infty$$

4.6 Varianssi, keskihajonta ja kovarianssi

Määritelmä (Varianssi ja keskihajonta)

Satunnaismuuttujan X varianssi $\operatorname{var} X$ on sen 2. keskusmomentti, eli

$$\operatorname{var} X = \operatorname{var}(X) = E[(X - EX)^2],$$

mikäli kyseinen odotusarvo on hyvin määritelty. Tällöin X:n keskihajonta (engl. *standard deviation*) on varianssin $\operatorname{var} X$ (positiivinen) neliöjuuri. Usein X:n varianssia merkitään σ^2 (tai σ_X^2). Tällöin varianssin positiivinen neliöjuuri $\sigma \geq 0$ (tai $\sigma_X \geq 0$) on X:n keskihajonta.

Huomautuksia varianssista

- Varianssi voidaan määritellä vain integroituvalle satunnaismuuttujalle, jolle siis $EX \in \mathbb{R}$. Jos EX on äärellinen, mutta $E[(X - EX)^2] = \infty$, niin voidaan sanoa, että $\operatorname{var} X = \infty$.
- Varianssille $\operatorname{var} X$ käytetään yleisesti myös merkintää $D^2(X)$.
- Mikäli EX^2 on äärellinen, niin myös EX ja $\operatorname{var} X$ ovat äärellisiä, ja

$$\operatorname{var} X = EX^2 - (EX)^2. \tag{6}$$

- *Johto*: Jos $\mu = EX$, niin

$$E(X - \mu)^2 = E(X^2 - 2\,\mu\,X + \mu^2) = EX^2 - 2\,\mu^2 + \mu^2.$$

Varianssin ominaisuuksia

Lause

(i) Mikäli $\operatorname{var} X$ on olemassa, niin $\operatorname{var} X \geq 0$. Jos $\operatorname{var} X = 0$, niin X on vakio.

(ii) Jos $a, b \in \mathbb{R}$ ja $\operatorname{var} X$ on olemassa, niin $\operatorname{var}(aX + b) = a^2 \operatorname{var} X$.

Todistus kohdalle (i): taululla.
Todistus kohdalle (ii): $E(aX + b) = aEX + b$, joten

$$\operatorname{var}(aX + b) = E[(a(X - EX))^2] = a^2 E[(X - EX)^2].$$

Kovarianssi

Määritelmä

Satunnaismuuttujien X ja Y (välinen) kovarianssi $\mathrm{cov}(X, Y)$ on odotusarvo

$$\mathrm{cov}(X, Y) = E[(X - EX)(Y - EY)],$$

mikäli kyseinen odotusarvo on olemassa.

- Myöhemmin nähdään (Cauchyn–Schwarzin epäyhtälön avulla), että odotusarvo $\mathrm{cov}(X, Y)$ on äärellinen ainakin silloin, jos $EX^2 < \infty$ ja $EY^2 < \infty$.

Kovarianssin ominaisuuksia

- Mikäli kyseiset suureet ovat äärellisiä, niin
  - $\operatorname{cov}(Y, X) = \operatorname{cov}(X, Y)$,
  - $\operatorname{var} X = \operatorname{cov}(X, X)$,
- Kun kerrotaan sulut auki, ja lasketaan odotusarvo termi termiltä, saadaan

$$\begin{aligned}
\operatorname{cov}(X, Y) &= E[(X - EX)(Y - EY)] \\
&= E[XY - X\,(EY) - (EX)\,Y + (EX)\,(EY)] \\
&= E(XY) + (-1 - 1 + 1)\,(EX)\,(EY),
\end{aligned}$$

joten kovarianssi voidaan laskea myös kaavalla

$$\operatorname{cov}(X, Y) = E(XY) - (EX)(EY). \tag{7}$$

Kovarianssin bilineaarisuus

Lause

Kovarianssi on bilineaarinen, eli molempien argumenttiensa suhteen lineaarinen operaattori. Ts. jos $a_1, a_2 \in \mathbb{R}$ ovat vakioita ja X_1, X_2 ja Z ovat satunnaismuuttujia, niin

$$\begin{aligned} \operatorname{cov}(a_1X_1 + a_2X_2, Z) &= a_1 \operatorname{cov}(X_1, Z) + a_2 \operatorname{cov}(X_2, Z) \\ \operatorname{cov}(Z, a_1X_1 + a_2X_2) &= a_1 \operatorname{cov}(Z, X_1) + a_2 \operatorname{cov}(Z, X_2) \end{aligned}$$

