CHAPTER 8

Statistical Inference

- Point Estimation
- Confidence Interval Estimation
- Hypothesis Testing

CHAPTER 8

Chapter 8: Inferences Based on a Single Sample: *Tests of Hypotheses*

Hypothesis Testing

Example: We might need to decide whether the mean weight of all bags of ecclairs packaged by a particular company differs from the advertised weight of 454 grams.

One of the most commonly used methods for making such decisions or judgments about a specific value of a population parameter is to perform a hypothesis test.

Hypothesis:

A hypothesis is a statement that something is true.

Example: The statement "the mean weight of all bags of pretzels packaged differs from the advertised weights of 454 grams" is a hypothesis.

Statistical Hypothesis:

A statistical hypothesis is a conjecture about a population parameter.

A Test is a rule to check a hypothesis.

C H A P T E R 8

Setting up the hypotheses:

Clearly state both the null hypothesis and the alternative hypothesis using the appropriate symbols and also describing the claims in words expressing the context of the current problem. (You must have both the verbal and symbolic form.)

Remember, hypotheses are always in terms of the parameter (eg., μ, p, etc.) **NOT** the statistic (eg., $\bar{x}$, $\hat{p}$, etc.)

C H A P T E R 8

Null Hypothesis:

- It is denoted by H_0.

- It is a hypothesis to be tested, and the decision is based on the null hypothesis.

- It is a statement of "no effect" , or "no difference", or "the difference is null".

- If the original claim includes equality (=) it is the null hypothesis. If the original claim does not include equality ($<$, $\neq$, $>$), then the null hypothesis is the complement of the original claim. The null hypothesis always includes the equal sign.

Alternative Hypothesis:

- It is denoted by H_a or H_1
- It is a statement which is true if the null hypothesis is false

-Alternative hypothesis can be <,>,≠ type.

- The type of test (left-tail, right-tail, or two-tail is based on the alternative hypothesis

- It is also called research hypothesis.

CHAPTER 8

The null hypothesis is usually stated as an equality …

$H_0: \mu = \mu_0$

$H_a: \mu < \mu_0$

$H_a: \mu \neq \mu_0$

$H_a: \mu > \mu_0$

… even though the alternative hypothesis can be either an equality or an inequality.

The form of a null hypothesis is

H_0: population characteristic = hypothesized value (e.g., H_0: μ = 454 gram)

where the hypothesized value is a specific number determined by the problem context.

The alternative hypothesis will have one of the following three forms:

H_a: population characteristic > hypothesized value (e.g., H_a: μ >454)

H_a: population characteristic < hypothesized value (e.g., H_a: μ < 454)

H_a: population characteristic ≠ hypothesized value (e.g., H_a: $\mu \neq$ 454)

Identifying Hypotheses

Example: Teenagers (age 15 to 20) make up 7% of the driving population. The article "More States Demand Teens Pass Rigorous Driving Tests" (*San Luis Obispo Tribune*, January 27, 2000) described a study of auto accidents conducted by the Insurance Institute for Highway Safety. The Institute found that 14% of the accidents studied involved teenage drivers. Suppose that this percentage was based on examining records from 500 randomly selected accidents. Does the study provide convincing evidence that the proportion of accidents involving teenage drivers differs from .7, the proportion of teens in the driving population? Use $\alpha = .05$.

