Problem set 6
607: Applied Macroeconometrics
Fall 2016
Jon Faust

The following is due at the beginning of next class. You can turn in any paper in my mailbox or in class; email me and requested computer work. You may work in groups; hand in a single submission for the group. The submission should list those who contributed.

1. Thinking about Monte Carlos, samples, EDFs, and urns. Throughout this problem set, drawing from a urn is taken to mean drawing uniformly from the items in the urn.

   Given a sample of data, Y, $(T \times 1)$, the empirical distribution function (EDF) is defined as

   $$EDF(c) = \frac{\#(y_t < c)}{T}$$

   where # means 'number of.' That is, the EDF is the share of the ys that are less than c.

   (a) Give some conditions under which the EDF will be consistent for the true CDF. That is, $EDF(c)$ converges to the true $F(c)$ with sample size for any fixed c.
   (b) Suppose we have a sample from an iid random variable Y $(T \times 1)$. We write each value on a piece of paper, throw them in an urn and form new samples by drawing with replacement from the urn. What is the relation between the EDF and the exact distribution function for a single draw from this urn?
   (c) Given the vector `Y`, how would you compute $EDF(c)$ in Matlab?
   (d) You find yourself in Matlab with only acess to the `mean` and `randn` functions. Write a single command that returns the value of the standard normal CDF at c.

2. More urns.

Our data are $Y \sim P_\theta$, where Y is $(T \times 1)$ and the elements of Y are iid. We are interested in the distribution of a scalar statistic $\psi(Y)$, but P_θ is unknown. Specifically, we are looking for a good approximation to

$$\pi = P_{\theta^*}(\psi(Y) > c)$$

under an unknown P_{θ^*}.

We follow the urn procedure desribed in problem 1, and draw a zillion samples of size T. We compute $\psi(Y)$ on each sample, and use the EDF of the resulting ψs as our proxy for the true distribution.

Formally or informally give an argument as to why $\hat{\pi} = 1 - EDF(c)$ will be consistent for π?

3. More urns, again. $Y \sim P_\theta$, and we want to know the distribution of, $\psi(Y)$, where Y is $(T \times 1)$. We know $Ey_t = 0$ and that y_t follows a covariance stationary MA(1) driven by iid shocks, but we don't know the shock distribution.

   [I've simplified the setup relative to the question you were asked in order to make this more focussed. As I warned, I'm trying to refine these problem sets]

   Consider the following 'urn' procedure. Partition the T observations into B blocks of b contiguous observations. (Suppose that we are only interested in samples of sizes satisfying $T = Bb$ for integer B, b. This integer-constraint issue is a bit of a pain in practice, but not for the basic theory argument).

   Write each block on a piece of paper, throw them in the urn, and form new samples of size T by drawing B blocks with replacement.

   Once again, draw a zillion such samples, compute ψ on each, and use the resulting EDF as a proxy for the true CDF of ψ under the unknown P_θ.

   Formally or informally, give an explanation as to why the EDF of $\psi(Y)$ will be a good approximation to the true distribution of ψ under P_θ as T, B, and b get large.

4. One more urn. Suppose the model explains $Z = [Y X] \sim P_\theta$ where Y and X are $(T \times 1)$ and the data have no time dependence (the rows of Z are mutually independent). You'd like to understand the distribution of the OLS $\hat{\beta}$ from a regression of Y on X.

You write each observation (each row of Z) on a piece of paper. Put them in an urn and draw a zillion samples of size T, drawing with replacement. You compute $\hat{\beta}$ on each and take the EDF of the resulting $\hat{\beta}$s as a proxy for the exact distribution of $\hat{\beta}$.

Formally or informally give an argument as to why the EDF of the $\hat{\beta}$ will converge to the distribution of $\hat{\beta}$ under the true DGP.

5. The subject of the classroom presentation: Suppose you have $Z = [Y X]$ $(T \times 2)$ where $(y_t, x_t) \sim iidN(0, \Omega)$ with Ω full rank. Define the correlation of x and y to be ρ and $\hat{\rho}$ to be the natural sample estimate—the sample covariance divided by the square root of the product of the two sample variances.

   (a) Show that $\hat{\rho}$ is consistent.

   (b) Show that $\sqrt{T}(\hat{\rho} - \rho) \overset{a}{\sim} N(0, \sigma^2)$ for some σ^2 and give an expression for σ^2 in terms of the parameters driving the Zs.

   (c) Run a Monte Carlo for various sample sizes and values of ρ. For an interesting assortment of T and ρ show a histogram approximating the distribution of $\hat{\rho}$. (Print a matrix of histograms, each row a different T; each column a different ρ). Print analogous tables in which each cell gives the mean, variance, and skewness of $\hat{\rho}$ for different (T, ρ) pairs.

   (d) Show that the t-statistic, $\tau = \sqrt{T}(\hat{\rho} - \rho_0)^2/\hat{\sigma}^2$, is asymptotically $N(0, 1)$ when constructed with ρ_0 equal to the true ρ.

   (e) Using the same Monte Carlo as above, compute the distribution of the marginal significance of the t-test just described. More specifically, on each Monte Carlo sample, compute the marginal significance of both one-tailed tests and the two tailed test. Produce a matrix of histograms of those p values.

   Note: On each sample, compute the statistic $\hat{\tau}$ basing it on the true ρ for that sample. The marginal significance under the three different tests are then,

$$\text{pr}(z < \hat{\tau}), \text{pr}(z > \hat{\tau}), \text{ and } \text{pr}(|z| > |\hat{\tau}|)$$

   where $\hat{\tau}$ is fixed at the sample value and z is a standard normal random variable.

   What do you see?

(f) What is Fisher's-z transform in this context; describe informally why it makes sense.

Notes: The distribution of estimates of correlation has been discussed immensely since the early days of statistics. This may be the simplest case that gives rise to versions of many of the complexities that drive us to advanced methods. It provides a nice context in which to think about the stochastic expansions we will be discussing and their practical importance. If you are into theory, mastering this literature will give you sense of the topics we'll be grappling with.

A very nice summary of the highlights in this area is provided in the introduction of Ogasawara (2006) cited below. If you read the article, you'll also note how explicit expressions in this expansions literature get pretty messy even in simple cases.

We'll show that the bootstrap can at times capture higher order terms in expansions using Monte Carlo methods of the type you played with in the earlier questions. The Monte Carlo methods are conceptually straightforward and easy to implement, which explains why the bootstrap has become such an important technique in practice.

Haruhiko Ogasawara, Asymptotic expansion of the sample correlation coefficient under nonnormality, Computational Statistics & Data Analysis, Volume 50, Issue 4, 24 February 2006, Pages 891-910,

http://www.sciencedirect.com/science/article/pii/S0167947304003251