4.1 Uskottavuusfunktio

Kun aineisto $\mathbf{y}$ on havaittu, ja havaittua arvoa käytetään funktion $f(\mathbf{y};\theta)$ ensimmäisenä argumenttina, niin parametriavaruudella määriteltyä funktiota

$$\theta \mapsto f(\mathbf{y};\theta)$$

kutsutaan uskottavuusfunktioksi (engl. *likelihood function*).
Sitä merkitään

$$L(\theta) = f(\mathbf{y};\theta).$$

Joskus tahdotaan kirjata näkyviin, että uskottavuusfunktio riippuu myös aineistosta $\mathbf{y}$, ja tällöin voidaan käyttää merkintää

$$L(\theta;\mathbf{y}) = f(\mathbf{y};\theta).$$

Haluttaessa voidaan sanoa tarkemmin, että kyseessä on havaintoa $\mathbf{y}$ vastaava parametrin θ uskottavuusfunktio.

Kommentteja uskottavuusfunktion määritelmästä

- Uskottavuusfunktion yhteydessä θ on vapaa muuttuja, ei parametrin todellinen arvo.
- Funktio

$$\mathbf{y} \mapsto f(\mathbf{y}; \theta)$$

  on satunnaisvektorin $\mathbf{Y}$ yptnf tai ytf.
- Uskottavuusfunktiossa argumentti $\mathbf{y}$ kiinnitetään sijoittamalla siihen havaitut arvot. Tätä lauseketta tarkastellaan parametrin funktiona. Uskottavuusfunktio saa pisteessä θ arvon

$$L(\theta) = f(\mathbf{y}; \theta), \qquad \theta \in \Theta.$$

- Uskottavuusfunktio ei ole pistetodennäköisyysfunktio eikä tiheysfunktio.

Pallot kulhossa — uskottavuusfunktio

Esimerkki 1. Oletetaan pallot kulhossa -esimerkissä, että kulhossa on $N = 5$ palloa ja että nostot tehdään palauttaen ja että tulokset ovat $\mathbf{y} = (1, 0, 0, 0, 1, 0, 0)$. Tällöin valkoisten pallojen lukumäärä $n = 7$ nostossa (eli onnistumisten lukumäärä) on 2, ja uskottavuusfunktio on

$$L(\theta) = \left(\frac{\theta}{5}\right)^2 \left(1 - \frac{\theta}{5}\right)^5, \qquad \theta = 0, 1, 2, 3, 4 \text{ tai } 5,$$

jossa onnistumistodennäköisyys on θ/N, joka on valkoisten pallojen suhteellinen osuus kulhossa. △

Nasta purkissa — uskottavuusfunktio

Jos sama havainto $\mathbf{y} = (1, 0, 0, 0, 1, 0, 0)$ saadaan nasta purkissa -esimerkissä, niin tällöin onnistumistodennäköisyys on θ, ja uskottavuusfunktio on

$$L(\theta) = \theta^2 (1 - \theta)^5.$$

Tässä parametriavaruudeksi ja uskottavuusfunktion määrittelyjoukoksi voidaan valita joko suljettu väli $[0, 1]$ tai avoin väli $(0, 1)$.

Laajennettu määritelmä

- Laajennamme määritelmää sillä tavalla, että uskottavuusfunktioksi kelpuutetaan myös mikä tahansa muotoa

$$L(\theta) = k(\mathbf{y})\, f(\mathbf{y}; \theta) \tag{1}$$

  oleva lauseke, jossa positiivinen vakio $k(\mathbf{y}) > 0$ saa riippua aineistosta $\mathbf{y}$, mutta ei uskottavuusfunktion argumentista θ.
- Tilastotieteessä suositaan sellaisia menetelmiä, joiden kannalta on saman tekevää, mitä verrannollisuuskerrointa $k = k(\mathbf{y}) > 0$ uskottavuusfunktion määritelmässä käytetään.
- Esimerkiksi uskottavuusfunktion maksimikohta pysyy samana vaikka kerrointa $k > 0$ muutetaan.
- Uskottavuusfunktion voidaan ajatella sisältävän kaiken aineistoon liittyvän informaation parametrin θ arvosta.

Uskottavuusfunktion logaritmi

Määritelmä (Logaritminen uskottavuusfunktio)

Uskottavuusfunktion logaritmia

$$\ell(\theta) = \log L(\theta)$$

kutsutaan logaritmiseksi uskottavuusfunktioksi tai uskottavuusfunktion logaritmiksi tai log-uskottavuusfunktioksi (engl. *log-likelihood*). Tässä log tarkoittaa luonnollista logaritmia.

Kun tarkastellaan logaritmista uskottavuusfunktiota $\ell(\theta) = \log L(\theta)$, niin tavallisesti oletetaan, että $L(\theta) > 0$ koko parametriavaruudessa, jolloin $\ell(\theta)$ on hyvin määritelty reaalifunktio: $\log(0)$ ei ole reaaliluku. Vaihtoehtoinen tapa selvitä tästä pulmasta on sopia, että $\log(0) = -\infty$, joka on pienempi kuin mikään reaaliluku.

Miksi logaritmi?

- Jos uskottavuusfunktio on tulomuotoa, niin logaritmin otto muuttaa sen summaksi, sillä

$$\log\left(\prod_{i=1}^{n} f_{Y_i}(y_i;\theta)\right) = \sum_{i=1}^{n} \log\left(f_{Y_i}(y_i;\theta)\right).$$

- Tässä sovellettiin kaavaa

$$\log(ab) = \log(a) + \log(b), \qquad \text{kun } a > 0 \text{ ja } b > 0.$$

- Tietokoneella laskettaessa logaritmointi on tärkeää, sillä uskottavuusfunktiossa esiintyvät tulon termit ovat usein erittäin pieniä positiivisia lukuja, jolloin itse uskottavuusfunktion arvoksi saattaa tietokoneohjelmassa tulla tasan nolla, vaikka kyseessä olisi aidosti positiivinen luku. Logaritmin ottaminen ratkaisee tavallisesti tämän ongelman.

4.2 Suurimman uskottavuuden estimaatti

Tunnetuin frekventistisen päättelyn estimointiperiaate on ns. suurimman uskottavuuden, eli SU-periaate (engl. *maximum likelihood, ML*), jonka mukaan parametrin parhaana estimaattina pidetään sitä parametriavaruuden arvoa $\hat{\theta}$, joka maksimoi uskottavuusfunktion. Sitä kutsutaan suurimman uskottavuuden estimaatiksi (eli SU-estimaatiksi) (engl. *maximum likelihood estimate, ML estimate, MLE*).

