Hypotheses testing

Rosario Barone

Tor Vergata University of Rome

Statistical tools for decision making

Undergraduate Degree in Global Governance

A.Y. 2024/2025

Hypotheses Testing

- Fundamental statistical technique used to make decisions based on sample data.
- Allows us to evaluate competing claims or hypotheses about a population parameter.

Hypotheses Testing: step 1 and 2

1) State the null hypothesis (H_0) and the alternative hypothesis (H_1).

Null Hypothesis (H_0)

The null hypothesis is the default claim or assumption that we want to test. It may represent no effect, no difference, or no relationship between variables.

Alternative Hypothesis (H_1 or H_a)

The alternative hypothesis is the claim we are interested in establishing if the null hypothesis is rejected. It represents an effect, difference, or relationship between variables.

2) Choose a significance level (α) to determine the threshold for rejecting the null hypothesis.

Hypotheses Testing: step 3 and 4

3) Collect sample data and calculate the test statistic.

Test Statistic

The test statistic is a calculated value that is used to assess the evidence against the null hypothesis. It quantifies the discrepancy between the observed data and what is expected under the null hypothesis.

4) Compare the p-value to the significance level and make a decision:

- If p-value $\leq \alpha$, reject H_0 in favor of H_1.
- If p-value $> \alpha$, fail to reject H_0.

p-value

The p-value is the probability of observing a test statistic as extreme as, or more extreme than, the one calculated from the sample data, assuming that the null hypothesis is true.

Steps in Hypothesis Testing

1. State the null hypothesis (H_0) and the alternative hypothesis (H_1).
2. Choose a significance level (α) to determine the threshold for rejecting the null hypothesis.
3. Collect sample data and calculate the test statistic.
4. Determine the p-value associated with the test statistic.
5. Compare the p-value to the significance level and make a decision:
   - If p-value $\leq \alpha$, reject H_0 in favor of H_1.
   - If p-value $> \alpha$, fail to reject H_0.
6. Interpret the results in the context of the problem.

Formally

Suppose we have some observed data X and want to test a hypothesis. The data may provide evidence against the hypothesis or support its validity.
To perform this test, we select a test statistic $T = t(X)$, which is a function of the observed data. Large values of T cast doubt on the hypothesis, suggesting that the observed data may not align with its assumptions.

The hypothesis being tested, known as the **Null Hypothesis (H_0)**, imposes restrictions on the distribution of X. Using this, we calculate a **p-value**, which is the probability, under the null hypothesis, of observing a test statistic T as extreme or more extreme than the observed value t_{obs}:

$$p_{obs} = P_0(T > t_{obs})$$

Formally

The *null distribution* is the distribution of a test statistic computed under the assumption that the **Null Hypothesis (H_0)** is true.

If there is a specific idea about what situation holds when the null hypothesis is false, this leads to a clearly specified **Alternative Hypothesis (H_1)**. The alternative hypothesis describes how the data are expected to deviate from H_0.

To test these hypotheses, we choose a test statistic $T = t(X)$ that has a high probability of detecting departures from H_0 in the direction of H_1.

Significance level and Power

Significance level

The significance level of a binary hypothesis test, denoted by α , is the probability of the study rejecting the null hypothesis, given that the null hypothesis is true.

Power of the test

The power of a binary hypothesis test $(1 - \beta)$, is the probability that the test correctly rejects the null hypothesis (H_0) when a specific alternative hypothesis (H_1) is true.

Significance level and Power

- $\beta =$ probability of a **Type II error**, known as a "false negative"
- $1-\beta =$ probability of a "true positive", i.e., correctly rejecting the null hypothesis. $1-\beta$ is also known as the power of the test.
- $\alpha =$ probability of a **Type I error**, known as a "false positive"
- $1-\alpha=$ probability of a "true negative", i.e., correctly not rejecting the null hypothesis

	H_0 **is True**	H_0 **is False**
Test Rejects H_0	$\boldsymbol{\alpha}$	$1-\boldsymbol{\beta}$
Test Doesn't Reject H_0	$1-\boldsymbol{\alpha}$	$\boldsymbol{\beta}$

Formulation of Hypotheses

- Hypotheses should be formulated based on prior knowledge, research questions, or practical considerations.
- The null hypothesis typically represents a position of no effect, no difference, or no relationship.
- The alternative hypothesis represents the claim or effect that we want to establish.
- Hypotheses can be one-sided (directional) or two-sided (non-directional), depending on the research question and the nature of the claim.
- Test statistics are used to measure the strength of evidence against the null hypothesis.
- The choice of test statistic depends on the type of data and the hypothesis being tested: common test statistics include the t-statistic, z-statistic, chi-square statistic, and F-statistic.

