Lineaariset mallit, kl 2015, Harjoitus 4, viikko 16

1. Tarkastellaan yhden selittäjän lineaarista regressiomallia $Y_1, ..., Y_n \perp\!\!\!\perp$, $Y_i \sim \mathsf{N}\left(\beta_1 + \beta_2 x_i, \sigma^2\right)$. Olkoon $\hat{\boldsymbol{\beta}} = [\hat{\beta}_1 \quad \hat{\beta}_2]'$ parametrin $\boldsymbol{\beta} = [\beta_1 \quad \beta_2]'$ PNS-estimaattori ja $\bar{x} = (x_1 + \cdots + x_n)/n$. (i) Muodosta Lauseen 2.1(i) tulosta käyttäen $\mathsf{Cov}(\hat{\boldsymbol{\beta}})$ ja edelleen $\mathsf{Var}(\hat{\beta}_2)$ ja $\mathsf{Var}(\hat{\beta}_1 + \hat{\beta}_2 \bar{x})$.

(ii) Oletetaan, että selittävien muuttujien arvot $x_1, ..., x_n$ voidaan valita vapaasti väliltä $[c, d]$. Miten ne on valittava, jos $\mathsf{Var}(\hat{\beta}_2)$ halutaan minimoida? Onko tämä valinta muuten järkevä? (*Huom.*: Kohdassa (ii) ei vaadita yksityiskohtaisia matemaattisia perusteluja. Voit myös olettaa, että siinä n on parillinen.)

2. Tarkastellaan kahta riippumatonta lineaarista mallia

$$\mathbf{Y}_i = \mathbf{X}_i \boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i, \quad \boldsymbol{\varepsilon}_i \sim \mathsf{N}_{n_i}\left(\mathbf{0}, \sigma^2 \mathbf{I}_{n_i}\right), \quad \boldsymbol{\varepsilon}_1 \perp\!\!\!\perp \boldsymbol{\varepsilon}_2, \quad \boldsymbol{\beta}_i \in \mathbb{R}^p, \ \sigma^2 > \mathbf{0}, \ \mathsf{r}(\mathbf{X}_i) = p, \ i = 1, 2.$$

Muodosta näistä matriiseja käyttäen yksi malli ja esitä parametrin $\boldsymbol{\beta} = \left[\boldsymbol{\beta}_1' \quad \boldsymbol{\beta}_2'\right]'$ PNS-estimaattorin $\hat{\boldsymbol{\beta}} = [\hat{\boldsymbol{\beta}}_1' \quad \hat{\boldsymbol{\beta}}_2']'$ jakauma. Ovatko PNS-estimaattorit $\hat{\boldsymbol{\beta}}_1$ ja $\hat{\boldsymbol{\beta}}_2$ riippumattomia?

Aputulos: Jos $\mathbf{A}$ ja $\mathbf{B}$ ovat epäsingulaarisia neliömatriiseja, niin

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^{-1} \end{bmatrix}.$$

3. Jatkoa edelliselle. Estimoi parametri $\boldsymbol{\beta} = \left[\boldsymbol{\beta}_1' \quad \boldsymbol{\beta}_2'\right]'$ ehdolla $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$. Mikä on $\boldsymbol{\beta}_1$:n (ja $\boldsymbol{\beta}_2$:n) rajoitetun PNS-estimaattorin jakauma? (*Huom.*: Käytännön kannalta oletuksena on, että rajoite $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$ on muuttujien luonne huomioon ottaen mielekäs, mikä se on mahdollisesti todettu myös formaalissa testauksessa. Voit käyttää kumpaa tahansa monisteessa s. 16-17 esitettyä estimointitapaa, mutta toinen niistä on suoraviivaisempi.)

4. Tarkastellaan kahden riippumattoman normaalisen otoksen mallia

$$Y_1, ..., Y_n \perp\!\!\!\perp, \ Y_i \sim \begin{cases} \mathsf{N}\left(\mu_1, \sigma^2\right), & \text{kun } i = 1, ..., n_1 \\ \mathsf{N}\left(\mu_2, \sigma^2\right), & \text{kun } i = n_1 + 1, ..., n_1 + n_2 = n \end{cases}$$

(μ_1, $\mu_2 \in \mathbb{R}$, $\sigma^2 > 0$, n_1, $n_2 > 1$). (i) Harjoitustehtävän 2.5 ratkaisusta saadaan parametrin $\boldsymbol{\mu} = [\mu_1 \quad \mu_2]'$ PNS-estimaattori $\hat{\boldsymbol{\mu}}$. Selvitä $\hat{\boldsymbol{\mu}}$:n jakauma ja parametrin σ^2 harhattoman estimaattorin jakauma.

(ii) Estimoi nyt parametrit μ_1 ja μ_2 ehdolla $\mu_1 = \mu_2$ ja selvitä μ_1:n (ja μ_2:n) rajoitetun PNS-estimaattorin jakauma. Mikä on tässä tapauksessa parametrin σ^2 harhaton estimaattori ja sen jakauma? (*Huom.*: Voit käyttää tässäkin tapauksessa kumpaa tahansa monisteessa s. 16-17 esitettyä estimointitapaa, mutta toinen niistä on suoraviivaisempi.)