ECON 3150/4150 spring term 2014: Exercise set for the first seminar and DIY exercises for the first few weeks of the course

Ragnar Nymoen

13 January 2014.

Exercise set to seminar 1 (week 6, 3-7 Feb)

This first exercise set is too long for a single seminar. Therefore, it is only Question A which will be discussed at the seminar meeting. The other exercises is for your own work (alone or together with others) while you wait for the seminars to begin in calendar week 6. Answers to the other exercises will be given during the first six lectures, or they will be posted on the web page of ECON 4150. Note about this will be given!

Question A

- Stock and Watson Chapter 2: 2.4; 2.8; 2.17; 2.24;2.26
- Stock and Watson Chapter 3: 3.17;3.20; 3.21

As noted, the rest are DIY exercises (for calendar week 3,4 and 5)

Question B

1. Consider the three stochastic variables X, Y and Z. The three variables are connected by the linear function

$$Y = a + bX + Z$$

where a and b are parameters. Assume that $E(Z) = 0$ and $Cov(X, Z) = 0$ and show that the parameter b can be written as

$$b = \frac{Cov(X, Y)}{Var(X)}$$

Note: This is an example of a population regression. An interpretation of the regression model with stochastic regressor in econometrics is that is a method that by conditioning allows us to obtain probabilistic knowledge about the population slope parameter, written as b here, without having prior knowledge about $Cov(X, Y)$ and $Var(X)$

Table 2: Conditional distribution for Y given X, based on Table 1.

		X		
		-8	0	8
Y	-2	$\frac{0.1}{0.1}$	$\frac{0.5}{0.7}$	$\frac{0.1}{0.2}$
	6	$\frac{0}{0.1}$	$\frac{0.2}{0.7}$	$\frac{0.1}{0.2}$

2. Assume that X and Y are two stochastic variables. The law of iterated expectations (also called law of double expectation) states that

$$E\left[E(Y \mid X)\right] = E(Y) \tag{1}$$

   (a) Interpret the two expectations operators on the left hand side of equality sign.

   (b) Try to prove (1) for the case where X and Y are discrete stochastic variables.

3. Consider the two stochastic variables X and Y. The conditional expectation $E\left(Y \mid x\right)$ is deterministic for any given value of $X = x$. But we can also regard the expectation of Y for all possible values of X. In this interpretation $E\left(Y \mid X\right)$ is a stochastic variable with realization $E\left(Y \mid x\right)$ when $X = x$. Hence, the conditional expectation function $E\left(Y \mid X\right)$ is a function of the stochastic variable X so that we can write $E\left(Y \mid X\right) = g_X\left(X\right)$ where $g_X(X)$ is a function.

   Consider the discrete distribution function in table 1.

Table 1: A discrete probability distribution for X and Y

		X			
		-8	0	8	$f_Y(y_i)$
	-2	0.1	0.5	0.1	0.7
Y	6	0	0.2	0.1	0.3
	$f_X(x_i)$	0.1	0.7	0.2	

   Note that the marginal probabilities are in the last row (X) and in the column on the right right (Y) of the table.

   (a) Obtain the conditional distribution for Y shown in table 2.

   (b) Show that the conditional expectation function $E(Y \mid X)$ becomes:

$$E(Y \mid X = -8) = -2$$
$$E(Y \mid X = 0) = \frac{2}{7}$$
$$E(Y \mid X = 8) = 2$$

(c) Calculate $E\left[E(Y \mid X)\right]$ by using the law of iterated expectation, i.e., by summing the three conditional expectations multiplied by the marginal probabilities $f_x(x_i)$.

(d) Use the marginal distribution for Y to show that $E(Y) = E\left[E(Y \mid X)\right]$, as a check of the law of iterated expectations.

4. Consider the stochastic variables X and Y and the conditional expectations function $E(Y \mid X)$. Let the stochastic variable ε be implicitly defined by the equation

$$Y = E(Y \mid X) + \varepsilon$$

Show that $E(\varepsilon \mid X) = 0$.

We know that $Cov(X, \varepsilon) = E(X\varepsilon) = 0$ is a consequence of $E(\varepsilon \mid X) = 0$. Assume that $E(Y \mid X)$ is linear: $E(Y \mid X) = \beta_0 + \beta_1 X$. What are $E(\varepsilon \mid X)$ and $Cov(X, \varepsilon)$ for the stochastic variable ε defined by

$$Y = \beta_0 + \beta_1 X + \varepsilon?$$

Question C

1. For a given set of data observations x_i $(i = 1, 2, \ldots, n)$, and y_i $(i = 1, 2, \ldots, n)$, show that:

$$\sum_{i=1}^{n} (x_i - \overline{x}) = 0 \tag{2}$$

$$\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y}) = \sum_{i=1}^{n} (x_i - \overline{x}) y_i \tag{3}$$

$$\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y}) = \sum_{i=1}^{n} x_i (y_i - \overline{y}) \tag{4}$$

$$\sum_{i=1}^{n} (x_i - \overline{x})^2 = \sum_{i=1}^{n} x_i (x_i - \overline{x}) \tag{5}$$

where $\overline{x}$ and $\overline{y}$ denote the arithmetic means.

2. The least squares estimation principle entails that we chose the estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ that minimize the following criterion function;

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 \tag{6}$$

where $\{x_i, y_i;\ i = 1, 2, \ldots, n\}$ are interpreted as given numbers (data).

