Luennoija: Pentti Saikkonen

Laajuus: 5 op.

Tyyppi: Syventävät opinnot

Luentoajat: I periodi ma 10–12, ti 12–14 B322

Suoritustapa: Kurssikoe ti 1.11. klo 9.00-12.00 D123 ja yleistentit, ensimmäinen 17.11.

Kurssimateriaali: Luentomoniste, jossa mainintoja oheislukemistosta.

Laskuharjoitukset: Ti 10–12, C122 (alkaen 13.9.), Pekka Nieminen

- Lisäpisteitä harjoitustehtävien ratkaisemisesta saa kahdessa ensimmäisessä loppukokeessa 1, 2, 3 tai 4, jos ratkaistujen tehtävien määrä on vastaavasti 20, 40, 60 tai 80 prosenttia.
- Lisäpisteet ainoastaan hyväksytysti suoritettujen tenttien arvosanan korotukseen. Edellyttää läsnäoloa harjoitustilaisuudessa.

Esitietovaatimukset

- Todennäköisyyslaskennan kurssi
- Tilastollisen päättelyn kurssi
- Lineaaristen mallien kurssi
- Edellä mainittujen edellyttämät matematiikan kurssit (yhden ja useamman muuttujan differentiaali- ja integraalilaskenta sekä lineaarialgebra ja matriisilaskenta)

Asema opetuksessa

- Tilastollisen päättelyn jatkokurssi on tilastotieteen syventävien opintojen pakollinen kurssi.
- Kurssilla syvennetään ja laajennetaan aineopintojen tilastollista päättelyn kurssilla esitettyä uskottavuuspäättelyä.
- Keskeinen sisältö muodostuu suurimman uskottavuuden estimaattorin asymptoottisten ominaisuuksien ja uskottavuusteorian testien tarkastelusta vektoriparametrisissa malleissa.

- Jono satunnaismuuttujia (sm) $X_1, X_2, \ldots$ **konvergoi stokastisesti kohti sm:aa** Z, merkintään, $X_n \xrightarrow{p} Z$ tai $plim_{n\to\infty} X_n = Z$, jos

$$\mathrm{P}\left\{|X_n - Z| > \varepsilon\right\} \underset{n\to\infty}{\longrightarrow} 0 \quad \text{kaikilla } \varepsilon > 0. \qquad (*)$$

- Jono satunnaismuuttujia (sm) $X_1, X_2, \ldots$ **konvergoi stokastisesti kohti sm:aa** Z, merkintään, $X_n \xrightarrow{p} Z$ tai $plim_{n\to\infty} X_n = Z$, jos

$$\mathrm{P}\left\{|X_n - Z| > \varepsilon\right\} \underset{n\to\infty}{\longrightarrow} 0 \quad \text{kaikilla } \varepsilon > 0. \qquad (*)$$

- Määritelmässä voi vaihtoehtoisesti olla tapahtuma $\left\{|X_n - Z| \geq \varepsilon\right\}$.

- Jono satunnaismuuttujia (sm) $X_1, X_2, \ldots$ **konvergoi stokastisesti kohti sm:aa** Z, merkintään, $X_n \xrightarrow{p} Z$ tai $plim_{n\to\infty} X_n = Z$, jos

$$\mathrm{P}\left\{|X_n - Z| > \varepsilon\right\} \underset{n\to\infty}{\rightarrow} 0 \quad \text{kaikilla } \varepsilon > 0. \qquad (*)$$

- Määritelmässä voi vaihtoehtoisesti olla tapahtuma $\left\{|X_n - Z| \geq \varepsilon\right\}$.
- $X_1, X_2, \ldots$ ja Z ovat $k \times 1$ satunnaisvektoreita (sv), määritelmä on

$$X_n \xrightarrow{p} Z \quad \Leftrightarrow \quad X_{a,n} \xrightarrow{p} Z_a \quad \text{kaikilla } a = 1, \ldots, k,$$

kun esimerkiksi $X_{a,n}$ on X_n:n a. komponentti.

- Jono satunnaismuuttujia (sm) $X_1, X_2, \ldots$ **konvergoi stokastisesti kohti sm:aa** Z, merkintään, $X_n \xrightarrow{p} Z$ tai $plim_{n \to \infty} X_n = Z$, jos

$$\mathrm{P}\left\{|X_n - Z| > \varepsilon\right\} \underset{n \to \infty}{\longrightarrow} 0 \quad \text{kaikilla } \varepsilon > 0. \qquad (*)$$

- Määritelmässä voi vaihtoehtoisesti olla tapahtuma $\left\{|X_n - Z| \geq \varepsilon\right\}$.
- $X_1, X_2, \ldots$ ja Z ovat $k \times 1$ satunnaisvektoreita (sv), määritelmä on

$$X_n \xrightarrow{p} Z \quad \Leftrightarrow \quad X_{a,n} \xrightarrow{p} Z_a \quad \text{kaikilla } a = 1, \ldots, k,$$

kun esimerkiksi $X_{a,n}$ on X_n:n a. komponentti.

- Vaihtoehtoisesti $(*)$:ssä itseisarvo korvataan vektorin normilla $\|\cdot\|$; $x \in \mathbb{R}^k$, $\|x\| = \sqrt{x_1^2 + \cdots + x_k^2}$.

- Jono satunnaismuuttujia (sm) $X_1, X_2, \ldots$ **konvergoi stokastisesti kohti sm:aa** Z, merkintään, $X_n \xrightarrow{p} Z$ tai $plim_{n\to\infty} X_n = Z$, jos

$$\mathrm{P}\left\{|X_n - Z| > \varepsilon\right\} \underset{n\to\infty}{\rightarrow} 0 \quad \text{kaikilla } \varepsilon > 0. \qquad (*)$$

- Määritelmässä voi vaihtoehtoisesti olla tapahtuma $\left\{|X_n - Z| \geq \varepsilon\right\}$.
- $X_1, X_2, \ldots$ ja Z ovat $k \times 1$ satunnaisvektoreita (sv), määritelmä on

$$X_n \xrightarrow{p} Z \quad \Leftrightarrow \quad X_{a,n} \xrightarrow{p} Z_a \quad \text{kaikilla } a = 1, \ldots, k,$$

kun esimerkiksi $X_{a,n}$ on X_n:n a. komponentti.
- Vaihtoehtoisesti $(*)$:ssä itseisarvo korvataan vektorin normilla $\|\cdot\|$; $x \in \mathbb{R}^k$, $\|x\| = \sqrt{x_1^2 + \cdots + x_k^2}$.
- Satunnaismatriisiin tapauksessa määritelmä voidaan palauttaa satunnaisvektorin määritelmään suorittamalla "vektorointi"

- Seuraavassa tarkoitetaan vakiolla c ei-satunnaista (äärellistä) suuretta. Tilastollisissa sovelluksissa on usein tarpeen tietää milloin pätee implikaatio

$$X_n \xrightarrow{p} c \Rightarrow g\left(X_n\right) \xrightarrow{p} g(c).$$

- Seuraavassa tarkoitetaan vakiolla c ei-satunnaista (äärellistä) suuretta. Tilastollisissa sovelluksissa on usein tarpeen tietää milloin pätee implikaatio

$$X_n \xrightarrow{p} c \Rightarrow g\left(X_n\right) \xrightarrow{p} g(c).$$

- Liittyy ilmeisesti jotenkin funktion g jatkuvuuteen.

- Seuraavassa tarkoitetaan vakiolla c ei-satunnaista (äärellistä) suuretta. Tilastollisisissa sovelluksissa on usein tarpeen tietää milloin pätee implikaatio

$$X_n \xrightarrow{p} c \Rightarrow g\left(X_n\right) \xrightarrow{p} g(c).$$

- Liittyy ilmeisesti jotenkin funktion g jatkuvuuteen.

- **Lause 1.1.** Olkoon $X_1, X_2, \ldots$ ja c $(k \times 1)$. Jos $X_n \xrightarrow{p} c$ ja $g : \mathbb{R}^k \to \mathbb{R}^l$ jatkuva pisteessä c, niin $g\left(X_n\right) \xrightarrow{p} g(c)$.

- **Lause 1.1.** Olkoon $X_1, X_2, \ldots$ ja c $(k \times 1)$. Jos $X_n \xrightarrow{p} c$ ja $g : \mathbb{R}^k \to \mathbb{R}^l$ jatkuva pisteessä c, niin $g(X_n) \xrightarrow{p} g(c)$.

- **Lause 1.1.** Olkoon $X_1, X_2, \ldots$ ja c $(k \times 1)$. Jos $X_n \xrightarrow{p} c$ ja $g : \mathbb{R}^k \to \mathbb{R}^l$ jatkuva pisteessä c, niin $g(X_n) \xrightarrow{p} g(c)$.
- Esimerkkitapauksia: $g(x_1, x_2) = x_1 + x_2$ ja $g(x_1, x_2) = x_1 x_2$ (olettaen summa ja tulo määritellyiksi).

- **Lause 1.1.** Olkoon $X_1, X_2, ...$ ja c $(k \times 1)$. Jos $X_n \xrightarrow{p} c$ ja $g : \mathbb{R}^k \to \mathbb{R}^l$ jatkuva pisteessä c, niin $g(X_n) \xrightarrow{p} g(c)$.
- Esimerkkitapauksia: $g(x_1, x_2) = x_1 + x_2$ ja $g(x_1, x_2) = x_1 x_2$ (olettaen summa ja tulo määritellyiksi).
- Lauseen tulos pitää paikkansa myös, kun raja-arvo on satunnainen, mutta todistus ei onnistu kopioimalla edellä esitettyä todistusta aivan suoraan. Jatkossa tätä yleistystä ei kuitenkaan tarvita.

- Olkoon F_{X_1}, F_{X_2}, , ... jono sm:ien $X_1, X_2, \ldots$ kertymäfunktioita ja F_Z sm:n Z kertymäfunktio (esim. $F_Z(x) = \mathrm{P}\{Z \leq x\}$).

- Olkoon $F_{X_1}, F_{X_2}, , \ldots$ jono sm:ien $X_1, X_2, \ldots$ kertymäfunktioita ja F_Z sm:n Z kertymäfunktio (esim. $F_Z(x) = \mathrm{P}\{Z \leq x\}$).
- Jonon $X_1, X_2, \ldots$ sanotaan **konvergoivan jakaumaltaan kohti sm:aa** Z, merkintään $X_n \xrightarrow{d} Z$, jos

$$\lim_{n\to\infty} F_{X_n}(x) = F_Z(x) \quad \text{kaikissa} \quad F_Z\text{:n jatkuvuuspisteissä.}$$

- Olkoon $F_{X_1}, F_{X_2}, , \ldots$ jono sm:ien $X_1, X_2, \ldots$ kertymäfunktioita ja F_Z sm:n Z kertymäfunktio (esim. $F_Z(x) = \mathrm{P}\{Z \leq x\}$).
- Jonon $X_1, X_2, \ldots$ sanotaan **konvergoivan jakaumaltaan kohti sm:aa** Z, merkintään $X_n \xrightarrow{d} Z$, jos

$$\lim_{n\to\infty} F_{X_n}(x) = F_Z(x) \quad \text{kaikissa} \quad F_Z\text{:n jatkuvuuspisteissä.}$$

- Yleistys sv:lle sujuu käyttämällä sv:ien kertymäfunktioita (esim. $F_Z(x) = \mathrm{P}\{Z_1 \leq x_1, ..., Z_k \leq x_k\}$, kun $Z = (Z_1, ..., Z_k)$).

$X_n \xrightarrow{d} Z$, jos $\lim_{n\to\infty} F_{X_n}(x) = F_Z(x)$ kaikissa F_Z:n jatkuvuuspisteissä

- Tilastotieteessä rajajakauma on useimmiten jatkuva, joten Z:n kf $F_Z(x)$ on jatkuva kaikilla x.

Konvergenssikäsitteitä ja tuloksia

Jakaumakonvergenssi:

$X_n \xrightarrow{d} Z$, jos $\lim_{n\to\infty} F_{X_n}(x) = F_Z(x)$ kaikissa F_Z:n jatkuvuuspisteissä

- Tilastotieteessä rajajakauma on useimmiten jatkuva, joten Z:n kf $F_Z(x)$ on jatkuva kaikilla x.
- Epäjatkuvuuspisteitä voi olla korkeintaan numeroituva määrä.

$X_n \xrightarrow{d} Z$, jos $\lim_{n\to\infty} F_{X_n}(x) = F_Z(x)$ kaikissa F_Z:n jatkuvuuspisteissä

- Tilastotieteessä rajajakauma on useimmiten jatkuva, joten Z:n kf $F_Z(x)$ on jatkuva kaikilla x.
- Epäjatkuvuuspisteitä voi olla korkeintaan numeroituva määrä.
- Kun F_Z on jatkuva, niin "suurilla" n:n arvoilla ja kaikilla x pätee

$$\mathrm{P}\{X_n \le x\} \approx \mathrm{P}\{Z \le x\}.$$

$X_n \xrightarrow{d} Z$, jos $\lim_{n\to\infty} F_{X_n}(x) = F_Z(x)$ kaikissa F_Z:n jatkuvuuspisteissä

- Tilastotieteessä rajajakauma on useimmiten jatkuva, joten Z:n kf $F_Z(x)$ on jatkuva kaikilla x.
- Epäjatkuvuuspisteitä voi olla korkeintaan numeroituva määrä.
- Kun F_Z on jatkuva, niin "suurilla" n:n arvoilla ja kaikilla x pätee

$$\mathrm{P}\{X_n \leq x\} \approx \mathrm{P}\{Z \leq x\}.$$

- Kaikkien sm:aa X_n koskevien tapahtumien todennäköisyyksiä voidaan approksimoida käyttäen sm:n Z jakaumaa.

$X_n \xrightarrow{d} Z$, jos $\lim_{n\to\infty} F_{X_n}(x) = F_Z(x)$ kaikissa F_Z:n jatkuvuuspisteissä

- Tilastotieteessä rajajakauma on useimmiten jatkuva, joten Z:n kf $F_Z(x)$ on jatkuva kaikilla x.
- Epäjatkuvuuspisteitä voi olla korkeintaan numeroituva määrä.
- Kun F_Z on jatkuva, niin "suurilla" n:n arvoilla ja kaikilla x pätee

$$\mathrm{P}\{X_n \le x\} \approx \mathrm{P}\{Z \le x\}.$$

- Kaikkien sm:aa X_n koskevien tapahtumien todennäköisyyksiä voidaan approksimoida käyttäen sm:n Z jakaumaa.
- Tyypillisessä tilastollisessa sovelluksessa X_n on testisuure tai (muunnettu) estimaattori.

- Jakaumakonvergenssin tai heikon konvergenssin määrittelyehto $F_{X_n}(x) \to F_Z(x)$ on havainnollinen, mutta matemaattisesti hankala.

- Jakaumakonvergenssin tai heikon konvergenssin määrittelyehto $F_{X_n}(x) \to F_Z(x)$ on havainnollinen, mutta matemaattisesti hankala.

- Sv:n $Z = (Z_1, ..., Z_k)$ **karakteristinen funktio** (kf)

$$\varphi_Z(t) = \mathrm{E}\left[\exp\left(it'Z\right)\right] = \mathrm{E}\left[\cos\left(t'Z\right)\right] + i\mathrm{E}\left[\sin\left(t'Z\right)\right],$$

jossa $i = \sqrt{-1}$, $t = (t_1, ..., t_k)$ ja $t'Z = \sum_{j=1}^{k} t_j Z_j$.

- Jakaumakonvergenssin tai heikon konvergenssin määrittelyehto $F_{X_n}(x) \to F_Z(x)$ on havainnollinen, mutta matemaattisesti hankala.

- Sv:n $Z = (Z_1, ..., Z_k)$ **karakteristinen funktio** (kf)

$$\varphi_Z(t) = \mathrm{E}\left[\exp\left(it'Z\right)\right] = \mathrm{E}\left[\cos\left(t'Z\right)\right] + i\mathrm{E}\left[\sin\left(t'Z\right)\right],$$

jossa $i = \sqrt{-1}$, $t = (t_1, ..., t_k)$ ja $t'Z = \sum_{j=1}^{k} t_j Z_j$.

- kf määrää jakauman yksikäsitteisesti

Lause 1.2. Olkoon $X_1, X_2, \ldots$ ja Z sv:ta $(k \times 1)$. Seuraavat kolme konvergenssia pätevät yhtäpitävästi.

(i) $X_n \xrightarrow{d} Z$

(ii) $\varphi_{X_n}(t) \underset{n \to \infty}{\longrightarrow} \varphi_Z(t)$ jokaisella $t \in \mathbb{R}^k$.

(iii) $\mathrm{E}\left[h(X_n)\right] \underset{n \to \infty}{\longrightarrow} \mathrm{E}\left[h(Z)\right]$ kaikilla reaaliarvoisilla rajoitetuilla ja jatkuvilla funktioilla $h : \mathbb{R}^k \to \mathbb{R}$.

- Jakaumakonvergenssin "jatkuvan kuvauksen" lause saadaan helposti ekvivalenssista $X_n \xrightarrow{d} Z \Leftrightarrow \mathrm{E}\left[h\left(X_n\right)\right] \underset{n\to\infty}{\longrightarrow} \mathrm{E}\left[h\left(Z\right)\right]$ kaikilla reaaliarvoisilla rajoitetuilla ja jatkuvilla funktioilla h.

- Jakaumakonvergenssin "jatkuvan kuvauksen" lause saadaan helposti ekvivalenssista $X_n \xrightarrow{d} Z \Leftrightarrow \mathrm{E}\left[h\left(X_n\right)\right] \underset{n\to\infty}{\longrightarrow} \mathrm{E}\left[h\left(Z\right)\right]$ kaikilla reaaliarvoisilla rajoitetuilla ja jatkuvilla funktioilla h.

- **Lause 1.3.** Olkoon $X_1, X_2, \ldots$ ja Z $(k \times 1)$. Jos $X_n \xrightarrow{d} Z$ ja $g : \mathbb{R}^k \to \mathbb{R}^l$ on jatkuva, niin $g\left(X_n\right) \xrightarrow{d} g(Z)$.

- Jakaumakonvergenssin "jatkuvan kuvauksen" lause saadaan helposti ekvivalenssista $X_n \xrightarrow{d} Z \Leftrightarrow \mathrm{E}\left[h\left(X_n\right)\right] \underset{n\to\infty}{\longrightarrow} \mathrm{E}\left[h\left(Z\right)\right]$ kaikilla reaaliarvoisilla rajoitetuilla ja jatkuvilla funktioilla h.
- **Lause 1.3.** Olkoon $X_1, X_2, \ldots$ ja $Z\ (k \times 1)$. Jos $X_n \xrightarrow{d} Z$ ja $g : \mathbb{R}^k \longrightarrow \mathbb{R}^l$ on jatkuva, niin $g\left(X_n\right) \xrightarrow{d} g(Z)$.
- Lievennys: Riittää, että g on jatkuva vain joukossa C_g, jolle pätee $\mathrm{P}\left\{Z \in C_g\right\} = 1$.

- Asymptoottisessa tilastollisessa päättelyssä joudutaan usein käyttämään samanaikaisesti sekä jakaumakonvergenssia että stokastista konvergenssia eli tilanne on $X_n \xrightarrow{d} Z$ ja $Y_n \xrightarrow{p} c$.

  Tällöin seuraava lause on hyödyllinen.

- Asymptoottisessa tilastollisessa päättelyssä joudutaan usein käyttämään samanaikaisesti sekä jakaumakonvergenssia että stokastista konvergenssia eli tilanne on $X_n \xrightarrow{d} Z$ ja $Y_n \xrightarrow{p} c$.

  Tällöin seuraava lause on hyödyllinen.

- **Lause 1.4.** Oletetaan, että $X_n \xrightarrow{d} Z \quad (k \times 1)$ ja $Y_n \xrightarrow{p} c \quad (l \times 1)$, jossa c on vakiovektori. Tällöin $(X_n, Y_n) \xrightarrow{d} (Z, c)$.

- Asymptoottisessa tilastollisessa päättelyssä joudutaan usein käyttämään samanaikaisesti sekä jakaumakonvergenssia että stokastista konvergenssia eli tilanne on $X_n \xrightarrow{d} Z$ ja $Y_n \xrightarrow{p} c$.

  Tällöin seuraava lause on hyödyllinen.

- **Lause 1.4.** Oletetaan, että $X_n \xrightarrow{d} Z \quad (k \times 1)$ ja $Y_n \xrightarrow{p} c \quad (l \times 1)$, jossa c on vakiovektori. Tällöin $(X_n, Y_n) \xrightarrow{d} (Z, c)$.

- Yhdistettynä jakaumakonvergenssin jatkuvan kuvauksen lauseeseen seuraa tästä

$$g\left(X_n, Y_n\right) \xrightarrow{d} g(Z, c), \quad \text{kun funktio} \quad (z, c) \mapsto g\left(z, c\right) \quad \text{on jatkuva.}$$

$$X_n \xrightarrow{d} Z \text{ ja } \; Y_n \xrightarrow{p} c \Longrightarrow (X_n, Y_n) \xrightarrow{d} (Z, c)$$

- **Seuraus 1.1.** $X_n \xrightarrow{d} Z$ ja $Y_n \xrightarrow{p} c$ (vakio) reaalisia. Tällöin,

$$X_n \xrightarrow{d} Z \text{ ja } \; Y_n \xrightarrow{p} c \Longrightarrow (X_n, Y_n) \xrightarrow{d} (Z, c)$$

- **Seuraus 1.1.** $X_n \xrightarrow{d} Z$ ja $Y_n \xrightarrow{p} c$ (vakio) reaalisia. Tällöin,
- (i) $X_n + Y_n \xrightarrow{d} Z + c$ [$(z, c) \mapsto z + c$ jatkuva]

$$X_n \xrightarrow{d} Z \text{ ja } \; Y_n \xrightarrow{p} c \Longrightarrow (X_n, Y_n) \xrightarrow{d} (Z, c)$$

- **Seuraus 1.1.** $X_n \xrightarrow{d} Z$ ja $Y_n \xrightarrow{p} c$ (vakio) reaalisia. Tällöin,
- (i) $X_n + Y_n \xrightarrow{d} Z + c$ $\qquad$ $[(z, c) \mapsto z + c$ jatkuva$]$
- (ii) $X_n Y_n \xrightarrow{d} Zc$ $\qquad$ $[(x, c) \mapsto zc$ jatkuva$]$

$$X_n \xrightarrow{d} Z \text{ ja } \ Y_n \xrightarrow{p} c \Longrightarrow (X_n, Y_n) \xrightarrow{d} (Z, c)$$

- **Seuraus 1.1.** $X_n \xrightarrow{d} Z$ ja $Y_n \xrightarrow{p} c$ (vakio) reaalisia. Tällöin,
- (i) $X_n + Y_n \xrightarrow{d} Z + c$ $\quad [(z, c) \mapsto z + c$ jatkuva$]$
- (ii) $X_n Y_n \xrightarrow{d} Zc$ $\quad [(x, c) \mapsto zc$ jatkuva$]$
- (iii) $X_n / Y_n \xrightarrow{d} Z/c, \quad c \neq 0$ $\quad [(z, c) \mapsto z/c$ jatkuva, kun $c \neq 0]$

$$X_n \xrightarrow{d} Z \text{ ja } \; Y_n \xrightarrow{p} c \Longrightarrow (X_n, Y_n) \xrightarrow{d} (Z, c)$$

- **Seuraus 1.2.** $X_n \xrightarrow{d} Z \;\; (k \times 1)$ ja $Y_n \xrightarrow{p} A \; (l \times k)$, A vakiomatriisi. Tällöin,

$$X_n \xrightarrow{d} Z \text{ ja } \; Y_n \xrightarrow{p} c \Longrightarrow (X_n, Y_n) \xrightarrow{d} (Z, c)$$

- **Seuraus 1.2.** $X_n \xrightarrow{d} Z \;\; (k \times 1)$ ja $Y_n \xrightarrow{p} A \; (l \times k)$, A vakiomatriisi. Tällöin,
- (i) $\; Y_n X_n \xrightarrow{d} AZ$

$$X_n \xrightarrow{d} Z \text{ ja } \; Y_n \xrightarrow{p} c \Longrightarrow (X_n, Y_n) \xrightarrow{d} (Z, c)$$

- **Seuraus 1.2.** $X_n \xrightarrow{d} Z \;\; (k \times 1)$ ja $Y_n \xrightarrow{p} A \; (l \times k)$, A vakiomatriisi. Tällöin,
- (i) $\;Y_n X_n \xrightarrow{d} AZ$
- (ii) $\;X_n' Y_n X_n \xrightarrow{d} Z' A Z$, kun $k = l$

$$X_n \xrightarrow{d} Z \text{ ja } \; Y_n \xrightarrow{p} c \Longrightarrow (X_n, Y_n) \xrightarrow{d} (Z, c)$$

- **Seuraus 1.2.** $X_n \xrightarrow{d} Z \;\; (k \times 1)$ ja $Y_n \xrightarrow{p} A \; (l \times k)$, A vakiomatriisi. Tällöin,
- (i) $Y_n X_n \xrightarrow{d} AZ$
- (ii) $X_n' Y_n X_n \xrightarrow{d} Z' A Z$, kun $k = l$
- (iii) $X_n' Y_n^{-1} X_n \xrightarrow{d} Z' A^{-1} Z$, $k = l$ ja A on epäsingulaarinen.

$$X_n \xrightarrow{d} Z \text{ ja } \ Y_n \xrightarrow{p} c \Longrightarrow (X_n, Y_n) \xrightarrow{d} (Z, c)$$

- **Seuraus 1.2.** $X_n \xrightarrow{d} Z \ \ (k \times 1)$ ja $Y_n \xrightarrow{p} A \ (l \times k)$, A vakiomatriisi. Tällöin,
- (i) $Y_n X_n \xrightarrow{d} AZ$
- (ii) $X_n' Y_n X_n \xrightarrow{d} Z' AZ$, kun $k = l$
- (iii) $X_n' Y_n^{-1} X_n \xrightarrow{d} Z' A^{-1} Z$, $k = l$ ja A on epäsingulaarinen.
- Perustelu: Kuvaukset $(A, z) \mapsto Az$, $(A, z) \mapsto z' Az$ ja $A \mapsto A^{-1}$ (A epäsing.) jatkuvia.

- Olkoon $\theta \in \Theta \subseteq \mathbb{R}$ estimoitava parametri, θ_0 sen "todellinen" arvo ja parametrivaruus Θ avoin väli. Tarkastellaan θ_0:n asymptoottisesti normaalia estimaattoria $\hat{\theta}_n$:

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \sigma^2\left(\theta_0\right)\right) \qquad \left(\sigma^2\left(\theta_0\right) > 0\right).$$

- Olkoon $\theta \in \Theta \subseteq \mathbb{R}$ estimoitava parametri, θ_0 sen "todellinen" arvo ja parametrivaruus Θ avoin väli. Tarkastellaan θ_0:n asymptoottisesti normaalia estimaattoria $\hat{\theta}_n$:

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \sigma^2\left(\theta_0\right)\right) \qquad \left(\sigma^2\left(\theta_0\right) > 0\right).$$

- Jos $h : \Theta \to \mathbb{R}$ on jatkuvasti derivoituva ja $h'(\theta_0) \neq 0$, niin

$$\sqrt{n}(h(\hat{\theta}_n) - h(\theta_0)) \xrightarrow{d} \mathrm{N}\left(0, \left[h'(\theta_0)\right]^2 \sigma^2\left(\theta_0\right)\right).$$

- Olkoon $\theta \in \Theta \subseteq \mathbb{R}$ estimoitava parametri, θ_0 sen "todellinen" arvo ja parametrivaruus Θ avoin väli. Tarkastellaan θ_0:n asymptoottisesti normaalia estimaattoria $\hat{\theta}_n$:

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \sigma^2\left(\theta_0\right)\right) \qquad \left(\sigma^2\left(\theta_0\right) > 0\right).$$

- Jos $h : \Theta \to \mathbb{R}$ on jatkuvasti derivoituva ja $h'(\theta_0) \neq 0$, niin

$$\sqrt{n}(h(\hat{\theta}_n) - h(\theta_0)) \xrightarrow{d} \mathrm{N}\left(0, \left[h'\left(\theta_0\right)\right]^2 \sigma^2\left(\theta_0\right)\right).$$

- Perustelu (HT 1.3) nojaa väliarvolauseeseen ja Lauseisiin 1.1, 1.3 ja 1.4 (tai jälkimmäisten seuraukseen).

- Olkoon $\theta \in \Theta \subseteq \mathbb{R}$ estimoitava parametri, θ_0 sen "todellinen" arvo ja parametrivaruus Θ avoin väli. Tarkastellaan θ_0:n asymptoottisesti normaalia estimaattoria $\hat{\theta}_n$:

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \sigma^2\left(\theta_0\right)\right) \qquad \left(\sigma^2\left(\theta_0\right) > 0\right).$$

- Jos $h : \Theta \to \mathbb{R}$ on jatkuvasti derivoituva ja $h'(\theta_0) \neq 0$, niin

$$\sqrt{n}(h(\hat{\theta}_n) - h(\theta_0)) \xrightarrow{d} \mathrm{N}\left(0, [h'(\theta_0)]^2\sigma^2\left(\theta_0\right)\right).$$

- Perustelu (HT 1.3) nojaa väliarvolauseeseen ja Lauseisiin 1.1, 1.3 ja 1.4 (tai jälkimmäisten seuraukseen).

- Tiedoksi lievennys: Funktion h jatkuva derivoituvuus riittää olettaa vain jollain (pienellä) θ_0:n sisältävällä välillä.

- Tilastollisessa päättelyssä tarkastellaan usein parametrista riippuvia sm:ia ja niistä muodostettuja jonoja:

$$X_1(\theta), X_2(\theta), \ldots, \quad \theta \in \Theta.$$

- Tilastollisessa päättelyssä tarkastellaan usein parametrista riippuvia sm:ia ja niistä muodostettuja jonoja:

$$X_1(\theta), X_2(\theta), \ldots, \quad \theta \in \Theta.$$

- Esimerkiksi

$$X_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} q(Y_i;\theta), \quad \text{jossa} \quad q(Y_i;\theta) = \log f(Y_i;\theta)$$

eli $n^{-1}\times$ log-uskottavuusfunktio riippumattomien ja samoin jakautuneiden havaintojen $Y_1, \ldots, Y_n$ tapauksessa (iid–tapauksessa).

- Tilastollisessa päättelyssä tarkastellaan usein parametrista riippuvia sm:ia ja niistä muodostettuja jonoja:

$$X_1(\theta), X_2(\theta), \ldots, \quad \theta \in \Theta.$$

- Esimerkiksi

$$X_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} q(Y_i;\theta), \quad \text{jossa} \quad q(Y_i;\theta) = \log f(Y_i;\theta)$$

eli $n^{-1}\times$ log-uskottavuusfunktio riippumattomien ja samoin jakautuneiden havaintojen $Y_1, \ldots, Y_n$ tapauksessa (iid–tapauksessa).

  - Todennäköisyydet lasketaan olettaen "todellinen" parametriarvo θ_0 tai havaintojen jakauma $\prod_{i=1}^{n} f(Y_i;\theta_0)$

- Jono reaaliarvoisia satunnaisfunktioita $X_1(\theta), X_2(\theta), \ldots$, $\theta \in \Theta$, **konvergoi stokastisesti ja tasaisesti** kohti ei-satunnaista funktiota $c(\theta)$, jos

$$\sup_{\theta \in \Theta} |X_n(\theta) - c(\theta)| \xrightarrow{p} 0.$$

- Jono reaaliarvoisia satunnaisfunktioita $X_1(\theta), X_2(\theta), \ldots, \theta \in \Theta$, **konvergoi stokastisesti ja tasaisesti** kohti ei-satunnaista funktiota $c(\theta)$, jos

$$\sup_{\theta \in \Theta} |X_n(\theta) - c(\theta)| \xrightarrow{p} 0.$$

- Tällöin $X_n(\theta)$ on "suurilla" n ja "lähes ykkösen" todennäköisyydellä "lähellä" funktiota $c(\theta)$ eikä "läheisyys" riipu siitä mikä parametriarvo on kysymyksessä.

- Jono reaaliarvoisia satunnaisfunktioita $X_1(\theta), X_2(\theta), \ldots, \theta \in \Theta$, **konvergoi stokastisesti ja tasaisesti** kohti ei-satunnaista funktiota $c(\theta)$, jos

$$\sup_{\theta \in \Theta} |X_n(\theta) - c(\theta)| \xrightarrow{p} 0.$$

- Tällöin $X_n(\theta)$ on "suurilla" n ja "lähes ykkösen" todennäköisyydellä "lähellä" funktiota $c(\theta)$ eikä "läheisyys" riipu siitä mikä parametriarvo on kysymyksessä.

- Vaatii enemmän kuin **pisteittäinen stokastinen konvergenssi** $X_n(\theta) \xrightarrow{p} c(\theta)$ jokaisella $\theta \in \Theta$ eli

$$|X_n(\theta) - c(\theta)| \xrightarrow{p} 0 \quad \text{jokaisella } \theta \in \Theta.$$

Tasaisen stokastisen konvergenssin sovelluksia.

- Suurimman uskottavuuden (SU) estimaattorin (ja muiden saman tyyppisten estimaattoreiden) tarkentuvuuden toteaminen.

Tasaisen stokastisen konvergenssin sovelluksia.

- Suurimman uskottavuuden (SU) estimaattorin (ja muiden saman tyyppisten estimaattoreiden) tarkentuvuuden toteaminen.
- Olkoon $\hat{\theta}_n$ parametrin θ_0 tarkentuva estimaattori eli $\hat{\theta}_n \xrightarrow{p} \theta_0$ ja $X_n(\theta_0) \xrightarrow{p} c(\theta_0)$ tai $X_n(\theta) \xrightarrow{p} c(\theta)$ jokaisella $\theta \in \Theta$ pätee.

Tasaisen stokastisen konvergenssin sovelluksia.

- Suurimman uskottavuuden (SU) estimaattorin (ja muiden saman tyyppisten estimaattoreiden) tarkentuvuuden toteaminen.
- Olkoon $\hat{\theta}_n$ parametrin θ_0 tarkentuva estimaattori eli $\hat{\theta}_n \xrightarrow{p} \theta_0$ ja $X_n(\theta_0) \xrightarrow{p} c(\theta_0)$ tai $X_n(\theta) \xrightarrow{p} c(\theta)$ jokaisella $\theta \in \Theta$ pätee.
  - Päteekö $X_n(\hat{\theta}_n) \xrightarrow{p} c(\theta_0)$?

Tasaisen stokastisen konvergenssin sovelluksia.

- Suurimman uskottavuuden (SU) estimaattorin (ja muiden saman tyyppisten estimaattoreiden) tarkentuvuuden toteaminen.
- Olkoon $\hat{\theta}_n$ parametrin θ_0 tarkentuva estimaattori eli $\hat{\theta}_n \xrightarrow{p} \theta_0$ ja $X_n(\theta_0) \xrightarrow{p} c(\theta_0)$ tai $X_n(\theta) \xrightarrow{p} c(\theta)$ jokaisella $\theta \in \Theta$ pätee.
  - Päteekö $X_n(\hat{\theta}_n) \xrightarrow{p} c(\theta_0)$?
  - Pätee, jos stokastinen konvergenssi $X_n(\theta) \xrightarrow{p} c(\theta)$ on tasainen, mutta ei välttämättä muuten.

- **Lause 1.5.** Oletetaan, että

- **Lause 1.5.** Oletetaan, että
  - $X_n(\theta) \xrightarrow{p} c(\theta)$ tasaisesti joukossa Θ

- **Lause 1.5.** Oletetaan, että
  - $X_n(\theta) \xrightarrow{p} c(\theta)$ tasaisesti joukosssa Θ
  - funktio $c(\theta)$ on jatkuva pisteessä θ_0 ja

- **Lause 1.5.** Oletetaan, että
  - $X_n(\theta) \xrightarrow{p} c(\theta)$ tasaisesti joukosssa Θ
  - funktio $c(\theta)$ on jatkuva pisteessä θ_0 ja
  - $\hat{\theta}_n \xrightarrow{p} \theta_0$.

- **Lause 1.5.** Oletetaan, että
    - $X_n(\theta) \xrightarrow{p} c(\theta)$ tasaisesti joukossa Θ
    - funktio $c(\theta)$ on jatkuva pisteessä θ_0 ja
    - $\hat{\theta}_n \xrightarrow{p} \theta_0$.
- Tällöin, $X_n(\hat{\theta}_n) \xrightarrow{p} c(\theta_0)$.

- **Lause 1.5.** Oletetaan, että
  - $X_n(\theta) \xrightarrow{p} c(\theta)$ tasaisesti joukosssa Θ
  - funktio $c(\theta)$ on jatkuva pisteessä θ_0 ja
  - $\hat{\theta}_n \xrightarrow{p} \theta_0$.
- Tällöin, $X_n(\hat{\theta}_n) \xrightarrow{p} c(\theta_0)$.
- Sovelletaan myöhemmin, kun $X_n(\theta)$ on $n^{-1}\times$ havaittu informaatiomatriisi ja siten tyyppiä

$$X_n(\theta) = n^{-1}\sum_{i=1}^{n} q(Y_i;\theta) \qquad \text{tai} \qquad X_n(\theta) = n^{-1}\sum_{i=1}^{n} q_i(Y_i;\theta).$$

Tn–laskennan kurssilla on esitetty iid–jonojen suurten lukujen laki.

- Jos $X_1, X_2, \ldots$ on reaalinen iid–jono, $\mathrm{E}\left(X_i\right) = \mu$ ja $\mathrm{Var}\left(X_i\right) = \sigma^2 < \infty$, niin

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \xrightarrow{p} \mu.$$

Todistus seuraa helposti Tšebyševin (tai Markovin) epäyhtälöstä.

Tn–laskennan kurssilla on esitetty iid–jonojen suurten lukujen laki.

- Jos $X_1, X_2, \ldots$ on reaalinen iid–jono, $\mathrm{E}\left(X_i\right) = \mu$ ja $\mathrm{Var}\left(X_i\right) = \sigma^2 < \infty$, niin

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mu.$$

  Todistus seuraa helposti Tšebyševin (tai Markovin) epäyhtälöstä.

- Tilastollisen päättelyn sovelluksissa riippumattomuus ja/tai samoin jakautuneisuus eivät useinkaan päde.

- Riippumattomuutta ei yleensä voida perustella aikasarja-aineistoissa (esim. päivittäinen lämpötila Kaisaniemessä kesäkuussa 2011).

- Riippumattomuutta ei yleensä voida perustella aikasarja-aineistoissa (esim. päivittäinen lämpötila Kaisaniemessä kesäkuussa 2011).
- Samoin lääketieteellisissä kokeissa, joissa havaintoina on mittaustuloksia samoista koehenkilöistä eri ajankohtina. Tällöin eri koehenkilöiden mittaustuloksia voidaan kuitenkin pitää yleensä riippumattomina.

- Riippumattomuutta ei yleensä voida perustella aikasarja-aineistoissa (esim. päivittäinen lämpötila Kaisaniemessä kesäkuussa 2011).
- Samoin lääketieteellisissä kokeissa, joissa havaintoina on mittaustuloksia samoista koehenkilöistä eri ajankohtina. Tällöin eri koehenkilöiden mittaustuloksia voidaan kuitenkin pitää yleensä riippumattomina.
- Regressiomallien tyyppisissä malleissa havaintoja ei voida useinkaan pitää samoin jakautuneina (pätee esim. edelliseen tilanteeseen).

- Riippumattomuutta ei yleensä voida perustella aikasarja-aineistoissa (esim. päivittäinen lämpötila Kaisaniemessä kesäkuussa 2011).
- Samoin lääketieteellisissä kokeissa, joissa havaintoina on mittaustuloksia samoista koehenkilöistä eri ajankohtina. Tällöin eri koehenkilöiden mittaustuloksia voidaan kuitenkin pitää yleensä riippumattomina.
- Regressiomallien tyyppisissä malleissa havaintoja ei voida useinkaan pitää samoin jakautuneina (pätee esim. edelliseen tilanteeseen).
- Miten SLL toimii tällaisissa tapauksissa?

- Oletetaan ensin, että jono $X_1, X_2, \ldots$ on $\underline{\perp\!\!\!\perp}$, mutta $\mathrm{E}\left(X_i\right) = \mu_i$ ja $\mathrm{Var}\left(X_i\right) = \sigma_i^2 < \infty$ vaihtelevat.

- Oletetaan ensin, että jono $X_1, X_2, \ldots$ on $\underline{\ \perp\!\!\!\perp\ }$, mutta $\mathrm{E}\left(X_i\right) = \mu_i$ ja $\mathrm{Var}\left(X_i\right) = \sigma_i^2 < \infty$ vaihtelevat.
- Tällöin $\mathrm{E}\left(\bar{X}_n\right) = n^{-1} \sum_{i=1}^n \mu_i$ ja $\mathrm{Var}\left(\bar{X}_n\right) = n^{-2} \sum_{i=1}^n \sigma_i^2$ ($\underline{\ \perp\!\!\!\perp\ }$).

- Oletetaan ensin, että jono $X_1, X_2, \ldots$ on $\underline{\perp\!\!\!\perp}$, mutta $\mathrm{E}(X_i) = \mu_i$ ja $\mathrm{Var}(X_i) = \sigma_i^2 < \infty$ vaihtelevat.
- Tällöin $\mathrm{E}(\bar{X}_n) = n^{-1}\sum_{i=1}^n \mu_i$ ja $\mathrm{Var}(\bar{X}_n) = n^{-2}\sum_{i=1}^n \sigma_i^2$ ($\underline{\perp\!\!\!\perp}$).
- Tšebyševin epäyhtälö $\Rightarrow$

$$\mathrm{P}\left\{\left|\bar{X}_n - \frac{1}{n}\sum_{i=1}^n \mu_i\right| > \varepsilon\right\} \le \frac{1}{\varepsilon^2 n^2}\sum_{i=1}^n \sigma_i^2 .$$

- Oletetaan ensin, että jono $X_1, X_2, \ldots$ on $\underline{\perp\!\!\!\perp}$, mutta $\mathrm{E}\,(X_i) = \mu_i$ ja $\mathrm{Var}\,(X_i) = \sigma_i^2 < \infty$ vaihtelevat.
- Tällöin $\mathrm{E}\,(\bar{X}_n) = n^{-1} \sum_{i=1}^{n} \mu_i$ ja $\mathrm{Var}\,(\bar{X}_n) = n^{-2} \sum_{i=1}^{n} \sigma_i^2$ ($\underline{\perp\!\!\!\perp}$).
- Tšebyševin epäyhtälö $\Rightarrow$

$$\mathrm{P}\left\{ \left| \bar{X}_n - \frac{1}{n} \sum_{i=1}^{n} \mu_i \right| > \varepsilon \right\} \leq \frac{1}{\varepsilon^2 n^2} \sum_{i=1}^{n} \sigma_i^2 .$$

- Jos (esim.) $\sigma_i^2 \leq C < \infty$ kaikilla i, pätee SLL muodossa

$$\bar{X}_n - \frac{1}{n} \sum_{i=1}^{n} \mu_i \xrightarrow{p} 0.$$

- Oletetaan nyt, että

$$\bar{X}_n - \frac{1}{n}\sum_{i=1}^{n}\mu_i \xrightarrow{p} 0.$$

- Oletetaan nyt, että

$$\bar{X}_n - \frac{1}{n}\sum_{i=1}^{n}\mu_i \xrightarrow{p} 0.$$

- Jos oletetaan lisäksi konvergenssi $n^{-1}\sum_{i=1}^{n}\mu_i \to \mu$, saadaan (ks. Lause 1.1)

$$\bar{X}_n \xrightarrow{p} \mu.$$

- Oletetaan nyt, että

$$\bar{X}_n - \frac{1}{n}\sum_{i=1}^{n}\mu_i \xrightarrow{p} 0.$$

- Jos oletetaan lisäksi konvergenssi $n^{-1}\sum_{i=1}^{n}\mu_i \to \mu$, saadaan (ks. Lause 1.1)

$$\bar{X}_n \xrightarrow{p} \mu.$$

- Tulkitsemalla tässä $\bar{X}_n$ satunnaisvektorin komponentista muodostetuksi keskiarvoksi saadaan perustelu vastaavalle moniulotteiselle yleistykselle.

