Tilastollinen päättely

3. Piste-estimointi

3.1. Johdanto

Estimaattori, Estimaatti, Estimointi, Havainto, Havaintopiste, Otos, Otostunnusluku, Parametri, Piste-estimointi, Pistetodennäköisyysfunktio, Riippumattomuus, Tiheysfunktio

3.2. Estimointimenetelmät

Asymptoottinen normaalisuus, Bernoulli-jakauma, Bijektio, χ^2-jakauma, Estimaattori, Estimaatti, Estimointi, Harhattomuus, Havainto, Havaintopiste, Logaritminen uskottavuusfunktio, Maksimointi, Momenttiestimaattori, Momenttimenetelmä, Normaalijakauma, Normaalisuus, Otos, Otostunnusluku, Parametri, Pistetodennäköisyysfunktio, Riippumattomuus, Suurimman uskottavuuden estimaattori, Suurimman uskottavuuden menetelmä, Tarkentuvuus, Tehokkuus, Tiheysfunktio, *t*-jakauma, Tyhjentävyys, Uskottavuusfunktio, Yhteisjakauma

3.3. Estimaattoreiden ominaisuudet

Asymptoottinen normaalisuus, Cramérin ja Raon alaraja, Cramérin ja Raon epäyhtälö, Estimaattori, Estimaatti, Estimointi, Fisher-informaatio, Harha, Harhattomuus, Havainto, Havaintopiste, Informaatio, Keskineliövirhe, Logaritminen uskottavuusfunktio, Maksimointi, Minimivarianssisuus, Normaalijakauma, Normaalisuus, Otos, Otostunnusluku, Parametri, Paras harhaton estimaattori, Pistetodennäköisyysfunktio, Raon ja Blackwellin teoreema, Riippumattomuus, Schwarzin epäyhtälö, Suurimman uskottavuuden estimaattori, Suurimman uskottavuuden menetelmä, Tarkentuvuus, Tehokkuus, Tiheysfunktio, Tyhjentävyys, Uskottavuusfunktio, Yhteisjakauma

3.1. Johdanto

Piste-estimointi

Olkoon

$$f(x;\theta)$$

satunnaismuuttujan X *pistetodennäköisyys*- tai *tiheysfunktio*, joka riippuu *tuntemattomasta parametrista* θ. Haluamme määrätä parametrille θ mahdollisimman hyvän *estimaatin* eli *arvion* jakaumasta $f(x;\theta)$ poimitun *otoksen* (so. *havaintojen*) perusteella. Kutsumme tätä tehtävää **estimoinniksi**.

Olkoon

$$X_1, X_2, \ldots , X_n$$

(*yksinkertainen*) *satunnaisotos* satunnaismuuttujan X jakaumasta, jonka *pistetodennäköisyys*- tai *tiheysfunktio* $f(x;\theta)$ riippuu parametrista θ. Tällöin *havainnot* $X_1, X_2, \ldots , X_n$ ovat *riippumattomia*, *identtisesti jakautuneita satunnaismuuttujia*, joilla on *sama pistetodennäköisyys*- tai *tiheysfunktio* $f(x;\theta)$:

$$X_1, X_2, \ldots, X_n \perp$$
$$X_i \sim f(x;\theta) \text{ , } i = 1, 2, \ldots, n$$

Olkoon

$$\mathbf{X} = (X_1, X_2, \ldots , X_n)$$

satunnaismuuttujien (havaintojen) $X_1, X_2, \ldots , X_n$ muodostama n-vektori.

Olkoot satunnaismuuttujien $X_1, X_2, \ldots , X_n$ *havaitut arvot*

$$x_1, x_2, \ldots , x_n$$

Merkitään tätä:

$$X_1 = x_1, X_2 = x_2, \ldots , X_n = x_n$$

Satunnaismuuttujien $X_1, X_2, \ldots , X_n$ havaitut arvot $x_1, x_2, \ldots , x_n$ määräävät *havaintopisteen*

$$\mathbf{x} = (x_1, x_2, \ldots , x_n)$$

Kutsumme *satunnaismuuttujien* $X_1, X_2, \ldots , X_n$ (mitallista) *funktiota*

$$W = W(\mathbf{X}) = W(X_1, X_2, \ldots, X_n)$$

(**piste**-) **estimaattoriksi**. Tämä merkitsee sitä, että estimaattori on **otostunnusluku** (ks. luku 1).

Jos satunnaismuuttujien $X_1, X_2, \ldots , X_n$ havaitut arvot ovat $x_1, x_2, \ldots , x_n$, niin *estimaattori*

$$W = W(\mathbf{X}) = W(X_1, X_2, \ldots, X_n)$$

saa havaituksi arvokseen w funktion $W(\cdot)$ arvon havaintopisteessä $\mathbf{x} = (x_1, x_2, \ldots , x_n)$:

$$w = W(\mathbf{x}) = W(x_1, x_2, \ldots, x_n)$$

Kutsumme estimaattorin W havaintoarvoista $x_1, x_2, \ldots , x_n$ määrättyä arvoa w **estimaatiksi**.

Huomautus:

Koska haluamme käyttää estimaattoria W parametrin θ arvojen estimointiin, tuntuisi luonnolliselta sisällyttää estimaattorin määritelmään vaatimus, että estimaattorin pitää *jollakin tavalla* liittyä estimoitavaan parametriin. On kuitenkin osoittautunut, että tällaista vaatimusta ei pidä liittää estimaattorin määritelmään.

3.2. Estimointimenetelmät

Johdanto

Miten todennäköisyysjakauman parametreille löydetään *estimaattorit*?

Tarkastelemme tässä kappaleessa kahta *estimointimenetelmää*: **momenttimenetelmää** ja **suurimman uskottavuuden menetelmää**.

Momenttimenetelmä

Satunnaisotos

Olkoon

$$X_1, X_2, \ldots , X_n$$

satunnaisotos jakaumasta, jonka *pistetodennäköisyys*- tai *tiheysfunktio*

$$f(x;\theta_1, \theta_2, \ldots , \theta_p)$$

riippuu *parametreista*

$$\theta_1, \theta_2, \ldots , \theta_p$$

Tällöin havainnot $X_1, X_2, \ldots , X_n$ ovat *riippumattomia*, *identtisesti jakautuneita satunnaismuuttujia*, joilla on *sama pistetodennäköisyys*- tai *tiheysfunktio* $f(x;\theta_1, \theta_2, \ldots , \theta_p)$:

$$X_1, X_2, \ldots, X_n \perp$$
$$X_i \sim f(x;\theta_1, \theta_2, \ldots, \theta_p)\,,\ i = 1, 2, \ldots, n$$

Momenttiestimaattori

Oletetaan, että jakaumalla $f(x;\theta_1, \theta_2, \ldots , \theta_p)$ on kaikki (*origo*-) *momentit* kertalukuun p saakka:

$$\mathrm{E}(X^k) = \alpha_k\,,\ k = 1, 2, \ldots, p$$

Oletetaan, että momenttien $\alpha_1, \alpha_2 , \ldots , \alpha_p$ ja parametrien $\theta_1, \theta_2, \ldots , \theta_p$ välillä on jatkuva *bijektio* eli kääntäen yksikäsitteinen kuvaus:

$$(1)\qquad \begin{cases} \alpha_1 = g_1(\theta_1, \theta_2, \ldots, \theta_p) \\ \alpha_2 = g_2(\theta_1, \theta_2, \ldots, \theta_p) \\ \vdots \\ \alpha_p = g_p(\theta_1, \theta_2, \ldots, \theta_p) \end{cases}$$

Tällöin parametrit θ_1, θ_2, … , θ_p voidaan esittää momenttien α_1, α_2 , … , α_p funktioina:

$$\text{(2)}\qquad \begin{cases} \theta_1 = h_1(\alpha_1, \alpha_2, \ldots, \alpha_p) \\ \theta_2 = h_2(\alpha_1, \alpha_2, \ldots, \alpha_p) \\ \vdots \\ \theta_p = h_p(\alpha_1, \alpha_2, \ldots, \alpha_p) \end{cases}$$

Estimoidaan momentit α_1, α_2 , … , α_p vastaavilla *otosmomenteilla*:

$$a_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \text{ , } k = 1, 2, \ldots, p$$

Sijoittamalla estimaattorit a_1, a_2 , … , a_p momenttien α_1, α_2 , … , α_p paikalle yhtälöihin (2), saadaan parametrien θ_1, θ_2, … , θ_p **momenttiestimaattorit** eli **MM-estimaattorit**

$$\begin{cases} \hat{\theta}_1 = h_1(a_1, a_2, \ldots, a_p) \\ \hat{\theta}_2 = h_2(a_1, a_2, \ldots, a_p) \\ \vdots \\ \hat{\theta}_p = h_p(a_1, a_2, \ldots, a_p) \end{cases}$$

Monet todennäköisyysjakaumat on *parametroitu jakauman* (origo-) *momenteilla* tai *keskus-momenteilla*:

(i) Jos jakauma on parametroitu jakauman (origo-) momenteilla, ko. parametrien *momentti-estimaattoreita* ovat *vastaavat otosorigomomentit*.

