KTH | **Project assignment VT 2008**
Inst. för Matematik | **5B1541 Sannolikhetsteori och**
Harald Lang | **Linjära modeller**

Generate 3×51 independent $N(0,1)$ variables (with some generator of pseudo random numbers) $z_{0,i}, z_{1,i}, z_{2,i}$, $i = 1, \ldots, 51$.

With these numbers You then create 3×51 values $y_i, x_{1,i}, x_{2,i}$ with

$$\begin{pmatrix} y_i \\ x_{1,i} \\ x_{2,i} \end{pmatrix} = \begin{pmatrix} 4 & 7 & -2 \\ 3 & 3 & 3 \\ 0 & 3 & 6 \end{pmatrix} \begin{pmatrix} z_{0,i} \\ z_{1,i} \\ z_{2,i} \end{pmatrix} + \begin{pmatrix} 50 \\ 30 \\ 20 \end{pmatrix}, \quad i = 1, \ldots, 51.$$

1. Determine $\theta_0, \theta_1, \theta_2$ so that

$$E\left(y - \theta_0 + x_1\theta_1 + x_2\theta_2\right)^2$$

is minimized. In other words, determine the coefficients in a theoretical regression function

$$y_i = \theta_0 + x_{1,i}\theta_1 + x_{2,i}\theta_2 + \epsilon_i.$$

Determine also (the optimal residual variance) $\sigma^2 = \mathrm{Var}(\epsilon)$. *Hint*: Extend the technique in Theorem 5.2. and Theorem 5.3 in Gut p. 52 to three coefficients.
Report these values and show Your calculations!

2. Next forget the previous values of $\theta_0, \theta_1, \theta_2$. Consider the x-variables as given (i.e., condition w.r.t the observed x-variables) and *estimate* $\theta_0, \theta_1, \theta_2$ och σ^2 according to the LSE theory in the lecture handout, **but use only the 50 first values !** The final 51:st values are retained for further purposes.
Report the value of Your estimates!

3. Determine an (approximative) 95% confidence interval for θ_0, θ_1 och θ_2. Are the "true" values found in project assignment 1. above inside these confidence intervals ?
Report these three confidence intervals !

4. Test the simultaneous hypothesis that
H_0: $\quad \theta_0 = 10$ and $\theta_1 + 2\theta_2 = 0$.

on significance level 95%. Do you accept the hypothesis ?
Report the result of the test !

5. Make a prediction on y when $x_1 = x_{1,51}$ och $x_2 = x_{2,51}$ and make a prediction interval with the degree of confidence 95%. Does the value of y_{51} lie in the prediction interval ?
Report the value of the prediction of y.
Report Your observed value of y_{51} and prediction interval !