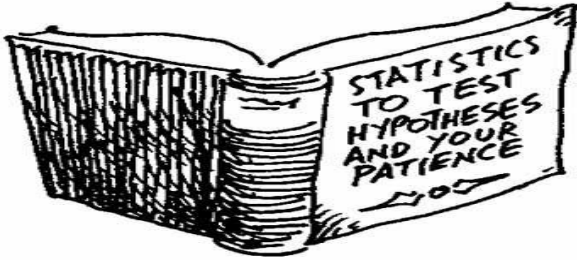


Chapter 11: Introduction to Hypothesis Testing

Chapter 11: Introduction to Hypothesis Testing



1

Introduction

- In this chapter we present the second general procedure of making inferences about a population – hypothesis testing.
- The purpose of hypothesis testing is to determine whether we have enough statistical evidence to conclude that a belief or hypothesis about a parameter is supported by the data.
- This chapter represents a critical contribution to your development as a statistics practitioner because it lays the foundation for the rest of the properties.
- First, we introduce the concepts of hypothesis testing, and then we develop a method to test a hypothesis about a population mean when the population standard deviation is known.

2

2

Concepts of Hypothesis Testing

The critical concepts in **hypothesis testing** may be summarized as follows:

1. There are two hypotheses
 - The **null hypothesis** (H_o – pronounced “H nought”)
 - The **alternative** (or research) **hypothesis** (H_i).
2. The test begins with the assumption that H_o is true.
3. The goal of the test is to determine whether there is enough evidence to infer that H_i is true.
4. There are two possible decisions
 - Reject H_o . Conclude that there is enough evidence to support H_i .
 - Do not reject H_o . Conclude that there is not enough evidence to support H_i .

3

3

Nonstatistical Application of Hypothesis Testing

In a criminal trial, a defendant stands accused of a crime in front of a jury.

It can be said that the jury conducts a test of hypotheses based on the evidence presented by both the prosecution and the defense, that is:

- H_o : The defendant is innocent.
- H_i : The defendant is guilty.

The jury can only make two decisions (i.e., render two verdicts):

- Reject H_o . There is enough evidence to conclude that the defendant is guilty: a conviction.
- Do not reject H_o . There is not enough evidence to conclude that the defendant is guilty: an acquittal.

Under no circumstances, the decision that H_o is true, and the defendant is innocent, is ever made.

4

4

Chapter 11: Introduction to Hypothesis Testing

Error Types

Two possible errors can be made in a test:

- A **Type I error** occurs when we reject a true null hypothesis.
 $P(\text{Type I error}) = \alpha$, also called the **significance level**.
- A **Type II error** occurs when we do not reject a false null hypothesis.

$$P(\text{Type II error}) = \beta$$

The error probabilities α and β are inversely related, and any attempt to reduce one will increase the other. The table on the next slide summarizes the terminology and the concepts.

5

The Blackstone Ratio

- Justice system (in UK and Canada) is arranged by placing the burden of proof on the prosecution.
- The judges instruct the jury to find the defendant guilty only if there is "evidence beyond a reasonable doubt", and in the absence of enough evidence, the jury must acquit.
- It follows that the probability of acquitting guilty people (Type II error) is greater than the probability of convicting innocent people (Type I error.)
- Sir William Blackstone (1723–1780), an English legal scholar, famously quipped: "Better for 10 guilty persons escape than that one innocent suffers." This is known as the **Blackstone Ratio**.

$$\text{The Blackstone Ratio: } \frac{P(\text{Type I error})}{P(\text{Type II error})} = \frac{\alpha}{\beta} = \frac{1}{10}$$

Benjamin Franklin (1706–1790) thought that a 1 to 10 ratio was not enough when he stated: "it is better 100 guilty persons should escape than that one innocent person should suffer, is a maxim that has been long and generally approved."

6

Statistical Application of Hypothesis Testing

Recollect the example of the operations manager at Doll's Computers wanted to estimate the mean demand during lead time. Suppose that the manager wants to know instead whether the mean demand is *different* from 350.

Now the question becomes: "Is there enough evidence to conclude that μ is *not equal* to 350?"

The hypotheses are set up to reflect a manager's decision problem wherein the null hypothesis represents the *status quo*, whereas the *burden of proof* is placed on the alternative hypothesis.

Thus, we write: $H_0: \mu = 350$
 $H_1: \mu \neq 350$

7

Type I and Type II Errors

Conclusion	Population Condition	
	H_0 True ($\mu \leq 12$)	H_0 False ($\mu > 12$)
Accept H_0 (Conclude $\mu \leq 12$)	Correct Decision	Type II Error
Reject H_0 (Conclude $\mu > 12$)	Type I Error	Correct Decision

8

Chapter 11: Introduction to Hypothesis Testing

Statistical Application of Hypothesis Testing

And what if the manager wants to know whether the mean demand has *increased from 350*?

The new question is now: "Is there enough evidence to conclude that μ is *greater than 350*?"

We can write the hypotheses as:

$$H_0: \mu = 350$$

$$H_1: \mu > 350$$

What if the manager wants to know whether the mean demand has *decreased from 350*?

We can phrase the question as: "Is there enough evidence to conclude that μ is *less than 350*?"