(ii) Jos jakauma on parametroitu jakauman keskusmomenteilla, ko. parametrien *momentti-estimaattoreita* ovat *vastaavat otoskeskusmomentit*.

Momenttiestimaattorin ominaisuudet

Hyvä estimaattori on *harhaton*, *tyhjentävä*, *tehokas* ja *tarkentuva* (ks. lukua 2, kappaletta 3.3 tässä luvussa ja lukua 7). MM-estimaattori ei välttämättä täytä yhtäkään hyvän estimaattorin kriteereistä, joten momenttimenetelmää käytettäessä on aina erikseen varmistettava tuloksena saadun estimaattorin hyvyys.

Esimerkki 2.1: Normaalijakauman parametrien momenttiestimointi

Satunnaismuuttuja X noudattaa *normaalijakaumaa* parametrein

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

jos sen tiheysfunktio on muotoa

$$f(x;\mu,\sigma^2) = (2\pi\sigma^{-2})^{-1/2}\exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right]$$
$$-\infty < \mu < +\infty \text{ , } \sigma > 0 \text{ , } -\infty < x < +\infty$$

Parametri μ on normaalijakauman *odotusarvo* ja parametri σ^2 on normaalijakauman *varianssi*.

Oletetaan, että *havainnot*

$X_1, X_2, \ldots, X_n$

muodostavat satunnaisotoksen normaalijakaumasta $\mathrm{N}(\mu,\sigma^2)$. Tällöin

$$X_1, X_2, \ldots, X_n \perp$$
$$X_i \sim \mathrm{N}(\mu,\sigma^2)\,,\ i = 1,2,\ldots,n$$

Parametrien μ ja σ^2 **MM-estimaattorit** ovat havaintojen $X_1, X_2, \ldots, X_n$ *aritmeettinen keskiarvo*

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

ja *otosvarianssi*

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Perustelu:

Oletetaan, että satunnaismuuttuja X noudattaa *normaalijakaumaa* parametrein μ ja σ^2:

$$X \sim \mathrm{N}(\mu,\sigma^2)$$

Jakauman parametrien ja satunnaismuuttujan origomomenttien välillä on seuraava *bijektio*:

(i) Parametrit μ ja σ^2 lausuttuina 1. ja 2. origomomentin funktioina:

$$\begin{cases} \mu = \mathrm{E}(X) = \alpha_1 \\ \sigma^2 = \mathrm{Var}(X) = \mathrm{E}[(X-\mu)^2] = \mathrm{E}(X^2) - \mu^2 = \alpha_2 - \alpha_1^2 \end{cases}$$

(ii) 1. ja 2. origomomentti lausuttuina parametrien μ ja σ^2 funktioina:

$$\begin{cases} \alpha_1 = \mathrm{E}(X) = \mu \\ \alpha_2 = \mathrm{E}(X^2) = \sigma^2 + \mu^2 \end{cases}$$

Olkoon $X_1, X_2, \ldots, X_n$ *satunnaisotos* normaalijakaumasta $\mathrm{N}(\mu,\sigma^2)$.

Havaintojen $X_1, X_2, \ldots, X_n$ 1. ja 2. *otosorigomomentti* saadaan kaavoilla

$$a_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k\,,\ k = 1,2$$

Siten normaalijakauman $\mathrm{N}(\mu,\sigma^2)$ parametrien μ ja σ^2 *MM-estimaattorit* ovat

$$\begin{cases} \hat{\mu} = a_1 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i \\ \hat{\sigma}^2 = a_2 - a_1^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i^2 - \left(\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i\right)^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(X_i - \bar{X})^2 \end{cases}$$

Huomaa, että $\hat{\mu}$ on havaintojen *aritmeettinen keskiarvo* ja $\hat{\sigma}^2$ on havaintojen 2. *otoskeskusmomentti* eli havaintojen otosvarianssi, jossa neliösumman $\sum(X_i - \bar{X})^2$ jakajana on käytetty havaintojen lukumäärää n.

Estimaattoreiden ominaisuudet: ks. esimerkkiä 2.2.

Suurimman uskottavuuden menetelmä

Satunnaisotos

Olkoon

$$X_1, X_2, \dots, X_n$$

satunnaisotos satunnaismuuttujan X jakaumasta, jonka *pistetodennäköisyys*- tai *tiheysfunktio* $f(x;\theta)$ riippuu parametrista θ. Tällöin havainnot $X_1, X_2, \dots, X_n$ ovat *riippumattomia*, *identtisesti jakautuneita satunnaismuuttujia*, joilla on *sama pistetodennäköisyys*- tai *tiheysfunktio* $f(x;\theta)$:

$$X_1, X_2, \ldots, X_n \perp$$
$$X_i \sim f(x;\theta)\,,\; i = 1, 2, \ldots, n$$

Uskottavuusfunktio

Koska havainnot $X_1, X_2, \dots, X_n$ on oletettu *riippumattomiksi*, niin otoksen

$$\mathbf{X} = (X_1, X_2, \ldots, X_n)$$

yhteisjakauman pistetodennäköisyys- tai *tiheysfunktiolla* on tulomuotoinen esitys

$$f(x_1, x_2, \ldots, x_n;\theta) = f(x_1;\theta) f(x_2;\theta) \cdots f(x_n;\theta)$$

jossa

$$f(x_i;\theta)\,,\; i = 1, 2, \ldots, n$$

on yksittäiseen havaintoon X_i , $i = 1, 2, \dots, n$ liittyvä pistetodennäköisyys- tai tiheysfunktio.

Otoksen $X_1, X_2, \dots, X_n$ **uskottavuusfunktio**

$$L(\theta; x_1, x_2, \ldots, x_n) = f(x_1, x_2, \ldots, x_n;\theta)$$

on havaintojen $X_1, X_2, \dots, X_n$ *yhteisjakauman pistetodennäköisyys*- tai *tiheysfunktion* f arvo pisteessä

$$x_1, x_2, \dots, x_n$$

tulkittuna *parametrin* θ *arvojen funktioksi*. Uskottavuusfunktio L sisältää *kaiken* (stokastisen) *informaation otoksesta*.

Suurimman uskottavuuden estimaattori

Olkoon

$$t = g(x_1, x_2, \ldots, x_n)$$

parametrin θ arvo, joka *maksimoi otoksen* $X_1, X_2, \dots, X_n$ *uskottavuusfunktion*

$$L(\theta; x_1, x_2, \ldots, x_n)$$

parametrin θ suhteen. Huomaa, että uskottavuusfunktion L maksimin antava parametrin θ arvo t on muuttujien (havaintoarvojen) $x_1, x_2, \dots, x_n$ funktio.

Sijoittamalla uskottavuusfunktion L maksimin parametrin θ suhteen antavassa lausekkeessa

$$t = t(x_1, x_2, \ldots, x_n)$$

muuttujien

$$x_1, x_2, \ldots, x_n$$

paikalle satunnaismuuttujat (havainnot)

$$X_1, X_2, \ldots, X_n$$

saadaan parametrin θ **suurimman uskottavuuden estimaattori** eli **SU-estimaattori**

$$\hat{\theta} = g(X_1, X_2, \ldots, X_n)$$

Parametrin θ suurimman uskottavuuden estimaattori $\hat{\theta}$ tuottaa parametrille θ arvon, joka *maksimoi poimitun otoksen* eli *saatujen havaintoarvojen uskottavuuden* (*todennäköisyyden*). Siten suurimman uskottavuuden estimaattorin otoskohtainen arvo *maksimoi uskottavuuden* (*todennäköisyyden*) *saada juuri se otos, joka on saatu*.

Suurimman uskottavuuden estimointimenetelmää sovellettaessa oletetaan implisiittisesti, että **uskottavuusperiaate** pätee (ks. luku 2).

Suurimman uskottavuuden estimaattorin määrääminen

Parametrin θ suurimman uskottavuuden estimaattori määrätään *etsimällä uskottavuusfunktion*

$$L(\theta; x_1, x_2, \ldots, x_n)$$

(*globaali*) *maksimi* parametrin θ suhteen.

Tavallisesti uskottavuusfunktio on parametrin θ funktiona niin *säännöllinen*, että sen ääriarvot voidaan löytää *derivoimalla*: Jos uskottavuusfunktiolla on maksimi jossakin parametriavaruuden *sisäpisteessä*, se löydetään kaikissa tavanomaisissa tilanteissa merkitsemällä uskottavuusfunktion 1. *derivaatta* $L'(\theta)$ *nollaksi* ja ratkaisemalla θ saadusta *normaaliyhtälöstä*

$$L'(\theta) = 0$$

Olkoon $\hat{\theta}$ jokin normaaliyhtälön ratkaisu. Piste $\hat{\theta}$ *vastaa uskottavuusfunktion* (*lokaalia*) *maksimia*, jos uskottavuusfunktion 2. derivaatta $L''(\theta)$ on *negatiivinen*:

$$L''(\hat{\theta}) < 0$$

Uskottavuusfunktion käyttäytyminen parametriavaruuden *reunalla* on tarkistettava erikseen, koska on mahdollista, että uskottavuusfunktio saavuttaa maksiminsa parametriavaruuden reunalla pisteessä, jossa

$$L'(\theta) \neq 0$$

Logaritminen uskottavuusfunktio

Uskottavuusfunktion maksimi voidaan etsiä maksimoimalla uskottavuusfunktion sijasta *logaritminen uskottavuusfunktio* (uskottavuusfunktion logaritmi)

$$l(\theta; x_1, x_2, \ldots, x_n) = \log L(\theta; x_1, x_2, \ldots, x_n)$$

parametrin θ suhteen. Tämä johtuu siitä, että logaritmifunktio on *aidosti monotoninen*, jolloin uskottavuusfunktio ja sen logaritmi saavuttavat ääriarvonsa samoissa pisteissä.

