STAT/MATH 415 HW#4

September 29, 2017

EXERCISES

7.2.3 Independent random samples of the heights of adult males living in two countries yielded the following results: $n = 12$, $\bar{x} = 65.7$ inches, $s_x = 4$ inches and $m = 15$, $\bar{y} = 68.2$ inches, $s_y = 3$ inches. Find an approximate 98% confidence interval for the difference $\mu_x - \mu_y$ of the mean of the populations of heights. Assume that ${\sigma_x}^2 = {\sigma_y}^2$

Answer: Since ${\sigma_x}^2$, ${\sigma_y}^2$ are unknown, we can use ${S_p}^2$ to estimate

$${s_x}^2 = \frac{1}{12-1}\sum_{i=1}^{12}(X_i - \bar{X})^2 \qquad \Rightarrow \qquad \sum_{i=1}^{12}(X_i - \bar{X})^2 = 176$$

$${s_y}^2 = \frac{1}{15-1}\sum_{j=1}^{15}(Y_j - \bar{Y})^2 \qquad \Rightarrow \qquad \sum_{j=1}^{15}(Y_j - \bar{Y})^2 = 126$$

$${S_p}^2 = \frac{1}{12-1+15-1}\Big(\sum_{i=1}^{12}(X_i - \bar{X})^2 + \sum_{j=1}^{15}(Y_j - \bar{Y})^2\Big) = 12.08$$

Then $\alpha = 0.02$, $t_{0.01(25)} = 2.485$, the two sided confidence interval is

$$(\bar{X} - \bar{Y}) \pm t_{0.01(25)} S_p \sqrt{\frac{1}{12} + \frac{1}{15}} = 65.7 - 68.2 \pm 2.485 \times \sqrt{12.08} \times \frac{\sqrt{15}}{10} = -2.5 \pm 3.345$$

The 98% confidence interval for $\mu_x - \mu_y$ is $[-5.845, 0.845]$

7.2.6 A test was conducted to determine whether a wedge on the end of a plug fitting designed to hold a seal onto the plug was doing its job. The data taken were in the form of measurements of the force required to remove a seal from the plug with the wedge in place (say X) and the force required without the plug (say Y). Assume the distributions of X and Y are $N(\mu_X, \sigma^2)$ and $N(\mu_Y, \sigma^2)$ respectively. Ten independent observations of X are

3.26 2.26 2.62 2.62 2.36 3.00 2.62 2.40 2.30 2.40

Ten independent observations of Y are

1.80 1.46 1.54 1.42 1.32 1.56 1.36 1.64 2.00 1.54

a) Find a 95% confidence interval for $\mu_X - \mu_Y$

Answer: Use ${S_p}^2$ to estimate σ^2

$$\bar{X} = 2.584 \quad \sum_{i=1}^{10}(X_i - \bar{X})^2 = 0.93744 \qquad \text{and} \qquad \bar{Y} = 1.564 \quad \sum_{j=1}^{10}(Y_j - \bar{Y})^2 = 0.38544$$

$${S_p}^2 = \frac{1}{10-1+10-1}(0.93744 + 0.38544) = 0.07349$$

Then $\alpha = 0.05,\ t_{0.025(18)} = 2.101$

$$\bar{X} - \bar{Y} \pm t_{0.025(18)} S_p \sqrt{\frac{1}{10} + \frac{1}{10}} = 1.02 \pm 2.101 \times 0.2711 \times 0.4472 = 1.02 \pm 0.2547$$

Thus the 95% confidence interval for $\mu_X - \mu_Y$ is [0.7653, 1.2747]

c) Is the wedge necessary?

Answer: Since there is 95% probability that $\mu_X - \mu_Y$ will fall within [0.7653, 1.2747], so we can conclude that μ_X is significantly greater than μ_Y, that is, the wedge is necessary.