- Jono sv:ta $X_1, X_2, \ldots$ on *m–riippuva*, jos kaikilla $k \geq 1$ pätee

$$(X_1, \ldots, X_k) \perp\!\!\!\perp (X_{k+l}, X_{k+l+1}, \ldots), \quad \text{kun} \quad l > m.$$

- Jono sv:ta $X_1, X_2, \dots$ on *m–riippuva*, jos kaikilla $k \geq 1$ pätee

$$(X_1, \dots, X_k) \perp\!\!\!\perp (X_{k+l}, X_{k+l+1}, \dots), \quad \text{kun} \quad l > m.$$

- Kun alaindeksi viittaa havaintoyksikköön, niin havainnot, joiden välinen "etäisyys" on suurempi kuin m ovat riippumattomia.

- Jono sv:ta $X_1, X_2, \dots$ on *m–riippuva*, jos kaikilla $k \geq 1$ pätee

$$(X_1, \dots, X_k) \perp\!\!\!\perp (X_{k+l}, X_{k+l+1}, \dots), \quad \text{kun} \quad l > m.$$

- Kun alaindeksi viittaa havaintoyksikköön, niin havainnot, joiden välinen "etäisyys" on suurempi kuin m ovat riippumattomia.
- Luonteva oletus, kun havainnot saadaan riippumattomista koehenkilöistä ja kustakin koehenkilöstä on korkeintaan m mittausta.

- Jono sv:ta $X_1, X_2, \ldots$ on *m–riippuva*, jos kaikilla $k \geq 1$ pätee

$$(X_1, \ldots, X_k) \perp\!\!\!\perp (X_{k+l}, X_{k+l+1}, \ldots), \quad \text{kun} \quad l > m.$$

- Kun alaindeksi viittaa havaintoyksikköön, niin havainnot, joiden välinen "etäisyys" on suurempi kuin m ovat riippumattomia.
- Luonteva oletus, kun havainnot saadaan riippumattomista koehenkilöistä ja kustakin koehenkilöstä on korkeintaan m mittausta.
- Aikasarjatapauksessa riippumattomuus pätee kaikille tapahtumille, joiden välinen ajallinen etäisyys on vähintään m aikayksikköä (ei välttämättä luontevaa mainitussa lämpötilaesimerkissä).

- Olkoon sm–jono $X_1, X_2, \ldots$ *m–riippuva* ja $\mathrm{E}\left(X_i\right) = \mu_i$ ja $\mathrm{Var}\left(X_i\right) = \sigma_i^2 \leq C < \infty$ kaikilla i.

- Olkoon sm–jono $X_1, X_2, \ldots$ *m–riippuva* ja $\mathrm{E}\left(X_i\right) = \mu_i$ ja $\mathrm{Var}\left(X_i\right) = \sigma_i^2 \leq C < \infty$ kaikilla i.

- Kuten riippumattomassa tapauksessa

$$\bar{X}_n - \frac{1}{n}\sum_{i=1}^{n}\mu_i \xrightarrow{p} 0 \qquad \text{(vas. odotusarvo } = 0\text{ !)}$$

pätee, jos

$$\mathrm{Var}\left(\bar{X}_n\right) \underset{n\to\infty}{\longrightarrow} 0.$$

- Olkoon sm–jono $X_1, X_2, \ldots$ *m–riippuva* ja $\mathrm{E}\left(X_i\right) = \mu_i$ ja $\mathrm{Var}\left(X_i\right) = \sigma_i^2 \leq C < \infty$ kaikilla i.

- Kuten riippumattomassa tapauksessa

$$\bar{X}_n - \frac{1}{n}\sum_{i=1}^{n} \mu_i \xrightarrow{p} 0 \qquad (\text{vas. odotusarvo } = 0\ !)$$

  pätee, jos

$$\mathrm{Var}\left(\bar{X}_n\right) \underset{n\to\infty}{\longrightarrow} 0.$$

- Tämä todetaan suoralla laskulla.

- Oletettiin siis, että sm–jono $X_1, X_2, \ldots$ on *m–riippuva* ja $\mathrm{E}\left(X_i\right) = \mu_i$ ja $\mathrm{Var}\left(X_i\right) = \sigma_i^2 \leq C < \infty$ kaikilla i.

- Oletettiin siis, että sm–jono $X_1, X_2, \ldots$ on *m–riippuva* ja $\mathrm{E}(X_i) = \mu_i$ ja $\mathrm{Var}(X_i) = \sigma_i^2 \leq C < \infty$ kaikilla i.

- Todettiin, että tällöin

$$\bar{X}_n - \frac{1}{n}\sum_{i=1}^{n}\mu_i \xrightarrow{p} 0.$$

- Oletettiin siis, että sm–jono $X_1, X_2, \ldots$ on *m–riippuva* ja $\mathrm{E}(X_i) = \mu_i$ ja $\mathrm{Var}(X_i) = \sigma_i^2 \leq C < \infty$ kaikilla i.

- Todettiin, että tällöin

$$\bar{X}_n - \frac{1}{n}\sum_{i=1}^{n} \mu_i \xrightarrow{p} 0.$$

- Jos oletetaan lisäksi konvergenssi $n^{-1}\sum_{i=1}^{n} \mu_i \to \mu$, saadaan

$$\bar{X}_n \xrightarrow{p} \mu.$$

- Oletettiin siis, että sm–jono $X_1, X_2, \ldots$ on *m–riippuva* ja $\mathrm{E}\left(X_i\right) = \mu_i$ ja $\mathrm{Var}\left(X_i\right) = \sigma_i^2 \leq C < \infty$ kaikilla i.

- Todettiin, että tällöin
$$\bar{X}_n - \frac{1}{n}\sum_{i=1}^{n}\mu_i \xrightarrow{p} 0.$$

- Jos oletetaan lisäksi konvergenssi $n^{-1}\sum_{i=1}^{n}\mu_i \to \mu$, saadaan
$$\bar{X}_n \xrightarrow{p} \mu.$$

- Tulkitsemalla taas $\bar{X}_n$ satunnaisvektorin komponentista muodostetuksi keskiarvoksi saadaan perustelu vastaavalle moniulotteiselle SLL:lle.

- SLL voidaan osoittaa myös tapauksissa, joissa $X_i \perp\!\!\!\perp X_j$ ei päde olipa $|i - j|$ kuinka suuri tahansa.

- SLL voidaan osoittaa myös tapauksissa, joissa $X_i \perp\!\!\!\perp X_j$ ei päde olipa $|i-j|$ kuinka suuri tahansa.

- Vaatimuksena on kuitenkin, että X_i:n ja X_j:n välinen riippuvuus (tai korrelaatio) "häviää", kun $|i-j| \to \infty$.

- SLL voidaan osoittaa myös tapauksissa, joissa $X_i \perp\!\!\!\perp X_j$ ei päde olipa $|i-j|$ kuinka suuri tahansa.
- Vaatimuksena on kuitenkin, että X_i:n ja X_j:n välinen riippuvuus (tai korrelaatio) "häviää", kun $|i-j| \to \infty$.
- Tälle intuitiiviselle luonnehdinnalle on useita matemaattisia muotoilija, joita käyttäen SLL:ja on todistettu sallimalla riippuvuuden lisäksi mahdollisesti myös jakaumien heterogeenisuus.

- SLL voidaan osoittaa myös tapauksissa, joissa $X_i \perp\!\!\!\perp X_j$ ei päde olipa $|i-j|$ kuinka suuri tahansa.
- Vaatimuksena on kuitenkin, että X_i:n ja X_j:n välinen riippuvuus (tai korrelaatio) "häviää", kun $|i-j| \to \infty$.
- Tälle intuitiiviselle luonnehdinnalle on useita matemaattisia muotoilija, joita käyttäen SLL:ja on todistettu sallimalla riippuvuuden lisäksi mahdollisesti myös jakaumien heterogeenisuus.
- SLL ei kuitenkaan päde, jos riippuvuus ja/tai jakaumien heterogeenisuus on "liian" voimakasta.

- Tasainen SLL pätee, jos jonolle satunnaisfunktioita $X_1(\theta), X_2(\theta), \ldots$, $\theta \in \Theta$, pätee

$$\bar{X}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} X_i(\theta) \xrightarrow{p} \mu(\theta) \quad \text{tasaisesti joukossa } \Theta.$$

Tässä $\mu(\theta)$ tulkitaan funktiojonon $n^{-1}\sum_{i=1}^{n} \mathrm{E}[X_i(\theta)]$ raja-arvoksi.

- Tasainen SLL pätee, jos jonolle satunnaisfunktioita $X_1(\theta), X_2(\theta), \ldots$, $\theta \in \Theta$, pätee

$$\bar{X}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} X_i(\theta) \xrightarrow{p} \mu(\theta) \quad \text{tasaisesti joukossa } \Theta.$$

Tässä $\mu(\theta)$ tulkitaan funktiojonon $n^{-1}\sum_{i=1}^{n} \mathrm{E}[X_i(\theta)]$ raja-arvoksi.

- Ts.,

$$\sup_{\theta \in \Theta} \left|\bar{X}_n(\theta) - \mu(\theta)\right| \xrightarrow{p} 0.$$

- Vaikka pisteittäinen SLL $\bar{X}_n(\theta) \xrightarrow{p} \mu(\theta)$ (jokaisella $\theta \in \Theta$) voidaan usein todeta kohtuullisen helposti, on tasaisten SSL:ien todistaminen tai edes riittävien ehtojen esittäminen matemaattisesti paljon hankalampaa.

- Vaikka pisteittäinen SLL $\bar{X}_n(\theta) \xrightarrow{p} \mu(\theta)$ (jokaisella $\theta \in \Theta$) voidaan usein todeta kohtuullisen helposti, on tasaisten SSL:ien todistaminen tai edes riittävien ehtojen esittäminen matemaattisesti paljon hankalampaa.
- iid–tapauksessa ja kun $X_i(\theta) = q(Y_i;\theta)$ riittää, että

- Vaikka pisteittäinen SLL $\bar{X}_n(\theta) \xrightarrow{p} \mu(\theta)$ (jokaisella $\theta \in \Theta$) voidaan usein todeta kohtuullisen helposti, on tasaisten SSL:ien todistaminen tai edes riittävien ehtojen esittäminen matemaattisesti paljon hankalampaa.

- iid–tapauksessa ja kun $X_i(\theta) = q(Y_i; \theta)$ riittää, että
    - (i) funktio $\theta \mapsto q(y; \theta)$ on jatkuva kaikilla Y_i:n mahdollisilla arvoilla y,

- Vaikka pisteittäinen SLL $\bar{X}_n(\theta) \xrightarrow{p} \mu(\theta)$ (jokaisella $\theta \in \Theta$) voidaan usein todeta kohtuullisen helposti, on tasaisten SSL:ien todistaminen tai edes riittävien ehtojen esittäminen matemaattisesti paljon hankalampaa.

- iid–tapauksessa ja kun $X_i(\theta) = q(Y_i; \theta)$ riittää, että
  - (i) funktio $\theta \mapsto q(y; \theta)$ on jatkuva kaikilla Y_i:n mahdollisilla arvoilla y,
  - (ii) joukko Θ on kompakti (eli suljettu ja rajoitettu) ja

- Vaikka pisteittäinen SLL $\bar{X}_n(\theta) \xrightarrow{p} \mu(\theta)$ (jokaisella $\theta \in \Theta$) voidaan usein todeta kohtuullisen helposti, on tasaisten SSL:ien todistaminen tai edes riittävien ehtojen esittäminen matemaattisesti paljon hankalampaa.

- iid–tapauksessa ja kun $X_i(\theta) = q(Y_i;\theta)$ riittää, että
  - (i) funktio $\theta \mapsto q(y;\theta)$ on jatkuva kaikilla Y_i:n mahdollisilla arvoilla y,
  - (ii) joukko Θ on kompakti (eli suljettu ja rajoitettu) ja
  - (iii) $\mathrm{E}\left(\sup_{\theta\in\Theta} |q(Y_1;\theta)|\right) < \infty$.

Tn–laskennan kurssilla on esitetty reaalisten iid–jonojen KRL.

- Jos $X_1, X_2, \ldots$ on reaalinen iid–jono, $\mathrm{E}\left(X_i\right) = \mu$ ja $\mathrm{Var}\left(X_i\right) = \sigma^2 < \infty$, niin

  $$\sqrt{n}\left(\bar{X}_n - \mu\right) \xrightarrow{d} \mathrm{N}\left(0, \sigma^2\right)$$

  tai vaihtoehtoisesti esitettynä (ks. Lause 1.3)

  $$\sqrt{n}\left(\bar{X}_n - \mu\right)/\sigma \xrightarrow{d} \mathrm{N}\left(0, 1\right).$$

Tn–laskennan kurssilla on esitetty reaalisten iid–jonojen KRL.

- Jos $X_1, X_2, \ldots$ on reaalinen iid–jono, $\mathrm{E}(X_i) = \mu$ ja $\mathrm{Var}(X_i) = \sigma^2 < \infty$, niin

$$\sqrt{n}\left(\bar{X}_n - \mu\right) \xrightarrow{d} \mathrm{N}\left(0, \sigma^2\right)$$

tai vaihtoehtoisesti esitettynä (ks. Lause 1.3)

$$\sqrt{n}\left(\bar{X}_n - \mu\right)/\sigma \xrightarrow{d} \mathrm{N}(0, 1).$$

- Käytännön merkitys: Otoskeskiarvon $\bar{X}_n$ (tuntematonta) jakaumaa voidaan "suurilla" n:n arvoilla approksimoida $\mathrm{N}\left(\mu, \sigma^2/n\right)$–jakaumalla tai "merkinnällisesti"

$$\bar{X}_n \underset{as}{\sim} \mathrm{N}\left(\mu, \sigma^2/n\right).$$

- Jos $X_1, X_2, \ldots$ on iid–jono sv:ta $(k \times 1)$, $\mathrm{E}(X_i) = \mu$ $(k \times 1)$ ja $\mathrm{Cov}(X_i) = \Sigma$ $(k \times k)$ (äärellinen ja nollasta poikkeava), niin

$$\sqrt{n}\left(\bar{X}_n - \mu\right) \xrightarrow{d} \mathrm{N}\left(0, \Sigma\right).$$

- Jos $X_1, X_2, \ldots$ on iid–jono sv:ta $(k \times 1)$, $\mathrm{E}\,(X_i) = \mu$ $(k \times 1)$ ja $\mathrm{Cov}\,(X_i) = \Sigma$ $(k \times k)$ (äärellinen ja nollasta poikkeava), niin

$$\sqrt{n}\,(\bar{X}_n - \mu) \xrightarrow{d} \mathrm{N}\,(0, \Sigma)\,.$$

- Myöhemmin Σ on (ainakin yleensä) epäsingulaarinen. Joissakin sovelluksissa voidaan kuitenkin päätyä singulaariseen raja-jakaumaan.

Raja-arvolauseita

Keskeinen raja-arvolause (KRL); ei-iid–tapaus

- KRL:sta on olemassa lukuisia versioita, joissa sallitaan jakaumien heterogeenisuus ja muuttujien riippuvuus kuten esim. m–riippuvuus.

- KRL:sta on olemassa lukuisia versioita, joissa sallitaan jakaumien heterogeenisuus ja muuttujien riippuvuus kuten esim. *m*–riippuvuus.
- Tällaiset KRL:t esitetään (reaaliarvoisessa tapauksessa) usein muodossa

$$\left(\bar{X}_n - \mu_n\right)/\sigma_n \xrightarrow{d} \mathrm{N}\left(0,1\right),$$

jossa $\mu_n = \mathrm{E}\left(\bar{X}_n\right)$ ja $\sigma_n^2 = \mathrm{Var}\left(\bar{X}_n\right)$ eli "merkinnällisesti"

$$\bar{X}_n \underset{as}{\sim} \mathrm{N}\left(\mu_n, \sigma_n^2\right).$$

- KRL:sta on olemassa lukuisia versioita, joissa sallitaan jakaumien heterogeenisuus ja muuttujien riippuvuus kuten esim. *m*–riippuvuus.
- Tällaiset KRL:t esitetään (reaaliarvoisessa tapauksessa) usein muodossa

$$\left(\bar{X}_n - \mu_n\right)/\sigma_n \xrightarrow{d} \mathrm{N}\left(0,1\right),$$

jossa $\mu_n = \mathrm{E}\left(\bar{X}_n\right)$ ja $\sigma_n^2 = \mathrm{Var}\left(\bar{X}_n\right)$ eli "merkinnällisesti"

$$\bar{X}_n \underset{as}{\sim} \mathrm{N}\left(\mu_n, \sigma_n^2\right).$$

- KRL ei päde, jos "poikkeavuus iid–tapauksesta on liiallista".

- KRL:sta on olemassa lukuisia versioita, joissa sallitaan jakaumien heterogeenisuus ja muuttujien riippuvuus kuten esim. m–riippuvuus.
- Tällaiset KRL:t esitetään (reaaliarvoisessa tapauksessa) usein muodossa

$$\left(\bar{X}_n - \mu_n\right)/\sigma_n \xrightarrow{d} \mathrm{N}\left(0,1\right),$$

jossa $\mu_n = \mathrm{E}\left(\bar{X}_n\right)$ ja $\sigma_n^2 = \mathrm{Var}\left(\bar{X}_n\right)$ eli "merkinnällisesti"

$$\bar{X}_n \underset{as}{\sim} \mathrm{N}\left(\mu_n, \sigma_n^2\right).$$

- KRL ei päde, jos "poikkeavuus iid–tapauksesta on liiallista".
- Seuraavassa esitettävä ns. **martingaalien KRL** ja sen eri versiot ovat käyttökelppoisia (erityisesti uskottavuusfunktioon perustuvassa) tilastollisessa päättelyssä.

- Jos sv:lla Y ja X on jatkuva yhteisjakauma ja $\mathrm{E}(|Y|) < \infty$, määritellään Y:n ehdollinen odotusarvo ehdolla $X = x$ yhtälöllä

$$\mathrm{E}(Y|X = x) = \int_{-\infty}^{\infty} y \frac{f_{Y,X}(y,x)}{f_X(x)} dy = \int_{-\infty}^{\infty} y f_{Y|X}(y;x)\, dy.$$

- Jos sv:lla Y ja X on jatkuva yhteisjakauma ja $\mathrm{E}(|Y|) < \infty$, määritellään Y:n ehdollinen odotusarvo ehdolla $X = x$ yhtälöllä

$$\mathrm{E}(Y|X = x) = \int_{-\infty}^{\infty} y \frac{f_{Y,X}(y,x)}{f_X(x)} dy = \int_{-\infty}^{\infty} y f_{Y|X}(y;x)\, dy.$$

  - $f_{Y,X}(y,x)$ on (Y,X):n yhteistiheysfunktio

- Jos sv:lla Y ja X on jatkuva yhteisjakauma ja $\mathrm{E}\left(|Y|\right) < \infty$, määritellään Y:n ehdollinen odotusarvo ehdolla $X = x$ yhtälöllä

$$\mathrm{E}\left(Y \,|X = x\right) = \int_{-\infty}^{\infty} y \frac{f_{Y,X}\left(y,x\right)}{f_X\left(x\right)} dy = \int_{-\infty}^{\infty} y f_{Y|X}\left(y;x\right) dy.$$

  - $f_{Y,X}\left(y,x\right)$ on $\left(Y,X\right)$:n yhteistiheysfunktio
  - $f_X\left(x\right) = \int_{-\infty}^{\infty} f_{Y,X}\left(y,x\right) dy$ on X:n reunajakauman tf ja

- Jos sv:lla Y ja X on jatkuva yhteisjakauma ja $\mathrm{E}\left(|Y|\right) < \infty$, määritellään Y:n ehdollinen odotusarvo ehdolla $X = x$ yhtälöllä

$$\mathrm{E}\left(Y \,|X = x\right) = \int_{-\infty}^{\infty} y \frac{f_{Y,X}\left(y,x\right)}{f_X\left(x\right)} dy = \int_{-\infty}^{\infty} y f_{Y|X}\left(y;x\right) dy.$$

  - $f_{Y,X}\left(y,x\right)$ on $\left(Y,X\right)$:n yhteistiheysfunktio
  - $f_X\left(x\right) = \int_{-\infty}^{\infty} f_{Y,X}\left(y,x\right) dy$ on X:n reunajakauman tf ja
  - $f_{Y|X}\left(y|x\right) = f_{Y,X}\left(y,x\right) / f_X\left(x\right)$ on Y:n ehdollinen tiheysfunktio ehdolla $X = x$.

- Jos sv:lla Y ja X on jatkuva yhteisjakauma ja $\mathrm{E}(|Y|) < \infty$, määritellään Y:n ehdollinen odotusarvo ehdolla $X = x$ yhtälöllä

$$\mathrm{E}(Y|X = x) = \int_{-\infty}^{\infty} y \frac{f_{Y,X}(y,x)}{f_X(x)} dy = \int_{-\infty}^{\infty} y f_{Y|X}(y;x)\, dy.$$

  - $f_{Y,X}(y,x)$ on (Y,X):n yhteistiheysfunktio
  - $f_X(x) = \int_{-\infty}^{\infty} f_{Y,X}(y,x)\, dy$ on X:n reunajakauman tf ja
  - $f_{Y|X}(y|x) = f_{Y,X}(y,x) / f_X(x)$ on Y:n ehdollinen tiheysfunktio ehdolla $X = x$.

- Kun x vaihtelee, määrittelee x:n funktio $\mathrm{E}(Y|X = x)$ sv:n, jonka luonteva merkintä on $\mathrm{E}(Y|X)$. Tätä merkintää käytetään seuraavassa.

Ehdollisella odotusarvolla on tunnetusti mm. seuraavat ominaisuudet.

- **EO1** $\mathrm{E}\left(aY_1 + bY_2 \,|X\right) = a\mathrm{E}\left(Y_1 \,|X\right) + b\mathrm{E}\left(Y_2 \,|X\right)$, kun a ja b ovat vakioita (lineaarisuus).

Ehdollisella odotusarvolla on tunnetusti mm. seuraavat ominaisuudet.

- **EO1** $\mathrm{E}\left(aY_1 + bY_2 \mid X\right) = a\mathrm{E}\left(Y_1 \mid X\right) + b\mathrm{E}\left(Y_2 \mid X\right)$, kun a ja b ovat vakioita (lineaarisuus).
- **EO2** $\mathrm{E}\left(Y \mid X\right) = \mathrm{E}\left(Y\right)$, kun $Y \perp\!\!\!\perp X$.

Ehdollisella odotusarvolla on tunnetusti mm. seuraavat ominaisuudet.

- **EO1** $\mathrm{E}(aY_1 + bY_2 \,|X) = a\mathrm{E}(Y_1 \,|X) + b\mathrm{E}(Y_2 \,|X)$, kun a ja b ovat vakioita (lineaarisuus).
- **EO2** $\mathrm{E}(Y \,|X) = \mathrm{E}(Y)$, kun $Y \perp\!\!\!\perp X$.
- **EO3** $\mathrm{E}(Y) = \mathrm{E}[\mathrm{E}(Y \,|X)]$ (ns. iteroidun odotusarvon laki)

Ehdollisella odotusarvolla on tunnetusti mm. seuraavat ominaisuudet.

- **EO1** $\mathrm{E}\left(aY_1 + bY_2 \,|X\right) = a\mathrm{E}\left(Y_1 \,|X\right) + b\mathrm{E}\left(Y_2 \,|X\right)$, kun a ja b ovat vakioita (lineaarisuus).
- **EO2** $\mathrm{E}\left(Y \,|X\right) = \mathrm{E}\left(Y\right)$, kun $Y \perp\!\!\!\perp X$.
- **EO3** $\mathrm{E}\left(Y\right) = \mathrm{E}\left[\mathrm{E}\left(Y \,|X\right)\right]$ (ns. iteroidun odotusarvon laki)
- **EO4** $\mathrm{E}\left(g\left(X\right) Y \,|X\right) = g\left(X\right)\mathrm{E}\left(Y \,|X\right)$ mille tahansa funktiolle g (jolla vasen puoli on hyvin määritelty).

- **Määritelmä 2.1.** Olkoon $Z_1, Z_2, \ldots$ jono sm:ia ja $\mathcal{F}_i = (Z_1, \ldots, Z_i)$. Sm–jonoa $M_1, M_2, \ldots$ sanotaan **martingaaliksi informaatiojoukon** $\mathcal{F}_i$ **suhteen**, jos

- **Määritelmä 2.1.** Olkoon $Z_1, Z_2, \ldots$ jono sm:ia ja $\mathcal{F}_i = (Z_1, \ldots, Z_i)$. Sm–jonoa $M_1, M_2, \ldots$ sanotaan **martingaaliksi informaatiojoukon** $\mathcal{F}_i$ **suhteen**, jos
- **(i)** $M_i = g_i(\mathcal{F}_i)$ jollain funktiolla g_i ja kaikilla $i = 1, 2, \ldots$

- **Määritelmä 2.1.** Olkoon $Z_1, Z_2, ...$ jono sm:ia ja $\mathcal{F}_i = (Z_1, ..., Z_i)$. Sm–jonoa $M_1, M_2, ...$ sanotaan **martingaaliksi informaatiojoukon** $\mathcal{F}_i$ **suhteen**, jos

- **(i)** $M_i = g_i(\mathcal{F}_i)$ jollain funktiolla g_i ja kaikilla $i = 1, 2, ...$

- **(ii)** $\mathrm{E}(|M_i|) < \infty$ kaikilla $i = 1, 2, ...$

- **Määritelmä 2.1.** Olkoon $Z_1, Z_2, ...$ jono sm:ia ja $\mathcal{F}_i = (Z_1, ..., Z_i)$. Sm–jonoa $M_1, M_2, ...$ sanotaan **martingaaliksi informaatiojoukon** $\mathcal{F}_i$ **suhteen**, jos
- **(i)** $M_i = g_i(\mathcal{F}_i)$ jollain funktiolla g_i ja kaikilla $i = 1, 2, ...$
- **(ii)** $\mathrm{E}(|M_i|) < \infty$ kaikilla $i = 1, 2, ...$
- **(iii)** $\mathrm{E}(M_i \,|\mathcal{F}_{i-1}) = M_{i-1}$ kaikilla $i = 2, 3, ...$

- **Määritelmä 2.1.** Olkoon $Z_1, Z_2, ...$ jono sm:ia ja $\mathcal{F}_i = (Z_1, ..., Z_i)$. Sm–jonoa $M_1, M_2, ...$ sanotaan **martingaaliksi informaatiojoukon** $\mathcal{F}_i$ **suhteen**, jos

- **(i)** $M_i = g_i(\mathcal{F}_i)$ jollain funktiolla g_i ja kaikilla $i = 1, 2, \ldots$

- **(ii)** $\mathrm{E}(|M_i|) < \infty$ kaikilla $i = 1, 2, \ldots$

- **(iii)** $\mathrm{E}(M_i \,|\mathcal{F}_{i-1}) = M_{i-1}$ kaikilla $i = 2, 3, \ldots$

- Yleistys tapaukseen, jossa Z_i ja M_i ovat vektoreita on samanlainen, kun kohdassa (ii) itseisarvo korvataan normilla $\|M_i\|$ (tai odotusarvon äärellisyys vaaditaan kaikilta sv:n M_i komponenteilta).

Martingaalin asemesta lähtökohtana on usein martingaalidifferenssi. Sm–jonoa $X_1, X_2, \ldots$ sanotaan **martingaalidifferenssiksi informaatiojoukon** $\mathcal{F}_i$ **suhteen**, jos

- **(i)** $X_i = h_i(\mathcal{F}_i)$ jollain funktiolla h_i ja kaikilla $i = 1, 2, \ldots$

Martingaalin asemesta lähtökohtana on usein martingaalidifferenssi. Sm–jonoa $X_1, X_2, \ldots$ sanotaan **martingaalidifferenssiksi informaatiojoukon** $\mathcal{F}_i$ **suhteen**, jos

- **(i)** $X_i = h_i(\mathcal{F}_i)$ jollain funktiolla h_i ja kaikilla $i = 1, 2, \ldots$
- **(ii)** $\mathrm{E}(|X_i|) < \infty$ kaikilla $i = 1, 2, \ldots$

Martingaalin asemesta lähtökohtana on usein martingaalidifferenssi. Sm–jonoa $X_1, X_2, \ldots$ sanotaan **martingaalidifferenssiksi informaatiojoukon $\mathcal{F}_i$ suhteen**, jos

- **(i)** $X_i = h_i(\mathcal{F}_i)$ jollain funktiolla h_i ja kaikilla $i = 1, 2, \ldots$
- **(ii)** $\mathrm{E}(|X_i|) < \infty$ kaikilla $i = 1, 2, \ldots$
- **(iii)** $\mathrm{E}(X_i \,|\mathcal{F}_{i-1}) = 0$ kaikilla $i = 2, 3, \ldots$

Martingaalin asemesta lähtökohtana on usein martingaalidifferenssi. Sm–jonoa $X_1, X_2, \ldots$ sanotaan **martingaalidifferenssiksi informaatiojoukon $\mathcal{F}_i$ suhteen**, jos

- **(i)** $X_i = h_i(\mathcal{F}_i)$ jollain funktiolla h_i ja kaikilla $i = 1, 2, \ldots$
- **(ii)** $\mathrm{E}(|X_i|) < \infty$ kaikilla $i = 1, 2, \ldots$
- **(iii)** $\mathrm{E}(X_i \mid \mathcal{F}_{i-1}) = 0$ kaikilla $i = 2, 3, \ldots$

- Tällöin

$$M_i = X_1 + \cdots + X_i = M_{i-1} + X_i \qquad (M_0 = 0)$$

on martingaali ja, jos $\mathrm{E}(X_1) = 0$, niin $\mathrm{E}(X_i) = 0$ kaikilla $i \geq 1$.

Martingaalin asemesta lähtökohtana on usein martingaalidifferenssi. Sm–jonoa $X_1, X_2, \ldots$ sanotaan **martingaalidifferenssiksi informaatiojoukon $\mathcal{F}_i$ suhteen**, jos

- **(i)** $X_i = h_i(\mathcal{F}_i)$ jollain funktiolla h_i ja kaikilla $i = 1, 2, \ldots$
- **(ii)** $\mathrm{E}(|X_i|) < \infty$ kaikilla $i = 1, 2, \ldots$
- **(iii)** $\mathrm{E}(X_i \,|\mathcal{F}_{i-1}) = 0$ kaikilla $i = 2, 3, \ldots$

- Tällöin

$$M_i = X_1 + \cdots + X_i = M_{i-1} + X_i \qquad (M_0 = 0)$$

  on martingaali ja, jos $\mathrm{E}(X_1) = 0$, niin $\mathrm{E}(X_i) = 0$ kaikilla $i \geq 1$.
- MD–jonon korreloimattomuudesta seuraa, että reaalisessa tapauksessa $\mathrm{Var}(M_n) = \sum_{i=1}^n \mathrm{Var}(X_i)$.

- **Esimerkki 1.1. (i)** Reaalinen iid–jono $\{X_i\}_{i=1}^{\infty}$ on MD, kun $\mathrm{E}(X_1) = 0$ ja $\mathcal{F}_i = (X_1, ..., X_i)$.

- **Esimerkki 1.1. (i)** Reaalinen iid–jono $\{X_i\}_{i=1}^{\infty}$ on MD, kun $\mathrm{E}(X_1) = 0$ ja $\mathcal{F}_i = (X_1, ..., X_i)$.
- **(ii)** $\{X_i\}_{i=1}^{\infty}$ ja $\{Z_i\}_{i=1}^{\infty}$ $\perp\!\!\!\perp$ ja $\{X_i\}_{i=1}^{\infty}$ kuten kohdassa (i).

- **Esimerkki 1.1. (i)** Reaalinen iid–jono $\{X_i\}_{i=1}^{\infty}$ on MD, kun $\mathrm{E}\left(X_1\right) = 0$ ja $\mathcal{F}_i = (X_1, ..., X_i)$.
- **(ii)** $\{X_i\}_{i=1}^{\infty}$ ja $\{Z_i\}_{i=1}^{\infty}$ $\perp\!\!\!\perp$ ja $\{X_i\}_{i=1}^{\infty}$ kuten kohdassa (i).
  - Tällöin $\{Z_i X_i\}_{i=1}^{\infty}$ on MD, kun $\mathrm{E}\left(|Z_i|\right) < \infty$ kaikilla $i \geq 1$ ja $\mathcal{F}_i = (X_1, ..., X_i, Z_1, ..., Z_{i+1})$.

- **Esimerkki 1.1. (i)** Reaalinen iid–jono $\{X_i\}_{i=1}^{\infty}$ on MD, kun $\mathrm{E}(X_1) = 0$ ja $\mathcal{F}_i = (X_1, ..., X_i)$.
- **(ii)** $\{X_i\}_{i=1}^{\infty}$ ja $\{Z_i\}_{i=1}^{\infty}$ $\perp\!\!\!\perp$ ja $\{X_i\}_{i=1}^{\infty}$ kuten kohdassa (i).
  - Tällöin $\{Z_i X_i\}_{i=1}^{\infty}$ on MD, kun $\mathrm{E}(|Z_i|) < \infty$ kaikilla $i \geq 1$ ja $\mathcal{F}_i = (X_1, ..., X_i, Z_1, ..., Z_{i+1})$.
- **(iii)** $\{X_i\}_{i=1}^{\infty}$ kuten kohdassa (i), $Z_1 = c$ (vakio) ja $Z_i = g_i(X_1, ..., X_{i-1})$, $i \geq 2$.

- **Esimerkki 1.1. (i)** Reaalinen iid–jono $\{X_i\}_{i=1}^{\infty}$ on MD, kun $\mathrm{E}(X_1) = 0$ ja $\mathcal{F}_i = (X_1, ..., X_i)$.
- **(ii)** $\{X_i\}_{i=1}^{\infty}$ ja $\{Z_i\}_{i=1}^{\infty}$ $\perp\!\!\!\perp$ ja $\{X_i\}_{i=1}^{\infty}$ kuten kohdassa (i).
  - Tällöin $\{Z_i X_i\}_{i=1}^{\infty}$ on MD, kun $\mathrm{E}(|Z_i|) < \infty$ kaikilla $i \geq 1$ ja $\mathcal{F}_i = (X_1, ..., X_i, Z_1, ..., Z_{i+1})$.
- **(iii)** $\{X_i\}_{i=1}^{\infty}$ kuten kohdassa (i), $Z_1 = c$ (vakio) ja $Z_i = g_i(X_1, ..., X_{i-1})$, $i \geq 2$.
  - Tällöin $\{Z_i X_i\}_{i=1}^{\infty}$ on MD, kun $\mathrm{E}(|Z_i|) < \infty$ kaikilla $i \geq 2$ ja $\mathcal{F}_i = (X_1, ..., X_i)$.

- **Lause 1.6.** Olkoon $X_1, X_2, \ldots$ $(k \times 1)$ MD–jono jonkun informaatiojoukon suhteen. Oletetaan

- **Lause 1.6.** Olkoon $X_1, X_2, \ldots$ $(k \times 1)$ MD–jono jonkun informaatiojoukon suhteen. Oletetaan
- (i) $\mathrm{E}(|X_{a,i}|^{2+\delta}) \leq C < \infty$ jollain $\delta > 0$, $\forall\ i \geq 1$ ja $a = 1, \ldots, k$

- **Lause 1.6.** Olkoon $X_1, X_2, ...$ $(k \times 1)$ MD–jono jonkun informaatiojoukon suhteen. Oletetaan
- (i) $\mathrm{E}(|X_{a,i}|^{2+\delta}) \leq C < \infty$ jollain $\delta > 0$, $\forall\ i \geq 1$ ja $a = 1, ..., k$
- (ii) $n^{-1} \sum_{i=1}^{n} X_i X_i' \xrightarrow{p} \Sigma$ (nollasta poikkeava ja äärellinen),

- **Lause 1.6.** Olkoon $X_1, X_2, \ldots$ $(k \times 1)$ MD–jono jonkun informaatiojoukon suhteen. Oletetaan
- (i) $\mathrm{E}(|X_{a,i}|^{2+\delta}) \leq C < \infty$ jollain $\delta > 0$, $\forall\ i \geq 1$ ja $a = 1, \ldots, k$
- (ii) $n^{-1} \sum_{i=1}^{n} X_i X_i' \xrightarrow{p} \Sigma$ (nollasta poikkeava ja äärellinen),
- Tällöin,

- **Lause 1.6.** Olkoon $X_1, X_2, \ldots$ $(k \times 1)$ MD–jono jonkun informaatiojoukon suhteen. Oletetaan
- (i) $\mathrm{E}(|X_{a,i}|^{2+\delta}) \leq C < \infty$ jollain $\delta > 0$, $\forall$ $i \geq 1$ ja $a = 1, \ldots, k$
- (ii) $n^{-1} \sum_{i=1}^{n} X_i X_i' \xrightarrow{p} \Sigma$ (nollasta poikkeava ja äärellinen),
- Tällöin,
  - $$\mathrm{Cov}\left(n^{-1/2} \sum_{i=1}^{n} X_i\right) \underset{n\to\infty}{\to} \Sigma \quad \Leftrightarrow \quad n^{-1} \sum_{i=1}^{n} \mathrm{E}\left(X_i X_i'\right) \underset{n\to\infty}{\to} \Sigma$$

- **Lause 1.6.** Olkoon $X_1, X_2, \ldots$ $(k \times 1)$ MD–jono jonkun informaatiojoukon suhteen. Oletetaan
- (i) $\mathrm{E}(|X_{a,i}|^{2+\delta}) \leq C < \infty$ jollain $\delta > 0$, $\forall\ i \geq 1$ ja $a = 1, \ldots, k$
- (ii) $n^{-1} \sum_{i=1}^{n} X_i X_i' \xrightarrow{p} \Sigma$ (nollasta poikkeava ja äärellinen),
- Tällöin,
  - $$\mathrm{Cov}\left(n^{-1/2} \sum_{i=1}^{n} X_i\right) \underset{n\to\infty}{\to} \Sigma \quad \Leftrightarrow \quad n^{-1} \sum_{i=1}^{n} \mathrm{E}\left(X_i X_i'\right) \underset{n\to\infty}{\to} \Sigma$$
  - ja $$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \xrightarrow{d} \mathrm{N}\left(0, \Sigma\right).$$

- Tilastollisen tutkimuksen lähtökohtana aineisto $\mathbf{y} = (y_1, ..., y_n)$, jossa y_i on havaintoyksikköön i liittyvä havainto.

- Tilastollisen tutkimuksen lähtökohtana aineisto $\mathbf{y} = (y_1, ..., y_n)$, jossa y_i on havaintoyksikköön i liittyvä havainto.
- Havainnot voivat olla joko reaaliarvoisia tai vektoriarvoisia ja ne tulkitaan sm:ien tai sv:ien $Y_1, ..., Y_n$ havaituiksi arvoiksi.

Uskottavuuspäättelyä

Tilastollinen malli

- Tilastollisen tutkimuksen lähtökohtana aineisto $\mathbf{y} = (y_1, ..., y_n)$, jossa y_i on havaintoyksikköön i liittyvä havainto.
- Havainnot voivat olla joko reaaliarvoisia tai vektoriarvoisia ja ne tulkitaan sm:ien tai sv:ien $Y_1, ..., Y_n$ havaituiksi arvoiksi.
- Esimerkki vektoritapauksesta on $Y_1, ..., Y_n \sim \mathrm{N}_k(\mu, \Sigma) \perp\!\!\!\perp$.

- Tilastollisen tutkimuksen lähtökohtana aineisto $\mathbf{y} = (y_1, \ldots, y_n)$, jossa y_i on havaintoyksikköön i liittyvä havainto.
- Havainnot voivat olla joko reaaliarvoisia tai vektoriarvoisia ja ne tulkitaan sm:ien tai sv:ien $Y_1, \ldots, Y_n$ havaituiksi arvoiksi.
- Esimerkki vektoritapauksesta on $Y_1, \ldots, Y_n \sim \mathrm{N}_k(\mu, \Sigma) \perp\!\!\!\perp$.
- Tilastollinen malli on sv:n $\mathbf{Y} = (Y_1, \ldots, Y_n)$ yhteistiheysfunktio (ytf) tai yhteispistetodennäköisyysfunktio (yptf), joka rajataan käytettävissä olevan taustatiedon perusteella.

- Tarkasteltavassa parametrisessa tilastollisessa päättelyssä oletetaan, että ytf/yptf $f_{\mathbf{Y}}\left(\mathbf{y};\theta\right)$ riippuu tuntemattomasta parametrista $\theta=\left(\theta_1,..,\theta_d\right)\in\Theta\subseteq\mathbb{R}^d$.

- Tarkasteltavassa parametrisessa tilastollisessa päättelyssä oletetaan, että ytf/yptf $f_{\mathbf{Y}}(\mathbf{y};\theta)$ riippuu tuntemattomasta parametrista $\theta = (\theta_1, .., \theta_d) \in \Theta \subseteq \mathbb{R}^d$.
- Mahdollisten parametriarvojen joukon Θ määrittely on osa mallin spesifiointia.