Interpretation of Results

- When the null hypothesis is rejected, it suggests evidence in favor of the alternative hypothesis.
- The rejection of the null hypothesis does not prove the alternative hypothesis; it only indicates that the data provide strong evidence against the null hypothesis.
- The interpretation of results should consider the context of the problem, the magnitude of the effect, and the implications of the findings.

Hypothesis Test for Population Mean: Formulating Hypotheses

- For the hypothesis test for population mean,

$$
\begin{aligned}
H_0 &: \mu = \mu_0 \\
H_1 &: \mu \neq \mu_0 \quad \text{(two-tailed test)}
\end{aligned}
$$

or

$$
\begin{aligned}
H_0 &: \mu \leq \mu_0 \\
H_1 &: \mu > \mu_0 \quad \text{(one-tailed test)}
\end{aligned}
$$

Hypothesis Test for Population Mean (σ known): Test Statistic and Critical Region

- The z-statistic is calculated as:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

  where $\bar{x}$ is the sample mean, μ_0 is the hypothesized population mean, σ is the population standard deviation, and n is the sample size.
- Null distribution: standard Normal $N(0, 1)$.
- The critical region is determined based on the chosen significance level (α) and standard Normal distribution.
- For a two-tailed test, the critical region is split evenly in both tails, each containing an area of $\alpha/2$; for a one-tailed test, the critical region is located entirely in one tail, containing an area of α.

Hypothesis Test for Population Mean (σ known): Test Statistic and Critical Region

- The z-statistic is calculated as:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

  where $\bar{x}$ is the sample mean, μ_0 is the hypothesized population mean, σ is the population standard deviation, and n is the sample size.
- Null distribution: standard Normal $N(0, 1)$.
- The critical region is determined based on the chosen significance level (α) and standard Normal distribution.
- For a two-tailed test, the critical region is split evenly in both tails, each containing an area of $\alpha/2$; for a one-tailed test, the critical region is located entirely in one tail, containing an area of α.

Hypothesis Test for Population Mean (σ unknown): Test Statistic and Critical Region

- The t-statistic is calculated as:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

  where s is the sample standard deviation.
- Null distribution: Student t with $n - 1$ degrees of freedom.
- The critical region is determined based on the chosen significance level (α) and the Student t distribution.
- For a two-tailed test, the critical region is split evenly in both tails, each containing an area of $\alpha/2$; for a one-tailed test, the critical region is located entirely in one tail, containing an area of α.

Hypothesis Test for Population Mean (σ unknown): Test Statistic and Critical Region

- The t-statistic is calculated as:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

  where s is the sample standard deviation.
- Null distribution: Student t with $n - 1$ degrees of freedom.
- The critical region is determined based on the chosen significance level (α) and the Student t distribution.
- For a two-tailed test, the critical region is split evenly in both tails, each containing an area of $\alpha/2$; for a one-tailed test, the critical region is located entirely in one tail, containing an area of α.

Hypothesis Test for Population Proportion: Formulating Hypotheses

For the hypothesis test for population proportion, the hypotheses can be stated as follows:

$$
\begin{aligned}
H_0 &: p = p_0 \\
H_1 &: p \neq p_0 \quad \text{(two-tailed test)}
\end{aligned}
$$

or

$$
\begin{aligned}
H_0 &: p \leq p_0 \\
H_1 &: p > p_0 \quad \text{(one-tailed test)}
\end{aligned}
$$

Hypothesis Test for Population Proportion: Test Statistic and Critical Region

- Null distribution: Standard Normal N(0,1).
- The z-statistic is calculated as:

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

where $\hat{p}$ is the sample proportion, p_0 is the hypothesized population proportion, and n is the sample size.

Hypothesis Test for Difference between Means: Formulating Hypotheses

For the hypothesis test for difference between means, the hypotheses can be stated as follows:

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_1 : \mu_1 - \mu_2 \neq 0 \quad \text{(two-tailed test)}$$

or

$$H_0 : \mu_1 - \mu_2 \leq 0$$
$$H_1 : \mu_1 - \mu_2 > 0 \quad \text{(one-tailed test)}$$

Hypothesis Test for Difference between Means: Test Statistic and Critical Region

- The t-statistic is calculated as:

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

  where $\bar{x}_1$ and $\bar{x}_2$ are the sample means, μ_1 and μ_2 are the population means, s_1 and s_2 are the sample standard deviations, n_1 and n_2 are the sample sizes.
- Null distribution: Student t with $n_1 + n_2 - 2$ degrees of freedom.

