AA9-5 Statistical Significance Name ________________________________ Per______

When a sample is properly taken, and is large enough, but not too large, we can use that sample to make inferences about the larger population.

First, we write two hypotheses (educated guesses) about what we think is happening. In AP Stats, we'll write these in symbols, but for this class, we'll just write in words.

Null hypothesis: This is what history or probability tells us should be happening

Alternative hypothesis: This is what we suspect may have changed

Next, we check the conditions. The conditions we check to create a sampling distribution for proportions are from AA9-4):

1.

2.

3.

We then create a model based on the sampling distribution of what should be happening. The model is:

$\mu =$ $\sigma =$

Next, we calculate the z-score of the samples observed proportion and use the calculator (2^{nd} VARS, normalcdf) to determine the probability of a sample like that occurring if the null hypothesis is true. This is called the P-value. If the P-value is less than .05, we say the sample is statistically significant. That means it was unlikely to occur due to natural differences between samples. (For future reference, .05 isn't always the number we compare to, but we're keeping it simple for this class)

Finally, we write a conclusion. If the P-value is less than .05, we reject the null hypothesis because we have evidence of the alternative and say our result is statistically significant. If the P-value is greater than .05, we don't have evidence to reject the null, and our result is not statistically significant.

Example:

In the 1980's, medical records determined that about 5% of the nation's children were affected by congenital abnormalities. A group concerned that an increase in the number of chemicals in the environment did a study with a sample of 384 randomly selected children across the US. The results showed that 12% of them were affected by congenital abnormalities. Is this evidence that the risk has increased?

Hypotheses:
Null: *The proportion of all US children with abnormalities is 5%*
Alternative: *The proportion of US children with abnormalities has increased (is more than 5%).*
Conditions:
Random: *It states it was a random sample.*
Success/Failure: *np=384(.05)=19.2, nq=384(.95)=364.8, both are greater than 10*

We use the p and q from the hypotheses, because that's what we're basing the model on

10% Condition: *384 is less than 10% of all US children.*
Conditions have been met.
Model and Math:

$\sigma = \sqrt{\frac{.05(.95)}{384}} = .011.$ $z = \frac{.12-.05}{.011} = 6.364$ *normalcdf(6.364,99)=*

Conclusion:
The P-value was nearly zero, so we can reject the null. We have evidence that the proportion of US children with congenital abnormalities has increased.

Practice. This is A/B level for AA9-4.

1. The National Center for Education Statistics monitors many aspects of elementary and secondary education nationwide. Their 1996 numbers of often used as a baseline to assess changes. In 1996, 31% of randomly selected students reported that their mothers had graduated from college. In 2000, responses from 8368 students found that this figure had grown to 32%. Is this evidence of a statistically significant change in education level among mothers?

Hypotheses:

Null:

Alternative:

Conditions:

Random:

Success/Failure:

10% Conditions:

Have conditions been met?

Model and Math:

Conclusion:

2. A company is criticized because only 13 of 43 people in executive-level positions are women. The company explains that although this proportion is lower than it might wish, it's not surprising given that only 40% of all its employees are women. Is the 13 of 43 people significantly lower than 40%?

Hypotheses:

Null:

Alternative:

Conditions:

Random:

Success/Failure:

10% Conditions:

Have conditions been met?

Model and Math:

Conclusion: