SUPPLEMENTARY EXERCISES, CHAPTER 5

INTRODUCTION TO INFERENCE: ESTIMATION AND PREDICTION

S5.1. Distortion product autoacoustic emission (DPOAE) is a method for testing human hearing by measuring emissions produced by the cochlea (a part of the ear) in response to an input signal. For one set of 20 subjects randomly-selected from adults with normal hearing, the signal-noise ratio of the response to a 5 kHz input signal had a mean of 21.11 and a variance of 61.03. The data showed no evidence of nonnormality or outliers.

a. What is the target population to which inference will be made?

b. Compute a 99% confidence interval for the true mean signal-noise ratio of this population.

c. Explain what 99% confidence means for the interval in b.

d. A new subject has her hearing tested with a 5 kHz input signal, and a signal-noise ratio of 5.19 is obtained. Construct an appropriate level 0.90 interval based on the given data to decide if her hearing is normal.

e. Obtain a range of signal-noise ratios that you are 99% certain contain 95% of all signal-noise ratios for adults with normal hearing. (NOTE: negative values of the signal-noise ratio are legitimate since the ratio is in logged units. A negative value merely means that the noise is larger than the signal.)

S5.2. It seems as if many, if not most, retail prices end in a 9 (e.g., \$29.99). To see if this is true, students in a business course took a random sample of 100 retail prices, and found that 19 ended in a 9. Obtain a 99% approximate score confidence interval for the proportion p of all retail prices that end in a 9. Is your result consistent with what you found in exercise? Justify your conclusion.

S5.3. The Rockwell hardness of a random sample of 15 shearing pins taken from a large production lot has mean 49.12 and standard deviation 1.29. The data appear stationary and normal.

a. Obtain a point estimate for the mean Rockwell hardness of the production lot.

b. Obtain a 95% confidence interval for the mean Rockwell hardness of the production lot. Tell what "95% confidence" means here.

c. Obtain an interval you are 95% confident will contain the Rockwell hardness measurement of the next shearing pin sampled from the lot.

d. The quality supervisor is interested in an interval that she can be 90% confident will contain at least 99% of the Rockwell hardness measurements of all shearing pins in the lot. Construct such an interval.

e. Assuming you know the population standard deviation to be 1.3, how large a sample would be required to estimate the mean Rockwell hardness to within 0.1 with 95% confidence?

S5.4. Suppose we have a normal population with known variance σ^2, and that we want to estimate its mean μ to within one-tenth of a standard deviation (i.e., 0.1σ) with confidence 0.90. What sample size is needed?

S5.5. Piston rings for an automobile engine are produced by a forging process. To monitor quality, one piston ring is selected each hour from the production line for evaluation. Measurements of the inside diameters of 15 consecutively-selected rings have a mean of 74.998 mm, and a standard deviation of 0.007 mm. The data appear stationary and normal.

a. Obtain a point estimate for the process mean inner diameter.

b. Obtain a 95% confidence interval for the mean. Tell what "95% confidence" means here.

c. Obtain an interval you are 95% confident will contain the inner diameter measurement of the next piston ring produced.

d. The quality supervisor is interested in an interval that she can be 90% confident will contain at least 99% of the population of all possible measurements. Construct such an interval.

S5.6. In an experiment to test the effectiveness of nicotine patches in curbing cigarette smoking, a group of 200 volunteers was randomly divided into two groups of 100 each. Subjects in the first group were given a six week program of counseling, while those in the second group were given a six week program that included use of nicotine patches in addition to the counseling. Subjects were followed for one year after the program ended. Of the nicotine patch group, 26 remained non-smoking for the entire year, while of the group who received counseling alone, 86 resumed smoking within the year.

a. Let p_1 denote the population proportion of smokers who remain non-smoking for at least one year after the counseling and nicotine patch treatment, and p_2 the population proportion of smokers who remain non-smoking for at least one year after the counseling-only treatment. Obtain a point estimate of $p_1 - p_2$.

b. Construct an approximate 99% score confidence interval for $p_1 - p_2$. Make sure the necessary assumptions are satisfied to ensure a good approximation.

c. Does the interval you constructed provide convincing evidence the nicotine patch/counseling treatment is more effective than counseling treatment alone? Support your answer.

