**AP STATISTICS: Review for Unit 5 Exam ANSWERS**

1. We want to compare the proportion of men who are colorblind to the proportion of women. We take 2 independent random samples and find that there are 315 out of 1150 males that are colorblind and only 235 out of 1000 women.
	1. Find the standard error needed for a **confidence interval** comparing the males to the females.



* 1. Find the pooled sample proportion (pooled $\hat{p}$).



* 1. Find the standard error needed for a **test of significance** comparing the males to the females.



* 1. There are some doctors who believe that colorblindness occurs more frequently in males. Test this hypothesis at the 5% significance level.

*p*1 – proportion of males with colorblindness

*p*2 – proportion of females with colorblindness

Ho: *p*1 = *p*2

HA : *p*1 > *p*2

STATE: CHECK:

1. Random They are stated to be random samples
2. Independent It is safe to assume that the two groups are independent,

1. Success/failure 



1. 10% condition There are more than 11500 men and 10000 women

 All conditions have been met to use the Normal model for a 2-proportion z-test.

**Mechanics**:



P-Value = P(*z* > 2.063) = 0.0196

Conclusion:

We reject Ho b/c the P-Value of 0.0196 is less than alpha = 0.05. We have sufficient evidence that the true proportion of males that are colorblind is greater than for females.

* 1. Interpret the P-value in this context.

There is a 1.96% chance of getting samples with a difference of 3.9% or more between the percent of colorblind males and females, if really there is no difference between the two genders colorblindness.

* 1. What would a Type I error be in this context? What is its probability?

If we stated that the proportion of colorblindness in males was higher than for females when in fact it really isn’t. The probability of a type I error for this test was 5%.

* 1. What would a Type II error be in this context?

If we stated that the proportion of colorblindness in males is not greater than for females, but in reality it really is.

* 1. What would power be in this context?

The probability that we stated the proportion of colorblindness in males is higher than for females when in fact it really is.

* 1. Type II = 1-Power

Type II = 1 – 0.85 – 0.15 = 15%

* 1. Since we rejected our Ho in part (c), create a 95% confidence interval for the difference between the % of males and females who are colorblind.

Conditions met above to use a normal model for a 2-Prop Z Interval.

 

We are 95% confident that the true proportion of males with colorblindness is between 0.2% to 7.6% higher than the proportion of females with colorblindness.

OR

We are 95% confident that the difference between the true % of men and women with colorblindness is between 0.2% and 7.6%.

1. I perform a test of significance and I calculate a P-value of 0.06.
	1. Is this significant at the 5% level?

 No since it is higher than 0.05

* 1. Is this significant at the 1% level?

 No since it is higher than 0.01

* 1. Is this significant at the 10% level?

 Yes since it is lower than 0.10

1. What is the Z\* for a 91% confidence interval?

 z\* = invnorm(0.045, 0, 1) = 1.695

1. I have a 92% confidence interval that is (0.22, 0.26). Which of the following could be the 94% confidence interval?
	1. (0.20, 0.24)
	2. (0.20, 0.28)
	3. (0.23, 0.25)
	4. (0.23, 0.27)
2. Using the same info in #4, what could be the 90% confidence interval?

 c. (0.23, 0.25)

1. I have an interval that is (0.30, 0.39)
	1. What is my sample proportion ($\hat{p}$)?

 $\hat{p}=0.345$

* 1. What is my margin of error?

 MOE = 0.045

* 1. If my sample size is 200, what is my level of confidence?

  Level of Confidence = normalcdf(-1.339, 1.339, 0, 1) = **81.9%**

1. I want to sample HS seniors to see what percent of them plan to attend the senior prom. I want to have a 6% margin of error, and want to be 99% confident. What sample size should I take? Last year’s result was 86%.

 

1. Nationwide, it is estimated that 40% of gas stations have tanks that leak to some extent. A new program in California is designed to lessen the prevalence of these leaks. We want to assess the effectiveness of this program and take a random sample of 45 stations and find that 15 of them have leaks.
	1. Create a 94% confidence interval for the percent of stations that leak.

Conditions:

1. Random Stated as a random sample
2. Success/failure 
3. 10% condition There are more than 450 gas stations

 All conditions have been met to use the Normal model for a 1-proportion z-interval.

 

 We are 94% confident that the true proportion of gas stations with leaks is between 20.1% to 46.6%.

* 1. **Using this interval**, do you think that the percent of stations with leaks has decreased? Why or why not?

 No. 40% falls within the confidence interval we created.

* 1. Explain what 94% confidence means in this **context**.

 About 94% of all random samples of size 45 will produce confidence intervals that contain the true proportion of gas stations with tanks that leak.

1. Many doctors believe that teenagers do not get enough Vitamin C. Previous studies have indicated that up to 42% of teenagers are Vitamin C deficient. PA decides to implement a program to educate students about getting enough Vitamin C, in hopes of decreasing the percent of teenagers who are deficient in the state. After this program is in place for a year, researchers take a sample of 200 total high school students from randomly selected high schools across the state. They then test the students Vitamin C levels and find that only 76 of them are Vitamin C deficient.
	1. Is there sufficient evidence at the 5% significance level that the campaign worked (and the percent of HS students who are deficient has decreased)?

Hypothesis:

H0: *p* = 0.42

HA: *p* < 0.42

Conditions:

1. Random It is not a random sample, we will assume it is a

representative sample

2. success/failure *np* = 84 ≥ 10

 *nq =* 116 ≥ 10

3. 10% condition There are more than 2000 high school students

All conditions have been met to use the Normal model for a one-prop z-test.

Mechanics:

*n* = 200; *p* = 0.42; 



P-Value = P(z <-1.146) = 0.126

Conclusion:

We fail to reject the Ho b/c the P-Value of 0.126 is greater than alpha = 0.05.

There is insufficient evidence to say that the proportion of students with a vitamin deficiency is less than 42%. ***OR*** We have sufficient evidence that the true percent of students with a Vitamin deficiency is still equal to 42%.

* 1. Interpret the P-Value in this context.

 P-Value is the probability that if the true proportion of students with vitamin deficiency is 42%, we would

get a sample proportion this low or lower.

* 1. What would a Type I error be in this context? What is its probability?

 We state that the proportion of students with vitamin C deficiency is less than 42% when in fact it is not.

* 1. What would a Type II error be in this context?

 We state that the proportion of students with vitamin C deficiency is not less than 42% when in fact it is.

* 1. What would power be in this context?

 Power is the probability that we state that the proportion of students with vitamin C deficiency is less than 42% when in fact it is.

* 1. If we decreased our significance level to 3%, what would happen to the power, Type I error, and Type II error?

 Type I would decrease; Type II would increase; Power would decrease

* 1. If we kept the same level of significance of 5% but took a new sample of size 450, what would happen to the Type I Error, Type II Error, and Power?

Type I would stay the same; Type II would decrease; Power would increase