Stat 131 (Mathematical Statistics III)

Lesson 3.2 Z Test for One Population Proportion

Learning Outcomes

At the end of the lesson, students should be able to

1. derive the test statistic for testing hypothesis on one population proportion; and
2. conduct the z test for testing hypothesis on one population proportion.

Introduction

For categorical data, particularly binary data, the sample proportion ($\widehat{p}$) is often computed to summarize the data. Suppose that $Y_1, Y_2, \cdots, Y_n$ is distributed as Bernoulli(p). We know that $T = \sum_{i=1}^{n} Y_i$ has a binomial distribution with parameters n and p. Let the sample proportion be defined as $\widehat{p} = \frac{T}{n}$. An equivalent version of the Central Limit Theorem is the *Rule of Sample Proportions* which states that if $np \geq 5$ and $n(1-p) \geq 5$, then the sample proportions have an approximate normal distribution with mean $\mu = np$ and variance $\sigma^2 = np(1-p)$ (Ramachandan and Tsokos, 2009).

The Z test statistic

A hypothesis test for a proportion is used when you are comparing one group to a known or hypothesized population proportion value. In other words, you have one sample with one categorical variable. Suppose that our interest lies in testing $H_0 : p = p_0$ versus $H_1 : p \neq p_0$ (or any other H_1). In this situation, we identify

$$\begin{aligned}
\theta &= p \\
\widehat{\theta} &= \widehat{p} \\
\sigma_{\widehat{\theta}} &= \sqrt{\frac{p(1-p)}{n}} \\
\widehat{\sigma}_{\widehat{\theta}} &= \sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}
\end{aligned}$$

To perform the test for $H_0 : p = p_0$ versus $H_1 : p \neq p_0$, there are two candidate test statistics. First, we have

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

This test statistic arises from the theory we have just developed and assumes that we know p in advance. But, in practice p is unknown, thus we have the second test statistic given below.

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

For theoretical reasons, Z_W is called a Wald statistic and Z_S is called a score statistic. Both have large sample $N(0,1)$ distributions when $H_0 : p = p_0$ is true. The score statistic Z_S is known to have better properties in small (i.e., finite) samples; i.e., it possesses a true significance level which is often closer to the nominal level α. The Wald statistic is often liberal, possessing a true significance level larger than the nominal level. Thus, in most applications we shall use the score statistic Z_S.

To proceed with either test statistic we need to verify if the normal approximation holds by checking if $np \geq 5$ and $n(1-p) \geq 5$.

Example 3.2.1

A machine is considered to be unsatisfactory if it produces more than 8% defectives. It is suspected that the machine is unsatisfactory. A random sample of 120 items produced by the machine contains 14 defectives. Does the sample evidence support the claim that the machine is unsatisfactory? Use $\alpha = 0.01$.

SOLUTION

Let p be the true proportion of defective items. The appropriate hypotheses are

$$H_0 : p = 0.08$$
$$H_1 : p > 0.08$$

Test statistic:

Let Y be the number of observed defectives. This follows a binomial distribution. However, because $np_0 = 120 \times 0.08 = 9.6 > 5$ and $n(1 - p_0) = 120 \times (1 - 0.08) = 110.4 > 5$, we can use a normal approximation to the binomial distribution. Based on the sample we have

$\hat{p} = \frac{14}{120} \approx 0.1167$. Using the score statistic we have

$$\begin{aligned} Z_S &= \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \\ &= \frac{0.1167 - 0.08}{\sqrt{\frac{0.08(1-0.08)}{120}}} \\ &\approx 1.48 \end{aligned}$$

Rejection Region

Based on the form of the alternative hypothesis, the rejection region is given by $RR = \{z : z \geq z_{0.01} = 2.33\}$, while $p - value = P(Z \geq 1.48) = 0.0694$.

Decision on H_0

Since $Z = 1.48$ does not fall in the rejection region, we fail to reject the H_0. Similarly, since $p - value > \alpha = 0.01$, we fail to reject the H_0.

Conclusion

Therefore, at $\alpha = 0.01$, The data is not sufficient to conclude that the machine is unsatisfactory.

Example 3.2.2

An e-commerce research company claims that 60% or more graduate students have bought merchandise online. A consumer group is suspicious of the claim and thinks that the proportion is lower than 60%. A random sample of 80 graduate students shows that only 36 students have ever done so. Is there enough evidence to show that the true proportion is lower than 60%?

SOLUTION: Left as a classroom exercise!

Learning Tasks

Instruction: Answer the following as indicated. Follow the steps in hypothesis testing.

1. Let p be the proportion of drivers who use a seat belt in a state that does not have a mandatory seat belt law. It was claimed that $p = 0.14$. An advertising campaign was conducted to increase this proportion. Two months after the campaign, 104 out of a random sample of 590 drivers were wearing their seat belts. Was the campaign successful?

2. The Admissions Office of the Visayas State University (VSU) claims that more than 30% of VSU students are graduates of STEM track in their senior high school. A student researcher conducted a survey using a random sample of 320 students and finds that 124 of them are STEM graduates. Does the data support the claim of the Admissions Office? Use $\alpha = 0.05$.