

Probability & Statistics for IS

Chapter 2: Test Procedures

Lecturers



Dr. Ghandi Manasra
Dr. Monjed H. Samuh

Palestine Polytechnic University
(ghandi@ppu.edu, monjedsamuh@ppu.edu)



Term 191

Table of Contents

- 1 Learning Objectives
- 2 Introduction
- 3 Hypothesis Testing: A Single Population Mean
- 4 Hypothesis Testing: The Difference Between Two Population Means
- 5 Hypothesis Testing: A Single Population Proportion
- 6 Hypothesis Testing: The Difference Between Two Population Proportions

Learning Objectives

After studying this chapter, the student will:

- Understand how to correctly state a null and alternative hypothesis and carry out a structured hypothesis test.
- Understand the concepts of type I error, and type II error.
- Be able to calculate and interpret t test statistics for making statistical inference.
- Understand how to calculate and interpret p-values.

Introduction

- In hypothesis testing, we test a certain claim about a **population parameter**. We may want to find out, using some sample information, whether a given claim (or statement) about a population parameter is true or not.
- A **hypothesis test** is a process that uses sample statistics to test a claim about the population parameter.
- As an example, a soft-drink company may claim that, on average, its cans contain 12 ounces of soda. A government agency may want to test whether such cans do contain, on average, 12 ounces of soda or not.
- Suppose we take a sample of 100 cans of the soft drink under investigation. We then find out that the mean amount of soda in these 100 cans is 11.89 ounces. Based on this result, can we state that, on average, all such cans contain less than 12 ounces of soda and that the company is lying to the public?
- Not until we perform a test of hypothesis can we make such an accusation.

Introduction

- The reason is that the mean, ounces, is obtained from a sample.
- The difference between 12 ounces (the required average amount for the population) and 11.89 ounces (the observed average amount for the sample) may have occurred only because of the sampling error.
- Another sample of 100 cans may give us a mean of 12.04 ounces.
- Therefore, we perform a test of hypothesis to find out how large the difference between 12 ounces and 11.89 ounces is and to investigate whether this difference has occurred as a result of chance alone or not.
- Now, if 11.89 ounces is the mean for all cans and not for just 100 cans, then we do not need to make a test of hypothesis. Instead, we can immediately state that the mean amount of soda in all such cans is less than 12 ounces.
- We perform a test of hypothesis only when we are making a decision about a population parameter based on the value of a sample statistic.

Steps in hypothesis testing

- (1) State the statistical hypotheses.
- (2) Collect the data and calculate the proper test statistic for the problem at hand.
- (3) Identify the distribution of the test statistic and find the critical value using a preassigned level of significance α .
- (4) Make decision and draw conclusion.

Rules for stating the statistical hypotheses

- A verbal statement, or claim, about a population parameter is called a **statistical hypothesis**.
- To test a population parameter, you should carefully state a pair of hypotheses; one that represents the claim and the other, its complement.
- When one of these hypotheses is false, the other must be true.
- A **null hypothesis** H_0 is a statistical hypothesis that contains a statement of equality, such as \leq , \geq , or $=$.
- The **alternative hypothesis** H_A is the complement of the null hypothesis. It is a statement that must be true if H_0 is false and it contains a statement of inequality, such as $>$, $<$ or \neq .
- Some possible pairs of null and alternative hypotheses are:
 - $H_0: \mu \leq \mu_0$ versus $H_A: \mu > \mu_0$, Right-tailed test.
 - $H_0: \mu \geq \mu_0$ versus $H_A: \mu < \mu_0$, Left-tailed test.
 - $H_0: \mu = \mu_0$ versus $H_A: \mu \neq \mu_0$, Two-tailed test. where μ is the population parameter and μ_0 is the claim value.

Rules for stating the statistical hypotheses

- Suppose, for example, that we want to answer the question:

- Can we conclude that a certain population mean is not 50?

- The null and alternative hypotheses are

$$H_0 : \mu = 50 \text{ versus } H_A : \mu \neq 50.$$

- Can we conclude that the population mean is greater than 50?

- The null and alternative hypotheses are

$$H_0 : \mu \leq 50 \text{ versus } H_A : \mu > 50.$$

- Can we conclude that the population mean is less than 50?

