ECMA 31100: Problem Set 2

Due Feb 20 by 11:59PM

Question 1 Consider the potential outcomes framework

$$Y = Dy_1 + (1 - D)\, y_0$$

and suppose that $(y_0, y_1) \perp D|X$, where X is a vector of covariates. Define $ATE\,(x) := \mathrm{E}\,(y_1 - y_0|X = x)$. You are given an iid sample of $\{Y_i, D_i, X_i\}_{i=1}^{N}$. Assume that all relevant moments exist. Assume $\mu_d\,(x) := \mathrm{E}\,(y_d|X = x) = \alpha_d + x'\beta_d$ for $d = 0, 1$.

a) Show that running a regression of Y on X and a constant for observations with $D = 1$ produces unbiased estimates of α_1, β_1. Argue likewise for $D = 0$ and α_0, β_0.

b) Define $\hat{\mu}_d\,(x) := \hat{\alpha}_d + x'\hat{\beta}_d$ for $d = 0, 1$, where $\left\{\hat{\alpha}_d, \hat{\beta}_d\right\}_{d=0,1}$ denote the OLS estimates from part a). Construct an estimate of the ATE using

$$\widehat{ATE} = \frac{1}{n}\sum_{i=1}^{n} \hat{\mu}_1\,(X_i) - \hat{\mu}_0\,(X_i)\,.$$

Show that

$$\widehat{ATE} = \bar{Y}_1 - \bar{Y}_0 - \left(\bar{X}_1 - \bar{X}_0\right)'\left(\frac{N_0}{N}\hat{\beta}_1 + \frac{N_1}{N}\hat{\beta}_0\right),$$

where N_d is the number of observations for which $D = d$, $\bar{Y}_d = \frac{1}{N_d}\sum_{i=1}^{N} Y_i\mathbf{1}\,(D_i = d)$ and $\bar{X}_d = \frac{1}{N_d}\sum_{i=1}^{N} X_i\mathbf{1}\,(D_i = d)$. How does this estimate relate to your estimate of the ATE from problem 2 of problem set 1?

c) Estimate the standard deviation of $\widehat{ATE}$ using a bootstrap procedure:

- Draw N observations from your original dataset $\{Y_i, D_i, X_i\}_{i=1}^{N}$ at random and with replacement. Give each sample point equal probability of being drawn.
- Compute $\widehat{ATE}$ using your resampled data.
- Repeat 2,000 times to form a collection $\left\{\widehat{ATE}_j\right\}_{j=1}^{2000}$ of estimates.
- Compute the standard deviation of your estimates.

How does the bootstrap standard error compare to your consistent estimate of the asymptotic standard deviation from problem set 1?

d) Under what conditions does $\widehat{ATE}$ converge in probability to the naive comparison $\mathrm{E}(Y|D=1) - \mathrm{E}(Y|D=0)$? Provide intuition for the validity of the naive comparison when D is independent of X. Is it necessarily the case that $(y_0, y_1) \perp D$?

Question 2 Let the set of possible treatments be $\mathcal{D}$. For each $d \in \mathcal{D}$ let $y(d)$ be the potential outcome under treatment d.

a) Suppose that $y(d) = x(d,w)' \beta + u$ where x is a known function, w is a set of covariates and $\mathrm{E}(u|w) = 0$. Show that for all $d, d_j \in \mathcal{D}$:

1. $y(d) - y(d_j) = j(d,w)' \beta_j$ for some function j. (Constant treatment effect)
2. $\mathrm{E}(y(d)|w) = x(d,w)' \beta$. (Correctly specified conditional mean)

b) Now suppose 1. and 2. hold from part a). Show that there exists a u such that for every d:

- $y(d) = x(d,w)' \beta + u$
- $\mathrm{E}(u|w) = 0$.

c) Consider the simultaneous equations example from Week 2. Suppose the model is in fact given by

$$\begin{aligned} q_D &= \beta_0 + \beta_1 p + \beta_2 r + u; \quad \mathrm{E}(u) = \mathrm{E}(ru) = 0, \\ q_S &= \gamma_0 + \gamma_1 p + \gamma_2 z + v; \quad \mathrm{E}(v) = \mathrm{E}(vz) = 0. \end{aligned}$$

where z is an "exogenous supply shifter" and r is an "exogenous demand shifter". You observe a sample of $\{q_i, p_i, r_i, z_i\}_{i=1}^n$ from n markets, where q, p are observed prices and quantities that occur in equilibrium. Give conditions under which an IV strategy would allow you to identify β_1 and γ_1 and detail the strategy.