Koska havainnot $X_1, X_2, \ldots, X_n$ on oletettu *riippumattomiksi*, *logaritminen uskottavuusfunktio* voidaan kirjoittaa seuraavaan muotoon:

$$\begin{aligned} l(\theta) &= \log L(\theta) \\ &= \log\big(f(x_1;\theta)f(x_2;\theta)\cdots f(x_n;\theta)\big) \\ &= \log f(x_1;\theta) + \log f(x_2;\theta) + \cdots + \log f(x_n;\theta) \\ &= l(\theta;x_1) + l(\theta;x_2) + \cdots + l(\theta;x_n) \end{aligned}$$

Tässä

$$l(\theta;x_i) = \log f(x_i;\theta)\ ,\ i = 1, 2, \ldots, n$$

on havaintoarvoon x_i liittyvä logaritminen uskottavuusfunktio.

Suurimman uskottavuuden estimaattorin ominaisuudet

Hyvä estimaattori on *harhaton*, *tyhjentävä*, *tehokas* ja *tarkentuva* (ks. lukua 2, kappaletta 3.3 tässä luvussa ja lukua 7). SU-estimaattori ei välttämättä täytä yhtäkään hyvän estimaattorin kriteereistä, joten suurimman uskottavuuden menetelmää käytettäessä on aina erikseen varmistettava tuloksena saadun estimaattorin hyvyys.

Jos parametrin θ SU-estimaattori ei täytä *hyvän estimaattorin kriteereitä* (ks. kappaletta 3.3) *äärellisillä havaintojen lukumäärillä*, SU-estimaattorin käyttöä parametrin θ estimaattorina voidaan kuitenkin usein perustella sillä, että SU-estimaattorilla on hyvin yleisesti *hyvät* **asymptoottiset ominaisuudet** (ks. asymptoottista teoriaa koskevaa 7. lukua).

Voidaan osoittaa, että hyvin yleisin ehdoin pätee:

(i) SU-estimaattori $\hat{\theta}$ on **tarkentuva** eli

$$\Pr(\hat{\theta} \to \theta) = 1, \text{ kun } n \to +\infty$$

(ii) SU-estimaattori $\hat{\theta}$ on **asymptoottisesti normaalinen**.

SU-estimaattorin *tarkentuvuus* merkitsee sitä, että SU-estimaattori toteuttaa **suurten lukujen lain**. Suurten lukujen lain mukaan *SU-estimaattorin arvo lähestyy* (*melkein varmasti*) *parametrin oikeata arvoa*, kun otoskoko kasvaa.

SU-estimaattorin *asymptoottinen normaalisuus* merkitsee sitä, että SU-estimaattori toteuttaa **keskeisen raja-arvolauseen**. Keskeisen raja-arvolauseen mukaan *SU-estimaattorin jakaumaa voidaan suurissa otoksissa approksimoida normaalijakaumalla*.

Suurimman uskottavuuden estimaattori *vs* momenttiestimaattori

Monissa alkeellisissa tilanteissa suurimman uskottavuuden menetelmällä ja momenttimenetelmällä saadaan *samat* estimaattorit. *Yleisesti tämä ei kuitenkaan ole totta*.

Nykyaikaisessa tilastotieteessä *SU-menetelmä on hyvin pitkälti syrjäyttänyt momenttimenetelmän* estimaattoreiden johtamisen menetelmänä. Tämä johtuu siitä, että SU-menetelmällä on momenttimenetelmää vankempi teoreettinen perusta: *Uskottavuusperiaate* (ks. luku 2) tukee voimakkaasti uskottavuusfunktioon perustuvan menetelmän käyttämistä estimaattoreiden johtamiseen. Lisäksi SU-estimaattoreilla on *hyvät asymptoottiset ominaisuudet* (ks. tarkemmin lukua 7).

Esimerkki 2.2: Normaalijakauman parametrien suurimman uskottavuuden estimointi

Satunnaismuuttuja X noudattaa *normaalijakaumaa* parametrein

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

jos sen tiheysfunktio on muotoa

$$f(x;\mu,\sigma^2) = (2\pi\sigma^{-2})^{-1/2}\exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right]$$
$$-\infty < \mu < +\infty\,,\ \sigma > 0\,,\ -\infty < x < +\infty$$

Parametri μ on normaalijakauman *odotusarvo* ja parametri σ^2 on normaalijakauman *varianssi*.

Oletetaan, että *havainnot*

$$X_1, X_2, \ldots, X_n$$

muodostavat satunnaisotoksen normaalijakaumasta $\mathrm{N}(\mu,\sigma^2)$. Tällöin

$$X_1, X_2, \ldots, X_n \perp$$
$$X_i \sim \mathrm{N}(\mu,\sigma^2)\,,\ i = 1,2,\ldots,n$$

Parametrien μ ja σ^2 **SU-estimaattorit** ovat havaintojen $X_1, X_2, \ldots, X_n$ *aritmeettinen keskiarvo*

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

ja *otosvarianssi*

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Perustelu:

Olkoon $X_1, X_2, \ldots, X_n$ *satunnaisotos* normaalijakaumasta $\mathrm{N}(\mu,\sigma^2)$.

Otoksen $X_1, X_2, \ldots, X_n$ *uskottavuusfunktio* on

$$\begin{aligned}
&L(\mu,\sigma^2;x_1,x_2,\ldots,x_n)\\
&\quad = f(x_1;\mu,\sigma^2)f(x_2;\mu,\sigma^2)\cdots f(x_n;\mu,\sigma^2)\\
&\quad = \sigma^{-n}(2\pi)^{-\frac{1}{2}n}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}
\end{aligned}$$

Otoksen $X_1, X_2, \ldots, X_n$ *logaritminen uskottavuusfunktio* on

$$\begin{aligned}
&l(\mu,\sigma^2;x_1,x_2,\ldots,x_n)\\
&\quad = \log L(\mu,\sigma^2;x_1,x_2,\ldots,x_n)\\
&\quad = -\frac{n}{2}\log\sigma^2 - \frac{1}{2}n\log(2\pi) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2
\end{aligned}$$

Derivoidaan logaritminen uskottavuusfunktio $l(\mu,\sigma^2)$ parametrin μ suhteen ja merkitään derivaatta nollaksi:

$$\frac{\partial l(\mu,\sigma^2)}{\partial\mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) = 0$$

Derivaatan ainoa *nollakohta*

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

antaa log-uskottavuusfunktion *maksimin* parametrin μ suhteen.

Sijoitetaan ratkaisu logaritmiseen uskottavuusfunktioon $l(\mu,\sigma^2)$:

$$l(\bar{x},\sigma^2) = -\frac{n}{2}\log\sigma^2 - \frac{1}{2}n\log(2\pi) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Derivoidaan funktio $l(\bar{x},\sigma^2)$ parametrin σ^2 suhteen ja merkitään derivaatta nollaksi:

$$\frac{\partial l(\mu,\sigma^2)}{\partial\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \bar{x}) = 0$$

Derivaatan ainoa *nollakohta*

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

antaa log-uskottavuusfunktion *maksimin* parametrin σ^2 suhteen.

Yllä esitetystä seuraa, että normaalijakauman $\mathrm{N}(\mu,\sigma^2)$ parametrien μ ja σ^2 *SU-estimaattorit* ovat

$$\begin{cases} \hat{\mu} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i \\ \hat{\sigma}^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(X_i - \bar{X})^2 \end{cases}$$

Siten normaalijakauman $\mathrm{N}(\mu,\sigma^2)$ parametrien μ ja σ^2 SU-estimaattorit yhtyvät niiden *momenttiestimaattoreihin*

Voidaan osoittaa, että normaalijakauman $\mathrm{N}(\mu,\sigma^2)$ odotusarvon μ SU-estimaattorilla $\bar{X}$ on seuraavat ominaisuudet (ks. lukuja 2, 3 ja 7 sekä kappaletta 3.3):

(i) $\bar{X}$ on *harhaton*.

(ii) $\bar{X}$ ja $\hat{\sigma}^2$ ovat yhdessä *tyhjentäviä* parametreille μ ja σ^2.

(iii) $\bar{X}$ *on tehokas* eli minimivarianssinen estimaattori.

(iv) $\bar{X}$ on *tarkentuva*.