Jos lisäksi $b_1, b_2 \in \mathbb{R}$ ja Y_1, Y_2 ovat satunnaismuuttujia, niin

$$\operatorname{cov}(a_1X_1 + a_2X_2, b_1Y_1 + b_2Y_2) = \sum_{i=1}^{2}\sum_{j=1}^{2} a_i\, b_j \operatorname{cov}(X_i, Y_j).$$

Nämä kaavat pitävät paikkansa, mikäli kyseiset odotusarvot ovat olemassa.

Esimerkki: lineaarikombinaation varianssin laskeminen kovarianssin bilineaarisuuden avulla

$$
\begin{aligned}
&\operatorname{var}(aX + bY) \\
&\quad = \operatorname{cov}(aX + bY, aX + bY) \\
&\quad = a \operatorname{cov}(X, aX + bY) + b \operatorname{cov}(Y, aX + bY) \\
&\quad = a^2 \operatorname{cov}(X, X) + ab \operatorname{cov}(X, Y) + ab \operatorname{cov}(Y, X) + b^2 \operatorname{cov}(Y, Y) \\
&\quad = a^2 \operatorname{var}(X) + 2ab \operatorname{cov}(X, Y) + b^2 \operatorname{var}(Y).
\end{aligned}
$$

Summan ja erotuksen varianssi

- Edellisen erikoistapauksena

$$\operatorname{var}(X+Y) = \operatorname{var} X + \operatorname{var} Y + 2\operatorname{cov}(X,Y) \tag{8}$$
$$\operatorname{var}(X-Y) = \operatorname{var} X + \operatorname{var} Y - 2\operatorname{cov}(X,Y). \tag{9}$$

- Erityisesti, jos $\operatorname{cov}(X,Y) = 0$, niin

$$\operatorname{var}(X+Y) = \operatorname{var}(X-Y) = \operatorname{var} X + \operatorname{var} Y.$$

Korreloimattomuus ja riippumattomuus

- Summan $X + Y$ varianssi on yhtä kuin muuttujien varianssien summa täsmälleen silloin, kun $\mathrm{cov}(X, Y) = 0$.
- Jos $\mathrm{cov}(X, Y) = 0$, niin sanotaan, että X ja Y eivät korreloi eli että X ja Y ovat korreloimattomia.
-
- Jos $X \perp\!\!\!\perp Y$, niin $\mathrm{cov}(X, Y) = 0$, sillä tällöin

$$\mathrm{cov}(X, Y) = E[(X - EX)(Y - EY)] = E(X - EX)\, E(Y - EY) = 0.$$

- Siis riippumattomuudesta seuraa korreloimattomuus.
- Korreloimattomat satunnaismuuttujat eivät välttämättä ole riippumattomia.

Esimerkki: korreloimattomat muuttujat, jotka eivät ole riippumattomia

- Olkoon sm:lla X jatkuva jakauma tf:lla f, ja olkoon f parillinen funktio. Oletetaan lisäksi, että EX^4 on äärellinen.
- Tf:n parillisuuden ansiosta $EX = E(X^3) = 0$, joten

$$\operatorname{cov}(X, X^2) = E(X \cdot X^2) - (EX)(EX^2) = 0.$$

- Siis X ja X^2 eivät korreloi, mutta ne eivät tietenkään ole riippumattomia satunnaismuuttujia.

n:n riippumattoman satunnaismuuttujan summan varianssi

- Helpolla laskulla nähdään, että jos $X_1, \ldots, X_n$ ovat riippumattomia ja niillä on kaikilla äärellinen varianssi, niin tällöin
$$\mathrm{var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathrm{var}(X_i)$$

