Homework 3 Solutions
Chapter 6: #40; Chapter 15: #10 (Use R Commander, Excel, or by hand)
Chapter 3: #23b, 63a to d
Chapter 4: #5c, 39; Chapter 5: #51

Assigned Friday, October 9

6.40 Null hypothesis: Feeling about importance of religion and opinion on same-sex marriage are not related variables (in the national adult population).
Alternative hypothesis: Feeling about importance of religion and opinion on same-sex marriage are related variables (in the national adult population).

15.10 **a.** For females, 90.9% said "most times or always" and 9.1% said "rarely or never." For males, 78.4% said "most times or always" and 21.6% said "rarely or never." Females were more likely to wear a seatbelt.

b. Null: Gender and seatbelt use are not related for the population of 12th graders.

Alternative: Gender and seatbelt use are related for the population of 12th graders.

Parts c to f by hand, except using Excel to find the p-value:

c. Expected Count = $\frac{\text{Row Total} \times \text{Column Total}}{\text{Total } n} = \frac{1178 \times 1888}{2239} = 993.33.$

d. Female, Most times or always: 1888 – 993.33 = 894.67; Male, Rarely or never: 1178 – 993.33 = 184.67; Female, Rarely or never: 1061 – 894.67 = 166.33.

e. $\chi^2 = \frac{(964 - 894.67)^2}{894.67} + \frac{(97 - 166.33)^2}{166.33} + \frac{(924 - 993.33)^2}{993.33} + \frac{(254 - 184.67)^2}{184.67} = 65.14$

f. The *p*-value is essentially 0. With Excel, *p*-value = CHIDIST(65.14,1) = 6.98×10^{-16}.

Parts a and c to f using R Commander (showing results only, not the R commands):

```
> .Table  # Counts (These are the observed counts)
    1   2
1 964  97
2 924 254

> rowPercents(.Table) # Row Percentages (For Part a)
     1    2 Total Count
1 90.9  9.1   100  1061
2 78.4 21.6   100  1178

      Pearson's Chi-squared test
X-squared = 65.1365, df = 1, p-value = 6.989e-16

> .Test$expected # Expected Counts
         1        2
1 894.6708 166.3292
2 993.3292 184.6708

> round(.Test$residuals^2, 2) # Chi-square Components
     1     2
1 5.37 28.90
2 4.84 26.03
```

g. The null hypothesis can be rejected because p-value < .05. We conclude that there is a relationship between gender and seat belt use in the population of 12th graders.

Assigned Monday, October 12

3.23 **a.** $\frac{1}{\sqrt{n}} \times 100\% = \frac{1}{\sqrt{1031}} \times 100\% = 3.1\%$.

b. 66% ± 3.1%, or 62.9% to 69.1%.

3.63 **a.** The margin of error is $\frac{1}{\sqrt{n}} = \frac{1}{\sqrt{1016}} = .031$, or about 3.1%.

b. This can be done by focusing on the results for the *first survey question*. For example, "About two-thirds of American adults favor teaching creationism in the public schools."
c. The sentences would emphasize the results for the *second survey question*. For instance, "A majority of U.S. adults are opposed to teaching creationism in public schools, as indicated by a survey of 1016 adults in which 55% said they were opposed"
d. A simple statement of the survey percentages should be sufficient. For example, "About 68% of U.S. adults favor teaching creationism along with evolution, while 29% are opposed to doing so. About 40% favor teaching creationism instead of evolution, while 55% are opposed to that option."

Assigned Wednesday, October 14

4.5 **c.** A confounding variable affects the response variable: Mother's educational level (confounding variable) may affect Child's IQ at age 10 (response variable) because less educated mothers may not provide as much intellectual stimulation for their children. A confounding variable is related to the explanatory variable: Mother's educational level (confounding variable) is related to whether the mother smokes (explanatory variable) – less educated mothers are more likely to be smokers.

4.39 A case-control study would be best. In such a study, the cellular phone usage habits of people with brain cancer (cases) would be compared to the cellular phone use of people without brain cancer (controls). Matching cases to controls with regard to factors like age would help reduce the effects of any potential confounding variables. An observational study that did not find cases but just used a sample from the general population probably would have too few cases of brain cancer to make any conclusions. A randomized experiment would be inefficient, and perhaps impossible. If an experiment were done, some people would be assigned to use a cellular phone and others would be assigned to not do so. The experimenters would then observe who develops brain cancer. Such an experiment would take too long, many people would not comply if assigned to not use cellular phones, and there would be too few cases of brain cancer observed to make meaningful conclusions.

5.51 **a.** Both variables may be increasing due to an increasing population during this time.
b. There may be causation, but there is the possibility of confounding. Perhaps people who walk more also smoke less. And, if there is causation, it could go in either direction. Regular walking might lead to better health and better health might allow the men to walk more.
c. Both variables are related to the number of people at the ski resort on a given day.