- Tarkasteltavassa parametrisessa tilastollisessa päättelyssä oletetaan, että ytf/yptf $f_{\mathbf{Y}}(\mathbf{y};\theta)$ riippuu tuntemattomasta parametrista $\theta = (\theta_1, .., \theta_d) \in \Theta \subseteq \mathbb{R}^d$.
- Mahdollisten parametriarvojen joukon Θ määrittely on osa mallin spesifiointia.
- Tavoitteena on tehdä päätelmiä tuntemattomasta parametrista θ aineiston $\mathbf{y}$ perusteella. Joissakin yhteyksissä on hyödyllistä osoittaa havaintomäärä vektoreissa $\mathbf{y}$ ja $\mathbf{Y}$. Tällöin merkitään $\mathbf{y}_n$ ja $\mathbf{Y}_n$.

$$Y_i = \phi Y_{i-1} + \varepsilon_i, \quad i = 1, \ldots, n, \quad Y_0 = y_0, \quad \varepsilon_1, \ldots, \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \quad \underline{\|}$$

- Havainto Y_i riippuu lineaarisesti edellisestä havainnosta Y_{i-1} ja ei-havaittavasta virhetermistä ε_i aivan kuten lineaarisessa (regressio)mallissa. Havaintojen riippuvuus on ilmeistä (ellei $\phi = 0$).

$$Y_i = \phi Y_{i-1} + \varepsilon_i, \quad i = 1, \ldots, n, \quad Y_0 = y_0, \quad \varepsilon_1, \ldots, \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \quad \perp\!\!\!\perp$$

- Havainto Y_i riippuu lineaarisesti edellisestä havainnosta Y_{i-1} ja ei-havaittavasta virhetermistä ε_i aivan kuten lineaarisessa (regressio)mallissa. Havaintojen riippuvuus on ilmeistä (ellei $\phi = 0$).
- Havainnot eivät myöskään ole samoin jakautuneita: $Y_1 = \phi y_0 + \varepsilon_1$ ja $Y_2 = \phi^2 y_0 + \phi \varepsilon_1 + \varepsilon_2 \Rightarrow \mathrm{Var}\left(Y_1\right) \neq \mathrm{Var}\left(Y_2\right)$.

$$Y_i = \phi Y_{i-1} + \varepsilon_i, \quad i = 1, ..., n, \quad Y_0 = y_0, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \quad \underline{\|}$$

- Havainto Y_i riippuu lineaarisesti edellisestä havainnosta Y_{i-1} ja ei-havaittavasta virhetermistä ε_i aivan kuten lineaarisessa (regressio)mallissa. Havaintojen riippuvuus on ilmeistä (ellei $\phi = 0$).
- Havainnot eivät myöskään ole samoin jakautuneita: $Y_1 = \phi y_0 + \varepsilon_1$ ja $Y_2 = \phi^2 y_0 + \phi\varepsilon_1 + \varepsilon_2 \Rightarrow \mathrm{Var}\left(Y_1\right) \neq \mathrm{Var}\left(Y_2\right)$.
- Usein vaaditaan $|\phi| < 1$, kun taas $\sigma^2 > 0$ luonteva rajoite.

Uskottavuuspäättelyä

Tilastollinen malli: Autoregressiivinen aikasarjamalli

$$Y_i = \phi Y_{i-1} + \varepsilon_i, \quad i = 1, \ldots, n, \quad Y_0 = y_0, \quad \varepsilon_1, \ldots, \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \quad \underline{\perp\!\!\!\perp}$$

- Havainto Y_i riippuu lineaarisesti edellisestä havainnosta Y_{i-1} ja ei-havaittavasta virhetermistä ε_i aivan kuten lineaarisessa (regressio)mallissa. Havaintojen riippuvuus on ilmeistä (ellei $\phi = 0$).
- Havainnot eivät myöskään ole samoin jakautuneita: $Y_1 = \phi y_0 + \varepsilon_1$ ja $Y_2 = \phi^2 y_0 + \phi\varepsilon_1 + \varepsilon_2 \Rightarrow \mathrm{Var}\left(Y_1\right) \neq \mathrm{Var}\left(Y_2\right)$.
- Usein vaaditaan $|\phi| < 1$, kun taas $\sigma^2 > 0$ luonteva rajoite.
- Kun $|\phi| \geq 1$ parametrien estimointi- ja testiteoria ns. epästandardi, sillä tavanomaiset SLL:t ja KRL:t eivät päde.

- Havainnot saatu N:stä riippumattomasta "yksilöstä" (esim. koehenkilö tai kotitalous) ja yksilölle j spesifioidaan (aluksi) lineaarinen regressiomalli

$$\begin{bmatrix} Y_{j,1} \\ \vdots \\ Y_{j,n_j} \end{bmatrix} = \alpha_j \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} + \gamma \begin{bmatrix} x_{j,1} \\ \vdots \\ x_{j,n_j} \end{bmatrix} + \begin{bmatrix} \varepsilon_{j,1} \\ \vdots \\ \varepsilon_{j,n_j} \end{bmatrix}$$

- Havainnot saatu N:stä riippumattomasta "yksilöstä" (esim. koehenkilö tai kotitalous) ja yksilölle j spesifioidaan (aluksi) lineaarinen regressiomalli

$$\begin{bmatrix} Y_{j,1} \\ \vdots \\ Y_{j,n_j} \end{bmatrix} = \alpha_j \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} + \gamma \begin{bmatrix} x_{j,1} \\ \vdots \\ x_{j,n_j} \end{bmatrix} + \begin{bmatrix} \varepsilon_{j,1} \\ \vdots \\ \varepsilon_{j,n_j} \end{bmatrix}$$

- tai

$$Y_j = \alpha_j 1_{n_j} + \gamma x_j + \varepsilon_j, \quad j = 1, ..., N.$$

- Havainnot saatu N:stä riippumattomasta "yksilöstä" (esim. koehenkilö tai kotitalous) ja yksilölle j spesifioidaan (aluksi) lineaarinen regressiomalli

$$\begin{bmatrix} Y_{j,1} \\ \vdots \\ Y_{j,n_j} \end{bmatrix} = \alpha_j \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} + \gamma \begin{bmatrix} x_{j,1} \\ \vdots \\ x_{j,n_j} \end{bmatrix} + \begin{bmatrix} \varepsilon_{j,1} \\ \vdots \\ \varepsilon_{j,n_j} \end{bmatrix}$$

- tai

$$Y_j = \alpha_j 1_{n_j} + \gamma x_j + \varepsilon_j, \quad j = 1, ..., N.$$

- Oletetaan

$$\varepsilon_1, ..., \varepsilon_N \perp\!\!\!\perp, \quad \varepsilon_j \sim \mathrm{N}\left(0, \sigma^2 I_{n_j}\right),$$

jolloin kaikki havainnot ovat riippumattomia.

- Malli

$$Y_j = Z_j\beta + \eta_j, \quad \eta_j = (\alpha_j - \alpha)\,1_{n_j} + \varepsilon_j, \quad j = 1, \ldots, N,$$

jossa

$$Z_j = \left[1_{n_j} : x_j\right] \quad (n_j \times 2) \quad \text{ja} \quad \beta = [\alpha \;\; \gamma]'$$

$$\eta_j \sim \mathrm{N}\left(0, \omega^2 J_{n_j} + \sigma^2 I_{n_j}\right) \;\; \perp\!\!\!\perp\,, \quad J_{n_j} = 1_{n_j}1'_{n_j} \;\; \text{tunnettu}$$

- Malli

$$Y_j = Z_j\beta + \eta_j, \quad \eta_j = (\alpha_j - \alpha)\,1_{n_j} + \varepsilon_j, \quad j = 1, \ldots, N,$$

jossa

$$Z_j = \left[1_{n_j} : x_j\right] \quad (n_j \times 2) \quad \text{ja} \quad \beta = [\alpha \ \ \gamma]'$$

$$\eta_j \;\sim\; \mathrm{N}\left(0, \omega^2 J_{n_j} + \sigma^2 I_{n_j}\right) \;\; \perp\!\!\!\perp, \quad J_{n_j} = 1_{n_j} 1'_{n_j} \;\; \text{tunnettu}$$

- Parametriavaruus määräytyy ehdoista $\beta \in \mathbb{R}^2$, $\sigma^2 > 0$ ja $\omega^2 \geq 0$.

- $$Y_j = Z_j\beta + \eta_j, \quad \eta_j = (\alpha_j - \alpha)\, 1_{n_j} + \varepsilon_j, \quad j = 1, \ldots, N,$$

- $$Y_j = Z_j\beta + \eta_j, \quad \eta_j = (\alpha_j - \alpha)\, 1_{n_j} + \varepsilon_j, \quad j = 1, \ldots, N,$$
- $$\eta_j \sim \mathrm{N}\left(0, V_j\left(\omega^2, \sigma^2\right)\right) \perp\!\!\!\perp, \quad V_j\left(\omega^2, \sigma^2\right) = \omega^2 J_{n_j} + \sigma^2 I_{n_j}$$

- $$Y_j = Z_j\beta + \eta_j, \quad \eta_j = (\alpha_j - \alpha)\,1_{n_j} + \varepsilon_j, \quad j = 1, \ldots, N,$$

- $$\eta_j \sim \mathrm{N}\left(0, V_j\left(\omega^2, \sigma^2\right)\right) \;\perp\!\!\!\perp, \quad V_j\left(\omega^2, \sigma^2\right) = \omega^2 J_{n_j} + \sigma^2 I_{n_j}$$

- Tilastollinen malli

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y};\theta) &= \prod_{j=1}^{N} (2\pi)^{-n_j/2} \det\left(V_j\left(\omega^2, \sigma^2\right)\right)^{-1/2} \\ &\qquad \times \exp\left\{-\tfrac{1}{2}\left(y_j - Z_j\beta\right)' V_j\left(\omega^2, \sigma^2\right)^{-1}\left(y_j - Z_j\beta\right)\right\} \end{aligned}$$

jossa $\theta = \left(\beta, \sigma^2, \omega^2\right) \in \mathbb{R}^2 \times (0, \infty) \times [0, \infty)$.

- $$Y_j = Z_j\beta + \eta_j, \quad \eta_j = (\alpha_j - \alpha)\, 1_{n_j} + \varepsilon_j, \quad j = 1, \ldots, N,$$

Uskottavuuspäättelyä

Tilastollinen malli: Lineaarisen mallin yleistyksiä

- $$Y_j = Z_j\beta + \eta_j, \quad \eta_j = (\alpha_j - \alpha)\, 1_{n_j} + \varepsilon_j, \quad j = 1, ..., N,$$
- $$\eta_j \sim \mathrm{N}\left(0, V_j\left(\omega^2, \sigma^2\right)\right) \perp\!\!\!\perp\, , \quad V_j\left(\omega^2, \sigma^2\right) = \omega^2 J_{n_j} + \sigma^2 I_{n_j}$$

- $$Y_j = Z_j\beta + \eta_j, \quad \eta_j = (\alpha_j - \alpha)\,1_{n_j} + \varepsilon_j, \quad j = 1, ..., N,$$
- $$\eta_j \sim \mathrm{N}\left(0, V_j\left(\omega^2, \sigma^2\right)\right) \perp\!\!\!\perp, \quad V_j\left(\omega^2, \sigma^2\right) = \omega^2 J_{n_j} + \sigma^2 I_{n_j}$$
- Lisätään virhetermiin kolmas komponentti eli määritellään

$$\eta_j = (\alpha_j - \alpha)\,1_{n_j} + \xi_j + \varepsilon_j,$$

jossa $\xi_j \perp\!\!\!\perp$ kahdesta muusta komponentista ja $\perp\!\!\!\perp$ "yli j:n arvojen".

- $$Y_j = Z_j\beta + \eta_j, \quad \eta_j = (\alpha_j - \alpha)\,1_{n_j} + \varepsilon_j, \quad j = 1, \ldots, N,$$
- $$\eta_j \sim \mathrm{N}\left(0, V_j\left(\omega^2, \sigma^2\right)\right) \;\perp\!\!\!\perp\;, \quad V_j\left(\omega^2, \sigma^2\right) = \omega^2 J_{n_j} + \sigma^2 I_{n_j}$$
- Lisätään virhetermiin kolmas komponentti eli määritellään

$$\eta_j \;=\; (\alpha_j - \alpha)\,1_{n_j} + \xi_j + \varepsilon_j,$$

jossa ξ_j $\perp\!\!\!\perp$ kahdesta muusta komponentista ja $\perp\!\!\!\perp$ "yli j:n arvojen".

- Esim.

$$\xi_{j,i} = \phi\xi_{j,i-1} + \epsilon_{j,i}, i = 1, \ldots, n_j, \quad \epsilon_{j,i} \sim \mathrm{N}\left(0, \lambda^2\right), \quad \lambda \geq 0, \quad \xi_{j,0} = 0$$

- $Y_1, ..., Y_N \perp\!\!\!\perp, \quad Y_j \sim \mathrm{N}\left(Z_j\beta, V_j\left(\psi\right)\right)$ ja

  $V_j\left(\psi\right)$ on parametrivektorin ψ (tunnettua muotoa oleva) funktio $\Rightarrow$

- $Y_1, ..., Y_N \perp\!\!\!\perp, \quad Y_j \sim \mathrm{N}\left(Z_j\beta, V_j\left(\psi\right)\right)$ ja

  $V_j\left(\psi\right)$ on parametrivektorin ψ (tunnettua muotoa oleva) funktio $\Rightarrow$

-
$$\begin{aligned} f_{\mathbf{Y}}\left(\mathbf{y};\theta\right) &= \prod_{j=1}^{N}\left(2\pi\right)^{-n_j/2}\det\left(V_j\left(\psi\right)\right)^{-1/2} \\ &\quad\times\exp\left\{-\tfrac{1}{2}\left(y_j - Z_j\beta\right)' V_j\left(\psi\right)^{-1}\left(y_j - Z_j\beta\right)\right\}, \end{aligned}$$

  jossa $\theta = \left(\beta,\psi\right)$ ja matriisi Z_j voi sisältää useita selittäviä muuttujia.

- Aineisto $\mathbf{y} = (y_1, ..., y_n)$, vast. sv $\mathbf{Y} = (Y_1, ..., Y_n)$ ja ytf/yptf $f_{\mathbf{Y}}(\mathbf{y}; \theta)$, $\quad \theta \in \Theta \subseteq \mathbb{R}^d$

- Aineisto $\mathbf{y} = (y_1, ..., y_n)$, vast. sv $\mathbf{Y} = (Y_1, ..., Y_n)$ ja ytf/yptf $f_{\mathbf{Y}}(\mathbf{y}; \theta)$, $\quad \theta \in \Theta \subseteq \mathbb{R}^d$
- Määritelmän mukaan uskottavuusfunktio $L(\theta; \mathbf{y})$ on

$$L(\theta; \mathbf{y}) = c(\mathbf{y})\, f_{\mathbf{Y}}(\mathbf{y}; \theta), \quad \theta \in \Theta \subseteq \mathbb{R}^d,$$

jossa $c(\mathbf{y}) > 0$ (myöhemmin usein $c(\mathbf{y}) = 1$).

- Aineisto $\mathbf{y} = (y_1, ..., y_n)$, vast. sv $\mathbf{Y} = (Y_1, ..., Y_n)$ ja ytf/yptf $f_{\mathbf{Y}}(\mathbf{y}; \theta)$, $\quad \theta \in \Theta \subseteq \mathbb{R}^d$

- Määritelmän mukaan uskottavuusfunktio $L(\theta; \mathbf{y})$ on

$$L(\theta; \mathbf{y}) = c(\mathbf{y}) f_{\mathbf{Y}}(\mathbf{y}; \theta), \quad \theta \in \Theta \subseteq \mathbb{R}^d,$$

  jossa $c(\mathbf{y}) > 0$ (myöhemmin usein $c(\mathbf{y}) = 1$).

- Monissa tarkasteluissa käytetään logaritmoitua uskottavuusfunktiota

$$l(\theta; \mathbf{y}) \quad = \quad \log L(\theta; \mathbf{y}).$$

- Aineisto $\mathbf{y} = (y_1, ..., y_n)$, vast. sv $\mathbf{Y} = (Y_1, ..., Y_n)$ ja ytf/yptf $f_{\mathbf{Y}}(\mathbf{y}; \theta)$, $\quad \theta \in \Theta \subseteq \mathbb{R}^d$

- Määritelmän mukaan uskottavuusfunktio $L(\theta; \mathbf{y})$ on

$$L(\theta; \mathbf{y}) = c(\mathbf{y}) f_{\mathbf{Y}}(\mathbf{y}; \theta), \quad \theta \in \Theta \subseteq \mathbb{R}^d,$$

  jossa $c(\mathbf{y}) > 0$ (myöhemmin usein $c(\mathbf{y}) = 1$).

- Monissa tarkasteluissa käytetään logaritmoitua uskottavuusfunktiota

$$l(\theta; \mathbf{y}) \quad = \quad \log L(\theta; \mathbf{y}).$$

- Kun aiheellista korostaa riippuvuutta havaintojen lukumäärästä, merkitään $L_n(\theta; \mathbf{y}_n)$ ja $l_n(\theta; \mathbf{y}_n)$.

- $\mathbf{Y} = (Y_1, ..., Y_n)$ ja $Y_1, ..., Y_n \perp\!\!\!\perp$ ja $Y_i \sim f_{Y_i}(y_i; \theta) \Rightarrow$

$$f_{\mathbf{Y}}(\mathbf{y}; \theta) = f_{Y_1}(y_1; \theta) \cdots f_{Y_n}(y_n; \theta)$$

- $\mathbf{Y} = (Y_1, ..., Y_n)$ ja $Y_1, ..., Y_n \perp\!\!\!\perp$ ja $Y_i \sim f_{Y_i}(y_i; \theta) \Rightarrow$

$$f_{\mathbf{Y}}(\mathbf{y}; \theta) = f_{Y_1}(y_1; \theta) \cdots f_{Y_n}(y_n; \theta)$$

- Merkitään $\mathbf{Y}_i = (Y_1, ..., Y_i)$ ja $\mathbf{y}_i = (y_1, ..., y_i)$, $i = 1, ..., n$, joten aineistoa vastaavan sv:n $\mathbf{Y}_n$ ytf/yptf on $f_{\mathbf{Y}_n}$.

- $\mathbf{Y} = (Y_1, ..., Y_n)$ ja $Y_1, ..., Y_n \perp\!\!\!\perp$ ja $Y_i \sim f_{Y_i}(y_i; \theta) \Rightarrow$

$$f_{\mathbf{Y}}(\mathbf{y}; \theta) = f_{Y_1}(y_1; \theta) \cdots f_{Y_n}(y_n; \theta)$$

- Merkitään $\mathbf{Y}_i = (Y_1, ..., Y_i)$ ja $\mathbf{y}_i = (y_1, ..., y_i)$, $i = 1, ..., n$, joten aineistoa vastaavan sv:n $\mathbf{Y}_n$ ytf/yptf on $f_{\mathbf{Y}_n}$.

- $\mathbf{Y}_i = (\mathbf{Y}_{i-1}, Y_i) \Rightarrow \cdots \Rightarrow$ käyttäen ehd. tf:n/ptf:n määritelmää

$$f_{\mathbf{Y}}(\mathbf{y}; \theta) = f_{Y_1}(y_1; \theta) \cdot \prod_{i=2}^{n} f_{Y_i|\mathbf{Y}_{i-1}}(y_i|\mathbf{y}_{i-1}; \theta)$$

- $\mathbf{Y} = (Y_1, ..., Y_n)$ ja $Y_1, ..., Y_n \perp\!\!\!\perp$ ja $Y_i \sim f_{Y_i}(y_i; \theta) \Rightarrow$

$$f_{\mathbf{Y}}(\mathbf{y}; \theta) = f_{Y_1}(y_1; \theta) \cdots f_{Y_n}(y_n; \theta)$$

- Merkitään $\mathbf{Y}_i = (Y_1, ..., Y_i)$ ja $\mathbf{y}_i = (y_1, ..., y_i)$, $i = 1, ..., n$, joten aineistoa vastaavan sv:n $\mathbf{Y}_n$ ytf/yptf on $f_{\mathbf{Y}_n}$.

- $\mathbf{Y}_i = (\mathbf{Y}_{i-1}, Y_i) \Rightarrow \cdots \Rightarrow$ käyttäen ehd. tf:n/ptf:n määritelmää

$$f_{\mathbf{Y}}(\mathbf{y}; \theta) = f_{Y_1}(y_1; \theta) \cdot \prod_{i=2}^{n} f_{Y_i|\mathbf{Y}_{i-1}}(y_i|\mathbf{y}_{i-1}; \theta)$$

- Yleistää yo $\perp\!\!\!\perp$ tapauksen tuloesityksen. Toimii periaatteessa aina.

- $f_{\mathbf{Y}}\left(\mathbf{y};\theta\right) = f_{Y_1}\left(y_1;\theta\right) \cdot \prod_{i=2}^{n} f_{Y_i|\mathbf{Y}_{i-1}}\left(y_i|\mathbf{y}_{i-1};\theta\right)$

- $f_{\mathbf{Y}}\left(\mathbf{y};\theta\right) = f_{Y_1}\left(y_1;\theta\right) \cdot \prod_{i=2}^{n} f_{Y_i|\mathbf{Y}_{i-1}}\left(y_i|\mathbf{y}_{i-1};\theta\right)$
- ja, kun $f_{i-1}\left(y_i;\theta\right) = f_{Y_i|\mathbf{Y}_{i-1}}\left(y_i|\mathbf{y}_{i-1};\theta\right)$,

- $f_{\mathbf{Y}}(\mathbf{y};\theta) = f_{Y_1}(y_1;\theta) \cdot \prod_{i=2}^{n} f_{Y_i|\mathbf{Y}_{i-1}}(y_i|\mathbf{y}_{i-1};\theta)$
- ja, kun $f_{i-1}(y_i;\theta) = f_{Y_i|\mathbf{Y}_{i-1}}(y_i|\mathbf{y}_{i-1};\theta)$,
- $f_{\mathbf{Y}}(\mathbf{y};\theta) = f_{Y_1}(y_1;\theta) \cdot \prod_{i=2}^{n} f_{i-1}(y_i;\theta)$,

- $f_{\mathbf{Y}}\left(\mathbf{y};\theta\right) = f_{Y_1}\left(y_1;\theta\right) \cdot \prod_{i=2}^{n} f_{Y_i|\mathbf{Y}_{i-1}}\left(y_i|\mathbf{y}_{i-1};\theta\right)$
- ja, kun $f_{i-1}\left(y_i;\theta\right) = f_{Y_i|\mathbf{Y}_{i-1}}\left(y_i|\mathbf{y}_{i-1};\theta\right)$,
- $f_{\mathbf{Y}}\left(\mathbf{y};\theta\right) = f_{Y_1}\left(y_1;\theta\right) \cdot \prod_{i=2}^{n} f_{i-1}\left(y_i;\theta\right)$,
    - jossa merkintöjen ja ehdollisen tf:n/ptf:n määritelmän mukaan

- $f_{\mathbf{Y}}(\mathbf{y};\theta) = f_{Y_1}(y_1;\theta) \cdot \prod_{i=2}^{n} f_{Y_i|\mathbf{Y}_{i-1}}(y_i|\mathbf{y}_{i-1};\theta)$
- ja, kun $f_{i-1}(y_i;\theta) = f_{Y_i|\mathbf{Y}_{i-1}}(y_i|\mathbf{y}_{i-1};\theta)$,
- $f_{\mathbf{Y}}(\mathbf{y};\theta) = f_{Y_1}(y_1;\theta) \cdot \prod_{i=2}^{n} f_{i-1}(y_i;\theta)$,
    - jossa merkintöjen ja ehdollisen tf:n/ptf:n määritelmän mukaan
    - $f_{i-1}(y_i;\theta) = \dfrac{f_{\mathbf{Y}_{i-1},Y_i}(\mathbf{y}_{i-1},y_i;\theta)}{f_{\mathbf{Y}_{i-1}}(\mathbf{y}_{i-1};\theta)}$ tai

- $f_{\mathbf{Y}}(\mathbf{y};\theta) = f_{Y_1}(y_1;\theta) \cdot \prod_{i=2}^{n} f_{Y_i|\mathbf{Y}_{i-1}}(y_i|\mathbf{y}_{i-1};\theta)$
- ja, kun $f_{i-1}(y_i;\theta) = f_{Y_i|\mathbf{Y}_{i-1}}(y_i|\mathbf{y}_{i-1};\theta)$,
- $f_{\mathbf{Y}}(\mathbf{y};\theta) = f_{Y_1}(y_1;\theta) \cdot \prod_{i=2}^{n} f_{i-1}(y_i;\theta)$,
  - jossa merkintöjen ja ehdollisen tf:n/ptf:n määritelmän mukaan
  - $f_{i-1}(y_i;\theta) = \dfrac{f_{\mathbf{Y}_{i-1},Y_i}(\mathbf{y}_{i-1},y_i;\theta)}{f_{\mathbf{Y}_{i-1}}(\mathbf{y}_{i-1};\theta)}$ tai
  - $f_{i-1}(y_i;\theta) = \dfrac{f_{Y_1,\ldots,Y_i}(y_1,\ldots,y_i;\theta)}{f_{Y_1,\ldots,Y_{i-1}}(y_1,\ldots,y_{i-1};\theta)}$

- Koska

$$
\begin{aligned}
f_{\mathbf{Y}}\left(\mathbf{y};\theta\right) &= f_{Y_1}\left(y_1;\theta\right)\cdot\prod_{i=2}^{n} f_{Y_i|\mathbf{Y}_{i-1}}\left(y_i|\mathbf{y}_{i-1};\theta\right) \\
&= f_{Y_1}\left(y_1;\theta\right)\cdot\prod_{i=2}^{n} f_{i-1}\left(y_i;\theta\right),
\end{aligned}
$$

- Koska

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y};\theta) &= f_{Y_1}(y_1;\theta)\cdot\prod_{i=2}^{n} f_{Y_i|\mathbf{Y}_{i-1}}(y_i|\mathbf{y}_{i-1};\theta) \\ &= f_{Y_1}(y_1;\theta)\cdot\prod_{i=2}^{n} f_{i-1}(y_i;\theta), \end{aligned}$$

- niin uskottavuusfunktiolle saadaan tuloesitys

$$L(\theta;\mathbf{y}) = c(\mathbf{y})\cdot f_{Y_1}(y_1;\theta)\cdot\prod_{i=2}^{n} f_{i-1}(y_i;\theta)$$

- Koska

$$
\begin{aligned}
f_{\mathbf{Y}}(\mathbf{y};\theta) &= f_{Y_1}(y_1;\theta)\cdot\prod_{i=2}^{n} f_{Y_i|\mathbf{Y}_{i-1}}(y_i|\mathbf{y}_{i-1};\theta) \\
&= f_{Y_1}(y_1;\theta)\cdot\prod_{i=2}^{n} f_{i-1}(y_i;\theta),
\end{aligned}
$$

- niin uskottavuusfunktiolle saadaan tuloesitys

$$
L(\theta;\mathbf{y}) = c(\mathbf{y})\cdot f_{Y_1}(y_1;\theta)\cdot\prod_{i=2}^{n} f_{i-1}(y_i;\theta)
$$

- ja sen logaritmille vastaavasti summaesitys

$$
l(\theta;\mathbf{y}) = \log c(\mathbf{y}) + \log f_{Y_1}(y_1;\theta) + \sum_{i=2}^{n} \log f_{i-1}(y_i;\theta).
$$

- Tarkastellaan autoregresssiivistä aikasarjamallia

$$Y_i = \phi Y_{i-1} + \varepsilon_i, \quad i = 1, ..., n, \quad Y_0 = y_0, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \quad \perp\!\!\!\perp$$

- Tarkastellaan autoregresssiivistä aikasarjamallia

$$Y_i = \phi Y_{i-1} + \varepsilon_i, \quad i = 1, \ldots, n, \quad Y_0 = y_0, \quad \varepsilon_1, \ldots, \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \quad \perp\!\!\!\perp$$

- Havaintoje ytf:ksi saadaan $\left(\theta = \left(\phi, \sigma^2\right)\right)$

$$f_{\mathbf{Y}}(\mathbf{y}; \theta) = \prod_{i=1}^{n} \left(2\pi\sigma^2\right)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2}\left(y_i - \phi y_{i-1}\right)^2\right\}$$

- Tarkastellaan autoregresssiivistä aikasarjamallia

$$Y_i = \phi Y_{i-1} + \varepsilon_i, \quad i = 1, \ldots, n, \quad Y_0 = y_0, \quad \varepsilon_1, \ldots, \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \quad \perp\!\!\!\perp$$

- Havaintoje ytf:ksi saadaan $\left(\theta = \left(\phi, \sigma^2\right)\right)$

$$f_{\mathbf{Y}}\left(\mathbf{y}; \theta\right) = \prod_{i=1}^{n} \left(2\pi\sigma^2\right)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2}\left(y_i - \phi y_{i-1}\right)^2\right\}$$

- tai

$$f_{\mathbf{Y}}\left(\mathbf{y}; \theta\right) = \left(2\pi\sigma^2\right)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - \phi y_{i-1}\right)^2\right\}$$

- *SU–estimaatti* $\hat{\theta} = \hat{\theta}\,(\mathbf{y})$ toteuttaa määritelmän mukaan ehdon

$$L(\hat{\theta}; \mathbf{y}) \geq L\,(\theta; \mathbf{y}) \quad \text{kaikilla } \theta \in \Theta \subseteq \mathbb{R}^d$$

- *SU–estimaatti* $\hat{\theta} = \hat{\theta}\,(\mathbf{y})$ toteuttaa määritelmän mukaan ehdon

$$L(\hat{\theta};\mathbf{y}) \geq L\,(\theta;\mathbf{y}) \quad \text{kaikilla } \theta \in \Theta \subseteq \mathbb{R}^d$$

- tai yhtäpitävästi log-uskottavuusfunktion avulla ilmaistuna

$$l(\hat{\theta};\mathbf{y}) \geq l\,(\theta;\mathbf{y}) \quad \text{kaikilla } \theta \in \Theta \subseteq \mathbb{R}^d.$$

- *SU–estimaatti* $\hat{\theta} = \hat{\theta}\left(\mathbf{y}\right)$ toteuttaa määritelmän mukaan ehdon

$$L(\hat{\theta}; \mathbf{y}) \geq L\left(\theta; \mathbf{y}\right) \quad \text{kaikilla } \theta \in \Theta \subseteq \mathbb{R}^d$$

- tai yhtäpitävästi log-uskottavuusfunktion avulla ilmaistuna

$$l(\hat{\theta}; \mathbf{y}) \geq l\left(\theta; \mathbf{y}\right) \quad \text{kaikilla } \theta \in \Theta \subseteq \mathbb{R}^d.$$

- Tulkittuna satunnaisena eli $\hat{\theta} = \hat{\theta}\left(\mathbf{Y}\right)$ on $\hat{\theta}$ *SU–estimaattori*.

- *SU–estimaatti* $\hat{\theta} = \hat{\theta}\,(\mathbf{y})$ toteuttaa määritelmän mukaan ehdon

$$L(\hat{\theta}; \mathbf{y}) \geq L\,(\theta; \mathbf{y}) \quad \text{kaikilla } \theta \in \Theta \subseteq \mathbb{R}^d$$

- tai yhtäpitävästi log-uskottavuusfunktion avulla ilmaistuna

$$l(\hat{\theta}; \mathbf{y}) \geq l\,(\theta; \mathbf{y}) \quad \text{kaikilla } \theta \in \Theta \subseteq \mathbb{R}^d.$$

- Tulkittuna satunnaisena eli $\hat{\theta} = \hat{\theta}\,(\mathbf{Y})$ on $\hat{\theta}$ *SU–estimaattori*.
- Jatkossa käytetään yleensä merkintää $\hat{\theta}$ tai $\hat{\theta}_n$.

- Edellä oletettiin, että uskottavuusfunktio todella saavuttaa maksimiarvonsa pisteessä $\hat{\theta}\left(\mathbf{y}\right)$ ja että $\hat{\theta}\left(\mathbf{Y}\right)$ on sm.

- Edellä oletettiin, että uskottavuusfunktio todella saavuttaa maksimiarvonsa pisteessä $\hat{\theta}\left(\mathbf{y}\right)$ ja että $\hat{\theta}\left(\mathbf{Y}\right)$ on sm.
- Matemaattisesti tämä ei ole itsestään selvää.

- Edellä oletettiin, että uskottavuusfunktio todella saavuttaa maksimiarvonsa pisteessä $\hat{\theta}(\mathbf{y})$ ja että $\hat{\theta}(\mathbf{Y})$ on sm.
- Matemaattisesti tämä ei ole itsestään selvää.
- Mainitaan tiedoksi, että (eräät) riittävät ehdot ovat, että

- Edellä oletettiin, että uskottavuusfunktio todella saavuttaa maksimiarvonsa pisteessä $\hat{\theta}(\mathbf{y})$ ja että $\hat{\theta}(\mathbf{Y})$ on sm.
- Matemaattisesti tämä ei ole itsestään selvää.
- Mainitaan tiedoksi, että (eräät) riittävät ehdot ovat, että
    - uskottavuusfunktio $\theta \mapsto L(\theta; \mathbf{y})$ on jatkuva kaikilla mahdollisilla aineistoilla $\mathbf{y}$ ja

- Edellä oletettiin, että uskottavuusfunktio todella saavuttaa maksimiarvonsa pisteessä $\hat{\theta}(\mathbf{y})$ ja että $\hat{\theta}(\mathbf{Y})$ on sm.
- Matemaattisesti tämä ei ole itsestään selvää.
- Mainitaan tiedoksi, että (eräät) riittävät ehdot ovat, että
  - uskottavuusfunktio $\theta \mapsto L(\theta; \mathbf{y})$ on jatkuva kaikilla mahdollisilla aineistoilla $\mathbf{y}$ ja
  - parametriavaruus Θ on kompakti (eli suljettu ja rajoitettu).

Uskottavuuspäättelyä

Uskottavuusfunktio ja SU–estimaattori: SU–estimaattori

- Edellä oletettiin, että uskottavuusfunktio todella saavuttaa maksimiarvonsa pisteessä $\hat{\theta}(\mathbf{y})$ ja että $\hat{\theta}(\mathbf{Y})$ on sm.
- Matemaattisesti tämä ei ole itsestään selvää.
- Mainitaan tiedoksi, että (eräät) riittävät ehdot ovat, että
  - uskottavuusfunktio $\theta \mapsto L(\theta; \mathbf{y})$ on jatkuva kaikilla mahdollisilla aineistoilla $\mathbf{y}$ ja
  - parametriavaruus Θ on kompakti (eli suljettu ja rajoitettu).
- Jatkossa SU–estimaattorin olemassaolo tullaan olettamaan, mikä ei kuitenkaan takaa yksikäsiteisyyttä.

- Edellä oletettiin, että uskottavuusfunktio todella saavuttaa maksimiarvonsa pisteessä $\hat{\theta}(\mathbf{y})$ ja että $\hat{\theta}(\mathbf{Y})$ on sm.
- Matemaattisesti tämä ei ole itsestään selvää.
- Mainitaan tiedoksi, että (eräät) riittävät ehdot ovat, että
  - uskottavuusfunktio $\theta \mapsto L(\theta; \mathbf{y})$ on jatkuva kaikilla mahdollisilla aineistoilla $\mathbf{y}$ ja
  - parametriavaruus Θ on kompakti (eli suljettu ja rajoitettu).
- Jatkossa SU–estimaattorin olemassaolo tullaan olettamaan, mikä ei kuitenkaan takaa yksikäsiteisyyttä.
- Useimmissa monimutkaisemmissa malleissa SU–estimaattia ei voida ratkaista suljetussa muodossa, vaan estimointi joudutaan suorittamaan numeerisia menetelmiä käyttäen.

Ehdollinen malli ja margnaalimalli

- Usein aineiston havainnot on luontevaa jakaa kahteen komponenttiin eli aineisto on $\mathbf{w} = (w_1, ..., w_n)$, jossa $w_i = (y_i, x_i)$, $i = 1, ..., n$.

Ehdollinen malli ja margnaalimalli

- Usein aineiston havainnot on luontevaa jakaa kahteen komponenttiin eli aineisto on $\mathbf{w} = (w_1, ..., w_n)$, jossa $w_i = (y_i, x_i)$, $i = 1, ..., n$.
- Jako perustuu taustatietoon tai -teoriaan ja ideana on, että y–muuttujat ovat regressioanalyysin tapaan selitettäviä muuttujia ja x–muuttujat niiden vaihtelua selittäviä ”syymuuttujia”.

Ehdollinen malli ja margnaalimalli

- Usein aineiston havainnot on luontevaa jakaa kahteen komponenttiin eli aineisto on $\mathbf{w} = (w_1, ..., w_n)$, jossa $w_i = (y_i, x_i)$, $i = 1, ..., n$.
- Jako perustuu taustatietoon tai -teoriaan ja ideana on, että y–muuttujat ovat regressioanalyysin tapaan selitettäviä muuttujia ja x–muuttujat niiden vaihtelua selittäviä "syymuuttujia".
- Kumpaankin muuttujaan ajatellaan liittyvän satunnaisvaihtelua (molempien havainnot on esimerkiksi saatu satunnaisotantaa käyttäen). Merkitään jälleen vastaavia satunnaismuuttujia suurilla kirjaimilla $\mathbf{W}$, W_i, Y_i ja X_i.

Ehdollinen malli ja margnaalimalli

- Usein aineiston havainnot on luontevaa jakaa kahteen komponenttiin eli aineisto on $\mathbf{w} = (w_1, \ldots, w_n)$, jossa $w_i = (y_i, x_i)$, $i = 1, \ldots, n$.
- Jako perustuu taustatietoon tai -teoriaan ja ideana on, että y–muuttujat ovat regressioanalyysin tapaan selitettäviä muuttujia ja x–muuttujat niiden vaihtelua selittäviä "syymuuttujia".
- Kumpaankin muuttujaan ajatellaan liittyvän satunnaisvaihtelua (molempien havainnot on esimerkiksi saatu satunnaisotantaa käyttäen). Merkitään jälleen vastaavia satunnaismuuttujia suurilla kirjaimilla $\mathbf{W}$, W_i, Y_i ja X_i.
- Tällaisessa tilanteessa pääasiallinen mielenkiinto kohdistuu usein y–muuttujien riippuvuuteen x–muuttujista, jolloin y–muuttujia tuntuu luontevalta tarkastella ehdollistamalla x–muuttujien suhteen.

Ehdollinen malli ja margnaalimalli

Multinormaalijakaumaan perustuva esimerkki

- Olkoon $W_1, ..., W_n \sim \mathrm{N}_k(\mu, \Sigma) \underline{\ \|\ }$. Ositetaan

$$W_i = \begin{bmatrix} Y_i \\ X_i \end{bmatrix} \begin{matrix} 1 \\ k-1 \end{matrix} \qquad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \qquad \text{ja} \qquad \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

- Olkoon $W_1, \ldots, W_n \sim \mathrm{N}_k\left(\mu, \Sigma\right) \perp\!\!\!\perp$. Ositetaan

$$W_i = \begin{bmatrix} Y_i \\ X_i \end{bmatrix} \begin{matrix} 1 \\ k-1 \end{matrix} \qquad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \quad \text{ja} \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

- W_i:n ytf on $f_{W_i} = f_{Y_i|X_i} \cdot f_{X_i}$.

Ehdollinen malli ja margnaalimalli

Multinormaalijakaumaan perustuva esimerkki

- Olkoon $W_1, ..., W_n \sim \mathrm{N}_k\left(\mu, \Sigma\right) \perp\!\!\!\perp$. Ositetaan

$$W_i = \begin{array}{cc} \begin{bmatrix} Y_i \\ X_i \end{bmatrix} & \begin{array}{c} 1 \\ k-1 \end{array} \end{array} \qquad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \quad \text{ja} \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

- W_i:n ytf on $f_{W_i} = f_{Y_i|X_i} \cdot f_{X_i}$.
- MN–jakauman ominaisuuksien nojalla,

Ehdollinen malli ja margnaalimalli

Multinormaalijakaumaan perustuva esimerkki

- Olkoon $W_1, ..., W_n \sim \mathrm{N}_k(\mu, \Sigma) \perp\!\!\!\perp$. Ositetaan

$$W_i = \begin{bmatrix} Y_i \\ X_i \end{bmatrix} \begin{matrix} 1 \\ k-1 \end{matrix} \qquad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \quad \text{ja} \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

- W_i:n ytf on $f_{W_i} = f_{Y_i|X_i} \cdot f_{X_i}$.
- MN–jakauman ominaisuuksien nojalla,
  - $X_i \sim \mathrm{N}_{k-1}(\mu_2, \Sigma_{22})$ ja

- Olkoon $W_1, \dots, W_n \sim \mathrm{N}_k(\mu, \Sigma) \perp\!\!\!\perp$. Ositetaan

$$W_i = \begin{bmatrix} Y_i \\ X_i \end{bmatrix} \begin{array}{c} 1 \\ k-1 \end{array} \qquad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \quad \text{ja} \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

- W_i:n ytf on $f_{W_i} = f_{Y_i|X_i} \cdot f_{X_i}$.
- MN–jakauman ominaisuuksien nojalla,
  - $X_i \sim \mathrm{N}_{k-1}(\mu_2, \Sigma_{22})$ ja
  - $Y_i | (X_i = x_i) \sim \mathrm{N}\left(\mu_1 - \sigma_{12}\Sigma_{22}^{-1}(x_i - \mu_2), \sigma^2\right)$,

    jossa $\sigma^2 = \sigma_1^2 - \sigma_{12}\Sigma_{22}^{-1}\sigma_{21}$.

- Olkoon $W_1, ..., W_n \sim \mathrm{N}_k(\mu, \Sigma) \underline{\perp\!\!\!\perp}$. Ositetaan

$$W_i = \begin{bmatrix} Y_i \\ X_i \end{bmatrix} \begin{matrix} 1 \\ k-1 \end{matrix} \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \quad \text{ja} \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

- W_i:n ytf on $f_{W_i} = f_{Y_i|X_i} \cdot f_{X_i}$.
- MN–jakauman ominaisuuksien nojalla,
  - $X_i \sim \mathrm{N}_{k-1}(\mu_2, \Sigma_{22})$ ja
  - $Y_i | (X_i = x_i) \sim \mathrm{N}\left(\mu_1 - \sigma_{12}\Sigma_{22}^{-1}(x_i - \mu_2), \sigma^2\right)$,

    jossa $\sigma^2 = \sigma_1^2 - \sigma_{12}\Sigma_{22}^{-1}\sigma_{21}$.
- Mielenkiinto kohdistuu ehdolliseen jakaumaan $Y_i | (X_i = x_i)$.

