# 1 Econometrics for Business Informatics

## 1.1 Exercises 1: Date 22.3.2013

1. The random variable $Y$ has a mean of 15 and a variance of 9. Let $Z = \frac{1}{3}(Y - 15)$. Show that $Z$ has mean zero and variance one.

2. Let $X$ and $Y$ be two random variables with means $\mu_X$ and $\mu_Y$, variances $\sigma_Y^2$, $\sigma_X^2$ and covariance $\sigma_{XY}$. Show that $E(XY) = \sigma_{XY} + \mu_X \mu_Y$.

3. Let $a$ and $b$ be constants and $X$ and $Y$ be two random variables with means $\mu_X$ and $\mu_Y$, variances $\sigma_Y^2$, $\sigma_X^2$ and covariance $\sigma_{XY}$. Show that $Var(aX + bY) = a^2\sigma_X^2 + 2ab\sigma_{XY} + b^2\sigma_Y^2$.

4. Suppose you have some money to invest (for simplicity 1$) and you are planning to put a fraction $w$ into a stock market fund and the rest $(1 - w)$ into a bond fund.The stock fund yields $R_s$ after one year and the bond fund $R_b$. $R_s$ is random with mean 0.08 and standard deviation 0.07. $R_b$ is random with mean 0.05 and standard deviation 0.04. The correlation between $R_b$ and $R_s$ is 0.25. If you place a fraction $w$ of your money in the stock fund and the rest $(1 - w)$ in the bond fund then the return on your investment is $R = wR_s + (1 - w)R_b$.
   (a) Suppose that $w = 0.5$. Compute the mean and standard deviation of $R$.
   (b) Suppose that $w = 0.75$. Compute the mean and standard deviation of $R$.
   (c) What value of $w$ makes the mean of $R$ as large as possible? What is the standard deviation of $R$ for this value of $w$.

5. A very simple model to explain the variable $Y_t$ is $Y_t = \beta + u_t$ with $\beta$ a constant and $t = 1, ...T$. Derive the least squares estimator for $\beta$, i.e. minimize $\sum_{t=1}^{T}(Y_t - \beta)^2$ with respect to $\beta$.

6. Define the variables in deviation from the mean form $y_t = Y_t - \bar{Y}$, $x_t = X_t - \bar{X}$ and show that for the simple regression model $y_t = \beta x_t + u_t$
$$\hat{\beta} - \beta = \frac{\sum x_t u_t}{\sum x_t^2}. \quad (1)$$
   with $\hat{\beta}$ the least squares estimator.

7. Show that the variance of the least squares estimator in the simple regression model $y_t = \beta x_t + u_t$ with variables in deviation form is
$$V(\hat{\beta}) = E[(\hat{\beta} - \beta)^2] = {\sigma^2}_u / \sum {x_t}^2 \quad (2)$$
   where ${\sigma^2}_u = E(u_t^2)$. Hint: Use the result of exercise 6.

8. Show that the matrix $M = I - X(X'X)^{-1}X'$ is idempotent, where $X$ is a matrix with $T$ rows and $k$ columns and $I$ the unit matrix of dimension $T$. Determine the rank of this matrix by using the properties of the trace of the matrix.

9. Use the data file **comprice.xls** which contains data on the development of average monthly oilprices (BRENT) and an index of non-energy world market commodity prices (HWWI). Consider the annual rate of change of these series as the random variable of interest. Transform the data into monthly series of the annual rates of change by calculating the percentage change of the series with respect to the month of the previous year (i.e. $y_t = (Y_t - Y_{t-12})/Y_{t-12} * 100$ for $t = Jan.1987, ..., Dec.2012$, altogether 312 observations). Use the data to answer the following questions:

(a) Compute the sample mean, the sample variance and the standard error, the skewness and the curtosis for the oilprice changes and the commodity price index changes.

(b) Draw a histogram for each random variable. How would you characterize the distributions.

(c) Compute the sample covariance and the correlation coefficient between the two random variables.

(d) How do these distributions compare with a normal distribution?

(e) Construct a 95%-confidence interval for the population means of the monthly oil and commodity price changes.

10. Apply **one** of the series in annual percentages of exercise 9 to the simple model $y_t = \beta + u_t$ and calculate the estimated residual series

$$\hat{u}_t = y_t - b_{OLS} \tag{3}$$

using the least squares estimator $b_{OLS}$. Calculate the sample variance of $u_t$ by computing the variance beginning with the first observation to the respective point of time, thus obtaining a "moving variance". Additionally compute an alternative moving variance by calculating the variances for the sequence of twelve observations of $\hat{u}_t$. Do you think that the assumption of a constant variance of $u_t$ is justified considering the plots obtained.