Kaava SU-estimaatille

- SU-estimaatti voidaan esittää kaavalla

$$\hat{\theta} = \arg\max_{\theta \in \Theta} L(\theta). \tag{2}$$

- Merkintä $\arg\max L(\theta)$ tarkoittaa lausekkeen $L(\theta)$ maksimoivaa argumenttia (ts. maksimipistettä).
- Merkintä $\max L(\theta)$ tarkoittaa lausekkeen $L(\theta)$ maksimiarvoa.
- Kun näitä merkintöjä käytetään, niin tällöin hiljaisesti oletetaan, että parametriavaruudessa on olemassa yksikäsitteinen maksimipiste $\hat{\theta}$, jolle

$$L(\hat{\theta}) \geq L(\theta), \quad \text{kaikille } \theta \in \Theta.$$

SU-estimaatti on log-uskottavuuden maksimipiste

Koska logaritmi on aidosti kasvava funktio, on uskottavuusfunktiolla $L(\theta)$ ja logaritmisella uskottavuusfunktiolla $\ell(\theta)$ samat maksimipisteet. Tämän takia SU-estimaatti voidaan yhtä hyvin määrittää kaavalla

$$\hat{\theta} = \arg\max_{\theta \in \Theta} \ell(\theta). \tag{3}$$

Miksi SU-periaate on järkevä

- Mikäli aineistoa vastaavan satunnaisvektorin $\mathbf{Y}$ jakauma on diskreetti, niin SU-estimaatti on se parametrialueen piste, joka tekee havaitun aineiston (mallin puitteissa) mahdollisimman todennäköiseksi, eli

$$P_{\hat{\theta}}(\mathbf{Y} = \mathbf{y}) \geq P_{\theta}(\mathbf{Y} = \mathbf{y}), \qquad \text{kaikilla } \theta \in \Theta.$$

- Tuntuu järkevältä suosia sellaisia parametrin arvioita, joille havainnot ovat todennäköisiä eikä sellaisia, joille ne ovat epätodennäköisiä.
- Jatkuvan yhteisjakauman tapauksessa motivointi on monimutkaisempi, sillä jatkuvalle jakaumalle

$$P_{\theta}(\mathbf{Y} = \mathbf{y}) = 0 \qquad \text{kaikilla } \theta.$$

Järkevän motivoinnin saa aikaan sopivasti integroimalla (ks. moniste).

Jatkoa esimerkille pallot kulhossa

Esimerkki 2. Valkoisia palloja nostettiin 2 ja nostoja oli yhteensä $n = 7$. Valkoisten pallojen lukumäärä θ on yksi luvuista $0, 1, 2, 3, 4$ tai 5, joten uskottavuusfunktio on

$$L(\theta) = \left(\frac{\theta}{5}\right)^2 \left(1 - \frac{\theta}{5}\right)^5$$

$$= \begin{cases} 0, & \text{jos } \theta = 0, \\ 1024/5^7, & \text{jos } \theta = 1, \\ 972/5^7, & \text{jos } \theta = 2, \\ 288/5^7, & \text{jos } \theta = 3, \\ 16/5^7, & \text{jos } \theta = 4, \\ 0, & \text{jos } \theta = 5. \end{cases}$$

SU-estimaatti on $\hat{\theta} = 1$. $\triangle$

Miksi uusi termi: uskottavuus?

- SU-estimaatti $\hat{\theta}$ on edellisessä esimerkissä se arvo, joka tekee havainnot (mallin puitteissa) mahdollisimman todennäköisiksi.
- Olisi väärin väittää, että $\hat{\theta}$ eli uskottavin parametrin olisi parametrin todennäköisin arvo.
- Frekventistisen tilastotieteen puitteissa tällainen lausuma on mieltä vailla, koska parametrin arvoa koskevia todennäköisyyksiä ei frekventistisessä mallissa ole määriteltynä.
- Tästä syystä Fisher otti käyttöön uuden termin uskottavuus.
- Arkikielessä probability ja likelihood ovat suunnilleen synonyymejä, mutta tilastotieteen termeinä ne ovat kaikkea muuta kuin synonyymeja.

SU-estimaatin hakeminen, kun parametriavaruus on diskreetti

- Pallot kulhossa -esimerkissä parametriavaruus on diskreetti.
- Koska se ei esimerkissä koostu kovin monesta pisteestä, pystymme laskemaan uskottavuusfunktion arvon jokaisessa parametriavaruuden pisteessä.
- Tämän jälkeen valitsemme sen pisteen, jossa suurin arvo saavutetaan.

SU-estimaatin hakeminen, kun parametriavaruus on jatkuva

- Jos parametriavaruus on jatkuva, niin maksimoinnissa käytetään hyväksi derivaattaa. Tarkastelemme ensin yhden parametrin θ tapausta.
- Usein SU-estimaatti saadaan ratkaistua etsimällä logaritmisen uskottavuusfunktion derivaatan nollakohdat, eli ratkaisemalla ns. uskottavuusyhtälö

$$\ell'(\theta) = 0.$$

- Tämä perustuu siihen, että mikäli (jatkuvasti derivoituva) yhden muuttujan funktio saavuttaa maksimin jossakin määrittelyjoukkonsa sisäpisteessä, niin kyseisessä pisteessä funktion derivaatta saa arvon nolla.

SU-estimaatin haku, jatkuva parametriavaruus

- Tämän jälkeen pitää funktion ℓ kriittisistä pisteistä eli derivaatan nollakohdista valita ne, jotka ovat maksimipisteitä.
- Tämä onnistuu joko tarkastelemalla derivaatan merkkikaaviota tai ℓ:n toista derivaattaa (jos $\ell'(\theta_0) = 0$ ja $\ell''(\theta_0) < 0$, niin θ_0 on maksimipiste).
- Lisäksi pitää kiinnittä huomiota (log-)uskottavuusfunktion käyttäytmiseen, kun lähestytään parametriavaruuden reunapisteitä, sillä maksimipiste voi löytyä parametriavaruuden reunalta.
- Tällä tavalla löydetään kaikki paikalliset maksimipisteet, ja lopulta niistä valitaan globaali maksimi, eli se piste, jossa ℓ saavuttaa suurimman arvonsa koko parametriavaruudessa.

SU-estimaatin haku, jatkuva parametriavaruus

- Jos parametreja on useita, niin kaikki ℓ:n ensimmäisen kertaluvun osittaisderivaatat häviävät maksimipisteessä, joten tällöin uskottavuusyhtälö on yhtälöryhmä.
- Esim. kahden parametrin $\theta = (\mu, \phi)$ tapauksessa pitäisi etsiä ne pisteet, joissa molemmat yhtälöt

$$\frac{\partial}{\partial \mu}\ell(\mu, \phi) = 0, \qquad \frac{\partial}{\partial \phi}\ell(\mu, \phi) = 0$$

  toteutuvat. Kriittisen pisteen laadun voi tarkistaa toisen kertaluvun osittaisderivaattojen avulla.
- Monimutkaisemmissa tapauksissa maksimipisteitä ei pystytä määrittämään algebrallisesti, vaan ne haetaan tietokoneella soveltamalla jotakin numeerista maksimointimenetelmää.

4.3 SU-estimaatti binomikokeessa

- Johdamme seuraavaksi SU-estimaatin kaavan binomikokeen tapauksessa silloin, kun n toistossa onnistutaan x kertaa, ja onnistumistodennäköisyys yhdessä toistossa on θ.
- Oletamme, että parametriavaruus Θ on joko avoin väli $(0, 1)$ tai suljettu väli $[0, 1]$. Nasta purkissa -esimerkissä on tällainen tilanne.
- Uskottavuusfunktio on

$$L(\theta) = \theta^x (1 - \theta)^{n-x}, \qquad \theta \in \Theta.$$

- Tässä ei tarvitse tietää, missä järjestyksessä onnistumiset ja epäonnistumiset sattuivat.

