Lineaariset mallit I -kurssi, MAT22004, kurssikoe 08.05.2018

1. (6 p.) a) Määrittele yleinen lineaarinen malli (matriisiesitys $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$) *kaikkine oletuksineen* ja siihen liittyen käsitteet (pienimmän neliösumman) sovite ja residuaali. b) Selvitä mallin sekä sovitteen että residuaalin tulkinta. c) Laske $\mathsf{Cov}(\hat{\mathbf{Y}})$ ja $\mathsf{Cov}(\hat{\boldsymbol{\varepsilon}}, \hat{\mathbf{Y}})$, kun $\hat{\mathbf{Y}}$ on aineistosta laskettua sovitevektoria vastaava satunnaisvektori ja $\hat{\boldsymbol{\varepsilon}}$ on aineistosta laskettua residuaalivektoria vastaava satunnaisvektori.

2. (6 p.) Olkoon $Y_1, ..., Y_5$ riippumattomia normaalisti jakautuneita satunnaismuuttujia, joille pätee $Y_i \sim \mathsf{N}(\mu_i, \sigma^2)$ ja

$$\mu_i = \begin{cases} \beta_1, & \text{kun } i = 1, 2 \\ \beta_1 + 2\beta_2, & \text{kun } i = 3 \\ \beta_1 - \beta_2, & \text{kun } i = 4, 5. \end{cases}$$

Muotoile tilanne lineaarisena mallina ja estimoi parametrit β_1 ja β_2 soveltamalla lineaarisen mallin estimointiteoriaa. Kuvaa lyhyesti käyttämäsi estimointiperiaate ja selvitä parametrivektorin $\boldsymbol{\beta} = [\beta_1 \;\; \beta_2]'$ estimaattorin $\hat{\boldsymbol{\beta}} = [\hat{\beta}_1 \;\; \hat{\beta}_2]'$ jakauma. Ovatko $\hat{\beta}_1$ ja $\hat{\beta}_2$ riippumattomia?

3. (6 p.) Yksisuuntainen varianssianalyysi: Ongelmanasettelu, tilastollinen malli, tavallisimmin testattava hypoteesi ja sen testaaminen. Selitä myös kuinka johdettua testisuuretta käyttäen tehdään johtopäätös hypoteesin suhteen.

4. (6 p.) Oletetaan, että n riippumatonta havaintoa $Y_1, ..., Y_n$ $(n > 2)$ on poimittu normaalijakaumasta, jonka varianssi on tuntematon σ^2. Ensimmäisen $n - 1$ havainnon tiedetään olevan peräisin samasta normaalijakaumasta, mutta ennen viimeisen havainnon poimimista epäillään, että havaintojen odotusarvo on muuttunut äkillisesti ja saattaa poiketa ensimmäisen $n - 1$ havainnon odotusarvosta. Muotoile tilanne lineaarisen mallin avulla ja johda testi hypoteesille, jonka mukaan n. havainto on peräisin samasta normaalijakaumasta kuin $n - 1$ ensimmäistä havaintoa. Muista selittää kuinka johdettua testisuuretta käyttäen tehdään johtopäätös hypoteesin suhteen.

Muistin tueksi

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

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

  jossa $\det(\boldsymbol{\Sigma})$ on kovarianssimatriisin $\boldsymbol{\Sigma}$ determinantti ja yksiulotteinen tapaus saadaan asettamalle $k = 1$.

- $F_{k,m}$-jakauman määrittelee satunnaismuuttuja $m\chi_k^2 / k\chi_m^2$, jossa $\chi_k^2 \perp\!\!\!\perp \chi_m^2$. Lisäksi $E(\chi_k^2) = k$, $Var\left(\chi_k^2\right) = 2k$.

- t_k-jakauman määrittelee satunnaismuuttuja $Z / \sqrt{\frac{1}{k}\chi_k^2}$, jossa $Z \sim N(0, 1)$ ja $Z \perp\!\!\!\perp \chi_k^2$.

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

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