STA 2210 Homework 7 (Due on Mon. 7/13 by 11:59pm)

Directions: Write your R codes, in addition to your answer, to the following problems. You may find the “prop.test()” function to be especially helpful on this assignment. Refer to the R reference card for more helpful commands.

1. **A random sample of 40 American adults was asked the question “Do you support the idea of a federally funded health care system?” The data set HEALTH, a .csv file, contains the results.** 
   1. **Do the data provide significant evidence that the majority of Americans support such a system? Perform a hypothesis test using . Be sure to check that the conditions for inference are satisfied.**

Ho: Majority of American adults do not support the health care system (p <= 0.5).

Ha: Majority of American adults do support the health care system (p > 0.5).

The condition of inference is checked because the sample size is less than 10% of the overall population.

table(HEALTH$UHC)

Does not support Support

13 27

The condition of success and failure is met because there are 27 successes and 13 failures, which are both greater than 10. This means we can continue with the 1-prop Z test.

prop.test(27, 40, p = 0.5, alternative = "greater", correct = TRUE)

1-sample proportions test with continuity correction

data: 27 out of 40, null probability 0.5

X-squared = 4.225, df = 1, p-value = 0.01992

alternative hypothesis: true p is greater than 0.5

95 percent confidence interval:

0.532742 1.000000

sample estimates:

p

0.675

The p value is 0.01992, which is less than 0.05, so we reject the null hypothesis.

* 1. **Interpret the decision of the test in the context. In your interpretation, also describe the type of error that may have been made and what the error means in the context.**

Since the null hypothesis is rejected, we support the alternative hypothesis and say that a majority of American adults do support the federally funded health care system. We could have made a Type 1 error, where we say that the null hypothesis is false, when it really is true. This would mean we think a majority of Americans believe in the federally funded health care system, but in reality, a minority or exactly 50% of Americans would believe in it if we made an error in our test.

1. **The data set CES11.csv contains data from a study titled “2011 Canadian National Election Study, With Attitude Toward Abortion.” Assume the data represents a random sample of Canadian adults.** 
   1. **Do the data provide significant evidence that the proportion of Canadian females who think that abortion should not be banned is different that the proportion of Canadian males? Perform the hypothesis test using . Be sure to check that the conditions for inference are satisfied.**

Ho: No difference between male’s and female’s belief that abortion should not be banned in Canada (pf = pm).

Ha: Difference between male’s and female’s belief that abortion should not be banned in Canada (pf =/= pm).

female <- subset(CES11, CES11$gender == 'Female')

male <- subset(CES11, CES11$gender == 'Male')

table(CES11$abortion, CES11$gender)

Female Male

No 1012 806

Yes 232 181

The condition of independence is met because it is a random sample and the number of samples of males and females is less than 10% of the population of Canada. The condition of success-failure is met because the numbers of successes and failures for both genders is greater than 10. We don’t need to calculate this with the p value because we aren’t given a p value.

> res <- prop.test(x = c(1012, 806), n = c(1244, 987))

> res

2-sample test for equality of proportions with continuity correction

data: c(1012, 806) out of c(1244, 987)

X-squared = 0.017701, df = 1, p-value = 0.8942

alternative hypothesis: two.sided

95 percent confidence interval: -0.03644416 0.03022179

sample estimates:

prop 1 prop 2

0.8135048 0.8166160

The p-value is 0.8942 which is greater than the alpha value of 0.05, so we fail to reject the null hypothesis. This means there no difference between men and women in Canada who believe we should not ban abortion.

* 1. **If you were to construct a 95% confidence interval for the true difference in proportions, would you expect “0” to be in the interval? Explain.**

Yes, I would expect 0 to be included in the 95% confidence interval. This is because we agreed with the null hypothesis, which tells us that there is no difference in proportion between men and women who oppose or support banning abortion. If the null hypothesis of a two-sided test is supported, then it tells us the difference in the value of 2 variables on average is 0. The 2-prop Z test gave us a 95% confidence interval automatically of -0.036 to 0.030. This means we are 95% confident that the difference in proportion between all Canadian men and women’s belief of supporting abortion is between -0.036 to 0.030, which includes 0.