Exercises for ECON 4150, Spring 09

In order to become more familiar with regression and the potential use of regression analysis, students are recommended to attend the seminars arranged as part of the course. Students are encouraged to do the exercises on their own and then discuss their achievements at the seminars. A number of exercises are taken from the textbook "Principles of Econometrics" by Hill et. al. Good luck!

Problem set 1

Exercise 1.1

In any year, the weather can inflict storm damage to a home. From year to year, the damage is random. Let Y denote the dollar value of damage in any given year: Suppose that in 95% of the years $Y = \$0$, but in 5% of the years $Y = \$20000$.

(a) What is the mean and the standard deviation of the damage to a home in any given year?

(b) Consider an "insurance pool" of 100 people whose homes are sufficiently dispersed so that, in any given year, the damage to different homes are viewed as independently distributed random variables. Let $\bar{Y}$ denote the average damage to these homes 100 homes in a year. (i) What is the expected value of the average damage $\bar{Y}$? (ii) What is the probability that $\bar{Y}$ exceeds $\$2000$?

Exercise 1.2

Let $\bar{X}$ denote the arithmetic mean of a set of sample values $\{X_1, X_2, X_3, ..., X_N\}$. Show that

$$\sum_{i=1}^{N}(X_i - \bar{X}) = 0$$

Let $\bar{Y}$ denote the arithmetic of the sample values $\{Y_1, Y_2, Y_3, ..., Y_N\}$. Show that

$$\sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y}) = \sum (X_i - \bar{X})Y_i$$

In addition to this problem solve the following exercises from the textbook:
(2.7), (2.8) and (2.10)

Problem set 2.

Exercise 2.1

Consider two random variables X *and* Y. Suppose that Y *takes on k values* $y_1, ..., y_k$ and that X *takes on l values* $x_1, ..., x_l.$

(a) Show that $\Pr(Y = y_j) = \sum_{i=1}^{l} \Pr(Y = y_j \mid X = x_i) \Pr(X = x_i)$

(b) Show that $E(Y) = E(E(Y \mid X))$

(c) Suppose that X *and* Y are independent. Show then that the covariance between X *and* Y *are zero.*

Exercise 2.2

The simple regression model is usually written on the form

(1) $$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad i = 1,2,...,N$$

However, some authors prefer to write it on form

(2) $$Y_i = \alpha + \beta_1 (X_i - \overline{X}) + \varepsilon_i \quad i = 1,2,...,N$$

where $\alpha = \beta_0 + \beta_1 \overline{X}$

(a) Show that the least square estimates of β_1 and α and are given by

$$\hat{\alpha} = \overline{Y}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{N} (X_i - \overline{X}) Y_i}{\sum_{i=1}^{N} (X_i - \overline{X})^2}$$

(b) Show that $\hat{\alpha}$ *and* $\hat{\beta}_1$ are uncorrelated.

© Discuss the Gauss-Markov theorem

Problem set 3. The exercises here ask you to run simple regressions and perform testing of hypotheses.

Solve the exercises: (3.5), (3.7) and (3.12)

Problem set 4. The exercises here are concerned about 'testing' if the disturbances in the regression equations are normally distributed and applying the regression model for doing predictions.

Solve the exercises: (4.8), (4.12) and (4.14)

Problem set 5. The step from simple to multiple regression is in many ways immediate. The problems listed below is meant to give you some training in using multiple regression.

Solve the exercises: (5.10), (5.12) and (5.15)