SU-estimaatti binomikokeessa, kun $1 \leq x \leq n-1$

- Logaritminen uskottavuusfunktio on

$$\ell(\theta) = \log\left(\theta^x (1-\theta)^{n-x}\right) = x \log\theta + (n-x)\log(1-\theta),$$

joka on hyvin määritelty, kun $0 < \theta < 1$.

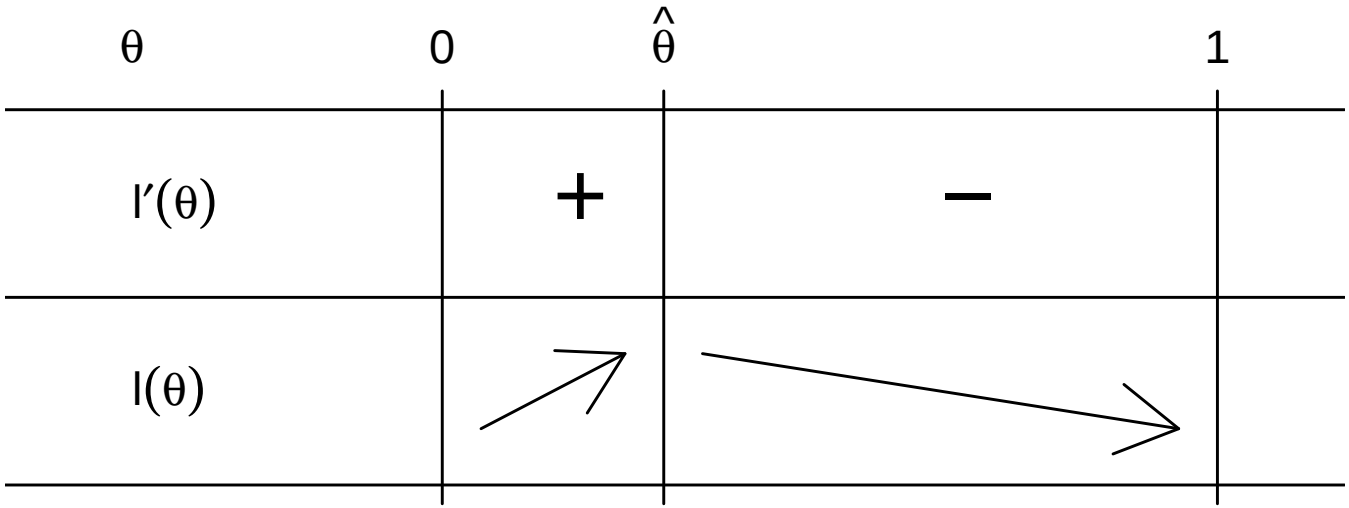
- Ratkaisemme logaritmisen uskottavuusfunktion derivaatan nollakohdat. Kun $0 < \theta < 1$, niin

$$\ell'(\theta) = \frac{x}{\theta} - \frac{n-x}{1-\theta} = \frac{x - n\theta}{\theta(1-\theta)}$$

Derivaatan ainoa nollakohta on $\hat{\theta} = x/n$, ja kyseessä on maksimipiste, sillä derivaatan merkki vaihtuu siinä positiivisesta negatiiviseksi. (Nimittäjä $\theta(1-\theta)$ on positiivinen.)

Log-uskottavuusfunktion kulkukaavio

Kuva: Logaritmisen uskottavuusfunktion kulkukaavio binomikokeessa, kun $n = 7$ toistossa havaitaan $x = 2$ onnistumista.



SU-estimaatti binomikokeessa, kun $x = n$

- Tapauksessa $x = n$ uskottavuusfunktio on

$$L(\theta) = \theta^n,$$

  ja tämä on selvästi aidosti kasvava funktio välillä $(0, 1)$. Jos parametriavaruus on $[0, 1]$, niin SU-estimaatti on $\hat{\theta} = 1 = x/n$.
- SU-estimaatti ei tässä tapauksessa löydy derivaatan nollakohdasta, vaan parametriavaruuden reunalta.
- Jos parametriavaruus kuitenkin on avoin väli $(0, 1)$, niin tällöin joudumme toteamaan, että SU-estimaattia ei ole olemassa, koska uskottavuusfunktio ei saavuta missään parametriavaruuden pisteessä maksimiarvoaan.

SU-estimaatti binomikokeessa, kun $\Theta = [0, 1]$

- Tapauksessa $x = 0$ nädään vastaavasti, että SU-estimaatti on $\hat{\theta} = 0 = x/n$, mikäli parametriavaruus on $[0, 1]$. Jos parametriavaruus kuitenkin on $(0, 1)$, niin SU-estimaattia ei ole olemassa.
- Mikäli binomikokeessa tahdotaan käyttää SU-estimointia, niin tästä syystä on kätevää valita parametriavaruudeksi suljettu väli $[0, 1]$.
- Tällöin SU-estimaatti saadaan kaikissa tapauksissa kaavalla

$$\hat{\theta} = \frac{x}{n} \tag{4}$$

eli SU-estimaatti on onnistumisten x suhteellinen osuus n toistossa.

Keskivirhe onnistumisten suhteelliselle osuudelle binomikokeessa

Olemme jo aikaisemmin nähneet, että vastaava estimaattori on harhaton ja että sen (otantajakauman) varianssi on

$$\frac{1}{n}\,\theta\,(1-\theta).$$

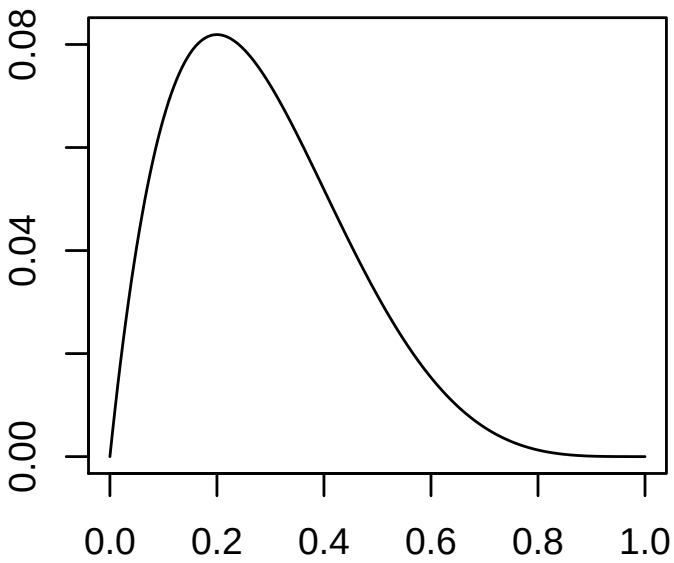
SU-estimaatin $\hat{\theta}$ keskivirhe voidaan laskea kaavalla

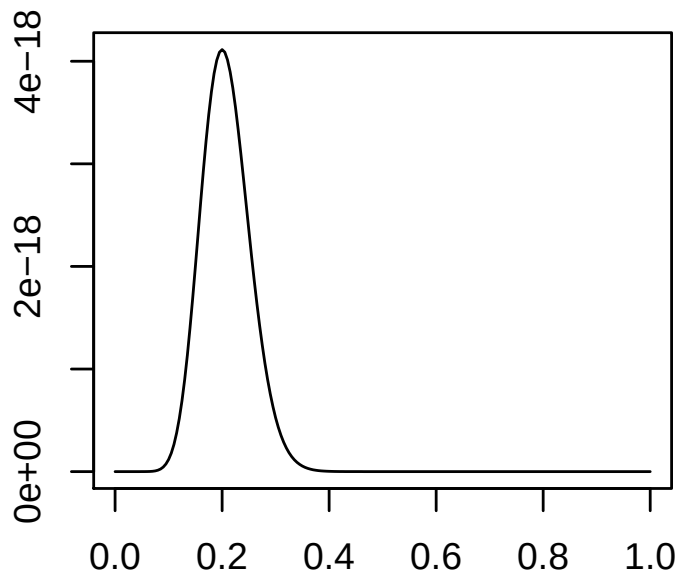
$$\sqrt{\frac{1}{n}\,\hat{\theta}\,(1-\hat{\theta})} \tag{5}$$

jossa tuntematon parametri θ on korvattu sen estimaatilla $\hat{\theta}$.