And write the hypotheses as:

$$H_0: \mu = 350$$

$$H_1: \mu < 350$$

9

9

A Statistical Application of Hypothesis Testing

Finally, suppose that the manager does not know the actual mean demand, but the current inventory policy is based on the assumption that the mean is *less or equal to 350*. He wants to know whether the mean demand has *increased from 350*?

And write the hypotheses as:

$$H_0: \mu \leq 350$$

$$H_1: \mu > 350$$

We always assume the actual form:

$$H_0: \mu = 350$$

$$H_1: \mu > 350$$

Reason: If the alternative hypothesis is true when $\mu = 350$ then it is true when $\mu \leq 350$

10

10

The Test Statistic

The **test statistic** is based on the best estimator of the population parameter, and it is the criterion used to make our decision about the hypotheses.

In the Doll example, the test statistic is the mean of a random sample drawn from the population.

We use the test statistic to make one of the following two decisions:

- If the test statistic's value is inconsistent with the hypothesized mean of 350, we reject H_0 and conclude that there is *sufficient evidence* that H_1 is true.
- If the test statistic's value is close to the hypothesized mean of 350, we do not reject H_0 and conclude that there is no *sufficient evidence* that H_1 is true.

In the next section, we show how to use the test statistic's sampling distribution to define "sufficient evidence."

11

11

Estimating the Population Mean when the Population Standard Deviation is Known

Example – Department Store's New Billing System

After a thorough financial analysis, the manager of a department store has determined that a new billing system will be cost-effective only if the mean monthly account is more than \$170.

A random sample of 400 monthly accounts has a sample mean of \$178. The manager knows that the accounts are approximately normally distributed with a standard deviation of \$65.

Question: "can the manager conclude that the new system will be cost-effective?"

The null hypothesis can be expressed as:

$$H_0: \mu \leq 170$$

However, as discussed in Section 11-1, we actually test $\mu = 170$, so we set the hypotheses as:

$$H_0: \mu = 170 \quad (\text{system is not cost-effective: do not install new system})$$

$$H_1: \mu > 170 \quad (\text{system is cost-effective: install new system})$$

12

12

Chapter 11: Introduction to Hypothesis Testing

Defining the Rejection Region

Example– Department Store's New Billing System

To conduct the test, we ask: "is a sample mean of 178 sufficiently greater than 170 to conclude that the population mean is greater than 170?"

To make a decision about the sample mean, we set up the *rejection region*:

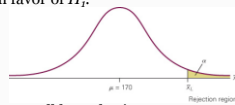
The **rejection region** is a range of values such that if the test statistic falls into that range, we decide to reject H_0 in favor of H_a .

We define the rejection region such that:

$$\bar{x} > \bar{x}_L$$

Because a Type I error is defined as rejecting a true null hypothesis, we can write that:

$$P(\text{Type I error}) = P(\bar{x} > \bar{x}_L | H_0 \text{ is true}) = \alpha$$



13

13

Calculating the Test Statistic from the Rejection Region

Example – Department Store's New Billing System

From before, we know that $\bar{X}$ is normally distributed. As a result, we can standardize $\bar{x}$:

$$P\left(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} > \frac{\bar{x}_L - \mu}{\sigma/\sqrt{n}}\right) = P\left(Z > \frac{\bar{x}_L - \mu}{\sigma/\sqrt{n}}\right) = \alpha$$

From Section 8.2, we defined z_α such that:

$$P(Z > z_\alpha) = \alpha$$

Because the two probabilities have the same distribution, they also have the same limit:

$$\frac{\bar{x}_L - \mu}{\sigma/\sqrt{n}} = z_\alpha$$

Solving for $\bar{x}_L$ yields:

$$\bar{x}_L = \mu + z_\alpha \sigma/\sqrt{n}$$

14

14

Calculating the Test Statistic from the Rejection Region

Example – Department Store's New Billing System

We can choose a significance level, $\alpha = 5\%$, that is, $z_\alpha = 1.645$.

If now we plug in $\mu = 170$, $\sigma = 65$, $n = 400$, we obtain:

$$\bar{x}_L = 170 + 1.645(65/\sqrt{400}) = 175.34$$

Therefore, the rejection region is:

$$\bar{x} > 175.34$$

Because the test statistic (sample mean) is 178, it is in the rejection region.

Thus, we reject the null hypothesis, and conclude that there is sufficient evidence that the mean monthly account is greater than \$170.

15

15

Standardized Test Statistic

- One limitation of the previous method is that the rejection region is set up in terms of $\bar{x}$. It is easier to use the **standardized test statistic** instead:

$$z > \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

- Algebraically, the rejection region can be written as: $z > z_\alpha$

- In our example we have: $z > z_\alpha = z_{0.05} = 1.645$

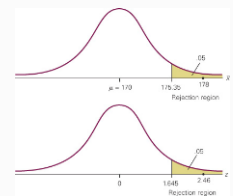
- The standardized test statistic is:

$$z > \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{178 - 170}{65/\sqrt{400}} = 2.46$$

- Because 2.46 is greater than 1.645, we reject the null hypothesis and conclude that there is enough evidence to infer that the mean monthly account is greater than \$170.

- For simplicity, we will refer to the *standardized test statistic* as the *test statistic* for the rest of this course.

