Econ 325 Section 003
Assignment 6

The due date is Wednesday, November 8 before the class starts.

1. State whether each of the following is true or false with brief explanation.

   (a) Let $\{X_1, X_2, ..., X_n\}$ be n observations, each of which is randomly drawn from a distribution with mean μ and variance σ^2. Let $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$ and $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Then, the distribution of a statistic $\frac{\bar{X}-\mu}{s/\sqrt{n}}$ is always given by t-distribution with the degree of freedom $n-1$.

   (b) Given a realized sample, the confidence interval contains the population parameter with probability either one or zero.

   (c) As the sample size increases to infinity, the variance of the sample mean approaches zero.

2. Suppose that X_1 and X_2 are random samples of observations from a population with mean μ and variance σ^2. Consider the following three point estimators of μ:

$$\hat{\mu}_1 = \frac{1}{2}X_1 + \frac{1}{2}X_2,$$
$$\hat{\mu}_2 = \frac{1}{4}X_1 + \frac{3}{4}X_2,$$
$$\hat{\mu}_3 = \frac{1}{3}X_1 + \frac{2}{3}X_2$$

   (a) Show that all three estimators are unbiased.

   (b) Which of the estimators is the most efficient, i.e., has the smallest variance?

   (c) In general, we may consider an estimator of μ given by

$$\hat{\mu} = \beta_1 X_1 + \beta_2 X_2.$$

      1. What is the condition of β_1 and β_2 for $\hat{\mu}$ to be an unbiased estimator of μ.
      2. What values of β_1 and β_2 leads to an unbiased estimator of μ that has the smallest variance? Prove your claim.

3. Assume that the yield per acre for a particular variety of soybeans is $N(\mu, \sigma^2)$. For a random sample of $n = 4$ plots, the yields per acre were 37.4, 48.8, 46.9, and 55.0. Find a 90 percent confidence interval for μ.

4. A business school placement director wants to estimate the mean annual salaries 5 years after students graduate. A random sample of 25 such graduates found a sample mean of \$42,740 and a sample standard deviation of \$4,780. Find a 90% confidence interval for the population mean, assuming that the population distribution is normal.

5. Let $\{X_1, X_2, ..., X_n\}$ be $n = 400$ observations, each of which is an independent Bernoulli random variable taking the value of 0 or 1 with probability $P(X_i = 1) = p$ and $P(X_i = 0) = 1 - p$. Here, p is a unknown population parameter. Suppose that the sample proportion of $X_i = 1$ is given by $\bar{X} = 0.52$. Construct a 95 percent confidence interval for the population parameter p.

6. Suppose that you are a plant manager for producing electrical devices operated by a thermostatic control. According to the engineering specifications, the standard deviation of the temperature at which these controls actually operate should not exceed 2.0 degrees Fahrenheit. As a plant manager, you would like to know how large the (population) standard deviation σ is. We assume that the temperature is normally distributed. Suppose that you randomly sampled 25 of these controls, and the sample variance of operating temperatures was $s_n^2 = 2.36$ degrees Fahrenheit. Compute the 95 percent confidence interval for the population standard deviation σ.

7. Table 1 reports the number of non-smokers and the numbers of smokers with the daily average of 1-14 cigarettes, 15-24 cigarettes, and more than 25 cigarettes among 1357 patients with lung-cancer and the number of smokers among 1357 patients with other diseases. Suppose that Doll and Hill randomly sampled 1357 patients with lung-cancer from a population of patients with lung-cancer and 1357 patients with other diseases from a population of patients with other diseases.

   Denote the proportions of smokers with more than 25 cigarettes per day for patients with lung-cancer and for patients with other diseases by p_x and p_y. Our concern is whether heavy smoking is associated with lung cancer in population so that we are interested in the population difference $p_x - p_y$.

   (a) What is the estimator of the difference between two population proportions of smokers, $p_x - p_y$?

   (b) Compute the 95 percent confidence interval of the difference between two population proportions of smokers.

Table 1: Average Amount of Tobacco Smoked Daily Over the 10 years

	No. of Non-Smokers	No. of Smokers with the Daily Average of		
Disease Group		1-14 Cigs.	15-24 Cigs.	25+ Cigs.
Men:				
1357 lung-cancer	7	55	964	331
1357 other diseases	61	129	1001	166

Notes: Computed from Table V of Doll and Hill (1952).