Esimerkki: binomijakauman varianssi

- Jos $X \sim \mathrm{Bin}(n, p)$, niin se voidaan esittää summana $\sum_{i=1}^{n} Y_i$, jossa $Y_i \sim \mathrm{Bernoulli}(p)$ ja Y_i:t ovat riippumattomia.
- Nyt $EY_i = p$, joten

$$\mathrm{var}\, Y_i = E[(Y_i - p)^2] = (0-p)^2(1-p) + (1-p)^2 p = p(1-p).$$

- Riippumattomuuden nojalla

$$\mathrm{var}\, X = \sum_{i=1}^{n} \mathrm{var}\, Y_i = np(1-p).$$

4.7 Momenttiemäfunktio ja kumulanttiemäfunktio

Määritelmä (Momenttiemäfunktio)

Satunnaismuuttujan X (jakauman) momenttiemäfunktio eli momenttigeneroivafunktio eli momentit generoiva funktio (engl. *moment generating function, mgf*) on funktio

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

joka on (reaalifunktiona) olemassa niissä pisteissä t, joissa kyseinen odotusarvo on (äärellisenä) olemassa.

- Huomaa, että $M(0) = E\mathrm{e}^0 = 1$.
- Momenttiemäfunktio on (argumentin merkkiä vaille) sama asia kuin Laplacen muunnos (engl. *Laplace transform*).

Momentit momenttiemäfunktion avulla: heuristinen johto

- Jos seuraavassa laskussa derivoinnin ja odotusarvon järjestyksen vaihto pystytään jotenkin perustelemaan, niin

$$M_X'(t) = \frac{\mathrm{d}}{\mathrm{d}t} M_X(t) = \frac{\mathrm{d}}{\mathrm{d}t} E\mathrm{e}^{tX} = E\frac{\partial}{\partial t}\mathrm{e}^{tX} = EX\mathrm{e}^{tX},$$

  joten $EX = M_X'(0)$.

- Laskemalla toinen derivaatta nähdään, että

$$M_X''(t) = \frac{\mathrm{d}}{\mathrm{d}t} EX\mathrm{e}^{tX} = E\frac{\partial}{\partial t}X\mathrm{e}^{tX} = EX^2\mathrm{e}^{tX},$$

  joten $EX^2 = M_X''(0)$.

- Samalla tavalla edeten nähdään, että aina $EX^k = M^{(k)}(0)$.

Momentit momenttiemäfunktion avulla: lause

Lause

Jos momenttiemäfunktio $M_X(t)$ on reaalisena olemassa jossakin origon ympäristössä $|t| < h$, jossa $h > 0$, niin $M_X(t)$ on äärettömän monta kertaa derivoituva origossa, X:n kaikki momentit ovat olemassa, ja M_X voidaan esittää suppenevana Taylorin sarjana

$$M_X(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} EX^k, \qquad |t| < h.$$

Erityisesti $\quad EX^k = M_X^{(k)}(0), \qquad k \geq 1.$

Määritelmä (Kumulanttiemäfunktio)

Jos momenttiemäfunktio $M_X(t)$ on määritelty jossakin origon ympäristössä, niin X:n (jakauman) kumulanttiemäfunktio $K(t) = K_X(t)$ (engl. *cumulant generating function, cgf*) määritellään kyseisessä ympäristössä momenttiemäfunktion logaritmina,

$$K(t) = K_X(t) = \ln M_X(t).$$

Odotusarvo ja varianssi kumulanttiemäfunktion avulla

- Odotusarvo ja varianssi saadaan laskettua derivoimalla kumulanttiemäfunktiota origossa:

$$K_X'(0) = \frac{M_X'(0)}{M_X(0)} = \frac{EX}{1} = EX,$$

$$K_X''(0) = \frac{M_X''(0)M_X(0) - M'(0)M'(0)}{M_X(0)^2} = EX^2 - (EX)^2 = \operatorname{var} X.$$

- Tämä tarjoaa näppärän tavan johtaa odotusarvo ja varianssi joillekin jakaumille, mutta toisille jakaumille sattaa olla kätevämpää johtaa odotusarvo ja varianssi suoraan määritelmistä.