Example – Department Store's New Billing System



16

16

Chapter 11: Introduction to Hypothesis Testing

p-Value

The main limitation of the rejection region method is that it produces a *yes* or *no* answer to the question of whether to reject H_0 , but offers no further information on the statistical significance of the test.

One of the implications of using the rejection region method is the impossibility of calculating the exact possibility of making an error (or the exact probability that a Type I error will occur).

A better approach that provides a measure of the amount of statistical evidence supporting H_1 uses the *p-value*.

The **p-value** of a test is the probability of observing a test statistic at least as extreme as the one computed given that the null hypothesis is true.

In our example, the p-value is the probability of observing $\bar{x} \geq 178$ when $\mu = 170$ is:

$$p\text{-value} = P(\bar{X} > 178) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > \frac{178 - 170}{65/\sqrt{400}}\right) = P(Z > 2.46) = 1 - P(Z < 2.46) = 1 - .9931 = .0069$$



17

Interpreting the p-Value

Example – Department Store's New Billing System

We just concluded that the probability of observing $\bar{x} \geq 178$ when $\mu = 170$ is .0069, which is a very unlikely event.

Consequently, we have reason to reject H_0 and support H_1 . But that we cannot make a probability statement about a parameter that is not a random variable.

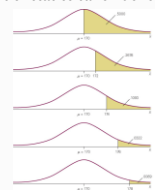
Incorrect interpretation: the p-value is the probability that the H_0 is true.

Correct interpretation: the p-value measures the amount of statistical evidence that supports H_1 .

The table and the graph to the right show several values of $\bar{x}$, their z-statistics, and p-values.

Notice how the farther $\bar{x}$ is from the hypothesized mean of 170, the smaller the p-value is, and the more the statistical evidence supports H_1 .

SAMPLE MEAN ($\bar{x}$)	TEST STATISTIC (Z)	p-VALUE
170	0	.5000
172	0.62	.2676
174	1.23	.1093
176	1.85	.0322
178	2.46	.0069



18

Summary of Suggested Guidelines for Interpreting p-Values

- **Less than .01**
Overwhelming evidence to conclude H_a is true.
- **Between .01 and .05**
Strong evidence to conclude H_a is true.
- **Between .05 and .10**
Weak evidence to conclude H_a is true.
- **Greater than .10**
Insufficient evidence to conclude H_a is true.

19

Describing the p-Value

Example – Department Store's New Billing System

But how small does the p-value have to be to infer that the alternative hypothesis is true? The answer depends on several factors, including the costs of making Type I and Type II errors.

In the example, a Type I error would occur if the manager adopts the new billing system when it is not cost-effective.

If such cost is high, then we want to reject H_0 when the p-value is quite low by setting a very low significance level, such as $\alpha = 1\%$.



In general, the following descriptive terms apply:

- If $p\text{-value} \leq .01$, there is *overwhelming* evidence that H_1 is true. The test is **highly significant**.
- If $.01 < p\text{-value} \leq .05$, there is *strong* evidence that H_1 is true. The test is **significant**.
- If $.05 < p\text{-value} \leq .10$, there is *weak* evidence that H_1 is true. The test is **weakly significant**.
- If $p\text{-value} > .10$, there is *little to no* evidence that H_1 is true. The test is **not significant**.

20

Chapter 11: Introduction to Hypothesis Testing

The p-Value and Rejection Region Methods

We can make similar decisions using either method.

When using the *rejection region* method, the following rules apply:

- If the test statistic is within the rejection region, we reject H_0 .
- If the test statistic is outside the rejection region, we do not reject H_0 .

When using the *p-value* method, the following rules apply:

- If $p\text{-value} \leq \alpha$, we reject H_0 .
- If $p\text{-value} > \alpha$, we do not reject H_0 .

21

21

One- and Two-Tail Tests: Identify

Example – The Effects of Kiosks on Sales in Fast-Food Restaurants

The example before was called a **one-tail test** because the rejection region is located only in one tail of the sampling distribution. We now present an example of a **two-tail test**.

A restaurant franchise, concerned about possible changes in sales due to the installation of self-serve kiosks, takes a random sample of 100 customers who use the new machines at McDonald's.

Question: Has the mean expenditure changed from \$6.03?

And we know that $\mu = 6.03$, $\sigma = 0.91$, and $\bar{x} = 5.91$

It is known that before the machine's installation, the average customer at McDonald's spent \$6.03 with a standard deviation of \$0.91. The sample mean is found to be equal to \$5.91.

We set the hypotheses as:
 $H_0: \mu = 6.03$
 $H_1: \mu \neq 6.03$

Do the data provide enough evidence at the 5% significance level of a change in the size of the transaction for an individual customer?

22

22

One- and Two-Tail Tests: Compute and Interpret

Because we can reject the null hypothesis when the test statistic is large or when it is small, and the total area in the rejection region must be α , we divide this probability by 2:

$$z < z_{\alpha/2} \text{ or } z > z_{\alpha/2}$$

Because $\alpha = .05$, then $\alpha/2 = .025$, and $z_{0.025} = 1.96$.

Thus, we can write the rejection region as:

$$z < -1.96 \text{ or } z > 1.96$$

The value of the test statistic is:

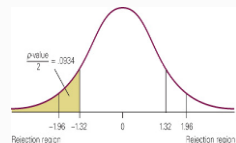
$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{5.91 - 6.03}{0.91/\sqrt{100}} = -1.32$$

Because -1.32 is not in the rejection region, we cannot reject H_0 .

