Chapter 12: Notes

Evan Hunt

October 2, 2020

Section 12.1

Population Proportion and Sample Proportion

Consider a population in which each member either has or does not have a specified attribute. Then we use the following notation and terminology.

Population proportion, p: The proportion (percentage) of the entire population that has the specified attribute.

Sample proportion, $\hat{p}$: The proportion (percentage) of a sample from the population that has the specified attribute.

Sample Population

A sample proportion, $\hat{p}$, is computed by using the formula

$$\hat{p} = \frac{x}{n},$$

where x denotes the number of members in the sample that have the specified attribute and, as usual, n denotes the sample size.

Note:

For convenience, we sometimes refer to x (the number of members in the sample that have the specified attribute) as the number of successes and to $n - x$ (the number of member in the sample that do not have the specified attribute) as the number of failures. In this context, the words success and failure may not have their ordinary meanings.

The sampling Distribution of the Sample Proportion‘

For samples of size n,

- the mean of $\hat{p}$ equals the population proportion: $\mu_{\hat{p}} = p$ (i.e., the sample proportion is an unbiased estimator of the population proportion);
- the standard deviation of $\hat{p}$ equals the square root of the product of the population proportion and one mine the population proportion divided by sample size: $\sigma_{\hat{p}=\sqrt{p(1-p)/n}}$; and
- $\hat{p}$ is approximately normally distributed for large n.

If n is large, the possible samples of size n have approximately a normal distribution with mean p and standard deviation $\sqrt{p(1-p)/n}$.

0.1 One-Proportion z-Interval Procedure

Purpose To find confidence interval for a population proportion, p

Assumptions
1. Simple random sample
2. The number of successes, x, and the number of failures, $n - x$, are both 5 or greater.

Step 1 For a confidence level of $1 - \alpha$, use a z-table to find $z_{\alpha/2}$.

Step 2 The confidence interval for p is from

$$[\ \hat{p} - z_{\alpha/2} * \sqrt{\hat{p}(1-\hat{p})/n},\ \hat{p} + z_{\alpha/2} * \sqrt{\hat{p}(1-\hat{p})/n}\ ],$$

where $z_{\alpha/2}$ is found in Step 1, n is the sample size, and $\hat{p} = x/n$ is the sample proportion.

Step 3 Interpret the confidence interval.

Margin of Error for the Estimate of p

The margin of error for the estimate of p is

$$E = z_{\alpha/2} * \sqrt{\hat{p}(1-\hat{p})/n}.$$

The margin of error for the estimate of a population proportion indicates the accuracy with which a sample proportion estimates the unknown population proportion at the specified confidence level.

Sample Size for Estimating p

- A $(1-\alpha)$-level confidence interval for population proportion that has a margin of error of at most E can be obtained by choosing

$$n = \frac{(z_{\alpha/2}/E)^2}{4}$$

rounded up to the nearest whole number.

- If you can make an educated guess, $\hat{p}_g$ (g for guess), for the observed value of $\hat{p}$, then you should instead choose

$$n = \hat{p}_g(1-\hat{p}_g)(z_{\alpha/2}/E)^2$$

rounded up to the nearest whole number.

- If you have in mind a likely range for the observed value of $\hat{p}$, then you should apply the preceding formula with you educated guess for observed value of $\hat{p}$ being the value in range closest to 0.5.

Section 12.2

One-Proportion z-Test

Purpose To perform a hypothesis test for a population proportion, p.

Assumption 1. Simple random sample

2. Both np_0 and $n(1-p0)$ are 5 or greater

Step 1 The null hypothesis is H_0: $p = p_0$, and the alternative hypothesis is

$$\begin{array}{ccccc} H_a : p! = p_0 & or & H_a : p < p_0 & or & H_a : p > p_0 \\ (Two\ Tailed) & & (Left\ Tailed) & & (Right\ Tailed) \end{array}$$

Step 2 Decide the significance level, α.

Step 3 Compute the value of the test statistic

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$$

and denote that value z_0.

Critical-Value Approach

Step 4 The critical value(s) are

$$\begin{array}{ccccc} \pm z_{\alpha/2} & or & -z_{\alpha} & or & z_{\alpha} \\ (Two\ tailed) & & (Left\ tailed) & & (Right\ tailed) \end{array}$$

Step 5 If the value of the test statistic falls in the rejection region, reject H_0: otherwise, do not reject H_0.

Step 6 Interpret the results of the hypothesis test.

P-Value Approach

Step 4 Use a z-table to obtain the P-value.

Step 5 If $P \leq \alpha$, reject H_0: otherwise, do not reject H_0.

Step 6 Interpret the results of the hypothesis test.