- Kiinnostavassa ehdollisessa jakaumassa

$$Y_i | \left(X_i = x_i\right) \sim \mathrm{N}\left(\mu_1 - \sigma_{12}\Sigma_{22}^{-1}\left(x_i - \mu_2\right), \sigma^2\right)$$

on varianssille otettu jo käyttöön uudelleen parametrointi

$$\sigma^2 = \sigma_1^2 - \sigma_{12}\Sigma_{22}^{-1}\sigma_{21}.$$

- Kiinnostavassa ehdollisessa jakaumassa

$$Y_i | (X_i = x_i) \sim \mathrm{N}\left(\mu_1 - \sigma_{12}\Sigma_{22}^{-1}(x_i - \mu_2), \sigma^2\right)$$

on varianssille otettu jo käyttöön uudelleen parametrointi

$$\sigma^2 = \sigma_1^2 - \sigma_{12}\Sigma_{22}^{-1}\sigma_{21}.$$

- Samoin on luontevaa käyttää parametrointia

$$\alpha = \mu_1 + \sigma_{12}\Sigma_{22}^{-1}\mu_2 \quad \text{ja} \quad \gamma' = -\sigma_{12}\Sigma_{22}^{-1},$$

Ehdollinen malli ja margnaalimalli

Multinormaalijakaumaan perustuva esimerkki

- Kiinnostavassa ehdollisessa jakaumassa

$$Y_i | \left(X_i = x_i\right) \sim \mathrm{N}\left(\mu_1 - \sigma_{12}\Sigma_{22}^{-1}\left(x_i - \mu_2\right), \sigma^2\right)$$

  on varianssille otettu jo käyttöön uudelleen parametrointi

$$\sigma^2 = \sigma_1^2 - \sigma_{12}\Sigma_{22}^{-1}\sigma_{21}.$$

- Samoin on luontevaa käyttää parametrointia

$$\alpha = \mu_1 + \sigma_{12}\Sigma_{22}^{-1}\mu_2 \quad \text{ja} \quad \gamma' = -\sigma_{12}\Sigma_{22}^{-1},$$

- jolloin

$$Y_i | \left(X_i = x_i\right) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right),$$

  jossa $z_i = [1 \;\; x_i']'$ ja $\beta = [\alpha \;\; \gamma']'$ ovat $k \times 1$ vektoreita.

Ehdollinen malli ja margnaalimalli

Multinormaalijakaumaan perustuva esimerkki

- Olkoon ehd. jakauman $Y_i | (X_i = x_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right)$ parametri $\psi = \left(\beta, \sigma^2\right)$ ja ehtomuuttujan $X_i \sim \mathrm{N}_{k-1}\left(\mu_2, \Sigma_{22}\right)$ parametri λ.

Ehdollinen malli ja margnaalimalli

Multinormaalijakaumaan perustuva esimerkki

- Olkoon ehd. jakauman $Y_i | (X_i = x_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right)$ parametri $\psi = \left(\beta, \sigma^2\right)$ ja ehtomuuttujan $X_i \sim \mathrm{N}_{k-1}(\mu_2, \Sigma_{22})$ parametri λ.
- Riippumattomuuden nojalla saadaan tilastolliseksi malliksi

$$f_{\mathbf{W}}(\mathbf{w}; \psi, \lambda) = \prod_{i=1}^{n} f_{Y_i|X_i}(y_i|x_i; \psi) \times f_{X_i}(x_i; \lambda)$$

- Olkoon ehd. jakauman $Y_i | (X_i = x_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right)$ parametri $\psi = \left(\beta, \sigma^2\right)$ ja ehtomuuttujan $X_i \sim \mathrm{N}_{k-1}\left(\mu_2, \Sigma_{22}\right)$ parametri λ.
- Riippumattomuuden nojalla saadaan tilastolliseksi malliksi

$$f_{\mathbf{W}}\left(\mathbf{w}; \psi, \lambda\right) = \prod_{i=1}^{n} f_{Y_i|X_i}\left(y_i | x_i; \psi\right) \times f_{X_i}\left(x_i; \lambda\right)$$

- eli

$$\begin{aligned} f_{\mathbf{W}}\left(\mathbf{w}; \psi, \lambda\right) &= \left(2\pi\sigma^2\right)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - z_i'\beta\right)^2\right\} \\ &\qquad \times \prod_{i=1}^{n} f_{X_i}\left(x_i; \lambda\right), \end{aligned}$$

jossa parametrit ψ ja λ vaihtelevat vapaasti ($\sigma^2 > 0$ ja $\Sigma_{22} > 0$).

Ehdollinen malli ja margnaalimalli

Multinormaalijakaumaan perustuva esimerkki

- $f_{\mathbf{W}}\left(\mathbf{w};\psi,\lambda\right) = \left(2\pi\sigma^2\right)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - z_i'\beta\right)^2\right\} \times \prod_{i=1}^{n} f_{X_i}\left(x_i;\lambda\right)$

Ehdollinen malli ja margnaalimalli

Multinormaalijakaumaan perustuva esimerkki

- $f_{\mathbf{W}}(\mathbf{w}; \psi, \lambda) =$
  $(2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - z_i'\beta)^2\right\} \times \prod_{i=1}^{n} f_{X_i}(x_i; \lambda)$
- Jälkimmäinen tekijä oikealla eli sv:n $\mathbf{X} = (X_1, ..., X_n)$ ytf on ns. *marginaalimalli*, josta ei olla kiinnostuneita.

Ehdollinen malli ja margnaalimalli

Multinormaalijakaumaan perustuva esimerkki

- $f_{\mathbf{W}}\left(\mathbf{w};\psi,\lambda\right) =$
  $\left(2\pi\sigma^2\right)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n\left(y_i - z_i'\beta\right)^2\right\}\times\prod_{i=1}^n f_{X_i}\left(x_i;\lambda\right)$
- Jälkimmäinen tekijä oikealla eli sv:n $\mathbf{X} = \left(X_1, ..., X_n\right)$ ytf on ns. *marginaalimalli*, josta ei olla kiinnostuneita.
- Edellinen, ehd. jakaumiin $Y_i|\left(X_i = x_i\right) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right)$ perustuva tekijä, on ns. *ehdollinen malli*.

- $f_{\mathbf{W}}\left(\mathbf{w};\psi,\lambda\right) = \left(2\pi\sigma^2\right)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - z_i'\beta\right)^2\right\} \times \prod_{i=1}^{n} f_{X_i}\left(x_i;\lambda\right)$
- Jälkimmäinen tekijä oikealla eli sv:n $\mathbf{X} = \left(X_1, \ldots, X_n\right)$ ytf on ns. *marginaalimalli*, josta ei olla kiinnostuneita.
- Edellinen, ehd. jakaumiin $Y_i|\left(X_i = x_i\right) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right)$ perustuva tekijä, on ns. *ehdollinen malli*.
- Kun mielenkiinto kohdistuu ehd. mallin parametriin $\psi = \left(\beta, \sigma^2\right)$, voidaan marginaalimalli sivuuttaa ψ:sta riippumattomana tekijänä.

Ehdollinen malli ja margnaalimalli

Multinormaalijakaumaan perustuva esimerkki

- $f_{\mathbf{W}}\left(\mathbf{w};\psi,\lambda\right) = \left(2\pi\sigma^2\right)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n\left(y_i - z_i'\beta\right)^2\right\}\times\prod_{i=1}^n f_{X_i}\left(x_i;\lambda\right)$
- Jälkimmäinen tekijä oikealla eli sv:n $\mathbf{X} = \left(X_1, ..., X_n\right)$ ytf on ns. *marginaalimalli*, josta ei olla kiinnostuneita.
- Edellinen, ehd. jakaumiin $Y_i|\left(X_i = x_i\right) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right)$ perustuva tekijä, on ns. *ehdollinen malli*.
- Kun mielenkiinto kohdistuu ehd. mallin parametriin $\psi = \left(\beta, \sigma^2\right)$, voidaan marginaalimalli sivuuttaa ψ:sta riippumattomana tekijänä.
- ψ:aa koskeva päättely voidaan tällöin perustaa pelkästään ehd. malliin.

Ehdollinen malli ja margnaalimalli

Multinormaalijakaumaan perustuva esimerkki

- $f_{\mathbf{W}}(\mathbf{w}; \psi, \lambda) =$
  $(2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - z_i'\beta)^2\right\} \times \prod_{i=1}^{n} f_{X_i}(x_i; \lambda)$
- Jälkimmäinen tekijä oikealla eli sv:n $\mathbf{X} = (X_1, \ldots, X_n)$ ytf on ns. *marginaalimalli*, josta ei olla kiinnostuneita.
- Edellinen, ehd. jakaumiin $Y_i | (X_i = x_i) \sim \mathrm{N}(z_i'\beta, \sigma^2)$ perustuva tekijä, on ns. *ehdollinen malli*.
- Kun mielenkiinto kohdistuu ehd. mallin parametriin $\psi = (\beta, \sigma^2)$, voidaan marginaalimalli sivuuttaa ψ:sta riippumattomana tekijänä.
- ψ:aa koskeva päättely voidaan tällöin perustaa pelkästään ehd. malliin.
- Lisäksi, sv:n $\mathbf{X}$ havainnot voidaan tulkita kiinteiksi ei–satunnaisiksi selittäviksi muuttujiksi kuten perinteisessä lineaarisessa mallissa.

- Havaintojen jako komponentteihin perustuu ei–tilastollisiin argumentteihin.

Ehdollinen malli ja margnaalimalli

Multinormaalijakaumaan perustuva esimerkki

- Havaintojen jako komponentteihin perustuu ei–tilastollisiin argumentteihin.
- Selitettävää y–muuttujaa ei ole relevanttia tulkita x–muuttujaa tai mitään sen komponenttia selittäväksi "syymuuttujaksi".

- Havaintojen jako komponentteihin perustuu ei–tilastollisiin argumentteihin.
- Selitettävää y–muuttujaa ei ole relevanttia tulkita x–muuttujaa tai mitään sen komponenttia selittäväksi "syymuuttujaksi".
- Jos y–muuttuja on x:n "syymuuttujaksi" eli "kausaalisuus" sekä suuntaan $x \longrightarrow y$ että suuntaan $y \longrightarrow x$ pätevät, johtaa edellä käytetty ehdollistaminen virheellisiin johtopäätöksiin.

Ehdollinen malli ja margnaalimalli

Multinormaalijakaumaan perustuva esimerkki

- Havaintojen jako komponentteihin perustuu ei–tilastollisiin argumentteihin.
- Selitettävää y–muuttujaa ei ole relevanttia tulkita x–muuttujaa tai mitään sen komponenttia selittäväksi "syymuuttujaksi".
- Jos y–muuttuja on x:n "syymuuttujaksi" eli "kausaalisuus" sekä suuntaan $x \rightarrow y$ että suuntaan $y \rightarrow x$ pätevät, johtaa edellä käytetty ehdollistaminen virheellisiin johtopäätöksiin.
- Tällaisessa tilanteessa y–muuttujaan tulisi sisällyttää ne x:n komponentit, jotka on aiheellista tulkita "selitettäviksi" muuttujiksi.

Ehdollinen malli ja margnaalimalli

Yleinen tapaus

- Aineisto $\mathbf{w} = (w_1, ..., w_n)$, vastaava sv $\mathbf{W} = (W_1, ..., W_n)$ ja $f_{\mathbf{W}}(\mathbf{w}; \theta)$ sen ytf/yptf, jossa θ tuntematon parametrivektori.

- Aineisto $\mathbf{w} = (w_1, ..., w_n)$, vastaava sv $\mathbf{W} = (W_1, ..., W_n)$ ja $f_{\mathbf{W}}(\mathbf{w}; \theta)$ sen ytf/yptf, jossa θ tuntematon parametrivektori.
- Kuten jaksossa 2.2,

$$f_{\mathbf{W}}(\mathbf{w}; \theta) = f_{W_1}(w_1; \theta) \prod_{i=2}^{n} f_{W_i|\mathbf{W}_{i-1}}(w_i|\mathbf{w}_{i-1}; \theta).$$

- Aineisto $\mathbf{w} = (w_1, \ldots, w_n)$, vastaava sv $\mathbf{W} = (W_1, \ldots, W_n)$ ja $f_{\mathbf{W}}(\mathbf{w}; \theta)$ sen ytf/yptf, jossa θ tuntematon parametrivektori.
- Kuten jaksossa 2.2,

$$f_{\mathbf{W}}(\mathbf{w}; \theta) = f_{W_1}(w_1; \theta) \prod_{i=2}^{n} f_{W_i|\mathbf{W}_{i-1}}(w_i | \mathbf{w}_{i-1}; \theta).$$

- Olkoon $W_i = (Y_i, X_i)$, jossa Y_i ja X_i kuten MN–esimerkissä.

Ehdollinen malli ja margnaalimalli

- Aineisto $\mathbf{w} = (w_1, ..., w_n)$, vastaava sv $\mathbf{W} = (W_1, ..., W_n)$ ja $f_{\mathbf{W}}(\mathbf{w};\theta)$ sen ytf/yptf, jossa θ tuntematon parametrivektori.
- Kuten jaksossa 2.2,

$$f_{\mathbf{W}}(\mathbf{w};\theta) = f_{W_1}(w_1;\theta) \prod_{i=2}^{n} f_{W_i|\mathbf{W}_{i-1}}(w_i|\mathbf{w}_{i-1};\theta).$$

- Olkoon $W_i = (Y_i, X_i)$, jossa Y_i ja X_i kuten MN–esimerkissä.
- Tällöin $f_{W_1} = f_{Y_1|X_1} f_{X_1}$ ja

$$\begin{aligned} f_{\mathbf{W}}(\mathbf{w};\theta) &= f_{Y_1|X_1}(y_1|x_1;\theta) \prod_{i=2}^{n} f_{Y_i|(\mathbf{W}_{i-1},X_i)}(y_i|\mathbf{w}_{i-1}, x_i;\theta) \\ &\quad \times f_{X_1}(x_1;\theta) \prod_{i=2}^{n} f_{X_i|\mathbf{W}_{i-1}}(x_i|\mathbf{w}_{i-1};\theta). \end{aligned}$$

Ehdollinen malli ja margnaalimalli

Yleinen tapaus

- Aineisto $\mathbf{w} = (w_1, \ldots, w_n)$, vastaava sv $\mathbf{W} = (W_1, \ldots, W_n)$ ja $f_{\mathbf{W}}(\mathbf{w}; \theta)$ sen ytf/yptf, jossa θ tuntematon parametrivektori.
- Kuten jaksossa 2.2,

$$f_{\mathbf{W}}(\mathbf{w}; \theta) = f_{W_1}(w_1; \theta) \prod_{i=2}^{n} f_{W_i|\mathbf{W}_{i-1}}(w_i|\mathbf{w}_{i-1}; \theta).$$

- Olkoon $W_i = (Y_i, X_i)$, jossa Y_i ja X_i kuten MN–esimerkissä.
- Tällöin $f_{W_1} = f_{Y_1|X_1} f_{X_1}$ ja

$$\begin{aligned} f_{\mathbf{W}}(\mathbf{w}; \theta) &= f_{Y_1|X_1}(y_1|x_1; \theta) \prod_{i=2}^{n} f_{Y_i|(\mathbf{W}_{i-1}, X_i)}(y_i|\mathbf{w}_{i-1}, x_i; \theta) \\ &\quad \times f_{X_1}(x_1; \theta) \prod_{i=2}^{n} f_{X_i|\mathbf{W}_{i-1}}(x_i|\mathbf{w}_{i-1}; \theta). \end{aligned}$$

- Oikealla *ehdollinen malli* $\times$ *marginaalimalli*.

- *Ehdollinen malli* $\times$ *Marginaalimalli*:

$$
\begin{aligned}
f_{\mathbf{W}}\left(\mathbf{w};\theta\right) &= f_{Y_1|X_1}\left(y_1|x_1;\theta\right)\prod_{i=2}^{n} f_{Y_i|(\mathbf{W}_{i-1},X_i)}\left(y_i|\mathbf{w}_{i-1},x_i;\theta\right) \\
&\quad \times f_{X_1}\left(x_1;\theta\right)\prod_{i=2}^{n} f_{X_i|\mathbf{W}_{i-1}}\left(x_i|\mathbf{w}_{i-1};\theta\right).
\end{aligned}
$$

- *Ehdollinen malli* × *Marginaalimalli*:

$$
\begin{aligned}
f_{\mathbf{W}}\left(\mathbf{w};\theta\right) &= f_{Y_1|X_1}\left(y_1|x_1;\theta\right)\prod_{i=2}^{n} f_{Y_i|(\mathbf{W}_{i-1},X_i)}\left(y_i|\mathbf{w}_{i-1},x_i;\theta\right) \\
&\quad \times f_{X_1}\left(x_1;\theta\right)\prod_{i=2}^{n} f_{X_i|\mathbf{W}_{i-1}}\left(x_i|\mathbf{w}_{i-1};\theta\right).
\end{aligned}
$$

- Parametrivektori θ esiintyy molemmissa $\Rightarrow$ tilastollista päättelyä ei voida yleisesti ottaen perustaa pelkästään ehdolliseen malliin menettämättä informaatiota.

Ehdollinen malli ja margnaalimalli

Yleinen tapaus

- *Ehdollinen malli* $\times$ *Marginaalimalli*:

$$
\begin{aligned}
f_{\mathbf{W}}\left(\mathbf{w};\theta\right) &= f_{Y_1|X_1}\left(y_1|x_1;\theta\right)\prod_{i=2}^{n} f_{Y_i|(\mathbf{W}_{i-1},X_i)}\left(y_i|\mathbf{w}_{i-1},x_i;\theta\right) \\
&\quad \times f_{X_1}\left(x_1;\theta\right)\prod_{i=2}^{n} f_{X_i|\mathbf{W}_{i-1}}\left(x_i|\mathbf{w}_{i-1};\theta\right).
\end{aligned}
$$

- Parametrivektori θ esiintyy molemmissa $\Rightarrow$ tilastollista päättelyä ei voida yleisesti ottaen perustaa pelkästään ehdolliseen malliin menettämättä informaatiota.
- Oletetaan MN–esimerkin mukainen tilanne: $\theta = (\psi, \lambda)$, jossa λ esiintyy vain marginaalimallissa ja ψ ehdollisessa mallissa.

- *Ehdollinen malli* $\times$ *Marginaalimalli*:

$$
\begin{aligned}
f_{\mathbf{W}}\left(\mathbf{w};\theta\right) &= f_{Y_1|X_1}\left(y_1|x_1;\theta\right)\prod_{i=2}^{n} f_{Y_i|(\mathbf{W}_{i-1},X_i)}\left(y_i|\mathbf{w}_{i-1},x_i;\theta\right) \\
&\quad \times f_{X_1}\left(x_1;\theta\right)\prod_{i=2}^{n} f_{X_i|\mathbf{W}_{i-1}}\left(x_i|\mathbf{w}_{i-1};\theta\right).
\end{aligned}
$$

- Parametrivektori θ esiintyy molemmissa $\Rightarrow$ tilastollista päättelyä ei voida yleisesti ottaen perustaa pelkästään ehdolliseen malliin menettämättä informaatiota.
- Oletetaan MN–esimerkin mukainen tilanne: $\theta = (\psi, \lambda)$, jossa λ esiintyy vain marginaalimallissa ja ψ ehdollisessa mallissa.
- ψ ei saa riippua λ:sta, mutta ψ:n määritelmän yhteydessä on voitu käyttää uudelleen parametrointia.

- *Ehdollinen malli* $\times$ *Marginaalimalli*:

$$f_{\mathbf{W}}(\mathbf{w};\theta) = f_{Y_1|X_1}(y_1|x_1;\theta)\prod_{i=2}^{n} f_{Y_i|(\mathbf{W}_{i-1},X_i)}(y_i|\mathbf{w}_{i-1},x_i;\theta) \times f_{X_1}(x_1;\theta)\prod_{i=2}^{n} f_{X_i|\mathbf{W}_{i-1}}(x_i|\mathbf{w}_{i-1};\theta).$$

- Parametrivektori θ esiintyy molemmissa $\Rightarrow$ tilastollista päättelyä ei voida yleisesti ottaen perustaa pelkästään ehdolliseen malliin menettämättä informaatiota.
- Oletetaan MN–esimerkin mukainen tilanne: $\theta = (\psi, \lambda)$, jossa λ esiintyy vain marginaalimallissa ja ψ ehdollisessa mallissa.
- ψ ei saa riippua λ:sta, mutta ψ:n määritelmän yhteydessä on voitu käyttää uudelleen parametrointia.
- Parametriavaruus on siten $\Psi \times \Lambda$

Ehdollinen malli ja margnaalimalli

Yleinen tapaus

- Kun ψ esiintyy vain ehd. mallissa ja ja λ vain marginaalimallissa, saadaan malliksi

$$f_{\mathbf{W}}\left(\mathbf{w};\psi,\lambda\right) = f_{Y_1|X_1}\left(y_1|x_1;\psi\right)\prod_{i=2}^{n} f_{Y_i|(\mathbf{W}_{i-1},X_i)}\left(y_i|\mathbf{w}_{i-1},x_i;\psi\right)$$
$$\times f_{X_1}\left(x_1;\lambda\right)\prod_{i=2}^{n} f_{X_i|\mathbf{W}_{i-1}}\left(x_i|\mathbf{w}_{i-1};\lambda\right),\ (\psi,\lambda)\in\Psi\times\Lambda.$$

Ehdollinen malli ja margnaalimalli

Yleinen tapaus

- Kun ψ esiintyy vain ehd. mallissa ja ja λ vain marginaalimallissa, saadaan malliksi

$$f_{\mathbf{W}}\left(\mathbf{w};\psi,\lambda\right) = f_{Y_1|X_1}\left(y_1|x_1;\psi\right)\prod_{i=2}^{n} f_{Y_i|(\mathbf{W}_{i-1},X_i)}\left(y_i|\mathbf{w}_{i-1},x_i;\psi\right)$$
$$\times f_{X_1}\left(x_1;\lambda\right)\prod_{i=2}^{n} f_{X_i|\mathbf{W}_{i-1}}\left(x_i|\mathbf{w}_{i-1};\lambda\right),\ (\psi,\lambda)\in\Psi\times\Lambda.$$

- Kun mielenkiinto kohdistuu vain ehd. mallin parametriin ψ, voidaan marginaalimalli tulkita parametrista riippumatomaksi vakioksi.

- Kun ψ esiintyy vain ehd. mallissa ja ja λ vain marginaalimallissa, saadaan malliksi

$$
\begin{aligned}
f_{\mathbf{W}}(\mathbf{w};\psi,\lambda) &= f_{Y_1|X_1}(y_1|x_1;\psi)\prod_{i=2}^{n} f_{Y_i|(\mathbf{W}_{i-1},X_i)}(y_i|\mathbf{w}_{i-1},x_i;\psi)\\
&\times f_{X_1}(x_1;\lambda)\prod_{i=2}^{n} f_{X_i|\mathbf{W}_{i-1}}(x_i|\mathbf{w}_{i-1};\lambda),\ (\psi,\lambda)\in\Psi\times\Lambda.
\end{aligned}
$$

- Kun mielenkiinto kohdistuu vain ehd. mallin parametriin ψ, voidaan marginaalimalli tulkita parametrista riippumatomaksi vakioksi.
- Tällöin päädytään ehdolliseen uskottavuusfunktioon (ilman vakiota)

$$
L^{(c)}(\psi;\mathbf{w}) = f_{Y_1|X_1}(y_1|x_1;\psi)\prod_{i=2}^{n} f_{Y_i|(\mathbf{W}_{i-1},X_i)}(y_i|\mathbf{w}_{i-1},x_i;\psi),
$$

johon ψ:aa koskeva päättely voidaan perustaa menettämättä infoa.

- *Ehdollinen malli* $\times$ *Marginaalimalli*:

$$f_{\mathbf{W}}\left(\mathbf{w};\psi,\lambda\right)=f_{Y_1|X_1}\left(y_1|x_1;\psi\right)\prod_{i=2}^{n}f_{Y_i|(\mathbf{W}_{i-1},X_i)}\left(y_i|\mathbf{w}_{i-1},x_i;\psi\right)$$

$$\times f_{X_1}\left(x_1;\lambda\right)\prod_{i=2}^{n}f_{X_i|\mathbf{W}_{i-1}}\left(x_i|\mathbf{w}_{i-1};\lambda\right),\ \left(\psi,\lambda\right)\in\Psi\times\Lambda.$$

- *Ehdollinen malli* $\times$ *Marginaalimalli*:

$$f_{\mathbf{W}}\left(\mathbf{w};\psi,\lambda\right)=f_{Y_1|X_1}\left(y_1|x_1;\psi\right)\prod_{i=2}^{n}f_{Y_i|(\mathbf{W}_{i-1},X_i)}\left(y_i|\mathbf{w}_{i-1},x_i;\psi\right)$$

$$\times f_{X_1}\left(x_1;\lambda\right)\prod_{i=2}^{n}f_{X_i|\mathbf{W}_{i-1}}\left(x_i|\mathbf{w}_{i-1};\lambda\right),\ \ (\psi,\lambda)\in\Psi\times\Lambda.$$

- Kuten MN–esimerkissä, x–muuttujat a priori y–muuttujien vaihtelua selittäviä "syymuuttujia", mutta ei päinvastoin.

- *Ehdollinen malli* $\times$ *Marginaalimalli*:

$$f_{\mathbf{W}}\left(\mathbf{w};\psi,\lambda\right)=f_{Y_1|X_1}\left(y_1|x_1;\psi\right)\prod_{i=2}^{n}f_{Y_i|(\mathbf{W}_{i-1},X_i)}\left(y_i|\mathbf{w}_{i-1},x_i;\psi\right)$$
$$\times f_{X_1}\left(x_1;\lambda\right)\prod_{i=2}^{n}f_{X_i|\mathbf{W}_{i-1}}\left(x_i|\mathbf{w}_{i-1};\lambda\right),\ \ (\psi,\lambda)\in\Psi\times\Lambda.$$

- Kuten MN–esimerkissä, x–muuttujat a priori y–muuttujien vaihtelua selittäviä "syymuuttujia", mutta ei päinvastoin.
- "Syymuuttujan" tulkinta ei yhtä selkeä kuin MN–esimerkissä, sillä $f_{X_i|\mathbf{W}_{i-1}}\left(x_i|\mathbf{w}_{i-1};\lambda\right)$ voi riippua muuttujista $y_1,\ldots,y_{i-1}$.

Ehdollinen malli ja margnaalimalli
Yleinen tapaus

- *Ehdollinen malli* $\times$ *Marginaalimalli*:

$$f_{\mathbf{W}}(\mathbf{w};\psi,\lambda) = f_{Y_1|X_1}(y_1|x_1;\psi)\prod_{i=2}^{n} f_{Y_i|(\mathbf{W}_{i-1},X_i)}(y_i|\mathbf{w}_{i-1},x_i;\psi)$$
$$\times f_{X_1}(x_1;\lambda)\prod_{i=2}^{n} f_{X_i|\mathbf{W}_{i-1}}(x_i|\mathbf{w}_{i-1};\lambda),\ (\psi,\lambda)\in\Psi\times\Lambda.$$

- Kuten MN–esimerkissä, x–muuttujat a priori y–muuttujien vaihtelua selittäviä "syymuuttujia", mutta ei päinvastoin.
- "Syymuuttujan" tulkinta ei yhtä selkeä kuin MN–esimerkissä, sillä $f_{X_i|\mathbf{W}_{i-1}}(x_i|\mathbf{w}_{i-1};\lambda)$ voi riippua muuttujista $y_1,\ldots,y_{i-1}$.
- Näin ei ole, jos $X_i|\mathbf{W}_{i-1} \stackrel{d}{=} X_i|\mathbf{X}_{i-1}$, $\mathbf{X}_{i-1} = (X_1,\ldots,X_{i-1})$ $(i \geq 2)$, jolloin sv:n $\mathbf{X}_n$ havainnot voidaan tulkita ehdollisessa mallissa kiinteiksi selittäviksi muuttujiksi ja marginaalimallin tarkempi täsmennys jätetään usein tekemättä.

- Seuraavassa ehdollista mallia sovelletaan lineaariseen malliin.

Ehdollinen malli ja margnaalimalli

Yleinen tapaus

- Seuraavassa ehdollista mallia sovelletaan lineaariseen malliin.
- Siinä kuten useissa muissakin erikoistapauksissa ehdolliset ytf:t $f_{Y_i|(\mathbf{w}_{i-1},x_i)}(y_i|\mathbf{w}_{i-1},x_i;\psi)$ (tai yptf:t) riippuvat vain kiinteästä määrästä vektorin $(\mathbf{w}_{i-1},x_i)=(y_1,x_1,...,y_{i-1},x_{i-1},x_i)$ komponentteja (esim. vain x_i:stä ja y_{i-1}:stä).

Ehdollinen malli ja margnaalimalli

Yleinen tapaus

- Seuraavassa ehdollista mallia sovelletaan lineaariseen malliin.
- Siinä kuten useissa muissakin erikoistapauksissa ehdolliset ytf:t $f_{Y_i|(\mathbf{w}_{i-1},x_i)}(y_i|\mathbf{w}_{i-1},x_i;\psi)$ (tai yptf:t) riippuvat vain kiinteästä määrästä vektorin $(\mathbf{w}_{i-1},x_i)=(y_1,x_1,...,y_{i-1},x_{i-1},x_i)$ komponentteja (esim. vain x_i:stä ja y_{i-1}:stä).
- Esitetyt periaatteet eivät rajoitu lineaariseen malliin, vaan ovat relevantteja myös muissa malleissa, joissa on "selitettäviä" ja "selittäviä" muuttujia.

Ehdollinen malli ja margnaalimalli

Yleinen tapaus

- Seuraavassa ehdollista mallia sovelletaan lineaariseen malliin.
- Siinä kuten useissa muissakin erikoistapauksissa ehdolliset ytf:t $f_{Y_i|(\mathbf{w}_{i-1},x_i)}(y_i|\mathbf{w}_{i-1},x_i;\psi)$ (tai yptf:t) riippuvat vain kiinteästä määrästä vektorin $(\mathbf{w}_{i-1},x_i) = (y_1,x_1,...,y_{i-1},x_{i-1},x_i)$ komponentteja (esim. vain x_i:stä ja y_{i-1}:stä).
- Esitetyt periaatteet eivät rajoitu lineaariseen malliin, vaan ovat relevantteja myös muissa malleissa, joissa on "selitettäviä" ja "selittäviä" muuttujia.
- Esimerkkejä muista malleista ovat epälineaarinen regressiomalli ja yleistetyn lineaarisen mallin erikoistapaukset (mukaan lukien tapaukset, jossa y–muuttuja on vektori).

- Oletetaan tilanne *ehdollinen malli* × *marginaalimalli*:

$$f_{\mathbf{W}}\left(\mathbf{w};\psi,\lambda\right)=f_{Y_1|X_1}\left(y_1|x_1;\psi\right)\prod_{i=2}^{n}f_{Y_i|(\mathbf{W}_{i-1},X_i)}\left(y_i|\mathbf{w}_{i-1},x_i;\psi\right)$$
$$\times f_{X_1}\left(x_1;\lambda\right)\prod_{i=2}^{n}f_{X_i|\mathbf{W}_{i-1}}\left(x_i|\mathbf{w}_{i-1};\lambda\right),\ \ (\psi,\lambda)\in\Psi\times\Lambda$$

- Oletetaan tilanne *ehdollinen malli* $\times$ *marginaalimalli*:

$$f_{\mathbf{W}}\left(\mathbf{w};\psi,\lambda\right)=f_{Y_1|X_1}\left(y_1|x_1;\psi\right)\prod_{i=2}^{n}f_{Y_i|(\mathbf{W}_{i-1},X_i)}\left(y_i|\mathbf{w}_{i-1},x_i;\psi\right)$$
$$\times f_{X_1}\left(x_1;\lambda\right)\prod_{i=2}^{n}f_{X_i|\mathbf{W}_{i-1}}\left(x_i|\mathbf{w}_{i-1};\lambda\right),\ \left(\psi,\lambda\right)\in\Psi\times\Lambda$$

- ja

$$Y_i|(\mathbf{W}_{i-1},X_i)\overset{d}{=}Y_i|Z_i\quad\text{ja}\quad Y_i|(Z_i=z_i)\sim\mathrm{N}\left(z_i'\beta,\sigma^2\right),$$

jossa Z_i $(p\times 1)$ sisältää $(\mathbf{W}_{i-1},X_i)$:n "relevantit" komponentit.

- Oletetaan tilanne *ehdollinen malli* $\times$ *marginaalimalli*:

$$f_{\mathbf{W}}\left(\mathbf{w};\psi,\lambda\right) = f_{Y_1|X_1}\left(y_1|x_1;\psi\right)\prod_{i=2}^{n} f_{Y_i|(\mathbf{W}_{i-1},X_i)}\left(y_i|\mathbf{w}_{i-1},x_i;\psi\right)$$

$$\times f_{X_1}\left(x_1;\lambda\right)\prod_{i=2}^{n} f_{X_i|\mathbf{W}_{i-1}}\left(x_i|\mathbf{w}_{i-1};\lambda\right),\ \ (\psi,\lambda)\in\Psi\times\Lambda$$

- ja

$$Y_i|(\mathbf{W}_{i-1},X_i)\overset{d}{=}Y_i|Z_i \quad \text{ja} \quad Y_i|(Z_i=z_i)\sim \mathrm{N}\left(z_i'\beta,\sigma^2\right),$$

jossa Z_i $(p\times 1)$ sisältää $(\mathbf{W}_{i-1},X_i)$:n "relevantit" komponentit.

- Kysymyksessä on siis normaalinen lineaarinen regressiomalli.

- Oletetaan tilanne *ehdollinen malli* $\times$ *marginaalimalli*:

$$f_{\mathbf{W}}\left(\mathbf{w};\psi,\lambda\right)=f_{Y_1|X_1}\left(y_1|x_1;\psi\right)\prod_{i=2}^{n}f_{Y_i|(\mathbf{W}_{i-1},X_i)}\left(y_i|\mathbf{w}_{i-1},x_i;\psi\right)$$

$$\times f_{X_1}\left(x_1;\lambda\right)\prod_{i=2}^{n}f_{X_i|\mathbf{W}_{i-1}}\left(x_i|\mathbf{w}_{i-1};\lambda\right),\ \ (\psi,\lambda)\in\Psi\times\Lambda$$

- ja

$$Y_i|(\mathbf{W}_{i-1},X_i)\overset{d}{=}Y_i|Z_i\quad\text{ja}\quad Y_i|(Z_i=z_i)\sim\mathrm{N}\left(z_i'\beta,\sigma^2\right),$$

  jossa Z_i $(p\times 1)$ sisältää $(\mathbf{W}_{i-1},X_i)$:n "relevantit" komponentit.
- Kysymyksessä on siis normaalinen lineaarinen regressiomalli.
- Konkreettisuuden vuoksi tarkastellaan kahta tapausta:

- Oletetaan tilanne *ehdollinen malli* $\times$ *marginaalimalli*:

$$f_{\mathbf{W}}(\mathbf{w};\psi,\lambda) = f_{Y_1|X_1}(y_1|x_1;\psi)\prod_{i=2}^{n} f_{Y_i|(\mathbf{W}_{i-1},X_i)}(y_i|\mathbf{w}_{i-1},x_i;\psi)$$

$$\times f_{X_1}(x_1;\lambda)\prod_{i=2}^{n} f_{X_i|\mathbf{W}_{i-1}}(x_i|\mathbf{w}_{i-1};\lambda),\ (\psi,\lambda)\in\Psi\times\Lambda$$

- ja

$$Y_i|(\mathbf{W}_{i-1},X_i) \overset{d}{=} Y_i|Z_i \quad \text{ja} \quad Y_i|(Z_i = z_i) \sim \mathrm{N}\left(z_i'\beta,\sigma^2\right),$$

  jossa Z_i $(p\times 1)$ sisältää $(\mathbf{W}_{i-1},X_i)$:n "relevantit" komponentit.
- Kysymyksessä on siis normaalinen lineaarinen regressiomalli.
- Konkreettisuuden vuoksi tarkastellaan kahta tapausta:
  - $Z_i = (1, X_i)$ eli "tavanomaiseen" regressiomallia ja

- Oletetaan tilanne *ehdollinen malli* $\times$ *marginaalimalli*:

$$f_{\mathbf{W}}\left(\mathbf{w};\psi,\lambda\right)=f_{Y_1|X_1}\left(y_1|x_1;\psi\right)\prod_{i=2}^{n}f_{Y_i|(\mathbf{W}_{i-1},X_i)}\left(y_i|\mathbf{w}_{i-1},x_i;\psi\right)$$
$$\times f_{X_1}\left(x_1;\lambda\right)\prod_{i=2}^{n}f_{X_i|\mathbf{W}_{i-1}}\left(x_i|\mathbf{w}_{i-1};\lambda\right),\ \ (\psi,\lambda)\in\Psi\times\Lambda$$

- ja

$$Y_i|(\mathbf{W}_{i-1},X_i)\overset{d}{=}Y_i|Z_i\quad\text{ja}\quad Y_i|(Z_i=z_i)\sim\mathrm{N}\left(z_i'\beta,\sigma^2\right),$$

jossa Z_i $(p\times 1)$ sisältää $(\mathbf{W}_{i-1},X_i)$:n "relevantit" komponentit.
- Kysymyksessä on siis normaalinen lineaarinen regressiomalli.
- Konkreettisuuden vuoksi tarkastellaan kahta tapausta:
  - $Z_i=(1,X_i)$ eli "tavanomaiseen" regressiomallia ja
  - $Z_i=(1,Y_{i-1},X_i)$ eli autoregressiivinen aikasarjamalli ulkopuolisin selittävin muuttujin ($Y_0=y_0$ tunnettu vakio).

- Koska

$$Y_i|(\mathbf{W}_{i-1}, X_i) \overset{d}{=} Y_i|Z_i \quad \text{ja} \quad Y_i|(Z_i = z_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right),$$

saadaan ehdolliseksi uskottavuusfunktioksi $\left(\beta \in \mathbb{R}^p,\ \sigma^2 > 0\right)$

$$L^{(c)}\left(\beta, \sigma^2; \mathbf{w}\right) = \prod_{i=1}^{n} \left(2\pi\sigma^2\right)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2}\left(y_i - z_i'\beta\right)^2\right\}.$$

- Koska

$$Y_i|(\mathbf{W}_{i-1}, X_i) \overset{d}{=} Y_i|Z_i \quad \text{ja} \quad Y_i|(Z_i = z_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right),$$

saadaan ehdolliseksi uskottavuusfunktioksi $\left(\beta \in \mathbb{R}^p,\ \sigma^2 > 0\right)$

$$L^{(c)}\left(\beta, \sigma^2; \mathbf{w}\right) = \prod_{i=1}^{n} \left(2\pi\sigma^2\right)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2}\left(y_i - z_i'\beta\right)^2\right\}.$$

- Nähdään, että ehdollinen malli voidaan esittää myös käyttäen yhtälöä

$$Y_i = Z_i'\beta + \varepsilon_i, \quad i = 1, ..., n,$$

jossa $\varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp$ ja $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$, $i = 1, ..., n$.

- Koska

$$Y_i|(\mathbf{W}_{i-1}, X_i) \overset{d}{=} Y_i|Z_i \quad \text{ja} \quad Y_i|(Z_i = z_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right),$$

saadaan ehdolliseksi uskottavuusfunktioksi $\left(\beta \in \mathbb{R}^p,\ \sigma^2 > 0\right)$

$$L^{(c)}\left(\beta, \sigma^2; \mathbf{w}\right) = \prod_{i=1}^{n} \left(2\pi\sigma^2\right)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2}\left(y_i - z_i'\beta\right)^2\right\}.$$

- Nähdään, että ehdollinen malli voidaan esittää myös käyttäen yhtälöä

$$Y_i = Z_i'\beta + \varepsilon_i, \quad i = 1, ..., n,$$

jossa $\varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp$ ja $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$, $i = 1, ..., n$.

- Tämä otetaan usein mallin lähtökohdaksi.

- Mallin määrittely voidaan siis perustaa yhtälöön

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.

Ehdollinen malli ja margnaalimalli

Sovellus lineaariseen malliin

- Mallin määrittely voidaan siis perustaa yhtälöön

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.

- Mallin johtaminen ehdollista- ja marginaalimallia käyttäen osoittaa kuitenkin, mitä satunnaisten selittäjien tapauksessa täytyy olettaa, jotta mallia voidaan käyttää ajatellulla tavalla.

- Mallin määrittely voidaan siis perustaa yhtälöön

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, \ldots, \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

  jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, \ldots, n$.
- Mallin johtaminen ehdollista- ja marginaalimallia käyttäen osoittaa kuitenkin, mitä satunnaisten selittäjien tapauksessa täytyy olettaa, jotta mallia voidaan käyttää ajatellulla tavalla.
- Esimerkiksi tapauksessa $Z_i = (1, X_i)$ oletetaan selittävät muuttujat X_i usein kiinteiksi perinteisen lineaarisen mallin tapaan.

- Mallin määrittely voidaan siis perustaa yhtälöön

$$Y_i = Z_i' \beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.

- Mallin johtaminen ehdollista- ja marginaalimallia käyttäen osoittaa kuitenkin, mitä satunnaisten selittäjien tapauksessa täytyy olettaa, jotta mallia voidaan käyttää ajatellulla tavalla.
- Esimerkiksi tapauksessa $Z_i = (1, X_i)$ oletetaan selittävät muuttujat X_i usein kiinteiksi perinteisen lineaarisen mallin tapaan.
    - Tämä on loogista, jos $X_i | \mathbf{W}_{i-1} \stackrel{d}{=} X_i | \mathbf{X}_{i-1}$,

Ehdollinen malli ja margnaalimalli

Sovellus lineaariseen malliin

- Mallin määrittely voidaan siis perustaa yhtälöön

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.

- Mallin johtaminen ehdollista- ja marginaalimallia käyttäen osoittaa kuitenkin, mitä satunnaisten selittäjien tapauksessa täytyy olettaa, jotta mallia voidaan käyttää ajatellulla tavalla.
- Esimerkiksi tapauksessa $Z_i = (1, X_i)$ oletetaan selittävät muuttujat X_i usein kiinteiksi perinteisen lineaarisen mallin tapaan.
    - Tämä on loogista, jos $X_i|\mathbf{W}_{i-1} \stackrel{d}{=} X_i|\mathbf{X}_{i-1}$,
    - mutta ei, jos X_i riippuu esimerkiksi satunnaiseksi tulkitusta Y_{i-1}:stä ja siten myös ε_{i-1}:stä. Tällöin tavanomainen lineaarisen mallin eksakti jakaumateoria ei päde.

