Statistical Inference and Hypothesis Testing for Population Proportion

Kory Illenye

6/11/2018

Statistical Inference of Proportions

Statistical inference is the process of making generalizations about a population from sampled information. In order to make use of proportions our test statistic must be binomial in nature. Meaning the test statistic can be described as successes and failures.

Key Conditions:

1. Random: Data must come from a random sample, experiment or simulation
2. Normal: The sampling distribution of $\hat{p}$ needs to be relatively normal.
    - n represents sample size and p represents the population's proportion of successes.
    - $np \geq 10$
    - $n(1-p) \geq 10$
3. Independence: If the sample size is less than 10% of population size then we can treat each observation as independent since removing each observation doesn't significantly change the population as it is sampled.

What can we do if the Key Conditions are met?

Once key conditions are made we can begin to make generalizations of a population based on the random sample. Such as create confidence intervals and conduct hypothesis testing. These are just a couple of tools we use to generalize a population.

Hypothesis testing of population proportion

Can you think of some places where this might be used?

- How about manufacturing?
- What about politically?
- Education?
- Research?

There are many uses for Hypothesis Testing

5 Step Process for Testing a Hypothesis of Population Proportion

Step 1: State the null hypothesis (H_0) and the alternate hypothesis (H_a).

Step 2: Identify the test statistic

Step 3: Calculate the Rejection Region

Step 4: State the statistical conclusion

Step 5: State the English conclusion

5 Step Process Example

A report claimed that 20% of all college graduates find a job in their chosen field of study within one year of graduation. A random sample of 500 graduates found that 90 obtained work in there field within one year of graduation. On a significance level of 0.05 ($\alpha = 0.05$) is there statistical evidence to refute the claim of this report?

- Step 1:
  - $H_0 = 0.20$
  - $H_a \neq 0.20$
- Step 2: $\hat{p} = \frac{90}{500} = 0.18$
- Step 3: Compare p-value to $\alpha = 0.05$. The rejection region is when p-value is less than 0.05.
  - $\text{sd}(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0.2*0.8}{500}} = 0.018$
  - $z = \frac{\hat{p}-p}{sd(\hat{p})} = \frac{0.18-.2}{0.01789} = -1.12$
  - p-value $= 2^*\text{P}(z < -1.12) = 0.2628$

5 Step Process Example (Continued)

- Step 4: Since the p-value of 0.2628 is greater than the α value of 0.05 we fail to reject H_0
- Step 5: There is not significant enough evidence to suggest the proportion of college graduates finding work in there chosen field is something other than 20%.

Questions

Questions?

References

De Veaux; Velleman; and Bock; *Stats: Data and Models*, Pearson Education, 2016.

Ugarte, Maria D.; Militino, Ana F.; Arnholt, Alan T.; *Probability and Statistics with R, Second Edition*, CRC Press, 2016.

Networked Knowledge - Law Report, *R. v. Sally Clark [2003] EWCA Crim 1020 [Part Two]*, http://netk.net.au/UK/SallyClark2.asp