STA 570 — Spring 2012 — Dr. Charnigo

Written Assignment 3 Solutions

1a. Using notation as in Lecture 7, the test statistic is

$$\chi^2 = \frac{(n_1-1)s_x^2}{900} = \frac{(48)1172.6}{900} = 62.54.$$

The upper critical value is

$$\chi^2_{n_1-1,1-\alpha/2} = \chi^2_{48,0.975} = 69.02,$$

and the lower critical value is

$$\chi^2_{n_1-1,\alpha/2} = \chi^2_{48,0.025} = 30.75.$$

Since 62.54 is neither greater than 69.02 nor less than 30.75, we do not reject $H_0 : \sigma_1^2 = 900$ in favor of $H_1 : \sigma_1^2 \neq 900$.

1b. We have

$$s^2 = \frac{(n_1-1)s_x^2 + (n_2-1)s_y^2}{n_1+n_2-2} = \frac{(48)1172.6 + (49)887.4}{97} = 1028.5.$$

The test statistic is

$$t = \frac{\bar{x}-\bar{y}}{\sqrt{s^2(1/n_1+1/n_2)}} = \frac{112.6-136.5}{\sqrt{1028.5(1/49+1/50)}} = \frac{-23.9}{6.45} = -3.71.$$

The (lower) critical value is

$$-t_{n_1+n_2-2,1-\alpha/2} = -t_{97,0.975} = -1.985.$$

Since $-3.71 < -1.985$ we reject $H_0 : \mu_1 = \mu_2$ in favor of $H_1 : \mu_1 \neq \mu_2$.

1c. The 95% confidence interval for $\mu_1 - \mu_2$ assuming $\sigma_1^2 = \sigma_2^2$ is

$$\bar{x}-\bar{y} \pm t_{n_1+n_2-2,1-\alpha/2}\sqrt{s^2(1/n_1+1/n_2)},$$

which is

$$-23.9 \pm 1.985(6.45)$$

or -36.7 to -11.1. Exclusion of 0 from the 95% confidence interval for $\mu_1-\mu_2$ assuming $\sigma_1^2 = \sigma_2^2$ corresponds to rejection of $H_0 : \mu_1 - \mu_2 = 0$ in favor of $H_1 : \mu_1 - \mu_2 \neq 0$ at level $\alpha = 0.05$ assuming $\sigma_1^2 = \sigma_2^2$.

1d. We have

$$df = \frac{(s_x^2/n_1 + s_y^2/n_2)^2}{(s_x^2/n_1)^2/(n_1-1) + (s_y^2/n_2)^2/(n_2-1)} = \frac{(1172.6/49+887.4/50)^2}{(1172.6/49)^2/(48)+(887.4/50)^2/(49)} = \frac{1737}{18.36} = 94.6.$$

The test statistic is

$$t = \frac{\bar{x}-\bar{y}}{\sqrt{s_x^2/n_1 + s_y^2/n_2}} = \frac{112.6-136.5}{\sqrt{1172.6/49+887.4/50}} = \frac{-23.9}{6.46} = -3.70.$$

The (lower) critical value is

$$-t_{df,1-\alpha/2} = -t_{94.6,0.975} = -1.985.$$

Since $-3.70 < -1.985$ we reject $H_0 : \mu_1 = \mu_2$ in favor of $H_1 : \mu_1 \neq \mu_2$.

1e. The 95% confidence interval for $\mu_1 - \mu_2$ assuming $\sigma_1^2 \neq \sigma_2^2$ is

$$\bar{x} - \bar{y} \pm t_{df,1-\alpha/2}\sqrt{s_x^2/n_1 + s_y^2/n_2},$$

which is

$$-23.9 \pm 1.985(6.46)$$

or -36.7 to -11.1. Exclusion of 0 from the 95% confidence interval for $\mu_1 - \mu_2$ assuming $\sigma_1^2 \neq \sigma_2^2$ corresponds to rejection of $H_0 : \mu_1 - \mu_2 = 0$ in favor of $H_1 : \mu_1 - \mu_2 \neq 0$ at level $\alpha = 0.05$ assuming $\sigma_1^2 \neq \sigma_2^2$.

1f. The test statistic is

$$f = \frac{s_x^2}{s_y^2} = \frac{1172.6}{887.4} = 1.32.$$

The upper critical value is

$$f_{n_1-1,n_2-1,1-\alpha/2} = f_{48,49,0.975} = 1.766,$$

while the lower critical value is

$$f_{n_1-1,n_2-1,\alpha/2} = f_{48,49,0.025} = 0.565.$$

(This lower critical value can also be written as $1/1.769 = 1/f_{49,48,0.975} = 1/f_{n_2-1,n_1-1,1-\alpha/2}$.) Since $0.565 < 1.32 < 1.766$ we accept $H_0 : \sigma_1^2 = \sigma_2^2$. Given that we have accepted $H_0 : \sigma_1^2 = \sigma_2^2$, most people would say that the test and confidence interval in parts b and c are more appropriate than the test and confidence interval in parts d and e.

2. Suppose that large samples are drawn from the two populations of interest. The proposed testing procedure amounts to rejecting $H_0 : \mu_1 = \mu_2$ if the test statistic $z := (\bar{x} - \bar{y})/\sqrt{s_x^2/n_1 + s_y^2/n_2}$ is either greater than $z_{1-\alpha}$ or less than $-z_{1-\alpha}$. However, a standard normal random variable has probability 2α of being either greater than $z_{1-\alpha}$ or less than $-z_{1-\alpha}$. Thus, since the random conceptualization of the test statistic z has an approximate standard normal distribution under $H_0 : \mu_1 = \mu_2$, the proposed testing procedure has approximate significance level 2α.

Similar reasoning (but with obvious modifications, such as replacing z with t) and a similar conclusion apply if we suppose that small samples are drawn from the two populations of interest (and that the populations are normal).

3a. Using notation as in Lecture 6, the test statistic is

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{48/74 - 0.60}{\sqrt{0.60(1-0.60)/74}} = \frac{0.0486}{0.0569} = 0.85.$$

The (upper) critical value is

$$z_{1-\alpha/2} = z_{0.975} = 1.96.$$

Since $-1.96 < 0.85 < 1.96$ we do not reject $H_0 : p = 0.60$ in favor of $H_1 : p \neq 0.60$.

3b. The power with a sample size of 100 (number of persons with diabetes) is

$$\Phi\left[\sqrt{\frac{p_0(1-p_0)}{p_1(1-p_1)}}\left(-z_{1-\alpha/2} + \frac{\sqrt{n}|p_0 - p_1|}{\sqrt{p_0(1-p_0)}}\right)\right] = \Phi\left[\sqrt{\frac{0.60(1-0.60)}{48/74(26/74)}}\left(-1.96 + \frac{\sqrt{100}|0.60 - 48/74|}{\sqrt{0.60(1-0.60)}}\right)\right] =$$

$$\Phi\left[1.026(-0.967)\right] = \Phi[-0.992] \approx 16\%.$$

3c. For 90% power the sample size (number of persons with diabetes) is

$$\frac{p_0(1-p_0)\left(z_{1-\alpha/2} + z_{1-\beta}\sqrt{\frac{p_1(1-p_1)}{p_0(1-p_0)}}\right)^2}{(p_1-p_0)^2} = \frac{0.60(1-0.60)\left(1.96 + 1.282\sqrt{\frac{48/74(26/74)}{0.60(1-0.60)}}\right)^2}{(48/74-0.60)^2} \approx 1045.$$