MRU Mastering Econometrics **M. Joshway**
Spring 2020

Problem Set 1

1. Table 3 in Woodbury and R.G. Spiegelman (1987) reports the results of two social experiments meant to encourage Unemployment Insurance (UI) recipients to return to work. 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) For each experiment, test the hypothesis that bonuses decreased the proportion of UI claimants who exhausted their benefits. Compute the test statistic under two scenarios: (i) the experiment has no effect and (ii) the experiment has an effect.

   (b) For each experiment, pick a significance level and test the hypothesis that the experiment reduced weeks of insured unemployment in the first spell using a one-tailed and two-tailed test. Which test seems to make more sense in this case?

2. This problem asks you to conduct a series of sampling experiments - use Stata for this

   (a) Draw 500 random samples each with a sample size of 8 from a random number generator for a standard Normal distribution. Then increase the sample size to 32. Finally, increase the sample size to 128. Plot histograms of the sampling distributions of (i) the sample mean and (ii) the sample variance, for each of these three sample sizes. Now repeat your experiments (and plots) for three samples drawn from another parametric distribution of your choice (e.g., a uniform distribution).

   (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 seem to matter for this?

3. Table 1.1 in Mastering 'Metrics compares the health and demographic characteristics of insured and uninsured couples in the NHIS. Panel A compares the health across husbands in this sample with and without health insurance.

   (a) Calculate the t-statistic for the null hypothesis that there is no difference between the health of husbands with and without health insurance in this sample. Is the difference significantly different from zero?

   (b) Panel B of Table 1.1 shows that husbands with and without health insurance differ along many demographic dimensions. It is possible that the difference in health between the "Some HI" and "No HI" groups may be smaller if we compare across groups that are more homogeneous. To investigate this, go to http://masteringmetrics.com/resources/ and download the Stata data to produce MM Table 1.1. Use the updated Stata code posted with the module (NHIS2009_hicompare_v2.do). Execute the file through line 35 to make sure that you use the same selection criteria to produce Table 1.1.

      Is the difference between the health of husbands with Some and No HI significantly different from zero if you restrict to men who:

      i. are employed?
      ii. are employed and have at least 12 years of education?
      iii. are employed, have at least 12 years of education, and earn income of at least $80,000?

   Briefly explain how and why these restrictions affect the estimated coefficient and standard errors. Why is this relevant for a causal interpretation of the contrasts in the table?