University of Illinois
Spring 2009

Final Exam
Economics 574

Department of Economics
Professor Roger Koenker

The ultimate objective of field courses is (presumably) to prepare you to digest new material in future research. Thus, the exam is structured to allow you to demonstrate some facility for this, and in the process to learn some (marginally) fun, new stuff. The exam is due friday, May 15 at 5pm. Feel free to email or stop by my office if you feel that there are points that need clarification; I'll post corrections or clarifications on the course webpage, if necessary.

1. (Extreme Econometrics) In Problem Set 1 we saw that for iid standard Gaussians $X_1, \ldots, X_n$:

$$M_n = \max_i \{X_1, \ldots, X_n\} \sim \sqrt{2 \log n},$$

but inquiring minds may want to know more. Could we normalize M_n in some way so that say, $a_n(M_n - b_n)$ converged *in distribution*. In particular, we would like to show that by appropriate choices of a_n and b_n our normalized M_n has a Gumbel, or Type I extreme value distribution.

Theorem 1 *If $\{X_1, \ldots, X_n\}$ are iid $\mathcal{N}(0,1)$, then $M_n = \max\{X_1, \ldots, X_n\}$ satisfies*

$$P(a_n(M_n - b_n) \le x) \to e^{-e^{-x}}$$

where $a_n = \sqrt{2 \log n}$ and $b_n = \sqrt{2 \log n} - \frac{1}{2}(\log \log n + \log 4\pi)/\sqrt{2 \log n}$.

(a) Recall that the limiting distribution appearing in this theorem is the one appearing in the McFadden multinomial choice model. If individual i gets utility U_{ij} from choice j given by

$$U_{ij} = \mu_{ij} + \varepsilon_{ij}$$

where μ_{ij} are some constants depending upon characteristics of the individual i and choice j, and the ε_{ij} are independent random variables with distribution function $F(x) = e^{-e^{-x}}$, then the utility maximizing choices take the multinomial logit form

$$P(Y_i = j) = \frac{e^{\mu_{ij}}}{\sum_k e^{\mu_{ik}}}$$

where $Y_i = \arg\max_j \{U_{ij} : j = 1, \ldots, J\}$. Derive this result justifying each step, following if necessary Lecture 20 for my 508 course. Comment on whether you think that the interpretation of F as a limiting distribution for a maximum of normals makes the McFadden model more appealing.

(b) To establish the Theorem consider the following

Lemma 1 *Suppose $\{\xi_1, \ldots, \xi_n\}$ are iid with df F, and $M_n = \max\{\xi_i\}$. Let $0 \le \tau < \infty$ and $\{u_n\}$ be a sequence of real numbers such that*

$$n(1 - F(u_n)) \to \tau \quad \text{as } n \to \infty.$$

Then

$$P(M_n \le u_n) \to e^{-\tau} \quad \text{as } n \to \infty.$$

Complete the proof of the lemma by first noting that,

$$P(M_n \leq u_n) = F^n(u_n) \to (1 - (1 - F(u_n)))^n.$$

(c) To complete the proof of the theorem, take $\tau = e^{-x}$ in the lemma so

$$1 - \Phi(u_n) = n^{-1}e^{-x}.$$

Using Feller's inequality from Lecture 3 we have

$$1 - \Phi(u_n) \sim \phi(u_n)/u_n$$

so for $u_n \to \infty$

$$n^{-1}e^{-x}u_n/\phi(u_n) \to 1$$

or

$$-2\log n - 2x + 2\log u_n + \log 2\pi + u_n^2 \to 0$$

hence (why?) $u_n^2/(2\log n) \to 1$, so

$$2\log u_n - \log 2 - \log\log n \to 0.$$

Now, substituting back we have

$$\begin{aligned} u_n &= \frac{x}{a_n} + b_n + \sigma((\log n)^{1/2}) \\ &= \frac{x}{a_n} + b_n + \sigma(a_n^{-1}) \end{aligned}$$

Fill in the missing details.

(d) As a check on the foregoing try simulating some moderately large number of Gaussian maxima. In R you can use, `Mn <- apply(matrix(rnorm(n*m),n,m),2,max)`, and plotting their empirical cdf against the cdf of the limiting distribution.

(e) Suppose we wanted to find the asymptotic distribution of the sample *minimum* how would this modify the result of the theorem?