The *p-value* corresponds to the area of both tails:

$$p\text{-value} = P(z < -1.32) + P(z > 1.32) = .0934 + .0934 = .1868 > \alpha$$

There is not enough evidence to conclude that the self-serve kiosks result in a change from the amounts incurred with a cashier.



23

23

When Do We Conduct One- and Two-Tail Tests?

We conduct a **two-tail test** whenever we want to know whether there is enough evidence to infer that the mean is *not equal to* the value stated in the null hypothesis:

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

There are two **one-tail tests**.

We conduct a **right one-tail test** whenever we want to know whether there is enough evidence to infer that the mean is *greater than* the quantity specified by the null hypothesis:

$$H_0: \mu = \mu_0$$

$$H_1: \mu > \mu_0$$

We conduct a **left one-tail test** whenever we want to know whether there is enough evidence to infer that the mean is *less than* the quantity specified by the null hypothesis:

$$H_0: \mu = \mu_0$$

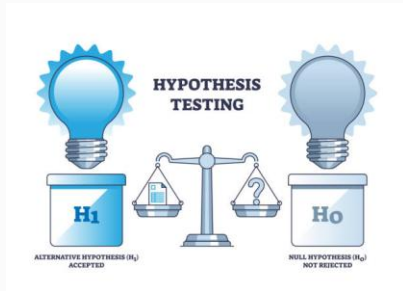
$$H_1: \mu < \mu_0$$

24

24

Chapter 11: Introduction to Hypothesis Testing

Summary of Hypothesis Testing



25

Critical Value Approach to Two-Tailed Hypothesis Testing

- The critical values will occur in both the lower and upper tails of the standard normal curve.
- Use the standard normal probability distribution table to find $z_{\alpha/2}$ (the z-value with an area of $\alpha/2$ in the upper tail of the distribution).
- The rejection rule is:
Reject H_0 if $z \leq -z_{\alpha/2}$ or $z \geq z_{\alpha/2}$.

26

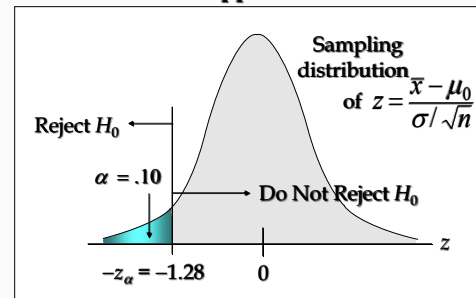
Critical Value Approach to One-Tailed Hypothesis Testing

- The test statistic z has a standard normal probability distribution.
- We can use the standard normal probability distribution table to find the z-value with an area of α in the lower (or upper) tail of the distribution.
- The value of the test statistic that established the boundary of the rejection region is called the critical value for the test.
 - The rejection rule is:
 - Lower tail: Reject H_0 if $z \leq -z_\alpha$
 - Upper tail: Reject H_0 if $z \geq z_\alpha$

27

Lower-Tailed Test About a Population Mean: σ Known

■ Critical Value Approach

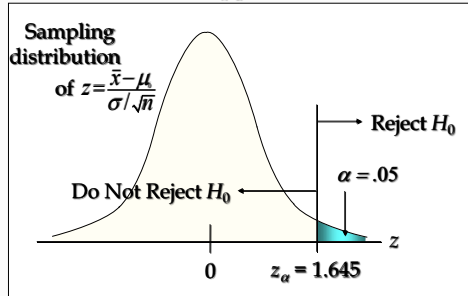


28

Chapter 11: Introduction to Hypothesis Testing

Upper-Tailed Test About a Population Mean: σ Known

■ Critical Value Approach



29

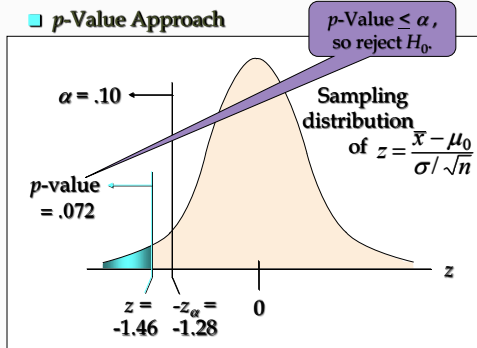
p-Value Approach to Two-Tailed Hypothesis Testing

- Compute the p-value using the following three steps:
 1. Compute the value of the test statistic z .
 2. If z is in the upper tail ($z > 0$), find the area under the standard normal curve to the right of z .
If z is in the lower tail ($z < 0$), find the area under the standard normal curve to the left of z .
 3. Double the tail area obtained in step 2 to obtain the p -value.
- The rejection rule:
Reject H_0 if the p -value $\leq \alpha$.

