Stockholm University
Department of Statistics
Ellinor Fackle Fornius, Jessica Franzén

HT 2008

EXAM IN ECONOMETRICS, PART I
2008-12-03

Exam hours: 9.00-14.00
Calculator and one book of your choice are allowed.
The exam consists of five problems which are worth 10 points each. For the maximum of 10 points detailed and clear solutions are required.

Problem 1. (10 points)

The following table shows the regression output, with some numbers erased, when a simple regression model relating the response variable Y to a predictor variable X is fitted based on 20 observations. Complete the 12 missing numbers, then compute $Var(Y)$ and $Var(X)$.

ANOVA Table				
Source	Sum of Squares	df	Mean Square	F-test
Regression	1848.76	?	?	?
Residuals	?	?	?	
Coefficients Table				
Variable	Coefficient	$se\left(\widehat{\beta}\right)$	t-test	p-value
Constant	-23.43	12.74	?	0.0824
X	?	0.15	8.32	< 0.0001
$n =?$	$R^2 =?$	$\overline{R}^2 =?$	$\widehat{\sigma} =?$	

Problem 2. (10 points)

Describe one formal test that can be used to evaluate each of the following assumptions. For each test state the null and alternative hypotheses and the test statistic. In addition, specify when to reject the null hypothesis for each test.

a) The model is correctly specified
b) The error terms have constant variance
c) The error terms are uncorrelated
d) The error terms are normally distributed
e) There is not severe multicollinearity

Problem 3. (10 points)

a) Explain the term autocorrelation and why we must be cautious in the presence of autocorrelation.

b) A researcher suspects that the residuals in his model might be autocorrelated. He computes the Durbin-Watson test statistic and obtains a value of 0.86. The regression is based on 80 observations and has 3 explanatory variables (plus an intercept term). Perform the Durbin-Watson test. Is there evidence of positive or negative autocorrelation? What is your conclusion?

c) Can the Durbin-Watson test be used to detect autocorrelation in the following models? If your answer is no, explain why.

$$
\begin{aligned}
(1) \qquad Y_t &= \beta_1 X_{1,t} + \beta_2 X_{2,t} + \beta_3 X_{3,t} + \beta_4 X_{4,t} + u_t \\[1em]
(2) \qquad Y_t &= \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \beta_3 X_{3,t} + \beta_4 X_{4,t} + v_t \\
\text{where } v_t &= \rho_1 v_{t-1} + \rho_2 v_{t-2} + \rho_3 v_{t-3} + \varepsilon_t \\[1em]
(3) \qquad Y_t &= \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \beta_3 X_{2,t-1} + \varepsilon_t \\[1em]
(4) \qquad Y_t &= \beta_1 X_{1,t} + \beta_2 X_{2,t} + \beta_3 X_{3,t} + \beta_4 X_{4,t} + u_t \\
\text{where } u_t &\sim N\left(0, \frac{\sigma^2}{\pi^3}\right)
\end{aligned}
$$

Problem 4. (10 points)

Consider

$$\begin{aligned} \varepsilon_t &= u_t - \rho_1 u_{t-1} - \rho_2 u_{t-2} \\ &\quad \text{where} \\ E(u_t) &= 0 \qquad \text{for all } t \\ Var(u_t) &= \sigma^2 \qquad \text{for all } t \\ E(u_t u_s) &= 0 \qquad \text{for all } t \neq s \end{aligned}$$

What are $E(\varepsilon_t)$ and $Var(\varepsilon_t)$?

Problem 5. (10 points)

a) Let the response variable Y be a binary variable where

$$Y = \begin{cases} 1 & \text{denotes success} \\ 0 & \text{denotes failure} \end{cases}$$

The following logistic regression model is fitted

$$\ln\left(\frac{P_i}{1-P_i}\right) = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i}$$

where

$$P_i = P(Y_i = 1).$$

The resulting output is given in the table:

Variable	Coefficient	Standard error	p-value	Odds-ratio
Constant	-0.550	0.951	0.563	
X_2	0.157	0.075	0.036	1.17
X_3	0.195	0.122	0.112	1.21
Likelihood Ratio (LR) Statistic:	$LR = 82.024$	$df = 2$		

Interpret the regression coefficient $\widehat{\beta}_2$ and the odds-ratio $e^{\widehat{\beta}_2}$.

b) Test the overall fit of the model, that is test the null hypothesis that β_2 and β_3 are both equal to zero. What is your conclusion?

c) Consider estimating the logistic regression model

$$\widehat{P}_i = \frac{1}{1 + e^{-\left(\widehat{\beta}_1 + \widehat{\beta}_2 X_i\right)}}$$

for the following data:

x_i	-2	1	4	9
y_i	1	0	1	0

You are given two alternatives for the estimates of the regression coefficients β_1 and β_2 as follows

$$\textbf{Alternative A} \quad : \qquad \widehat{\beta}_1^A = 0.1 \quad \text{and} \quad \widehat{\beta}_2^A = 0.2$$

$$\textbf{Alternative B} \quad : \qquad \widehat{\beta}_1^B = 0.8 \quad \text{and} \quad \widehat{\beta}_2^B = -0.3$$

Which of the two alternatives A and B gives the best parameter estimates?
Hint: Use the maximum likelihood method.