PhD in Economics
1st year Econometrics test (2018-05-14)

Name: ____

1. Say if the following statements are unambiguously true (TRUE), unambiguously false (FALSE) or impossible to classify the way they are stated (NOT NECESSARILY). Write the motivations to your answers **only** in the space provided. A "Not necessarily" answer with no motivations will be considered wrong.

   (a) If X is a random variable whose support is the $[-1, 1]$ interval and $E(X) = 0$, then $1 < E(e^X) < e$.
   True ◯ False ◯ Not necessarily ◯

   (b) Suppose you have a sequence of random variables such that $Y_n \sim N(0, 1/n)$; then, $Z_n = e^{Y_n} \xrightarrow{p} 0$.
   True ◯ False ◯ Not necessarily ◯

   (c) Maximum likelihood estimators are unbiased.
   True ◯ False ◯ Not necessarily ◯

   (d) GMM estimators are unbiased.
   True ◯ False ◯ Not necessarily ◯

   (e) When estimating a model via ML, a Wald test for the restriction $H_0 : g(\theta) = 0$ can be computed by using only the unrestricted estimator $\hat{\theta}$, while the restricted estimator $\tilde{\theta}$ is not necessary.
   True ◯ False ◯ Not necessarily ◯

2. You have a sample of 807 individuals living in the US and containing information on the number of cigarettes smoked per day (*cigs*), and the following covariates:

Variable	Description
lnprice	average price of cigarettes in the state of residence (in logs)
lninc	income (in logs)
resta	dummy; 1 if smoking in restaurants is banned in the country of residence
educ	years of education

Table 1 contains a few descriptive statistics.

Table 1: Descriptive statistics of the dependent variable and of the covariates

	Mean	Median	S.D.	Min	Max	1st quartile	3rd quartile
cigs	8.686	0.000	13.72	0.000	80.00	0.000	20.000
lcigpric	4.096	4.112	0.08292	3.784	4.250	4.063	4.146
lincome	9.687	9.903	0.7127	6.215	10.31	9.433	10.309
educ	12.47	12.00	3.057	6.000	18.00	10.000	13.500
restaurn	0.2466	0.000	0.4313	0.000	1.000	0.000	0.000

(a) In the rest of the exercise, a Poisson regression model will be used to understand the impact of these covariates on the number of cigarettes smoked per day. However, an OLS regression could have been used with no substantial differences. Do you agree with this statement? Motivate your answer in a separate sheet.

Agree ○ Disagree ○

(b) Table 2 displays the estimation results obtained by Poisson Maximum Likelihood (ML) and Poisson Quasi ML. At the bottom of the table, an estimate of the variance-mean ratio of the dependent variable conditional on covariates, $\frac{V(cigs|\mathbf{x})}{E(cigs|\mathbf{x})}$, is reported. How would you estimate the variance-mean ratio? Write your answer in the space below:

(c) On the basis of the reported estimate of the variance-mean ratio, would you make inference with the standard errors from the Poisson ML or with those from the Poisson Quasi ML? Motivate your answer in a separate sheet.

ML ○ Quasi ML ○

(d) Are cigarettes an inferior, a normal, or a luxury good? More in detail, what is the percentage variation in the number of cigarettes smoked per day if income increases by 1%?

Inferior ○ normal ○ luxury ○

$$\frac{\Delta Y}{Y} = 0.01 \Longrightarrow \frac{\Delta cig}{cig} = \_\_\_\_\_\_$$

(e) Estimate the relative variation in the consumption of cigarettes induced by restaurant smoking bans. Is it significantly different from zero at the usual 5% level?

$$Ban = 0 \rightarrow Ban = 1 \Longrightarrow \frac{\Delta cig}{cig} = \quad \_\_\_\_\_\_\_\_$$