C

H

A

P

T

E

R

8

Identifying Hypotheses

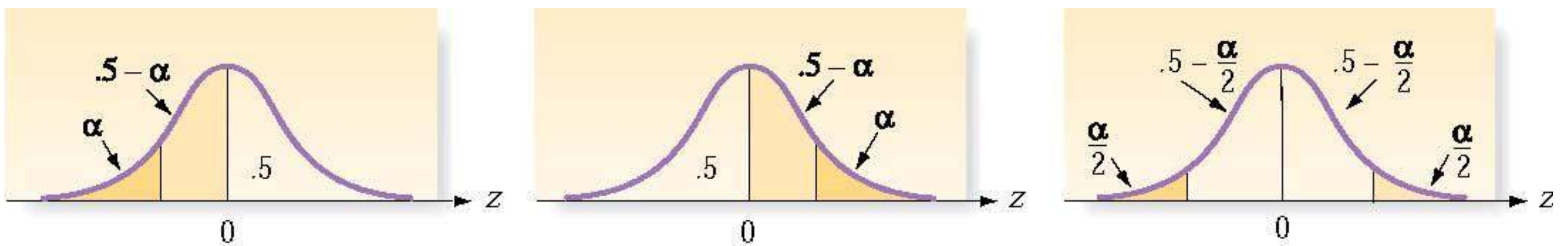
Example : A researcher is interested in determining if children have average cholesterol levels that are higher than the national average. Suppose that the distribution of cholesterol levels in children is normal with a population standard deviation of 15. If the national average cholesterol is 190 and a sample of 100 children yields sample mean cholesterol of 196.2, determine if children have mean cholesterol levels higher than the national average. Use a significance level of $\alpha = 0.01$.

Identifying Hypotheses

A metal lathe is checked periodically by quality control inspectors to determine whether it is producing machine bearings with a mean diameter of .5 inch. If the mean diameter of the bearings is larger or smaller than .5 inch, then the process is out of control and must be adjusted. Formulate the null and alternative hypotheses for a test to determine whether the bearing production process is out of control.

A statistical test is:

- **Left-tailed** if H_1 states that the parameter is less than the value claimed in H_0
- **Right-tailed** if H_1 states that the parameter is greater than the value claimed in H_0
- **Two-tailed** if H_1 states that the parameter is different from (or not equal to) the value claimed in H_0



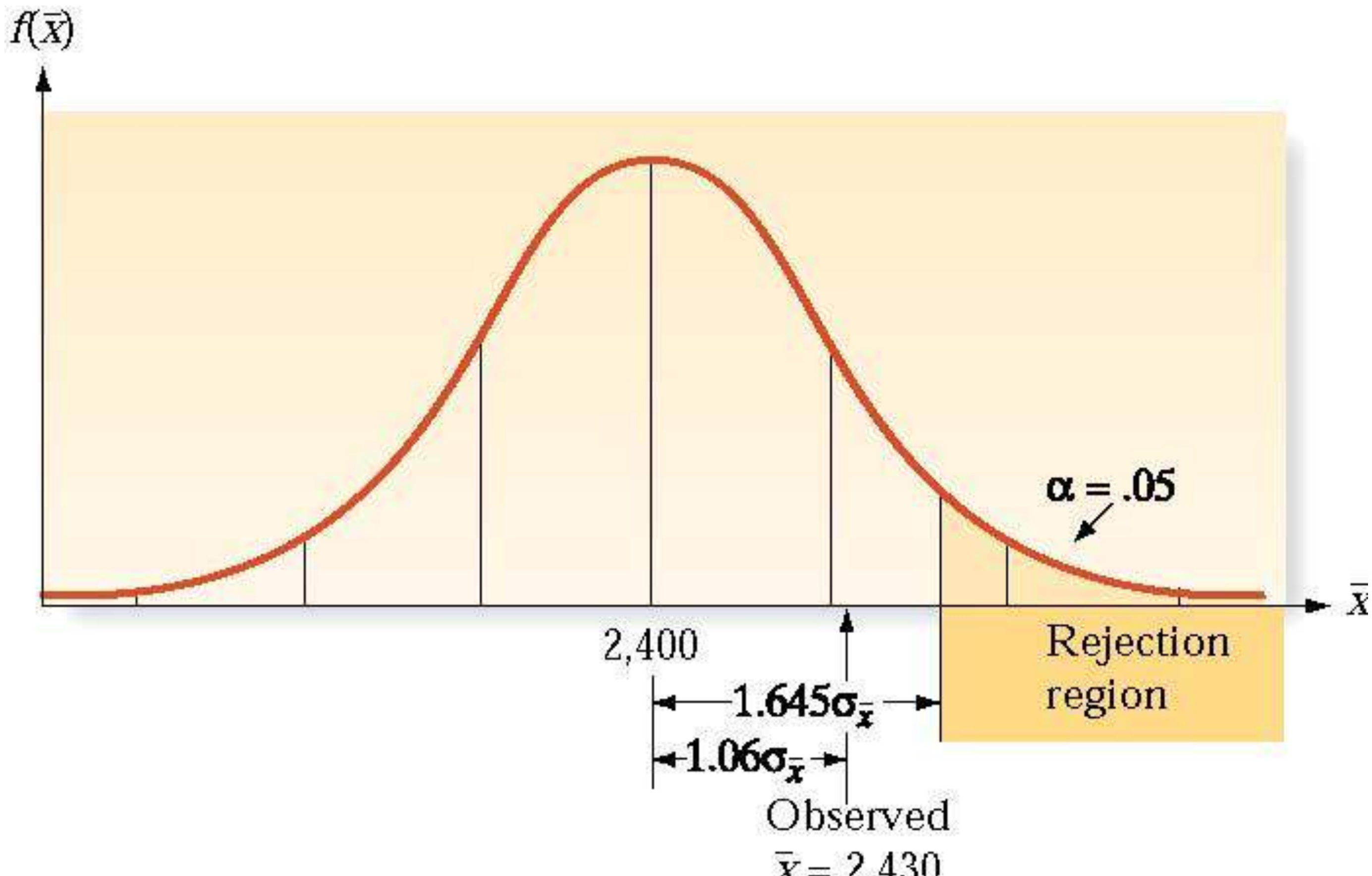
a. Form of H_a: <   b. Form of H_a: >   c. Form of H_a: ≠