- Ehdollinen uskottavuusfunktio on muodollisesti samanlainen kuin tavanomaisessa lineaarisessa mallissa eli

$$L^{(c)}\left(\beta, \sigma^2; \mathbf{w}\right) = \left(\sigma^2\right)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - z_i'\beta\right)^2\right\},$$

- Ehdollinen uskottavuusfunktio on muodollisesti samanlainen kuin tavanomaisessa lineaarisessa mallissa eli

$$L^{(c)}\left(\beta,\sigma^2;\mathbf{w}\right) = \left(\sigma^2\right)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - z_i'\beta\right)^2\right\},$$

- josta ehd. log-uskottavuusfunktio $l^{(c)}\left(\beta,\sigma^2;\mathbf{w}\right) = \log L^{(c)}\left(\beta,\sigma^2;\mathbf{w}\right)$:

$$l^{(c)}\left(\beta,\sigma^2;\mathbf{w}\right) = -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - z_i'\beta\right)^2,\ \beta\in\mathbb{R}^p,\ \sigma^2>0.$$

- Ehdollinen uskottavuusfunktio on muodollisesti samanlainen kuin tavanomaisessa lineaarisessa mallissa eli

$$L^{(c)}\left(\beta,\sigma^2;\mathbf{w}\right)=\left(\sigma^2\right)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i-z_i'\beta\right)^2\right\},$$

- josta ehd. log-uskottavuusfunktio $l^{(c)}\left(\beta,\sigma^2;\mathbf{w}\right)=\log L^{(c)}\left(\beta,\sigma^2;\mathbf{w}\right)$:

$$l^{(c)}\left(\beta,\sigma^2;\mathbf{w}\right)=-\frac{n}{2}\log\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i-z_i'\beta\right)^2,\ \beta\in\mathbb{R}^p,\ \sigma^2>0.$$

- Lineaaristen mallien kurssi $\Rightarrow$ SU–estimaattorit

$$\hat{\beta}=\left(\sum_{i=1}^{n}Z_iZ_i'\right)^{-1}\sum_{i=1}^{n}Z_iY_i\quad\text{ja}\quad\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(Y_i-Z_i'\hat{\beta})^2.$$

- Oletetaan, että uskottavuusfunktio $\theta \mapsto L(\theta; \mathbf{y})$ on kahdesti jatkuvasti derivoituva jokaisella $\mathbf{y}$. $[l(\theta; \mathbf{y}) = \log L(\theta; \mathbf{y})]$

- Oletetaan, että uskottavuusfunktio $\theta \mapsto L(\theta; \mathbf{y})$ on kahdesti jatkuvasti derivoituva jokaisella $\mathbf{y}$. $[l(\theta; \mathbf{y}) = \log L(\theta; \mathbf{y})]$
- *Pistemääräfunktio* tai *pistemäärä* on

$$s(\theta; \mathbf{y}) := \frac{\partial}{\partial \theta} l(\theta; \mathbf{y}) = \left(\frac{\partial}{\partial \theta_1} l(\theta; \mathbf{y}), \ldots, \frac{\partial}{\partial \theta_d} l(\theta; \mathbf{y}) \right)$$

- Oletetaan, että uskottavuusfunktio $\theta \mapsto L(\theta; \mathbf{y})$ on kahdesti jatkuvasti derivoituva jokaisella $\mathbf{y}$. $[l(\theta; \mathbf{y}) = \log L(\theta; \mathbf{y})]$
- *Pistemääräfunktio* tai *pistemäärä* on

$$s(\theta; \mathbf{y}) := \frac{\partial}{\partial \theta} l(\theta; \mathbf{y}) = \left(\frac{\partial}{\partial \theta_1} l(\theta; \mathbf{y}), \ldots, \frac{\partial}{\partial \theta_d} l(\theta; \mathbf{y}) \right)$$

- *Havaittu informaatio(matriisi)*

$$\mathcal{J}(\theta; \mathbf{y}) := -\frac{\partial^2}{\partial \theta \partial \theta'} l(\theta; \mathbf{y}) = \left[-\frac{\partial^2}{\partial \theta_a \partial \theta_b} l(\theta; \mathbf{y}) \right]_{a,b=1}^{d}$$

- Oletetaan, että uskottavuusfunktio $\theta \mapsto L(\theta; \mathbf{y})$ on kahdesti jatkuvasti derivoituva jokaisella $\mathbf{y}$. $[l(\theta; \mathbf{y}) = \log L(\theta; \mathbf{y})]$
- *Pistemääräfunktio* tai *pistemäärä* on

$$s(\theta; \mathbf{y}) := \frac{\partial}{\partial \theta} l(\theta; \mathbf{y}) = \left(\frac{\partial}{\partial \theta_1} l(\theta; \mathbf{y}), \dots, \frac{\partial}{\partial \theta_d} l(\theta; \mathbf{y}) \right)$$

- *Havaittu informaatio(matriisi)*

$$\mathcal{J}(\theta; \mathbf{y}) := -\frac{\partial^2}{\partial \theta \partial \theta'} l(\theta; \mathbf{y}) = \left[-\frac{\partial^2}{\partial \theta_a \partial \theta_b} l(\theta; \mathbf{y}) \right]_{a,b=1}^{d}$$

- *Fisherin informaatio(matriisi)*

$$\mathcal{I}(\theta) := \mathrm{E}\left[\mathcal{J}(\theta; \mathbf{Y})\right] = \mathrm{E}\left[-\frac{\partial^2}{\partial \theta \partial \theta'} l(\theta; \mathbf{Y}) \right],$$

jossa odotusarvon äärellisyys sisältyy määritelmään.

- Log-uskottavuusfunktiolle todettu summaesitys $(c(\mathbf{y}) = 1)$

$$l(\theta; \mathbf{y}) = \sum_{i=1}^{n} \log f_{i-1}(y_i; \theta), \quad f_0(y_1; \theta) = f_{Y_1}(y_1; \theta)$$

- Log-uskottavuusfunktiolle todettu summaesitys $(c(\mathbf{y}) = 1)$

$$l(\theta; \mathbf{y}) = \sum_{i=1}^{n} \log f_{i-1}(y_i; \theta), \quad f_0(y_1; \theta) = f_{Y_1}(y_1; \theta)$$

- Siis, pistemäärä $s(\theta; \mathbf{y}) = \partial l(\theta; \mathbf{y}) / \partial\theta$ voidaan kirjoittaa

$$s(\theta; \mathbf{y}) = \sum_{i=1}^{n} u_i(\theta; y_i), \quad u_i(\theta; y_i) = \frac{\partial}{\partial\theta} \log f_{i-1}(y_i; \theta).$$

- Log-uskottavuusfunktiolle todettu summaesitys $(c(\mathbf{y}) = 1)$

$$l(\theta; \mathbf{y}) = \sum_{i=1}^{n} \log f_{i-1}(y_i; \theta), \quad f_0(y_1; \theta) = f_{Y_1}(y_1; \theta)$$

- Siis, pistemäärä $s(\theta; \mathbf{y}) = \partial l(\theta; \mathbf{y}) / \partial\theta$ voidaan kirjoittaa

$$s(\theta; \mathbf{y}) = \sum_{i=1}^{n} u_i(\theta; y_i), \quad u_i(\theta; y_i) = \frac{\partial}{\partial\theta} \log f_{i-1}(y_i; \theta).$$

- Vastaavasti havaittu informaatio

$$\mathcal{J}(\theta; \mathbf{y}) = -\sum_{i=1}^{n} \frac{\partial^2}{\partial\theta\partial\theta'} \log f_{i-1}(y_i; \theta)$$

- Log-uskottavuusfunktiolle todettu summaesitys $(c(\mathbf{y}) = 1)$

$$l(\theta; \mathbf{y}) = \sum_{i=1}^{n} \log f_{i-1}(y_i; \theta), \quad f_0(y_1; \theta) = f_{Y_1}(y_1; \theta)$$

- Siis, pistemäärä $s(\theta; \mathbf{y}) = \partial l(\theta; \mathbf{y}) / \partial\theta$ voidaan kirjoittaa

$$s(\theta; \mathbf{y}) = \sum_{i=1}^{n} u_i(\theta; y_i), \quad u_i(\theta; y_i) = \frac{\partial}{\partial\theta} \log f_{i-1}(y_i; \theta).$$

- Vastaavasti havaittu informaatio

$$\mathcal{J}(\theta; \mathbf{y}) = -\sum_{i=1}^{n} \frac{\partial^2}{\partial\theta\partial\theta'} \log f_{i-1}(y_i; \theta)$$

- ja Fisherin informaatio

$$\mathcal{I}(\theta) = -\sum_{i=1}^{n} \mathrm{E}\left[\frac{\partial^2}{\partial\theta\partial\theta'} \log f_{i-1}(\mathbf{Y}_i; \theta)\right].$$

- Tilastollisen päättelyn kurssilla on todettu, että ns. säännöllisisissä malleissa pistemäärän $s\left(\theta; \mathbf{Y}\right)$ odotusarvo on nolla eli

$$\mathrm{E}\left[s\left(\theta; \mathbf{Y}\right)\right] = 0$$

- Tilastollisen päättelyn kurssilla on todettu, että ns. säännöllisissä malleissa pistemäärän $s(\theta; \mathbf{Y})$ odotusarvo on nolla eli

$$\mathrm{E}\left[s(\theta; \mathbf{Y})\right] = 0$$

- ja kovarianssimatriisi on Fisherin informaatiomatriisi $\mathcal{I}(\theta)$ eli

$$\mathrm{Cov}\left[s(\theta; \mathbf{Y})\right] = \mathrm{E}\left[s(\theta; \mathbf{Y})\, s(\theta; \mathbf{Y})'\right] = \mathcal{I}(\theta).$$

- Tilastollisen päättelyn kurssilla on todettu, että ns. säännöllisissä malleissa pistemäärän $s(\theta; \mathbf{Y})$ odotusarvo on nolla eli

$$\mathrm{E}\left[s(\theta; \mathbf{Y})\right] = 0$$

- ja kovarianssimatriisi on Fisherin informaatiomatriisi $\mathcal{I}(\theta)$ eli

$$\mathrm{Cov}\left[s(\theta; \mathbf{Y})\right] = \mathrm{E}\left[s(\theta; \mathbf{Y})\, s(\theta; \mathbf{Y})'\right] = \mathcal{I}(\theta).$$

- Lisäksi, Fisherin informaatiomatriisin avulla voidaan vastata kysymykseen kuinka tarkasti mallin parametria θ tai sen funktiota voidaan estimoida.

- *Informaatioepäyhtälö*: Parametrivektorin θ mille tahansa harhattomalle estimaattorille $T = t(\mathbf{Y})$ pätee

$$\mathrm{Var}\left(a'T\right) = a'\mathrm{Cov}\,(T)\,a \geq a'\mathcal{I}\,(\theta)^{-1}\,a \quad \text{kaikilla } a \in \mathbb{R}^d$$

tai

$$\mathrm{Cov}\,(T) - \mathcal{I}\,(\theta)^{-1} \geq 0 \quad \text{(posit. semidefiniitti).}$$

- *Informaatioepäyhtälö*: Parametrivektorin θ mille tahansa harhattomalle estimaattorille $T = t(\mathbf{Y})$ pätee

$$\mathrm{Var}\left(a'T\right) = a'\mathrm{Cov}\left(T\right)a \geq a'\mathcal{I}\left(\theta\right)^{-1}a \quad \text{kaikilla } a \in \mathbb{R}^d$$

tai

$$\mathrm{Cov}\left(T\right) - \mathcal{I}\left(\theta\right)^{-1} \geq 0 \quad \text{(posit. semidefiniitti).}$$

- Jos epäyhtälön paikalla on yhtälö, pätee $\mathrm{Cov}\left(T\right) = \mathcal{I}\left(\theta\right)^{-1}$ ja estimaattori T on täystehokas.

- *Informaatioepäyhtälö*: Parametrivektorin θ mille tahansa harhattomalle estimaattorille $T = t\left(\mathbf{Y}\right)$ pätee

$$\mathrm{Var}\left(a'T\right) = a'\mathrm{Cov}\left(T\right)a \geq a'\mathcal{I}\left(\theta\right)^{-1}a \quad \text{kaikilla } a \in \mathbb{R}^d$$

tai

$$\mathrm{Cov}\left(T\right) - \mathcal{I}\left(\theta\right)^{-1} \geq 0 \quad \text{(posit. semidefiniitti).}$$

- Jos epäyhtälön paikalla on yhtälö, pätee $\mathrm{Cov}\left(T\right) = \mathcal{I}\left(\theta\right)^{-1}$ ja estimaattori T on täystehokas.
- SU–estimaattori on todettu täystehokkaaksi "suurissa otoksissa" ja "riittävien säännöllisyysehtojen vallitessa", sillä tällöin

$$\hat{\theta} \underset{as}{\sim} \mathrm{N}\left(\theta, \mathcal{I}\left(\theta\right)^{-1}\right).$$

- Ositetaan mallin $f_{\mathbf{Y}}(\mathbf{y};\theta)$ parametrivektori kahteen (mahdollisesti vektoriarvoiseen) komponenttiin: $\theta = (\psi, \lambda)$.

- Ositetaan mallin $f_{\mathbf{Y}}(\mathbf{y};\theta)$ parametrivektori kahteen (mahdollisesti vektoriarvoiseen) komponenttiin: $\theta = (\psi, \lambda)$.
- Kuten tilastollisen päättelyn kurssilla, ovat parametrit ψ ja λ määritelmän mukaan *ortogonaalisia*, jos Fisherin informaatiomatriisi

$$\mathcal{I}(\theta) = \begin{bmatrix} \mathcal{I}_{\psi\psi}(\theta) & \mathcal{I}_{\psi\lambda}(\theta) \\ \mathcal{I}_{\lambda\psi}(\theta) & \mathcal{I}_{\lambda\lambda}(\theta) \end{bmatrix}$$

on lohkodiagonaalinen.

- Ositetaan mallin $f_{\mathbf{Y}}(\mathbf{y};\theta)$ parametrivektori kahteen (mahdollisesti vektoriarvoiseen) komponenttiin: $\theta = (\psi, \lambda)$.
- Kuten tilastollisen päättelyn kurssilla, ovat parametrit ψ ja λ määritelmän mukaan *ortogonaalisia*, jos Fisherin informaatiomatriisi

$$\mathcal{I}(\theta) = \begin{bmatrix} \mathcal{I}_{\psi\psi}(\theta) & \mathcal{I}_{\psi\lambda}(\theta) \\ \mathcal{I}_{\lambda\psi}(\theta) & \mathcal{I}_{\lambda\lambda}(\theta) \end{bmatrix}$$

  on lohkodiagonaalinen.
- Toisin sanoen, ortogonallisuus pätee, kun

$$\mathcal{I}_{\psi\lambda}(\theta) = \mathcal{I}_{\psi\lambda}(\theta)' = \mathsf{E}\left[-\frac{\partial^2}{\partial\psi\partial\lambda'} l(\theta;\mathbf{Y})\right] = 0.$$

- Parametrien ortogonalisuus on hyödyllinen ominaisuus, sillä sen voimassa ollessa parametreja ψ ja λ koskevat estimointi- ja testaustarkastelut ovat (suurissa otoksissa) likimain riippumattomia.

- Parametrien ortogonalisuus on hyödyllinen ominaisuus, sillä sen voimassa ollessa parametreja ψ ja λ koskevat estimointi- ja testaustarkastelut ovat (suurissa otoksissa) likimain riippumattomia.
- Intuitiivisesti: Parametria λ koskevat päätelmät ovat eivät vaikuta parametria ψ koskeviin päätelmiin (ja päinvastoin).

- Parametrien ortogonalisuus on hyödyllinen ominaisuus, sillä sen voimassa ollessa parametreja ψ ja λ koskevat estimointi- ja testaustarkastelut ovat (suurissa otoksissa) likimain riippumattomia.
- Intuitiivisesti: Parametria λ koskevat päätelmät ovat eivät vaikuta parametria ψ koskeviin päätelmiin (ja päinvastoin).
- Usein toinen parametreista kiinnostava ja toinen kiusa- tai haittaparametri.

- Parametrien ortogonalisuus on hyödyllinen ominaisuus, sillä sen voimassa ollessa parametreja ψ ja λ koskevat estimointi- ja testaustarkastelut ovat (suurissa otoksissa) likimain riippumattomia.
- Intuitiivisesti: Parametria λ koskevat päätelmät ovat eivät vaikuta parametria ψ koskeviin päätelmiin (ja päinvastoin).
- Usein toinen parametreista kiinnostava ja toinen kiusa- tai haittaparametri.
- Ortogonaalisuuden määritelmää käyttäen nähdään helposti, että ehdollisen mallin ja marginaalimallin parametrit ψ ja λ ovat ortogonalliset (perustelu jätetään tehtäväksi).

- Edellä tarkastellun lineaarisen mallin uskottavuusfunktio on

$$L^{(c)}\left(\beta,\sigma^2;\mathbf{w}\right)=\left(\sigma^2\right)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i-z_i'\beta\right)^2\right\}\quad\left(\beta\in\mathbb{R}^p,\ \sigma^2\ \right.$$

josta log-uskottavuusfunktio

$$l^{(c)}\left(\beta,\sigma^2;\mathbf{w}\right)=-\frac{n}{2}\log\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i-z_i'\beta\right)^2.$$

- Edellä tarkastellun lineaarisen mallin uskottavuusfunktio on

$$L^{(c)}\left(\beta, \sigma^2; \mathbf{w}\right) = \left(\sigma^2\right)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - z_i'\beta\right)^2\right\} \quad (\beta \in \mathbb{R}^p,\ \sigma^2 >$$

josta log-uskottavuusfunktio

$$l^{(c)}\left(\beta, \sigma^2; \mathbf{w}\right) = -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - z_i'\beta\right)^2.$$

- Derivointi:

$$\begin{aligned}
\frac{\partial}{\partial\beta} l^{(c)}\left(\beta, \sigma^2; \mathbf{w}\right) &= \frac{1}{\sigma^2}\sum_{i=1}^{n} z_i\left(y_i - z_i'\beta\right) \\
\frac{\partial}{\partial\sigma^2} l^{(c)}\left(\beta, \sigma^2; \mathbf{w}\right) &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}\left(y_i - z_i'\beta\right)^2
\end{aligned}$$

- Jatkuva malli $f_{\mathbf{Y}}(\mathbf{y};\theta)$, $\theta \in \Theta \subseteq \mathbb{R}^d$, on *säännöllinen*, jos $(\forall\ n \geq 1)$:

- Jatkuva malli $f_{\mathbf{Y}}(\mathbf{y};\theta)$, $\theta \in \Theta \subseteq \mathbb{R}^d$, on *säännöllinen*, jos ($\forall\ n \geq 1$):
- (a) jakauman alusta eli joukko $A = \{\mathbf{y} : f_Y(\mathbf{y};\theta) > 0\}$ ei riipu θ:sta,

- Jatkuva malli $f_{\mathbf{Y}}(\mathbf{y};\theta)$, $\theta \in \Theta \subseteq \mathbb{R}^d$, on *säännöllinen*, jos $(\forall\ n \geq 1)$:
- (a) jakauman alusta eli joukko $A = \{\mathbf{y} : f_Y(\mathbf{y};\theta) > 0\}$ ei riipu θ:sta,
- (b) funktiolla $\theta \mapsto f_{\mathbf{Y}}(\mathbf{y};\theta)$ on jatkuvat 2. os. derivaatat jokaisella $\mathbf{y}$,

- Jatkuva malli $f_{\mathbf{Y}}(\mathbf{y};\theta)$, $\theta \in \Theta \subseteq \mathbb{R}^d$, on *säännöllinen*, jos $(\forall\ n \geq 1)$:
- (a) jakauman alusta eli joukko $A = \{\mathbf{y} : f_Y(\mathbf{y};\theta) > 0\}$ ei riipu θ:sta,
- (b) funktiolla $\theta \mapsto f_{\mathbf{Y}}(\mathbf{y};\theta)$ on jatkuvat 2. os. derivaatat jokaisella $\mathbf{y}$,
- (c) aina kun $T = t(\mathbf{Y})$ on tunnusluku, jolle $\mathrm{E}_\theta(T)$ on olemassa kaikilla θ, pätee

- Jatkuva malli $f_{\mathbf{Y}}(\mathbf{y};\theta)$, $\theta \in \Theta \subseteq \mathbb{R}^d$, on *säännöllinen*, jos ($\forall\ n \geq 1$):
- (a) jakauman alusta eli joukko $A = \{\mathbf{y} : f_Y(\mathbf{y};\theta) > 0\}$ ei riipu θ:sta,
- (b) funktiolla $\theta \mapsto f_{\mathbf{Y}}(\mathbf{y};\theta)$ on jatkuvat 2. os. derivaatat jokaisella $\mathbf{y}$,
- (c) aina kun $T = t(\mathbf{Y})$ on tunnusluku, jolle $\mathrm{E}_\theta(T)$ on olemassa kaikilla θ, pätee
    - $$\frac{\partial}{\partial\theta_a}\int t(\mathbf{y}) f_{\mathbf{Y}}(\mathbf{y};\theta)\,d\mathbf{y} = \int t(\mathbf{y}) \frac{\partial}{\partial\theta_a} f_{\mathbf{Y}}(\mathbf{y};\theta)\,d\mathbf{y}, \quad a = 1, ..., d,$$

- Jatkuva malli $f_{\mathbf{Y}}(\mathbf{y};\theta)$, $\theta \in \Theta \subseteq \mathbb{R}^d$, on *säännöllinen*, jos ($\forall\ n \geq 1$):
- (a) jakauman alusta eli joukko $A = \{\mathbf{y} : f_Y(\mathbf{y};\theta) > 0\}$ ei riipu θ:sta,
- (b) funktiolla $\theta \mapsto f_{\mathbf{Y}}(\mathbf{y};\theta)$ on jatkuvat 2. os. derivaatat jokaisella $\mathbf{y}$,
- (c) aina kun $T = t(\mathbf{Y})$ on tunnusluku, jolle $\mathrm{E}_\theta(T)$ on olemassa kaikilla θ, pätee
  -
  $$\frac{\partial}{\partial\theta_a}\int t(\mathbf{y})\, f_{\mathbf{Y}}(\mathbf{y};\theta)\, d\mathbf{y} = \int t(\mathbf{y})\frac{\partial}{\partial\theta_a} f_{\mathbf{Y}}(\mathbf{y};\theta)\, d\mathbf{y}, \quad a = 1, \ldots, d,$$
  - ja

(d)
$$\frac{\partial^2}{\partial\theta_a\partial\theta_b}\int f_{\mathbf{Y}}(\mathbf{y};\theta)\, d\mathbf{y} = \int \frac{\partial^2}{\partial\theta_a\partial\theta_b} f_{\mathbf{Y}}(\mathbf{y};\theta)\, d\mathbf{y}, \quad a, b = 1, \ldots, d,$$

- Käytetään merkintöjä $s_n(\theta; \mathbf{y}_n) = \partial l_n(\theta; \mathbf{y}_n) / \partial\theta$ ja $\mathcal{I}_n(\theta)$.

- Käytetään merkintöjä $s_n\left(\theta;\mathbf{y}_n\right)=\partial l_n\left(\theta;\mathbf{y}_n\right)/\partial\theta$ ja $\mathcal{I}_n\left(\theta\right)$.
- **Lause 2.1.** Jos malli on säännöllinen, niin

- Käytetään merkintöjä $s_n(\theta; \mathbf{y}_n) = \partial l_n(\theta; \mathbf{y}_n) / \partial\theta$ ja $\mathcal{I}_n(\theta)$.
- **Lause 2.1.** Jos malli on säännöllinen, niin
  - (i) $\mathrm{E}_\theta [s_n(\theta; \mathbf{Y}_n)] = 0$ kaikilla $n \geq 1$ ja

- Käytetään merkintöjä $s_n(\theta; \mathbf{y}_n) = \partial l_n(\theta; \mathbf{y}_n) / \partial\theta$ ja $\mathcal{I}_n(\theta)$.
- **Lause 2.1.** Jos malli on säännöllinen, niin
  - (i) $\mathrm{E}_\theta[s_n(\theta; \mathbf{Y}_n)] = 0$ kaikilla $n \geq 1$ ja
  - (ii) $\mathrm{E}_\theta[s_n(\theta; \mathbf{Y}_n) | \mathbf{Y}_{n-1}] = s_{n-1}(\theta; \mathbf{Y}_{n-1})$ kaikilla $n \geq 2$, ja

- Käytetään merkintöjä $s_n(\theta; \mathbf{y}_n) = \partial l_n(\theta; \mathbf{y}_n) / \partial\theta$ ja $\mathcal{I}_n(\theta)$.
- **Lause 2.1.** Jos malli on säännöllinen, niin
  - (i) $\mathrm{E}_\theta[s_n(\theta; \mathbf{Y}_n)] = 0$ kaikilla $n \geq 1$ ja
  - (ii) $\mathrm{E}_\theta[s_n(\theta; \mathbf{Y}_n) | \mathbf{Y}_{n-1}] = s_{n-1}(\theta; \mathbf{Y}_{n-1})$ kaikilla $n \geq 2$, ja
  - (iii) $\mathrm{Cov}_\theta[s_n(\theta; \mathbf{Y}_n)] = \mathcal{I}_n(\theta)$.

- Käytetään merkintöjä $s_n(\theta; \mathbf{y}_n) = \partial l_n(\theta; \mathbf{y}_n) / \partial\theta$ ja $\mathcal{I}_n(\theta)$.
- **Lause 2.1.** Jos malli on säännöllinen, niin
  - (i) $\mathrm{E}_\theta[s_n(\theta; \mathbf{Y}_n)] = 0$ kaikilla $n \geq 1$ ja
  - (ii) $\mathrm{E}_\theta[s_n(\theta; \mathbf{Y}_n) | \mathbf{Y}_{n-1}] = s_{n-1}(\theta; \mathbf{Y}_{n-1})$ kaikilla $n \geq 2$, ja
  - (iii) $\mathrm{Cov}_\theta[s_n(\theta; \mathbf{Y}_n)] = \mathcal{I}_n(\theta)$.
- (ii) ja (iii) $\Rightarrow$ jono $s_n(\theta; \mathbf{Y}_n) = \sum_{i=1}^n u_i(\theta; Y_i)$ on martingaali ja $u_i(\theta; Y_i) = \partial \log f_{i-1}(Y_i; \theta) / \partial\theta$ on MD $(n = 1, 2, ...)$ informaatiojoukon $\mathbf{Y}_n = (Y_1, ..., Y_n)$ suhteen.

SU–estimaattorin asymptotiikka

Tarkentuvuus

- Tarkasteltaessa tarkentuvuutta yleisessä SU–kehikossa otettava käyttöön merkintä θ_0 "todelliselle" parametriarvolle.

- Tarkasteltaessa tarkentuvuutta yleisessä SU–kehikossa otettava käyttöön merkintä θ_0 "todelliselle" parametriarvolle.
- Ts., θ_0 on se parametriavaruuden Θ piste, joka vastaa aineiston tuottanutta todennäköisyysmekanismia.

- Tarkasteltaessa tarkentuvuutta yleisessä SU–kehikossa otettava käyttöön merkintä θ_0 "todelliselle" parametriarvolle.
- Ts., θ_0 on se parametriavaruuden Θ piste, joka vastaa aineiston tuottanutta todennäköisyysmekanismia.
- θ esiintyy mallissa $f_{\mathbf{Y}_n}(\mathbf{y}_n; \theta)$ "vapaana" muuttujana ja kuvaa mahdollisia vaihtoehtoisia parametriarvoja.

SU–estimaattorin asymptotiikka

- Tarkasteltaessa tarkentuvuutta yleisessä SU–kehikossa otettava käyttöön merkintä θ_0 "todelliselle" parametriarvolle.
- Ts., θ_0 on se parametriavaruuden Θ piste, joka vastaa aineiston tuottanutta todennäköisyysmekanismia.
- θ esiintyy mallissa $f_{\mathbf{Y}_n}(\mathbf{y}_n;\theta)$ "vapaana" muuttujana ja kuvaa mahdollisia vaihtoehtoisia parametriarvoja.
- Aineistoa vastaavan sv:n $\mathbf{Y}_n$ oletetaan noudattavan mallia $f_{\mathbf{Y}_n}(\mathbf{y}_n;\theta_0)$, johon kaikki todennäköisyyspäätelmät perustuvat.

- Tarkastellaan lineaarista mallia

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.

- Tarkastellaan lineaarista mallia

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.

- Parametrin β SU–estimaattori on PNS–estimaattori

$$\hat{\beta} = \left(\sum_{i=1}^{n} Z_i Z_i'\right)^{-1} \sum_{i=1}^{n} Z_i Y_i,$$

- Tarkastellaan lineaarista mallia

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.

- Parametrin β SU–estimaattori on PNS–estimaattori

$$\hat{\beta} = \left(\sum_{i=1}^{n} Z_i Z_i'\right)^{-1} \sum_{i=1}^{n} Z_i Y_i,$$

- joka voidaan kirjoittaa (tässä β on todellinen parametriarvo)

$$\hat{\beta} = \left(\sum_{i=1}^{n} Z_i Z_i'\right)^{-1} \sum_{i=1}^{n} Z_i Y_i = \beta + \left(\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i'\right)^{-1} \frac{1}{n}\sum_{i=1}^{n} Z_i \varepsilon_i.$$

- Oletetaan

- Oletetaan
  - $\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \xrightarrow{p} Q, \quad Q$ ei–satunnainen ja positiivisesti definiitti

- Oletetaan
  - $\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \xrightarrow{p} Q, \quad Q$ ei–satunnainen ja positiivisesti definiitti
  - $\frac{1}{n}\sum_{i=1}^{n} Z_i \varepsilon_i \xrightarrow{p} 0.$

- Oletetaan
  - $\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \xrightarrow{p} Q, \quad Q$ ei–satunnainen ja positiivisesti definiitti
  - $\frac{1}{n}\sum_{i=1}^{n} Z_i \varepsilon_i \xrightarrow{p} 0.$
- Nämä pätevät, kun $\{Z_i\}_{i=1}^{\infty}$ on esimerkiksi iid tai yleisemmin m–riippuva ja $\mathrm{E}(Z_{a,i}^4) \leq C < \infty$ kaikilla a ja i.

- Oletetaan
  - $\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \xrightarrow{p} Q, \quad Q$ ei–satunnainen ja positiivisesti definiitti
  - $\frac{1}{n}\sum_{i=1}^{n} Z_i \varepsilon_i \xrightarrow{p} 0.$
- Nämä pätevät, kun $\{Z_i\}_{i=1}^{\infty}$ on esimerkiksi iid tai yleisemmin m–riippuva ja $\mathrm{E}(Z_{a,i}^4) \leq C < \infty$ kaikilla a ja i.
- Tällöin,

$$\hat{\beta} \text{ on tarkentuva} \quad \text{eli} \quad \hat{\beta} \xrightarrow{p} \beta.$$

- Oletetaan
  - $\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \xrightarrow{p} Q, \quad Q$ ei–satunnainen ja positiivisesti definiitti
  - $\frac{1}{n}\sum_{i=1}^{n} Z_i \varepsilon_i \xrightarrow{p} 0.$
- Nämä pätevät, kun $\{Z_i\}_{i=1}^{\infty}$ on esimerkiksi iid tai yleisemmin m–riippuva ja $\mathrm{E}(Z_{a,i}^4) \leq C < \infty$ kaikilla a ja i.
- Tällöin,

$$\hat{\beta} \text{ on tarkentuva eli } \hat{\beta} \xrightarrow{p} \beta.$$

- Ei vaadi normaalisuusoletusta, joten riittää, että $\{\varepsilon_i\}_{i=1}^{\infty}$ iid–jono, jolla $\mathrm{E}(\varepsilon_i) = 0$ ja $\mathrm{E}(\varepsilon_i^2) = \sigma^2 < \infty$.

- Kaksi yleisluonteista riittävää ehtoa SUE:n $\hat{\theta}_n$ tarkentuvuudelle $\left(f_0\left(y_1;\theta\right) = f_{Y_1}\left(y_1;\theta\right)\right)$.

- Kaksi yleisluonteista riittävää ehtoa SUE:n $\hat{\theta}_n$ tarkentuvuudelle $\left(f_0\left(y_1;\theta\right) = f_{Y_1}\left(y_1;\theta\right)\right)$.

- **(i)** $n^{-1} l_n\left(\theta; \mathbf{Y}_n\right) = n^{-1} \sum_{i=1}^{n} \log f_{i-1}\left(Y_i;\theta\right) \xrightarrow{p} \bar{l}\left(\theta\right)$ tasaisesti Θ:ssa.

SU–estimaattorin asymptotiikka

Tarkentuvuus: Yleinen tapaus

- Kaksi yleisluonteista riittävää ehtoa SUE:n $\hat{\theta}_n$ tarkentuvuudelle $(f_0(y_1;\theta) = f_{Y_1}(y_1;\theta))$.
- **(i)** $n^{-1} l_n(\theta; \mathbf{Y}_n) = n^{-1} \sum_{i=1}^{n} \log f_{i-1}(Y_i;\theta) \xrightarrow{p} \bar{l}(\theta)$ tasaisesti Θ:ssa.
- Tulkitaan $\bar{l}(\theta) = \lim_{n\to\infty} \dfrac{1}{n} \sum_{i=1}^{n} \mathsf{E}\left[\log f_{i-1}(Y_i;\theta)\right].$

- Kaksi yleisluonteista riittävää ehtoa SUE:n $\hat{\theta}_n$ tarkentuvuudelle $\left(f_0\left(y_1;\theta\right) = f_{Y_1}\left(y_1;\theta\right)\right)$.
- **(i)** $n^{-1} l_n\left(\theta; \mathbf{Y}_n\right) = n^{-1} \sum_{i=1}^{n} \log f_{i-1}\left(Y_i;\theta\right) \xrightarrow{p} \bar{l}\left(\theta\right)$ tasaisesti Θ:ssa.
- Tulkitaan $\bar{l}\left(\theta\right) = \lim_{n\to\infty} \dfrac{1}{n} \sum_{i=1}^{n} \mathrm{E}\left[\log f_{i-1}\left(Y_i;\theta\right)\right]$.
    - Ts., "suurilla" havaintomäärillä ja todennäköisyydellä, joka on lähes yksi, $n^{-1} l_n\left(\theta; \mathbf{Y}_n\right)$ on "lähellä" rajafunktiota $\bar{l}\left(\theta\right)$ eikä "läheisyys" riipu siitä mikä parametriarvo on kysymyksessä.

- Kaksi yleisluonteista riittävää ehtoa SUE:n $\hat{\theta}_n$ tarkentuvuudelle $(f_0(y_1;\theta) = f_{Y_1}(y_1;\theta))$.
- **(i)** $n^{-1} l_n(\theta; \mathbf{Y}_n) = n^{-1} \sum_{i=1}^{n} \log f_{i-1}(Y_i;\theta) \xrightarrow{p} \bar{l}(\theta)$ tasaisesti Θ:ssa.
- Tulkitaan $\bar{l}(\theta) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}\left[\log f_{i-1}(Y_i;\theta)\right]$.
    - Ts., "suurilla" havaintomäärillä ja todennäköisyydellä, joka on lähes yksi, $n^{-1} l_n(\theta; \mathbf{Y}_n)$ on "lähellä" rajafunktiota $\bar{l}(\theta)$ eikä "läheisyys" riipu siitä mikä parametriarvo on kysymyksessä.
    - Tällöin intuitiivisesti $n^{-1} l_n(\theta; \mathbf{Y}_n)$:n maksimipiste $\hat{\theta}_n$ konvergoi stokastisesti kohti $\bar{l}(\theta)$:n yhtä ainoaa maksimipistettä, jos sellainen on.

- Kaksi yleisluonteista riittävää ehtoa SUE:n $\hat{\theta}_n$ tarkentuvuudelle $(f_0(y_1;\theta) = f_{Y_1}(y_1;\theta))$.

- Kaksi yleisluonteista riittävää ehtoa SUE:n $\hat{\theta}_n$ tarkentuvuudelle $(f_0(y_1;\theta) = f_{Y_1}(y_1;\theta))$.
- **(i)** $n^{-1} l_n(\theta;\mathbf{Y}_n) = n^{-1} \sum_{i=1}^{n} \log f_{i-1}(Y_i;\theta) \xrightarrow{p} \bar{l}(\theta)$ tasaisesti Θ:ssa.

- Kaksi yleisluonteista riittävää ehtoa SUE:n $\hat{\theta}_n$ tarkentuvuudelle $(f_0(y_1;\theta) = f_{Y_1}(y_1;\theta))$.
- **(i)** $n^{-1} l_n(\theta; \mathbf{Y}_n) = n^{-1} \sum_{i=1}^{n} \log f_{i-1}(Y_i;\theta) \xrightarrow{p} \bar{l}(\theta)$ tasaisesti Θ:ssa.
- Tällöin intuitiivisesti $n^{-1} l_n(\theta; \mathbf{Y}_n)$:n maksimipiste $\hat{\theta}_n$ konvergoi stokastisesti kohti $\bar{l}(\theta)$:n yhtä ainoaa maksimipistettä, jos sellainen on.

- Kaksi yleisluonteista riittävää ehtoa SUE:n $\hat{\theta}_n$ tarkentuvuudelle $(f_0(y_1;\theta) = f_{Y_1}(y_1;\theta))$.
- **(i)** $n^{-1} l_n(\theta; \mathbf{Y}_n) = n^{-1} \sum_{i=1}^{n} \log f_{i-1}(Y_i;\theta) \xrightarrow{p} \bar{l}(\theta)$ tasaisesti Θ:ssa.
- Tällöin intuitiivisesti $n^{-1} l_n(\theta; \mathbf{Y}_n)$:n maksimipiste $\hat{\theta}_n$ konvergoi stokastisesti kohti $\bar{l}(\theta)$:n yhtä ainoaa maksimipistettä, jos sellainen on.
- **(ii)** Asetetaan ehto, joka takaa, että rajafunktiolla on yksikäsitteinen maksimi todellisessa parametriarvossa θ_0.

- Kaksi yleisluonteista riittävää ehtoa SUE:n $\hat{\theta}_n$ tarkentuvuudelle $(f_0(y_1;\theta) = f_{Y_1}(y_1;\theta))$.
- **(i)** $n^{-1} l_n(\theta; \mathbf{Y}_n) = n^{-1} \sum_{i=1}^{n} \log f_{i-1}(Y_i;\theta) \xrightarrow{p} \bar{l}(\theta)$ tasaisesti Θ:ssa.
- Tällöin intuitiivisesti $n^{-1} l_n(\theta; \mathbf{Y}_n)$:n maksimipiste $\hat{\theta}_n$ konvergoi stokastisesti kohti $\bar{l}(\theta)$:n yhtä ainoaa maksimipistettä, jos sellainen on.
- **(ii)** Asetetaan ehto, joka takaa, että rajafunktiolla on yksikäsitteinen maksimi todellisessa parametriarvossa θ_0.
- Tällöin tarkentuvuus seuraa.

Lause 2.2. Olkoon $\bar{l} : \Theta \to \mathbb{R}$ ei-satunnainen funktio. Oletetaan, että

- (i) $\dfrac{1}{n} l_n(\theta; \mathbf{Y}_n) \xrightarrow{p} \bar{l}(\theta)$ tasaisesti joukossa Θ ja

Lause 2.2. Olkoon $\bar{l} : \Theta \to \mathbb{R}$ ei-satunnainen funktio. Oletetaan, että

- (i) $\frac{1}{n} l_n(\theta; \mathbf{Y}_n) \xrightarrow{p} \bar{l}(\theta)$ tasaisesti joukossa Θ ja
- (ii) jokaisella $\varepsilon > 0$, $\sup_{\|\theta-\theta_0\|\geq\varepsilon} \bar{l}(\theta) < \bar{l}(\theta_0)$.

Lause 2.2. Olkoon $\bar{l} : \Theta \to \mathbb{R}$ ei-satunnainen funktio. Oletetaan, että

- (i) $\frac{1}{n} l_n(\theta; \mathbf{Y}_n) \xrightarrow{p} \bar{l}(\theta)$ tasaisesti joukossa Θ ja
- (ii) jokaisella $\varepsilon > 0$, $\sup_{\|\theta-\theta_0\| \geq \varepsilon} \bar{l}(\theta) < \bar{l}(\theta_0)$.
- Tällöin SU–estimaattori on tarkentuva eli $\hat{\theta}_n \xrightarrow{p} \theta_0$.

Lause 2.2. Olkoon $\bar{l} : \Theta \to \mathbb{R}$ ei-satunnainen funktio. Oletetaan, että

- (i) $\frac{1}{n} l_n(\theta; \mathbf{Y}_n) \xrightarrow{p} \bar{l}(\theta)$ tasaisesti joukossa Θ ja
- (ii) jokaisella $\varepsilon > 0$, $\sup_{\|\theta - \theta_0\| \geq \varepsilon} \bar{l}(\theta) < \bar{l}(\theta_0)$.
- Tällöin SU–estimaattori on tarkentuva eli $\hat{\theta}_n \xrightarrow{p} \theta_0$.
    - Huom.: Mallin säännöllisyyttä ei vaadita.

- Oletusten yleisluonteisuuden vuoksi tarkentuvuustuloksen merkitys on pitkälti periaatteellinen.

- Oletusten yleisluonteisuuden vuoksi tarkentuvuustuloksen merkitys on pitkälti periaatteellinen.
- **(i)** $n^{-1} l_n(\theta; \mathbf{Y}_n) = n^{-1} \sum_{i=1}^{n} \log f_{i-1}(Y_i; \theta) \xrightarrow{p} \bar{l}(\theta)$ tasaisesti Θ:ssa.

- Oletusten yleisluonteisuuden vuoksi tarkentuvuustuloksen merkitys on pitkälti periaatteellinen.
- **(i)** $n^{-1} l_n(\theta; \mathbf{Y}_n) = n^{-1} \sum_{i=1}^{n} \log f_{i-1}(Y_i; \theta) \xrightarrow{p} \bar{l}(\theta)$ tasaisesti Θ:ssa.
    - Todetaan konkreettisissa tilanteissa soveltamalla jotain useista tasaisista SLL:sta – ei välttämättä yksinkertaista.

SU–estimaattorin asymptotiikka

Tarkentuvuus: Yleinen tapaus

- Oletusten yleisluonteisuuden vuoksi tarkentuvuustuloksen merkitys on pitkälti periaatteellinen.
- **(i)** $n^{-1} l_n(\theta; \mathbf{Y}_n) = n^{-1} \sum_{i=1}^{n} \log f_{i-1}(Y_i; \theta) \xrightarrow{p} \bar{l}(\theta)$ tasaisesti Θ:ssa.
  - Todetaan konkreettisissa tilanteissa soveltamalla jotain useista tasaisista SLL:sta – ei välttämättä yksinkertaista.
  - Toisaalta tasaisia SLL:ja on esitetty usein erilaisin oletuksin sallien havaintojen riippuvuus ja jakaumien heterogeenisuus.

- Oletusten yleisluonteisuuden vuoksi tarkentuvuustuloksen merkitys on pitkälti periaatteellinen.
- **(i)** $n^{-1} l_n(\theta; \mathbf{Y}_n) = n^{-1} \sum_{i=1}^{n} \log f_{i-1}(Y_i; \theta) \xrightarrow{p} \bar{l}(\theta)$ tasaisesti Θ:ssa.
  - Todetaan konkreettisissa tilanteissa soveltamalla jotain useista tasaisista SLL:sta – ei välttämättä yksinkertaista.
  - Toisaalta tasaisia SLL:ja on esitetty usein erilaisin oletuksin sallien havaintojen riippuvuus ja jakaumien heterogeenisuus.
- **(ii)** jokaisella $\varepsilon > 0$, $\sup_{\|\theta-\theta_0\|\geq\varepsilon} \bar{l}(\theta) < \bar{l}(\theta_0)$.

