SUPPLEMENTARY EXERCISE SOLUTIONS, 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?

ANS: *All adults with normal hearing.* **(5 points)**

b. **(5 points)** Compute a 99% confidence interval for the true mean signal-noise ratio of this population.

ANS: *Since* $t_{19,0.995} = 2.8609$, *the interval is*

$$21.11 \pm \left(\sqrt{\frac{61.03}{20}}\right)(2.8609) = (16.11, 26.11)$$

c. **(5 points)** Explain what 99% confidence means for the interval in b.

ANS: *If repeated samples are taken from this population, if in each sample the signal-noise ratio in response to a 5 kHz input signal is recorded for each subject, and if a 99% confidence interval for the population mean is calculated from each sample, then approximately 99% of all intervals will contain the population mean.*

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.

ANS: *Since* $t_{19,0.95} = 1.7291$, *the interval (a level 0.90 prediction interval for a new observation) is*

$$21.11 \pm \left(\sqrt{61.03\left(1+\frac{1}{20}\right)}\right)(1.7291) = (7.27, 34.95) \textbf{ (5 points)}$$

Her hearing does not appear normal, since 5.19 lies outside the interval. **(5 points)**

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.)

ANS: **(5 points)** *The constant for a normal-theory tolerance interval is 3.168. Therefore, the interval is*

$$21.11 \pm (\sqrt{61.03})(3.168) = (-3.63, 45.86)$$

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 S4.4? Justify your conclusion.

ANS: *Since* $z_{.995} = 2.5758$,

$$\tilde{p} = \frac{19 + 0.5 \cdot 2.5758^2}{100 + 2.5758} = 0.2093$$

The interval is then

$$0.2093 \pm 2.5758\sqrt{\frac{(0.2093)(1-0.2093)}{100}} = (0.1045, 0.3141) \textbf{ (5 points)}$$

The result is consistent with that of problem 3 since the interval contains only values greater than 0.1, the probability of obtaining a 9 if it is no more likely to occur than any other integer. **(5 points)**

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?

ANS: (5 points) *The formula for sample size is* $n = (\sigma^2 \cdot z^2_{(1+L)/2})/d^2$. *Here* $L = 0.90$, *so* $z_{(1+L)/2} = z_{0.95} = 1.6449$. *Also,* $d = 0.1\sigma$, *so* $n = (\sigma^2 \cdot 1.6449^2)/(0.1\sigma)^2 = 270.5696$, *so take* $n = 271$ *observations.*

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 $p_1 - p_2$.

ANS: $\hat{p}_1 - \hat{p}_2 = 26/100 - 14/100 = 0.12$ **(5 points)**

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.

ANS: *Since* $z_{.995} = 2.5758$,

$$\tilde{p}_1 = \frac{26 + 0.25 \cdot 2.5758^2}{100 + 0.5 \cdot 2.5758^2} = 0.2677$$

and

$$\tilde{p}_2 = \frac{14 + 0.25 \cdot 2.5758^2}{100 + 0.5 \cdot 2.5758^2} = 0.1466$$

so the interval is

$$0.2677 - 0.1466 \pm 2.5758\sqrt{\frac{0.2677(1-0.2677)}{100} + \frac{0.1466(1-0.1466)}{100}} = (-0.0306, 0.2629)$$

(5 points)

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

ANS: *No, since the interval contains 0, we cannot conclude that nicotine patch/counseling treatment is more effective than counseling treatment alone.* **(5 points)**

5.4. Can use S_p interval since S_1 and S_2 are close. Since $s_p = \sqrt{\frac{(9)(0.086)+(7)(0.082)}{16}} = 0.29$, a level L interval for $\mu_1 - \mu_2$ is

$$(3.62 - 3.18) \pm (0.29)\left[\sqrt{\frac{1}{10} + \frac{1}{8}}\right] t_{16,\frac{1+L}{2}}.$$

$$\left.\begin{aligned} L &= 0.90,\ t_{16,.95} = 1.7459 \Rightarrow (0.20, 0.68) \\ L &= 0.95,\ t_{16,.975} = 2.1199 \Rightarrow (0.15, 0.73) \\ L &= 0.99,\ t_{16,.995} = 2.9208 \Rightarrow (0.04, 0.84) \end{aligned}\right\} \text{Most likely one of these. } \textbf{(5 points)}$$

Interpretation

(a) Level L: In repeated sampling, a proportion L of all intervals will contain $\mu_1 - \mu_2$.**(5 points)**

(b) In all cases above, the interval is contained in $(0, \infty)$, which leads us to conclude $\mu_1 > \mu_2$.**(5 points)**