Lineaariset mallit, kl 2011, Harjoitus 5, viikko 18 (Huom. aika)

1. Jatkoa HT:lle 4.3. Estimoi parametri $\boldsymbol{\beta} = \begin{bmatrix}\boldsymbol{\beta}_1' & \boldsymbol{\beta}_2'\end{bmatrix}'$ ehdolla $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$. Mikä on estimaattorin $\hat{\boldsymbol{\beta}}_1$ jakauma?

2. Jatkoa HT:lle 4.4. 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*: Mallin matriisiesitys (ks. monisteen s. 5), F-testisuureen yleinen lauseke (ks. monisteen s. 18) ja tulos $\sqrt{F_{1,k}} \sim t_k$).

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

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

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

4. Jatkoa edelliselle ja HT:lle 2.3. (i) Osoita, että edellisen tehtävän mallissa

$$\mathrm{SSE} = \sum_{i=1}^n (y_i - \bar{y})^2 - \hat{\beta}_2^2 \sum_{i=1}^n (x_i - \bar{x})^2 = \left(1 - r_{xy}^2\right)(n-1)\, s_y^2,$$

jossa r_{xy} on havainnoista (y_i, x_i), $i = 1, ..., n$, laskettu korrelaatiokerroin (ks. monisteen alaviite s. 11), $s_y^2 = (n-1)^{-1} \sum_{i=1}^n (y_i - \bar{y})^2$ ja residuaalineliösumma SSE on kuten monisteen s. 10. (ii) Osoita tämän avulla, että edellisen tehtävän F-testi on yhtäpitävä seuraavan t-testin kanssa

$$\sqrt{n-2}\, r_{xy} \Big/ \sqrt{1 - r_{xy}^2} \overset{H}{\sim} t_{n-2}.$$

Minkä hypoteesin testiksi tämä voidaan myös tulkita? (*Vihje*: Monisteen s. 10 oleva yhtälö SST = SSR + SSE, jossa SSR:ää voidaan muokata HT:n 2.3 tuloksen avulla. Muista myös, että nyt $S^2 = \mathrm{SSE}/(n-2)$).