30

Lower-Tailed Test About a Population Mean: σ Known

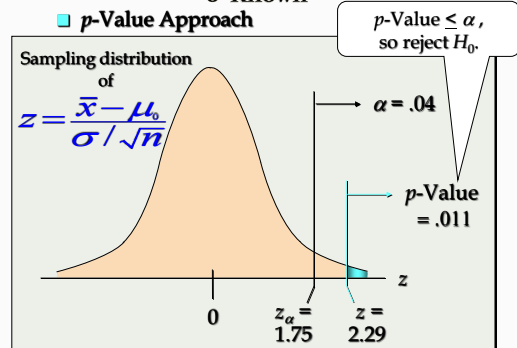
■ p-Value Approach



31

Upper-Tailed Test About a Population Mean: σ Known

■ p-Value Approach



32

Chapter 11: Introduction to Hypothesis Testing

Summary of Forms for Null and Alternative Hypotheses about a Population Mean

- The equality part of the hypotheses always appears in the null hypothesis.
- In general, a hypothesis test about the value of a population mean μ must take one of the following three forms (where μ_0 is the hypothesized value of the population mean).

$H_0: \mu \geq \mu_0$ $H_a: \mu < \mu_0$	$H_0: \mu \leq \mu_0$ $H_a: \mu > \mu_0$	$H_0: \mu = \mu_0$ $H_a: \mu \neq \mu_0$
One-tailed (lower-tail)	One-tailed (upper-tail)	Two-tailed

33

Confidence Interval Approach to Two-Tailed Tests About a Population Mean

- Select a simple random sample from the population and use the value of the sample mean $\bar{x}$ to develop the confidence interval for the population mean μ .
- If the confidence interval contains the hypothesized value μ_0 , do not reject H_0 . Otherwise, reject H_0 . (Actually, H_0 should be rejected if μ_0 happens to be equal to one of the end points of the confidence interval.)

34

33

34

Testing Hypotheses and Confidence Interval

Example – The Effects of Kiosks on Sales in Fast-Food Restaurants

The test statistic and the confidence interval estimator are both derived from the sampling distribution. It follows that we can use the confidence interval estimator to test hypotheses.

The 95% confidence interval estimate of the population mean is:

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 5.91 \pm 1.96 \frac{0.91}{\sqrt{100}} = 5.91 \pm 0.18$$

LCL = 5.73 and UCL = 6.09

We estimate that μ lies between \$5.73 and \$6.09. Because the interval includes \$6.03, we cannot conclude that there is sufficient evidence to infer that the mean has changed.

This process is equivalent to the rejection region approach.

35

35

Steps of Hypothesis Testing Once Again

- Step 1.** Develop the null and alternative hypotheses.
- Step 2.** Specify the level of significance α .
- Step 3.** Collect the sample data and compute the test statistic.

p-Value Approach

- Step 4.** Use the value of the test statistic to compute the p-value.
- Step 5.** Reject H_0 if $p\text{-value} \leq \alpha$.

Critical Value Approach

- Step 4.** Use the level of significance to determine the critical value and the rejection rule.
- Step 5.** Use the value of the test statistic and the rejection rule to determine whether to reject H_0 .

36

36

Chapter 11: Introduction to Hypothesis Testing

Null and Alternative Hypotheses Example

■ Metro EMS

A major west coast city provides one of the most comprehensive emergency medical services in the world. Operating in a multiple hospital system with approximately 20 mobile medical units, the service goal is to respond to medical emergencies with a mean time of 12 minutes or less.

The director of medical services wants to formulate a hypothesis test that could use a sample of emergency response times to determine whether or not the service goal of 12 minutes or less is being achieved.

37

37

Null and Alternative Hypotheses

$$H_0: \mu \leq 12$$

The emergency service is meeting the response goal; no follow-up action is necessary.

$$H_a: \mu > 12$$

The emergency service is not meeting the response goal; appropriate follow-up action is necessary.

where: μ = mean response time for the population of medical emergency requests

38

38

One-Tailed Tests About a Population Mean: σ Known

■ Metro EMS

The response times for a random sample of 40 medical emergencies were tabulated. The sample mean is 13.25 minutes. The population standard deviation is believed to be 3.2 minutes.

The EMS director wants to perform a hypothesis test, with a .05 level of significance, to determine whether the service goal of 12 minutes or less is being achieved.

39

39

One-Tailed Tests About a Population Mean: σ Known

■ p -Value and Critical Value Approaches

1. Develop the hypotheses. $H_0: \mu \leq 12$
 $H_a: \mu > 12$

2. Specify the level of significance. $\alpha = .05$

3. Compute the value of the test statistic.

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{13.25 - 12}{3.2 / \sqrt{40}} = 2.47$$

40

40

Chapter 11: Introduction to Hypothesis Testing

One-Tailed Tests About a Population Mean: σ Known

■ p -Value Approach

4. Compute the p -value.

For $z = 2.47$, cumulative probability = .9932.

$$p\text{-value} = 1 - .9932 = .0068$$

5. Determine whether to reject H_0 .

Because $p\text{-value} = .0068 \leq \alpha = .05$, we reject H_0 .

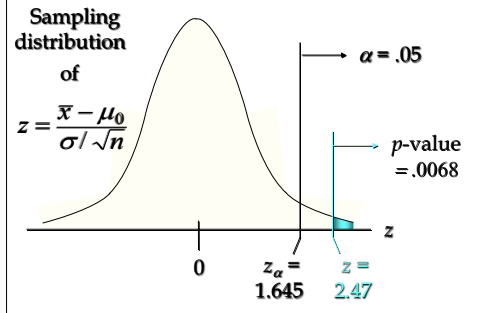
There is sufficient statistical evidence to infer that Metro EMS is not meeting the response goal of 12 minutes.