- The null and alternative hypotheses are

$$H_0 : \mu \geq 50 \text{ versus } H_A : \mu < 50.$$

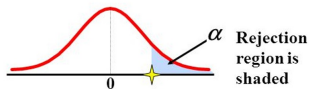
Test Statistic

- A **test statistic** is a value, determined from sample information, used to determine whether to reject the null hypothesis or not.
- If the test statistic falls within a range of values called **acceptance region**, then the null hypothesis is not rejected.
- In hypothesis testing for the mean μ when σ is known or the sample size is large, the test statistic is computed by:

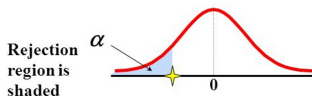
$$Z_{cal} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}.$$

Critical Value(s)

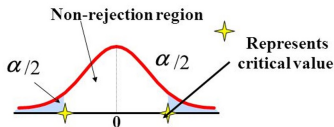
- A **critical value** is a value, determined from the distribution of the test statistic.
 - For right-tailed test, we have the following:



- For left-tailed test, we have the following:



- For two-tailed test, we have the following:



Decision Rule

- **Reject the null hypothesis** if the test statistic (ex. Z_{cal}) falls in the **rejection region**; otherwise, **do not reject** the null hypothesis.
- **Notes:**
 - We should emphasize that if the null hypothesis is not rejected on the basis of the sample data, we can't say that the null hypothesis is true.
 - To put in another way, failing to reject the null hypothesis does not prove that H_0 is true, it means we have failed to disprove H_0 .
 - When you perform a hypothesis test, you make one of two decisions: Reject the null hypothesis or fail to reject the null hypothesis.

Types of Error

- Because your decision is based on a sample rather than the entire population, there is always the possibility you will make the wrong decision.
- The following table shows the four possible outcomes of a hypothesis test.

	Truth of H_0	
Decision	H_0 is true.	H_0 is false.
Reject H_0	<i>Type I error</i>	<i>Correct decision</i>
Do not reject H_0	<i>Correct decision</i>	<i>Type II error</i>

- A **type I error** occurs if the null hypothesis is rejected when it is true.
- A **type II error** occurs if the null hypothesis is not rejected when it is false.

Types of Error

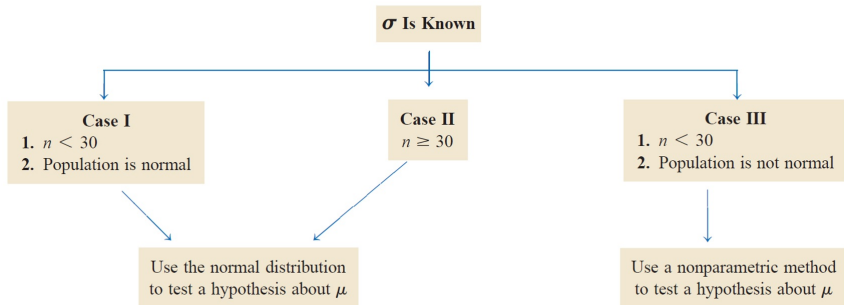
- Because there is variation from sample to sample, there is always a possibility that you will reject a null hypothesis when it is actually true.
- You can decrease the probability of doing so by lowering the level of significance.
- In a hypothesis test, the level of significance is your maximum allowable probability of making a type I error. It is denoted by α .
- By setting the level of significance at a small value, you are saying that you want the probability of rejecting a true null hypothesis to be small.
- Three commonly used levels of significance are $\alpha = 0.10$, $\alpha = 0.05$, $\alpha = 0.01$.

In this chapter

- We will study the following problems:
 - 1 Hypothesis Testing: A Single Population Mean.
 - 2 Hypothesis Testing: The Difference Between Two Population Means.
 - 3 Hypothesis Testing: A Single Population Proportion.
 - 4 Hypothesis Testing: The Difference Between Two Population Proportions.

Hypothesis Testing: A Single Population Mean

Sampling from Normally Distributed Population: **Population Variance Known**



The formula for the test statistic is

$$Z_{cal} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}.$$

Hypothesis Testing: A Single Population Mean

Example

A research team is willing to assume that systolic blood pressures in a certain population of males are approximately normally distributed with a standard deviation of 16. A simple random sample of 64 males from the population had a mean systolic blood pressure reading of 133. At the 0.05 level of significance, do these data provide sufficient evidence for us to conclude that the population mean is greater than 130?