Show that the first order conditions for a minimum of $S(\beta_0, \beta_1)$ are:

$$\overline{y} - \hat{\beta}_0 - \hat{\beta}_1 \overline{x} = 0 \tag{7}$$

$$\sum_{i=1}^{n} x_i y_i - \hat{\beta}_0 \sum_{i=1}^{n} x_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0. \tag{8}$$

3. Solve (7) and (8) for $\hat{\beta}_0$ and $\hat{\beta}_1$.

4. Show that the alternative formulation of the criterion function

$$S(\alpha, \beta_1) = \sum_{i=1}^{n} (y_i - \alpha - \beta_1(x_i - \overline{x}))^2 \tag{9}$$

   gives rise to the same OLS estimate $\hat{\beta}_1$ as in you answer to (3) above.

5. What is the algebraic relationship between the two OLS estimates $\hat{\beta}_0$ and $\hat{\alpha}$?

6. Assume that the empirical correlation coefficient between x and y, $r_{X,Y}$, is 0.5. Does this imply that the regression coefficient $\hat{\beta}_1$ in the regression where X is the regressor is positive? Explain.

7. Consider the ("inverse") regression where X is the regressand and Y is the regressor. What is the expression for the OLS estimate of the slope coefficient (call it $\hat{\beta}_1'$) in that regression? Are the two regression coefficients the same?

8. Assume that you run the two regressions on two sub-samples with time series data and that you observe the following for the OLS estimate $\hat{\beta}_1$ (in the regression with X as the regressor) and the squared empirical correlation coefficient:

	$\hat{\beta}_1$	$r^2_{X,Y}$
First sample	0.5	$0.5^2 = 0.25$
Second sample	0.5	$0.9^2 = 0.81$

   What are the two estimates for $\hat{\beta}_1'$ in the inverse regression (where Y is the regressor)?

9. How does the invariance of $\hat{\beta}_1$ with respect to the break in the correlation structure between the two sample periods affect your thinking about a possible causal relationship between X and Y?

Question D

Assume that Y_i is explained by a single dummy (or indicator) variable X_i that takes either the value 0 or 1. For concreteness think of Y_i as the number of recorded damages to the eye during New Year celebrations in year i, and define X_i as

$$X_i = \begin{cases} 1, & \text{if year } i \text{ is with the law against rockets in private fireworks} \\ 0, & \text{if not} \end{cases}$$

1. What are the algebraic expressions for the OLS estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ in this case?

2. In Norway, the use of rockets in private fireworks became illegal in 2008. Before this policy intervention the number of damages during New Year festivities averaged 20. The numbers of damages during the New Year celebrations after rockets were prohibited have been: 10, 10,17,16,17,15 (2013 celebration) What is the estimated values of $\hat{\beta}_0$ and $\hat{\beta}_1$ based on these data?[1]

3. Using the model above: What is your predicted number of damages for the 2014 New Year celebration in Norway? What do you regard to be the main sources of forecast uncertainty in this case? (No calculations expected/required!)

Question E

Define the OLS fitted values for Y_i by

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i, \quad i = 1, 2, \ldots, n$$

where $\hat{\beta}_0$ and $\hat{\beta}_1$ are the OLS estimates given by (7) and (8) above.

1. Show that, equivalently, $\hat{Y}_i$ can be defined by

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta}_1 (X_i - \overline{X}), \quad i = 1, 2, \ldots, n$$

2. Define the OLS residuals by

$$\hat{\varepsilon}_i = Y_i - \widehat{Y}_i,\ i = 1, 2, ..., n \tag{10}$$

and show that

$$\overline{\hat{\varepsilon}} = \frac{1}{n} \sum_{i=1}^{n} \hat{\varepsilon}_i = 0 \tag{11}$$

$$\frac{1}{n} \sum_{i=1}^{n} (\hat{\varepsilon}_i - \overline{\hat{\varepsilon}})(X_i - \overline{X}) = 0 \tag{12}$$

$$\overline{\hat{Y}} = \frac{1}{n} \sum_{i=1}^{n} \hat{Y}_i = \overline{Y} \tag{13}$$

3. Explain intuitively why the following decomposition of the total variation in Y (the regressand) is true:

$$\underbrace{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}_{\text{total variation}} = \underbrace{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}_{\text{residual variation}} + \underbrace{\sum_{i=1}^{n} (\widehat{Y}_i - \overline{\hat{Y}})^2}_{\text{explained variation}} . \tag{14}$$

4. What is the relationship between the empirical correlation coefficient $r_{X,Y}$ between X and Y and the coefficient of determination (R^2)?

[1] http://www.helse-bergen.no/aktuelt/nyheter/Sider/17-personar-med-alvorlege-augeskadar–.aspx

5. Do (11), (12) and (14) hold if the fitted values are instead from an OLS estimated linear relationship with no intercept (i.e. $\hat{\beta}_0$ is dropped)? What does this imply for the conventional R^2 statistic?

6. If we scale the variables Y_i and X_i in the way explained in Lecture 2, prior to OLS estimation. Will R^2 be the same as in the "unscaled case"?

Question F

1. Assume that Y_i and X_i are n random variables, $i = 1, 2, \ldots, n$. Assume that the conditional expectations function $E(Y_i \mid X_i)$ is linear: $E(Y_i \mid X_i) = \beta_0 + \beta_1 X_i$. Show that the (OLS) estimator $\hat{\beta}_1$ is unbiased:

$$E(\hat{\beta}_1 - \beta_1) = 0.$$

THE REGULAR EXERCISES FOR SEMINAR 2-10 WILL BE POSTED SEPARATELY.