7.2.9 Students in a semester-long health-fitness program have their percentage of body fat measured at the beginning and the end of the semester. The following measurements give these percentages for 10 men and 10 women:

Males		Females	
Pre-program %	Post-program %	Pre-program %	Post-program %
11.10	9.97	22.90	22.89
19.50	15.80	31.60	33.47
14.00	13.02	27.70	25.75
8.30	9.28	21.70	19.80
12.40	11.51	19.36	18.00
7.89	7.40	25.03	22.33
12.10	10.70	26.90	25.26
8.30	10.40	25.75	24.90
12.31	11.40	23.63	21.80
10.00	11.95	25.06	24.28

a) Find a 90% confidence interval for the mean of the difference in the percentages for the males

Answer: Since this is a pairwise measurement, we should turn it into a one-sample problem. Let $M = Male_{pre} - Male_{post}$, ten observations of M are

$$1.13 \quad 3.7 \quad 0.98 \quad -0.98 \quad 0.89 \quad 0.49 \quad 1.4 \quad -2.1 \quad 0.91 \quad -1.95$$

$$\bar{M} = 0.447 \qquad {s_M}^2 = 2.991 \qquad \alpha = 0.1 \qquad t_{0.05(9)} = 1.833$$

$$\bar{M} \pm t_{0.05(9)} \frac{s_M}{\sqrt{10}} = 0.447 \pm 1.002$$

The 90% confidence interval for the mean of the difference is [-0.555, 1.449]

b) Find a 90% confidence interval for the mean of the difference in the percentages for the females

Answer: Similarly, let $F = Female_{pre} - Female_{post}$, ten observations of F are

$$0.01 \quad -1.87 \quad 1.95 \quad 1.9 \quad 1.36 \quad 2.7 \quad 1.64 \quad 0.85 \quad 1.83 \quad 0.78$$

$$\bar{F} = 1.115 \qquad {s_F}^2 = 1.665 \qquad \alpha = 0.1 \qquad t_{0.05(9)} = 1.833$$

$$\bar{F} \pm t_{0.05(9)} \frac{s_F}{\sqrt{10}} = 1.115 \pm 0.748$$

The 90% confidence interval for the mean of the difference is [0.367, 1.863]

c) On the basis of these data, have these percentages decreased?

Answer: We are 90% sure that women's body fat percentages decreased after the program, but not for men.

d) If possible, check whether each set of differences comes from a normal distribution.

Answer: Here shows the normality plot for men's and women's difference in Minitab. Since both p-values are greater than 0.05, we can take the differences as normally distributed.

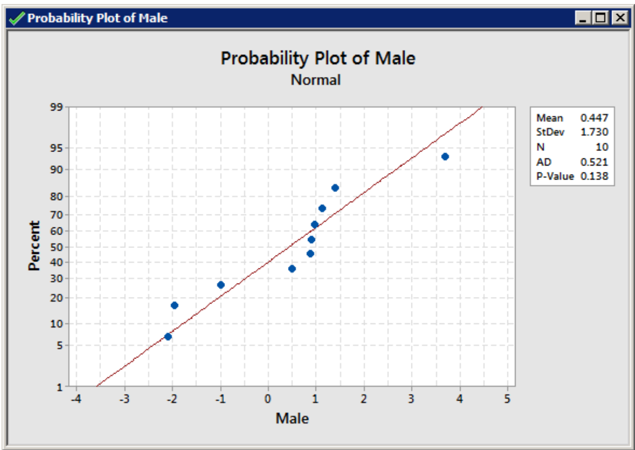

Figure 1: Normality Plot for M

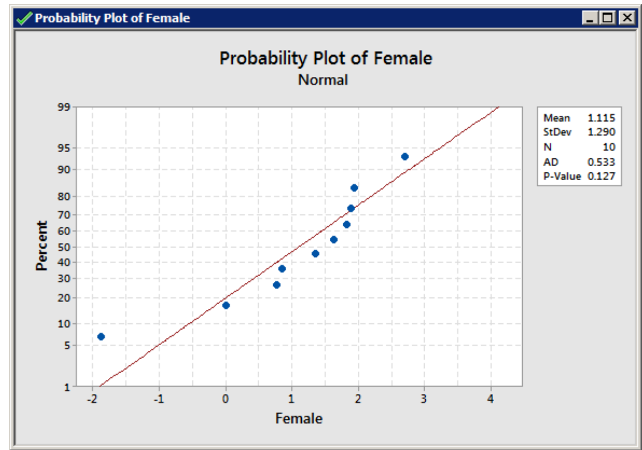

Figure 2: Normality Plot for F

7.2.11 The biomechanics lab tested healthy old women and healthy young women to discover whether lower extremity response time to a stimulus is a function of age. Let X and Y respectively equal the independent response times for these two groups. Find a one-sided 95% confidence interval that is a lower bound for $\mu_X - \mu_Y$ if $n = 60$ observations of X yielded $\bar{x} = 671$ and $s_x = 129$, while $m = 60$ observations of Y yielded $\bar{y} = 480$ and $s_y = 93$

