Lineaariset mallit, kl 2016, Harjoitus 5, viikko 17

1. Jatkoa HT:lle 4.2. Johda F-testi nollahypoteesille $\mu_1 = \mu_2$ ja osoita, että yhtäpitävä testi voidaan perustaa t_{n-2}-jakaumaa noudattavaan t-testisuureeseen ($n = n_1 + n_2$).

Vihje: Käytä monisteen s. 18 olevaa F-testisuureen lauseketta.

2. Tarkastellaan monisteen jakson 3.2 alun tilannetta (s. 18-19), jossa malliyhtälö on $Y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon_i$, $i = 1, ..., n$, ja testattava hypoteesi $H : \beta_2 = \cdots = \beta_p = 0$. Osoita ensin, että $\text{SSE} = \left(1 - R^2\right) \sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2$, jossa R^2 on selitysaste ja SSE on residuaalineliösumma (ks. R^2:n määritelmä monisteen s. 10). Osoita tätä tulosta käyttäen edelleen, että F-testisuure edellä mainitulle hypoteesille voidaan kirjoittaa

$$F = (n - p) \, R^2 / \, (p - 1) \left(1 - R^2\right).$$

3. Oletetaan, että edellisessä tehtävässä $p = 2$. Osoita, että F-testin asemesta voidaan käyttää t-testiä, joka perustuu testisuureeseen

$$t = \sqrt{n-2} \, r_{yx} / \sqrt{1 - r_{yx}^2} \overset{H}{\sim} t_{n-2},$$

jossa merkitsemällä $x_{i2} = x_i$, $\bar{x} = n^{-1} \sum_{i=1}^{n} x_i$ ja $\bar{y} = n^{-1} \sum_{i=1}^{n} y_i$

$$r_{yx} = \frac{\sum_{i=1}^{n} \left(y_i - \bar{y}\right)\left(x_i - \bar{x}\right)}{\sqrt{\sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2} \sqrt{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}}$$

eli selitettävän muuttujan ja selittävän muuttujan välinen otoskorrelaatiokerroin. Mitä hypoteesia tämän testin voidaan myös tulkita testaavan?

Vihje: Käytä edellistä tehtävää ja monisteen s. 11 esitettyä identiteettiä $R^2 = r_{y\hat{y}}^2$, jossa voit käyttää edelleen tehtävästä 2.4 saatavaa tulosta $\hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \bar{x}$.

4. Jatkoa tehtäville 4.3 ja 4.4. Johda F-testi hypoteesille $H : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$.

5. Jatkoa tehtävälle 1. Oletetaan, että hypoteesi $\mu_1 = \mu_2$ on tullut hylätyksi. Muodosta $100(1 - \alpha)\%$:n luottamusväli erotukselle $\mu_1 - \mu_2$.