(f) Suppose $X_1, \ldots, X_n$ were 3 parameter lognormal, i.e., $\log(X_i - \alpha) \sim \mathcal{N}(0,1)$ how does this change things? Hint: For any monotone transformation, g, so $Y_i = g(X_i)$ $i = 1, \ldots, n$

$$M_n(Y_1, \ldots, Y_n) = g(M_n(X_1, \ldots, X_n))$$

so since

$$P\left(M_n(X_1, \ldots, X_n) \leq \frac{x}{a_n} + b_n\right) \to e^{-e^{-x}}$$

it follows that

$$P\left(M_n(Y_1, \ldots, Y_n) \leq g\left(\frac{x}{a_n} + b_n\right)\right) \to e^{-e^{-x}}.$$

In the lognormal case $g(x) = e^x + \alpha$ so one can find new normalizing constants by expanding as follows

$$\begin{aligned} g\left(\frac{x}{a_n} + b_n\right) &= e^{b_n} \exp\left(\frac{x}{a_n}\right) + \alpha \\ &= e^{b_n}\left(1 + \frac{x}{a_n} + \nu\left(\frac{x}{a_n}\right)^2\right) + \alpha, \quad \nu \in (0,1) \\ &= e^{b_n} + \frac{e^{b_n}}{a_n} x\left(1 + \nu\frac{x}{a_n}\right) + \alpha \end{aligned}$$

but $a_n \to \infty$ so the term in parentheses converges to 1, and we have new constants $b_n' = \alpha + e^{b_n}$, and $a_n' = a_n e^{-b_n}$. Explain briefly steps that look unclear, or incorrect.

(g) Now suppose that we would like to use the foregoing theory to estimate α using the "method of moments." Note that this means we are really interested in the theory for the sample minimum not maximum. Find an asymptotically median unbiased estimator of α based on $X_{(1)}$, the smallest order statistic.

(h) Compare performance of your $\hat{\alpha}$ with the pseudo-maximum likelihood estimator of α based on maximizing the likelihood as a function of α and conscientiously staying away from limit $\alpha \to X_{(1)}$, which we know causes problems. In R the function `optimize()` is useful in such situations since you can specify an interval to search for the optimizer.

2. (Harmless Econometrics) Consider a model with a covariate f with many levels $f \in \{1, \ldots, J\}$ say, indicating industry or occupation. In R such covariates are referred to as "factors". We have a response variable y and a covariate of primary interest z which we regard as "endogenous." Perhaps we also have some other covariates X. Note that in above formulation the we will define a matrix F denoting the matrix of "dummy" (indicator) variables corresponding to the levels of the factor variable f, i.e., $F_{ij} = 1$ if $f_i = j$ and $F_{ij} = 0$ otherwise. And we will assume that X includes an intercept, so one level of f_i is excluded, and consequently values of the parameters γ_j and δ_j below can be interpreted as differences between the conditional means of jth level and that of the omitted reference level.

We proceed as follows:

(a) Estimate the two "reduced form" regressions

$$\begin{aligned} \hat{y} &= X\hat{\beta} + F\hat{\gamma} \\ \hat{z} &= X\hat{\alpha} + F\hat{\delta} \end{aligned}$$

to obtain the two $(J-1)$-vectors $\hat{\gamma}$ and $\hat{\delta}$.

(b) Plot these pairs of coefficients $\hat{\gamma}$ vs. $\hat{\delta}$ and now estimate, by least squares, the model

$$\hat{\gamma}_i = a + b\hat{\delta}_i + u_i$$

(c) Show that $\hat{b}_{OLS}$ in this regression has the same limiting value as the 2SLS estimator

$$\hat{b}_{2SLS} = (\tilde{z}' P \tilde{z})^{-1} \tilde{z} P \tilde{y}$$

where $\tilde{z}_i = z_i - \bar{z}$, $\tilde{y}_i = y_i - \bar{y}$ and P is the projection matrix onto the columns of $[X \vdots F]$.

(d) Try to characterize the asymptotic relative efficiency of the two procedures. Note that the claim in part (c.) is only that the two unnormalized estimators have the same asymptotic limit, not that their normalized versions have the same asymptotic distribution.

(e) Discuss assumptions under which the proposed estimator might be entitled to a claim of being a "causal effect."

Hint: To get started with this question I recommend first simplifying drastically so there is just a binary variable f, and X is just an intercept. Then we have (explain)

$$\begin{aligned} \hat{\gamma} &= \bar{y}_1 - \bar{y}_0 \\ \hat{\delta} &= \bar{z}_1 - \bar{z}_0 \end{aligned}$$

and $\hat{\gamma}/\hat{\delta}$ is the usual Wald estimator. Well, it is when the sample sizes are the same in both samples, but perhaps not otherwise. Note that in this case there is only one point to plot so what is meant by least squares must be fitting a line "through the origin." Now generalize. This question is based on a suggestion appearing in Angrist's Ph.d. thesis and in an (unpublished?) paper by Holzer, Katz and Krueger (1988) where the plotting technique is referred to as "visual IV." The idea has been recently revived by Angrist and was mentioned in his recent departmental seminar.