Answer: $\alpha = 0.05$, $t_{0.05(118)} = 1.645$

$$s_x = 129 \quad \Rightarrow \quad \sum_{i=1}^{60}(X_i - \bar{X})^2 = 981819 \qquad s_y = 93 \quad \Rightarrow \quad \sum_{i=1}^{60}(X_i - \bar{X})^2 = 510291$$

$$S_p{}^2 = \frac{1}{60-1+60-1}(981819 + 510291) = 12645$$

$$\bar{X} - \bar{Y} - t_{0.05(118)} S_p \sqrt{\frac{1}{60} + \frac{1}{60}} = 191 - 33.773$$

Thus the one-sided 95% confidence interval for $\mu_X - \mu_Y$ is $[157.227, +\infty)$

7.3.3 Let p equal the proportion of triathletes who suffered a training-related overuse injury during the past year. Out of 330 triathletes who responded to a survey, 167 indicated that they had suffered such an injury during the past year.

a) Use these data to give a point estimate of p.

Answer: The point estimate is $\hat{p} = {}^{167}/_{330} = 0.5061$

b) Use these data to find an approximate 90% confidence interval for p.

Answer: $z_{0.05} = 1.645$, so the interval is $[0.4608, 0.5514]$

$$\hat{p} \pm z_{0.05}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.5061 \pm 1.645 \times 0.0275 = 0.5061 \pm 0.0453$$

7.3.5 In order to estimate the proportion, p, of a large class of college freshmen that had high school GPAs from 3.2 to 3.6 inclusive, a sample of $n = 50$ students was taken. It was found that $y = 9$ students fell into this interval.

a) Give a point estimate of p.

Answer: The point estimate of p is $9/50 = 0.18$

b) Use equation 7.3.2 to find an approximate 95% confidence interval for p.

Answer: $\alpha = 0.05,\ \alpha = 0.05,\ z_{0.025} = 1.96$

$$\frac{y}{n} \pm z_{\alpha/2}\sqrt{\frac{(y/n)(1-y/n)}{n}} = 0.18 \pm 1.96\sqrt{\frac{0.18 \times 0.82}{50}} = 0.18 \pm 0.1065$$

The approximate 95% confidence interval is [0.0735, 0.2865]

c) Use equation 7.3.4 to find an approximate 95% confidence interval for p.

Answer: Let $z = z_{0.025}$

$$\frac{\hat{p} + z^2/(2n) \pm z\sqrt{\hat{p}(1-\hat{p})/n + z^2/(4n^2)}}{1 + z^2/n} = \frac{0.218416 \pm 1.96 \times 0.05776}{1.076832} = [0.098, 0.308]$$

d) Use equation 7.3.5 to find an approximate 95% confidence interval for p.

Answer: Let $\tilde{p} = (y+2)/(n+4) = 0.2037$

$$\tilde{p} \pm z_{0.025}\sqrt{\tilde{p}(1-\tilde{p})/(n+4)} = 0.2037 \pm 1.96 \times 0.0548 = 0.2037 \pm 0.1074 = [0.0963, 0.3111]$$

7.3.9 Consider the following two groups of women: group 1 consists of women who spend less than \$500 annually on clothes; group 2 comprises women who spend over \$1000 annually on clothes. Let p_1 and p_2 equal the proportions of women in these two groups respectively, who believes that clothes are too expensive. If 1009 out of a random sample of 1230 women from group 1 ad 207 out of a random sample of 340 from group 2 believe that clothes are two expensive.

a) Give a point estimate of $p_1 - p_2$

Answer: The point estimate is 0.2115

$$\hat{p_1} - \hat{p_2} = \frac{1009}{1230} - \frac{207}{340} = 0.2115$$

b) Find an approximate 95% confidence interval for $p_1 - p_2$

Answer: $\alpha = 0.05,\ z_{0.025} = 1.96$

$$\hat{p_1} - \hat{p_2} \pm z_{0.025}\sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1} + \frac{\hat{p_2}(1-\hat{p_2})}{n_2}} = 0.2115 \pm 1.96\sqrt{0.00012 + 0.0007} = 0.2115 \pm 0.0561$$

Thus the 95% confidence interval for $p_1 - p_2$ is [0.1554, 0.2676]