**MASTER 1**

**ECONOMETRICS**
*(durée* 3h00

**Wednesday, January 9th 2013 ~ 09h00 -12h00**

C. BONTEMPS
P. LAVERGNE

*****************

**EXERCISE 1.** Let $Y_1, \ldots, Y_n$ be a random sample from $Y$, which is distributed as a log-normal variable, that is $\log Y$ is normal with mean $\mu$ et variance $\sigma^2$. The density of $Y$ is given by

$$f(y; \mu, \sigma^2) = \left(2\pi\sigma^2\right)^{-1/2} y^{-1} \exp\left( - \frac{(\log y - \mu)^2}{\sigma^2} \right).$$

1. Write the likelihood of the sample.
2. Write the first order conditions (F.O.C.) of the maximization procedure and deduce the maximum likelihood estimators $\widehat{\mu}$ and $\widehat{\sigma}^2$ for $\mu$ and $\sigma^2$.
3. Compute the matrix of second-order derivatives

$$\begin{bmatrix} \frac{\partial^2}{\partial \mu^2} \log f(y; \mu, \sigma^2) & \frac{\partial^2}{\partial \mu \partial \sigma^2} \log f(y; \mu, \sigma^2) \\ \frac{\partial^2}{\partial \mu \partial \sigma^2} \log f(y; \mu, \sigma^2) & \frac{\partial^2}{\partial (\sigma^2)^2} \log f(y; \mu, \sigma^2) \end{bmatrix}.$$

   From the previous result, determine the information matrix.
4. What is the joint asymptotic distribution of the maximum likelihood estimators? *(A statement is sufficient for this question, no proof is expected)*
5. We want to test $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$. What is the Wald statistic for the test? What is its asymptotic distribution? Give the rejection rule of the test. *(Statements are sufficient for this question, no proof is expected)*
6. Show that the test is convergent, i.e. if $\mu \neq \mu_0$, the test rejects $H_0$ with probability tending to 1.

**EXERCISE 2.** Consider the heteroscedastic regression model

$$y_i = x_i\beta + \varepsilon_i \qquad \mathrm{E}(\varepsilon_i|x_i) = 0,\ \mathrm{Var}(\varepsilon_i|x_i) = \sigma_i^2 \qquad i = 1, \ldots n,$$

where $y_i$ and $x_i$ are scalar, and observations are independent, and $n^{-2}\sum_{i=1}^n x_i^6\sigma_i^2 \to 0$. Consider the estimator

$$\widehat{\beta} = \frac{n^{-1}\sum_{i=1}^n x_i^3 y_i}{n^{-1}\sum_{i=1}^n x_i^4}.$$

We denote by $\boldsymbol{x}$ the vector of observations $x_i$.

1. Find E $\left(\widehat{\beta}|x\right)$.
2. Find Var $\left(\widehat{\beta}|x\right)$.
3. Show that $\widehat{\beta}$ is consistent for the true $\beta$.
4. In which situation do you expect this estimator to be efficient? Explain. *(No proof is expected)*

**EXERCISE 3.** In this exercise, we try to explain the decision of mortgage denial. The dependent variable $y_i$ is equal to 1 if the mortgage application is denied, 0 if it is accepted. The first two explanatory variables are direct measures of the financial burden the proposed loan would place on the applicant measured in terms of his/her income. The first is **P/I ratio**, i.e. the ratio of total monthly debt payment to total monthly income. The second is **H/I ratio**, i.e. the ratio of housing-related expenses to income. We then have three dummy variables. The variable **black** is equal to 1 if the application is black, 0 otherwise, **self** is 1 if the applicant is self-employed, 0 otherwise, and **hsdipl** is equal to 1 if the applicant is graduated from high-school, 0 otherwise. value of the parameters.

Table 1 presents the results of the estimation of a logit model based on these variables (we also report the log-likelihood of the full sample evaluated at $\hat{\beta}_{MLE}$).

Table 1: Logit estimation of Mortgage Denial

| Regressor | Estim. | Std. error |
|---|---|---|
| intercept | -2.71 | 0.48 |
| P/I ratio | 4.76 | 1.33 |
| H/I ratio | -0.11 | 1.29 |
| black | 0.69 | 0.18 |
| self | 0.55 | 0.19 |
| hsdipl | -0.93 | 0.36 |
| Log-lik. sample | -64.1 | |

1. For this question, we denote by $x_i$ the vector of explanatory variables for observation $i$ and $\beta^0$ the true value of the parameters. Explain why we use a logit model instead of a linear model. Give the theoretical expression of $P[y_i = 1|x_i]$.

2. Which variables are significant ? Comment on the signs of the estimated parameters corresponding to these significant variables (except the intercept).

3. Calculate the expected probability of mortgage denial for a white applicant, with high school diploma, who is not self-employed, with **H/I ratio=0.2** and **P/I ratio=0.3**.

4. In which interval does $P[y_i = 1|x_i]$ vary for an individual whose characteristics are the ones of the previous question but P/I ratio varies between 0 and 1 ?

5. The econometrician estimates again a similar logit model adding a supplementary dummy variable **single** (equal to one is the applicant is single, 0 otherwise). The log-likelihood calculated at the (new) MLE is equal to -61.83. We want to test the significance of the variable **single**.

   (a) Write **explicitly** the null hypothesis, the alternative hypothesis, the test statistic you use and its (asymptotic) behavior under the null and the alternative.

   (b) When an econometrician does hypothesis testing, what is the level of the test ?

   (c) Conclude about the test of significance for **single** after having chosen your level equal to 5%.

**EXERCISE 4.** The goal of the following regression is to estimate the elasticity of the demand for cigarettes using data from 48 U.S. States.

$$\log(Q_i^{cig}) = \beta_1 + \beta_2 \log(P_i^{cig}) + \beta_3 \log(Inc_i) + u_i,$$

where $Q_i^{cig}$ is the quantity of packs sold per capita in the state in a given year, $P_i^{cig}$ the average price per pack including all taxes in dollars, $Inc_i$ the real per capita state income (in thousands of dollars). We assume that the asymptotic theory is valid for this sample size.

1. Explain briefly why a OLS regression does not provide a consistent estimator of $\beta_2$.

2. We observe the variable SaleTax which corresponds to the portion of the tax on cigarettes arising from the general sales tax and measured in dollars per pack. Explain what is required for an instrument to be valid and then, why SaleTax is a valid one for instrumenting the variable $P_i^{cig}$.

3. Derive the expression of a consistent estimator of $\beta = [\beta_1, \beta_2, \beta_3]^\top$ as a function of the usual matrices $X$ (the $n \times 3$ that stacks the observations $x_i$), $Y$, the vector of observations of the dependent variable and another matrix that would be determined.

4. Two-stage least squares estimation yields

$$\log(Q_i^{cig}) = \underset{(1.26)}{4.20} - \underset{(0.31)}{1.14}\widehat{\log(P_i^{cig})} + \underset{(0.37)}{0.21}\log(Inc_i).$$

The standard errors (corrected to account for the two-stage procedure) are in parenthesis below the estimated values. Give a 95% confidence interval for the price elasticity of the demand of cigarettes.

5. Assume the population of the state is equal to one million. The price of a pack (all taxes included) is 2 dollars. The average income per capita is 20 thousands dollars.

   (a) Following the results, estimate the number of packs sold in the state in a given year.

   (b) Calculate the annual revenue generated by a tax increase of 0.1 dollar per pack (don't forget that the equilibrium changes !).