Kuva uskottavuusfunktiosta binomikokeessa

Kuva: Uskottavuusfunktio binomikokeessa kahdella eri otoskoolla, kun parametriavaruus on jatkuva. Vasemmalla $n = 5$ ja oikealla $n = 80$. Molemmissa tapauksissa onnistumisten suhteellinen osuus $x/n = 0.2$.





4.4 Normaalijakauman parametrien estimointi

- Mallinnamme aineiston $\mathbf{y} = (y_1, \ldots, y_n)$ siten, että vastaavat satunnaismuuttujat $Y_1, \ldots, Y_n$ ovat satunnaisotos normaalijakaumasta $N(\mu, \sigma^2)$.
- Ts. oletamme, että satunnaismuuttujat Y_i ovat riippumattomia, ja kukin niistä noudattaa normaalijakaumaa $N(\mu, \sigma^2)$.
- Tässä $\mu \in \mathbb{R}$ ja $\sigma^2 > 0$ voivat molemmat olla tuntemattomia parametreja, tai sitten toinen niistä voi olla tunnettu vakio ja toinen tuntematon parametri.

Normaalijakauma $N(\mu, \sigma^2)$

- Kunkin yksittäisen satunnaismuuttujan Y_i tiheysfunktio on

$$g(y; \mu, \sigma^2) = \frac{1}{\sqrt{2\,\pi\,\sigma^2}} \exp\left(-\frac{1}{2\,\sigma^2}(y-\mu)^2\right).$$

- Tässä exp tarkoittaa eksponenttifunktiota, eli

$$\exp(x) = \mathrm{e}^x, \qquad \text{kun } x \in \mathbb{R}.$$

- Parametrien μ ja σ^2 merkitys on se, että kullakin i

$$EY_i = \mu, \qquad \operatorname{var} Y_i = \sigma^2.$$

- Parametri μ on paitsi normaalijakauman $N(\mu, \sigma^2)$ odotusarvo, myös sen moodi ja mediaani. Normaalijakauman tiheysfunktio on symmetrinen odotusarvon suhteen.
- Varianssiparametri kuvaa sitä, miten tiukasti jakauma on keskittynyt keskikohtansa ympärille: mitä pienempi varianssi, sitä keskittyneempi jakauma.

Satunnaisotos normaalijakaumasta $N(\mu, \sigma^2)$

Havaintosatunnaisvektorin $\mathbf{Y}$ yhteistiheysfunktio on

$$\begin{aligned} f(\mathbf{y}; \mu, \sigma^2) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\,\pi\,\sigma^2}} \exp\left(-\frac{1}{2\,\sigma^2}(y_i - \mu)^2\right). \\ &= (2\,\pi)^{-n/2}\,(\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\,\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2\right) \end{aligned} \tag{6}$$

Johdossa sovellettiin tuttua kaavaa

$$\mathrm{e}^{a}\,\mathrm{e}^{b} = \mathrm{e}^{a+b}, \quad \text{eli} \quad \exp(a)\,\exp(b) = \exp(a+b),$$

joka pätee kaikille reaaliluvuille a ja b.

Logaritminen uskottavuusfunktio

Jätetään uskottavuusfunktiosta 2π:n potenssit pois, jolloin kaavasta (6) saadaan havaintoa **y** vastaavalle logaritmiselle uskottavuusfunktiolle lauseke

$$\begin{aligned} \ell(\mu, \sigma^2) &= \log f(\mathbf{y}; \mu, \sigma^2) \\ &= -\frac{n}{2}\log(\sigma^2) - \frac{1}{2\,\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2 \end{aligned} \tag{7}$$

Ylläolevassa kaavassa voidaan neliöiden summa hajottaa kahteen osaan

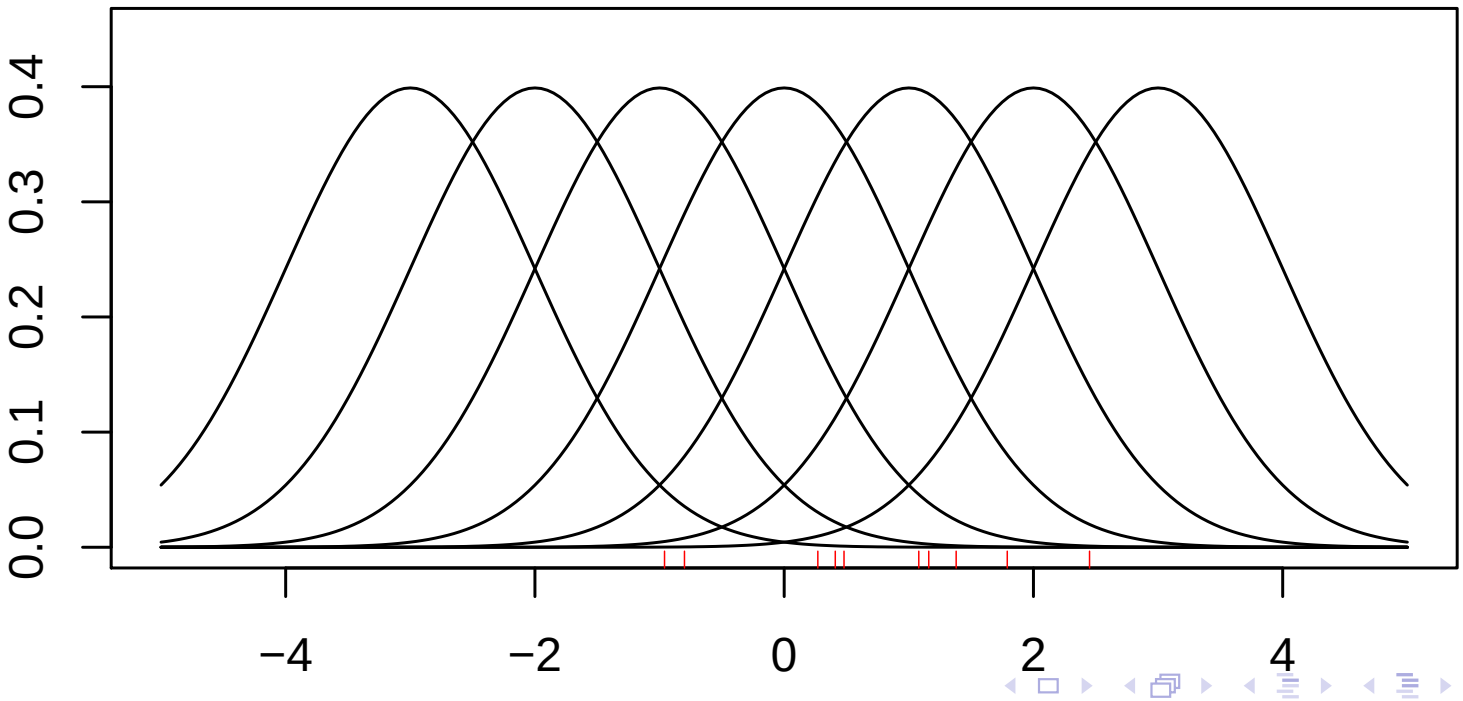
$$\sum_{i=1}^{n}(y_i - \mu)^2 = \sum_{i=1}^{n}(y_i - \bar{y})^2 + n\,(\bar{y} - \mu)^2, \tag{8}$$

