Nov 21, 2022

# Theoretical Hypothesis Testing

## Week 7: Tests on proportions

# Introduction

In this lecture,

We focus on the **tests which are used for the parameter of a binomial distribution**:

- test on the population proportion in the one-sample case
- tests for the difference between of two proportions using the pooled as well as the unpooled variances.
- tests for the equality of proportions for the multi-sample case.

# One sample test

# 1. Binomial test

Binomial test is used to test that if a proportion of population $p$ differs from a predefined value $p_0$.

<u>Assumptions:</u>

- Data are randomly sampled from a large population with two possible outcome.
- The parameter of interest $p$ is given by the proportion of successes in the population.
- The number of successes in a random sample of size $n$ follows a binomial distribution $B(n, p)$

# 1. Binomial test

Hypotheses:

- $H_0 : p = p_0$ vs. $H_A : p \neq p_0$
- $H_0 : p \leq p_0$ vs. $H_A : p > p_0$
- $H_0 : p \geq p_0$ vs. $H_A : p < p_0$

# 1. Binomial test

Test statistics:

$$Z = \frac{\sum_{i=1}^{n} X_i - np_0}{\sqrt{np_0(1-p_0)}}$$

- This test is a large sample test. If the sample size is large, $np(1-p) > 9$, the test statistics $Z$ is approximately a standard normal distribution. For small samples an exact test with $Y = \sum_{i=1}^{n} X_i$ as test statistic and critical regions based on the binomial distribution is used.
- If the sample size is small $np(1-p) < 9$, use prop.test( )* instead of binom.test( ) in R

*Fisher's exact test

# Application

To test the hypothesis that the proportion of defective (0: non-defective, 1:defective) workpieces of a machine equals 50%. The available dataset contains 40 observations.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 |
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 |

# Application

Hypotheses:

$H_0 :$

$H_A :$

Decision:

Interpretation:

# Two sample test

# 2. Two sample tests

In this section, **we deal with the question is that if proportions of two independent populations differ from each other**. There are two tests for this problem:

- In the first case, the standard deviations of both distributions may differ from each other.
- In the second one, the standard deviations of both distributions are equal, or pooled to obtain better estimate of the standard deviation if they are unknown.

# 2.1. Z-test with unpooled variance

**This test is used to test if two population proportions $p_1$ and $p_2$ differ from a specific value $d_0$.**

<u>Assumptions:</u>

- Data are randomly sampled with two possible outcomes.
- The number of success for each samples follow binomial distributions $B(n_1, p_1)$ and $B(n_1, p_2)$.

# 2.1. Z-test with unpooled variance

Hypotheses:

$H_0: p_1 = p_2$ vs. $H_A: p_1 \neq p_2$ or $H_0: p_1 - p_2 = p_0$ vs. $H_A: p_1 - p_2 \neq p_0$

$H_0: p_1 \geq p_2$ vs. $H_A: p_1 < p_2$ or $H_0: p_1 - p_2 \geq p_0$ vs. $H_A: p_1 - p_2 < p_0$

$H_0: p_1 \leq p_2$ vs. $H_A: p_1 > p_2$ or $H_0: p_1 - p_2 \leq p_0$ vs. $H_A: p_1 - p_2 > p_0$

# 2.1. Z-test with unpooled variance

Test statistics:

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - d_0}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

This is a large sample test. If the sample size is large enough the test statistics $Z$ follows a standard normal distribution. As a rule of thumb, $n_1 p_1$, $n_1(1 - p_1)$, $n_2 p_2$, and $n_2(1 - p_2)$ should all be higher than 5.

# 2.2. Z-test with pooled variance

**This test is used to test if two population proportions $p_1$ and $p_2$ differ from each other.**

<u>Assumptions:</u>

- Data are randomly sampled with two possible outcomes.
- The number of success for each samples follow binomial distributions $B(n_1, p_1)$ and $B(n_1, p_2)$.

<u>Test statistics:</u>

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{where } \hat{p} = \frac{\hat{p}_1 n_1 + \hat{p}_2 n_2}{n_1 + n_2}$$

This is a large sample test. If the sample size is large enough the test statistics $Z$ follows a standard normal distribution. As a rule of thumb, $n_1 p_1$, $n_1(1-p_1)$, $n_2 p_2$, and $n_2(1-p_2)$ should all be higher than 5.

# Application

To test the hypothesis that the proportion of defective (0: non-defective, 1:defective) workpieces of machines in company A is equal to the proportion of the defective machines in company B. The available dataset contains 20 observations for each company.

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| company | A | | | | | | | | | | | | | | | | | | | |
| malfunction | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 |
| company | B | | | | | | | | | | | | | | | | | | | |
| malfunction | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

# Application

Hypotheses:

$H_0$ :

$H_A$ :

Decision:

Interpretation:

# A/B testing

# 3. A/B Testing

A/B testing is used to choose the best alternative from two possible alternatives A and B.

A/B testing is a framework for you to test different ideas for how to improve upon an existing design

# 3. A/B Testing

For example, consider that there are two alternative designs A and B of a website. Visitors are randomly served with one of these designs. The data about this process can be collected by web analytics. Using this data, it can be possible to determine which design is better by applying statistical tests.

# 3. A/B Testing

To measure the efficiency of website design, there are some popular metrics:

- **Click-through-rate:** The ratio of clicking on advertisement by a customer after it is seen.
- **Conversion rate:** The ratio of purchasing any item by a customer after the advertisement is seen.
- **Bounce rate:** The ratio of re-visit the website by a customer after visiting.

# Application

When we test two variants we get the observed that the variant V1 which has 125 Clicks and 800 Impressions, and the variant V2 which has 85 Clicks and 500 Impressions. Test that if the actual CTR (Click Through Rate) of the V1 and V2 is whether equal.

Hypotheses:

$H_0$ :

$H_A$ :

Decision:

Interpretation:

# k-sample test

# 4. k-sample Binomial test

Tests if $k$ population proportions differ from each other.

<u>Assumptions:</u>

- Data are randomly sampled with two possible outcomes.
- The number of successes for each samples follow binomial distribution $B(n_i, p_i)$.

# 4. k-sample Binomial test

Hypotheses:

$$H_0 : p_1 = p_2 = \ldots = p_k$$

$H_A : p_i \neq p_j$ for at least one $i \neq j$

# Application

The proportions of male carp in three ponds are tested for equality. The observed relative frequency of male carp in pond one is 10/19, in pond two 12/20, and in pond three 14/21.

<u>Hypotheses:</u>

$H_0$ :

$H_A$ :