Lineaaristen mallien kurssi - yleistentti 14.12.2012

1. Define the (full rank) linear model with all assumptions and explain briefly the interpretation of the model. Define also the concepts of (least squares) residual and fitted value and explain briefly their interpretation. Finally, show that $\mathbf{y}'\mathbf{y} = \hat{\mathbf{y}}'\hat{\mathbf{y}} + \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}$, where $\hat{\mathbf{y}}$ $(n \times 1)$ is the vector of fitted values, $\hat{\boldsymbol{\varepsilon}} = [\hat{\varepsilon}_1 \ \cdots \ \hat{\varepsilon}_n]'$ is the vector of residuals, and n is the sample size.

2. Let $Y_1, ..., Y_5$ be independent normally distributed random variables such that $Y_i \sim N(\mu_i, \sigma^2)$ $(i = 1, ..., 5)$, where

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

Formulate the situation as a linear model and estimate the parameters β_1 and β_2 by applying the estimation theory of the linear model. Describe briefly the estimation principle and find out the probability distribution of $\hat{\boldsymbol{\beta}} = [\hat{\beta}_1 \ \ \hat{\beta}_2]'$, the estimator of the parameter vector $\boldsymbol{\beta} = [\beta_1 \ \ \beta_2]'$.

3. Let $Y_1, ..., Y_4$ be independent normally distributed random variables such that $Y_i \sim \mathsf{N}(i\beta, \sigma^2)$ $(i = 1, ..., 4)$. Formulate the situation as a linear model and derive a test for the hypothesis $H : \beta = 0$ against the alternative $\beta \neq 0$.

4. Suppose that the random variables $Y_{11}, ..., Y_{1n_1}, Y_{21}, ..., Y_{2n_2}, Y_{31}, ..., Y_{3n_3}$ are independent with $Y_{ji} \sim \mathsf{N}\left(\mu_j, \sigma^2\right)$, $i = 1, ..., n_j$, $j = 1, 2, 3$. Formulate the situation as a linear model and test the hypothesis $H : \mu_1 = \mu_2 = \mu_3$ by applying the test theory of the linear model.

In case you happen to need:

- The density function of a random vector $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is

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

  where $\det(\boldsymbol{\Sigma})$ is the determinant of the covariance matrix $\boldsymbol{\Sigma}$ and the real valued case is obtained by setting $k = 1$.

- The $F_{k,m}$-distribution is defined by the random variable $m\chi_k^2 / k\chi_m^2$ where χ_k^2 and χ_m^2 are independent. Moreover, $\mathsf{E}(\chi_k^2) = k$, $\mathsf{Var}\left(\chi_k^2\right) = 2k$.

- The t_k-distribution is defined by the random variable $Z/\sqrt{\frac{1}{k}\chi_k^2}$ where $Z \sim \mathsf{N}(0, 1)$ and Z and χ_k^2 are independent.

- If $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ then $(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_k^2$.

- If $\mathbf{X} \sim \mathsf{N}_k\left(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_k\right)$ and the matrix $\mathbf{P}$ $(k \times k)$ is an orthogonal projection of rank r then $(\mathbf{X} - \boldsymbol{\mu})' \mathbf{P} (\mathbf{X} - \boldsymbol{\mu}) / \sigma^2 \sim \chi_r^2$.