M. Sc. Program in International Economics and Finance
Otto-von-Guericke University, Magdeburg
Prof. Dr. Karl-Heinz Paqué

ADVANCED ECONOMETRICS: A GUIDED TOUR
(20308 Econometrics)

Retake Exam

Winter term 2011/12 - July 27, 2012

Antonio Fidalgo, Ph.D. candidate
Faculty of Business and Economics
University of Lausanne
Lausanne, Switzerland

Instructions

- No textbook, lecture note or any other piece of documentation is permitted. A dictionary for language issues is allowed.
- This exam contains 7 problems. You must solve all of them.
- Each problem is worth 14 points except for problem 4 which has 16 points. Hence, you can get a maximum of 100 points.
- For your answers, use exclusively the working sheets provided. Nothing you would write on these problems' sheets will be considered in the grading.
- The duration of the exam is 120 minutes ($\approx$ 15 minutes per problem). Allocate your time wisely.

Good luck!

Problem 1. Maximum likelihood estimate of the parameters of a normal distribution

Suppose you have a sample $x_1, x_2, .., x_n$ of n independent observations, all drawn from what you believe is a normal distribution. Remember that the density function for a normal $\mathcal{N}(\mu, \sigma^2)$ distribution is given by:

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\ \sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right). \tag{1}$$

Required. Write down $L(\mu, \sigma^2) = f(x_1, ..., x_n|\mu, \sigma^2)$, the log-likelihood function for that sample. Derive $\hat{\mu}^{MLE}$, the maximum likelihood estimate of μ to show

$$\hat{\mu}^{MLE} = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{2}$$

Problem 2. Common cause

Suppose you have a thousand samples of hundred observations on three variables, x, y and z, this latter being drawn from a $\mathcal{N}(0, 1)$ distribution. You know that the following relationships hold in reality

$$x = 20 + \frac{1}{15}z + \theta u \quad \text{and} \quad y = 80 + \frac{1}{3}z + v \tag{3}$$

where u and v are random $\mathcal{N}(0, 1)$ shocks and θ is a scalar of interest.

Required. Suppose you estimate by OLS the following model on the 1000 samples:

$$y = \alpha + \beta x + \varepsilon \tag{4}$$

For three relevant values of θ –two limiting cases and a standard one, give the resulting sampling distributions of β. Explain.

Problem 3. Uniformit model

In a model for a probability, we've seen that, in principle, any continuous probability distribution over the real line would make the job. The normal distribution gave rise to the *probit* model, the logistic function to the *logit* model, the distribution considered by Burr to the *burrit*, the Gompertz curve to the *gompit* (see Maddala (1983) *Limited-Dependent and Qualitative Variables in Econometrics*, pp. 31-32 for those examples), and so on...

Required. Explain with sufficient detail the case we invent here of the *uniformit*, a probability model based on the uniform distribution.

Problem 4. Dummies fit and quadratic fit

Thanks to the Canadian cps71 dataset, we were able to estimate an earnings profile over age: remember, we had information for the hourly log-wage of an individual along with his/her age.
We now build five dummies for the continuous variable *age*, $D0 - D4$, for the age groups <30, 30-39, 40-49, 50-59 and >60 respectively.
Then we estimate two models. Model 1 links *logwage* to the dummies and model 2 links *logwage* to the variable *age* and the square of that variable (quadratic model). The table below shows the results of the two OLS regressions.

Required. For each model, draw as precisely as possible —you might want to write values on the graph– the resulting fitted lines implied by the estimations. For each model, give the predicted age for which the log-wage is the highest.

Table 1: Alternative models for the earnings profile; dependent variable *logwage* ($N = 205$)

Indep. Variable	Model 1	Model 2
Age	.	.17***[7.26]
Age^2	.	-.002***[-6.82]
D1	.57***[5.13]	.
D2	.47***[4.11]	.
D3	.60***[4.96]	.
D4	.02[0.11]	.
Constant	13.13***[174.01]	10.04***[22.02]
R^2	0.1725	0.2308

*,** and *** for $p < .05, .01$ and $.001$ resp.
t-values in brackets next to coefficients

Problem 5. Nonsmooth, nonparametric estimator

Suppose you want to get an estimate, $\hat{f}(x)$, of a univariate density function $f(x)$ for which you neither have any prior knowledge nor are you inclined to assume a particular form.

Required. Explain and illustrate with graphs how you would proceed if you were to use a nonsmooth, nonparametric estimator. Explain what is the key "parameter" of such an estimation and underline the relevance of the mean square error (MSE) value as a criterion to make the optimal choice.

Problem 6. Diagoras the Atheist's insight on the mis-practice of econometrics

In thanks to the sea god Neptune, Greek sailors who escaped from shipwrecks or were saved from drowning at sea used to display pictures of themselves in a votive temple on the Aegean island of Samothrace. While visiting that temple, Diagoras of Melos (500 BC) was once challenged by a believer:

> "Advise now, you that think it folly to invocate Neptune in tempest. [...] Surely these portraits are proof that the gods really do intervene in human affairs".

To which Diagoras replied:

> "Yea, but... where are they painted that are drowned?"

Required. Considering that reply in the broad context of econometrics, of which concept could we say that Diagoras is the father? Explain.

Problem 7. White better than White

Remember that White's (1980) "A Heteroskedasticity-Consistent Covariance-Matrix Estimator and a Direct Test for Heteroskedasticity" (*Econometrica*, 48) is the most cited paper in economics and, hence, his estimator is surely the most used estimator in presence heteroskedasticity.
But then, we can read in Davidson and McKinnon (1993) (*Estimation and Inference in Econometrics*, p. 554):

> "One should *never* use [the White estimator] because [another estimator presented] *always* performs better."

A similar conclusion is reached by White himself in his joint work with McKinnon ("Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved XXX", *Journal of Econometrics*, 29, 1985)

> "We have examined the performance of three modified versions of White's (1980) heteroskedasticity-consistent covariance matrix estimator. All of them can be thought of as in some way derived from the jackknife, and the one which is explicitly the jackknife covariance estimator, *HC3*, always performs better than the other two, which in turn always outperform the original."

Required. Explain the room for improvement of the original White estimator. Notice that the "XXX" above gives the key to this problem.