- Oletusten yleisluonteisuuden vuoksi tarkentuvuustuloksen merkitys on pitkälti periaatteellinen.
- **(i)** $n^{-1} l_n(\theta; \mathbf{Y}_n) = n^{-1} \sum_{i=1}^{n} \log f_{i-1}(Y_i; \theta) \xrightarrow{p} \bar{l}(\theta)$ tasaisesti Θ:ssa.
  - Todetaan konkreettisissa tilanteissa soveltamalla jotain useista tasaisista SLL:sta – ei välttämättä yksinkertaista.
  - Toisaalta tasaisia SLL:ja on esitetty usein erilaisin oletuksin sallien havaintojen riippuvuus ja jakaumien heterogeenisuus.
- **(ii)** jokaisella $\varepsilon > 0$, $\sup_{\|\theta-\theta_0\| \geq \varepsilon} \bar{l}(\theta) < \bar{l}(\theta_0)$.
  - Takaa, että funktiolla $\bar{l}$ on yksikäsitteinen maksimi pisteessä θ_0.

- Oletusten yleisluonteisuuden vuoksi tarkentuvuustuloksen merkitys on pitkälti periaatteellinen.
- **(i)** $n^{-1} l_n(\theta; \mathbf{Y}_n) = n^{-1} \sum_{i=1}^{n} \log f_{i-1}(Y_i; \theta) \xrightarrow{p} \bar{l}(\theta)$ tasaisesti Θ:ssa.
  - Todetaan konkreettisissa tilanteissa soveltamalla jotain useista tasaisista SLL:sta – ei välttämättä yksinkertaista.
  - Toisaalta tasaisia SLL:ja on esitetty usein erilaisin oletuksin sallien havaintojen riippuvuus ja jakaumien heterogeenisuus.
- **(ii)** jokaisella $\varepsilon > 0$, $\sup_{\|\theta - \theta_0\| \geq \varepsilon} \bar{l}(\theta) < \bar{l}(\theta_0)$.
  - Takaa, että funktiolla $\bar{l}$ on yksikäsitteinen maksimi pisteessä θ_0.
  - Voidaan tulkita parametria θ koskevaksi identifiointiehdoksi.

- Kun SU–estimaattorin tarkentuvuus on todettu, voidaan asymptoottinen normaalisuus osoittaa käyttäen pistemäärän Taylorin kehitelmää tai väliarvolausetta.

- Kun SU–estimaattorin tarkentuvuus on todettu, voidaan asymptoottinen normaalisuus osoittaa käyttäen pistemäärän Taylorin kehitelmää tai väliarvolausetta.
- Tämä vaatii uskottavuusfunktion toisten derivaattojen olemassaolon, joka seuraa mallin oletetusta säännöllisyydestä (jakso 2.3).

- Kun SU–estimaattorin tarkentuvuus on todettu, voidaan asymptoottinen normaalisuus osoittaa käyttäen pistemäärän Taylorin kehitelmää tai väliarvolausetta.
- Tämä vaatii uskottavuusfunktion toisten derivaattojen olemassaolon, joka seuraa mallin oletetusta säännöllisyydestä (jakso 2.3).
- Keskeinen elementti SU–estimaattorin asymptoottisen normaalisuuden osoittamisessa on säännöllisyydestä seuraava pistemäärän martingaaliominaisuus (ks. Lause 2.1), joka sopivin lisäehdoin mahdollistaa martingaalien KRL:een soveltamisen (ks. Lause 1.6).

- Kun SU–estimaattorin tarkentuvuus on todettu, voidaan asymptoottinen normaalisuus osoittaa käyttäen pistemäärän Taylorin kehitelmää tai väliarvolausetta.
- Tämä vaatii uskottavuusfunktion toisten derivaattojen olemassaolon, joka seuraa mallin oletetusta säännöllisyydestä (jakso 2.3).
- Keskeinen elementti SU–estimaattorin asymptoottisen normaalisuuden osoittamisessa on säännöllisyydestä seuraava pistemäärän martingaaliominaisuus (ks. Lause 2.1), joka sopivin lisäehdoin mahdollistaa martingaalien KRL:een soveltamisen (ks. Lause 1.6).
- PNS–estimaattoria koskeva esimerkki havainnollistaa yleisessä tapauksessa vaadittavia tarkasteluja.

- Tarkastellaan lineaarista mallia

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.

- Tarkastellaan lineaarista mallia

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.

- Malliyhtälöstä seuraa

$$\sqrt{n}(\hat{\beta} - \beta) = \left(n^{-1}\sum_{i=1}^{n} Z_i Z_i'\right)^{-1} n^{-1/2}\sum_{i=1}^{n} Z_i\varepsilon_i,$$

jossa oletetaan jälleen

- Tarkastellaan lineaarista mallia

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.

- Malliyhtälöstä seuraa

$$\sqrt{n}(\hat{\beta} - \beta) = \left(n^{-1} \sum_{i=1}^{n} Z_i Z_i'\right)^{-1} n^{-1/2} \sum_{i=1}^{n} Z_i \varepsilon_i,$$

jossa oletetaan jälleen

  - $n^{-1} \sum_{i=1}^{n} Z_i Z_i' \xrightarrow{p} Q, \quad Q$ ei–satunnainen ja positiivisesti definiitti

- Tarkastellaan lineaarista mallia

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.

- Malliyhtälöstä seuraa

$$\sqrt{n}(\hat{\beta} - \beta) = \left(n^{-1} \sum_{i=1}^{n} Z_i Z_i'\right)^{-1} n^{-1/2} \sum_{i=1}^{n} Z_i \varepsilon_i,$$

jossa oletetaan jälleen

  - $n^{-1} \sum_{i=1}^{n} Z_i Z_i' \xrightarrow{p} Q, \quad Q$ ei–satunnainen ja positiivisesti definiitti

- Kun $\mathrm{E}\left(|Z_{i,a}|\right) < \infty$), on $\{Z_i \varepsilon_i\}_{i=1}^{\infty}$ on MD–jono (ks. HT 3.1) ja voidaan olettaa (ks. lause 1.6)

- Tarkastellaan lineaarista mallia

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.
- Malliyhtälöstä seuraa

$$\sqrt{n}(\hat{\beta} - \beta) = \left(n^{-1} \sum_{i=1}^{n} Z_i Z_i'\right)^{-1} n^{-1/2} \sum_{i=1}^{n} Z_i \varepsilon_i,$$

jossa oletetaan jälleen
  - $n^{-1} \sum_{i=1}^{n} Z_i Z_i' \xrightarrow{p} Q, \quad Q$ ei–satunnainen ja positiivisesti definiitti
- Kun $\mathrm{E}\left(|Z_{i,a}|\right) < \infty$), on $\{Z_i\varepsilon_i\}_{i=1}^{\infty}$ on MD–jono (ks. HT 3.1) ja voidaan olettaa (ks. lause 1.6)
  - $n^{-1/2} \sum_{i=1}^{n} Z_i \varepsilon_i \xrightarrow{d} \mathrm{N}\left(0, \sigma^2 Q\right)$

- Tarkastellaan lineaarista mallia

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.

- Malliyhtälöstä seuraa

$$\sqrt{n}(\hat{\beta} - \beta) = \left(n^{-1} \sum_{i=1}^{n} Z_i Z_i'\right)^{-1} n^{-1/2} \sum_{i=1}^{n} Z_i \varepsilon_i,$$

jossa oletetaan jälleen

  - $n^{-1} \sum_{i=1}^{n} Z_i Z_i' \xrightarrow{p} Q$, $\quad Q$ ei–satunnainen ja positiivisesti definiitti

- Kun $\mathrm{E}\left(|Z_{i,a}|\right) < \infty)$, on $\{Z_i\varepsilon_i\}_{i=1}^{\infty}$ on MD–jono (ks. HT 3.1) ja voidaan olettaa (ks. lause 1.6)

  - $n^{-1/2} \sum_{i=1}^{n} Z_i \varepsilon_i \xrightarrow{d} \mathrm{N}\left(0, \sigma^2 Q\right)$

- Saadaan tulos

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} \mathrm{N}\left(0, \sigma^2 Q^{-1}\right)$$

SU–estimaattorin asymptotiikka

Asymptoottinen normaalisuus: Yleinen tapaus

- Aineistoa vastaava sv $\mathbf{Y}_n$ noudattaa jälleen mallia $f_{\mathbf{Y}_n}(\mathbf{y}_n; \theta_0)$.

- Aineistoa vastaava sv $\mathbf{Y}_n$ noudattaa jälleen mallia $f_{\mathbf{Y}_n}(\mathbf{y}_n; \theta_0)$.
- Pistemäärä $s_n(\theta; \mathbf{y}_n)$, havaittu info $\mathcal{J}_n(\theta; \mathbf{y}_n)$, Fisherin info $\mathcal{I}_n(\theta)$:

$$
\begin{aligned}
s_n(\theta; \mathbf{y}_n) &= \frac{\partial}{\partial \theta} l_n(\theta; \mathbf{y}_n) \quad (d \times 1) \\
s_{a,n}(\theta; \mathbf{y}_n) &= \frac{\partial}{\partial \theta_a} l_n(\theta; \mathbf{y}_n) \quad (a = 1, ..., d) \\
\frac{\partial}{\partial \theta'} s_{a,n}(\theta; \mathbf{y}_n) &= \left[\frac{\partial}{\partial \theta_1} s_{a,n}(\theta; \mathbf{y}_n) \ \cdots \ \frac{\partial}{\partial \theta_d} s_{a,n}(\theta; \mathbf{y}_n) \right] \quad (1 \times d) \\
\mathcal{I}_n(\theta) &= \mathrm{E}_\theta \left[\mathcal{J}_n(\theta; \mathbf{Y}_n) \right] = -\mathrm{E}_\theta \left[\partial^2 l_n(\theta; \mathbf{Y}_n) / \partial \theta \partial \theta' \right]
\end{aligned}
$$

- Aineistoa vastaava sv $\mathbf{Y}_n$ noudattaa jälleen mallia $f_{\mathbf{Y}_n}(\mathbf{y}_n; \theta_0)$.
- Pistemäärä $s_n(\theta; \mathbf{y}_n)$, havaittu info $\mathcal{J}_n(\theta; \mathbf{y}_n)$, Fisherin info $\mathcal{I}_n(\theta)$:

$$
\begin{aligned}
s_n(\theta; \mathbf{y}_n) &= \frac{\partial}{\partial \theta} l_n(\theta; \mathbf{y}_n) \quad (d \times 1) \\
s_{a,n}(\theta; \mathbf{y}_n) &= \frac{\partial}{\partial \theta_a} l_n(\theta; \mathbf{y}_n) \quad (a = 1, ..., d) \\
\frac{\partial}{\partial \theta'} s_{a,n}(\theta; \mathbf{y}_n) &= \left[\frac{\partial}{\partial \theta_1} s_{a,n}(\theta; \mathbf{y}_n) \ \cdots \ \frac{\partial}{\partial \theta_d} s_{a,n}(\theta; \mathbf{y}_n) \right] \quad (1 \times d) \\
\mathcal{I}_n(\theta) &= \mathrm{E}_\theta \left[\mathcal{J}_n(\theta; \mathbf{Y}_n) \right] = -\mathrm{E}_\theta \left[\partial^2 l_n(\theta; \mathbf{Y}_n) / \partial \theta \theta' \right]
\end{aligned}
$$

- $\frac{\partial}{\partial \theta'} s_{a,n}(\theta; \mathbf{y}_n) =$ matriisin $\partial^2 l_n(\theta; \mathbf{y}_n) / \partial \theta \theta' = -\mathcal{J}_n(\theta; \mathbf{y}_n)$ a. rivi.

- Aineistoa vastaava sv $\mathbf{Y}_n$ noudattaa jälleen mallia $f_{\mathbf{Y}_n}(\mathbf{y}_n;\theta_0)$.
- Pistemäärä $s_n(\theta;\mathbf{y}_n)$, havaittu info $\mathcal{J}_n(\theta;\mathbf{y}_n)$, Fisherin info $\mathcal{I}_n(\theta)$:

$$
\begin{aligned}
s_n(\theta;\mathbf{y}_n) &= \frac{\partial}{\partial\theta} l_n(\theta;\mathbf{y}_n) \quad (d \times 1) \\
s_{a,n}(\theta;\mathbf{y}_n) &= \frac{\partial}{\partial\theta_a} l_n(\theta;\mathbf{y}_n) \quad (a = 1, ..., d) \\
\frac{\partial}{\partial\theta'} s_{a,n}(\theta;\mathbf{y}_n) &= \left[\frac{\partial}{\partial\theta_1} s_{a,n}(\theta;\mathbf{y}_n) \ \cdots \ \frac{\partial}{\partial\theta_d} s_{a,n}(\theta;\mathbf{y}_n) \right] \quad (1 \times d) \\
\mathcal{I}_n(\theta) &= \mathrm{E}_\theta\left[\mathcal{J}_n(\theta;\mathbf{Y}_n)\right] = -\mathrm{E}_\theta\left[\partial^2 l_n(\theta;\mathbf{Y}_n)/\partial\theta\theta'\right]
\end{aligned}
$$

- $\frac{\partial}{\partial\theta'} s_{a,n}(\theta;\mathbf{y}_n) =$ matriisin $\partial^2 l_n(\theta;\mathbf{y}_n)/\partial\theta\theta' = -\mathcal{J}_n(\theta;\mathbf{y}_n)$ a. rivi.
- Parametriavaruuden $\Theta \subseteq \mathbb{R}^d$ oletetaan olevan avoin ja konveksi, mikä takaa tarvittavan väliarvolauseen soveltamisen.

- Usean muuttujan funktion väliarvolause $\Rightarrow$

$$s_{a,n}(\hat{\theta}_n; \mathbf{Y}_n) = s_{a,n}(\theta_0; \mathbf{Y}_n) + \frac{\partial}{\partial \theta'} s_{a,n}\left(\bar{\theta}_{a,n}; \mathbf{Y}_n\right)\left(\hat{\theta}_n - \theta_0\right), \quad a = 1, ..., d,$$

jossa

$$\bar{\theta}_{a,n} = c_a \hat{\theta}_n + (1 - c_a)\,\theta_0,\ \ 0 \leq c_a \leq 1, \quad a = 1, ..., d.$$

- Usean muuttujan funktion väliarvolause $\Rightarrow$

$$s_{a,n}(\hat{\theta}_n;\mathbf{Y}_n) = s_{a,n}\left(\theta_0;\mathbf{Y}_n\right) + \frac{\partial}{\partial\theta'}s_{a,n}\left(\bar{\theta}_{a,n};\mathbf{Y}_n\right)\left(\hat{\theta}_n - \theta_0\right), \quad a = 1, ..., d,$$

jossa

$$\bar{\theta}_{a,n} = c_a\hat{\theta}_n + \left(1 - c_a\right)\theta_0,\ 0 \leq c_a \leq 1, \quad a = 1, ..., d.$$

- Välipisteille $\bar{\theta}_{a,n}$ (jotka satunnaisia) pätee

$$||\bar{\theta}_{a,n} - \theta_0|| \leq ||\hat{\theta}_n - \theta_0||, \quad a = 1, ..., d.$$

- Usean muuttujan funktion väliarvolause $\Rightarrow$

$$s_{a,n}(\hat{\theta}_n; \mathbf{Y}_n) = s_{a,n}\left(\theta_0; \mathbf{Y}_n\right) + \frac{\partial}{\partial \theta'} s_{a,n}\left(\bar{\theta}_{a,n}; \mathbf{Y}_n\right)\left(\hat{\theta}_n - \theta_0\right), \quad a = 1, \ldots, d,$$

  jossa

$$\bar{\theta}_{a,n} = c_a \hat{\theta}_n + (1 - c_a)\,\theta_0,\ 0 \leq c_a \leq 1, \quad a = 1, \ldots, d.$$

- Välipisteille $\bar{\theta}_{a,n}$ (jotka satunnaisia) pätee

$$||\bar{\theta}_{a,n} - \theta_0|| \leq ||\hat{\theta}_n - \theta_0||, \quad a = 1, \ldots, d.$$

- Oletus $\hat{\theta}_n \xrightarrow{p} \theta_0 \Rightarrow \bar{\theta}_{a,n} \xrightarrow{p} \theta_0$ (HT 1.1)

- $$s_n(\hat{\theta}_n;\mathbf{Y}_n) = s_n\left(\theta_0;\mathbf{Y}_n\right) - \bar{\mathcal{J}}_n\left(\bar{\theta}_n;\mathbf{Y}_n\right)\left(\hat{\theta}_n - \theta_0\right), \qquad (*)$$

  jossa

  $$\bar{\mathcal{J}}_n\left(\bar{\theta}_n;\mathbf{Y}_n\right) \text{:n } a\text{. rivi} = \mathcal{J}_n\left(\bar{\theta}_{a,n};\mathbf{Y}_n\right) \text{:n } a\text{. rivi}$$

  eli matriisin $\bar{\mathcal{J}}_n\left(\bar{\theta}_n;\mathbf{Y}_n\right)$ eri riveillä on eri välipisteet.

- $$s_n(\hat{\theta}_n; \mathbf{Y}_n) = s_n\left(\theta_0; \mathbf{Y}_n\right) - \bar{\mathcal{J}}_n\left(\bar{\theta}_n; \mathbf{Y}_n\right)\left(\hat{\theta}_n - \theta_0\right), \qquad (*)$$

  jossa

  $$\bar{\mathcal{J}}_n\left(\bar{\theta}_n; \mathbf{Y}_n\right) \text{:n } a\text{. rivi} = \mathcal{J}_n\left(\bar{\theta}_{a,n}; \mathbf{Y}_n\right) \text{:n } a\text{. rivi}$$

  eli matriisin $\bar{\mathcal{J}}_n\left(\bar{\theta}_n; \mathbf{Y}_n\right)$ eri riveillä on eri välipisteet.
- SUE:n $\hat{\theta}_n$ oletettu olemassaolo $\Rightarrow s_n(\hat{\theta}_n; \mathbf{Y}_n) = 0$.

- $$s_n(\hat{\theta}_n; \mathbf{Y}_n) = s_n(\theta_0; \mathbf{Y}_n) - \bar{\mathcal{J}}_n(\bar{\theta}_n; \mathbf{Y}_n)(\hat{\theta}_n - \theta_0), \qquad (*)$$

  jossa

  $$\bar{\mathcal{J}}_n(\bar{\theta}_n; \mathbf{Y}_n) \text{ :n } a. \text{ rivi} = \mathcal{J}_n(\bar{\theta}_{a,n}; \mathbf{Y}_n) \text{ :n } a. \text{ rivi}$$

  eli matriisin $\bar{\mathcal{J}}_n(\bar{\theta}_n; \mathbf{Y}_n)$ eri riveillä on eri välipisteet.
- SUE:n $\hat{\theta}_n$ oletettu olemassaolo $\Rightarrow s_n(\hat{\theta}_n; \mathbf{Y}_n) = 0$.
- Jos *oletetaan* $\bar{\mathcal{J}}_n(\bar{\theta}_n; \mathbf{Y}_n)$ epäsingulaariseksi, saadaan $(*)$:stä

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = \left(\frac{1}{n}\bar{\mathcal{J}}_n(\bar{\theta}_n; \mathbf{Y}_n)\right)^{-1} \frac{1}{\sqrt{n}} s_n(\theta_0; \mathbf{Y}_n).$$

- $\sqrt{n}(\hat{\theta}_n - \theta_0) = \left(\frac{1}{n}\bar{\mathcal{J}}_n\left(\bar{\theta}_n; \mathbf{Y}_n\right)\right)^{-1} \frac{1}{\sqrt{n}} s_n\left(\theta_0; \mathbf{Y}_n\right).$

- $\sqrt{n}(\hat{\theta}_n - \theta_0) = \left(\frac{1}{n}\bar{\mathcal{J}}_n\left(\bar{\theta}_n; \mathbf{Y}_n\right)\right)^{-1}\frac{1}{\sqrt{n}}s_n\left(\theta_0; \mathbf{Y}_n\right).$
- $\bar{\mathcal{J}}_n\left(\bar{\theta}_n; \mathbf{Y}_n\right)$:n a. rivi $= \mathcal{J}_n\left(\bar{\theta}_{a,n}; \mathbf{Y}_n\right)$:n a. rivi

- $\sqrt{n}(\hat{\theta}_n - \theta_0) = \left(\frac{1}{n}\bar{\mathcal{J}}_n\left(\bar{\theta}_n;\mathbf{Y}_n\right)\right)^{-1}\frac{1}{\sqrt{n}}s_n\left(\theta_0;\mathbf{Y}_n\right).$
- $\bar{\mathcal{J}}_n\left(\bar{\theta}_n;\mathbf{Y}_n\right)$:n a. rivi $= \mathcal{J}_n\left(\bar{\theta}_{a,n};\mathbf{Y}_n\right)$:n a. rivi
- $\frac{1}{n}\mathcal{J}_n\left(\bar{\theta}_{a,n};\mathbf{Y}_n\right) = -\frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2}{\partial\theta\partial\theta'}\log f_{i-1}\left(Y_i;\bar{\theta}_{a,n}\right)$ on otoskeskiarvo, johon tavoitteena soveltaa SLL:ia.

- $\sqrt{n}(\hat{\theta}_n - \theta_0) = \left(\frac{1}{n}\bar{\mathcal{J}}_n\left(\bar{\theta}_n; \mathbf{Y}_n\right)\right)^{-1} \frac{1}{\sqrt{n}} s_n\left(\theta_0; \mathbf{Y}_n\right).$
- $\bar{\mathcal{J}}_n\left(\bar{\theta}_n; \mathbf{Y}_n\right)$:n a. rivi $= \mathcal{J}_n\left(\bar{\theta}_{a,n}; \mathbf{Y}_n\right)$:n a. rivi
- $\frac{1}{n}\mathcal{J}_n\left(\bar{\theta}_{a,n}; \mathbf{Y}_n\right) = -\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2}{\partial\theta\partial\theta'} \log f_{i-1}\left(Y_i; \bar{\theta}_{a,n}\right)$ on otoskeskiarvo, johon tavoitteena soveltaa SLL:ia.
- Koska argumentti $\bar{\theta}_{a,n}$ on satunnainen, tarvitaan skaalatulle havaitulle informaatiomatriisille $\frac{1}{n}\mathcal{J}_n\left(\theta; \mathbf{Y}_n\right)$ tasainen SLL (vrt. Lause 1.5).

- Oletetaan

$$\frac{1}{n}\mathcal{J}_n(\theta; \mathbf{Y}_n) \xrightarrow{p} \bar{\mathcal{I}}(\theta) \quad \text{tasaisesti joukossa } \Theta,$$

jossa

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\mathcal{I}_n(\theta) &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathsf{E}\left[-\frac{\partial^2}{\partial\theta\partial\theta'}\log f_{i-1}(\mathbf{Y}_i;\theta)\right] \\ &= \lim_{n\to\infty}\frac{1}{n}\mathsf{E}\left[\mathcal{J}_n(\theta;\mathbf{Y}_n)\right] = \bar{\mathcal{I}}(\theta), \quad \theta\in\Theta,\end{aligned}$$

- Oletetaan

$$\frac{1}{n}\mathcal{J}_n\left(\theta;\mathbf{Y}_n\right)\xrightarrow{p}\bar{\mathcal{I}}\left(\theta\right)\quad\text{tasaisesti joukossa }\Theta,$$

jossa

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\mathcal{I}_n\left(\theta\right) &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathsf{E}\left[-\frac{\partial^2}{\partial\theta\partial\theta'}\log f_{i-1}\left(\mathbf{Y}_i;\theta\right)\right]\\ &= \lim_{n\to\infty}\frac{1}{n}\mathsf{E}\left[\mathcal{J}_n\left(\theta;\mathbf{Y}_n\right)\right]=\bar{\mathcal{I}}\left(\theta\right),\quad\theta\in\Theta,\end{aligned}$$

- ja raja-arvon olemassaolo oletetaan.

- Oletetaan

$$\frac{1}{n}\mathcal{J}_n(\theta;\mathbf{Y}_n) \xrightarrow{p} \bar{\mathcal{I}}(\theta) \quad \text{tasaisesti joukossa } \Theta,$$

jossa

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\mathcal{I}_n(\theta) &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathsf{E}\left[-\frac{\partial^2}{\partial\theta\partial\theta'}\log f_{i-1}(\mathbf{Y}_i;\theta)\right]\\ &= \lim_{n\to\infty}\frac{1}{n}\mathsf{E}\left[\mathcal{J}_n(\theta;\mathbf{Y}_n)\right] = \bar{\mathcal{I}}(\theta),\quad \theta\in\Theta,\end{aligned}$$

- ja raja-arvon olemassaolo oletetaan.
- Samoin $\bar{\mathcal{I}}(\theta)$:n jatkossa tarvittava jatkuvuus ja positiivisdefiniittisyys (ja siten epäsingulaarisuus) pisteessä θ_0.

- $\sqrt{n}(\hat{\theta}_n - \theta_0) = \left(\frac{1}{n}\bar{\mathcal{J}}_n\left(\bar{\theta}_n; \mathbf{Y}_n\right)\right)^{-1} \frac{1}{\sqrt{n}} s_n\left(\theta_0; \mathbf{Y}_n\right).$

- $\sqrt{n}(\hat{\theta}_n - \theta_0) = \left(\frac{1}{n}\bar{\mathcal{J}}_n\left(\bar{\theta}_n; \mathbf{Y}_n\right)\right)^{-1} \frac{1}{\sqrt{n}} s_n\left(\theta_0; \mathbf{Y}_n\right).$
- $s_n\left(\theta_0; \mathbf{Y}_n\right) = \sum_{i=1}^{n} u_i\left(\theta_0; \mathbf{Y}_i\right)$ on martingaali (Lause 2.1) ja (ks. (2.23))

$$\mathrm{Cov}_{\theta_0}\left[n^{-1/2} s_n\left(\theta_0; \mathbf{Y}_n\right)\right] = n^{-1}\mathcal{I}_n\left(\theta_0\right) = n^{-1} \times \text{Fisherin info}\left(\theta_0\right)$$

- $\sqrt{n}(\hat{\theta}_n - \theta_0) = \left(\frac{1}{n}\bar{\mathcal{J}}_n\left(\bar{\theta}_n; \mathbf{Y}_n\right)\right)^{-1} \frac{1}{\sqrt{n}} s_n\left(\theta_0; \mathbf{Y}_n\right).$

- $s_n\left(\theta_0; \mathbf{Y}_n\right) = \sum_{i=1}^n u_i\left(\theta_0; \mathbf{Y}_i\right)$ on martingaali (Lause 2.1) ja (ks. (2.23))

$$\mathrm{Cov}_{\theta_0}\left[n^{-1/2} s_n\left(\theta_0; \mathbf{Y}_n\right)\right] = n^{-1}\mathcal{I}_n\left(\theta_0\right) = \ n^{-1} \times \text{Fisherin info}\left(\theta_0\right)$$

- Kun MD–jonon $u_i\left(\theta_0; \mathbf{Y}_i\right)$, $i \geq 1$, oletetaan toteuttavan Lauseen 1.6 ehdot (i) ja (ii), saadaan

$$\frac{1}{\sqrt{n}} s_n\left(\theta_0; \mathbf{Y}_n\right) \xrightarrow{d} \mathrm{N}\left(0, \bar{\mathcal{I}}\left(\theta_0\right)\right), \qquad \bar{\mathcal{I}}\left(\theta\right) = \lim_{n\to\infty} \frac{1}{n}\mathcal{I}_n\left(\theta\right).$$

- Oletukset

$$\frac{1}{n}\mathcal{J}_n\left(\theta;\mathbf{Y}_n\right) \xrightarrow{p} \bar{\mathcal{I}}\left(\theta\right) \quad \text{tasaisesti joukossa } \Theta, \quad (2.30)$$

$$\lim_{n\to\infty}\frac{1}{n}\mathrm{E}\left[\mathcal{J}_n(\theta;\mathbf{Y}_n)\right] = \bar{\mathcal{I}}\left(\theta\right), \quad \theta\in\Theta. \quad (2.31)$$

$$\frac{1}{\sqrt{n}}s_n\left(\theta_0;\mathbf{Y}_n\right) \xrightarrow{d} \mathrm{N}\left(0,\bar{\mathcal{I}}\left(\theta_0\right)\right) \quad (2.32)$$

- Oletukset

$$\frac{1}{n}\mathcal{J}_n\left(\theta;\mathbf{Y}_n\right) \xrightarrow{p} \bar{\mathcal{I}}\left(\theta\right) \quad \text{tasaisesti joukossa } \Theta, \quad (2.30)$$

$$\lim_{n\to\infty}\frac{1}{n}\mathrm{E}\left[\mathcal{J}_n(\theta;\mathbf{Y}_n)\right] = \bar{\mathcal{I}}\left(\theta\right), \quad \theta\in\Theta. \quad (2.31)$$

$$\frac{1}{\sqrt{n}}s_n\left(\theta_0;\mathbf{Y}_n\right) \xrightarrow{d} \mathrm{N}\left(0,\bar{\mathcal{I}}\left(\theta_0\right)\right) \quad (2.32)$$

- **Lause 2.3.** Olkoon $f_{\mathbf{Y}_n}\left(\mathbf{y}_n;\theta\right)$, $\theta\in\Theta$, säännöllinen ja Θ avoin ja konveksi. Oletetaan, (2.30), (2.31) ja (2.32) ja että $\bar{\mathcal{I}}\left(\theta\right)$ on positiivisesti definiitti ja jatkuva pisteessä θ_0. Jos SU–estimaattori $\hat{\theta}_n$ on lisäksi tarkentuva, niin

$$\sqrt{n}(\hat{\theta}_n-\theta_0) \xrightarrow{d} \mathrm{N}\left(0,\bar{\mathcal{I}}\left(\theta_0\right)^{-1}\right) \quad \text{ja} \quad n^{-1}\mathcal{J}_n(\hat{\theta}_n;\mathbf{Y}_n) \xrightarrow{p} \bar{\mathcal{I}}\left(\theta_0\right).$$

- $$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \bar{\mathcal{I}}\left(\theta_0\right)^{-1}\right)$$

- $$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \bar{\mathcal{I}}\left(\theta_0\right)^{-1}\right)$$

- iid–tapauksessa

$$\begin{aligned}
\bar{\mathcal{I}}\left(\theta_0\right) &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}\left[-\frac{\partial^2}{\partial\theta\partial\theta'} \log f_{i-1}\left(Y_i;\theta\right)\right] \\
&= \mathrm{E}\left[-\frac{\partial^2}{\partial\theta\partial\theta'} \log f_{Y_1}\left(Y_1;\theta\right)\right]
\end{aligned}$$

on yksittäisen havainnon Fisherin informaatio.

-
$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \bar{\mathcal{I}}\left(\theta_0\right)^{-1}\right)$$

- iid–tapauksessa
$$\begin{aligned}\bar{\mathcal{I}}\left(\theta_0\right) &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}\left[-\frac{\partial^2}{\partial\theta\partial\theta'} \log f_{i-1}\left(Y_i;\theta\right)\right] \\ &= \mathrm{E}\left[-\frac{\partial^2}{\partial\theta\partial\theta'} \log f_{Y_1}\left(Y_1;\theta\right)\right]\end{aligned}$$
on yksittäisen havainnon Fisherin informaatio.

- Yleisesti, $\bar{\mathcal{I}}\left(\theta_0\right)$ on "keskimääräinen" Fisherin informaatioksi.

- $$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \bar{\mathcal{I}}\left(\theta_0\right)^{-1}\right)$$

- iid–tapauksessa
$$\begin{aligned}\bar{\mathcal{I}}\left(\theta_0\right) &= \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\left[-\frac{\partial^2}{\partial\theta\partial\theta'} \log f_{i-1}\left(Y_i;\theta\right)\right] \\ &= \mathrm{E}\left[-\frac{\partial^2}{\partial\theta\partial\theta'} \log f_{Y_1}\left(Y_1;\theta\right)\right]\end{aligned}$$
on yksittäisen havainnon Fisherin informaatio.

- Yleisesti, $\bar{\mathcal{I}}\left(\theta_0\right)$ on "keskimääräinen" Fisherin informaatioksi.

- Lauseen 2.3 tulos osoittaa SU–estimaattorin asymptoottisen täystehokkuuden.

- $$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \bar{\mathcal{I}}\left(\theta_0\right)^{-1}\right) \quad \text{ja} \quad n^{-1}\mathcal{J}_n(\hat{\theta}_n; \mathbf{Y}_n) \xrightarrow{p} \bar{\mathcal{I}}\left(\theta_0\right).$$

SU–estimaattorin asymptotiikka

Asymptoottinen normaalisuus: Yleinen tapaus

- $$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \bar{\mathcal{I}}\left(\theta_0\right)^{-1}\right) \quad \text{ja} \quad n^{-1}\mathcal{J}_n(\hat{\theta}_n; \mathbf{Y}_n) \xrightarrow{p} \bar{\mathcal{I}}\left(\theta_0\right).$$
- As. jakauman kovarianssimatriisin $\bar{\mathcal{I}}\left(\theta_0\right)^{-1}$ tarkentuva estimaattori saadaan kääntämällä skaalattu havaittu informaatiomatriisi $n^{-1}\mathcal{J}_n(\hat{\theta}_n; \mathbf{Y}_n)$.

- $$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \bar{\mathcal{I}}\left(\theta_0\right)^{-1}\right) \quad \text{ja} \quad n^{-1}\mathcal{J}_n(\hat{\theta}_n; \mathbf{Y}_n) \xrightarrow{p} \bar{\mathcal{I}}\left(\theta_0\right).$$
- As. jakauman kovarianssimatriisin $\bar{\mathcal{I}}\left(\theta_0\right)^{-1}$ tarkentuva estimaattori saadaan kääntämällä skaalattu havaittu informaatiomatriisi $n^{-1}\mathcal{J}_n(\hat{\theta}_n; \mathbf{Y}_n)$.
- Tätä tulosta voidaan käyttää SU–estimaattorin $\hat{\theta}_n$ komponenttien keskivirheiden laskemisessa tavanomaiseen tapaan.

- Lauseen 2.3 merkitys on oletusten yleisluonteisuuden vuoksi pitkälti periaatteellinen.

- Lauseen 2.3 merkitys on oletusten yleisluonteisuuden vuoksi pitkälti periaatteellinen.
- Konkreettisissa tilanteissa oletetut konvergenssit täytyy tarkistaa käyttäen tarkasteltavan mallin erityispiirteitä ja oletuksia.

- Lauseen 2.3 merkitys on oletusten yleisluonteisuuden vuoksi pitkälti periaatteellinen.
- Konkreettisissa tilanteissa oletetut konvergenssit täytyy tarkistaa käyttäen tarkasteltavan mallin erityispiirteitä ja oletuksia.
- Yleensä hankalin on tarkistaa tasaista SLL:ia koskeva konvergenssi

$$\frac{1}{n}\mathcal{J}_n(\theta; \mathbf{Y}_n) \xrightarrow{p} \bar{\mathcal{I}}(\theta) \quad \text{tasaisesti joukossa } \Theta.$$

- Lauseen 2.3 merkitys on oletusten yleisluonteisuuden vuoksi pitkälti periaatteellinen.
- Konkreettisissa tilanteissa oletetut konvergenssit täytyy tarkistaa käyttäen tarkasteltavan mallin erityispiirteitä ja oletuksia.
- Yleensä hankalin on tarkistaa tasaista SLL:ia koskeva konvergenssi

$$\frac{1}{n}\mathcal{J}_n(\theta; \mathbf{Y}_n) \xrightarrow{p} \bar{\mathcal{I}}(\theta) \quad \text{tasaisesti joukossa } \Theta.$$

- Teknisiä yksityiskohtia olennaisempaa on kuitenkin ymmärtää SU–estimaattorin asymptoottisen normaalisuuden todistamisessa käytettävät perusideat niiden toimiminen tapauksissa, joissa havainnot eivät ole riippumattomia ja samoin jakautuneita.

- Parametriavaruuden avoimuutta ja konveksisuutta ei välttämättä tarvita, mutta käytetty väliarvolause vaatii, että todellinen parametriarvo θ_0 on parametriavaruuden sisäpiste.

- Parametriavaruuden avoimuutta ja konveksisuutta ei välttämättä tarvita, mutta käytetty väliarvolause vaatii, että todellinen parametriarvo θ_0 on parametriavaruuden sisäpiste.
- Ei aina aivan harmiton oletus. Esimerkiksi malli

$$Y_j = \alpha_j 1_{n_j} + \gamma x_j + \varepsilon_j, \quad j = 1, \ldots, N,$$

jossa

$$\begin{aligned} &\varepsilon_1, \ldots, \varepsilon_N \ \perp\!\!\!\perp\ , \quad \varepsilon_j \sim \mathrm{N}\left(0, \sigma^2 I_{n_j}\right), \\ &\alpha_1, \ldots, \alpha_N \sim \mathrm{N}\left(\alpha, \omega^2\right) \ \perp\!\!\!\perp\ , \quad \omega^2 \geq 0 \\ &\left(\varepsilon_1, \ldots, \varepsilon_N\right) \ \perp\!\!\!\perp\ \left(\alpha_1, \ldots, \alpha_N\right). \end{aligned}$$

- Parametriavaruuden avoimuutta ja konveksisuutta ei välttämättä tarvita, mutta käytetty väliarvolause vaatii, että todellinen parametriarvo θ_0 on parametriavaruuden sisäpiste.
- Ei aina aivan harmiton oletus. Esimerkiksi malli

$$Y_j = \alpha_j 1_{n_j} + \gamma x_j + \varepsilon_j, \quad j = 1, \ldots, N,$$

jossa

$$\begin{aligned} &\varepsilon_1, \ldots, \varepsilon_N \ \perp\!\!\!\perp \ , \quad \varepsilon_j \sim \mathrm{N}\left(0, \sigma^2 I_{n_j}\right), \\ &\alpha_1, \ldots, \alpha_N \sim \mathrm{N}\left(\alpha, \omega^2\right) \ \perp\!\!\!\perp \ , \quad \omega^2 \geq 0 \\ &\left(\varepsilon_1, \ldots, \varepsilon_N\right) \ \perp\!\!\!\perp \ \left(\alpha_1, \ldots, \alpha_N\right). \end{aligned}$$

- Hypoteesi $\omega^2 = 0$ usein kiinnostava.

- Lauseen 2.3. oletuksin pätee

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \bar{\mathcal{I}}(\theta_0)^{-1}\right) \quad \text{ja} \quad n^{-1}\mathcal{J}_n(\hat{\theta}_n; \mathbf{Y}_n) \xrightarrow{p} \bar{\mathcal{I}}(\theta_0).$$

Oletuksiin kuului

$$\lim_{n\to\infty} \frac{1}{n}\mathcal{I}_n(\theta) = \lim_{n\to\infty} \frac{1}{n} \times \text{Fisherin info}(\theta) = \bar{\mathcal{I}}(\theta).$$

- Lauseen 2.3. oletuksin pätee

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \bar{\mathcal{I}}(\theta_0)^{-1}\right) \quad \text{ja} \quad n^{-1}\mathcal{J}_n(\hat{\theta}_n; \mathbf{Y}_n) \xrightarrow{p} \bar{\mathcal{I}}(\theta_0).$$

Oletuksiin kuului

$$\lim_{n\to\infty} \frac{1}{n}\mathcal{I}_n(\theta) = \lim_{n\to\infty} \frac{1}{n} \times \text{Fisherin info}(\theta) = \bar{\mathcal{I}}(\theta).$$

- Jos $\theta = (\psi, \lambda)$ ja ψ ja λ ovat ortogonaalisia, on $\mathcal{I}_n(\theta)$ lohkodiagonaalinen ja siten

$$\bar{\mathcal{I}}(\theta) = \begin{bmatrix} \bar{\mathcal{I}}_{\psi\psi}(\theta) & 0 \\ 0 & \bar{\mathcal{I}}_{\lambda\lambda}(\theta) \end{bmatrix}.$$

- Lauseen 2.3. oletuksin pätee

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \bar{\mathcal{I}}(\theta_0)^{-1}\right) \quad \text{ja} \quad n^{-1}\mathcal{J}_n(\hat{\theta}_n; \mathbf{Y}_n) \xrightarrow{p} \bar{\mathcal{I}}(\theta_0).$$

  Oletuksiin kuului

$$\lim_{n\to\infty} \frac{1}{n}\mathcal{I}_n(\theta) = \lim_{n\to\infty} \frac{1}{n} \times \text{Fisherin info}(\theta) = \bar{\mathcal{I}}(\theta).$$

- Jos $\theta = (\psi, \lambda)$ ja ψ ja λ ovat ortogonaalisia, on $\mathcal{I}_n(\theta)$ lohkodiagonaalinen ja siten

$$\bar{\mathcal{I}}(\theta) = \begin{bmatrix} \bar{\mathcal{I}}_{\psi\psi}(\theta) & 0 \\ 0 & \bar{\mathcal{I}}_{\lambda\lambda}(\theta) \end{bmatrix}.$$

- Tällöin SU–estimaattorit $\hat{\psi}_n$ ja $\hat{\lambda}_n$ ovat asymptoottisesti riippumattomia ja esimerkiksi

$$\sqrt{n}(\hat{\psi}_n - \psi_0) \xrightarrow{d} \mathrm{N}\left(0, \bar{\mathcal{I}}_{\psi\psi}(\theta_0)^{-1}\right).$$

Uskottavuusfunktioon perustuvia testejä

- Tarkastellaan (todelliseen parametriarvoon) liitettyä lineaarista nollahypoteesia

$$\mathrm{H}_0 : A\theta_0 = c,$$

jossa matriisi A $(q \times d)$ ja vektori c $(q \times 1)$ ovat tunnettuja ja A:n aste on q eli $\mathrm{r}(A) = q$.

Uskottavuusfunktioon perustuvia testejä

- Tarkastellaan (todelliseen parametriarvoon) liitettyä lineaarista nollahypoteesia

  $$\mathrm{H}_0 : A\theta_0 = c,$$

  jossa matriisi A $(q \times d)$ ja vektori c $(q \times 1)$ ovat tunnettuja ja A:n aste on q eli $\mathrm{r}(A) = q$.
- Vaihtoehtoinen hypoteesi on $A\theta_0 \neq c$.

Uskottavuusfunktioon perustuvia testejä

- Tarkastellaan (todelliseen parametriarvoon) liitettyä lineaarista nollahypoteesia

$$H_0 : A\theta_0 = c,$$

jossa matriisi A $(q \times d)$ ja vektori c $(q \times 1)$ ovat tunnettuja ja A:n aste on q eli $\mathrm{r}(A) = q$.

- Vaihtoehtoinen hypoteesi on $A\theta_0 \neq c$.
- Esimerkiksi,

$$A = [I_q : 0],$$

jolloin $H_0 \Rightarrow \theta_0$:n q ensimmäistä komponenttia $=$ tunnettu c.

- Waldin testi perustuu SU–estimaattoriin $\hat{\theta}$ ja Lauseen 2.3 tulokseen

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \bar{\mathcal{I}}\left(\theta_0\right)^{-1}\right) \quad \text{ja} \quad n^{-1}\mathcal{J}_n(\hat{\theta}_n; \mathbf{Y}_n) \xrightarrow{p} \bar{\mathcal{I}}\left(\theta_0\right).$$

- Waldin testi perustuu SU–estimaattoriin $\hat{\theta}$ ja Lauseen 2.3 tulokseen

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \bar{\mathcal{I}}\left(\theta_0\right)^{-1}\right) \quad \text{ja} \quad n^{-1}\mathcal{J}_n(\hat{\theta}_n; \mathbf{Y}_n) \xrightarrow{p} \bar{\mathcal{I}}\left(\theta_0\right).$$

- Testisuure perustuu erotukseen

$$A\hat{\theta} - c,$$

joka saa tyypillisesti "pieniä" arvoja, kun nollahypoteesi $A\theta_0 = c$ on tosi ja "suuria" arvoja, kun nollahypoteesi ei ole tosi.