jossa $\bar{y}$ on lukujen y_i otoskeskiarvo,

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{9}$$

4.4.1 Varianssi tunnettu

Jos normaalijakaumaperheessä varianssi σ^2 on tunnettu luku, niin mallissa on jäljellä vain yksi tuntematon parametri μ. Alla näytetään muutama $N(\mu, \sigma^2)$-jakaumaperheen tiheysfunktio, kun $\sigma^2 = 1$. Yksi tällainen tiheysfunktio pitää valita kuvamaan x-akselille lyhyillä viivoilla merkittyä aineistoa.



SU-estimaatti

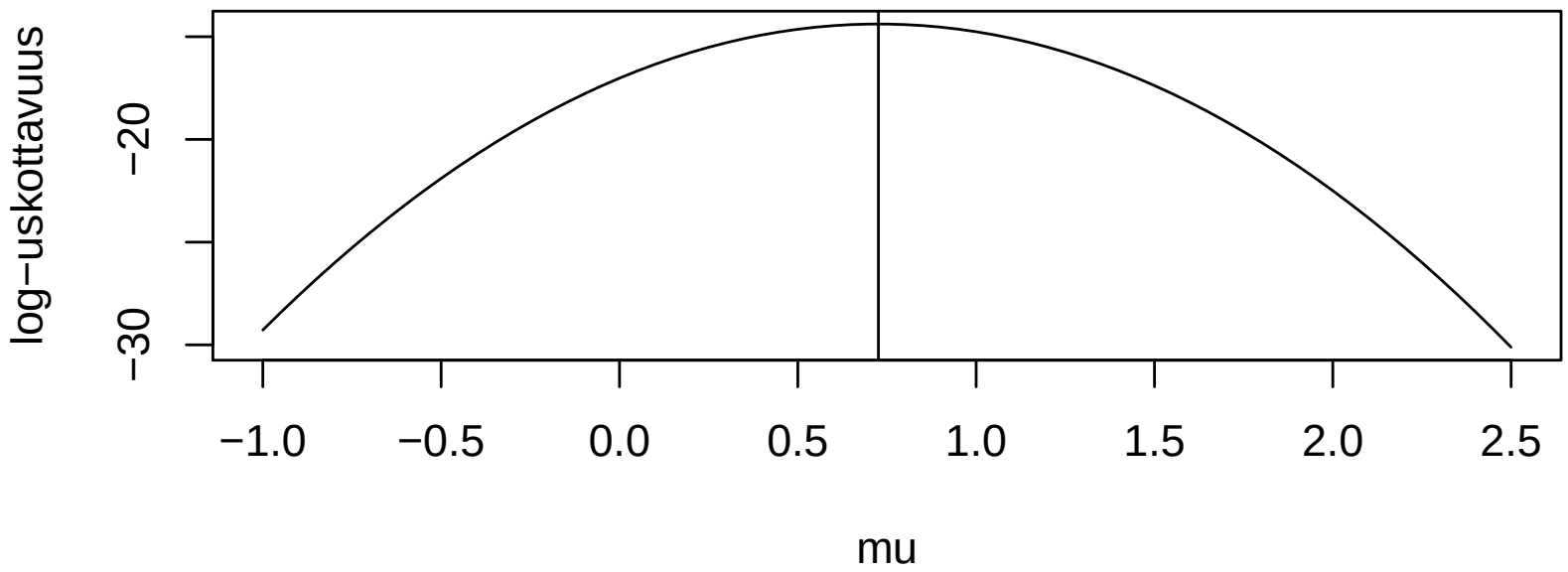
- Logaritminen uskottavuusfunktio on kaavojen (7) ja (8) mukaan

$$\ell(\mu) = -\frac{n}{2}\log(\sigma^2) - \frac{1}{2\,\sigma^2}\left(\sum_{i=1}^{n}(y_i - \bar{y})^2 + n\,(\bar{y} - \mu)^2\right)$$
$$= \text{vakio} - \frac{n}{2\,\sigma^2}(\bar{y} - \mu)^2$$

- Vakioksi merkitty termi ei riipu μ:stä.
- Koska kerroin $n/(2\,\sigma^2)$ on positiivinen, niin logaritminen uskottavuusfunktio maksimoituu täsmälleen silloin, kun lauseke $(\bar{y} - \mu)^2$ minimoituu, eli silloin, kun $\mu = \hat{\mu} = \bar{y}$.

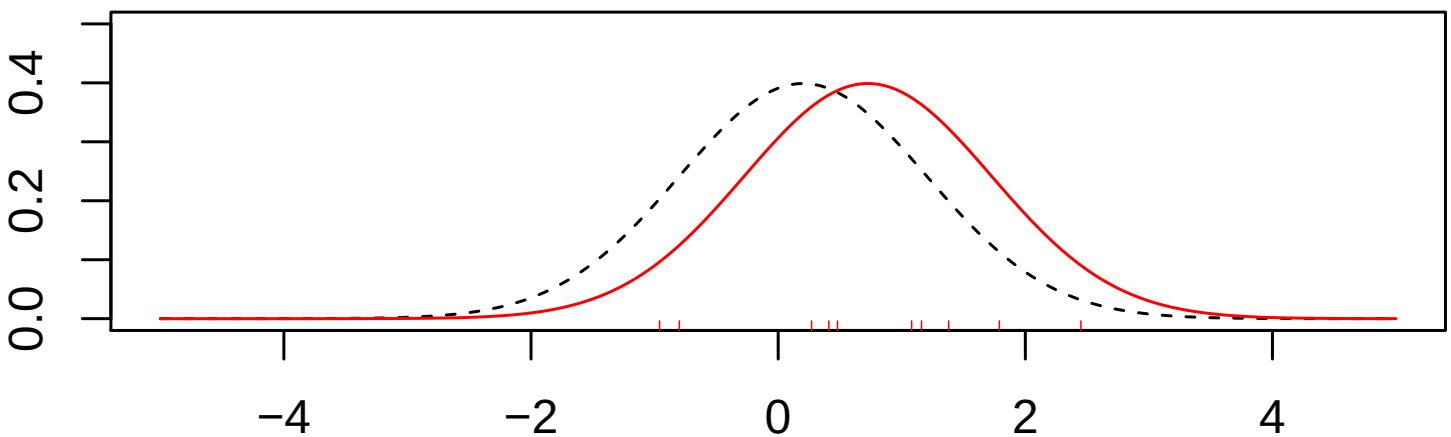
Kuva log-uskottavuusfunktiosta

Parametrin μ logaritminen uskottavuusfunktio edellisen kuvan aineistolle. SU-estimaatti on merkitty pystyviivalla.



