Stat 145: Exam 4 Review Answers

1. (a) The sample proportion is:

$$\begin{aligned} \hat{p} &= \frac{459}{850} \\ &= 0.54 \end{aligned}$$

(b) A 95% confidence interval for p is:

$$\begin{aligned} 0.54 \pm 1.96\sqrt{\frac{0.54(1-0.54)}{850}} &= 0.54 \pm 0.0335 \\ &= (0.5065, 0.5735) \end{aligned}$$

(c) The sample size should be:

$$\begin{aligned} n &= \left(\frac{1.96}{0.03}\right)^2 (0.54)(1-0.54) \\ &= 1060.2816 \end{aligned}$$

Rounding up to the next higher whole number, use a sample size of $n = 1061$.

2. (a) The sample mean is:

$$\begin{aligned} \bar{x} &= \frac{58+47+57+55}{4} \\ &= 54.25 \end{aligned}$$

and the sample standard deviation is:

$$\begin{aligned} s &= \sqrt{\frac{(58-54.25)^2+(47-54.25)^2+(57-54.25)^2+(55-54.25)^2}{4-1}} \\ &= \sqrt{\frac{74.75}{3}} \\ &= 4.99 \end{aligned}$$

(b) The hypotheses are:

$$\begin{aligned} H_0 &: \mu = 51 \\ H_A &: \mu > 51 \end{aligned}$$

The value of the test statistic is:

$$\begin{aligned} t &= \frac{54.25-51}{4.99/\sqrt{4}} \\ &\doteq 1.30 \end{aligned}$$

There are $4 - 1 = 3$ degrees of freedom, so the P-value is:

$$\begin{aligned} P &= P(T \geq 1.30) \\ &\rightarrow \text{ between .10 and .15} \end{aligned}$$

We fail to reject the null hypothesis and conclude that the current mean retail price for bananas is not greater than 51 cents per pound.

3. (a) The stemplot appears below.

2	79
3	45
4	26788
5	36
6	144
7	1

(b) Yes, because the distribution is roughly Normal with no apparent outliers.

4. (a) The degrees of freedom are equal to 31. This value is not on the table, so use the row with 30 df. A 90% confidence interval for $\mu_1 - \mu_2$ is:

$$\begin{aligned} (383.4 - 349.6) \pm 1.697\sqrt{\frac{37.2^2}{32} + \frac{34.6^2}{40}} &\doteq 33.8 \pm 14.5 \\ &= (19.3, 48.3) \end{aligned}$$

(b) The hypotheses are:

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 \\ H_A &: \mu_1 \neq \mu_2 \end{aligned}$$

The value of the test statistic is:

$$\begin{aligned} t &= \frac{383.4 - 349.6}{\sqrt{\frac{37.2^2}{32} + \frac{34.6^2}{40}}} \\ &\doteq 3.95 \end{aligned}$$

The P-value is:

$$\begin{aligned} P &= P(T \leq -3.95) + P(T \geq 3.95) \\ &= 2P(T \geq 3.95) \\ &\rightarrow \text{ less than } 2(.0005) = .001 \end{aligned}$$

We reject the null hypothesis and conclude that the mean potato yields for the two states differ.

5. The hypotheses are:

   H_0: there is not a relationship between gender and party identification
   H_A: there is a relationship between gender and party identification

   The expected cell counts if H_0 is true appear below the observed count in the following table:

	Party Identification			
Gender	Democrat	Independent	Republican	Total
Male	279	73	225	577
	261.4	70.7	244.9	
Female	165	47	191	403
	182.6	49.3	171.1	
Total	444	120	416	980

   For example, the expected count for the (Male, Democrat) cell is:

$$\frac{577 \times 444}{980} = 261.4$$

   The value of the chi-square statistic is:

$$\begin{aligned}
\chi^2 &= \frac{(279 - 261.4)^2}{261.4} \\
&+ \frac{(73 - 70.7)^2}{70.7} \\
&+ \frac{(225 - 244.9)^2}{244.9} \\
&+ \frac{(165 - 182.6)^2}{182.6} \\
&+ \frac{(47 - 49.3)^2}{49.3} \\
&+ \frac{(191 - 171.1)^2}{171.1} \\
&\doteq 7.1
\end{aligned}$$

The degrees of freedom are:

$$df = (2-1)(3-1) = (1)(2) = 2$$

The P-value is between .025 and .05. We reject the null hypothesis and conclude that there is a relationship between gender and party identification.