Economics 30330: Statistics for Economics
Problem Set 6
University of Notre Dame
Instructor: Julio Garín
Spring 2012

Note that there are two parts to this assignment, one of is a computer exercise. The homework is due at the beginning of class on **Wednesday, April 4**. Please complete the assignment in the allotted space. You may work in groups, but you need to turn your own work.

Linear Combinations of Random Variables and Sampling (100 points)

1. Four-part problem. Go get some coffee before starting these.

   (a) On the throw of a fair die, the expected value of the number showing is 3.5 and the standard deviation is 1.71. What is the expected value and standard deviation of the sum of the values from the throw of a pair of dice

   (b) Suppose Y_1 and Y_2 are independent, $\text{Var}(Y_1) = \text{Var}(Y_2) = \sigma_Y^2$ and $Z_1 = Y_1 + Y_2$. What is $\text{Var}(Z_1)$? How does this compare to the result found in part a)?

   (c) Generalize the previous results. Suppose $Y_1, Y_2, \cdots, Y_n$ are independent, $\text{Var}(Y_1) = \text{Var}(Y_2) = \cdots = \text{Var}(Y_n) = \sigma_Y^2$ and $Z_2 = Y_1 + Y_2 + \cdots + Y_n$. What is $\text{Var}(Z_2)$?

   (d) Again, let's keep generalizing these results. Suppose $Y_1, Y_2, \cdots, Y_n$ are independent, $\text{Var}(Y_1) = \text{Var}(Y_2) = \cdots = \text{Var}(Y_n) = \sigma_Y^2$ and $Z_2 = Y_1 + Y_2 + \cdots + Y_n$, and $Z_3 = \left(\frac{1}{n}\right)(Y_1 + Y_2 + \cdots + Y_n)$. What is $\text{Var}(Z_3)$?

2. In 2003, the average annual salary for 10 years after graduation was \$168,000 for men and \$117.000 for women. The standard deviation for male graduate salary is \$40,000 and for female salaries is \$25,000.

   (a) What is the probability that a random sample of 40 males will give a sample mean within \$10,000 of \$168,000?

   (b) What is the probability that a random sample of 40 females will give a sample mean within \$10,000 of \$117,000?

   (c) What do you prefer: giraffes or rhinos?

   (d) What is the probability that a random sample of 100 males will give a sample mean less than \$164,000?

3. Suppose in a population the math (M) and verbal (V) SAT scores have the following moments: $\mathbb{E}(M) = 510$, $\mathbb{E}(V) = 475$, $\text{Var}(M) = 750$, $\text{Var}(V) = 610$, and $\rho_{M,V} = 0.4$. What is the expected value and variance of the total SAT, $T = M + V$?

4. A random sample of size n is selected from a population with $\sigma = 10$. What is the standard error of the mean if

   (a) $n = 500$?

   (b) $n = 5,000$?

   (c) $n = \infty$? Compare this result with the previous parts.

5. Final grades in a class are a weighted average of the midterm (25%) and final (75%) exams. Each exam has 100 possible points. Suppose the average and standard deviation of scores on the midterm were 71 and 19 respectively, while the values for the final exam were 69 and 23. Suppose further that the correlation coefficient between the two exams is 0.50.

   (a) What is the mean and standard deviation of the final grades in class?

   (b) Suppose the final grades are normally distributed with mean and variance found in part a). What fraction of students will get an A if they need more than 93 points to obtain that grade?

6. Show that the sample variance is an unbiased estimator of the population variance.

7. A paint manufacturer advertises that its exterior paint will last 5 years. Assume paint life is normally distributed with a standard deviation of 0.5 years.

   (a) Suppose a local TV reporter tests this claim, paints one house, and notices that the house paint only lasts 4.5 years. Would you consider this evidence against the manufacturer's claim?

   (b) Suppose instead of the TV reporter's test, a consumer magazine paints 10 houses and finds the average life of the paint is 4.5 years. Would you consider this evidence against the manufacturer's claim?

8. A water bottler sells spring water in 1 liter bottles. The machines are set such that on average, 1.02 liters are dispensed with a known standard deviation of 0.06 liters. The firm routinely collects a random sample of bottles and tests whether the machine is dispensing correctly. Suppose the machine is working properly (μ and σ are known). What is the chance that in a random sample of 16 bottles, the sample mean will be within 0.05 liters of the specified level?

9. A student surveys 60 undergraduates to determine the average number of drinks consumed over the past two weeks (X). Based on previous surveys, the student believes the standard deviation of drinks per week in the population is 7. The researcher would like to reduce the standard error of by increasing the sample size.

   (a) If the student's beliefs are correct, what will be the standard error of X?

   (b) How many students would the student have to survey to cut the standard error of in half?

   (c) How many students would the research have to survey to cut the standard error by 75%?

10. The covariance between x and y represents the average of the products of the deviations of x and y from their respective means. In other words, it represents the average of the sum of the cross-products. Show that the sum of cross-product can be written either as $\sum_i (x_i - \bar{x}) y_i$ or $\sum_i (y_i - \bar{y}) x_i$.

Extra Problems

The following problems will not be counted for credit on this problem set, but you should know how to do them. They are good practice for the midterm!

1. In 2000, a *Time*/CNN polled 589 voters. If the population proportion for a candidate is $p = 0.5$ and $\bar{p}$ is the sample proportion of likely voters that favor this candidate,

   (a) Show the sampling distribution of $\bar{p}$

   (b) What is the probability that the poll will provide a sample proportion within ± 0.04 of the population proportion?

2. A random sample of size $n = 1,000$ is drawn from a population with $p = 0.4$, where p is the proportion of the population that has a certain characteristic.

   (a) What is the expected value of $\bar{p}$ and its standard error?

   (b) Illustrate the sampling distribution of $\bar{p}$.