#34 85

Econometrics
Comprehensive Exam
August 2002

Yoon/Basmann/Kumbhakar

Answer four of the following six questions. Justify your answers making appropriate additional assumptions if necessary.

1. Consider the following two equation model:

(i) $Y_1 + \gamma_{11} X_1 = u_1$
(ii) $\beta_{21} Y_1 + Y_2 + \gamma_{21} X_1 = u_2$

where $\text{cov}(u_1, u_2) = 0$. Y_1 and Y_2 are endogenous variables and X_1 is an exogenous variable; u_1 and u_2 are disturbance terms. Check the identifiability of (i) and (ii).

2. Let m_h' denote the sample moment and μ_h' the population moment around 0 of order h.

1) Find the covariance of m_h' and m_k'.

2) Prove that m_h' is asymptotically normal.

3) Give the definition of a consistent estimator and demonstrate that m_h' is a consistent estimator of μ_h' using the definition.

3.

TABLE 1
Year-to-year Change (pct) in Hours Worked
Non-farm Employment 1948-1992

(a) X Data in Temporal Order

1.7	-3.0	3.0	4.6	1.0	2.4	-3.4	3.9	2.5
-0.5	-4.2	4.3	0.6	-1.2	2.1	1.1	2.3	3.8
3.4	-0.1	1.7	2.9	-1.5	-0.3	3.3	4.0	0.2
-4.2	3.2	4.1	5.0	3.5	-0.8	0.7	-2.4	1.9
5.9	2.5	0.8	3.2	3.5	2.6	0.2	-2.4	-0.4

(b) Descriptive Statistics of X Data

Variable	N	Mean	Median	TrMean	StDev	SEMean
hrsnonf	45	1.347	1.900	1.417	2.574	0.384

Variable	Min	Max	Q1	Q3
hrsnonf	-4.200	5.900	-0.350	3.350

Sum(X)=60.6. SumSquares(X)=373.06

The economist estimates the population mean, μ, by the sample mean shown above. The economist proposes to employ the usual *t-statistic*, 44 degrees of freedom, to test null hypotheses of the form Ho:$\mu=\mu_0$ against the unrestricted alternative, H_{alt}:μ μ_0. *Fig. 1* plots the *p-values* for twenty different selections of μ_0 that the economist might make.

(a) Explain how the *p-values* are computed from the data and null hypotheses.

(b) Explain why the *p-value* cannot be "the probability that $\mu=\mu_0$ is true".

(c) *Fig. 2* displays the logarithm of the *likelihood* of the simple hypothesis (μ_0,σ^2_0). Explain how the likelihood is computed from the data and null hypothesis. Explain why the likelihood cannot be the "probability that the null hypothesis is true".

(d) Several other tests are performed: (i) The runs test (median-centered) with 23 runs and *p-value*=0.883; (ii) three different tests of normality of the distribution of X, (1) Anderson-Darling test with *p-value* = 0.065, (2) Shapiro-Wilk test, *p-value* > 0.1000 and (3) Kolmogorov-Smirnov test with *p-value* > 0.15. Explain the relevance of the outcomes of these tests to the reliability of the t-tests.

2 graphs follow this page.

4. 1) Describe *feasible generalized least squares*, the occasion of its use, its purpose and computation.

2) The undated data in **TABLE 2** are found in the *Handbook of Labor Statistics*, Bulletin 2175, December 1983, p. 407. The sample mean is 0.2367 and sample standard error of sample mean is 0.0286. For the alphabetical arrangement (a) the *Ljung-Box* statistic for testing zero autocorrelations is 30.67 (based on 20 correlations) and the *p-value* is 0.0597. For the temporal arrangement (b) the *Ljung-Box* statistic for testing zero autocorrelations is 17.31 (based on 20 correlations) and the *p-value* is 0.6326. An economist plans to test the null hypothesis that the true mean is 0.2 against the unrestricted alternative.

Discuss (pro and con) the economist's decision to re-arrange the original data (alphabetical order) as in **TABLE 2(b)** before estimating the standard error of the sample mean and computing a *p-value* from the *t-distribution* with 50 degrees of freedom.

3) If the economist does not publish test results from both arrangements, what is the major (nonstatistical) economic hypothesis she leaves untested?

TABLE 2
Work Stoppages by States 1974
(Percent of estimated private nonfarm working time)

(a) Data arranged in alphabetical order†

0.29	0.20	0.30	0.14	0.33	0.06	0.08	0.13	0.05
0.17	0.04	0.55	0.52	0.34	0.38	0.23	0.05	0.39
0.26	0.03	0.14	0.19	0.38	0.11	0.30	0.54	0.58
0.25	0.33	0.04	0.12	0.11	0.09	0.11	0.01	0.32
0.17	0.20	0.35	0.12	0.06	0.15	0.23	0.09	0.15
0.01	0.23	0.70	1.11	0.28	0.06			

† District of Columbia follows Delaware.

