STA 580 — Fall 2008 — Dr. Charnigo

Written Assignment 4 Solutions

1a. We have $\hat{p}_1 = 74/576 = 0.1285$, $\hat{p}_2 = 27/533 = 0.0507$, and $\hat{p} = 101/1109 = 0.0911$. The test statistic is

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(1/576 + 1/533)}} = \frac{0.0778}{0.0173} = 4.50,$$

which exceeds 1.96. So, we reject H_0 at level 0.05. [The p-value is less than 0.0001.]

1b. We have $E_{11} = 576 \times 101/1109 = 52.458$, $E_{21} = 48.542$, $E_{12} = 523.542$, and $E_{22} = 484.458$. The test statistic is

$$\chi^2 = \frac{(74 - 52.458)^2}{52.458} + \frac{(502 - 523.542)^2}{523.542} + \frac{(27 - 48.542)^2}{48.542} + \frac{(506 - 484.458)^2}{484.458} = 20.25,$$

which exceeds $3.84 = \chi^2_{1,0.95}$. So, we reject H_0 at level 0.05. [The p-value is less than 0.0001.]

2a. [Exhibit the box plots.] Applying the procedure in Section 8.4 to the length-of-stay measurements seems inappropriate:

• The distribution of length-of-stay measurements appears non-normal at the second hospital; although this impression is largely created by the observation with length of stay 238, we are reluctant to dismiss even a single observation when the sample size is 13. The procedure in Section 8.4 would require the distribution of length-of-stay measurements to be (approximately) normal in each hospital.

• Even without the 238 there appears to be greater variability in length-of-stay measurements at the second hospital. The procedure in Section 8.4 would require the variance in length-of-stay measurements at the first hospital to be (approximately) equal to that at the second hospital.

2b. The sum of ranks is 83.5 for the first hospital and 216.5 for the second hospital. If the median length of stay at the first hospital were equal to that at the second hospital, then we would have expected a rank sum of $11(11+13+1)/2 = 137.5$ for the first hospital and a rank sum of $13(11+13+1)/2 = 162.5$ for the second hospital. The absolute value of the difference between 83.5 and 137.5, or between 216.5 and 162.5, is 54. Using Equation 9.7, and noting the presence of three two-way ties, we obtain

$$z = \frac{54 - 1/2}{\sqrt{\left(\frac{11(13)}{12}\right)\left[25 - \frac{6+6+6}{24(23)}\right]}} = 3.10.$$

Since the test statistic exceeds 1.96, we conclude that median length of stay differs between the two hospitals. [The p-value is 0.002.]

2c. After logarithmic transformation the sample means are 3.006 for the first hospital and 4.128 for the second hospital. The sample variances are 0.594 and 0.605, yielding a pooled variance estimate of

$$\frac{0.594 \times 10 + 0.605 \times 12}{22} = 0.600.$$

So, the test statistic is

$$t = \frac{3.006 - 4.128}{\sqrt{0.600(1/11 + 1/13)}} = -3.54.$$

Since the test statistic is less than $-2.07 = -t_{22,0.975}$, we conclude that mean log-transformed length of stay differs between the two hospitals. [The p-value is 0.002.]

2d. Each sample should have size 11 since

$$\frac{(0.594 + 0.605)(1.960 + 1.282)^2}{(3.006 - 4.128)^2} = 10.01$$

rounds up to 11. [Since the hypothesis test itself would employ a t reference distribution rather than a z reference distribution, in practice we might be conservative and bump the 11 up to a 12 or a 13.]

3a. There are $n = 23$ nonzero change scores. Using Equation 9.1 we determine that more than $16.7 = 23/2 + 1/2 + 1.96\sqrt{23/4}$ or fewer than $6.3 = 23/2 - 1/2 - 1.96\sqrt{23/4}$ positive change scores would permit the conclusion of a nonzero median change in periodontal status six months after implementation of the dental education program. Since 15 of the nonzero change scores are positive, we are not entitled to that conclusion. [The p-value is 0.210.]

3b. The sum of the ranks for the positive change scores is $4 \times 20.5 + 5 \times 14.0 + 6 \times 5.5 = 185$. If the median change in periodontal status were zero, then we would have expected a rank sum of $138 = 23(24)/4$ for the positive change scores. Using Equation 9.5, and noting the presence of three multi-way ties, we obtain

$$z = \frac{|185 - 138| - 1/2}{\sqrt{23(24)(47)/24 - (210 + 336 + 990)/48}} = 1.44.$$

Since 1.44 falls between -1.96 and 1.96, we may not conclude that the median change in periodontal status is nonzero. [The p-value is 0.151.]

3c. The test statistic is

$$t = \frac{0.50}{1.77/\sqrt{28}} = 1.49.$$

Since 1.49 falls between $-2.05 = -t_{27,0.975}$ and $2.05 = t_{27,0.975}$, we may not conclude that the mean change in periodontal status is nonzero. [The p-value is 0.148.]

3d. All three testing procedures yield results that are in qualitative agreement. Strictly speaking the change scores cannot be normally distributed, as they are realizations of a discrete random variable. However, the discretization may be fine enough that we are comfortable with the paired t-test. If we are, then the paired t-test is preferred because the signed-rank test and sign test discard information. If we are not comfortable with the paired t-test, then the signed-rank test is preferred because it discards less information than the sign test.