41

41

One-Tailed Tests About a Population Mean: σ Known

■ p -Value Approach



42

42

One-Tailed Tests About a Population Mean: σ Known

■ Critical Value Approach

4. Determine the critical value and rejection rule.

For $\alpha = .05$, $z_{.05} = 1.645$

Reject H_0 if $z \geq 1.645$

5. Determine whether to reject H_0 .

Because $2.47 \geq 1.645$, we reject H_0 .

There is sufficient statistical evidence to infer that Metro EMS is not meeting the response goal of 12 minutes.

43

43

Two-Tailed Tests About a Population Mean: σ Known

■ Example: Glow Toothpaste

The production line for Glow toothpaste is designed to fill tubes with a mean weight of 6 oz. Periodically, a sample of 30 tubes will be selected in order to check the filling process.

Quality assurance procedures call for the continuation of the filling process if the sample results are consistent with the assumption that the mean filling weight for the population of toothpaste tubes is 6 oz.; otherwise the process will be adjusted.

44

44

Chapter 11: Introduction to Hypothesis Testing

Two-Tailed Tests About a Population Mean: σ Known

■ Example: Glow Toothpaste

Assume that a sample of 30 toothpaste tubes provides a sample mean of 6.1 oz. The population standard deviation is believed to be 0.2 oz.

Perform a hypothesis test, at the .03 level of significance, to help determine whether the filling process should continue operating or be stopped and corrected.

45

Two-Tailed Tests About a Population Mean: σ Known

■ p -Value and Critical Value Approaches

1. Determine the hypotheses.

$$H_0: \mu = 6$$

$$H_a: \mu \neq 6$$

2. Specify the level of significance. $\alpha = .03$

3. Compute the value of the test statistic.

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{6.1 - 6}{.2 / \sqrt{30}} = 2.74$$

46

Two-Tailed Tests About a Population Mean: σ Known

■ p -Value Approach

4. Compute the p -value.

For $z = 2.74$, cumulative probability = .9969

$$p\text{-value} = 2(1 - .9969) = .0062$$

5. Determine whether to reject H_0 .

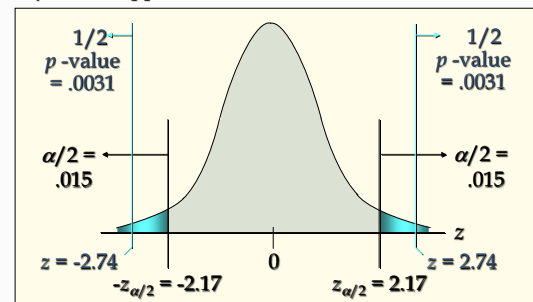
Because $p\text{-value} = .0062 \leq \alpha = .03$, we reject H_0 .

There is sufficient statistical evidence to infer that the alternative hypothesis is true (i.e. the mean filling weight is not 6 ounces).

47

Two-Tailed Tests About a Population Mean: σ Known

■ p -Value Approach



48

Chapter 11: Introduction to Hypothesis Testing

Two-Tailed Tests About a Population Mean: σ Known

■ Critical Value Approach

4. Determine the critical value and rejection rule.

For $\alpha/2 = .03/2 = .015$, $z_{.015} = 2.17$

Reject H_0 if $z \leq -2.17$ or $z \geq 2.17$

5. Determine whether to reject H_0 .

Because $2.74 \geq 2.17$, we reject H_0 .

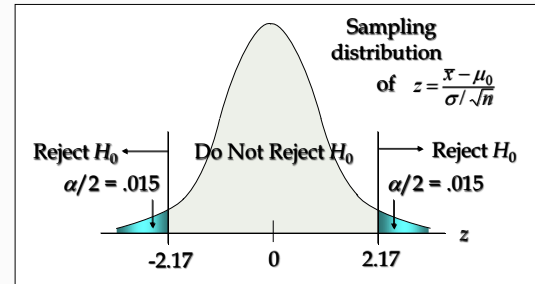
There is sufficient statistical evidence to infer that the alternative hypothesis is true (i.e. the mean filling weight is not 6 ounces).

49

49

Two-Tailed Tests About a Population Mean: σ Known

■ Critical Value Approach



50

50

Confidence Interval Approach to Two-Tailed Tests About a Population Mean

The 97% confidence interval for μ is

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 6.1 \pm 2.17(2/\sqrt{30}) = 6.1 \pm .07924$$

or 6.02076 to 6.17924

Because the hypothesized value for the population mean, $\mu_0 = 6$, is not in this interval, the hypothesis-testing conclusion is that the null hypothesis, $H_0: \mu = 6$, can be rejected.

51

51

A Summary of Forms for Null and Alternative Hypotheses About a Population Proportion

- The equality part of the hypotheses always appears in the null hypothesis.
- In general, a hypothesis test about the value of a population proportion p must take one of the following three forms (where p_0 is the hypothesized value of the population proportion).