Estimoitu normaalijakauma

SU-estimaattia vastaavan normaalijakauman tiheysfunktio. Todellinen populaation tiheysfunktio on merkitty katkoviivalla, ja sen piirtäminen käytännön tilanteessa olisi mahdotonta.



Estimaatin keskivirhe

Kun varianssi on tunnettu, SU-estimaatti on *otoskeskiarvo* (engl. *sample mean; average*), eli

$$\hat{\mu} = \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i. \tag{10}$$

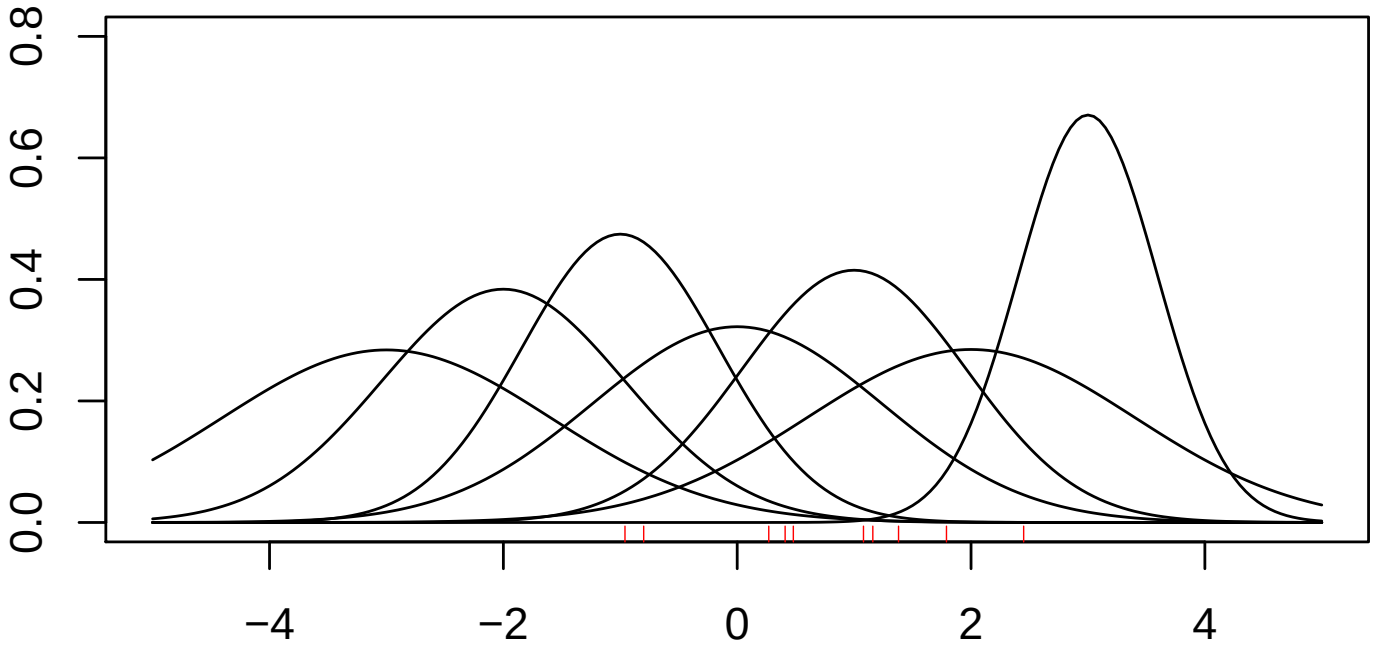
Vastaava estimaattori $\frac{1}{n} \sum_{i=1}^{n} Y_i$ on harhaton, ja sen varianssi on

$$\mathrm{var}_{\mu} \left(\frac{1}{n} \sum_{i=1}^{n} Y_i \right) = \frac{1}{n} \sigma^2,$$

joka on tässä mallissa tunnettu vakio. Tämän luvun neliöjuuri on SU-estimaatin keskivirhe.

4.4.2 Molemmat parametrit tuntemattomia

Nyt normaalijakauman $N(\mu, \sigma^2)$ molemmat ovat tuntemattomia, joten satunnaisvektorin $\mathbf{Y}$ jakauman kiinnittämiseksi pitäisi tuntea parametrivektorin $\boldsymbol{\theta} = (\mu, \sigma^2)$ arvo. Alla näytetään muutama $N(\mu, \sigma^2)$-jakaumaperheen tiheysfunktio. Yksi tällainen tiheysfunktio pitää jälleen valita aineiston perusteella.



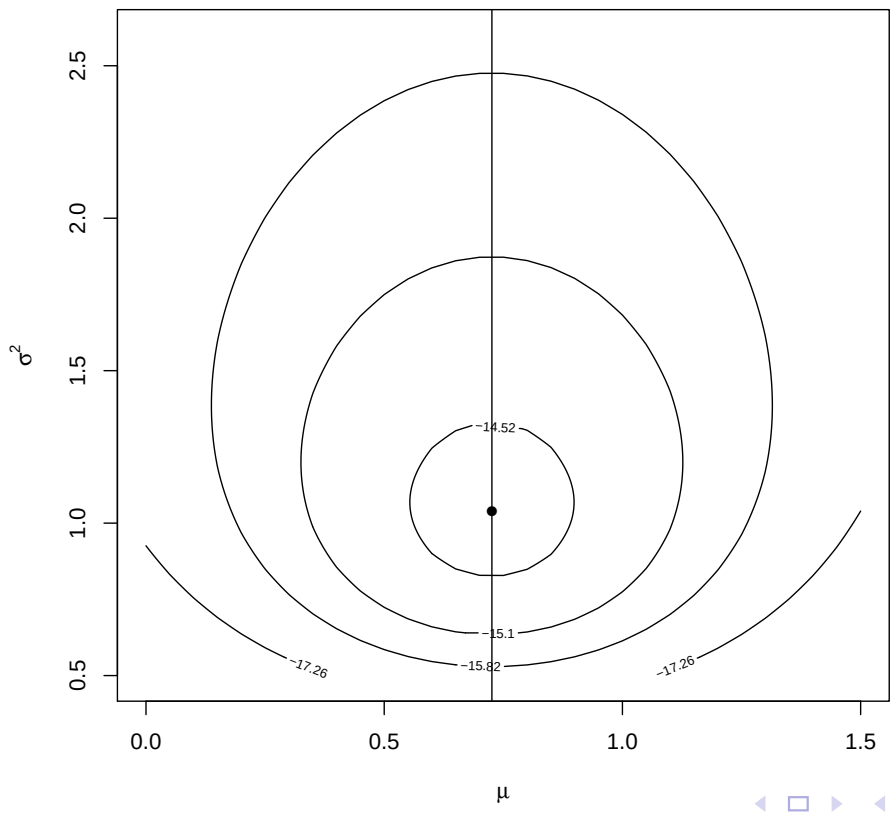
Log-uskottavuus

Logaritminen uskottavuusfunktio on kaavojen (7) ja (8) mukaan

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\log(\sigma^2) - \frac{1}{2\,\sigma^2}\sum_{i=1}^{n}(y_i - \bar{y})^2 - \frac{n}{2\,\sigma^2}(\bar{y} - \mu)^2$$

Kuva log-uskottavuusfunktiosta

Parametrivektorin (μ, σ^2) logaritminen uskottavuusfunktio $\ell(\mu, \sigma^2)$ esitettynä tasa-arvokäyriensä avulla. SU-piste on merkitty pallolla. Funktion $\mu \mapsto \ell(\mu, \sigma^2)$ maksimi löytyy pisteestä $\mu = \bar{y}$, joka on osoitettu suoralla.



