# APMA 3120: Guided Notes on Hypothesis Testing and P-Values

Midterm Study Guide

October 2023

## 1 Introduction to Hypothesis Testing

Hypothesis testing is a statistical method used to make inferences or draw conclusions about a population based on sample data. The core idea is to test a claim (hypothesis) about a population parameter (e.g., mean, proportion) by analyzing a sample.

### 1.1 Types of Hypotheses

- **Null Hypothesis ($H_0$)**: The statement we wish to test. It usually represents a status quo or no effect. For example, $H_0 : \mu = \mu_0$.
- **Alternative Hypothesis ($H_a$)**: The statement we want to support. It represents an effect or a difference. Examples include:
  - **Two-sided Test**: $H_a : \mu \neq \mu_0$
  - **One-sided Test**: $H_a : \mu > \mu_0$ or $H_a : \mu < \mu_0$

### 1.2 Test Statistic

A test statistic is calculated from the sample data and is used to decide whether to reject the null hypothesis. The form of the test statistic depends on the type of data and the sample size. Common test statistics include:

- **Z-Test**: Used when the population variance is known and the sample size is large ($n \geq 30$).
- **T-Test**: Used when the population variance is unknown and the sample size is small ($n < 30$). The test statistic is given by:

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \tag{1}$$

  where $\bar{X}$ is the sample mean, $\mu_0$ is the hypothesized population mean, $s$ is the sample standard deviation, and $n$ is the sample size.

### 1.3 Significance Level ($\alpha$)

The significance level, denoted as $\alpha$, is the threshold for rejecting the null hypothesis. Common values are $\alpha = 0.05$ or $\alpha = 0.01$. It represents the probability of making a Type I error (rejecting $H_0$ when it is true).

### 1.4 Decision Rule

The decision to reject or fail to reject $H_0$ depends on comparing the test statistic to a critical value or using a p-value approach:

- **Critical Value Approach**: Determine the critical value from statistical tables (e.g., Z or T distribution). Reject $H_0$ if the test statistic falls in the rejection region.

- **P-Value Approach**: The p-value is the probability of obtaining a test statistic at least as extreme as the one calculated, assuming $H_0$ is true.
  - If p-value $< \alpha$, reject $H_0$.
  - If p-value $\geq \alpha$, fail to reject $H_0$.

# 2 P-Value Generation

The p-value is a critical component in hypothesis testing, providing a measure of the strength of evidence against the null hypothesis.

## 2.1 Calculating P-Values for Z and T Tests

- **Z-Test**: Use the standard normal distribution table to find the area under the curve beyond the calculated Z value.
- **T-Test**: Use the t-distribution table, considering the degrees of freedom ($df = n - 1$ for one sample). The p-value is found by locating the test statistic in the table and identifying the corresponding probability.

## 2.2 Example: Finding a P-Value with T-Distribution

Assume a sample of size $n = 10$ with a sample mean of $\bar{X} = 21.9$ and a sample standard deviation of $s = 4.12$. We want to test if the true mean is $\mu_0 = 18$ at $\alpha = 0.05$.

1. **Step 1: State the Hypotheses**
   - $H_0 : \mu = 18$
   - $H_a : \mu > 18$
2. **Step 2: Calculate the Test Statistic**

$$t = \frac{21.9 - 18}{4.12/\sqrt{10}} = 3.00 \tag{2}$$

3. **Step 3: Determine the P-Value**
   - With $df = 9$, use the t-table to find the probability corresponding to $t = 3.00$.
   - The p-value is approximately 0.007, which is less than $\alpha = 0.05$. Thus, we reject $H_0$.
4. **Step 4: Conclusion**
   - Since the p-value is less than the significance level $\alpha = 0.05$, we reject the null hypothesis.
   - There is sufficient evidence to conclude that the true mean is greater than 18.

## 2.3 Example Problem 2: Two-Sided T-Test

Assume a sample of size $n = 15$ with a sample mean of $\bar{X} = 50$ and a sample standard deviation of $s = 5$. We want to test if the true mean is $\mu_0 = 52$ at $\alpha = 0.05$.

1. **Step 1: State the Hypotheses**
   - $H_0 : \mu = 52$
   - $H_a : \mu \neq 52$

2. **Step 2: Calculate the Test Statistic**

$$t = \frac{50 - 52}{5/\sqrt{15}} = -1.55 \tag{3}$$

3. **Step 3: Determine the P-Value**
   - With $df = 14$, use the t-table to find the probability corresponding to $t = -1.55$.
   - The two-tailed p-value is approximately 0.144.

4. **Step 4: Conclusion**
   - Since the p-value is greater than the significance level $\alpha = 0.05$, we fail to reject the null hypothesis.
   - There is not enough evidence to conclude that the true mean is different from 52.

# 3 Errors in Hypothesis Testing

- **Type I Error ($\alpha$)**: Rejecting $H_0$ when it is true. The probability of making this error is the significance level $\alpha$.
- **Type II Error ($\beta$)**: Failing to reject $H_0$ when it is false. The power of a test is $1-\beta$, which represents the probability of correctly rejecting a false $H_0$.

## 3.1 Example: Understanding Type I and Type II Errors

Suppose a pharmaceutical company is testing a new drug intended to lower blood pressure. Let:

- $H_0$: The drug has no effect on blood pressure ($\mu = 0$ decrease).
- $H_a$: The drug lowers blood pressure ($\mu > 0$).

**Type I Error**: The company concludes that the drug is effective when, in reality, it has no effect. This could lead to unnecessary costs and side effects for patients.

**Type II Error**: The company fails to conclude that the drug is effective when it actually is. This could result in a missed opportunity to provide a beneficial treatment to patients.