$H_0: p \geq p_0$ $H_a: p < p_0$	$H_0: p \leq p_0$ $H_a: p > p_0$	$H_0: p = p_0$ $H_a: p \neq p_0$
One-tailed (lower tail)	One-tailed (upper tail)	Two-tailed

52

52

Chapter 11: Introduction to Hypothesis Testing

Tests About a Population Proportion

■ Test Statistic

$$z = \frac{\bar{p} - p_0}{\sigma_{\bar{p}}}$$

where:

$$\sigma_{\bar{p}} = \sqrt{\frac{p_0(1-p_0)}{n}}$$

assuming $np \geq 5$ and $n(1-p) \geq 5$

53

Tests About a Population Proportion

■ Rejection Rule: p -Value Approach

Reject H_0 if $p\text{-value} \leq \alpha$

■ Rejection Rule: Critical Value Approach

$H_0: p \leq p_0$ Reject H_0 if $z \geq z_{\alpha}$

$H_0: p \geq p_0$ Reject H_0 if $z \leq -z_{\alpha}$

$H_0: p = p_0$ Reject H_0 if $z \leq -z_{\alpha/2}$ or $z \geq z_{\alpha/2}$

54

53

54

Two-Tailed Test About a Population Proportion

■ Example: National Safety Council (NSC)

For a Christmas and New Year's week, the National Safety Council estimated that 500 people would be killed and 25,000 injured on the nation's roads. The NSC claimed that 50% of the accidents would be caused by drunk driving.

A sample of 120 accidents showed that 67 were caused by drunk driving. Use these data to test the NSC's claim with $\alpha = .05$.

55

55

Two-Tailed Test About a Population Proportion

■ p -Value and Critical Value Approaches

1. Determine the hypotheses. $H_0: p = .5$ and $H_a: p \neq .5$

2. Specify the level of significance. $\alpha = .05$

3. Compute the value of the test statistic.

$$\sigma_{\bar{p}} = \sqrt{\frac{p_0(1-p_0)}{n}} = \sqrt{\frac{.5(1-.5)}{120}} = .045644$$

a common error is using $\bar{p}$ in this formula

$$z = \frac{\bar{p} - p_0}{\sigma_{\bar{p}}} = \frac{\left(\frac{67}{120}\right) - .5}{.045644} = 1.28$$

56

56

Chapter 11: Introduction to Hypothesis Testing

Two-Tailed Test About a Population Proportion

■ p -Value Approach

4. Compute the p -value.

For $z = 1.28$, cumulative probability = .8997

$$p\text{-value} = 2(1 - .8997) = .2006$$

5. Determine whether to reject H_0 .

Because $p\text{-value} = .2006 > \alpha = .05$,
we cannot reject H_0 .

57

Interval Estimate of a Population Proportion

The general form of an interval estimate of a population proportion is

$$\bar{p} \pm \text{Margin of Error}$$

The sampling distribution of $\bar{p}$ plays a key role in computing the margin of error for this interval estimate.

The sampling distribution of $\bar{p}$ can be approximated by a normal distribution whenever $np \geq 5$ and $n(1 - p) \geq 5$.

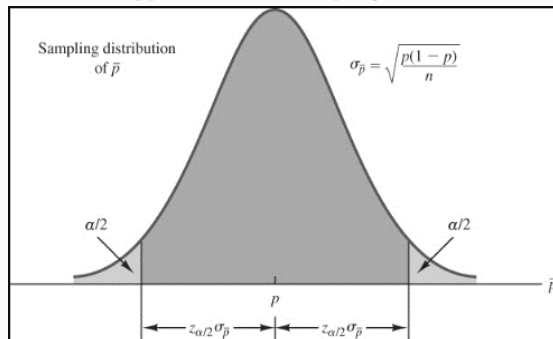
58

57

58

Interval Estimate of a Population Proportion

■ Normal Approximation of Sampling Distribution of $\bar{p}$



59

Interval Estimate of a Population Proportion

■ Interval Estimate

$$\bar{p} \pm z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

where: $1 - \alpha$ is the confidence coefficient

$z_{\alpha/2}$ is the z value providing an area of $\alpha/2$ in the upper tail of the standard normal probability distribution

$\bar{p}$ is the sample proportion

60

60

Chapter 11: Introduction to Hypothesis Testing

Interval Estimate of a Population Proportion

Example: Survey of women golfers

A national survey of 900 women golfers was conducted to learn how women golfers view their treatment at golf courses in United States. The survey found that 396 of the women golfers were satisfied with the availability of tee times.

Suppose one wants to develop a 95% confidence interval estimate for the proportion of the population of women golfers satisfied with the availability of tee times.

61

61

Interval Estimate of a Population Proportion

Example: Survey of women golfers

$$\bar{p} \pm z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

where: $n = 900$, $\bar{p} = 396/900 = .44$, $z_{\alpha/2} = 1.96$

$$.44 \pm 1.96 \sqrt{\frac{.44(1-.44)}{900}} = .44 \pm .0324$$

Survey results enable us to state with 95% confidence that between 40.76% and 47.24% of all women golfers are satisfied with the availability of tee times.

62

62

Sample Size for an Interval Estimate of a Population Proportion

Margin of Error

$$E = z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

Solving for the necessary sample size n , we get

$$n = \frac{(z_{\alpha/2})^2 \bar{p}(1-\bar{p})}{E^2}$$

However, $\bar{p}$ will not be known until after we have selected the sample. We will use the planning value p^* for $\bar{p}$.