SU-estimaatti μ:lle

- Logaritminen uskottavuusfunktio

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\log(\sigma^2) - \frac{1}{2\,\sigma^2}\sum_{i=1}^{n}(y_i - \bar{y})^2 - \frac{n}{2\,\sigma^2}(\bar{y} - \mu)^2$$

  riippuu μ:n arvosta vain sen viimeisen termin kautta.
- Oli varianssiparametrin $\sigma^2 > 0$ arvo mikä tahansa, niin funktion $\mu \mapsto \ell(\mu, \sigma^2)$ maksimoi arvo $\hat{\mu} = \bar{y}$ eli otoskeskiarvo.

SU-estimaatti σ^2:lle

Kahden muuttujan funktion $\ell(\mu, \sigma^2)$ maksimointi saadaan edellisen ansiosta palautettua yhdestä muuttujasta riippuvan funktion u maksimointitehtäväksi, jossa

$$\begin{aligned} u(\sigma^2) &= \max_{\mu} \ell(\mu, \sigma^2) = \ell(\bar{y}, \sigma^2) \\ &= -\frac{n}{2} \log(\sigma^2) - \frac{1}{2\,\sigma^2} \sum_{i=1}^{n} (y_i - \bar{y})^2 \end{aligned}$$

Tämän funktion maksimipiste on

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

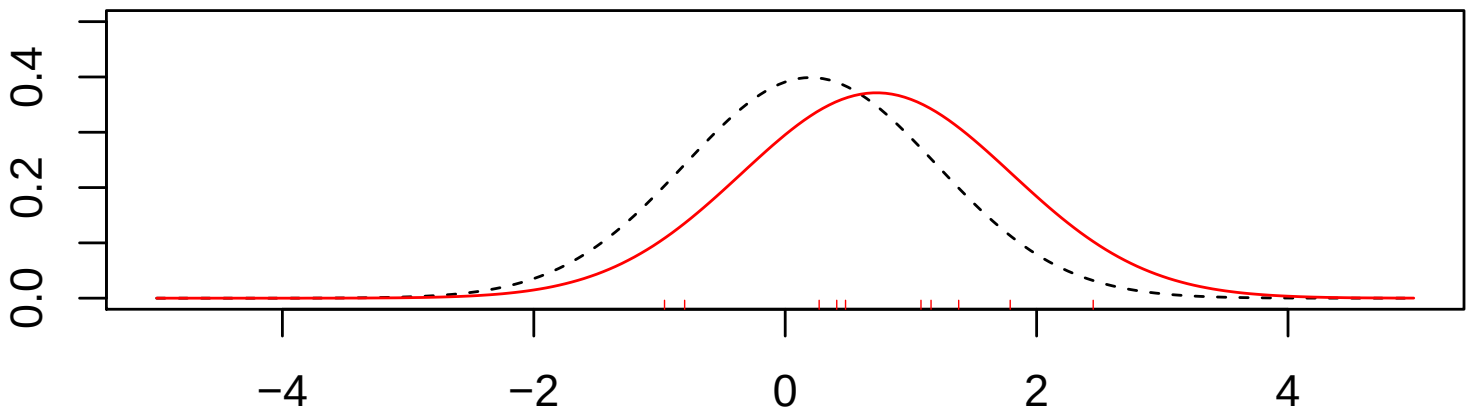
SU-estimaatti, kun molemmat parametrit tuntemattomia

Olemme saaneet selville, että parametrivektorin (μ, σ^2)
SU-estimaatti on

$$\hat{\mu} = \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i, \qquad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y})^2. \tag{11}$$

Estimoitu normaalijakauma

SU-estimaattia $(\hat{\mu}, \hat{\sigma}^2)$ vastaavan normaalijakauman tiheysfunktio. Todellinen populaation tiheysfunktio on merkitty katkoviivalla, ja sen piirtäminen käytännön tilanteessa olisi mahdotonta.



Onko SU-estimaattori kaikin puolin hyväksyttävä?

- Estimaattori

$$\hat{\mu}(\mathbf{Y}) = \frac{1}{n}\sum_{i=1}^{n} Y_i = \bar{Y}$$

  on harhaton.

- Varianssiparametrin SU-estimaattori

$$\hat{\sigma}^2(\mathbf{Y}) = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

  on harhainen, sillä

$$E_{(\mu,\sigma^2)}[\hat{\sigma}^2(\mathbf{Y})] = \frac{n-1}{n}\sigma^2.$$

- Harhan saa korjattua yksinkertaisella skaalauksella.

Varianssiparametrin harhaton estimaattori

- Koska harhan saa helposti korjattua, niin varianssin estimaattina käytetään tavallisesti SU-estimaatin sijasta otosvarianssia (engl. *sample variance*)

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2. \tag{12}$$

- Sitä vastaava estimaattori

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2. \tag{13}$$

on harhaton (varianssiparametrille σ^2), sillä

$$E_{(\mu,\sigma^2)}[S^2] = E_{(\mu,\sigma^2)}[\frac{n}{n-1}\hat{\sigma}^2(\mathbf{Y})] = \frac{n}{n-1}\,\frac{n-1}{n}\sigma^2 = \sigma^2.$$

Estimaattorien $(\bar{Y}, S^2)$ yhteisotantajakauma

Esim. aineopintojen todennäköisyyslaskennan kurssilla todistetaan, että kun $Y_1, \ldots, Y_n$ ovat riippumattomia normaalijakaumaa $N(\mu, \sigma^2)$ noudattavia satunnaismuuttujia, niin tällöin

$$\bar{Y} \text{ ja } S^2 \text{ ovat riippumattomia,} \tag{14}$$

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \sim N(\mu, \frac{1}{n}\sigma^2), \tag{15}$$

$$\frac{n-1}{\sigma^2}S^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - \bar{Y})^2 \sim \chi^2_{n-1}. \tag{16}$$

Tässä χ^2_{n-1} tarkoittaa khiin neliön jakaumaa vapausasteluvulla $n-1$, joka on eräs kuuluisa positiivisella reaaliakselilla määritelty jatkuva jakauma.