- Waldin testi perustuu SU–estimaattoriin $\hat{\theta}$ ja Lauseen 2.3 tulokseen

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \bar{\mathcal{I}}\left(\theta_0\right)^{-1}\right) \quad \text{ja} \quad n^{-1}\mathcal{J}_n(\hat{\theta}_n; \mathbf{Y}_n) \xrightarrow{p} \bar{\mathcal{I}}\left(\theta_0\right).$$

- Testisuure perustuu erotukseen

$$A\hat{\theta} - c,$$

joka saa tyypillisesti "pieniä" arvoja, kun nollahypoteesi $A\theta_0 = c$ on tosi ja "suuria" arvoja, kun nollahypoteesi ei ole tosi.

- Testisuure

$$\mathrm{W} = \sqrt{n}(A\hat{\theta} - c)' \left[A\left(n^{-1}\mathcal{J}(\hat{\theta}; \mathbf{Y})\right)^{-1} A'\right]^{-1} \sqrt{n}(A\hat{\theta} - c) \underset{\mathrm{H}_0}{\xrightarrow{d}} \chi_q^2.$$

- Waldin testi perustuu SU–estimaattoriin $\hat{\theta}$ ja Lauseen 2.3 tulokseen

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}\left(0, \bar{\mathcal{I}}\left(\theta_0\right)^{-1}\right) \quad \text{ja} \quad n^{-1}\mathcal{J}_n(\hat{\theta}_n; \mathbf{Y}_n) \xrightarrow{p} \bar{\mathcal{I}}\left(\theta_0\right).$$

- Testisuure perustuu erotukseen

$$A\hat{\theta} - c,$$

joka saa tyypillisesti "pieniä" arvoja, kun nollahypoteesi $A\theta_0 = c$ on tosi ja "suuria" arvoja, kun nollahypoteesi ei ole tosi.

- Testisuure

$$\mathrm{W} = \sqrt{n}(A\hat{\theta} - c)' \left[A\left(n^{-1}\mathcal{J}(\hat{\theta}; \mathbf{Y})\right)^{-1} A'\right]^{-1} \sqrt{n}(A\hat{\theta} - c) \xrightarrow[\mathrm{H}_0]{d} \chi^2_q.$$

- Approksimatiivinen P–arvo

$$\mathrm{P} = \mathrm{P}_{\mathrm{H}_0}\left\{\mathrm{W} \geq \mathrm{W}\left(\mathbf{y}\right)\right\} \approx \mathrm{P}\left\{\chi^2_q \geq \mathrm{W}\left(\mathbf{y}\right)\right\}.$$

- Waldin testi asymptoottisesti pätevä käytetäänpä mitä tahansa matriisin $\bar{\mathcal{I}}(\theta_0)$ tarkentuvaa estimaattoria.

- Waldin testi asymptoottisesti pätevä käytetäänpä mitä tahansa matriisin $\bar{\mathcal{I}}(\theta_0)$ tarkentuvaa estimaattoria.
- $s_n(\theta_0; \mathbf{Y}_n) = \sum_{i=1}^{n} u_i(\theta_0; Y_i)$, jossa $u_i(\theta_0; Y_i)$, $i \geq 1$, on MD–jono ja siten korreloimaton, joten oletuksen (2.31) perusteella,

$$\begin{aligned} \frac{1}{n}\mathrm{E}\left[s(\theta_0; \mathbf{Y}_n)\, s(\theta_0; \mathbf{Y}_n)'\right] &= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left[u_i(\theta_0; Y_i)\, u_i(\theta_0; Y_i)'\right] \\ &= \frac{1}{n}\mathcal{I}_n(\theta_0) \underset{n\to\infty}{\to} \bar{\mathcal{I}}(\theta_0), \end{aligned}$$

- Waldin testi asymptoottisesti pätevä käytetäänpä mitä tahansa matriisin $\bar{\mathcal{I}}(\theta_0)$ tarkentuvaa estimaattoria.
- $s_n(\theta_0;\mathbf{Y}_n) = \sum_{i=1}^n u_i(\theta_0;Y_i)$, jossa $u_i(\theta_0;Y_i)$, $i \geq 1$, on MD–jono ja siten korreloimaton, joten oletuksen (2.31) perusteella,

$$\begin{aligned}\frac{1}{n}\mathrm{E}\left[s(\theta_0;\mathbf{Y}_n)\,s(\theta_0;\mathbf{Y}_n)'\right] &= \frac{1}{n}\sum_{i=1}^n \mathrm{E}\left[u_i(\theta_0;Y_i)\,u_i(\theta_0;Y_i)'\right]\\ &= \frac{1}{n}\mathcal{I}_n(\theta_0) \underset{n\to\infty}{\to} \bar{\mathcal{I}}(\theta_0),\end{aligned}$$

- Siis, $\bar{\mathcal{I}}(\theta_0)$:n estimointi on luonteva perustaa myös ns. ulkotulomatriisiin

$$\mathrm{M}_n(\hat{\theta};\mathbf{Y}) = \sum_{i=1}^n u_i(\hat{\theta};Y_i)u_i(\hat{\theta};Y_i)'.$$

- Jos

$$\frac{1}{n}\mathrm{M}(\theta;\mathbf{Y}) = \frac{1}{n}\sum_{i=1}^{n} u_i(\theta;Y_i)u_i(\theta;Y_i)' \xrightarrow{p} \bar{\mathcal{I}}(\theta) \quad \text{tasaisesti joukossa } \Theta,$$

- Jos

$$\frac{1}{n}\mathrm{M}(\theta;\mathbf{Y}) = \frac{1}{n}\sum_{i=1}^{n} u_i(\theta;Y_i)u_i(\theta;Y_i)' \xrightarrow{p} \bar{\mathcal{I}}(\theta) \quad \text{tasaisesti joukossa } \Theta,$$

- niin $\hat{\theta} \xrightarrow{p} \theta_0$ ja Lause 1.5 $\Rightarrow$

$$\frac{1}{n}\mathrm{M}(\hat{\theta};\mathbf{Y}) = \frac{1}{n}\sum_{i=1}^{n} u_i(\hat{\theta};Y_i)u_i(\hat{\theta};Y_i)' \xrightarrow{p} \bar{\mathcal{I}}(\theta_0)\,.$$

- Jos

$$\frac{1}{n}\mathrm{M}(\theta;\mathbf{Y}) = \frac{1}{n}\sum_{i=1}^{n} u_i(\theta;Y_i)u_i(\theta;Y_i)' \xrightarrow{p} \bar{\mathcal{I}}(\theta) \quad \text{tasaisesti joukossa } \Theta,$$

- niin $\hat{\theta} \xrightarrow{p} \theta_0$ ja Lause 1.5 $\Rightarrow$

$$\frac{1}{n}\mathrm{M}(\hat{\theta};\mathbf{Y}) = \frac{1}{n}\sum_{i=1}^{n} u_i(\hat{\theta};Y_i)u_i(\hat{\theta};Y_i)' \xrightarrow{p} \bar{\mathcal{I}}(\theta_0)\,.$$

- Vaihtoehtoinen Waldin testisuure:

$$\mathrm{W} = \sqrt{n}(A\hat{\theta}-c)'\left[A\left(n^{-1}\mathrm{M}(\hat{\theta};\mathbf{Y})\right)^{-1}A'\right]^{-1}\sqrt{n}(A\hat{\theta}-c) \xrightarrow[\mathrm{H}_0]{d} \chi^2_q.$$

Uskottavuusfunktioon perustuvia testejä

Waldin testi: Vaihtoehtoinen testisuure

- Jos

$$\frac{1}{n}\mathrm{M}(\theta;\mathbf{Y}) = \frac{1}{n}\sum_{i=1}^{n} u_i(\theta;Y_i)u_i(\theta;Y_i)' \xrightarrow{p} \bar{\mathcal{I}}(\theta) \quad \text{tasaisesti joukossa } \Theta,$$

- niin $\hat{\theta} \xrightarrow{p} \theta_0$ ja Lause 1.5 $\Rightarrow$

$$\frac{1}{n}\mathrm{M}(\hat{\theta};\mathbf{Y}) = \frac{1}{n}\sum_{i=1}^{n} u_i(\hat{\theta};Y_i)u_i(\hat{\theta};Y_i)' \xrightarrow{p} \bar{\mathcal{I}}(\theta_0)\,.$$

- Vaihtoehtoinen Waldin testisuure:

$$\mathrm{W} = \sqrt{n}(A\hat{\theta}-c)'\left[A\left(n^{-1}\mathrm{M}(\hat{\theta};\mathbf{Y})\right)^{-1}A'\right]^{-1}\sqrt{n}(A\hat{\theta}-c) \xrightarrow[\mathrm{H}_0]{d} \chi_q^2.$$

- Toisin kuin Hessen matriisi $-\mathcal{J}(\theta;\mathbf{y})$ on ulkotulomatriisi $\mathrm{M}(\theta;\mathbf{Y})$ aina positiivisesti *semi*definiitti.

- Waldin testi voidaan yleistää epälineaarisille hypoteeseille

$$h(\theta_0) = 0,$$

- Waldin testi voidaan yleistää epälineaarisille hypoteeseille

$$h\left(\theta_0\right) = 0,$$

- jossa

- Waldin testi voidaan yleistää epälineaarisille hypoteeseille

$$h(\theta_0) = 0,$$

- jossa
  - $h : \Theta \to \mathbb{R}^q$ on jatkuvasti derivoituva ja

- Waldin testi voidaan yleistää epälineaarisille hypoteeseille

$$h(\theta_0) = 0,$$

- jossa
  - $h : \Theta \to \mathbb{R}^q$ on jatkuvasti derivoituva ja
  - $q \times d$ derivaattamatriisi

$$H(\theta) = \left[\partial h_a(\theta) / \partial \theta_b\right], a = 1, ..., q, b = 1, ..., d,$$

toteuttaa $\mathrm{r}(H(\theta_0)) = q$.

- Waldin testi voidaan yleistää epälineaarisille hypoteeseille

$$h(\theta_0) = 0,$$

- jossa
  - $h : \Theta \to \mathbb{R}^q$ on jatkuvasti derivoituva ja
  - $q \times d$ derivaattamatriisi

$$H(\theta) = [\partial h_a(\theta) / \partial \theta_b], a = 1, ..., q, b = 1, ..., d,$$

  toteuttaa $\mathrm{r}(H(\theta_0)) = q$.
- Lineaarinen hypoteesi erikoistapauksena valitsemalla $h(\theta) = A\theta - c$.

- Waldin testi voidaan yleistää epälineaarisille hypoteeseille

$$h(\theta_0) = 0,$$

- jossa
  - $h : \Theta \to \mathbb{R}^q$ on jatkuvasti derivoituva ja
  - $q \times d$ derivaattamatriisi

$$H(\theta) = \left[\partial h_a(\theta) / \partial \theta_b\right], a = 1, ..., q, b = 1, ..., d,$$

  toteuttaa $\mathrm{r}(H(\theta_0)) = q$.

- Lineaarinen hypoteesi erikoistapauksena valitsemalla $h(\theta) = A\theta - c$.
- Tässä tapauksessa Waldin testisuure on luentevaa perustaa suureeseen $h(\hat{\theta})$.

- Deltamenetelmä yleistys $\Rightarrow$ nollahypoteesin voimassa ollessa

$$\sqrt{n}h(\hat{\theta}) \stackrel{as}{=} H\left(\theta_0\right)\sqrt{n}(\hat{\theta}-\theta_0).$$

- Deltamenetelmä yleistys $\Rightarrow$ nollahypoteesin voimassa ollessa

$$\sqrt{n}h(\hat{\theta}) \overset{as}{=} H(\theta_0)\sqrt{n}(\hat{\theta} - \theta_0).$$

- Siis, Waldin testisuureeksi saadaan

$$\mathrm{W} = h(\hat{\theta})'[H(\hat{\theta})\mathcal{J}(\hat{\theta};\mathbf{Y})^{-1}H(\hat{\theta})']^{-1}h(\hat{\theta}) \xrightarrow[\mathrm{H}_0]{d} \chi^2_q,$$

jossa $\mathcal{J}(\hat{\theta};\mathbf{Y})$:n paikalla voi olla $\mathrm{M}(\hat{\theta};\mathbf{Y})$.

- Deltamenetelmä yleistys $\Rightarrow$ nollahypoteesin voimassa ollessa

$$\sqrt{n}h(\hat{\theta}) \stackrel{as}{=} H(\theta_0)\sqrt{n}(\hat{\theta} - \theta_0).$$

- Siis, Waldin testisuureeksi saadaan

$$\mathrm{W} = h(\hat{\theta})'[H(\hat{\theta})\mathcal{J}(\hat{\theta};\mathbf{Y})^{-1}H(\hat{\theta})']^{-1}h(\hat{\theta}) \xrightarrow[\mathrm{H}_0]{d} \chi^2_q,$$

jossa $\mathcal{J}(\hat{\theta};\mathbf{Y})$:n paikalla voi olla $\mathrm{M}(\hat{\theta};\mathbf{Y})$.

- Käytännössä testaus sujuu samaan tapaan kuin edellä lineaarisen hypoteesin tapauksessa.

- Ositetaan $\theta = (\psi, \lambda)$, jossa parametrit ψ ja λ ovat ortogonaaliset ja ψ on kiinnostava parametri.

- Ositetaan $\theta = (\psi, \lambda)$, jossa parametrit ψ ja λ ovat ortogonaaliset ja ψ on kiinnostava parametri.
- Testattava hypoteesi $\mathrm{H}_0 : A_\psi \psi = c \Leftrightarrow A\theta = c$, jossa $A = \left[A_\psi : 0\right]$.

- Ositetaan $\theta = (\psi, \lambda)$, jossa parametrit ψ ja λ ovat ortogonaaliset ja ψ on kiinnostava parametri.
- Testattava hypoteesi $\mathrm{H}_0 : A_\psi \psi = c \Leftrightarrow A\theta = c$, jossa $A = \left[A_\psi : 0\right]$.
- Nyt $\bar{\mathcal{I}}(\theta) = \mathrm{diag}[\bar{\mathcal{I}}_{\psi\psi}(\theta) \;\; \bar{\mathcal{I}}_{\lambda\lambda}(\theta)]$ ja $n^{-1}\mathcal{J}_o(\hat{\theta}; \mathbf{Y}) \xrightarrow{p} \bar{\mathcal{I}}(\theta_0)$, kun

$$\mathcal{J}_o(\hat{\theta}; \mathbf{Y}) = \mathrm{diag}[\mathcal{J}_{\psi\psi}(\hat{\theta}; \mathbf{Y}) \;\; \mathcal{J}_{\lambda\lambda}(\hat{\theta}; \mathbf{Y})].$$

- Ositetaan $\theta = (\psi, \lambda)$, jossa parametrit ψ ja λ ovat ortogonaaliset ja ψ on kiinnostava parametri.
- Testattava hypoteesi $\mathrm{H}_0 : A_\psi \psi = c \Leftrightarrow A\theta = c$, jossa $A = \left[A_\psi : 0\right]$.
- Nyt $\bar{\mathcal{I}}(\theta) = \mathrm{diag}[\bar{\mathcal{I}}_{\psi\psi}(\theta) \;\; \bar{\mathcal{I}}_{\lambda\lambda}(\theta)]$ ja $n^{-1}\mathcal{J}_\mathrm{o}(\hat{\theta}; \mathbf{Y}) \xrightarrow{p} \bar{\mathcal{I}}(\theta_0)$, kun

$$\mathcal{J}_\mathrm{o}(\hat{\theta}; \mathbf{Y}) = \mathrm{diag}[\mathcal{J}_{\psi\psi}(\hat{\theta}; \mathbf{Y}) \;\; \mathcal{J}_{\lambda\lambda}(\hat{\theta}; \mathbf{Y})].$$

- Lisäksi,

$$\left(n^{-1}\mathcal{J}_\mathrm{o}(\hat{\theta}; \mathbf{Y})\right)^{-1} \xrightarrow{p} \mathrm{diag}[\bar{\mathcal{I}}_{\psi\psi}^{-1}(\theta) \;\; \bar{\mathcal{I}}_{\lambda\lambda}^{-1}(\theta)].$$

- Ositetaan $\theta = (\psi, \lambda)$, jossa parametrit ψ ja λ ovat ortogonaaliset ja ψ on kiinnostava parametri.
- Testattava hypoteesi $\mathrm{H}_0 : A_\psi \psi = c \Leftrightarrow A\theta = c$, jossa $A = \left[A_\psi : 0\right]$.
- Nyt $\bar{\mathcal{I}}(\theta) = \mathrm{diag}[\bar{\mathcal{I}}_{\psi\psi}(\theta) \;\; \bar{\mathcal{I}}_{\lambda\lambda}(\theta)]$ ja $n^{-1}\mathcal{J}_o(\hat{\theta}; \mathbf{Y}) \xrightarrow{p} \bar{\mathcal{I}}(\theta_0)$, kun

$$\mathcal{J}_o(\hat{\theta}; \mathbf{Y}) = \mathrm{diag}[\mathcal{J}_{\psi\psi}(\hat{\theta}; \mathbf{Y}) \;\; \mathcal{J}_{\lambda\lambda}(\hat{\theta}; \mathbf{Y})].$$

- Lisäksi,

$$\left(n^{-1}\mathcal{J}_o(\hat{\theta}; \mathbf{Y})\right)^{-1} \xrightarrow{p} \mathrm{diag}[\bar{\mathcal{I}}_{\psi\psi}^{-1}(\theta) \;\; \bar{\mathcal{I}}_{\lambda\lambda}^{-1}(\theta)].$$

- Saadaan Waldin testisuure

$$\mathrm{W}_\psi = (A_\psi \hat{\psi} - c)'[A_\psi \mathcal{J}_{\psi\psi}^{-1}(\hat{\theta}; \mathbf{Y}) A_\psi']^{-1}(A_\psi \hat{\psi} - c) \xrightarrow[\mathrm{H}_0]{d} \chi_q^2.$$

- Tarkastellaan aluksi rajoitettua SU–estimaattoria $\tilde{\theta}$, joka maksimoi uskottavuusfunktion ehdolla $A\theta = c$ ja toteuttaa siten $A\tilde{\theta} = c$.

- Tarkastellaan aluksi rajoitettua SU–estimaattoria $\tilde{\theta}$, joka maksimoi uskottavuusfunktion ehdolla $A\theta = c$ ja toteuttaa siten $A\tilde{\theta} = c$.

- Lineaarialgebraa käyttäen voidaan todeta, että

$$\underset{q\times d}{A}\,\theta_0 = \underset{q\times 1}{c}, \;\; \mathrm{r}\left(A\right) = q \quad \Leftrightarrow \quad \theta_0 = \underset{d\times(d-q)}{B}\,\delta_0 + \underset{d\times 1}{e}, \;\; \mathrm{r}\left(B\right) = d - q$$

- Tarkastellaan aluksi rajoitettua SU–estimaattoria $\tilde{\theta}$, joka maksimoi uskottavuusfunktion ehdolla $A\theta = c$ ja toteuttaa siten $A\tilde{\theta} = c$.

- Lineaarialgebraa käyttäen voidaan todeta, että

$$\underset{q\times d}{A}\theta_0 = \underset{q\times 1}{c}, \ \ \mathrm{r}\left(A\right) = q \quad \Leftrightarrow \quad \theta_0 = \underset{d\times(d-q)}{B}\delta_0 + \underset{d\times 1}{e}, \ \ \mathrm{r}\left(B\right) = d-q$$

- Lisäksi, $AB = 0$ ja $\delta_0 = \left(B'B\right)^{-1}B'\left(\theta_0 - e\right)$, joten parametri δ saa arvoja joukossa $\Delta = \{\delta : \delta = \left(B'B\right)^{-1}B'\left(\theta - e\right), \theta \in \Theta\}$.

- Tarkastellaan aluksi rajoitettua SU–estimaattoria $\tilde{\theta}$, joka maksimoi uskottavuusfunktion ehdolla $A\theta = c$ ja toteuttaa siten $A\tilde{\theta} = c$.

- Lineaarialgebraa käyttäen voidaan todeta, että

$$\underset{q\times d}{A}\theta_0 = \underset{q\times 1}{c}, \;\; \mathrm{r}\left(A\right) = q \quad \Leftrightarrow \quad \theta_0 = \underset{d\times(d-q)}{B}\delta_0 + \underset{d\times 1}{e}, \;\; \mathrm{r}\left(B\right) = d-q$$

- Lisäksi, $AB = 0$ ja $\delta_0 = \left(B'B\right)^{-1} B'\left(\theta_0 - e\right)$, joten parametri δ saa arvoja joukossa $\Delta = \{\delta : \delta = \left(B'B\right)^{-1} B'\left(\theta - e\right), \theta \in \Theta\}$.

- Koska $\theta = B\delta + e$, on rajoitettu SU–estimaatti $\tilde{\theta} = B\tilde{\delta} + e$, jossa $\tilde{\delta}$ saadaan ratkaisemalla maksimointitehtävå

$$L^{(r)}(\tilde{\delta}; \mathbf{y}) = \max_{\delta\in\Delta} L^{(r)}\left(\delta; \mathbf{y}\right) = \max_{\delta\in\Delta} L\left(B\delta + e; \mathbf{y}\right).$$

- Raon testi perustuu pistemäärään $s(\tilde{\theta}; \mathbf{y}) = \partial l(\tilde{\theta}; \mathbf{y})/\partial\theta$, jossa rajoitettu SU–estimaattori $\tilde{\theta} = B\tilde{\delta} + e$ siis maksimoi uskottavuusfunktion ehdolla $A\theta = c$ ja toteuttaa siten $A\tilde{\theta} = c$.

- Raon testi perustuu pistemäärään $s(\tilde{\theta}; \mathbf{y}) = \partial l(\tilde{\theta}; \mathbf{y}) / \partial \theta$, jossa rajoitettu SU–estimaattori $\tilde{\theta} = B\tilde{\delta} + e$ siis maksimoi uskottavuusfunktion ehdolla $A\theta = c$ ja toteuttaa siten $A\tilde{\theta} = c$.

- Aikaisemmin esitettyjä tuloksia voidaan soveltaa myös estimaattoriin $\tilde{\delta}$ ja todeta erityisesti sen tarkentuvuus ja asymptoottinen normaalisuus.

- Raon testi perustuu pistemäärään $s(\tilde{\theta};\mathbf{y}) = \partial l(\tilde{\theta};\mathbf{y})/\partial\theta$, jossa rajoitettu SU–estimaattori $\tilde{\theta} = B\tilde{\delta} + e$ siis maksimoi uskottavuusfunktion ehdolla $A\theta = c$ ja toteuttaa siten $A\tilde{\theta} = c$.

- Aikaisemmin esitettyjä tuloksia voidaan soveltaa myös estimaattoriin $\tilde{\delta}$ ja todeta erityisesti sen tarkentuvuus ja asymptoottinen normaalisuus.

- Rajoitetun SU–estimaattorin $\tilde{\theta}$ vastaavat ominaisuudet voidaan johtaa tästä käyttäen yhtälöä $\tilde{\theta} = B\tilde{\delta} + e$.

- Raon testi perustuu pistemäärään $s(\tilde{\theta};\mathbf{y}) = \partial l(\tilde{\theta};\mathbf{y})/\partial\theta$, jossa rajoitettu SU–estimaattori $\tilde{\theta} = B\tilde{\delta} + e$ siis maksimoi uskottavuusfunktion ehdolla $A\theta = c$ ja toteuttaa siten $A\tilde{\theta} = c$.

- Aikaisemmin esitettyjä tuloksia voidaan soveltaa myös estimaattoriin $\tilde{\delta}$ ja todeta erityisesti sen tarkentuvuus ja asymptoottinen normaalisuus.

- Rajoitetun SU–estimaattorin $\tilde{\theta}$ vastaavat ominaisuudet voidaan johtaa tästä käyttäen yhtälöä $\tilde{\theta} = B\tilde{\delta} + e$.

- Tarkentuvuus, johon Raon testissä vedotaan, saadaan myös suoraan Lauseesta 2.2, kun parametriavaruus määritellään uudelleen korvaamalla Θ joukolla $\{\theta \in \Theta : A\theta = c\}$.

- Raon testin idea on tutkia poikkeaako $s(\tilde{\theta};\mathbf{y}) = \partial l(\tilde{\theta};\mathbf{y})/\partial\theta$ "liikaa" nollasta.

- Raon testin idea on tutkia poikkeaako $s(\tilde{\theta}; \mathbf{y}) = \partial l(\tilde{\theta}; \mathbf{y}) / \partial \theta$ "liikaa" nollasta.

- Tarkentuvuus $\Rightarrow$ H_0:n voimassa ollessa vapaa SUE $\hat{\theta}$ ja rajoitettu SUE $\tilde{\theta}$ toteuttavat $\hat{\theta} \approx \theta_0 \approx \tilde{\theta}$.

- Raon testin idea on tutkia poikkeaako $s(\tilde{\theta}; \mathbf{y}) = \partial l(\tilde{\theta}; \mathbf{y}) / \partial\theta$ "liikaa" nollasta.

- Tarkentuvuus $\Rightarrow$ H_0:n voimassa ollessa vapaa SUE $\hat{\theta}$ ja rajoitettu SUE $\tilde{\theta}$ toteuttavat $\hat{\theta} \approx \theta_0 \approx \tilde{\theta}$.

- Koska $\hat{\theta}$ toteuttaa uskottavuusyhtälöt eli $s(\hat{\theta}; \mathbf{y}) = 0$, on H_0:n voimassa ollessa odotettavaa, että $s(\tilde{\theta}; \mathbf{y}) \approx 0$.

- Raon testin idea on tutkia poikkeaako $s(\tilde{\theta}; \mathbf{y}) = \partial l(\tilde{\theta}; \mathbf{y}) / \partial \theta$ "liikaa" nollasta.
- Tarkentuvuus $\Rightarrow$ H_0:n voimassa ollessa vapaa SUE $\hat{\theta}$ ja rajoitettu SUE $\tilde{\theta}$ toteuttavat $\hat{\theta} \approx \theta_0 \approx \tilde{\theta}$.
- Koska $\hat{\theta}$ toteuttaa uskottavuusyhtälöt eli $s(\hat{\theta}; \mathbf{y}) = 0$, on H_0:n voimassa ollessa odotettavaa, että $s(\tilde{\theta}; \mathbf{y}) \approx 0$.
- Jos H_0 ei ole voimassa, ei ole mitään syytä miksi näin kävisi, joten pistemäärän $s(\tilde{\theta}; \mathbf{y})$ "suurien" arvojen voidaan tulkita viittaavan nollahypoteesin virheellisyyteen.

- Lausetta 2.3 perusteltaessa käytettiin kehitelmää

$$s(\hat{\theta};\mathbf{Y}) = s\left(\theta_0;\mathbf{Y}\right) - \bar{\mathcal{J}}\left(\bar{\theta};\mathbf{Y}\right)\left(\hat{\theta}-\theta_0\right), \qquad (*)$$

jossa

$$\bar{\mathcal{J}}\left(\bar{\theta};\mathbf{Y}\right) \text{:n } a. \text{ rivi} = \mathcal{J}\left(\bar{\theta}_a;\mathbf{Y}\right) \text{:n } a. \text{ rivi} \quad \text{ja}$$

$$||\bar{\theta}_a-\theta_0|| \quad \leq \quad ||\hat{\theta}-\theta_0||, \; a = 1, ..., d.$$

- Lausetta 2.3 perusteltaessa käytettiin kehitelmää

$$s(\hat{\theta};\mathbf{Y}) = s\left(\theta_0;\mathbf{Y}\right) - \bar{\mathcal{J}}\left(\bar{\theta};\mathbf{Y}\right)\left(\hat{\theta}-\theta_0\right), \qquad (*)$$

jossa

$$\bar{\mathcal{J}}\left(\bar{\theta};\mathbf{Y}\right) \text{:n } a. \text{ rivi} = \mathcal{J}\left(\bar{\theta}_a;\mathbf{Y}\right) \text{:n } a. \text{ rivi} \quad \text{ja}$$

$$||\bar{\theta}_a - \theta_0|| \quad \leq \quad ||\hat{\theta}-\theta_0||, \; a = 1, ..., d.$$

- Vastaava kehitelmä rajoitetun SUE:n $\tilde{\theta}$ tapauksessa on

$$s(\tilde{\theta};\mathbf{Y}) \quad = \quad s\left(\theta_0;\mathbf{Y}\right) - \bar{\mathcal{J}}\left(\bar{\theta};\mathbf{Y}\right)\left(\tilde{\theta}-\theta_0\right)$$

- Lausetta 2.3 perusteltaessa käytettiin kehitelmää

$$s(\hat{\theta};\mathbf{Y}) = s\left(\theta_0;\mathbf{Y}\right) - \bar{\mathcal{J}}\left(\bar{\theta};\mathbf{Y}\right)\left(\hat{\theta}-\theta_0\right), \quad (*)$$

jossa

$$\bar{\mathcal{J}}\left(\bar{\theta};\mathbf{Y}\right) \text{:n } a\text{. rivi} = \mathcal{J}\left(\bar{\theta}_a;\mathbf{Y}\right) \text{:n } a\text{. rivi ja}$$

$$||\bar{\theta}_a-\theta_0|| \quad \leq \quad ||\hat{\theta}-\theta_0||, \; a = 1, ..., d.$$

- Vastaava kehitelmä rajoitetun SUE:n $\tilde{\theta}$ tapauksessa on

$$s(\tilde{\theta};\mathbf{Y}) \quad = \quad s\left(\theta_0;\mathbf{Y}\right) - \bar{\mathcal{J}}\left(\bar{\theta};\mathbf{Y}\right)\left(\tilde{\theta}-\theta_0\right)$$

- tai, koska $\tilde{\theta}-\theta_0 = B(\tilde{\delta}-\delta_0)$,

$$s(\tilde{\theta};\mathbf{Y}) \quad = \quad s\left(\theta_0;\mathbf{Y}\right) - \bar{\mathcal{J}}\left(\bar{\theta};\mathbf{Y}\right) B(\tilde{\delta}-\delta_0)$$

- Raon testi perustuu testisuureeseen

$$\mathsf{S} = s(\tilde{\theta};\mathbf{Y})'\mathcal{J}(\tilde{\theta};\mathbf{Y})^{-1}A'\left[A\mathcal{J}(\tilde{\theta};\mathbf{Y})^{-1}A'\right]^{-1}A\mathcal{J}(\tilde{\theta};\mathbf{Y})^{-1}s(\tilde{\theta};\mathbf{Y}) \underset{\mathsf{H}_0}{\overset{d}{\longrightarrow}} \chi^2_q,$$

jolla on H_0:n voimassa ollessa sama asymptoottinen jakauma kuin Waldin testisuurella.

- Raon testi perustuu testisuureeseen

$$\mathrm{S} = s(\tilde{\theta};\mathbf{Y})'\mathcal{J}(\tilde{\theta};\mathbf{Y})^{-1}A'\left[A\mathcal{J}(\tilde{\theta};\mathbf{Y})^{-1}A'\right]^{-1}A\mathcal{J}(\tilde{\theta};\mathbf{Y})^{-1}s(\tilde{\theta};\mathbf{Y}) \underset{\mathrm{H}_0}{\xrightarrow{d}} \chi^2_q,$$

  jolla on H_0:n voimassa ollessa sama asymptoottinen jakauma kuin Waldin testisuurella.

- Havaitun informaatiomatriisin $\mathcal{J}(\tilde{\theta};\mathbf{Y})$ paikalla voidaan vaihtoehtoisesti käyttää ulkotulomatriisia $n\mathrm{M}(\tilde{\theta};\mathbf{Y})$.

- Raon testi perustuu testisuureeseen

$$\mathsf{S} = s(\tilde{\theta};\mathbf{Y})'\mathcal{J}(\tilde{\theta};\mathbf{Y})^{-1}A'\left[A\mathcal{J}(\tilde{\theta};\mathbf{Y})^{-1}A'\right]^{-1}A\mathcal{J}(\tilde{\theta};\mathbf{Y})^{-1}s(\tilde{\theta};\mathbf{Y}) \underset{\mathrm{H}_0}{\xrightarrow{d}} \chi^2_q,$$

  jolla on H_0:n voimassa ollessa sama asymptoottinen jakauma kuin Waldin testisuurella.

- Havaitun informaatiomatriisin $\mathcal{J}(\tilde{\theta};\mathbf{Y})$ paikalla voidaan vaihtoehtoisesti käyttää ulkotulomatriisia $n\mathrm{M}(\tilde{\theta};\mathbf{Y})$.

- Approksimatiiviset P–arvot

$$\mathrm{P} = \mathsf{P}_{\mathrm{H}_0}\left\{\mathsf{S} \geq \mathsf{S}(\mathbf{y})\right\} \approx \mathsf{P}\left\{\chi^2_q \geq \mathsf{S}(\mathbf{y})\right\}.$$

- Käyttäen matriisilaskentaa voidaan testisuure

$$\mathrm{S} = s(\tilde{\theta};\mathbf{Y})'\mathcal{J}(\tilde{\theta};\mathbf{Y})^{-1}A'\left[A\mathcal{J}(\tilde{\theta};\mathbf{Y})^{-1}A'\right]^{-1}A\mathcal{J}(\tilde{\theta};\mathbf{Y})^{-1}s(\tilde{\theta};\mathbf{Y})$$

sieventää yksinkertaiseen muotoon

$$\mathrm{S} = s(\tilde{\theta};\mathbf{Y})'\mathcal{J}(\tilde{\theta};\mathbf{Y})^{-1}s(\tilde{\theta};\mathbf{Y}) \xrightarrow[\mathrm{H}_0]{d} \chi^2_q.$$

- Kun $A = [I_q : 0]$, ovat pistemäärän $s(\tilde{\theta};\mathbf{Y})$ viimeiset $d-q$ komponenttia nollia ja testisuureen lauseke supistuu tilastollisen päättelyn kurssilla esitettyyn muotoon.

- Joissakin tapauksissa testisuureen lauseketta voidaan yksinkertaistaa myös käyttämällä havaitun informaatiomatriisin $\mathcal{J}(\tilde{\theta};\mathbf{Y})$ paikalla ulkotulomatriisia $\mathrm{M}(\tilde{\theta};\mathbf{Y})$.

- Rajoitettu SU–estimointi ehdolla $h(\theta) = 0$ Lagrangen kerroinmenettelyä käyttäen: Maksimoidaan funktio

$$l(\theta, \kappa; \mathbf{y}) = l(\theta; \mathbf{y}) + \kappa' h(\theta),$$

jossa vektori $\kappa = \begin{bmatrix} \kappa_1 & \cdots & \kappa_q \end{bmatrix}'$ sisältää Lagrangen kertoimet.

- Rajoitettu SU–estimointi ehdolla $h(\theta) = 0$ Lagrangen kerroinmenettelyä käyttäen: Maksimoidaan funktio

$$l(\theta, \kappa; \mathbf{y}) = l(\theta; \mathbf{y}) + \kappa' h(\theta),$$

jossa vektori $\kappa = \begin{bmatrix} \kappa_1 & \cdots & \kappa_q \end{bmatrix}'$ sisältää Lagrangen kertoimet.

- Derivointi θ:n suhteen ja gradientti $= 0 \Rightarrow$ yhtälö

$$\frac{\partial}{\partial \theta} l(\theta, \kappa; \mathbf{y}) = s(\theta; \mathbf{y}) + H(\theta)' \kappa = 0$$

- Rajoitettu SU–estimointi ehdolla $h(\theta) = 0$ Lagrangen kerroinmenettelyä käyttäen: Maksimoidaan funktio

$$l(\theta, \kappa; \mathbf{y}) = l(\theta; \mathbf{y}) + \kappa' h(\theta),$$

jossa vektori $\kappa = \begin{bmatrix} \kappa_1 & \cdots & \kappa_q \end{bmatrix}'$ sisältää Lagrangen kertoimet.

- Derivointi θ:n suhteen ja gradientti $= 0 \Rightarrow$ yhtälö

$$\frac{\partial}{\partial \theta} l(\theta, \kappa; \mathbf{y}) = s(\theta; \mathbf{y}) + H(\theta)' \kappa = 0$$

- Sijoittamalla $\theta = \tilde{\theta}$ päädytään κ:n suhteen ratkaisuun

$$\tilde{\kappa} = -[H(\tilde{\theta}) \mathcal{J}(\tilde{\theta}; \mathbf{y})^{-1} H(\tilde{\theta})']^{-1} H(\tilde{\theta}) \mathcal{J}(\tilde{\theta}; \mathbf{y})^{-1} s(\tilde{\theta}; \mathbf{y}).$$

- Rajoitettu SU–estimointi ehdolla $h(\theta) = 0$ Lagrangen kerroinmenettelyä käyttäen: Maksimoidaan funktio

$$l(\theta, \kappa; \mathbf{y}) = l(\theta; \mathbf{y}) + \kappa' h(\theta),$$

jossa vektori $\kappa = \begin{bmatrix} \kappa_1 & \cdots & \kappa_q \end{bmatrix}'$ sisältää Lagrangen kertoimet.

- Derivointi θ:n suhteen ja gradientti $= 0 \Rightarrow$ yhtälö

$$\frac{\partial}{\partial \theta} l(\theta, \kappa; \mathbf{y}) = s(\theta; \mathbf{y}) + H(\theta)' \kappa = 0$$

- Sijoittamalla $\theta = \tilde{\theta}$ päädytään κ:n suhteen ratkaisuun

$$\tilde{\kappa} = -[H(\tilde{\theta})\mathcal{J}(\tilde{\theta}; \mathbf{y})^{-1} H(\tilde{\theta})']^{-1} H(\tilde{\theta})\mathcal{J}(\tilde{\theta}; \mathbf{y})^{-1} s(\tilde{\theta}; \mathbf{y}).$$

- Nähdään, että

$$\mathsf{S} = \tilde{\kappa}' H(\tilde{\theta})\mathcal{J}(\tilde{\theta}; \mathbf{Y})^{-1} H(\tilde{\theta})' \tilde{\kappa}.$$

- Ositetaan $\theta = (\psi, \lambda)$, jossa parametrit ψ ja λ ovat ortogonaaliset ja ψ on kiinnostava parametri.

- Ositetaan $\theta = (\psi, \lambda)$, jossa parametrit ψ ja λ ovat ortogonaaliset ja ψ on kiinnostava parametri.
- Testattava hypoteesi $\mathrm{H}_0 : A_\psi \psi = c \Leftrightarrow A\theta = c$, jossa $A = \left[A_\psi : 0 \right]$.

- Ositetaan $\theta = (\psi, \lambda)$, jossa parametrit ψ ja λ ovat ortogonaaliset ja ψ on kiinnostava parametri.
- Testattava hypoteesi $\mathrm{H}_0 : A_\psi \psi = c \Leftrightarrow A\theta = c$, jossa $A = \left[A_\psi : 0\right]$.
- Nyt $\bar{\mathcal{I}}(\theta) = \mathrm{diag}[\bar{\mathcal{I}}_{\psi\psi}(\theta) \;\; \bar{\mathcal{I}}_{\lambda\lambda}(\theta)]$ ja sekä

$$n^{-1}\mathcal{J}(\tilde{\theta}; \mathbf{Y}) \quad \text{että} \quad n^{-1}\mathcal{J}_o(\tilde{\theta}; \mathbf{Y}) = n^{-1}\mathrm{diag}[\mathcal{J}_{\psi\psi}(\tilde{\theta}; \mathbf{Y}) \;\; \mathcal{J}_{\lambda\lambda}(\tilde{\theta}; \mathbf{Y})]$$

estimoivat tarkentuvasti matriisia $\bar{\mathcal{I}}(\theta)$.

- Ositetaan $\theta = (\psi, \lambda)$, jossa parametrit ψ ja λ ovat ortogonaaliset ja ψ on kiinnostava parametri.
- Testattava hypoteesi $H_0 : A_\psi \psi = c \Leftrightarrow A\theta = c$, jossa $A = \left[A_\psi : 0 \right]$.
- Nyt $\bar{\mathcal{I}}(\theta) = \mathrm{diag}[\bar{\mathcal{I}}_{\psi\psi}(\theta) \;\; \bar{\mathcal{I}}_{\lambda\lambda}(\theta)]$ ja sekä

$$n^{-1}\mathcal{J}(\tilde{\theta};\mathbf{Y}) \quad \text{että} \quad n^{-1}\mathcal{J}_o(\tilde{\theta};\mathbf{Y}) = n^{-1}\mathrm{diag}[\mathcal{J}_{\psi\psi}(\tilde{\theta};\mathbf{Y}) \;\; \mathcal{J}_{\lambda\lambda}(\tilde{\theta};\mathbf{Y})]$$

  estimoivat tarkentuvasti matriisia $\bar{\mathcal{I}}(\theta)$.
- Raon testissä voidaan havaittu informaatiomatriisi $\mathcal{J}(\tilde{\theta};\mathbf{Y})$ korvata siten matriisilla $\mathcal{J}_o(\tilde{\theta};\mathbf{Y})$.

$$\mathsf{S} = s(\tilde{\theta};\mathbf{Y})'\mathcal{J}_o(\tilde{\theta};\mathbf{Y})^{-1}A'\left[A\mathcal{J}_o(\tilde{\theta};\mathbf{Y})^{-1}A'\right]^{-1}A\mathcal{J}_o(\tilde{\theta};\mathbf{Y})^{-1}s(\tilde{\theta};\mathbf{Y})$$

jossa $A = \left[A_\psi : 0\right]$, $s(\tilde{\theta};\mathbf{Y}) = (s_\psi(\tilde{\theta};\mathbf{Y}), s_\lambda(\tilde{\theta};\mathbf{Y}))$ ja

$$\mathcal{J}_o(\tilde{\theta};\mathbf{Y}) = \mathrm{diag}[\mathcal{J}_{\psi\psi}(\tilde{\theta};\mathbf{Y}) \;\; \mathcal{J}_{\lambda\lambda}(\tilde{\theta};\mathbf{Y})]$$

Hieman matriisilaskentaa $\Rightarrow$ testisuure

$$\mathsf{S}_\psi = s_\psi(\tilde{\theta};\mathbf{Y})'\mathcal{J}_{\psi\psi}^{-1}(\tilde{\theta};\mathbf{Y})A_\psi'\left[A_\psi\mathcal{J}_{\psi\psi}^{-1}(\tilde{\theta};\mathbf{Y})A_\psi'\right]^{-1}A_\psi\mathcal{J}_{\psi\psi}^{-1}(\tilde{\theta};\mathbf{Y})s_\psi(\tilde{\theta};\mathbf{Y}$$

tai vaihtoehtoisesti

$$\mathsf{S}_\psi = s_\psi(\tilde{\theta};\mathbf{Y})'\mathcal{J}_{\psi\psi}^{-1}(\tilde{\theta};\mathbf{Y})s_\psi(\tilde{\theta};\mathbf{Y}) \xrightarrow[\mathsf{H}_0]{d} \chi_q^2.$$

- Uskottavuusosamäärätesti perustuu testisuureeseen

$$\mathrm{LR} = 2\left[l(\hat{\theta};\mathbf{Y}) - l(\tilde{\theta};\mathbf{Y})\right].$$

- Uskottavuusosamäärätesti perustuu testisuureeseen

$$\mathrm{LR} = 2\left[l(\hat{\theta};\mathbf{Y}) - l(\tilde{\theta};\mathbf{Y})\right].$$

- Intuitiivisesti selvää, että suuret LR:n arvot todistavat H_0:aa vastaan.