63

63

Sample Size for an Interval Estimate of a Population Proportion

Necessary Sample Size

$$n = \frac{(z_{\alpha/2})^2 p^*(1-p^*)}{E^2}$$

The planning value p^* can be chosen by:

1. Using the sample proportion from a previous sample of the same or similar units, or
2. Selecting a preliminary sample and using the sample proportion from this sample.
3. Using judgment or a "best guess" for a p^* value.
4. Otherwise, using .50 as the p^* value.

64

64

Chapter 11: Introduction to Hypothesis Testing

Sample Size for an Interval Estimate of a Population Proportion

Example: Survey of women golfers

Suppose the survey director wants to estimate the population proportion with a margin of error of .025 at 95% confidence.

How large a sample size is needed to meet the required precision? (A previous sample of similar units yielded .44 for the sample proportion.)

65

Sample Size for an Interval Estimate of a Population Proportion

Example: Survey of women golfers

$$E = z_{\alpha/2} \sqrt{\frac{p^*(1-p^*)}{n}} = .025$$

At 95% confidence, $z_{.0125} = 1.96$. Recall that $p^* = .44$.

$$n = \frac{(z_{\alpha/2})^2 p^*(1-p^*)}{E^2} = \frac{(1.96)^2 (.44)(.56)}{(.025)^2} = 1514.5$$

A sample of size 1515 is needed to reach a desired precision of $\pm .025$ at 95% confidence.

66

Interval Estimate of a Population Proportion (another example)

■ **Example:** Political Science, Inc.

Political Science, Inc. (PSI) specializes in voter polls and surveys designed to keep political office seekers informed of their position in a race. Using telephone surveys, PSI interviewers ask registered voters who they would vote for if the election were held that day.

In a current election campaign, PSI has just found that 200 registered voters, out of 500 contacted, favor a particular candidate. PSI wants to develop a 90% confidence interval estimate for the proportion of the population of registered voters that favor the candidate.

67

Interval Estimate of a Population Proportion

$$\bar{p} \pm z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

where: $n = 500$, $\bar{p} = 200/500 = .40$, $z_{.05} = 1.645$

$$.40 \pm 1.645 \sqrt{\frac{.40(1-.40)}{500}} = .40 \pm .036$$

PSI is 90% confident that the proportion of all voters that favor the candidate is between .364 and .436.

68

Chapter 11: Introduction to Hypothesis Testing

Sample Size for an Interval Estimate of a Population Proportion (another example)

■ Example: Political Science, Inc.

Suppose that PSI would like a .99 probability that the sample proportion is within $\pm .03$ of the population proportion.

How large a sample size is needed to meet the required precision? (A previous sample of similar units yielded .40 for the sample proportion.)

69

Sample Size for an Interval Estimate of a Population Proportion

$$z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}} = .03$$

At 99% confidence, $z_{.005} = 2.576$. Recall that $\bar{p} = .40$.

$$n = \frac{(z_{\alpha/2})^2 p(1-p)}{E^2} = \frac{(2.576)^2 (.40)(.60)}{(.03)^2} \approx 1769$$

A sample of size 1769 is needed to reach a desired precision of $\pm .03$ at 99% confidence.

70

Stata Output and Instructions

Example 11.1 – Testing a Population Mean: Standard Deviation Known

Instructions:

1. Import the data into one column. (Click File/Import/Excel spreadsheet("xls","xlsx)/Chapter11/Xm11-01.) Check **Import first row as variable names**.
2. Click **Statistics, Summaries, tables and tests, Classical tests of hypotheses, and z-test (mean comparison test known variance)**.
3. Select **One-sample**, select **Accounts** in the **Variable name:** box and type 170 in the **Hypothesized mean:** box. Type 65 for the **Standard deviation**.

One-sample z test					
Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]
Accounts	480	177.9965	3.25	65	171.6266 184.3664
mean = mean(Accounts) z = 2.4685					
Ho: mean = 170					
Ha: mean < 170 Ha: mean != 170 Ha: mean > 170					
Pr(Z < z) = 0.9931 Pr(Z > z) = 0.0139 Pr(Z > z) = 0.0069					

Keller, Gerald, Statistics for Management and Economics, 12th Edition. © 2023 Cengage. All Rights Reserved

71

Stata Output and Instructions

Example 11.2 – Testing a Population Mean: Standard Deviation Known

Instructions:

1. Import the data into one column. (Click File/Import/Excel spreadsheet("xls","xlsx)/Chapter11/Xm11-02.) Check **Import first row as variable names**.
2. Click **Statistics, Summaries, tables and tests, Classical tests of hypotheses, and z-test (mean comparison test known variance)**.
3. Select **One-sample**, select **AmountSpent** in the **Variable name:** box and type 6.03 in the **Hypothesized mean:** box. Type 0.91 for the **Standard deviation**.

One-sample z test					
Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]
Amount-t	100	5.91	.091	.91	5.731643 6.088357
mean = mean(AmountSpent) z = -1.3187					
Ho: mean = 6.03					
Ha: mean < 6.03 Ha: mean != 6.03 Ha: mean > 6.03					
Pr(Z < z) = 0.0936 Pr(Z > z) = 0.1873 Pr(Z > z) = 0.9064					

Keller, Gerald, Statistics for Management and Economics, 12th Edition. © 2023 Cengage. All Rights Reserved

72