ECON 626: Empirical Microeconomics

Problem Set 4

Department of Economics
University of Maryland
Fall 2016

Problem Set 4 is due at 5pm on Wednesday, November 9.

1. **Clusters.** Consider an error term, ϵ_i, under two scenarios.

   In the first scenario, $\epsilon_i \sim \mathcal{N}(0,1)$, with independent and identically distributed draws: $\epsilon_i \perp\!\!\!\perp \epsilon_j \ \forall i \neq j$. There are 800 observations of $Y_i = \beta T_i + \epsilon_i$, of which 400 are randomly assigned to treatment ($T_i = 1$) and the other 400 of which are randomly assigned to comparison ($T_i = 0$).

   In the second scenario, $\epsilon_i = \nu_g + \eta_i$. In this case, there are groups of observations within which there is a common component of the error term, ν_g, and there is an independent component, η_i. Let $\eta_i \sim \mathcal{N}(0, 0.64)$ (meaning SD=0.8), with independent and identically distributed draws: $\eta_i \perp\!\!\!\perp \eta_j \ \forall i \neq j$. Similarly, $\nu_g \sim \mathcal{N}(0, 0.36)$ (meaning SD=0.6), with independent and identically distributed draws: $\nu_g \perp\!\!\!\perp \nu_h \ \forall g \neq h$. Let groups be of size 16. There are 800 observations of $Y_i = \beta T_i + \epsilon_i$, of which 400 (that is, 25 groups of 16) are randomly assigned to treatment ($T_i = 1$) and the other 400 of which (again, the other 25 groups of 16) are randomly assigned to comparison ($T_i = 0$).

   Recall that the standard error of the coefficient on treatment should be (approximately) the standard deviation of the difference between the sample mean for treatment observations and the sample mean for comparison observations.

   (a) In the first scenario, what is the standard deviation of the sample mean of the 400 treatment observations? (Show analytically what it should be.)

   (b) In the first scenario, what is the standard deviation of the *difference* between treatment and comparison sample means? (Show analytically what it should be.)

   (c) In Stata, generate data according to the first DGP, and with $\beta = 0$ or any other value of β you choose, run the regression, showing that the standard error is very close to the value you worked out analytically.

   (d) What is the approximate MDE for power of 0.8 or 0.9, analytically (using the formula in the Lecture 7 slides)?

   (e) Use the `sampsi` command to confirm that the sample size required for the power and MDE you found above is roughly 800 (400 T, 400 C).

   (f) In the second scenario, what is the standard deviation of the sample mean of the 400 treatment observations? (Show analytically what it should be.)

   (g) In the second scenario, what is the standard deviation of the *difference* between treatment and comparison sample means? (Show analytically what it should be.)

(h) In Stata, generate data according to the second DGP, and with $\beta = 0$ or any other value of β you choose, run the regression, clustering standard errors at the group level, and showing that the standard error is very close to the value you worked out analytically.

(i) What should the intra-cluster correlation (ρ) be, according to the formula given in Lecture 7?

(j) Using the `loneway` command in Stata, confirm that ϵ has an intra-cluster correlation approximately equal to the value you calculated.

(k) What is the "design effect," following the Lecture 7 formula?

(l) What is the approximate MDE for power of 0.8 or 0.9, analytically (using the formula in the Lecture 7 slides)?

(m) Use the `sampsi` and `sampclus` commands to confirm that the sample size required for the power and MDE you found above is roughly 800 (400 T, 400 C, 50 clusters total, 25 per arm).

2. **Strata in Stata.** Extend the in-class stratification activity from Lecture 8 so that you collect three additional rejection frequencies (for at least 1000 iterations in each case):

   (a) The fraction of trials in which, under the simulated effect of 43, the null of no effect is rejected when the design is stratified (as in the original activity) <u>but</u> stratum indicators are <u>not</u> included in the regression. How does power compare to power in either the plain randomization or the stratified randomization with stratum indicators included (both of which were in the original activity)?

   (b) The fraction of trials in which, under a simulated effect of 0 (just using `math_fu` as the outcome), the null of no effect is (falsely) rejected when the design is stratified and stratum indicators are <u>included</u> in the regression? Is this test approximately correctly sized?

   (c) The fraction of trials in which, under a simulated effect of 0 (just using `math_fu` as the outcome), the null of no effect is (falsely) rejected when the design is stratified and stratum indicators are <u>not included</u> in the regression? Is this test approximately correctly sized? Which generates a smaller number of false rejections under the null: the approach in part 2b above, or the approach in this part?