Section 8: Estimators and statistics

ARE 210

October 24, 2017

1) (Midterm, Question 6) Let $Z = g(X)(Y - \mathbb{E}(Y|X))$ for some measurable function $g : \mathbb{R} \to \mathbb{R}$. Assume that all moments in question exist.

a) Compute $\mathbf{E}(Z|X)$.

$\mathbf{E}[Z|X] = g(X)\mathbf{E}[Y|X] - g(X)\mathbf{E}[Y|X] = 0.$

b) Compute $V(Z)$

$\text{Var}[Z] = \mathbf{E}[Z^2] - \mathbf{E}[Z]^2 = \mathbf{E}[Z^2]$, applying the law of iterated expectations. Next, $\mathbf{E}[Z^2|X] = g(X)^2(\mathbf{E}[Y^2|X] - 2\mathbf{E}[Y|X]^2 + \mathbf{E}[Y|X]^2) = g(X)^2(\mathbf{E}[Y^2|X] - \mathbf{E}[Y|X]^2)$. Taking the expectation of this, we're left with $\mathbf{E}[Z^2] = \mathbf{E}[g(X)^2(\mathbf{E}[Y^2|X] - \mathbf{E}[Y|X]^2)]$, which is our answer.

Note that we can rewrite this $\mathbf{E}[g(X)^2 V[Y|X]]$, which is a scaled weighted average of conditional variances of Y.

c) Now let Let $Z = \mathbf{E}(g(X))(Y - \mathbb{E}(Y|X))$. Compute $V(Z)$.

First, $\text{Var}[Z] = \mathbf{E}[g(X)]^2\text{Var}[Y - \mathbf{E}[Y|X]]$. Using the previous part, we get $\text{Var}[Z] = \mathbf{E}[g(X)]^2\mathbf{E}[(\mathbf{E}[Y^2|X] - \mathbf{E}[Y|X]^2)]$.

Similar to the last part, we can rewrite this $\mathbf{E}[g(X)]^2\mathbf{E}[V[Y|X]]$, which is a scaled *unweighted* average of conditional variances of Y.

2) (Midterm, Question 7) Let $Y_i^* \sim N(\theta, 1)$ and let $Y_i = \mathbb{I}\{Y_i^* > 0\}$. Suppose we observe a an i.i.d. sample $\{Y_i\}_{i=1}^n$.

a) Let Φ^{-1} denote the inverse of the cumulative distribution function (CDF) of the $N(0,1)$ distribution and let ϕ denote the probability density function of the $N(0,1)$ distribution. Find the MLE for θ and denote it by $\hat{\theta}_n$.

$Y_i \sim \text{Bern}[\Phi(\theta)]$, therefore $\hat{\Phi}(\theta)_{MLE} = \sum_{i=1}^n Y_i/n$. Applying the invariance principle, $\hat{\theta}_{MLE} \equiv \hat{\theta}_n = \Phi^{-1}(\sum_{i=1}^n Y_i/n)$.

b) Show whether $\hat{\theta}_n$ converges in probability to θ.

$\lim_{n\to\infty} \hat{\theta}_n = \lim_{n\to\infty} \Phi^{-1}(\sum_{i=1}^n Y_i/n) = \Phi^{-1}(\lim_{n\to\infty} \sum_{i=1}^n Y_i/n)$, by application of the CMT. By the WLLN, $\lim_{n\to\infty} \sum_{i=1}^n Y_i/n = \mathbf{E}[Y_i] = \Phi(\theta)$. Putting this together implies $\lim_{n\to\infty} \hat{\theta}_n = \theta$.

c) Derive the limiting distribution for (an appropriately normalized) $\hat{\theta}_n$

Applying the delta method and the inverse function theorem,
$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \to N\left(0, \frac{\Phi(\theta)(1-\Phi(\theta))}{\phi(\theta)^2}\right)$, where ϕ is the standard normal pdf.

3) (Problem Set 4, Question 4) (**Identification in an Endogenous Non-Parametric Regression Model**) Suppose that we have a model given by

$$y_i = \mu\left(x_i\right) + \epsilon_i$$

and that we assume $\mathbb{E}\left(\epsilon_i|z_i\right) = 0$ for some variable z_i and the model is non-parametric in the sense that we place no restrictions on the form of the (unknown) function $\mu\left(\cdot\right)$. Consider taking expectations conditional on z to obtain

$$\int y f\left(y|z\right) dy = \int \mu\left(x\right) f\left(x|z\right) dx$$

The above is an example of an integral equation. Note that since we observe (y, x, z) the density functions $f\left(y|z\right)$ and $f\left(x|z\right)$ are identified. We will show that this integral equation will have a unique solution if the distribution of x conditional on z is complete (in z).

Suppose that there exists another function $\tilde{\mu}\left(\cdot\right)$ that satisfies the equation, then we must have

$$\int \left(\mu\left(x\right) - \tilde{\mu}\left(x\right)\right) f\left(x|z\right) dx = 0$$

In order for their to be a unique solution therefore we must have the following hold (setting $\delta\left(x\right) = \mu\left(x\right) - \tilde{\mu}\left(x\right)$)

$$\int \delta\left(x\right) f\left(x|z\right) dx = 0 \Rightarrow \delta\left(x\right) = 0 \tag{1}$$

for almost all z. Suppose that the distribution of x conditional on z is given by

$$f\left(x|z\right) = h\left(x\right) \exp\left(\eta\left(z\right)\tau\left(x\right) - B(z)\right) \tag{2}$$

Suppose further that the function $\tau(\cdot)$ is one-to-one.[1] Show that this implies (1) so a sufficient condition for the integral equation to have a unique solution is the distributional assumption (2) (Hint: Use the result on Completeness for Exponential Families. See Newey & Powell (2003) for an extension to cases where the conditional distribution does not belong to an exponential family.)

First, note that τ is 1-to-1 implies that $\delta(x) = \delta(\tau^{-1}(\tau(x)))$, so δ is a function of τ. Second, we had the result that $\tau(x)$ is complete for z, since $f(x|z)$ belongs to the exponential family. Therefore, $\mathbf{E}_z[g(\tau)] = 0 \Rightarrow g(\tau) = 0$ almost everywhere. Substituting $g = \delta \circ \tau^{-1}$, we have that $\delta(x) = 0$ almost everywhere.

Note that the intuition here is in some ways similar to the intuition for how a characteristic function can encode all the information about a distribution - seeing how $\mathbf{E}[y|z]$ changes as $f(x|z)$ changes in a neighborhood of some z tells us the full function μ with the right functional form for f. Another assumption, that we need that $f(x|z)$ to be shifted by z, is equivalent to assuming there's a first stage relationship in standard instrumental variables.

[1]Formally we also need to assume that the support of $\mu(z)$ contains an open set and $h(x) > 0$.