- Uskottavuusosamäärätesti perustuu testisuureeseen

$$\mathrm{LR} = 2\left[l(\hat{\theta}; \mathbf{Y}) - l(\tilde{\theta}; \mathbf{Y})\right].$$

- Intuitiivisesti selvää, että suuret LR:n arvot todistavat H_0:aa vastaan.
- LR:n asymptoottinen jakauma voidaan johtaa käyttäen toisen asteen Taylorin kehitelmää.

- Uskottavuusosamäärätesti perustuu testisuureeseen

$$\mathrm{LR} = 2\left[l(\hat{\theta}; \mathbf{Y}) - l(\tilde{\theta}; \mathbf{Y})\right].$$

- Intuitiivisesti selvää, että suuret LR:n arvot todistavat H_0:aa vastaan.
- LR:n asymptoottinen jakauma voidaan johtaa käyttäen toisen asteen Taylorin kehitelmää.
- Kun $\mathrm{H}_0 : \theta_0 = c \ (d \times 1)$, saadaan $\cdots \ \cdots$

- Uskottavuusosamäärätesti perustuu testisuureeseen

$$\mathrm{LR} = 2\left[l(\hat{\theta};\mathbf{Y}) - l(\tilde{\theta};\mathbf{Y})\right].$$

- Intuitiivisesti selvää, että suuret LR:n arvot todistavat H_0:aa vastaan.
- LR:n asymptoottinen jakauma voidaan johtaa käyttäen toisen asteen Taylorin kehitelmää.
- Kun $\mathrm{H}_0 : \theta_0 = c \ (d \times 1)$, saadaan $\cdots \ \cdots$
-

$$\mathrm{LR} \overset{as}{=} (\hat{\theta} - c)'\mathcal{J}(\hat{\theta};\mathbf{Y})(\hat{\theta} - c) \xrightarrow[\mathrm{H}_0]{d} \chi^2_d$$

- Uskottavuusosamäärätesti perustuu testisuureeseen
$$\mathrm{LR} = 2\left[l(\hat{\theta};\mathbf{Y}) - l(\tilde{\theta};\mathbf{Y})\right].$$
- Intuitiivisesti selvää, että suuret LR:n arvot todistavat H_0:aa vastaan.
- LR:n asymptoottinen jakauma voidaan johtaa käyttäen toisen asteen Taylorin kehitelmää.
- Kun $\mathrm{H}_0 : \theta_0 = c \ (d \times 1)$, saadaan $\cdots\ \cdots$
- $$\mathrm{LR} \overset{as}{=} (\hat{\theta} - c)'\mathcal{J}(\hat{\theta};\mathbf{Y})(\hat{\theta} - c) \xrightarrow[\mathrm{H}_0]{d} \chi^2_d$$
- Yleiseinen tapaus algebraalisesti mutkikkaampi, mutta tulokseksi tulee
$$\mathrm{LR} \xrightarrow[\mathrm{H}_0]{d} \chi^2_q \quad (\mathrm{H}_0\!:\ A\theta_0 = c, \quad A \quad q \times d).$$

- Waldin testi perustuu vapaaseen SU–estimaattoriin ja Raon testi rajoitettuun SU–estimaattoriin.
- Waldin testi on siten usein kätevä tai luonteva, kun rajoittamaton malli on yksinkertainen ja sen parametrit on helppo estimoida.
  Raon testillä on puolestaan vastaava ominaisuus, kun rajoitettu malli on yksinkertainen ja sen parametrien estimointi on helppoa.
- Uskottavuusosamäärätesti on ”työläämpi” sikäli, että se vaatii sekä rajoitetun että rajoittamattoman SU–estimoinnin.
  Toisaalta, koska SU–estimaattori on tunnetusti invariantti vaihtoehtoisille parametroinneille, on uskottavuusosamäärätestillä vastaava invarianssiominaisuus, jota Waldin testillä ja Raon testillä ei yleisesti ole.

- Aineisto $\mathbf{w} = (w_1, ..., w_n)$, $w_i = (y_i, x_i)$, vastaava sv $\mathbf{W} = (W_1, ..., W_n)$, $W_i = (Y_i, X_i)$

- Aineisto $\mathbf{w} = (w_1, ..., w_n)$, $w_i = (y_i, x_i)$, vastaava sv $\mathbf{W} = (W_1, ..., W_n)$, $W_i = (Y_i, X_i)$
- Oletetaan, että analyysi voidaan perustaa ehdolliseen malliin

$$Y_i|(\mathbf{W}_{i-1}, X_i) \overset{d}{=} Y_i|Z_i \quad \text{ja} \quad Y_i|(Z_i = z_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right),$$

jossa $\mathbf{W}_i = (\mathbf{W}_1, ..., \mathbf{W}_i)$ ja Z_i $(p \times 1)$ sisältää $(\mathbf{W}_{i-1}, X_i)$:n "relevantit" komponentit eli joko

- Aineisto $\mathbf{w} = (w_1, ..., w_n)$, $w_i = (y_i, x_i)$, vastaava sv $\mathbf{W} = (W_1, ..., W_n)$, $W_i = (Y_i, X_i)$
- Oletetaan, että analyysi voidaan perustaa ehdolliseen malliin

$$Y_i|(\mathbf{W}_{i-1}, X_i) \stackrel{d}{=} Y_i|Z_i \quad \text{ja} \quad Y_i|(Z_i = z_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right),$$

jossa $\mathbf{W}_i = (\mathbf{W}_1, ..., \mathbf{W}_i)$ ja Z_i $(p \times 1)$ sisältää $(\mathbf{W}_{i-1}, X_i)$:n "relevantit" komponentit eli joko

  - $Z_i = (1, X_i)$ eli "tavanomaiseen" regressiomallia tai

- Aineisto $\mathbf{w} = (w_1, ..., w_n)$, $w_i = (y_i, x_i)$, vastaava sv $\mathbf{W} = (W_1, ..., W_n)$, $W_i = (Y_i, X_i)$
- Oletetaan, että analyysi voidaan perustaa ehdolliseen malliin

$$Y_i|(\mathbf{W}_{i-1}, X_i) \stackrel{d}{=} Y_i|Z_i \quad \text{ja} \quad Y_i|(Z_i = z_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right),$$

jossa $\mathbf{W}_i = (\mathbf{W}_1, ..., \mathbf{W}_i)$ ja Z_i $(p \times 1)$ sisältää $(\mathbf{W}_{i-1}, X_i)$:n "relevantit" komponentit eli joko

  - $Z_i = (1, X_i)$ eli "tavanomaiseen" regressiomallia tai
  - $Z_i = (1, Y_{i-1}, X_i)$ eli autoregressiivieen aikasarjamalli ulkopuolisin selittävin muuttujin ($Y_0 = y_0$ tunnettu vakio).

- Aineisto $\mathbf{w} = (w_1, ..., w_n)$, $w_i = (y_i, x_i)$, vastaava sv $\mathbf{W} = (W_1, ..., W_n)$, $W_i = (Y_i, X_i)$
- Oletetaan, että analyysi voidaan perustaa ehdolliseen malliin

$$Y_i | (\mathbf{W}_{i-1}, X_i) \overset{d}{=} Y_i | Z_i \quad \text{ja} \quad Y_i | (Z_i = z_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right),$$

jossa $\mathbf{W}_i = (\mathbf{W}_1, ..., \mathbf{W}_i)$ ja Z_i $(p \times 1)$ sisältää $(\mathbf{W}_{i-1}, X_i)$:n "relevantit" komponentit eli joko

- $Z_i = (1, X_i)$ eli "tavanomaiseen" regressiomallia tai
- $Z_i = (1, Y_{i-1}, X_i)$ eli autoregressiivieen aikasarjamalli ulkopuolisin selittävin muuttujin ($Y_0 = y_0$ tunnettu vakio).

- Vaihtoehtoinen muotoilu

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.

- Ehd. log-uskottavuusfunktio

$$l^{(c)}\left(\beta,\sigma^2;\mathbf{w}\right) = -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - z_i'\beta\right)^2, \ \beta\in\mathbb{R}^p,\ \sigma^2>0.$$

- Ehd. log-uskottavuusfunktio

$$l^{(c)}\left(\beta,\sigma^2;\mathbf{w}\right) = -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - z_i'\beta\right)^2,\ \beta\in\mathbb{R}^p,\ \sigma^2>0.$$

- $H_0 : h\left(\beta\right) = 0,\ \ h:\mathbb{R}^p \to \mathbb{R}^q$ jatkuvasti derivoituva ja $H\left(\beta\right) = \left[\partial h_a\left(\beta\right)/\partial\beta_b\right]$, $(q\times p)$ toteuttaa $\mathrm{r}\left(H\left(\beta\right)\right) = q$.

- Ehd. log-uskottavuusfunktio

$$l^{(c)}\left(\beta,\sigma^2;\mathbf{w}\right) = -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - z_i'\beta\right)^2, \ \beta\in\mathbb{R}^p,\ \sigma^2>0.$$

- $\mathrm{H}_0 : h\left(\beta\right) = 0, \ \ h:\mathbb{R}^p\to\mathbb{R}^q$ jatkuvasti derivoituva ja $H\left(\beta\right) = \left[\partial h_a\left(\beta\right)/\partial\beta_b\right]$, $(q\times p)$ toteuttaa $\mathrm{r}\left(H\left(\beta\right)\right) = q$.
- Konkreettinen esimerkki $\mathrm{H}_0 : \beta_3 = \beta_1\beta_2$.

- Ehd. log-uskottavuusfunktio

$$l^{(c)}\left(\beta,\sigma^2;\mathbf{w}\right) = -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - z_i'\beta\right)^2, \ \beta\in\mathbb{R}^p,\ \sigma^2>0.$$

- $\mathrm{H}_0 : h\left(\beta\right) = 0,\ \ h:\mathbb{R}^p \to \mathbb{R}^q$ jatkuvasti derivoituva ja $H\left(\beta\right) = \left[\partial h_a\left(\beta\right)/\partial\beta_b\right]$, $(q\times p)$ toteuttaa $\mathrm{r}\left(H\left(\beta\right)\right) = q$.
- Konkreettinen esimerkki $\mathrm{H}_0 : \beta_3 = \beta_1\beta_2$.
- $\mathrm{H}_0 \Rightarrow$ malli on lineaarinen ja SU–estimointi hoituu helposti PNS:llä. Siten Waldin testi on kätevä.

- Ehd. log-uskottavuusfunktio

$$l^{(c)}\left(\beta,\sigma^2;\mathbf{w}\right)=-\frac{n}{2}\log\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i-z_i'\beta\right)^2,\ \ \beta\in\mathbb{R}^p,\ \sigma^2>0.$$

- $\mathrm{H}_0 : h(\beta)=0,\ \ h:\mathbb{R}^p\to\mathbb{R}^q$ jatkuvasti derivoituva ja $H(\beta)=\left[\partial h_a(\beta)/\partial\beta_b\right]$, $(q\times p)$ toteuttaa $\mathrm{r}\left(H(\beta)\right)=q$.
- Konkreettinen esimerkki $\mathrm{H}_0 : \beta_3=\beta_1\beta_2$.
- $\mathrm{H}_0 \Rightarrow$ malli on lineaarinen ja SU–estimointi hoituu helposti PNS:llä. Siten Waldin testi on kätevä.
- Parametrit β ja σ^2 on todettu ortogonaalisiksi, joten Waldin testiä varten tarvitaan vapaan mallin SUE:t $\hat{\beta}$ ja $\hat{\sigma}^2$ sekä β:n havaittu informaatiomatriisiriisi

$$\mathcal{J}_{\beta\beta}(\beta,\sigma^2;\mathbf{W}):=\frac{1}{\sigma^2}\sum_{i=1}^{n}Z_iZ_i'=:\frac{1}{\sigma^2}Q_n.$$

- Yleinen Waldin testisuure epälin. hypoteesin tapauksessa:

$$\mathrm{W} = h(\hat{\theta})'[H(\hat{\theta})\mathcal{J}(\hat{\theta};\mathbf{Y})^{-1}H(\hat{\theta})']^{-1}h(\hat{\theta}) \xrightarrow[\mathrm{H}_0]{d} \chi^2_q,$$

- Yleinen Waldin testisuure epälin. hypoteesin tapauksessa:

$$\mathrm{W} = h(\hat{\theta})'[H(\hat{\theta})\mathcal{J}(\hat{\theta};\mathbf{Y})^{-1}H(\hat{\theta})']^{-1}h(\hat{\theta}) \xrightarrow[\mathrm{H}_0]{d} \chi^2_q,$$

- Yleinen Waldin testisuure lin. hypoteesin tapauksessa, kun parametrin $\theta = (\psi, \lambda)$ komponentit ortogonaaliset:

$$\mathrm{W}_\psi = (A_\psi\hat{\psi} - c)'[A_\psi\mathcal{J}^{-1}_{\psi\psi}(\hat{\theta};\mathbf{Y})A'_\psi]^{-1}(A_\psi\hat{\psi} - c) \xrightarrow[\mathrm{H}_0]{d} \chi^2_q.$$

Epälin. hypoteesin $\mathrm{H}_0 : h(\psi) = 0$ tapauksessa vastaava "sijoituksilla"

$A_\psi\hat{\psi} - c \to h(\hat{\psi})$ ja $A_\psi \to H(\hat{\psi}) = [\partial h_a(\psi)/\partial\psi_b]$.

- Yleinen Waldin testisuure epälin. hypoteesin tapauksessa:

$$\mathrm{W} = h(\hat{\theta})'[H(\hat{\theta})\mathcal{J}(\hat{\theta};\mathbf{Y})^{-1}H(\hat{\theta})']^{-1}h(\hat{\theta}) \xrightarrow[\mathrm{H}_0]{d} \chi^2_q,$$

- Yleinen Waldin testisuure lin. hypoteesin tapauksessa, kun parametrin $\theta = (\psi, \lambda)$ komponentit ortogonaaliset:

$$\mathrm{W}_\psi = (A_\psi\hat{\psi} - c)'[A_\psi\mathcal{J}^{-1}_{\psi\psi}(\hat{\theta};\mathbf{Y})A'_\psi]^{-1}(A_\psi\hat{\psi} - c) \xrightarrow[\mathrm{H}_0]{d} \chi^2_q.$$

Epälin. hypoteesin $\mathrm{H}_0 : h(\psi) = 0$ tapauksessa vastaava "sijoituksilla"

$$A_\psi\hat{\psi} - c \longrightarrow h\left(\hat{\psi}\right) \quad \text{ja} \quad A_\psi \longrightarrow H\left(\hat{\psi}\right) = \left[\partial h_a(\psi) / \partial\psi_b\right].$$

- Nyt $\mathcal{J}_{\psi\psi}(\hat{\theta};\mathbf{Y})$:n paikalla on $\frac{1}{\hat{\sigma}^2}\sum_{i=1}^n Z_i Z_i' =: \frac{1}{\hat{\sigma}^2}Q_n$, joten saadaan

$$\mathrm{W}_\beta = \frac{1}{\hat{\sigma}^2}h(\hat{\beta})'\left[H(\hat{\beta})Q_n^{-1}H(\hat{\beta})'\right]^{-1}h(\hat{\beta}) \xrightarrow[\mathrm{H}_0]{d} \chi^2_q.$$

- Lineaarisessa mallissa on aikaisemmin oletettu

$$Y_i|(\mathbf{W}_{i-1}, X_i) \stackrel{d}{=} Y_i|Z_i \quad \text{ja} \quad Y_i|(Z_i = z_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right),$$

- Lineaarisessa mallissa on aikaisemmin oletettu

$$Y_i|(\mathbf{W}_{i-1}, X_i) \stackrel{d}{=} Y_i|Z_i \quad \text{ja} \quad Y_i|(Z_i = z_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right),$$

- tai

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.

- Lineaarisessa mallissa on aikaisemmin oletettu

$$Y_i|(\mathbf{W}_{i-1}, X_i) \overset{d}{=} Y_i|Z_i \quad \text{ja} \quad Y_i|(Z_i = z_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right),$$

- tai

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.

- Ehdollisten varianssien olettaminen vakioksi σ^2 voi olla kyseenalaista.

- Lineaarisessa mallissa on aikaisemmin oletettu

$$Y_i|(\mathbf{W}_{i-1}, X_i) \overset{d}{=} Y_i|Z_i \quad \text{ja} \quad Y_i|(Z_i = z_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right),$$

- tai

$$Y_i = Z_i'\beta + \varepsilon_i, \quad \varepsilon_1, ..., \varepsilon_n \sim \mathrm{N}\left(0, \sigma^2\right) \perp\!\!\!\perp,$$

jossa $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \varepsilon_i$ ja siten $Z_i \perp\!\!\!\perp \varepsilon_i$ kaikilla $i = 1, ..., n$.

- Ehdollisten varianssien olettaminen vakioksi σ^2 voi olla kyseenalaista.
- Tarkastellaan tämän homoskedastisuusoletuksen testaamista Raon pistemäärätestillä, joka on tässä kätevä, koska nollahypoteen olettaa homoskedastisuuden ja SU–estimointi hoituu PNS:llä.

- Testiä varten yleistetään malli

$$Y_i \big| (\mathbf{W}_{i-1}, X_i) \stackrel{d}{=} Y_i \big| Z_i \quad \text{ja} \quad Y_i \big| (Z_i = z_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right).$$

- Testiä varten yleistetään malli

$$Y_i|(\mathbf{W}_{i-1}, X_i) \overset{d}{=} Y_i|Z_i \quad \text{ja} \quad Y_i|(Z_i = z_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right).$$

- Oletetaan $Y_i|(\mathbf{W}_{i-1}, X_i) \overset{d}{=} Y_i|(Z_i, V_i)$ ja

$$Y_i|(Z_i = z_i, V_i = v_i) \sim \mathrm{N}\left(z_i'\beta, \sigma_i^2(\delta)\right), \quad \sigma_i^2(\delta) = \exp\left\{v_i'\delta\right\},$$

jossa v_i $(k \times 1)$ ensimmäinen komponentti on vakio 1 ja muiksi komponenteiksi valitaan vektorin $(\mathbf{w}_{i-1}, x_i)$ "relevantit" komponentit.

- Testiä varten yleistetään malli

$$Y_i|(\mathbf{W}_{i-1}, X_i) \stackrel{d}{=} Y_i|Z_i \quad \text{ja} \quad Y_i|(Z_i = z_i) \sim \mathrm{N}\left(z_i'\beta, \sigma^2\right).$$

- Oletetaan $Y_i|(\mathbf{W}_{i-1}, X_i) \stackrel{d}{=} Y_i|(Z_i, V_i)$ ja

$$Y_i|(Z_i = z_i, V_i = v_i) \sim \mathrm{N}\left(z_i'\beta, \sigma_i^2(\delta)\right), \quad \sigma_i^2(\delta) = \exp\left\{v_i'\delta\right\},$$

  jossa v_i $(k \times 1)$ ensimmäinen komponentti on vakio 1 ja muiksi komponenteiksi valitaan vektorin $(\mathbf{w}_{i-1}, x_i)$ "relevantit" komponentit.

- Parametrivektorille $\delta = (\delta_1, \delta_2, ..., \delta_k)$ asetetaan nollahypoteesi

$$\mathrm{H}_0 : \delta_2 = \cdots = \delta_k = 0 \quad \Leftrightarrow \quad \sigma_i^2(\delta) = e^{\delta_1} := \sigma^2,$$

  jonka voimassa ollessa homoskedastinen malli on riittävä.

- Oletetaan siis $Y_i|(\mathbf{W}_{i-1}, X_i) \stackrel{d}{=} Y_i|(Z_i, V_i)$ ja

$$Y_i|(Z_i = z_i, V_i = v_i) \sim \mathrm{N}\left(z_i'\beta, \sigma_i^2(\delta)\right), \quad \sigma_i^2(\delta) = \exp\left\{v_i'\delta\right\}.$$

- Oletetaan siis $Y_i|(\mathbf{W}_{i-1}, X_i) \stackrel{d}{=} Y_i|(Z_i, V_i)$ ja

$$Y_i|(Z_i = z_i, V_i = v_i) \sim \mathrm{N}\left(z_i'\beta, \sigma_i^2(\delta)\right), \quad \sigma_i^2(\delta) = \exp\left\{v_i'\delta\right\}.$$

- Ehdolliseksi uskottavuusfunktioksi saadaan $\left(\theta = (\beta, \delta) \in \mathbb{R}^p \times \mathbb{R}^k\right)$

$$L^{(c)}(\theta; \mathbf{w}) = \prod_{i=1}^{n} \sigma_i^2(\delta)^{-1/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\frac{(y_i - z_i'\beta)^2}{\sigma_i^2(\delta)}\right\},$$

- Oletetaan siis $Y_i | (\mathbf{W}_{i-1}, X_i) \overset{d}{=} Y_i | (Z_i, V_i)$ ja

$$Y_i | (Z_i = z_i, V_i = v_i) \sim \mathrm{N}\left(z_i'\beta, \sigma_i^2(\delta)\right), \quad \sigma_i^2(\delta) = \exp\left\{v_i'\delta\right\}.$$

- Ehdolliseksi uskottavuusfunktioksi saadaan $\left(\theta = (\beta, \delta) \in \mathbb{R}^p \times \mathbb{R}^k\right)$

$$L^{(c)}(\theta; \mathbf{w}) = \prod_{i=1}^{n} \sigma_i^2(\delta)^{-1/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} \frac{(y_i - z_i'\beta)^2}{\sigma_i^2(\delta)}\right\},$$

- johon päädytään myös lähtemällä yhtälöstä

$$Y_i = Z_i'\beta + \sigma_i(\delta)\eta_i, \quad i = 1, ..., n,$$

jossa $\eta_1, ..., \eta_n \perp\!\!\!\perp$, $\eta_i \sim \mathrm{N}(0,1)$ ja $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \eta_i \;\; \forall \;\; i = 1, ..., n.$

- Oletetaan siis $Y_i | (\mathbf{W}_{i-1}, X_i) \overset{d}{=} Y_i | (Z_i, V_i)$ ja

$$Y_i | (Z_i = z_i, V_i = v_i) \sim \mathrm{N}\left(z_i'\beta, \sigma_i^2(\delta)\right), \quad \sigma_i^2(\delta) = \exp\left\{v_i'\delta\right\}.$$

- Ehdolliseksi uskottavuusfunktioksi saadaan $\left(\theta = (\beta, \delta) \in \mathbb{R}^p \times \mathbb{R}^k\right)$

$$L^{(c)}(\theta; \mathbf{w}) = \prod_{i=1}^{n} \sigma_i^2(\delta)^{-1/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\frac{(y_i - z_i'\beta)^2}{\sigma_i^2(\delta)}\right\},$$

- johon päädytään myös lähtemällä yhtälöstä

$$Y_i = Z_i'\beta + \sigma_i(\delta)\eta_i, \quad i = 1, ..., n,$$

jossa $\eta_1, ..., \eta_n \perp\!\!\!\perp$, $\eta_i \sim \mathrm{N}(0,1)$ ja $(\mathbf{W}_{i-1}, X_i) \perp\!\!\!\perp \eta_i \ \ \forall \ \ i = 1, ..., n.$

- $\mathrm{H}_0 \Rightarrow \sigma_i(\delta)\eta_i = \sigma\eta_i \sim \mathrm{N}\left(0, \sigma^2\right) \Rightarrow$ merkitsemällä $\varepsilon_i = \sigma\eta_i$ saadaan aikaisempi homoskedastinen tapaus

- Ehdollinen uskottavuusfunktio

$$L^{(c)}(\theta;\mathbf{w}) = \prod_{i=1}^{n} \sigma_i^2(\delta)^{-1/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\frac{(y_i - z_i'\beta)^2}{\sigma_i^2(\delta)}\right\}.$$

- Ehdollinen uskottavuusfunktio

$$L^{(c)}(\theta;\mathbf{w}) = \prod_{i=1}^{n} \sigma_i^2(\delta)^{-1/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\frac{(y_i - z_i'\beta)^2}{\sigma_i^2(\delta)}\right\}.$$

- $\sigma_i^2(\delta) = \exp\{v_i'\delta\}$ & logaritmointi $\Rightarrow$

$$l^{(c)}(\theta;\mathbf{w}) = -\frac{1}{2}\sum_{i=1}^{n} v_i'\delta - \frac{1}{2}\sum_{i=1}^{n}\exp\{-v_i'\delta\}(y_i - z_i'\beta)^2.$$

- Ehdollinen uskottavuusfunktio

$$L^{(c)}(\theta;\mathbf{w}) = \prod_{i=1}^{n} \sigma_i^2(\delta)^{-1/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\frac{(y_i - z_i'\beta)^2}{\sigma_i^2(\delta)}\right\}.$$

- $\sigma_i^2(\delta) = \exp\{v_i'\delta\}$ & logaritmointi $\Rightarrow$

$$l^{(c)}(\theta;\mathbf{w}) = -\frac{1}{2}\sum_{i=1}^{n} v_i'\delta - \frac{1}{2}\sum_{i=1}^{n}\exp\{-v_i'\delta\}(y_i - z_i'\beta)^2.$$

- Derivointi $\Rightarrow$

$$\frac{\partial}{\partial\beta} l^{(c)}(\theta;\mathbf{w}) = \sum_{i=1}^{n}\exp\{-v_i'\delta\}(y_i - z_i'\beta)z_i$$

ja

$$\frac{\partial}{\partial\delta} l^{(c)}(\theta;\mathbf{w}) = \frac{1}{2}\sum_{i=1}^{n}\left[\exp\{-v_i'\delta\}(y_i - z_i'\beta)^2 - 1\right]v_i.$$

- Suoraviivainen lasku $\Rightarrow$ parametrit β ja δ ovat ortogonaalisia $\Rightarrow$ Raon testi voidaan perustaa pelkästään δ:n pistemäärään ja havaittuun infoon.

- Suoraviivainen lasku $\Rightarrow$ parametrit β ja δ ovat ortogonaalisia $\Rightarrow$ Raon testi voidaan perustaa pelkästään δ:n pistemäärään ja havaittuun infoon.

- Edellinen laskettu. Jälkimmäistä varten lasketaan

$$\frac{\partial^2}{\partial\delta\partial\delta'} l^{(c)}(\theta; \mathbf{w}) = -\frac{1}{2}\sum_{i=1}^{n} \exp\left\{-v_i'\delta\right\}\left(y_i - z_i'\beta\right)^2 v_i v_i'.$$

- Suoraviivainen lasku $\Rightarrow$ parametrit β ja δ ovat ortogonaalisia $\Rightarrow$ Raon testi voidaan perustaa pelkästään δ:n pistemäärään ja havaittuun infoon.

- Edellinen laskettu. Jälkimmäistä varten lasketaan

$$\frac{\partial^2}{\partial\delta\partial\delta'} l^{(c)}(\theta; \mathbf{w}) = -\frac{1}{2}\sum_{i=1}^{n} \exp\left\{-v_i'\delta\right\}\left(y_i - z_i'\beta\right)^2 v_i v_i'.$$

- $Y_i = Z_i'\beta + \sigma_i(\delta)\,\eta_i, \quad i = 1, \ldots, n,$

- Suoraviivainen lasku $\Rightarrow$ parametrit β ja δ ovat ortogonaalisia $\Rightarrow$ Raon testi voidaan perustaa pelkästään δ:n pistemäärään ja havaittuun infoon.

- Edellinen laskettu. Jälkimmäistä varten lasketaan

$$\frac{\partial^2}{\partial\delta\partial\delta'} l^{(c)}(\theta; \mathbf{w}) = -\frac{1}{2}\sum_{i=1}^{n} \exp\left\{-v_i'\delta\right\}\left(y_i - z_i'\beta\right)^2 v_i v_i'.$$

- $Y_i = Z_i'\beta + \sigma_i(\delta)\,\eta_i, \quad i = 1, ..., n,$
- H_0:n voimassa ollessa $\sigma_i(\delta)\,\eta_i = \sigma\eta_i = \varepsilon_i \sim \mathrm{N}\left(0, \sigma^2\right)$

- Suoraviivainen lasku $\Rightarrow$ parametrit β ja δ ovat ortogonaalisia $\Rightarrow$ Raon testi voidaan perustaa pelkästään δ:n pistemäärään ja havaittuun infoon.

- Edellinen laskettu. Jälkimmäistä varten lasketaan

$$\frac{\partial^2}{\partial\delta\partial\delta'} l^{(c)}(\theta; \mathbf{w}) = -\frac{1}{2}\sum_{i=1}^{n} \exp\left\{-v_i'\delta\right\}\left(y_i - z_i'\beta\right)^2 v_i v_i'.$$

- $Y_i = Z_i'\beta + \sigma_i(\delta)\eta_i, \quad i = 1, ..., n,$

- H_0:n voimassa ollessa $\sigma_i(\delta)\eta_i = \sigma\eta_i = \varepsilon_i \sim N\left(0, \sigma^2\right)$

- Siis, $H_0 \Rightarrow$ lineaarinen mallin

$$Y_i = Z_i'\beta + \varepsilon_i, \quad i = 1, ..., n,$$

josta testissä tarvittavat rajoitetut SUE:t saadaan helposti.

- Rajoitettu SUE mallista $Y_i = Z_i'\beta + \varepsilon_i \;\Rightarrow\; \tilde{\theta} = (\tilde{\beta}, \tilde{\delta}) = \tilde{\beta} = \hat{\beta}$ ja $\tilde{\delta} = (\tilde{\delta}_1, 0, ..., 0)$, jossa $\tilde{\delta}_1 = \log \hat{\sigma}^2$ ja

$$\hat{\beta} = \left(\sum_{i=1}^{n} z_i z_i'\right)^{-1} \sum_{i=1}^{n} z_i y_i \quad \text{ja} \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - z_i'\hat{\beta})^2.$$

- Rajoitettu SUE mallista $Y_i = Z_i'\beta + \varepsilon_i \;\Rightarrow\; \tilde{\theta} = (\tilde{\beta}, \tilde{\delta}) = \tilde{\beta} = \hat{\beta}$ ja $\tilde{\delta} = (\tilde{\delta}_1, 0, ..., 0)$, jossa $\tilde{\delta}_1 = \log \hat{\sigma}^2$ ja

$$\hat{\beta} = \left(\sum_{i=1}^{n} z_i z_i'\right)^{-1} \sum_{i=1}^{n} z_i y_i \quad \text{ja} \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - z_i'\hat{\beta})^2.$$

- Tarvittava δ:n pistemäärä

$$s_\delta(\tilde{\theta}; \mathbf{w}) := \frac{\partial}{\partial \delta} l^{(c)}(\tilde{\theta}; \mathbf{w}) = \frac{1}{2}\sum_{i=1}^{n}\left(\frac{(y_i - z_i'\tilde{\beta})^2}{\tilde{\sigma}^2} - 1\right) v_i, \quad \tilde{\sigma}^2 = \hat{\sigma}^2.$$

- Rajoitettu SUE mallista $Y_i = Z_i'\beta + \varepsilon_i \;\Rightarrow\; \tilde{\theta} = (\tilde{\beta}, \tilde{\delta}) = \tilde{\beta} = \hat{\beta}$ ja $\tilde{\delta} = (\tilde{\delta}_1, 0, ..., 0)$, jossa $\tilde{\delta}_1 = \log \hat{\sigma}^2$ ja

$$\hat{\beta} = \left(\sum_{i=1}^{n} z_i z_i' \right)^{-1} \sum_{i=1}^{n} z_i y_i \quad \text{ja} \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (Y_i - z_i'\hat{\beta})^2.$$

- Tarvittava δ:n pistemäärä

$$s_\delta(\tilde{\theta}; \mathbf{w}) := \frac{\partial}{\partial \delta} l^{(c)}(\tilde{\theta}; \mathbf{w}) = \frac{1}{2} \sum_{i=1}^{n} \left(\frac{(y_i - z_i'\tilde{\beta})^2}{\tilde{\sigma}^2} - 1 \right) v_i, \quad \tilde{\sigma}^2 = \hat{\sigma}^2.$$

- Tarvittava δ:n havaittu informaatio

$$\mathcal{J}_{\delta\delta}(\tilde{\theta}; \mathbf{w}) := -\frac{\partial^2}{\partial \delta \partial \delta'} l^{(c)}(\tilde{\beta}, \tilde{\delta}; \mathbf{w}) = \frac{1}{2\tilde{\sigma}^2} \sum_{i=1}^{n} (y_i - z_i'\tilde{\beta})^2 v_i v_i'.$$

- $\tilde{\beta} = \hat{\beta}$ ja $\tilde{\sigma}^2 = \hat{\sigma}^2$ tavalliset lineaarisen mallin $Y_i = Z_i'\beta + \varepsilon_i$ estimaattorit ja

$$\tilde{s}_\delta(\tilde{\theta}) := \frac{\partial}{\partial \delta} l^{(c)}(\tilde{\theta}; \mathbf{W}) = \frac{1}{2} \sum_{i=1}^{n} \left(\frac{(Y_i - Z_i'\tilde{\beta})^2}{\tilde{\sigma}^2} - 1 \right) V_i$$

$$\tilde{\mathcal{J}}_{\delta\delta}(\tilde{\theta}) := -\frac{\partial^2}{\partial \delta \partial \delta'} l^{(c)}(\tilde{\theta}; \mathbf{W}) = \frac{1}{2\tilde{\sigma}^2} \sum_{i=1}^{n} (Y_i - Z_i'\tilde{\beta})^2 V_i V_i'.$$

- $\tilde{\beta} = \hat{\beta}$ ja $\tilde{\sigma}^2 = \hat{\sigma}^2$ tavalliset lineaarisen mallin $Y_i = Z_i'\beta + \varepsilon_i$ estimaattorit ja

$$\tilde{s}_\delta(\tilde{\theta}) := \frac{\partial}{\partial \delta} l^{(c)}(\tilde{\theta}; \mathbf{W}) = \frac{1}{2} \sum_{i=1}^{n} \left(\frac{(Y_i - Z_i'\tilde{\beta})^2}{\tilde{\sigma}^2} - 1 \right) V_i$$

$$\tilde{\mathcal{J}}_{\delta\delta}(\tilde{\theta}) := -\frac{\partial^2}{\partial \delta \partial \delta'} l^{(c)}(\tilde{\theta}; \mathbf{W}) = \frac{1}{2\tilde{\sigma}^2} \sum_{i=1}^{n} (Y_i - Z_i'\tilde{\beta})^2 V_i V_i'.$$

- $\mathrm{H}_0 : A\delta = 0$, jossa $A = [0 : I_{k-1}] \Rightarrow$ Raon testisuure

$$\begin{aligned} \mathrm{S}_\delta &= \tilde{s}_\delta(\tilde{\theta})' \tilde{\mathcal{J}}_{\delta\delta}(\tilde{\theta})^{-1} A' \left[A \tilde{\mathcal{J}}_{\delta\delta}(\tilde{\theta})^{-1} A' \right]^{-1} A \tilde{\mathcal{J}}_{\delta\delta}(\tilde{\theta})^{-1} \tilde{s}_\delta(\tilde{\theta}) \\ &= s_\delta(\tilde{\theta}; \mathbf{W})' \mathcal{J}_{\delta\delta}^{-1}(\tilde{\theta}; \mathbf{W}) s_\delta(\tilde{\theta}; \mathbf{W}) \\ &\xrightarrow[\mathrm{H}_0]{d} \chi^2_{k-1} \end{aligned}$$

- Koska $Y_i - Z_i'\beta = \sigma_i(\delta)\,\eta_i$ (olettaen todelliset parametriarvot) ja $\eta_i \perp\!\!\!\perp V_i$, niin H_0:n voimassa ollessa

$$
\begin{aligned}
\mathrm{E}\big[\big(Y_i - Z_i'\beta\big)^2\big] &= \sigma^2 \mathrm{E}(\eta_i^2) = \sigma^2 \\
\mathrm{E}\left[\mathcal{J}_{\delta\delta}(\theta; \mathbf{W})\right] &= \frac{1}{2}\sum_{i=1}^{n} \mathrm{E}\left(V_i V_i'\right).
\end{aligned}
$$

- Koska $Y_i - Z_i'\beta = \sigma_i(\delta)\,\eta_i$ (olettaen todelliset parametriarvot) ja $\eta_i \perp\!\!\!\perp V_i$, niin H_0:n voimassa ollessa
$$\begin{aligned} \mathrm{E}\big[\big(Y_i - Z_i'\beta\big)^2\big] &= \sigma^2 \mathrm{E}(\eta_i^2) = \sigma^2 \\ \mathrm{E}\left[\mathcal{J}_{\delta\delta}(\theta;\mathbf{W})\right] &= \frac{1}{2}\sum_{i=1}^{n} \mathrm{E}\left(V_i V_i'\right). \end{aligned}$$
- Siis, $\tilde{\mathcal{J}}_{\delta\delta}(\tilde{\theta}) = \mathcal{J}_{\delta\delta}(\tilde{\theta};\mathbf{W})$ voidaan testisuureessa S_δ korvata matriisilla $\tilde{\mathrm{M}}_{\delta\delta} = 2^{-1}\sum_{i=1}^{n} V_i V_i'$.

- Koska $Y_i - Z_i'\beta = \sigma_i(\delta)\,\eta_i$ (olettaen todelliset parametriarvot) ja $\eta_i \perp\!\!\!\perp V_i$, niin H_0:n voimassa ollessa

$$
\begin{aligned}
\mathrm{E}\big[\big(Y_i - Z_i'\beta\big)^2\big] &= \sigma^2 \mathrm{E}(\eta_i^2) = \sigma^2 \\
\mathrm{E}\left[\mathcal{J}_{\delta\delta}(\theta; \mathbf{W})\right] &= \frac{1}{2}\sum_{i=1}^{n} \mathrm{E}\left(V_i V_i'\right).
\end{aligned}
$$

- Siis, $\tilde{\mathcal{J}}_{\delta\delta}(\tilde\theta) = \mathcal{J}_{\delta\delta}(\tilde\theta; \mathbf{W})$ voidaan testisuureessa S_δ korvata matriisilla $\tilde{\mathrm{M}}_{\delta\delta} = 2^{-1}\sum_{i=1}^{n} V_i V_i'$.

- Saadaan testisuure

$$
\mathrm{S}_\delta^* = \tilde{\mathrm{s}}_\delta(\tilde\theta)'\tilde{\mathrm{M}}_{\delta\delta}^{-1}A'\left[A\tilde{\mathrm{M}}_{\delta\delta}^{-1}A'\right]^{-1}A\tilde{\mathrm{M}}_{\delta\delta}^{-1}\tilde{\mathrm{s}}_\delta(\tilde\theta) \xrightarrow{d} \chi^2_{k-1}
$$

H_0:n voimassa ollessa.

Suoraviivaisilla matriisilaskelmilla nähdään, että testisuure

$$\mathrm{S}_\delta^* = \tilde{\mathrm{s}}_\delta(\tilde{\theta})'\tilde{\mathrm{M}}_{\delta\delta}^{-1}A'\left[A\tilde{\mathrm{M}}_{\delta\delta}^{-1}A'\right]^{-1}A\tilde{\mathrm{M}}_{\delta\delta}^{-1}\tilde{\mathrm{s}}_\delta(\tilde{\theta}) \xrightarrow{d} \chi^2_{k-1}$$

voidaan kirjoittaa

$$\mathrm{S}_\delta^* = \frac{1}{2}\left(\sum_{i=1}^{n}\tilde{U}_i V_i'\right)\left(\sum_{i=1}^{n} V_i V_i'\right)^{-1}\left(\sum_{i=1}^{n}\tilde{U}_i V_i\right)$$

jossa $\tilde{U}_i = (Y_i - Z_i'\tilde{\beta})^2/\tilde{\sigma}^2 - 1$ ja kuten aikaisemmin V_i on ehdollista heteroskedastisuutta selittävä muuttuja.

- Johdettu testi perustui multiplikatiiviseen malliin

$$Y_i|(Z_i = z_i, V_i = v_i) \sim \mathrm{N}\left(z_i'\beta, \sigma_i^2(\delta)\right), \quad \sigma_i^2(\delta) = \exp\left\{v_i'\delta\right\},$$

jossa v_i $(k \times 1)$ ensimmäinen komponentti on vakio 1 ja muiksi komponenteiksi valitaan vektorin $(\mathbf{w}_{i-1}, x_i)$ "relevantit" komponentit.

- Johdettu testi perustui multiplikatiiviseen malliin

$$Y_i|(Z_i = z_i, V_i = v_i) \sim \mathrm{N}\left(z_i'\beta, \sigma_i^2(\delta)\right), \quad \sigma_i^2(\delta) = \exp\left\{v_i'\delta\right\},$$

jossa v_i $(k \times 1)$ ensimmäinen komponentti on vakio 1 ja muiksi komponenteiksi valitaan vektorin $\left(\mathbf{w}_{i-1}, x_i\right)$ "relevantit" komponentit.

- $v_i = (1, \log x_{a,i})$, $x_{a,i}$ selittävän muuttujan vektorin x_i (positiivinen) a. komponentti $\Rightarrow \sigma_i^2(\delta) = \sigma^2 x_{a,i}^{\delta_2}$.

- Johdettu testi perustui multiplikatiiviseen malliin

$$Y_i|(Z_i = z_i, V_i = v_i) \sim \mathrm{N}\left(z_i'\beta, \sigma_i^2(\delta)\right), \quad \sigma_i^2(\delta) = \exp\left\{v_i'\delta\right\},$$

  jossa v_i $(k \times 1)$ ensimmäinen komponentti on vakio 1 ja muiksi komponenteiksi valitaan vektorin $(\mathbf{w}_{i-1}, x_i)$ "relevantit" komponentit.

- $v_i = (1, \log x_{a,i})$, $x_{a,i}$ selittävän muuttujan vektorin x_i (positiivinen) a. komponentti $\Rightarrow \sigma_i^2(\delta) = \sigma^2 x_{a,i}^{\delta_2}$.

- Testisuureella S_δ^* lisäksi ns. invarianssiominaisuus eli samaan testisuureeseen päädytään olettamalla $\sigma_i^2(\delta) = h(v_i'\delta)$, jossa h on mikä tahansa kahdesti jatkuvasti derivoituva positiivinen funktio.