(v) $\bar{X}$ noudattaa *normaalijakaumaa*:

$$\bar{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

Voidaan osoittaa, että normaalijakauman $\mathrm{N}(\mu,\sigma^2)$ varianssin σ^2 SU-estimaattorilla $\hat{\sigma}^2$ on seuraavat ominaisuudet (ks. lukuja 2, 3 ja 7 sekä kappaletta 3.3):

(i) $\hat{\sigma}^2$ on *harhainen*, mutta estimaattori

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{n}{n-1}\hat{\sigma}^2$$

on *harhaton*.

(ii) $\bar{X}$ ja $\hat{\sigma}^2$ ovat yhdessä *tyhjentäviä* parametreille μ ja σ^2.

(iii) $\hat{\sigma}^2$ *ei ole tehokas* eli minimivarianssinen estimaattori.

(iv) $\hat{\sigma}^2$ on *tarkentuva*.

(v) $(n-1)\,s^2/\sigma^2$ noudattaa χ^2-*jakaumaa* vapausastein $(n-1)$:

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$$

Lisäksi voidaan osoittaa, että $\bar{X}$ ja $\hat{\sigma}^2$ ovat *riippumattomia* ja että (ks. lukua 1)

$$t = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t(n-1)$$

Suurimman uskottavuuden estimaattori voidaan usein johtaa usein vetoamalla suurimman uskottavuuden estimaattorin **invarianssiominaisuuteen**. Sen mukaan parametrin *funktioiden* suurimman uskottavuuden estimaattorit saadaan soveltamalla ko. funktioita ko. parametrin suurimman uskottavuuden estimaattoriin.

Ennen SU-estimaattorin invarianssiominaisuutta koskevan lauseen todistamista määrittelemme ns. *indusoidun uskottavuusfunktion*: Olkoon

$$\eta = \tau(\theta)$$

jokin parametrin θ funktio. Tällöin **indusoitu uskottavuusfunktio** L^* saadaan kaavasta

(1) $$L^*(\eta;\mathbf{x}) = \sup_{\{\theta|\tau(\theta)=\eta\}} L(\theta;\mathbf{x})$$

Olkoon $\hat{\eta}$ paramerin η arvo, joka *maksimoi* indusoidun uskottavuusfunktion L^*. Tällöin estimaattoria $\hat{\eta}$ kutsutaan parametrin $\eta = \tau(\theta)$ **suurimman uskottavuuden estimaattoriksi**.

Määritelmästä (1) nähdään suoraan, että uskottavuusfunktiolla L ja indusoidulla uskottavuusfunktiolla L^* on *sama* maksimi.

Lause:

Olkoon $\hat{\theta}$ parametrin θ suurimman uskottavuuden estimaattori. Tällöin $\tau(\hat{\theta})$ on parametrin $\tau(\theta)$ suurimman uskottavuuden estimaattori kaikille funktioille $\tau(\cdot)$.

Todistus:

Olkoon $\hat{\eta}$ paramerin η arvo, joka *maksimoi* indusoidun uskottavuusfunktion

$$L^*(\eta;\mathbf{x})$$

Lause tulee todistetuksi, jos osoitamme, että

$$L^*(\hat{\eta};\mathbf{x}) = \sup_{\eta} \sup_{\{\theta|\tau(\theta)=\eta\}} L(\theta;\mathbf{x})$$

Koska uskottavuusfunktiolla L ja indusoidulla uskottavuusfunktiolla L^* on *sama* maksimi, niin

$$\begin{aligned} L^*(\hat{\eta};\mathbf{x}) &= \sup_{\eta} \sup_{\{\theta|\tau(\theta)=\eta\}} L(\theta;\mathbf{x}) && |\text{ funktion } L^* \text{ määritelmä} \\ &= \sup_{\theta} L(\theta;\mathbf{x}) \\ &= L(\hat{\theta};\mathbf{x}) && |\text{ estimaattorin } \hat{\theta} \text{ määritelmä} \end{aligned}$$

jossa toinen = -merkki seuraa siitä, että iteroitu maksimointi yhtyy maksimointiin parametrin θ suhteen. Edelleen

$$\begin{aligned} L(\hat{\theta};\mathbf{x}) &= \sup_{\{\theta|\tau(\theta)=\tau(\hat{\theta})\}} L(\theta;\mathbf{x}) && |\ \hat{\theta} \text{ on SU-estimaattori} \\ &= \sup_{\theta} L(\theta;\mathbf{x}) \\ &= L^*(\tau(\hat{\theta});\mathbf{x}) && |\text{ funktion } L^* \text{ määritelmä} \end{aligned}$$

Siten

$$L^*(\hat{\eta};\mathbf{x}) = L^*(\tau(\hat{\theta});\mathbf{x})$$

ja $\tau(\hat{\theta})$ on parametrin $\tau(\theta)$ SU-estimaattori.

■

3.3. Estimaattoreiden ominaisuudet

Johdanto

Tässä kappaleessa tarkastellaan estimaattoreiden *hyvyysominaisuuksia*.

Keskineliövirhe

Olkoon W parametrin θ estimaattori. Estimaattorin W **keskineliövirhe** (*engl*. Mean Squared Error) on

$$\mathrm{MSE}(W) = \mathrm{E}_\theta[(W-\theta)^2]$$

On helppo nähdä, että

$$\mathrm{MSE}(W) = \mathrm{Var}_\theta(W) + [\mathrm{Bias}_\theta(W)]^2$$

jossa

$$\mathrm{Var}_\theta(W) = \mathrm{E}_\theta[(W - \mathrm{E}_\theta(W))^2]$$

on estimaattorin W **varianssi** ja

$$\mathrm{Bias}_\theta(W) = \mathrm{E}_\theta(W) - \theta$$

on estimaattorin W **harha**.

Perustelu:

$$\begin{aligned}\mathrm{MSE}(W) &= \mathrm{E}_\theta[(W-\theta)^2] \\ &= \mathrm{E}_\theta[(W - \mathrm{E}_\theta(W) + \mathrm{E}_\theta(W) - \theta)^2] \\ &= \mathrm{E}_\theta[(W - \mathrm{E}_\theta(W))^2] + 2\,\mathrm{E}_\theta[(W - \mathrm{E}_\theta(W))(\mathrm{E}_\theta(W) - \theta)] + [\mathrm{E}_\theta(W) - \theta]^2 \\ &= \mathrm{E}_\theta[(W - \mathrm{E}_\theta(W))^2] + [\mathrm{E}_\theta(W) - \theta]^2 \\ &= \mathrm{Var}_\theta(W) + [\mathrm{Bias}_\theta(W)]^2\end{aligned}$$

■

Harhattomuus

Olkoon W parametrin θ estimaattori. Estimaattorin W **harha** (*Bias*) on

$$\mathrm{Bias}_\theta(W) = \mathrm{E}_\theta(W) - \theta$$

Estimaattori W on **harhaton** parametrille θ, jos

$$\mathrm{Bias}_\theta(W) = 0$$

jolloin

$$\mathrm{E}_\theta(W) = \theta$$

ja

$$\mathrm{MSE}(W) = \mathrm{Var}_\theta(W)$$

Aritmeettisen keskiarvon ja otosvarianssin harhattomuus

Olkoon $X_1, X_2, \ldots, X_n$ *satunnaisotos* jakaumasta, jonka odotusarvo on μ ja varianssi on σ^2.

Määritellään havaintojen $X_1, X_2, \ldots, X_n$ *aritmeettinen keskiarvo* kaavalla

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

ja *otosvarianssi* kaavalla

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Aritmeettinen keskiarvo $\bar{X}$ ja otosvarianssi s^2 ovat parametrien μ ja σ^2 *harhattomat estimaattorit* (ks. lukua 1):

$$\mathrm{E}(\bar{X}) = \mu$$
$$\mathrm{E}(s^2) = \sigma^2$$

Jos $X_1, X_2, \ldots, X_n$ on *satunnaisotos* normaalijakaumasta $\mathrm{N}(\mu,\sigma^2)$, yllä todetusta seuraa, että varianssin σ^2 *suurimman uskottavuuden estimaattori* (tai *momenttiestimaattori*)

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{n-1}{n}s^2$$

on *harhainen*, mutta harha *häviää*, jos otoskoon n annetaan kasvaa rajatta.

Tämä nähdään seuraavalla tavalla: Koska

$$\hat{\sigma}^2 = \left(1 - \frac{1}{n}\right)s^2$$

niin

$$\mathrm{E}(\hat{\sigma}^2) = \left(1 - \frac{1}{n}\right)\mathrm{E}(s^2) = \left(1 - \frac{1}{n}\right)\sigma^2 \to \sigma^2$$

jos $n \to \infty$. Siten normaalijakauman varianssin σ^2 *suurimman uskottavuuden estimaattori* (tai *momenttiestimaattori*) on *asymptoottisesti harhaton*.

Paras harhaton estimaattori

Saattaisi tuntua houkuttelevalta pitää vaihtoehtoisista estimaattoreista parhaana sitä, jonka keskineliövirhe on pienin.