Scenarios for the null and alternative hypotheses for tests of the mean μ are given below:

Null hypothesis	**Alternative Hypotheses and Type of Test**		
Claim about μ or historical value of μ $H_0:\ \mu = k$	You believe that μ is less than the value stated in H_0 $H_1:\ \mu < k$ Left-tailed test	You believe that μ is more than the value stated in H_0 $H_1:\ \mu > k$ Right-tailed test	You believe that μ is different from the value stated in H_0 $H_1:\ \mu \neq k$ Two-tailed test

CHAPTER 8

- If the *test statistic* has a high probability when H_0 is true, then H_0 is *not rejected.*
- If the *test statistic* has a (very) low probability when H_0 is true, then H_0 is *rejected.*



Errors in Hypothesis Testing

Example 2: A Jury Trial

If on a jury, must presume defendant is innocent unless enough evidence to conclude is guilty.

Null hypothesis:	Defendant is ***innocent***.
Alternative hypothesis:	Defendant is ***guilty***.

- Trial held because prosecution believes status quo of innocence is incorrect.
- Prosecution collects evidence, like researchers collect data, in hope that jurors will be convinced that such evidence is extremely unlikely if the assumption of innocence were true.

Courtroom Analogy: *Potential choices and errors*

Choice 1: We cannot rule out that defendant is innocent, so he or she is set free without penalty.
Potential error: A criminal has been erroneously freed.

Choice 2: We believe enough evidence to conclude the defendant is guilty.
Potential error: An innocent person falsely convicted and guilty party remains free.

Choice 2 is usually seen as ***more serious***.

C
H
A
P
T
E
R

8

The Reasoning of Hypothesis Testing is Often Compared to that of a Jury Trial

		Defendant Is Actually	
		Innocent	**Guilty**
Jury's Decision	**Not Guilty**	Correct	Error
	Guilty	Worse Error	Correct

- *Each trial actually has 4 potential decisions – two are correct decisions, two are errors.*
- *Possible decisions are based on:*
  - *the reality about the defendant's innocence or guilt;*
  - *the decision that the jury makes based on the evidence*