Significant ◯ Not significant ◯

Table 2: Poisson ML and Poisson Quasi ML estimates of the number of cigarettes smoked per day

```
Model 1: Poisson ML, using observations 1-807
Dependent variable: cigs

              coefficient   std. err.  z-ratio   p-value
-----------------------------------------------------------
constant        0.8122817   0.5996325    1.35  0.176
lnprice        -0.1039017   0.1415866   -0.73  0.463
lninc           0.2409601   0.0198813   12.12  0.000 ***
resta          -0.3662369   0.0309915 -11.82  0.000 ***
educ           -0.0396276   0.0041741   -9.49  0.000 ***

Mean dependent var    8.686493   Sum squared resid     148927.2
Log-likelihood       -8479.165   McFadden R-squared    0.019281
Akaike criterion      16968.33   Variance-Mean ratio   21.43976

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

Model 2: Poisson Quasi ML, using observations 1-807
Dependent variable: cigs

              coefficient   std. err.  z-ratio   p-value
-----------------------------------------------------------
constant        0.8122817   2.8948590    0.28  0.779
lnprice        -0.1039017   0.6638025   -0.16  0.876
lninc           0.2409601   0.0833286    2.89  0.004 ***
resta          -0.3662369   0.1425355   -2.57  0.010 **
educ           -0.0396276   0.0189857   -2.09  0.037 **

Mean dependent var    8.686493   Sum squared resid     148927.2
Log-likelihood       -8479.165   McFadden R-squared    0.019281
Akaike criterion      16968.33   Variance-Mean ratio   21.43976
```

3. Consider the following DGP

$$y_t = \mu + \phi y_{t-1} + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \varepsilon_t, \tag{1}$$

where ε_t is a white noise process with unit variance.

(a) Suppose that x_t is an unobservable white noise process with variance 1, independent of ε_t at all lags and define $u_t = \beta_1 x_{t-1} + \beta_2 x_{t-2} + \varepsilon_t$. Calculate the first three autocovariances of u_t.

$\gamma_0 =$ ________ $\gamma_1 =$ ________ $\gamma_2 =$ ________

(b) prove that u_t is a MA(1) process in a separate sheet. *(Hint: consider the Wold representation of u_t)*

(c) Now assume, instead, that x_t is an observable process; the results of an OLS regression are provided in table 3. Write the ECM representation of the ADL(2,1) using the estimates in table 3.

$\Delta y_t =$

(d) Test the joint hypothesis $\beta_1 = 0, \beta_2 = 0$ using the estimates in table 3.

Test: ______________ Distribution: ____________ Test stat: ________

Decisione: DON'T REJECT ◯ REJECT ◯

(e) Write the estimated values for the following multipliers, using the estimates in table 3.

$$\frac{\partial y_t}{\partial x_t} = \_\_\_\_\_\_ \qquad \frac{\partial y_t}{\partial x_{t-1}} = \_\_\_\_\_\_ \qquad \frac{\partial y_t}{\partial x_{t-2}} = \_\_\_\_\_\_$$

(f) Now consider a bivariate VAR(2) in which the first equation is (1) and the second one is $x_t = \theta_1 x_{t-1} + u_t$, where $u_t \sim WN(0,1)$ is not correlated with ε_t. Write the companion matrix $\mathbf{A}$ for such a system as a function of ϕ, β_1, β_2 and θ.

$$\mathbf{A} = \left[\begin{array}{c|c|c|c} \_\_\_ & \_\_\_ & \_\_\_ & \_\_\_ \\ \_\_\_ & \_\_\_ & \_\_\_ & \_\_\_ \\ \_\_\_ & \_\_\_ & \_\_\_ & \_\_\_ \\ \_\_\_ & \_\_\_ & \_\_\_ & \_\_\_ \end{array} \right]$$

Table 3: OLS regression

	Coefficient	Std. Error	t-ratio	p-value
const	0.574	0.118	4.880	0.0000
y_{t-1}	0.350	0.090	3.899	0.0002
x_{t-1}	0.540	0.096	5.648	0.0000
x_{t-2}	−0.208	0.109	−1.919	0.0577

Parameter covariance matrix:

$$\hat{V} = 10^{-4} \times \begin{bmatrix} 138.13 & -69.43 & -0.42 & 34.28 \\ -69.43 & 80.79 & -3.24 & -44.85 \\ -0.42 & -3.24 & 91.41 & 25.44 \\ 34.28 & -44.85 & 25.44 & 117.96 \end{bmatrix}$$