Keskineliövirhettä ei voida kuitenkaan sellaisenaan käyttää parhaan estimaattorin valintaan. Tästä antaa esimerkin seuraava tilanne: Olkoon $\hat{\theta}$ parametrin θ estimaattori, jolla on vakioarvo a:

$$\hat{\theta} = a$$

Jos parametrin θ todellinen arvo sattuisi olemaan a, niin

$$\mathrm{MSE}(\hat{\theta}) = 0$$

mutta muulloin estimaattori $\hat{\theta} = a$ on tietysti muuten täysin käyttökelvoton.

Ongelma on siinä, että kaikkien estimaattoreiden luokka on *liian laaja* järkevien vertailujen tekemiseen. Sen sijaan, *jos tarkasteltavien estimaattoreiden luokkaa rajataan sopivasti*, *saatetaan ko. luokasta löytää paras estimaattori*.

Olkoot W_1 ja W_2 kaksi parametrin θ *harhatonta* estimaattoria. Tällöin

$$\mathrm{E}(W_1) = \mathrm{E}(W_2) = \theta$$

ja estimaattoreiden W_1 ja W_2 vertailuun voidaan käyttää niiden *varianssia*: Estimaattori W_1 on **parempi** kuin estimaattori W_2, jos

$$\mathrm{Var}(W_1) < \mathrm{Var}(W_2)$$

Sanomme usein, että estimaattori W_1 on tällöin **tehokkaampi** kuin estimaattori W_2.

Itse asiassa pätee seuraava: Olkoon W^* parametrin θ estimaattori, jolle

$$\mathrm{E}_\theta(W^*) = \tau(\theta) \neq \theta$$

Tarkastellaan estimaattoreiden luokkaa

$$C_\tau = \{W \mid \mathrm{E}_\theta(W) = \tau(\theta)\}$$

Olkoot

$$W_1,\, W_2 \in C_\tau$$

Tällöin

$$\mathrm{Bias}_\theta(W_1) = \mathrm{Bias}_\theta(W_2)$$

ja

$$\mathrm{MSE}(W_1) - \mathrm{MSE}(W_2) = \mathrm{Var}_\theta(W_1) - \mathrm{Var}_\theta(W_2)$$

joten estimaattoreiden keskineliövirhevertailut voidaan *luokassa* C_τ perustaa estimaattoreiden variansseihin.

Estimaattori W^* on parametrin $\tau(\theta)$ **paras harhaton estimaattori**, jos

$$\mathrm{E}_\theta(W^*) = \tau(\theta)$$

kaikille θ ja

$$\mathrm{Var}_\theta(W^*) \le \mathrm{Var}_\theta(W)$$

kaikille estimaattoreille W, jotka toteuttavat ehdon

$$\mathrm{E}_\theta(W) = \tau(\theta)$$

Parasta harhatonta estimaattoria kutsutaan usein **tasaisesti parhaaksi minimivarianssiseksi harhattomaksi estimaattoriksi** (UMVUE, *engl*. Uniform Minimum Variance Unbiased Estimator) tai **(täys-) tehokkaaksi estimaattoriksi**.

Jos kaikkien harhattomien estimaattoreiden varianssille voidaan löytää teoreettinen alaraja ja estimaattorin W varianssi saavuttaa ko. alarajan, tiedetään, että W on *paras harhaton estimaattori* eli *tehokas estimaattori*.

Harhattomien estimaattoreiden varianssien teoreettista alarajaa – ns. Cramérin ja Raon alarajaa – tarkastellaan seuraavassa kappaleessa.

Tehokkuus

Tarkastelemme tässä todennäköisyysjakauman parametrin estimaattorin varianssin teoreettista alarajaa. Todistettavat lauseet on muotoiltu pääasiassa jatkuville jakaumille, mutta ne pätevät sopivasti modifioituina myös diskreeteille jakaumille.

Muotoilemme ja todistamme ennen tämän kappaleen päätuloksen eli *Cramérin ja Raon epäyhtälön* muotoilemista ja todistamista sen todistamisessa tarvittavan *Schwarzin epäyhtälön*.

Schwarzin epäyhtälö:

Olkoot satunnaismuuttujien X ja Y *odotusarvot*, *varianssit* ja *kovarianssit*

$$\mathrm{E}(X) = \mu_X$$
$$\mathrm{E}(Y) = \mu_Y$$

$$\mathrm{Var}(X) = \mathrm{E}[(X-\mu_X)^2] = \mathrm{D}^2(X) = \sigma_X^2$$
$$\mathrm{Var}(Y) = \mathrm{E}[(Y-\mu_Y)^2] = \mathrm{D}^2(Y) = \sigma_Y^2$$
$$\mathrm{Cov}(X,Y) = \mathrm{E}[(X-\mu_X)(Y-\mu_Y)] = \sigma_{XY}$$

Tällöin

$$[\mathrm{Cov}(X,Y)]^2 \le \mathrm{Var}(X)\,\mathrm{Var}(Y)$$

ja lisäksi yhtäsuuruus pätee, jos ja vain jos on olemassa vakiot a ja b siten, että

$$\Pr(Y = a + bX) = 1$$

Todistus:

Tarkastellaan funktiota

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

Helposti nähdään, että

$$\begin{aligned} h(t) &= \mathrm{E}[((X-\mu_X)t + (Y-\mu_Y))^2] \\ &= \mathrm{E}[(X-\mu_X)^2]t^2 + 2\,\mathrm{E}[(X-\mu_X)(Y-\mu_Y)]t + \mathrm{E}[(Y-\mu_Y)^2] \\ &= \sigma_X^2\,t^2 + 2\sigma_{XY}\,t + \sigma_Y^2 \end{aligned}$$

Funktio $h(t)$ on muuttujan t funktiona *ylöspäin aukeava paraabeli*. Koska $h(t)$ on ei-negatiivisen satunnaismuuttujan odotusarvo, niin

$$h(t) \ge 0$$

kaikille t. Siten yhtälöllä

$$h(t) = 0$$

voi olla korkeintaan yksi reaalijuuri, joten yhtälön *diskriminantin* D on oltava ei-positiivinen:

$$D = (2\sigma_{XY})^2 - 4\sigma_X^2\sigma_Y^2 \le 0$$

eli

$$(\sigma_{XY})^2 \le \sigma_X^2\sigma_Y^2$$

Yhtälöllä

$$h(t) = 0$$

on yksi reaalijuuri, jos yhtälön diskriminantilla D on arvo nolla:

$$D = (2\sigma_{XY})^2 - 4\sigma_X^2\sigma_Y^2 = 0$$

Koska

$$((X-\mu_X)t + (Y-\mu_Y))^2 \ge 0$$

niin

$$h(t) = \mathrm{E}[((X-\mu_X)t + (Y-\mu_Y))^2] = 0$$

jos ja vain jos

$$\Pr\left([(X-\mu_X)t+(Y-\mu_Y)]^2=0\right)=1$$

eli

$$\Pr(Y = a + bX) = 1$$

jossa

$$b = -t$$

$$a = \mu_X t + \mu_Y$$

ja t on yhtälön $h(t) = 0$ juuri. Tämä juuri on

$$t = -\frac{\mathrm{Cov}(X,Y)}{\sigma_X^2}$$

Lisäksi tästä nähdään, että vakiolla b ja satunnaismuuttujien X ja Y kovarianssilla on sama merkki.

■

Huomaa, että Schwarzin epäyhtälöstä saadaan sivutuotteena satunnaismuuttujien X ja Y *Pearsonin* (*tulomomentti-*) *korrelaatiokerrointa*

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\mathrm{D}(X)\,\mathrm{D}(Y)} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}$$

koskeva epäyhtälö

$$-1 \le \rho_{XY} \le +1$$

Muotoillaan ja todistetaan seuraavaksi tämän kappaleen päätulos:

Lause: Cramérin ja Raon epäyhtälö

Olkoon

$$X_1, X_2, \ldots, X_n$$

satunnaisotos, jonka yhteisjakauman *tiheysfunktio* on $f(\mathbf{x};\theta)$ ja olkoon

$$W(\mathbf{X}) = W(X_1, X_2, \ldots, X_n)$$

parametrin θ estimaattori, jolle

$$\frac{d}{d\theta}\mathrm{E}_\theta[W(\mathbf{X})] = \int_{\chi}\frac{\partial}{\partial\theta}[W(\mathbf{x})f(\mathbf{x};\theta)]d\mathbf{x}$$

ja

$$\mathrm{Var}_\theta[W(\mathbf{X})] < \infty$$

Tällöin

$$\mathrm{Var}_\theta[W(\mathbf{X})] \ge \frac{\left[\dfrac{d}{d\theta}\mathrm{E}_\theta[W(\mathbf{X})]\right]^2}{\mathrm{E}_\theta\left(\left[\dfrac{\partial}{\partial\theta}\log(f(\mathbf{X};\theta))\right]^2\right)}$$

Todistus:

Sovelletaan todistuksessa Schwarzin epäyhtälöä

$$[\mathrm{Cov}(X,Y)]^2 \le \mathrm{Var}(X)\,\mathrm{Var}(Y)$$

Epäyhtälöstä seuraa satunnaismuuttujan X varianssille alaraja

$$\mathrm{Var}(X) \ge \frac{[\mathrm{Cov}(X,Y)]^2}{\mathrm{Var}(Y)}$$

Cramérin ja Raon epäyhtälön todistus perustuu tähän epäyhtälöön ja valintaan

$$X = W(\mathbf{X})$$

$$Y = \frac{\partial}{\partial\theta}\log(f(\mathbf{X};\theta))$$

Todetaan ensin, että

$$\begin{aligned}
\frac{d}{d\theta}\mathrm{E}_\theta[W(\mathbf{X})] &= \int_\chi W(\mathbf{x})\left[\frac{\partial}{\partial\theta}f(\mathbf{x};\theta)\right]d\mathbf{x} \\
&= \int_\chi W(\mathbf{x})\left[\frac{\frac{\partial}{\partial\theta}f(\mathbf{x};\theta)}{f(\mathbf{x};\theta)}\right]f(\mathbf{x};\theta)d\mathbf{x} \\
&= \mathrm{E}_\theta\left[W(\mathbf{X})\frac{\frac{\partial}{\partial\theta}f(\mathbf{X};\theta)}{f(\mathbf{X};\theta)}\right] \\
&= \mathrm{E}_\theta\left[W(\mathbf{X})\frac{\partial}{\partial\theta}\log(f(\mathbf{X};\theta))\right]
\end{aligned} \tag{1}$$

Sijoittamalla tähän

$$W(\mathbf{X}) = 1$$

saadaan yhtälö

$$\mathrm{E}_\theta\left[\frac{\partial}{\partial\theta}\log(f(\mathbf{X};\theta))\right] = \frac{d}{d\theta}\mathrm{E}_\theta[1] = \frac{d}{d\theta}1 = 0 \tag{2}$$

Yhdistämällä yhtälöt (1) ja (2) näemme, että

$$\begin{aligned}
&\mathrm{Cov}_\theta\left(W(\mathbf{X}),\frac{\partial}{\partial\theta}\log(f(\mathbf{X};\theta))\right) \\
&= \mathrm{E}_\theta\left[W(\mathbf{X})\frac{\partial}{\partial\theta}\log(f(\mathbf{X};\theta))\right] - \mathrm{E}_\theta\left[W(\mathbf{X})\right]\mathrm{E}_\theta\left[\frac{\partial}{\partial\theta}\log(f(\mathbf{X};\theta))\right] \\
&= \mathrm{E}_\theta\left[W(\mathbf{X})\frac{\partial}{\partial\theta}\log(f(\mathbf{X};\theta))\right] \\
&= \frac{d}{d\theta}\mathrm{E}_\theta[W(\mathbf{X})]
\end{aligned} \tag{3}$$

ja lisäksi

$$
\begin{aligned}
&\operatorname{Var}_\theta\left(\frac{\partial}{\partial\theta}\log(f(\mathbf{X};\theta))\right) \\
(4)\qquad &= \mathrm{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\log(f(\mathbf{X};\theta))\right]^2\right) - \left(\mathrm{E}_\theta\left[\frac{\partial}{\partial\theta}\log(f(\mathbf{X};\theta))\right]\right)^2 \\
&= \mathrm{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\log(f(\mathbf{X};\theta))\right]^2\right)
\end{aligned}
$$

Soveltamalla Schwarzin epäyhtälöä ja kaavoja (3) ja (4) saadaan haluttu tulos:

$$
\operatorname{Var}_\theta[W(\mathbf{X})] \geq \frac{\left[\dfrac{d}{d\theta}\mathrm{E}_\theta[W(\mathbf{X})]\right]^2}{\mathrm{E}_\theta\left(\left[\dfrac{\partial}{\partial\theta}\log(f(\mathbf{X};\theta))\right]^2\right)}
$$

■

Seuraus:

Jos edellisen lauseen oletukset pätevät, mutta lisäksi estimaattori $W(\mathbf{X})$ on *harhaton* parametrille θ, niin

$$
\operatorname{Var}_\theta[W(\mathbf{X})] \geq \frac{1}{\mathrm{E}_\theta\left(\left[\dfrac{\partial}{\partial\theta}\log(f(\mathbf{X};\theta))\right]^2\right)}
$$

Todistus:

Tulos seuraa suoraan Cramérin ja Raon epäyhtälön yleisestä muodosta, koska estimaattori $W(\mathbf{X})$ on *harhaton* parametrille θ, jos

$$
\mathrm{E}_\theta[W(\mathbf{X})] = \theta
$$

ja tällöin

$$
\frac{d}{d\theta}\mathrm{E}_\theta[W(\mathbf{X})] = \frac{d}{d\theta}\theta = 1
$$

■

Lause: Cramérin ja Raon epäyhtälö riippumattomille havainnoille

Olkoot havainnot

$$
X_1, X_2, \ldots, X_n
$$

riippumattomia ja *samaa jakaumaa noudattavia satunnaismuuttujia*. Jos lisäksi Cramérin ja Raon epäyhtälön yleistä muotoa koskevan lauseen oletukset pätevät, niin

$$
\operatorname{Var}_\theta[W(\mathbf{X})] \geq \frac{\left[\dfrac{d}{d\theta}\mathrm{E}_\theta[W(\mathbf{X})]\right]^2}{n\,\mathrm{E}_\theta\left(\left[\dfrac{\partial}{\partial\theta}\log(f(X;\theta))\right]^2\right)}
$$

jossa $f(x;\theta)$ on satunnaismuuttujien $X_1, X_2, \ldots, X_n$ yhteinen pistetodennäköisyys- tai tiheysfunktio.

Todistus:

Lause tulee todistetuksi, jos näytämme, että

$$\mathrm{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\log(f(\mathbf{X};\theta))\right]^2\right) = n\,\mathrm{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\log(f(X;\theta))\right]^2\right)$$

Nyt

$$\begin{aligned}
&\mathrm{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\log(f(\mathbf{X};\theta))\right]^2\right) \\
&\quad = \mathrm{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\log\left(\prod_{i=1}^n f(X_i;\theta)\right)\right]^2\right) \\
&\quad = \mathrm{E}_\theta\left(\left[\sum_{i=1}^n \frac{\partial}{\partial\theta}\log(f(X_i;\theta))\right]^2\right) \\
&\quad = \sum_{i=1}^n \mathrm{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\log(f(X_i;\theta))\right]^2\right) \\
&\qquad\quad + \sum_{i\neq j} \mathrm{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\log(f(X_i;\theta))\right]\left[\frac{\partial}{\partial\theta}\log(f(X_j;\theta))\right]\right) \\
&\quad = \sum_{i=1}^n \mathrm{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\log(f(X_i;\theta))\right]^2\right)
\end{aligned}$$

Viimeinen muoto seuraa siitä, että satunnaismuuttujien $X_1, X_2, \ldots, X_n$ *riippumattomuuden* takia

$$\begin{aligned}
&\mathrm{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\log(f(X_i;\theta))\right]\left[\frac{\partial}{\partial\theta}\log(f(X_j;\theta))\right]\right) \\
&\quad = \mathrm{E}_\theta\left[\frac{\partial}{\partial\theta}\log(f(X_i;\theta))\right]\mathrm{E}_\theta\left[\frac{\partial}{\partial\theta}\log(f(X_j;\theta))\right] \\
&\quad = 0\cdot 0 = 0
\end{aligned}$$

Koska satunnaismuuttujat $X_1, X_2, \ldots, X_n$ ovat samoin jakautuneita voimme kirjoittaa

$$\sum_{i=1}^n \mathrm{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\log(f(X_i;\theta))\right]^2\right) = n\,\mathrm{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\log(f(X;\theta))\right]^2\right)$$

■

Seuraus:

Jos edellisen lauseen oletukset pätevät, mutta lisäksi estimaattori $W(\mathbf{X})$ on *harhaton* parametrille θ, niin

$$\mathrm{Var}_\theta[W(\mathbf{X})] \geq \frac{1}{n\,\mathrm{E}_\theta\left(\left[\dfrac{\partial}{\partial\theta}\log(f(X;\theta))\right]^2\right)}$$

Todistus:

Tulos seuraa suoraan Cramérin ja Raon epäyhtälöstä riippumattomille havainnoille, koska estimaattori $W(\mathbf{X})$ on *harhaton* parametrille θ, jos

$$\mathrm{E}_\theta[W(\mathbf{X})] = \theta$$

ja tällöin

$$\frac{d}{d\theta}\mathrm{E}_\theta[W(\mathbf{X})] = \frac{d}{d\theta}\theta = 1$$

■

Apulause:

Jos tiheysfunktio $f(x;\theta)$ toteuttaa ehdon

$$\int_\chi \frac{\partial^2}{\partial\theta^2} f(x;\theta)dx = \frac{\partial^2}{\partial\theta^2}\int_\chi f(x;\theta)dx$$

niin

$$\mathrm{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\log(f(X;\theta))\right]^2\right) = -\mathrm{E}_\theta\left[\frac{\partial^2}{\partial\theta^2}\log(f(X;\theta))\right]$$

Todistus:

Lähdetään liikkeelle todistettavan yhtälön vasemmasta puolesta:

$$\begin{aligned}
&\mathrm{E}_\theta\left[\frac{\partial^2}{\partial\theta^2}\log(f(X;\theta))\right] \\
&\quad = \int_\chi \left[\frac{\partial^2}{\partial\theta^2}\log(f(x;\theta))\right] f(x;\theta)dx \\
&\quad = \int_\chi \left[\frac{\partial}{\partial\theta}\left(\frac{f'(x;\theta)}{f(x;\theta)}\right)\right] f(x;\theta)dx \\
&\quad = \int_\chi \left[\frac{f''(x;\theta)f(x;\theta) - [f'(x;\theta)]^2}{[f(x;\theta)]^2}\right] f(x;\theta)dx \\
&\quad = \int_\chi f''(x;\theta)dx - \int_\chi \left[\frac{f'(x;\theta)}{f(x;\theta)}\right]^2 f(x;\theta)dx
\end{aligned}$$

Väite seuraa tästä yhtälöstä, koska

$$\int_{\chi} f''(x;\theta)dx = \frac{\partial^2}{\partial\theta^2}\int_{\chi} f(x;\theta)dx = \frac{\partial^2}{\partial\theta^2}1 = 0$$

ja

$$\begin{aligned}
&\int_{\chi}\left[\frac{f'(x;\theta)}{f(x;\theta)}\right]^2 f(x;\theta)dx \\
&\quad = \int_{\chi}\left[\frac{\partial}{\partial\theta}\log(f(x;\theta))\right]^2 f(x;\theta)dx \\
&\quad = \mathrm{E}_\theta\left(\left[\frac{\partial}{\partial\theta}\log(f(x;\theta))\right]^2\right)
\end{aligned}$$

■

Voidaan osoittaa, että ns. *eksponenttiperheeseen* kuuluvat (jatkuvat) jakaumat toteuttavat apulauseen säännöllisyysehdon.

Huomautus:

Suuri osa tilastotieteen tavanomaisista jakaumista kuuluu eksponenttiperheeseen. Tällaisia ovat esimerkiksi sellaiset diskreetit jakaumat kuten *Bernoulli-jakauma*, *binomi-jakauma*, *geometrinen jakauma*, *negatiivinen binomijakauma* ja *Poisson-jakauma* sekä sellaiset jatkuvat jakaumat kuten *eksponenttijakauma*, *normaalijakauma*, *gamma-jakauma*, χ^2*-jakauma* ja *beta-jakauma*.

Seuraava lause antaa *välttämättömän ja riittävän ehdon* sille, että harhaton estimaattori saavuttaa Cramérin ja Raon lauseen alarajan.

Lause:

Olkoon

$$X_1, X_2, \ldots, X_n$$

satunnaisotos satunnaismuuttujan X jakaumasta, jonka pistetodennäköisyys- tai tiheysfunktio $f(x;\theta)$ riippuu parametrista θ. Olkoon

$$L(\theta;\mathbf{x}) = \prod_{i=1}^{n} f(x_i;\theta)$$

otoksen $X_1, X_2, \ldots, X_n$ *uskottavuusfunktio*. Olkoon

$$W(\mathbf{X}) = W(X_1, X_2, \ldots, X_n)$$

harhaton estimaattori parametrin θ funktiolle $\tau(\theta)$. Tällöin estimaattori $W(\mathbf{X})$ saavuttaa Cramérin ja Raon alarajan, jos ja vain jos on olemassa funktio $a(\theta)$ siten, että

$$\frac{\partial}{\partial\theta}\log L(\theta;\mathbf{x}) = a(\theta)[W(\mathbf{x}) - \tau(\theta)]$$

Todistus:

Esitetään *Cramérin ja Raon epäyhtälö Schwarzin epäyhtälön* mukaisessa muodossa.

Koska havainnot on tässä oletettu *riippumattomiksi* ja *samaa jakaumaa noudattaviksi satunnaismuuttujiksi*, niin saamme epäyhtälön

$$\mathrm{Cov}_\theta\left(W(\mathbf{X}), \frac{\partial}{\partial\theta}\log\left(\prod_{i=1}^{n} f(X_i;\theta)\right)\right) \le \mathrm{Var}_\theta[W(\mathbf{X})]\,\mathrm{Var}_\theta\left(\frac{\partial}{\partial\theta}\log\left(\prod_{i=1}^{n} f(X_i;\theta)\right)\right)$$

Koska W on oletettu harhattomaksi parametrille θ, niin

$$\mathrm{E}_\theta(W) = \tau(\theta)$$

Lisäksi *Cramérin ja Raon epäyhtälön* todistuksen kaavasta (2) seuraa, että

$$\mathrm{E}_\theta\left(\frac{\partial}{\partial\theta}\log\left(\prod_{i=1}^{n} f(X_i;\theta)\right)\right) = 0$$

Käyttämällä hyväksi edellä esitettyä Schwarzin epäyhtälön todistusta, näemme, että yllä esitetyssä epäyhtälössä vallitsee yhtäsuuruus, jos termi

$$W(\mathbf{x}) - \tau(\theta)$$

on (todennäköisyydellä 1) suhteessa termiin

$$\frac{\partial}{\partial\theta}\log\left(\prod_{i=1}^{n} f(X_i;\theta)\right)$$

■

Tyhjentävyys ja harhattomuus

Seuraava lause antaa menetelmän, jolla on mahdollista *parantaa* sellaisia harhattomia estimaattoreita, jotka *eivät ole* (*täys*-) *tehokkaita* eli *parhaita Cramérin ja Raon epäyhtälöm mielessä*.

Raon ja Blackwellin teoreema:

Olkoon W *harhaton estimaattori* parametrin θ funktiolle $\tau(\theta)$ ja olkoon T *tyhjentävä tunnusluku* parametrille θ. Määritellään tunnusluku

$$\phi(T) = \mathrm{E}(W \mid T)$$

Tällöin

$$\mathrm{E}_\theta[\phi(T)] = \tau(\theta)$$

ja

$$\mathrm{Var}_\theta[\phi(T)] \le \mathrm{Var}_\theta(W)$$

kaikille θ eli $\phi(T)$ on *tasaisesti* estimaattoria W *parempi harhaton estimaattori* parametrin θ funktiolle $\tau(\theta)$.

Todistus:

Todetaan ensin, että (olettaen, että ko. odotusarvot ovat olemassa)

(1) $$\mathrm{E}(X) = \mathrm{E}[\mathrm{E}(X \mid Y)]$$

ja

(2) $$\mathrm{Var}(X) = \mathrm{Var}[\mathrm{E}(X \mid Y)] + \mathrm{E}[\mathrm{Var}(X \mid Y)]$$

Yhtälöstä (1) seuraa, että

$$\tau(\theta) = \mathrm{E}_\theta(W) = \mathrm{E}_\theta[\mathrm{E}(W \mid T)] = \mathrm{E}_\theta[\phi(T)]$$

joten $\phi(T)$ on *harhaton* parametrin θ funktiolle $\tau(\theta)$.

Soveltamalla yhtälöä (2) saadaan arvio

$$\begin{aligned}\mathrm{Var}_\theta(W) &= \mathrm{Var}_\theta[\mathrm{E}(W \mid T)] + \mathrm{E}_\theta[\mathrm{Var}(W \mid T)] \\ &= \mathrm{Var}_\theta[\phi(T)] + \mathrm{E}_\theta[\mathrm{Var}(W \mid T)] \\ &\ge \mathrm{Var}_\theta[\phi(T)] \qquad \mid \mathrm{Var}(W \mid T) \ge 0\end{aligned}$$

Siten $\phi(T)$ on *tasaisesti parempi* kuin W.

Nyt on vielä näytettävä, että $\phi(T)$ on *estimaattori* eli että

$$\phi(T) = \mathrm{E}(W \mid T)$$

on otoksen funktio, joka *ei riipu* parametrista θ. Tämä seuraa siitä, että W on otoksen funktio, joka *ei riipu* parametrista θ ja siitä, että T on tyhjentävä parametrille θ, jolloin satunnaismuuttujan W ehdollinen jakauma ehdolla T *ei riipu* parametrista θ.

Siten olemme näyttäneet, että $\phi(T)$ on *tasaisesti* estimaattoria W *parempi harhaton estimaattori* parametrin θ funktiolle $\tau(\theta)$.

■

Raon ja Blackwellin teoreemasta seuraa, että ehdollistamalla mielivaltainen harhaton estimaattori jonkin *tyhjentävän* tunnusluvun suhteen saadaan alkuperäistä estimaattoria tasaisesti parempi estimaattori. Tämä merkitsee sitä, että parasta harhatonta estimaattoria etsittäessä riittää tarkastella tyhjentävien tunnuslukujen funktioita.

Tarkastellaan seuraavaa ongelmaa: Olkoon ϕ *harhaton* parametrin θ funktiolle $\tau(\theta)$ eli, että

$$\mathrm{E}_\theta[\phi] = \tau(\theta)$$

Oletetaan lisäksi, että ϕ perustuu tyhjentävään tunnuslukuun T. Miten saamme selville onko ϕ *paras* harhaton estimaattori?

