STAT 302H: Principles of Biostatistics

Homework 5 Solution (due 10/29/07)

Instructor: James A. Calvin *TA: Ying Sun*

11.5

(a) The samples are paired.

(b) $H_0 : \mu_{\text{corn}} - \mu_{\text{oats}} = 0$, $H_A : \mu_{\text{corn}} - \mu_{\text{oats}} \neq 0$.

(c) The difference in LDL cholesterol levels for each person is $\bar{d} = 0.363$, and $s_d = 0.406$.

The test statistic is $t = \frac{\bar{d}-\delta}{s_d/\sqrt{n}} = 3.35$, $df = 14 - 1 = 13$, $0.001 < p < 0.01$. We reject H_0 at the 0.05 level of significance.

(d) We conclude that the true difference in population mean cholesterol levels is not equal to 0. Mean LDL cholesterol is lower when individuals are adhering to the oat bran diet.

11.7

(a) $\bar{d} = 39.4$, and $s_d = 31.4$. $df = 7 - 1 = 6$, 95% of the values lie above -1.943. Therefore a one-sided 95% confidence interval for the true difference in population means $\delta = \mu_{12} - \mu_{24}$ is $\delta \geq \bar{d} - 1.943(s_d/\sqrt{n}) = 16.3$.

(b) $H_0 : \mu_{12} - \mu_{24} \leq 0$, $H_A : \mu_{12} - \mu_{24} > 0$. Given that $\delta = 0$, the test statistic is $t = 3.32$. Since $0.005 < p < 0.01$, we reject H_0 at the 0.05 level of significance and conclude that the true difference in population mean cotinine levels is not equal to 0. mean cotinine level decreased significantly between 12 and 24 hours after smoking.

3

(a)

$$s_p^2 = \frac{36 \times 5.6^2 + 18 \times 21.7^2}{36 + 18} = 177.87$$

$$TS = \frac{27.9 - 38.8}{\sqrt{177.87/37 + 177.87/19}} = -2.896$$

$df = 54$, $t_{54,0.025} = 2.009$, $|TS| > 2.009$, we reject H_0.

(b) We would not expect 95% CI to include 0, because in part(a), we reject the H_0, there is a difference between two means. Including 0 means Fail to reject H_0.

(c)

$$TS = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} = -2.153$$

$df = 19.24 = 19$, $t_{19,0.025} = 2.093$, $|TS| > 2.093$, Reject H_0.

11.13

(a) mean, median, min and max can be calculated.

(b) use stat tools $TS = 1.5396$, $p = 0.1268$, Fail to reject.

(c) $TS = 6.8721$, $p < 0.0001$ Reject

(d) We do not reach the same conclusion using the two different tests. The two-sample t-test assumes that the populations of interest are independent and suggests that the mean number of hospital beds in 1986 is not significantly different from the mean number of beds in 1980. On the other hand, the paired t-test indicates that the mean number of beds has decreased over the six-year period. Pairing helps to control for extraneous sources of variation and should be taken into account when applicable. If pairing is ignored, we might fail to reject a false null hypothesis.

(e) 95% CI is $(0.23, 0.42)$.

14.7

(a) $\hat{p} = 0.556$, CI is $\hat{p} \pm 1.96\sqrt{\hat{p}(1-\hat{p})/n}$, that is $(0.369, 0.743)$.

(b) $H_0 : \mu = 0.328$

(c) $H_A : \mu \neq 0.328$

(d) $z = \frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}} = 2.52$, Fail to reject.

(e) We conclude that for children with an oral cleft, there is no evidence that the proportion of mothers who smoked during pregnancy is different from the proportion of mothers who smoked for children with other types of malformations.

(f) $p_0 = 0.328$, $p_1 = 0.250$, $\alpha = 0.01$ for a two-sided test and $\beta = 0.10$, we have that $z_{\alpha/2} = 2.58$, and $z_\beta = 1.28$,

$$n = \left[\frac{2.58\sqrt{p_0(1-p_0)} + 1.28\sqrt{p_1(1-p_1)}}{p_1 - p_0}\right]^2 = 512.3$$

A sample of size 513 would be required.

14.12

(a) $\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1+n_2} = 0.20$

(b)

$$TS = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(1/n_1 + 1/n_2)}} = 1.417$$

$df = 256$

(c) Fail to reject

(d)

$$(\hat{p}_1 - \hat{p}_2) \pm 1.96\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

That is $(-0.031, 0.177)$.

14.13

(a) 0.21 (b) (0.135,0.303) (c) exact binomial interval, or (b) (0.130,0.290) (c) Normal approximation