John Chao
Spring 2015
Due Date: 3/10/2015

Economics 422

Econometrics I

Problem Set 2

1. Suppose Y_i, $i = 1, 2, ..., n$ are i.i.d. random variables, each distributed $N(10, 4)$.

   a. Let $\overline{Y} = n^{-1} \sum_{i=1}^{n} Y_i$ denote the sample mean. Compute $\Pr\left(9.6 \leq \overline{Y} \leq 10.4\right)$ when (i) $n = 20$, (ii) $n = 100$, and (iii) $n = 1000$.

   b. Suppose c is a positive number. Show that $\Pr\left(10 - c \leq \overline{Y} \leq 10 + c\right)$ becomes close to 1.0 as n grows large.

   c. Use your answer in (b) to argue that $\overline{Y}$ converges in probability to 10.

2. Y_i, $i = 1, 2, ..., n$ are i.i.d.Bernoulli random variables with $p = 0.4$. Again, let $\overline{Y}$ denote the sample mean.

   a. Use the central limit theorem to compute approximations for

      (i) $\Pr\left(\overline{Y} \geq 0.43\right)$ when $n = 100$.

      (ii) $\Pr\left(\overline{Y} \leq 0.37\right)$ when $n = 400$.

   b. How large would n need to be to ensure that $\Pr\left(0.39 \leq \overline{Y} \leq 0.41\right) \geq 0.95$? (Use the central limit theorem to compute an approximate answer.)

3. In a survey of 400 likely voters, 215 responded that they would vote for the incumbent and 185 responded that they would vote for the challenger. Let p denote the fraction of all likely voters who preferred the incumbent at the time of the survey, and let $\widehat{p}$ be the fraction of survey respondents who preferred the incumbent.

a. Use the survey results to estimate p.

b. Use the estimator of the variance of $\widehat{p}$, i.e., $\widehat{p}\left(1-\widehat{p}\right)/n$, to calculate the standard error of your estimator.

c. What is the p-value for the test H_0: $p = 0.5$ versus H_1: $p \neq 0.5$?

d. What is the p-value for the test H_0: $p = 0.5$ versus H_1: $p > 0.5$?

e. Do the results from (c) and (d) differ?

f. Did the survey contain statistically significant evidence that the incumbent was ahead of the challenger at the time of the survey? Explain.

4. Using the data in question 3 above:

a. Construct a 95% confidence interval for p.

b. Construct a 99% confidence interval for p.

c. Why is the interval in (b) wider than the interval in (a)?

d. Without doing any additional calculations, test the hypothesis H_0: $p = 0.5$ versus H_1: $p \neq 0.5$ at the 5% significance level.

5. A new version of the SAT test is given to 1000 randomly selected high school seniors. The sample mean test score is 1110 and the sample standard deviation is 123. Construct a 95% confidence interval for the population mean test score for high school seniors.