Jos estimaattorin ϕ varianssi saavuttaa *Cramérin ja Raon alarajan*, tiedämme, että ϕ on *paras harhaton estimaattori*. Entä jos näin ei tapahdu? Esimerkiksi, jos ϕ^* on *toinen* harhaton estimaattori parametrin θ funktiolle $\tau(\theta)$, niin miten

$$\mathrm{E}(\phi^* \mid T)$$

suhtautuu estimaattoriin ϕ?

Osittaisen ratkaisun tähän ongelmaan antaa seuraava lause:

Lause:

Jos W on *paras harhaton estimaattori* parametrin θ funktiolle $\tau(\theta)$, niin W on *yksikäsitteinen*.

Todistus:

Olkoon W *paras harhaton estimaattori* parametrin θ funktiolle $\tau(\theta)$ ja olkoon W' jokin *toinen* paras harhaton estimaattori parametrin θ funktiolle $\tau(\theta)$.

Tarkastellaan estimaattoria

$$W^* = \frac{1}{2}(W + W')$$

Estimaattori W^* on selvästi harhaton parametrin θ funktiolle $\tau(\theta)$:

$$\mathrm{E}_\theta(W^*) = \frac{1}{2}[\mathrm{E}_\theta(W) + \mathrm{E}_\theta(W')] = \frac{1}{2}[\tau(\theta) + \tau(\theta)] = \tau(\theta)$$

Schwarzin epäyhtälöstä ja siitä, että

$$\mathrm{Var}_\theta(W) = \mathrm{Var}_\theta(W')$$

seuraa epäyhtälö

$$\begin{aligned}
\mathrm{Var}_\theta(W^*) &= \mathrm{Var}_\theta\left(\frac{1}{2}W + \frac{1}{2}W'\right) \\
&= \frac{1}{4}\mathrm{Var}_\theta(W) + \frac{1}{4}\mathrm{Var}_\theta(W') + \frac{1}{2}\mathrm{Cov}_\theta(W, W') \\
&\le \frac{1}{4}\mathrm{Var}_\theta(W) + \frac{1}{4}\mathrm{Var}_\theta(W') + \frac{1}{2}[\mathrm{Var}_\theta(W)\,\mathrm{Var}_\theta(W')]^{1/2} \\
&= \mathrm{Var}_\theta(W)
\end{aligned}$$

Jos tämä epäyhtälö on aito, niin W *ei voi olla paras harhaton estimaattori*. *Schwarzin epäyhtälöstä* seuraa, että tässä epäyhtälössä vallitsee yhtäsuuruus vain, jos on olemassa sellaiset funktiot $a(\theta)$ ja $b(\theta)$, että (todennäköisyydellä 1)

$$W' = a(\theta) + b(\theta)W$$

Siten kovarianssin yleisistä ominaisuuksista seuraa, että

$$\begin{aligned}
\mathrm{Cov}_\theta(W, W') &= \mathrm{Cov}_\theta(W, a(\theta) + b(\theta)W) \\
&= \mathrm{Cov}_\theta(W, b(\theta)W) \\
&= b(\theta)\,\mathrm{Cov}_\theta(W, W) \\
&= b(\theta)\,\mathrm{Var}_\theta(W)
\end{aligned}$$

Koska yllä esitetyssä epäyhtälössä vallitsee yhtäsuuruus, niin välttämättä

$$b(\theta) = 1$$

Lisäksi, koska

$$\mathrm{E}(W') = a(\theta) + \mathrm{E}(W) = \tau(\theta) = \mathrm{E}(W)$$

niin välttämättä

$$a(\theta) = 0$$

Koska $b(\theta) = 1$ ja $a(\theta) = 0$, niin

$$W' = W$$

ja W on yksikäsitteinen.

■

Tarkastellaan vielä seuraavaa ongelmaa: Oletetaan, että olemme löytäneet harhattoman estimaattorin. Voimmeko parantaa sitä?

Oletetaan, että W on *harhaton estimaattori* parametrin θ funktiolle $\tau(\theta)$, ts.

$$\mathrm{E}_\theta(W) = \tau(\theta)$$

Olkoon U jokin *nollan harhaton estimaattori*, ts.

$$\mathrm{E}_\theta(U) = 0$$

kaikille θ. Määritellään estimaattori

$$\phi_a = W + aU$$

jossa a on vakio. Estimaattori ϕ_a on *harhaton* parametrin θ funktiolle $\tau(\theta)$, koska

$$\mathrm{E}_\theta(\phi_a) = \mathrm{E}_\theta(W) + a\,\mathrm{E}_\theta(U) = \tau(\theta) + a \cdot 0 = \tau(\theta)$$

Voisiko estimaattori ϕ_a olla *parempi* kuin estimaattori W?

Estimaattorin ϕ_a varianssi on

$$\mathrm{Var}_\theta(\phi_a) = \mathrm{Var}_\theta(W + aU) = \mathrm{Var}_\theta(W) + 2a\,\mathrm{Cov}_\theta(W,U) + a^2\,\mathrm{Var}_\theta(U)$$

Oletetaan, että

$$\mathrm{Cov}_\theta(W,U) < 0$$

jollekin $\theta = \theta_0$. Tällöin

$$2a\,\mathrm{Cov}_\theta(W,U) + a^2\,\mathrm{Var}_\theta(U) < 0$$

jos

$$a \in (0, -2\,\mathrm{Cov}_\theta(W,U)/\mathrm{Var}_\theta(U))$$

Siten estimaattori ϕ_a on parempi kuin estimaattori W pisteessä $\theta = \theta_0$ ja *W ei voi olla paras harhaton estimaattori*.

Vastaavalla tavalla näytetään, että *W ei voi olla paras harhaton estimaattori*, jos

$$\mathrm{Cov}_\theta(W,U) > 0$$

Siten yllä kuvattu suhde estimaattorin W ja *nollan harhattoman estimaattorin* U välillä ratkaisee sen, onko W *paras harhaton estimaattori*. Itse asiassa yllä kuvattu suhde *karakterisoi* parhaat harhattomat estimaattorit.

Lause:

Jos

$$\mathrm{E}_\theta(W) = \tau(\theta)$$

niin W on *paras harhaton estimaattori* parametrin θ funktiolle $\tau(\theta)$, jos ja vain jos W korreloimaton *kaikkien* nollan harhattomien estimaattoreiden kanssa.

Todistus:

Olkoon W on *paras harhaton estimaattori* parametrin θ funktiolle $\tau(\theta)$.

Yllä esitystä seuraa, että tällöin estimaattorin W on toteutettava ehto

$$\mathrm{Cov}_\theta(W, U) = 0$$

jokaiselle estimaattorille U, joka toteuttaa ehdon

$$\mathrm{E}_\theta(U) = 0$$

Siten ehdon *välttämättömyys* on todistettu.

Oletetaan nyt, että W on harhaton estimaattori, joka on korreloimaton kaikkien nollan harhattomien estimaattoreiden kanssa.

Olkoon W' mielivaltainen estimaattori, joka toteuttaa ehdon

$$\mathrm{E}_\theta(W') = \mathrm{E}_\theta(W) = \tau(\theta)$$

Näytämme, että estimaattori W on *parempi* kuin estimaattori W'.

Kirjoitetaan

$$W' = W + (W' - W)$$

ja määrätään estimaattorin W' varianssi:

$$\begin{aligned}\mathrm{Var}_\theta(W') &= \mathrm{Var}_\theta(W) + \mathrm{Var}_\theta(W' - W) + 2\,\mathrm{Cov}_\theta(W, W' - W) \\ &= \mathrm{Var}_\theta(W) + \mathrm{Var}_\theta(W' - W)\end{aligned}$$

koska $W' - W$ on nollan harhaton estimaattori ja olemme olettaneet, että kaikki nollan harhattomat estimaattorit ovat korreloimattomia estimaattorin W kanssa.

Koska

$$\mathrm{Var}_\theta(W' - W) \geq 0$$

niin

$$\mathrm{Var}_\theta(W') \geq \mathrm{Var}_\theta(W)$$

Koska W' oli mielivaltainen harhaton estimaattori, tästä seuraa, että W on *paras harhaton estimaattori* ja ehdon *riittävyys* on todistettu.

■

On syytä huomata, että edellisen lauseen käyttökelpoisuus on rajoitettu, koska on usein hyvin vaikeata näyttää, että annettu estimaattori on korreloimaton *kaikkien* nollan harhattomien estimaattoreiden kanssa.

Sen sijaan tarkasteltavan jakaumaperheen *täydellisyys* (ks. lukua 2) takaa sen, että nollalla *ei ole* muita harhattomia estimaattoreita kuin nolla itse. Koska aina pätee, että

$$\mathrm{Cov}_{\theta}(W, 0) = 0$$

niin W on *paras harhaton estimaattori*, jos se on *harhaton*.

Lause:

> Olkoon T *täydellinen* ja *tyhjentävä* estimaattori parametrille θ ja olkoon $\phi(T)$ mielivaltainen estimaattori, joka perustuu vain estimaattoriin T. Tällöin $\phi(T)$ on odotusarvonsa *paras harhaton estimaattori* ja lisäksi $\phi(T)$ on *yksikäsitteinen*.