# 4 Power of the Test

The power of a test is the probability that it correctly rejects a false null hypothesis. It is desirable to have a high power, as it reduces the likelihood of a Type II error. Factors affecting the power include:

- Sample size ($n$): Larger sample sizes generally increase power.
- Significance level ($\alpha$): Increasing $\alpha$ increases power, but also increases the risk of a Type I error.
- Effect size: Larger differences between the true population parameter and the null hypothesis value increase power.

## 4.1 Example: Calculating Power

Suppose we are conducting a one-sample t-test with $n = 20$, $\alpha = 0.05$, and we want to detect an effect size of $d = 0.8$. Using power tables or software, we find that the power is approximately 0.78. This means there is a 78% chance of correctly rejecting the null hypothesis if the true effect is $d = 0.8$.

# 5 Distribution Tables

Distribution tables are critical tools for hypothesis testing, providing the necessary critical values for determining the significance of test statistics.

## 5.1 Standard Normal (Z) Table

The standard normal (Z) table provides the area (probability) to the left of a given Z value in the standard normal distribution. It is used to determine p-values or critical values for hypothesis tests involving Z statistics.

**Example: Using the Z Table** Assume you have a Z value of 1.65. To find the p-value:

1. Locate the row corresponding to 1.6 and the column for 0.05.
2. The value at this intersection gives the cumulative probability, which is approximately 0.9505.
3. To find the p-value for a one-tailed test, subtract from 1: $1 - 0.9505 = 0.0495$.

## 5.2 T-Distribution Table

The t-distribution table is used for small sample sizes or when the population variance is unknown. The t-table provides critical values based on the degrees of freedom ($df = n - 1$ for a one-sample test).

extbfExample: Using the T Table Assume you have a t-statistic of 2.35 with $df = 14$.

1. Locate the row corresponding to $df = 14$.
2. Find the closest value to 2.35 in the row to determine the corresponding p-value range.
3. If the value falls between two columns (e.g., 0.025 and 0.01), it means the p-value is between these two probabilities.

## 5.3 Chi-Square Distribution Table

The chi-square distribution table is used primarily for tests of variance and goodness-of-fit. It provides critical values for different degrees of freedom.

**Example: Using the Chi-Square Table** Suppose you have $df = 10$ and want to test at $\alpha = 0.05$:

1. Locate the row for $df = 10$.
2. Find the critical value corresponding to $\alpha = 0.05$, which is approximately 18.31.
3. If your test statistic exceeds 18.31, you reject the null hypothesis.

## 5.4 F-Distribution Table

The F-distribution table is used in analysis of variance (ANOVA) to compare variances. It provides critical values based on numerator and denominator degrees of freedom.

**Example: Using the F Table** Assume you have numerator $df_1 = 5$ and denominator $df_2 = 10$ at $\alpha = 0.05$:

1. Locate the row for $df_1 = 5$ and the column for $df_2 = 10$.
2. The critical value at $\alpha = 0.05$ is approximately 3.33.
3. If your calculated F-statistic is greater than 3.33, reject the null hypothesis.

# 6 Categorical Data Tests

Categorical data tests are used to determine whether observed frequencies differ from expected frequencies in categorical data.

## 6.1 Chi-Square Goodness-of-Fit Test

The chi-square goodness-of-fit test is used to determine if a sample comes from a population with a specific distribution.

**Example: Testing if a Die is Fair** Suppose you roll a die 600 times and observe the following outcomes:

- 1: 151 times
- 2: 97 times
- 3: 93 times
- 4: 104 times
- 5: 105 times
- 6: 50 times

If the die is fair, we expect each number to appear 100 times. We use the chi-square test statistic:

$$\chi^2 = \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i} \tag{4}$$

where $O_i$ is the observed frequency and $E_i$ is the expected frequency.

**Step-by-Step Calculation**

1. Calculate the expected frequency: $E_i = 100$ for all $i$.
2. Compute the chi-square statistic:

$$\chi^2 = \frac{(151-100)^2}{100} + \frac{(97-100)^2}{100} + \cdots + \frac{(50-100)^2}{100} = 52.22 \tag{5}$$

3. Compare $\chi^2$ with the critical value from the chi-square table with $df = 5$ at $\alpha = 0.05$ (critical value is 11.07).
4. Since $\chi^2 = 52.22 > 11.07$, we reject the null hypothesis and conclude that the die is not fair.

## 6.2 Chi-Square Test for Independence

The chi-square test for independence is used to determine if there is an association between two categorical variables.

**Example: Medical Treatment and Institution Type** Suppose we have the following contingency table showing the type of medical treatment and the type of medical institution:

| | Diet | Oral Hypoglycemics | Insulin | Total |
|---|---|---|---|---|
| HMO | 700 | 580 | 420 | 1700 |
| UTH | 350 | 288 | 132 | 770 |
| IPA | 400 | 451 | 238 | 1089 |
| Total | 1450 | 1319 | 790 | 3559 |

We want to determine if there is independence between the type of medical institution and the type of treatment.

1. Calculate the expected frequency for each cell using:

$$E_{ij} = \frac{(\text{row total}) \times (\text{column total})}{\text{grand total}} \tag{6}$$

2. Compute the chi-square statistic:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \tag{7}$$

3. Compare the computed $\chi^2$ with the critical value from the chi-square distribution table with appropriate degrees of freedom.

4. Draw a conclusion based on the comparison.

# 7 Summary

Hypothesis testing involves making decisions about population parameters based on sample statistics. Key steps include defining hypotheses, choosing the appropriate test, calculating the test statistic, and making decisions based on p-values or critical values. Understanding the role of errors, the power of the test, and how to use distribution tables is crucial for proper interpretation.