CHAPTER 8

The Reasoning of Hypothesis Testing is Similar

Your Decision		Null Hypothesis Is Actually: True	Null Hypothesis Is Actually: False
	Don't Reject the Null	Correct	Type II Error
	Reject the Null	Type I Error	Correct

- *Each test has 4 potential decisions – two are correct decisions, two are errors.*
- *Possible decisions are based on:*
  - ☐ *the reality about the null hypothesis;*
  - ☐ *your decision based on the evidence from the sample.*

C
H
A
P
T
E
R
8

Type I & Type II Errors

- *Type I error:* *we reject H_0 when in fact H_0 is true*

  ☐ *The significance level, α, of a test is the probability of a Type I error.*

  Thus, a test with $\alpha = 0.01$ is said to have a level of significance of 0.01 or to be a level 0.01 test.

  ☐ *We control the probability of making a Type I error, which gives credibility to our decision if we decide to reject H_0 in favor of H_a.*

- *Type II error:* *we fail to reject H_0 when in fact H_0 is false.*

  ☐ *Decreasing the chance of a type I error can increase the chance of a type II error.*

  ☐ *The probability of a type II error is denoted by β.*

CHAPTER 8

Steps to do Hypothesis Testing:

Hypotheses: Formulate the null hypothesis (H_0) and alternative hypothesis (H_a).

Test Statistic: Clearly identify the test statistic to be used to test the hypothesis and why it is reasonable to use this test. All required assumptions must be satisfactorily addressed in this step.

Level of significance: It is denoted by α (alpha) and chosen to be α equal to 0.01, 0.05, or 0.10. It is the probability of rejecting the null hypothesis when it is true. α can be treated as the complement of the level of confidence in estimation.

Rejection region: The values for the test statistic which lead to rejection of the null hypothesis. OR, *p*-value (section 3.3)

Calculation of test statistic: The appropriate calculation for the test based on the sample data.

Decision and Conclusion: Reject the null hypothesis (with possible Type I error) or do not reject it (with possible Type II error)

Always provide a conclusion that is in the context of the problem and that answers the original research question which the hypothesis test was designed to answer.