Binomijakauman odotusarvo ja varianssi kumulanttiemäfunktion avulla

- Jos $X \sim \mathrm{Bin}(n, p)$, niin binomikaavan avulla nähdään helposti, että
$$M_X(t) = (p\,\mathrm{e}^t + 1 - p)^n.$$
- Siis
$$K_X(t) = n \ln(p\,\mathrm{e}^t + 1 - p).$$
- Laskemalla ensimmäinen ja toinen derivaatta origossa nähdään, että
$$EX = K_X'(0) = np, \qquad \mathrm{var}\, X = K_X''(0) = np(1 - p).$$

Momenttiemäfunktion ominaisuuksia

Lause

(a) Jos a, b ovat vakioita, niin $M_{aX+b}(t) = \mathrm{e}^{bt} M_X(at)$.
(b) Jos $X \perp\!\!\!\perp Y$, niin $M_{X+Y}(t) = M_X(t)\, M_Y(t)$.

Todistus:

$$M_{aX+b}(t) = E\exp(t(aX+b)) = \mathrm{e}^{bt}\, E\exp(atX) = \mathrm{e}^{bt}\, M_X(at),$$

ja kun $X \perp\!\!\!\perp Y$, niin

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

Yhtäsuuruus pätee joko muodossa $\infty = \infty$ tai reaalilukujen yhtäsuuruutena.

Momenttiemäfunktio määrää jakauman — tietyillä edellytyksillä

Lause

Jos X:n ja Y:n momenttiemäfunktiot ovat molemmat olemassa jossakin origon ympäristössä, ja ovat siellä yhtäsuuret, niin X:llä ja Y:llä on sama jakauma.

Momenttiemäfunktion käytön hankaluuksia

- Jos jokin X:n momentti ei ole olemassa, niin momenttiemäfunktio on määritelty origossa, mutta ei missään origon ympäristössä.
- On olemassa jakaumia (esim. log-normaalinen jakauma), joilla on kaikkien kertalukujen momentit, mutta joiden momenttiemäfunktio ei ole olemassa missään origon ympäristössä.
- Edistyneemmissä todennäköisyyslaskennan kirjoissa käytetään momenttiemäfunktion sijasta jakauman karakteristista funktiota.
- Momenttiemäfunktiota ja kumulanttiemäfunktio ovat tästä huolimatta oleellisia apuvälineitä tietyissä asymptoottisissa tarkasteluissa (esim. suurten poikkeamien teoriassa sekä satulapistekehitelmässä).

4.8 Karakteristinen funktio

- Tietyiltä vaikeuksilta vältytään, mikäli momenttiemäfunktion sijasta käytetään jakauman karakteristista funktiota, sillä karakteristinen funktio on kaikille jakaumille kaikkialla määritelty.
- Karakteristisen funktion määrittelyssä tarvitaan kompleksilukuja.
- Karakteristinen funktio on (argumentin skaalausta vaille) sama asia kuin Fourier'n muunnos (engl. *Fourier transform*).

Karakteristinen funktio

Määritelmä

Satunnaismuuttujan X karakteristinen funktio (engl. *characteristic function*) on koko reaaliakselilla määritelty kompleksiarvoinen funktio

$$\phi(t) = \phi_X(t) = E\mathrm{e}^{itX}, \quad \text{jossa } i = \sqrt{-1}.$$

- Kompleksinen eksponentti saadaan laskettua kaavalla
$$\mathrm{e}^{itX} = \cos(tX) + i\sin(tX),$$
- Sen odotusarvo määritellään laskemalla yhteen reaaliosan ja imaginääriosan odotusarvot,
$$E\mathrm{e}^{itX} = E\cos(tX) + i\,E\sin(tX).$$
- Karakteristinen funktio $\phi_X(t)$ on määritelty kaikilla X ja kaikilla t sen takia, että se on rajoitetun funktion odotusarvo.