Keskiarvon otantajakauma $N(\mu, \sigma^2)$-populaatiossa

- Keskiarvoa $\bar{Y}$ koskeva jakaumatulos (15) on helppo johtaa.
- Normaalijakauman yhteenlaskuominaisuuden mukaan riippumattomien satunnaismuuttujien Y_1 ja Y_2 summalla on normaalijakauma. Sen parametrit ovat

$$Y_1 + Y_2 \sim N(\mu + \mu, \sigma^2 + \sigma^2).$$

- Tätä päättelyä voidaan jatkaa, jolloin summan jakaumaksi saadaan

$$Y_1 + \cdots + Y_n \sim N(n\mu, n\sigma^2).$$

- Kun nyt muistetaan, että tässä ensimmäinen parametri on odotusarvo ja toinen varianssi, niin nähdään heti, että luvulla $1/n$ skaalatun summan jakauma on

$$\bar{Y} \sim N(\mu, \frac{1}{n}\sigma^2).$$

Keskiarvon keskivirhe

- Usein normaalijakaumamallissa ollaan kiinnostuneita vain populaation odotusarvosta μ, ja populaation varianssi σ^2 on ns. haittaparametri (engl. *nuisance parameter*), ts. parametri σ^2 tarvitaan mallin spesifioimiseksi, mutta sen arvosta ei olla kiinnostuneita. Tässä tapauksessa parametrin μ estimaatti on otoskeskiarvo $\bar{y}$.
- Vastaavan estimaattorin $\bar{Y}$ otantavarianssi on σ^2/n.
- Kun tähän kaavaan sijoitetaan tuntemattoman populaatiovarianssin σ^2 tilalle sen otosestimaatti s^2, päädytään siihen, että *keskiarvon keskivirhe* lasketaan kaavalla

$$\frac{1}{\sqrt{n}}\, s,$$

jossa otoskeskihajonta (engl. *sample standard deviation*) $s = \sqrt{s^2}$, ja s^2 on otosvarianssi (12).

Otoskeskihajonta vs. keskiarvon keskivirhe

Otoskeskihajonta

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}. \tag{17}$$

estimoi populaation keskihajontaa.
Sen sijaan *keskiarvon keskivirhe* (engl. *standard error of the mean*)

$$\frac{1}{\sqrt{n}}\, s = \frac{1}{\sqrt{n}} \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{18}$$

estimoi satunnaismuuttujan $\bar{Y}$ keskihajontaa $\sigma/\sqrt{n}$.

Otosvarianssin käyttö satunnaisotokselle muista populaatioista

- Jos populaation odotusarvo on tuntematon, niin myös muulloin kuin normaalijakautuneen populaation tapauksessa populaation varianssia usein estimoidaan otosvarianssilla s^2 (12).
- Sitä vastaava estimaattori S^2 (13) on populaation varianssin harhaton estimaattori.
- Populaation keskihajontaa $\sigma = \sqrt{\sigma^2}$ on myös tapana estimoida otoskeskihajonnalla, vaikka vastaava estimaattori $S = \sqrt{S^2}$ ei ole harhaton.

4.5 Momenttimenetelmä

- Momenttimenetelmä (engl. *method of moments*) on SU-menetelmää varhaisempi menetelmä estimaattorin määrittelemiseksi.
- Menetelmällä saadaan aikaan näppäriä kaavoja estimaateille joissakin sellaisissa tilanteissa, joissa SU-estimaatit jouduttaisiin määrittämään numeerisesti.
- Mallina tässä jaksossa on satunnaisotos $Y_1, \ldots, Y_n$ jakaumasta, jonka ptnf/tf on $g(y; \boldsymbol{\theta})$.
- Otamme käyttöön vielä satunnaismuuttujan Y jolla myöskin on on ptnf/tf $g(y; \boldsymbol{\theta})$.

Populaation momentit ja otosmomentit

Populaation k:s momentti ($k = 1, 2, \dots$) määritellään kaavalla

$$\mu_k(\boldsymbol{\theta}) = EY^k = \begin{cases} \sum_y y^k \, g(y; \boldsymbol{\theta}) & \text{jos jakauma on diskreetti,} \\ \int y^k \, g(y; \boldsymbol{\theta}) \, \mathrm{d}y & \text{jos jakauma on jatkuva.} \end{cases} \tag{19}$$

Momenttia $\mu_k(\theta)$ voidaan estimoida k:nnella otosmomentilla

$$m_k(\mathbf{Y}) = \frac{1}{n} \sum_{i=1}^{n} Y_i^k, \tag{20}$$

joka on populaatiomomentin $\mu_k(\boldsymbol{\theta})$ harhaton estimaattori.

Momenttimenetelmän idea

Momenttimenetelmässä estimaatti (tai estimaattori) muodostetaan ratkaisemalla yhtälöryhmästä

$$\begin{cases} \mu_1(\boldsymbol{\theta}) &= m_1 \\ \mu_2(\boldsymbol{\theta}) &= m_2 \\ &\vdots \\ \mu_r(\boldsymbol{\theta}) &= m_r \end{cases} \tag{21}$$

tuntematon suure $\boldsymbol{\theta}$, jossa otosmomentit $m_1, \ldots, m_r$ lasketaan aineistosta.
Ehtoja asetetaan niin monta, että yhtälöryhmällä on yksikäsitteinen ratkaisu parametriavaruudessa. Tavallisesti yhtälöitä asetetaan niin monta, kuin parametrivektorissa on komponentteja.

Eksponenttijakauman parametrin estimointi esimerkkinä momenttimenetelmästä

- Eksponenttijakaumaa noudattava satunnaismuuttuja Y voi saada kaikkia positiivisia reaaliarvoja, ja sillä on tiheysfunktio

$$g(y;\lambda) = \lambda \, \mathrm{e}^{-\lambda y}, \qquad y > 0. \tag{22}$$

  jossa jakauman parametria on merkitty kirjaimella $\lambda > 0$.
- Jakaumasta käytetään lyhennettä $\mathrm{Exp}(\lambda)$.
- Jos $Y \sim \mathrm{Exp}(\lambda)$, niin sen odotusarvo on tunnetusti

$$EY = \frac{1}{\lambda}.$$

Momenttimenetelmä jakaumalle $\mathrm{Exp}(\lambda)$

- Tarkastelemme satunnaisotosta $Y_1, \ldots, Y_n$ eksponenttijakaumasta $\mathrm{Exp}(\lambda)$.
- Havaitut arvot ovat $y_1, \ldots, y_n$.
- Koska parametreja on vain yksi, momenttimenetelmässä tarvitaan vain yksi yhtälö

$$EY = \frac{1}{\lambda} = m_1 = \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y}.$$

- Momenttimenetelmän mukainen parametrin λ estimaatti on

$$\hat{\lambda} = \frac{1}{\bar{y}}.$$

Momenttimenetelmä eksponenttijakaumalle (jatkoa)

- Vaihtoehtoisesti eksponenttijakauma $\mathrm{Exp}(\lambda)$ voitaisiin parametroida sen odotusarvolla $\theta = 1/\lambda$.
- Momenttimenetelmä antaa tälle parametrille estimaatin

$$\hat{\theta} = \bar{y}.$$

- Eksponenttijakauman parametrille voi helposti johtaa myös SU-estimaatin kummalla tahansa parametroinnilla.
- Tässä esimerkissä momenttimenetelmä antaa samat estimaatit kuin SU-menetelmä, mutta yleisesti ottaen nämä menetelmät voivat tuottaa erilaiset estimaatit.