Question 3 Consider the linear model

$$y = \beta_0 + \beta_1 x + u;$$

where x is an endogenous scalar regressor and z is an excluded instrument. Suppose that $\mathrm{E}(u|x,z) = 0$ and $\mathrm{E}(u^2|x,z) = \sigma^2$, and that $Cov(x,z) \neq 0$.

a) Show that if z is used as an instrument for x,

$$\sqrt{n}\left(\hat{\beta}_1^{IV} - \beta_1\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma^2}{Corr(x,z)^2 \, Var(x)}\right)$$

It might be easier to work out the joint asymptotic distribution of $\left(\hat{\beta}_0^{IV}, \hat{\beta}_1^{IV}\right)'$ and then figure out the $(2,2)$ entry of the limit variance matrix.

b) Now derive the limit distribution of $\hat{\beta}_1^{OLS}$ and compare the asymptotic variances of the IV and OLS estimates. What do you conclude?

c) Is your conclusion in part b) robust to heteroskedasticity? Prove it or find a counterexample.

Question 4 Consider the linear model

$$y = \beta x + u;$$

where x is an endogenous scalar regressor and z is an excluded instrument. Suppose that $\mathrm{E}(u|z) = 0$ and $\mathrm{E}(u^2|z) = \sigma_u^2$, and that $\mathrm{E}(xz) \neq 0$.

a) Show under weak instrument asymptotics that the OLS estimator is inconsistent and find its probability limit. That is, suppose that at sample size n the first stage is given by

$$x_n = \pi_n z + v; \quad \mathrm{E}(zv) = 0, \, \pi_n = \frac{\pi}{\sqrt{n}}, \, \pi \in \mathbb{R}.$$

Does the result depend on π? If not, why?

b) Do we really think the first stage is changing with the sample size? How do you interpret this result?

c) Find the limit distribution of the IV estimator once assuming $\pi_n = \pi \neq 0$ for all n, and again under weak instrument asymptotics, assuming homoskedasticity. Are your homoskedasticity assumptions covered by the information given in the question?

d) If you drop homoskedasticity, is the IV estimator still centered at the probability limit of the OLS estimator asymptotically when $\pi = 0$? Prove it or find a counterexample.

For the rest of this question, retain the homoskedasticity assumptions you stated in part c).

e) Reproduce figure 1a in Stock, Wright and Yogo (2002) (here) by simulating 10,000 draws of the IV estimator and plotting a kernel density estimate (Set $n = 500$, $z_i = 1$ for all $i \leq n$, and choose π_n to fix the concentration parameter accordingly).

f) Derive the t-test statistic based on the normal approximation of the IV estimator for the null hypothesis $\beta = \beta_0$. Find its limit distribution under H_0 assuming $\pi_n = \pi \neq 0$.

g) Suppose $\beta = \beta_0$ is true. Simluate (using the same design as in part e)) the rejection probability of the t-test for different values π under weak instrument asymptotics. Show that the anderson rubin test of the same null hypothesis has asymptotically correct rejection probability no matter the strength of the instruments and confirm this with a simulation.

h) Suppose you are interested in testing the validity condition $\mathrm{E}(zu) = 0$, but do not make any assumption about β. What is the value of the test statistic used for a test of overidentifying restrictions from the slides in this setting with $\dim(x) = \dim(z) = K$? Can you use this test statistic to construct a test with desirable properties under H_0 and H_1?