Example: Hypothesis Test for Difference between Means

```
# Generate two samples from normal distributions
n1 <- 50  # Sample size for population 1
n2 <- 50  # Sample size for population 2
mean1 <- 5  # Mean of population 1
mean2 <- 7  # Mean of population 2
sd1 <- 2  # Standard deviation of population 1
sd2 <- 3  # Standard deviation of population 2

sample1 <- rnorm(n1, mean = mean1, sd = sd1)
sample2 <- rnorm(n2, mean = mean2, sd = sd2)
```

```
# Perform the hypothesis test
alpha <- 0.05  # Significance level
xbar1 <- mean(sample1)
xbar2 <- mean(sample2)
s1 <- sd(sample1)
s2 <- sd(sample2)

# Calculate the test statistic
t_stat <-
(xbar1 - xbar2) / sqrt((s1^2/n1) + (s2^2/n2))
> t_stat
[1] -5.098568

# Calculate the degrees of freedom
df <- n1 + n2 - 2
df
[1] 98
```

```
# Calculate the critical value
(for a two-tailed test)
critical_value <- qt(1 - alpha/2, df)
> critical_value
[1] 1.984467

# Compare the test statistic with the critical value
if (abs(t_stat) > critical_value) {
  cat("Reject Null Hypothesis")
} else {
  cat("Fail to Reject Null Hypothesis")
}
"Reject Null Hypothesis"
```

Hypothesis Test for Difference between Proportions: Formulating Hypotheses

- The null hypothesis (H_0) represents the claim that we want to test. It is typically a statement of no effect or no difference.
- The alternative hypothesis (H_1 or H_a) represents the claim we suspect to be true if there is evidence against the null hypothesis.
- For the hypothesis test for difference between proportions, the hypotheses can be stated as follows:

$$\begin{aligned} H_0 &: p_1 - p_2 = 0 \\ H_1 &: p_1 - p_2 \neq 0 \quad \text{(two-tailed test)} \end{aligned}$$

or

$$\begin{aligned} H_0 &: p_1 - p_2 \leq 0 \\ H_1 &: p_1 - p_2 > 0 \quad \text{(one-tailed test)} \end{aligned}$$

Hypothesis Test for Difference between Proportions: Test Statistic and Critical Region

- The z-statistic is calculated as:

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

  where $\hat{p}_1$ and $\hat{p}_2$ are the sample proportions, $\hat{p}$ is the pooled proportion, n_1 and n_2 are the sample sizes.

- Null distribution: **Standard Normal**.

Chi-Square Test for Independence of categorical variables: Formulating Hypotheses

- The null hypothesis (H_0) represents the claim that the two categorical variables are independent.
- The alternative hypothesis (H_1) represents the claim that the two categorical variables are dependent or associated in some way.
- For the Chi-Square Test for Independence, the hypotheses can be stated as follows:

$$H_0 : \text{The two variables are independent}$$
$$H_1 : \text{The two variables are dependent}$$

Chi-Square Test for Independence: Test Statistic and Critical Region

- The chi-square statistic is calculated as:

$$\chi^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

  where O_{ij} represents the observed frequency and E_{ij} represents the expected frequency for each cell in the contingency table.
- Null distribution: Chi-Square (χ^2).

Chi-Square Test for Independence: Test Statistic and Critical Region

- The critical region is determined based on the chosen significance level (α) and the degrees of freedom associated with the test.
- The degrees of freedom for the Chi-Square Test for Independence is calculated as:

$$\text{df} = (r - 1) \times (c - 1)$$

where r is the number of rows and c is the number of columns in the contingency table.

Example: Chi-Square Test for Independence

```
# Create a contingency table
table <- matrix(c(15, 20, 10, 30, 25, 35),
nrow = 2, byrow = TRUE)
rownames(table) <- c("Group 1", "Group 2")
colnames(table) <- c("Category A", "Category B",
"Category C")

> print(table)
        Category A Category B Category C
Variable 1         15         20         10
Variable 2         30         25         35
```

```
# Calculate the row and column sums
row_sums <- rowSums(table)
col_sums <- colSums(table)

# Calculate the expected frequencies
under independence assumption
expected <- outer(row_sums, col_sums) / sum(table)
> print(expected)
        Category A Category B Category C
Variable 1        15         15         15
Variable 2        30         30         30

# Calculate the chi-squared test statistic
chi_sq_stat <- sum((table - expected)^2 / expected)
```

```
# Calculate the degrees of freedom
df <- (nrow(table) - 1) * (ncol(table) - 1)

# Calculate the p-value
p_value <- 1 - pchisq(chi_sq_stat, df)

> cat("Chi-Squared Statistic:", chi_sq_stat, "\n")
Chi-Squared Statistic: 5
> cat("Degrees of Freedom:", df, "\n")
Degrees of Freedom: 2
> cat("p-value:", p_value, "\n")
p-value: 0.082085
> cat("Reject Null Hypothesis:",
p_value < 0.05, "\n")
Reject Null Hypothesis: FALSE
```