S5.7. The proper operation of a nuclear power plant depends on the reactor feedwater system. If the concentrations of metals, such as iron or copper, are too high, reactor performance can be adversely affected. In order to monitor metal concentrations, measurements are taken each day. A recent set of copper concentration measurements, one taken on each of 15 successive days, had a mean of 0.2631 ppb and a standard deviation of 0.0129 ppb. It is desired to perform statistical inference for the process that produced these measurements.

a. If it is desired to estimate the process mean, what assumptions about the process should be checked?

b. Assume you have checked the assumptions in (a), and there is no evidence that the assumptions do not hold.

  i. Obtain a point estimate for the process mean copper concentration.

  ii. Obtain a 95% confidence interval for the mean. Tell what "95% confidence" means here.

  iii. Obtain an interval you are 95% confident will contain tomorrow's measurement.

  iv. Management is interested in an interval that it can be 95% confident will contain at least 99% of the population of all possible measurements. Construct such an interval.

S5.8. A survey is to be done to estimate the proportion of a population with a certain characteristic. How large a sample is needed to be 99% confident that the estimate will be within 0.02 of the true proportion?

S5.9. Infection with an apparently harmless, newly recognized virus called hepatitis G, seems to interfere with HIV, slowing its progress and prolonging survival of AIDS patients, according to an Associated Press article in the September 6, 2001 edition of the Boston Globe. A study by researchers at the Iowa City Veterans Affairs Medical Center and the University of Iowa looked at 362 HIV-infected patients treated between 1988 and 1999. The researchers found that 144 of these patients were also infected with hepatitis G. Of the 144 who were infected with hepatitis G, 41 died during four years of follow-up compared with 123 of the 218 who were not infected with hepatitis G.

Assume the 144 patients with both HIV and hepatitis G form a random sample from the population of all such individuals, and that the second sample of 218 patients form a random sample from the population of all patients with HIV but no hepatitis G.

(a) Obtain a point estimate of the difference between the proportion of the population of patients infected with both HIV and hepatitis G who survive four years and the proportion of the population of patients infected with HIV but not with hepatitis G who survive four years.

(b) Obtain a 95% approximate score confidence interval for the difference estimated in (a).

(c) Explain what "95% confidence" means here.

(d) Explain in layman's terms what the interval you computed in (a) means in terms of the population proportions involved.

S5.10. A computer manufacturer randomly samples computer mice from the production line and subjects them to wear tests. In one such test, the left mouse button is pressed repeatedly by a testing machine until it fails. Data on the number of press-release cycles until failure in 12 test mice has mean 6709.6 and standard deviation 117.2. The process is stationary, and the data show no evidence of nonnormality.

(a) Construct a level 0.99 confidence interval for the population mean cycles to failure.

(b) Construct an interval based on the 12 values in the data set which you are 99 percent confident will contain the number of cycles to failure of the next mouse tested. If the next mouse fails after 6208 cycles, what will you conclude based on the interval you calculated?

(c) Construct an interval you are 99% confident will contain at least 95% of all cycles to failure values in the population. Interpret this interval in terms of "99% confidence".

S5.11. Consider a level L confidence interval for the mean μ in the C+E model where σ is known. If I want to double the precision of the interval (i.e., make it half as wide), how much must I increase the sample size? How about if I want to triple the precision? Give a general formula for how much I must increase the sample size n to make the interval k times as precise.

S5.12. Market researchers want to estimate the proportion of the population that will respond to an advertizing flyer. To do so, they obtain a random sample of size 50 and find that 5 respond to the flyer.

(a) Obtain a point estimate of the proportion of the population who will respond to the flyer.

(b) Obtain a 95% approximate score confidence interval for the population proportion estimated in (a).

(c) Explain what "95% confidence" means here.

(d) Explain in layman's terms what the interval you computed in (b) means in terms of the population proportions involved.

(e) Suppose the market researchers want to estimate the population proportion to within ± 0.05 with 95% confidence. For planning purposes, they are willing to assume that the true population proportion p equals 0.10. How large a sample should they plan?

S5.13. A manufacturer of concrete beams analyzes the quality of the product by sampling randomly from production and measuring various characteristics. One of these is porosity of the concrete. Porosity measurements on 10 recently tested units give a mean porosity of 0.149 with standard deviation 0.017.

(a) Construct a level 0.90 confidence interval for the population mean porosity.

(b) Construct an interval based on the 10 values in the data set which you are 90 percent confident will contain the porosity of the next unit tested. If the next sample tested has porosity 0.201, what will you conclude based on the interval you calculated?

(c) Construct an interval you are 90% confident will contain at least 99% of the porosities of all units in the population. Interpret this interval in terms of "90% confidence".

(d) For all these intervals, what is being assumed about the population distribution of porosity measurements?