(b) Data arranged in order of dates of admission‡

0.13	0.35	0.12	0.04	0.08	0.19	0.14	0.06	0.04
0.23	0.09	0.11	0.12	0.01	0.39	0.23	0.32	0.26
0.38	0.30	0.34	0.29	0.03	0.54	0.14	0.38	0.17
0.09	0.23	0.28	0.33	0.11	0.20	0.05	1.11	0.33
0.25	0.06	0.01	0.15	0.58	0.70	0.52	0.06	0.15
0.17	0.11	0.30	0.20	0.55	0.05			

‡ District of Columbia listed last.

5. Consider the linear regression model $Y = \beta X + u$ which satisfies all the standard assumptions. Furthermore, assume that σ^2 is known. Also note that X is a scalar. You are given the following three estimators of β, viz.,

i. $b = \dfrac{\sum XY}{\sum X^2}$

ii. $b^* = \dfrac{\sum xy}{\sum x^2}$

iii. $\tilde{b} = \bar{Y} / \bar{X}$

where x and y are in deviation forms.

(a) Calculate variances of all three estimators and show that b is the most efficient estimator of β.

(b) Show that $\tilde{b}$ is the most efficient estimator of β, when $Var\ (u_i) = \sigma^2 X_i$.

(b) Which one is the most efficient estimator of β if the model is $Y = \alpha + \beta X + u$ in which all the classical assumptions are satisfied? Explain.

6. In the adaptive expectations model of capital stock growth, the level of capital is set at a multiple of anticipated output Y_t^*,

$$K_t = \delta Y_t^*,$$

where

$$Y_t^* - Y_{t-1}^* = \alpha(Y_t - Y_{t-1}^*) + u_t, \quad O < \alpha \leq 1$$

(a) Express K_t as a function of observed variables. (Note that Y_t^* and Y_{t-1}^* are unobserved.)

(b) Does OLS to the model in part (a) give consistent estimates of α and δ when u_t follows an AR(1) process? Explain.

Fig. 1. P-value of null hypothesis, Ho:mean=mu00, for HRSNONF sample.

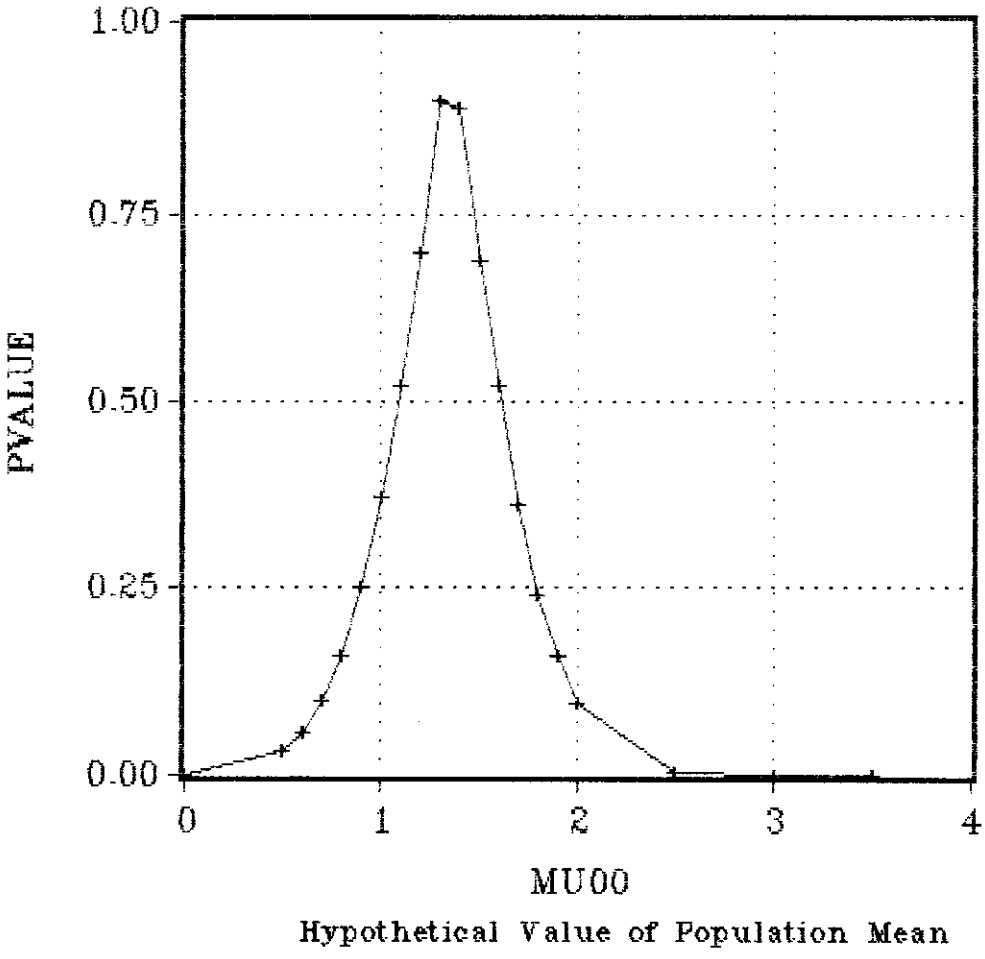

Fig. 2. Logarithm of likelihood function vs null hypothesis (mu=mu0)