

LINEAARISET MALLIT I -KURSSI, MAT22004, tentti 13.05.2022

EI SALLITA LASKIMIA TAI MUISTIINPANOJA

1. (6 p.) Olkoon Y_1, \dots, Y_n riippumattomia normaalisti jakautuneita satunnaismuuttujia, joille pätee $Y_i \sim N(\mu_i, \sigma^2)$ ja

$$\mu_i = \begin{cases} \beta_1, & \text{kun } i = 1, \dots, n_1, \\ \beta_1 - 2\beta_2, & \text{kun } i = n_1 + 1, \dots, n_2, \\ \beta_1 + \beta_2, & \text{kun } i = n_2 + 1, \dots, n. \end{cases}$$

(a) Muotoile tilanne lineaarisena mallina ja estimoi mallin parametrit soveltamalla lineaarisen mallin estimointiteoriaa.

(b) Kuvaa lyhyesti käyttämäsi estimointiperiaate ja selvitä mallin parametrivektorin estimaattorin jakauma.

(c) Ovatko estimaattorin komponentit riippumattomia?

2. (6 p.) Oletetaan, että tavanomainen lineaarinen malli $\mathbf{Y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}_n)$ ($\boldsymbol{\beta} \in \mathbb{R}^p$, $\sigma^2 > 0$, $r(\mathbf{X}) = p$). Johda tilastollinen testi hypoteesille $H_0: \mathbf{A}\boldsymbol{\beta} = \mathbf{c}$, jossa kiinteän matriisin \mathbf{A} aste on k , $k < p$ ja vektori \mathbf{c} on kiinteä.

3. (6 p.) Olkoon identifioituva lineaarinen malli sellainen, että mallin kerroinvektorin $\boldsymbol{\beta}$ dimensio on p . Johda F -testisuureen avulla $1 - \alpha$ luottamustasoinen luottamusjoukko vektorille $\boldsymbol{\beta}$. Selitä tekemäsi päättelyn askeleet niin yksityiskohtaisesti kuin mahdollista.

4. (6 p.) Oletetaan aineiston y_1, \dots, y_n olevan riippumattomien satunnaismuuttujien $Y_i \sim N(\beta_1 x_{i1} + \beta_2 x_{i1}, \sigma^2)$, kun $i = 1, \dots, n_1$, ja satunnaismuuttujien $Y_i \sim N(\beta_1 x_{i1} + \beta_2 x_{i1}, \sigma_i^2)$, kun $i = n_1 + 1, \dots, n$ generoima ja lisäksi oletetaan selittävien muuttujien x_{i1} , $i = 1, \dots, n$, olevan ei-satunnaisia. Voidaanko tällöin aineiston tilastolliseen mallintamiseen käyttää lineaarista mallia? Jos ei, niin miten oletuksia aineiston generoivasta prosessista pitäisi muuttaa, jotta lineaarisen mallin käyttö olisi mahdollista?

Muistin tueksi:

Satunnaisvektorin $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tiheysfunktio on

$$f(\mathbf{x}) = (2\pi)^{-k/2} \det(\boldsymbol{\Sigma})^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\},$$

jossa $\det(\boldsymbol{\Sigma})$ on kovarianssimatriisin $\boldsymbol{\Sigma}$ determinantti ja yksilutellinen tapaus saadaan asettamalla $k = 1$. $F_{k,m}$ -jakauman määrittelee satunnaismuuttujia $m\chi_k^2/k\chi_m^2$, jossa $\chi_k^2 \perp\!\!\!\perp \chi_m^2$. Lisäksi $E(\chi_k^2) = k$, $Var(\chi_k^2) = 2k$. t_k -jakauman määrittelee satunnaismuuttujia $Z/\sqrt{\frac{1}{k}\chi_k^2}$, jossa $Z \sim N(0, 1)$ ja $Z \perp\!\!\!\perp \chi_k^2$.

Jos $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, niin $(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu}) \sim \chi_k^2$.

Jos $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \sigma^2\mathbf{I}_k)$ ja matriisi \mathbf{P} ($k \times k$) astetta r oleva ortogonaalinen projektiio, niin $(\mathbf{X} - \boldsymbol{\mu})' \mathbf{P}(\mathbf{X} - \boldsymbol{\mu})/\sigma^2 \sim \chi_r^2$.

Oletetaan, että \mathbf{X} on ei-satunnainen $n \times p$ matriisi (täyttää sarakeastetta), ja SU estimaattorille pätee $\hat{\boldsymbol{\beta}} \sim N_p(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1})$. Jos lisäksi matriisi \mathbf{A} ($q \times p$) ja vektori \mathbf{c} ($q \times 1$) ovat tunnettuja ja $r(\mathbf{A}) = q$, niin $(\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c})'(\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}')^{-1}(\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c})/qS^2 \sim F_{q, n-p}$, kun parametriarvuudessa pätee tietty lineaarinen rajoite.