Lineaariset mallit, kevät 2014

Harjoitus 5, viikko 17

1. Jatkoa harjoituksen 4 tehtävälle 4. Estimoi parametrit μ_1 ja μ_2 ehdolla $\mu_1 = \mu_2$ ottamalla ehto $\mu_1 = \mu_2$ huomioon mallissa ja estimoimalla saadun mallin parametrit (eli käyttäen olennaisesti yhtälöön (2.9) perustuvaa vaihtoehtoista menettelyä).

2. Jatkoa harjoituksen 4 tehtävälle 3. Estimoi parametri $\boldsymbol{\beta} = [\boldsymbol{\beta}_1' \ \boldsymbol{\beta}_2']'$ ehdolla $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$ PNS-menetelmän avulla. Mikä on nyt estimaattorin jakauma?

3. Jatkoa harjoitustehtäville 4.3 ja 5.2. Johda F-testi hypoteesille $H : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$. (Vihje: Käytä F-testisuureen jälkimmäistä residuaalineliösummiin perustuvaa esitystä (ks. s. 19).)

4. Jatkoa harjoitustehtäville 4.3, 5.2 ja 5.3. (i) Muodosta $100(1-\alpha)\%$:n luottamusväli parametrivektorin $\boldsymbol{\beta}_1$ lineaarikombinaatiolle $\boldsymbol{a}'\boldsymbol{\beta}_1$ ($\boldsymbol{a} = [a_1 \cdots a_p]' \neq \boldsymbol{0}$). (ii) Tee sama olettaen, että $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$.

5. Tarkastellaan yhden selittäjän lineaarista regressiomallia $Y_1, \ldots, Y_n \perp\!\!\!\perp$, $Y_i \sim N(\beta_1 + \beta_2 x_i, \sigma^2)$. Johda F-testi nollahypoteesille $\beta_2 = 0$ ja osoita, että testisuure voidaan lausua (tavanomaisin merkinnöin) muodossa

$$F = \hat{\beta}_2^2 \Big/ S^2 \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

   (Vihje: Mallin matriisiesitys ja F-testisuureen yleinen lauseke (ks. monisteen s. 18). Testisuureen haetun lausekkeen johtamisessa tarvitset myös 2×2 matriisin käänteismatriisin laskukaavaa.)

6. Tarkastellaan jakson 3.2 alun tilannetta (s. 19), jossa malliyhtälö on $Y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon_i$, $i = 1, \cdots, n$, ja testattava hypoteesi $H : \beta_2 = \cdots = \beta_p = 0$. Osoita ensin, että $\mathrm{SSE} = (1 - R^2) \sum_{i=1}^{n} (y_i - \bar{y})^2$, jossa R^2 on selitysaste ja SSE on residuaalineliösumma (ks. s. 10), ja tämän perusteella edelleen, että F-testisuure edellä mainitulle hypoteesille voidaan kirjoittaa

$$F = (n-p)R^2/(p-1)(1-R^2).$$

7. Oletetaan, että tavanomaisessa lineaarisessa mallissa $\boldsymbol{Y} \sim N_n(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{I}_n)$ ($\boldsymbol{\beta} \in \mathbb{R}^p, \sigma^2 > 0, r(\boldsymbol{X}) = p$) selittäjät ovat ortogonaalisia eli $\boldsymbol{X}'\boldsymbol{X}$ on diagonaalimatriisi. Johda parametrin β_j PNS-estimaatti, t-testisuure hypoteesille $H : \beta_j = 0$ ja $100(1-\alpha)\%$:n luottamusväli parametrille β_j ($= \boldsymbol{\beta}$:n j. komponentti). Miten nämä muuttuvat, jos mallista poistetaan joku selittäjistä x_k ($k \neq j$)?