14.385 Recitation 4

Paul Schrimpf

October 3, 2008

1 MLE as efficient GMM

See lecture 5 notes and Newey and McFadden page 2163+.

2 Bootstrap

See MacKinnon's slides for some comparison of bootstrap methods.

See website for bootstrap code in R and Stata.

2.1 Examples of Failure

Recall the theorem for the consistency of the bootstrap:

Theorem 1. *$G_n(\cdot, F_n)$ is consistent if for any $\epsilon > 0$ and $F_0 \in mathcalF$:*

(i) $\lim P(\rho(F_n, F_0) > \epsilon) = 0$

(ii) $G_\infty(\tau, F)$ is continuous in τ for each F

(iii) For any τ and $\{H_n\}$ such that $\lim \rho(H_n, F_0) = 0$*, we have $G_n(\tau, H_n) \to G_\infty(\tau, F_0)$*

Interesting examples where the bootstrap fails usually involve condition (iii) being violated. Horowitz gives two simple examples where the bootstrap is inconsistent:[1]

- *Maximum of a sample:* we talked about this in class. Horowitz shows directly why it fails. It is also easy to see that it does not meet condition (iii). Take H_n to be the same as F_0, but truncated to have support $[0, \theta_0 - 1/n]$. Then $\rho(H_n, F_0) \leq F_0(\theta_0) - F_0(\theta_0 - 1/n) \to 0$, where ρ is the L^∞ metric. However, $T_n(H_n) \leq -1$, so $G_n(\tau, H_n)$ cannot converge to $G_\infty(\tau, F_0)$ for $\tau \in (-1, 0)$.

- *Parameter on the Boundary:* suppose we know the population mean is $\mu \geq 0$. Our estimator is $m_n = \bar{X}\mathbf{1}(\bar{X} > 0)$. The statistic is $T_n = \sqrt{n}(m_n - \mu)$. Suppose F_0 has $\mu = 0$. Then $G_\infty(\cdot, F_0)$ is a normal censored at 0. Take H_n equal to F_0 with the mean shifted by $1/\sqrt{n}$. Then $G_n(\cdot, H_n)$ converges to a normal censored at -1.

In both of these examples, the limiting distribution depends on how F_n approaches F_0. Two other common situations in econometrics with this feature are:

- Weak instruments, where if $\pi_n = C/\sqrt{n}$, $\hat{\beta} - \beta_0$ converges to a nonnormal limit distribution that depends on C

- Unit roots: $y_t = \rho y_{t-1} + e_t$. If $\rho < 1$, then $\hat{\rho}$ is asymptotically normal. If $\rho = 1$, $\hat{\rho}$ has a nonstandard distribution. If $\rho_T = 1 - a/\sqrt{T}$, $\hat{\rho}$ has another nonstandard distribution that depends on a.

[1]Note that in both of these examples, we only show that the above theorem is inapplicable. We would have to do more to show that the bootstrap is inconsistent.

2.2 Subsampling

Subsampling works in some of the situations where the bootstrap fails. Let $\{X_i\}_{i=1}^n$ be our data with true distribution F_0. We are interested in a statistic T_n for which $\tau_n(T_n - \theta_0)$ has distribution function $G_n(\cdot, F_0)$. Let $\{Y_k\}_{k=1}^{N_n}$ be the $\binom{N_n = n}{b}$ subsets of $\{X_i\}$ of size b. Let $S_{n,k}$ be the value of T computed using subset Y_k. The subsample approximation to $G_n(\cdot, F_0)$ is:

$$L_n(\tau) = N_n^{-1} \sum_{k=1}^{N_n} \mathbf{1}[\tau_b(S_{n,k} - T_n) \leq \tau$$

Politis and Romano (1994) give conditions for L_n to be consistent:

Theorem 2. *Assume $G_n(\cdot, F_0) \rightsquigarrow G_\infty(\cdot, F_0)$. Also assume $\frac{\tau_b}{\tau_n} \to 0$, $b \to \infty$ and $\frac{b}{n} \to 0$. Then*

1. *Wherever $G_\infty(\cdot, F_0)$ is continuous, $L_n(x) \xrightarrow{p} G_\infty(x, F_0)$*

2. *If $G_\infty(\cdot, F_0)$ is continuous, then* $\sup |L_n(x) - G_\infty(x, F_0)| \xrightarrow{p} 0$

Proof. Rewrite

$$L_n(\tau) = N_n^{-1} \sum_{k=1}^{N_n} \mathbf{1}[\tau_b(S_{n,k} - \theta_0) + \tau_b(\theta_0 - T_n) \leq \tau$$

Convince yourself that if

$$U_n(\tau) = N_n^{-1} \sum_{k=1}^{N_n} \mathbf{1}[\tau_b(S_{n,k} - \theta_0)]$$

converges to $G_\infty(\cdot, F_0)$ and $G_\infty(\cdot, F_0)$ is continuous, then $L_n \xrightarrow{p} G_\infty$ too.

U_n is a U-statistic of degree b. Hoeffding showed that U-statistics of degree r satisfy $P(U - EU \geq t) \leq e^{-2\frac{n}{r}t^2}$. Here $EU_n = G_b(\cdot, F_0)$ since $G_b(\cdot, F_0)$ is the distribution of $\tau_b(S_{n,k} - \theta_0)$. Thus we have

$$P(U_n(x) - G_b(x, F_0) \geq t) \leq e^{-2\frac{n}{b}t^2} \to 0$$

The second part of the theorem follows from Polya's lemma, just like in the proof of the bootstrap. □

Politis and Romano prove a similar theorem for dependent data. The only difference is that the subsamples must be continguous blocks of length b.

Remarks:

- The conditions for subsampling to be consistent are very weak. They are definitely met in the examples of the maximimum, parameter on a boundary, and unit root above. They may be met in the weak instruments case.

- Subsampling converges at best at an $n^{-1/3}$ rate

- The subsampling theorem above is about pointwise convergence. Subsampling often does not converge uniformly, i.e.

  $$\lim_{n \to \infty} \sup_{F \in \mathcal{F}} P(|L_n - G_\infty(\cdot, F)| < \epsilon) \neq 0$$

  In practice, this means that subsampling can give poor results in finite samples for some values of the parameters. Mikusheva (2007) talks about how subsampling is not uniformly valid in AR models. Andrews and coauthors have recent papers about how subsampling is not uniformly consistent for parameters on the boundary and weak instruments.