Hypothesis Testing: A Single Population Mean

Example

The management of Priority Health Club claims that its members lose an average of 10 pounds or more within the first month after joining the club. A consumer agency that wanted to check this claim took a random sample of 36 members of this health club and found that they lost an average of 9.2 pounds within the first month of membership. The population standard deviation is known to be 2.4 pounds. Test the claim at $\alpha = 0.01$.

p -Value Method for Hypothesis Testing

- If the null hypothesis is true, a p -value (or probability value) of a hypothesis test is the probability of obtaining a sample statistic with a value as extreme or more extreme than the one determined from the sample data.
- To calculate the p -value for a test of hypothesis:
 - If the test is right-tailed, then $p\text{-value} = P(Z > Z_{cal})$.
 - If the test is left-tailed, then $p\text{-value} = P(Z < Z_{cal})$.
 - If the test is two-tailed, we have two cases:
 - If $Z_{cal} > 0$ then $p\text{-value} = 2P(Z > Z_{cal})$.
 - If $Z_{cal} < 0$ then $p\text{-value} = 2P(Z < Z_{cal})$.
- If the $p\text{-value} < \alpha$, the null hypothesis H_0 is rejected.

p -Value Method for Hypothesis Testing

Example

The Environmental Protection Agency (EPA) recommends that the sodium content in public water supplies should be no more than 20 mg per liter. Forty samples were taken from a large reservoir, and the amount of sodium in each sample was measured. The sample average was 23.5 mg per liter. Assume that the population standard deviation is 5.6 mg per liter. The EPA is interested in knowing whether the average sodium content for the entire reservoir exceeds the recommended level. If so, the communities served by the reservoir will have to be made aware of the violation. Find the p -value for the test of hypothesis. Based on this p -value, would the communities need to be informed of an excessive average sodium level if the maximum probability of a Type I error is to be .05?

One-Sample Z test in R

```
> x <- rnorm(25, 100, 20)
> mean(x)
[1] 101.4046
> sd(x)
[1] 24.8942
> z.test(x, mu = 100, stdev=20, alternative = "two.sided", conf.level = 0.95)
```

One Sample z-test

```
data: x
z = 0.35115, n = 25, Std. Dev. = 20, Std. Dev. of the sample mean = 4,
p-value = 0.7255
alternative hypothesis: true mean is not equal to 100
95 percent confidence interval:
93.56475 109.24447
```

```
> z.test(x, mu = 105, stdev=20, alternative = "two.sided", conf.level = 0.95)
```

One Sample z-test

```
data: x
z = -0.89885, n = 25, Std. Dev. = 20, Std. Dev. of the sample mean = 4,
p-value = 0.3687
alternative hypothesis: true mean is not equal to 105
95 percent confidence interval:
93.56475 109.24447
```

One-Sample Z test in R

```
> z.test(x, mu = 110, stdev=20, alternative = "two.sided", conf.level = 0.95)
```

One Sample z-test

```
data: x
```

```
z = -2.1488, n = 25, Std. Dev. = 20, Std. Dev. of the sample mean = 4,
```

```
p-value = 0.03165
```

```
alternative hypothesis: true mean is not equal to 110
```

```
> z.test(x, mu = 120, stdev=20, alternative = "two.sided", conf.level = 0.95)
```

One Sample z-test

```
data: x
```

```
z = -4.6488, n = 25, Std. Dev. = 20, Std. Dev. of the sample mean = 4,
```

```
p-value = 3.338e-06
```

```
alternative hypothesis: true mean is not equal to 120
```

```
> z.test(x, mu = 130, stdev=20, alternative = "two.sided", conf.level = 0.95)
```

One Sample z-test

```
data: x
```

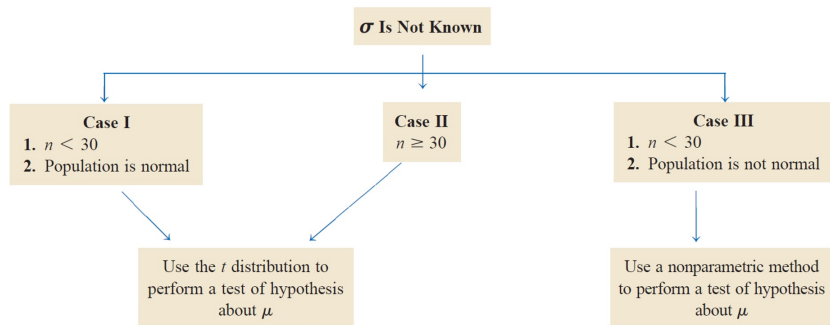
```
z = -7.1488, n = 25, Std. Dev. = 20, Std. Dev. of the sample mean = 4,
```

```
p-value = 8.751e-13
```

```
alternative hypothesis: true mean is not equal to 130
```

Hypothesis Testing: A Single Population Mean

Sampling from Normally Distributed Population: **Population Variance Unknown**



The formula for the test statistic is

$$t_{cal} = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

The degrees of freedom are $df = n - 1$.

