Problem Set I

MIT (14.32)
Spring 2003

A. From Wooldridge: C.4, C.5, C.7, C.9

B. Additional problems

1. This problem asks you to use SAS to conduct a series of sampling experiments.

a. First, draw 500 random samples of size 10 from a random number generator for a standard normal distribution. Then increase the sample size to 40. Finally, increase the sample size to 160. Plot the sampling distributions of (i) the sample mean and (ii) the sample variance for each these three sample sizes. Now repeat your experiments for three samples drawn from another parametric distribution of your choice (e.g., a uniform distribution). Discuss the results of your experiments in light of the central limit theorem.

b. Your experiments produce "samples of sample means." Compute the mean and variance of the sample means generated by each experiment and compare them to the mean and variance predicted by statistical theory. Does the variance of the sample means (i.e., the sampling variance) decrease with sample size at the rate predicted by the theory? Does Normality matter for this?

2. You are asked to conduct a social experiment to measure the effects of a Job Search Assistance program designed to help unemployed workers find jobs. You will do this by randomly choosing n_1 experimental subjects and n_2 control subjects from a pool of $n_1+n_2=n$ unemployed workers who were selected at random from the population of new Unemployment Insurance claimants in Massachussetts.

a. Find the choice of proportion treated, $p=n_1/n$, that minimizes the sampling variance of the difference in employment rates between treatment and controls. (Treat n as a known constant).

b. Now assume that it costs α dollars to collect data on anyone in your experiment and that the job search assistance provided to the experimental group costs β dollars. You can choose any sample size (n) but you must spend no more than R dollars on the experiment. Again, maintaining the assumption that there is no treatment effect, solve for the value of p which minimizes the variance of the treatment/control contrast given the experimenter's budget constraint. Interpret your result and compare to part (a).

c. Why is it useful to do exercises like (a) and (b) while assuming there is actually no treatment effect?

3. The attached table is taken from the Woodbury and Spiegelman (*American Economic Review* 1987) paper on the reading list. The table reports the results of two social experiments examining the relationship between Unemployment Insurance (UI) and the length of time unemployed. In the *Employer Experiment*, any UI recipient finding employment for at least 4 months received a voucher worth $500 to his or her employer. In the *Claimant Experiment*, any UI recipient finding employment for at least 4 months received $500 directly.

a. Test the hypothesis that the employer bonus increased the proportion of UI claimants returning to work (i.e., ending benefits) within 11 weeks.

b. Test the hypothesis that the claimant experiment increased the proportion of UI claimants returning to work within 11 weeks. Compute the test statistic under two scenarios: (i) the experiment has no effect; (ii) the experiment has an effect.

c. For each experiment, test the hypothesis (at the 5% level) that the experiment reduced weeks of insured unemployment in the benefit year using a one-tailed and a two-tailed test. Which test seems to make more sense in this case?

d. (Extra credit) Test the hypothesis that the claimant experiment decreased total benefits paid in the benefit year. Was the claimant experiment cost-effective (based on the reduction in UI payments)?

Refer to
Table 3. "Means of Program Variables By Experimental Group." The American Economic Review. (Sep. 1987): 520.