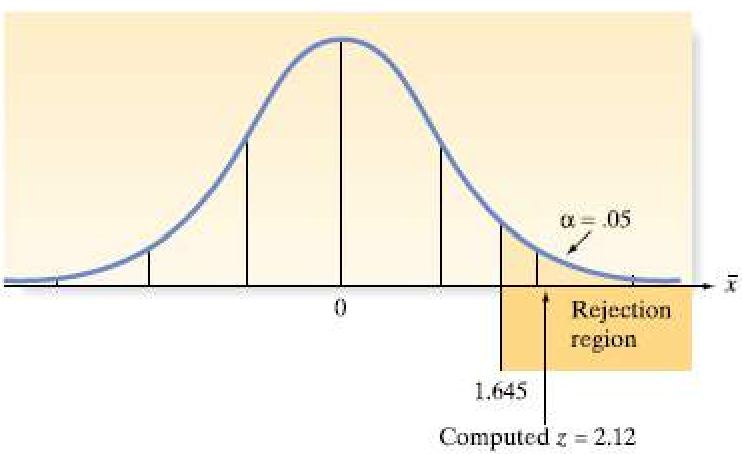
CHAPTER 8

Rejection Regions for Common Values of α for large sample z-test

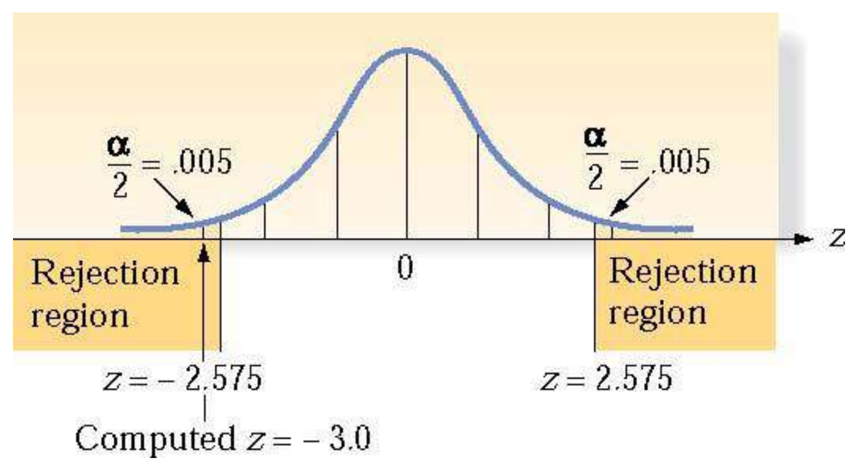
	Lower Tailed	Upper Tailed	Two tailed
$\alpha = .10$	$z < -1.28$	$z > 1.28$	$z < -1.645$ or $z > 1.645$
$\alpha = .05$	$z < -1.645$	$z > 1.645$	$z < -1.96$ or $z > 1.96$
$\alpha = .01$	$z < -2.33$	$z > 2.33$	$z < -2.575$ or $z > 2.575$

Examples

α =.05 (Upper Tailed Test)



α =.01 (Two Tailed Test)



How to write the conclusion?

Conclusions are based on the original claim, which may be the null or alternative hypotheses.

Decision	Original Claim	
	H_0 "Reject"	H_a "Support"
Reject H_0 "sufficient"	We conclude that there is sufficient evidence at …% level of significance to reject the claim that……	We conclude that there is sufficient evidence at …% level of significance to support the claim that……
Fail to reject H_0 "insufficient"	We conclude that there is insufficient evidence at …% level of significance to reject the claim that……	We conclude that there is insufficient evidence at …% level of significance to support the claim that……

CHAPTER 8

8.3 Test of Hypotheses about a Population Mean: Normal (z) Statistic

Hypotheses:

Let μ be the population mean.

Null Hypothesis, H_0: $\mu = \mu_0$, a specified value; (in words)

Alternative hypothesis,

Ha: $\mu > \mu_0$, a specified value; (in words)

or

Ha: $\mu < \mu_0$, a specified value; (in words)

or

Ha: $\mu \neq \mu_0$, a specified value; (in words)

Test Statistic

We will perform a large sample z-test for the population mean.

Conditions:

- A random sample is selected from the target population.
- The sample size *n* is large.

Verification of these assumptions makes it reasonable to assume the approximate normality of the sampling distribution of sample mean. Therefore, we can perform the z-test.

C
H
A
P
T
E
R

8

Level of significance

α = ??? (usually given, if not choose α =.05)

Rejection Region

$z > z_\alpha$ if upper tailed test (or Ha: $\mu > \mu_0$)

or

$z < -z_\alpha$ if lower tailed test (or Ha: $\mu < \mu_0$)

or

$z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$ if two tailed test (or Ha: $\mu \neq \mu_0$)

Calculation of test statistic

$$z_{cal} = \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

Decision and conclusion

If the calculated value of the proposed test statistic belongs to the rejection region, we reject H_0; otherwise we fail to reject H_0.

Write the conclusion in the context of question.

C
H
A
P
T
E
R

8

Example

The effect of drugs and alcohol on the nervous system has been the subject of considerable research. Suppose a research neurologist is testing the effect of a drug on response time by injecting 100 rats with a unit dose of the drug, subjecting each rat to a neurological stimulus, and recording the response time. The sample mean and standard deviation of the response times are 1.05 seconds and .5 second, respectively. The neurologist knows that the mean response time for rats not injected with the drug (the ·control" mean) is 1.2 seconds. She wishes to test whether the mean response time for drug-injected rats differ from 1.2 seconds. Set up the test of hypothesis for this experiment, using α=.01.

C
H
A
P
T
E
R

8

i) Hypotheses.

CHAPTER 8

ii) Test Statistic

CHAPTER 8

iii) Level of significance

iv) Rejection Region

v) Calculation of test statistic

C

H

A

P

T

E

R

8

vi) Decision and Conclusion