Hypothesis Testing: A Single Population Mean

Example

A psychologist claims that the mean age at which children start walking is 12.5 months. Dr. Samuh wanted to check if this claim is true. He took a random sample of 18 children and found that the mean age at which these children started walking was 12.9 months with a standard deviation of .80 month. It is known that the ages at which all children start walking are approximately normally distributed.

- 1 Do these data provide sufficient evidence for him to conclude that the mean age at which all children start walking is different from 12.5 months? Use $\alpha = 0.01$.
- 2 Find the p -value for the test that the mean age at which all children start walking is different from 12.5 months.

One-Sample T test in R

```
> t.test(x, mu = 100, alternative = "two.sided", conf.level = 0.95)
```

One Sample t-test

```
data: x  
t = 0.28212, df = 24, p-value = 0.7803  
alternative hypothesis: true mean is not equal to 100  
95 percent confidence interval:  
91.12879 111.68043
```

```
> t.test(x, mu = 110, alternative = "two.sided", conf.level = 0.95)
```

One Sample t-test

```
data: x  
t = -1.7264, df = 24, p-value = 0.09713  
alternative hypothesis: true mean is not equal to 110
```

```
> t.test(x, mu = 120, alternative = "two.sided", conf.level = 0.95)
```

One Sample t-test

```
data: x  
t = -3.7349, df = 24, p-value = 0.001027  
alternative hypothesis: true mean is not equal to 120
```


Hypothesis Testing: The Difference Between Two Population Means

- Hypothesis testing involving the difference between two population means is most frequently employed to determine whether or not it is reasonable to conclude that the two population means are unequal.
- In such cases, one or the other of the following hypotheses may be formulated:
 - 1 $H_0: \mu_1 - \mu_2 \leq 0$ versus $H_A: \mu_1 - \mu_2 > 0$, Right-tailed test.
 - 2 $H_0: \mu_1 - \mu_2 \geq 0$ versus $H_A: \mu_1 - \mu_2 < 0$, Left-tailed test.
 - 3 $H_0: \mu_1 - \mu_2 = 0$ versus $H_A: \mu_1 - \mu_2 \neq 0$, Two-tailed test.
- It is possible, however, to test the hypothesis that the difference is equal to, greater than or equal to, or less than or equal to some value other than zero, $(\mu_1 - \mu_2)_0$.

Test Statistic

- When σ_1 and σ_2 are known:

$$Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1).$$

- When σ_1 and σ_2 are unknown but equal:

$$t_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{(df)},$$

where $S_p = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}$ and $df = n_1 + n_2 - 2$.

- When σ_1 and σ_2 are unknown and unequal:

$$t_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim t_{(df)},$$

where $df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}}$ (round down).

Hypothesis Testing: The Difference Between Two Population Means

Example

The average length of "short hospital stays" for men is slightly longer than that for women. A random sample of recent hospital stays for 32 men revealed a mean of 5.5 days, and for 30 women revealed a mean of 4.2 days. At $\alpha = 0.01$, is there sufficient evidence to conclude that the average hospital stay for men is longer than the average hospital stays for women? Assume the population standard deviation for men is 1.2 days and for women is 1.5 days. Report the p -value.

Hypothesis Testing: The Difference Between Two Population Means

Example

Two competing headache remedies claim to give fast-acting relief. An experiment was performed to compare the mean lengths of time required for bodily absorption of brand A and brand B headache remedies. Twelve people were randomly selected and given an oral dosage of brand A. Another 12 were randomly selected and given an equal dosage of brand B. The lengths of time in minutes for the drugs to reach a specified level in the blood were recorded. For brand A: the mean is 21.8 min and the standard deviation is 8.7 min. For brand B: the mean is 18.9 min and the standard deviation is 7.5 min. Past experience with the drug composition of the two remedies permits researchers to assume that both distributions are approximately normal with equal variances. Let us use a 5% level of significance to test the claim that there is no difference in the mean time required for bodily absorption. Also, find or estimate the p -value of the sample test statistic.

Two-Sample T test in R

```
> x <- rnorm(25, 100, 20)
> y <- rnorm(30, 110, 20)
> t.test(x, y, alternative = "two.sided", paired = FALSE, var.equal = TRUE, conf.level = 0.95)
```

Two Sample t-test

```
data: x and y
t = -1.4413, df = 53, p-value = 0.1554
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -16.981117  2.780824
sample estimates:
mean of x mean of y
 108.0033  115.1035
```

```
> t.test(x, y, alternative = "two.sided", paired = FALSE, var.equal = FALSE, conf.level = 0.95)
```

Welch Two Sample t-test

```
data: x and y
t = -1.4881, df = 51.53, p-value = 0.1428
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -16.676659  2.476366
sample estimates:
mean of x mean of y
 108.0033  115.1035
```

Two-Sample T test in R

```
> u <- rnorm(25, 100, 10)
> v <- rnorm(30, 110, 10)
> t.test(u, v, alternative = "two.sided", paired = FALSE, var.equal = TRUE, conf.level = 0.95)
```

Two Sample t-test

```
data: u and v
t = -5.7061, df = 53, p-value = 5.293e-07
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -18.418654 -8.837745
sample estimates:
mean of x mean of y
 99.90204 113.53024
```

```
> t.test(u, v, alternative = "two.sided", paired = FALSE, var.equal = FALSE, conf.level = 0.95)
```

Welch Two Sample t-test

```
data: u and v
t = -5.9311, df = 49.966, p-value = 2.806e-07
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -18.243446 -9.012953
sample estimates:
mean of x mean of y
 99.90204 113.53024
```

Paired T test in R

```
> ## Paired t-test
> Sigma <- matrix(c(10,3,3,2),2,2)
> Sigma
      [,1] [,2]
[1,]  10   3
[2,]   3   2
> BVdata <- mvrnorm(n = 25, mu = c(100,110), Sigma)
> x1 <- BVdata[,1]
> x2 <- BVdata[,2]
> t.test(x1, x2, alternative = "two.sided", paired = TRUE, var.equal = FALSE, conf.level = 0.95
```

Paired t-test

```
data: x1 and x2
t = -21.207, df = 24, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -12.37730 -10.18183
sample estimates:
mean of the differences
 -11.27957
```

Hypothesis Testing: A Single Population Proportion

- Many situations arise that call for tests of proportions or percentages rather than means.
- Some possible pairs of null and alternative hypotheses are:
 - $H_0: p \leq p_0$ versus $H_A: p > p_0$, Right-tailed test.
 - $H_0: p \geq p_0$ versus $H_A: p < p_0$, Left-tailed test.
 - $H_0: p = p_0$ versus $H_A: p \neq p_0$, Two-tailed test. where p is the population proportion and p_0 is the claim value.
- The value of the test statistic is computed as:

$$Z_{cal} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim N(0, 1).$$

Hypothesis Testing: A Single Population Proportion

Example

A dietitian claims that 60% of people are trying to avoid trans fats in their diets. She randomly selected 200 people and found that 128 people stated that they were trying to avoid trans fats in their diets. At $\alpha = 0.05$, is there enough evidence to reject the dietitian's claim?

Hypothesis Testing: The Difference Between Two Population Proportions

- Some possible pairs of null and alternative hypotheses are:
 - $H_0: p_1 - p_2 \leq 0$ versus $H_A: p_1 - p_2 > 0$, Right-tailed test.
 - $H_0: p_1 - p_2 \geq 0$ versus $H_A: p_1 - p_2 < 0$, Left-tailed test.
 - $H_0: p_1 - p_2 = 0$ versus $H_A: p_1 - p_2 \neq 0$, Two-tailed test.
- The value of the test statistic is computed as:

$$Z_{cal} = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\frac{\bar{p}(1-\bar{p})}{n_1} + \frac{\bar{p}(1-\bar{p})}{n_2}}} \sim N(0, 1),$$

where \bar{p} is the pooled sample proportion, and is calculated by using one of the following two formulas:

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} \quad \text{or} \quad \bar{p} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2}.$$

Hypothesis Testing: The Difference Between Two Population Proportions

Example

In a study of obesity, the following results were obtained from samples of males and females between the ages of 20 and 75:

Gender	Sample size	Number overweight
Males	150	21
Females	200	48

Can we conclude from these data that in the sampled populations there is a difference in the proportions who are overweight? Let $\alpha = 0.05$. Report the p -value.