07. Hypothesis Testing: One Sample

Adapted From :

Probability & Statistics for Engineers & Scientists, 9th Ed.

Walpole/Myers/Myers/Ye (c)2010

Introduction to Business Statistics, 5e

Kvanli/Guynes/Pavur (c)2000

South-Western College Publishing

Statistics for Managers

Using Microsoft® Excel 4th Edition

Learning Objectives

After completing this session, you should be able to:

- Formulate null and alternative hypotheses for applications involving a single population mean or proportion
- Formulate a decision rule for testing a hypothesis
- Know how to use the critical value and p-value approaches to test the null hypothesis (for both mean and proportion problems)
- Know what Type I and Type II errors are

What is a Hypothesis?

- A hypothesis is a claim (assumption) about a population parameter:
  - population mean

Example: The mean monthly cell phone bill of this city is μ = $42

  - population proportion

Example: The proportion of adults in this city with cell phones is $p = .68$

The Null Hypothesis, H_0

- States the assumption (numerical) to be tested

> Example: The average number of TV sets in U.S. Homes is equal to three $(H_0 : \mu = 3)$

- Is always about a population parameter, not about a sample statistic

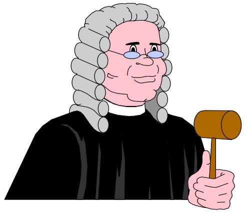
$$H_0 : \mu = 3$$

$$\cancel{H_0 : \bar{X} = 3}$$

The Null Hypothesis, H_0

(continued)

- Begin with the assumption that the null hypothesis is true
  - Similar to the notion of innocent until proven guilty
- Refers to the status quo
- Always contains "=" sign
- May or may not be rejected



The Alternative Hypothesis, H_1

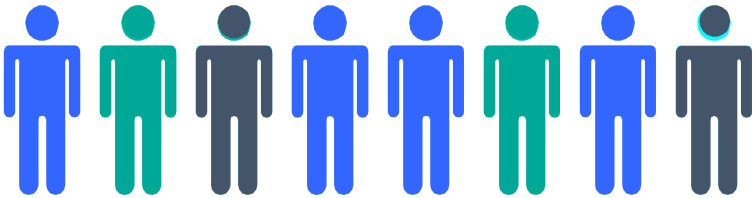
- Is the opposite of the null hypothesis
  - e.g.: The average number of TV sets in U.S. homes is not equal to 3 ($H_1: \mu \neq 3$)
- Challenges the status quo
- Never contains the "=" , "≤" or "≥" sign
- May or may not be accepted
- Is generally the hypothesis that is believed (or needs to be supported) by the researcher

Hypothesis Testing Process

Claim: **the population mean age is 50. (Null Hypothesis:**

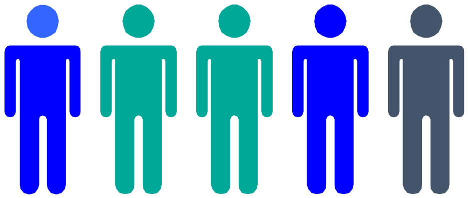
$$H_0: \mu = 50$$

)



Population

Now select a random sample



Sample

Suppose the sample mean age is 20: $\bar{X} = 20$

Is $\bar{X} = 20$ likely if $\mu = 50$?

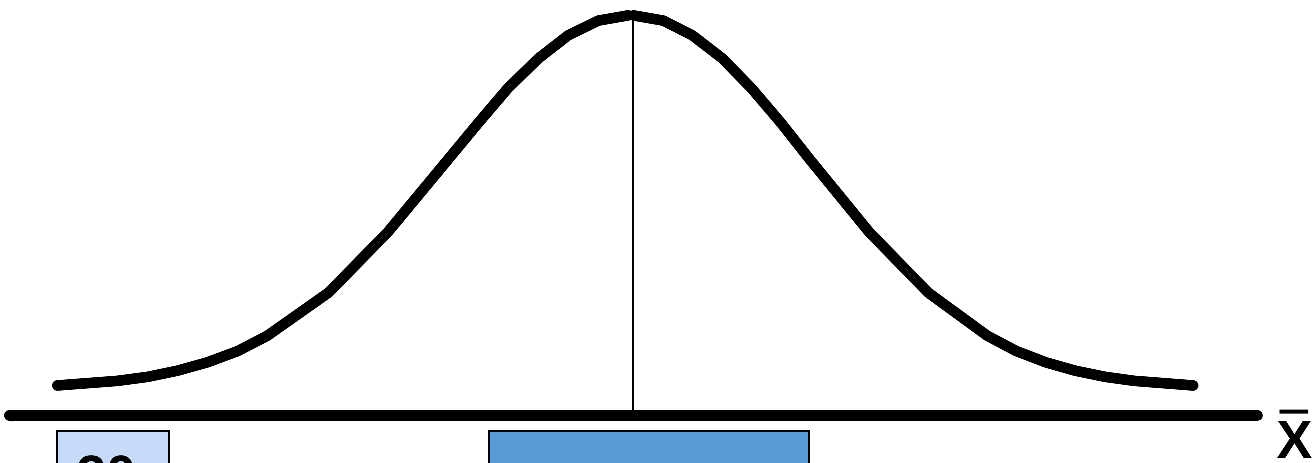
If not likely,

REJECT

Null Hypothesis

Reason for Rejecting H_0

Sampling Distribution of $\bar{X}$



20

$\mu = 50$
If H_0 is true

If it is unlikely that we would get a sample mean of this value ...

... if in fact this were the population mean…

... then we reject the null hypothesis that $\mu = 50$.

Level of Significance, α

- **Defines the unlikely values of the sample statistic if the null hypothesis is true**
  - Defines rejection region of the sampling distribution
- Is designated by α , (level of significance)
  - Typical values are .01, .05, or .10
- Is selected by the researcher at the beginning
- Provides the critical value(s) of the test

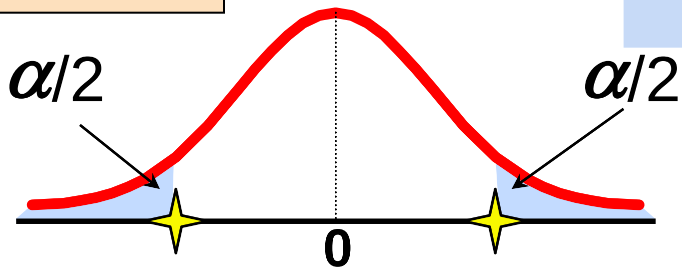
Level of Significance and the Rejection Region

Level of significance = α

Represents critical value

H_0: $\mu = 3$

H_1: $\mu \neq 3$

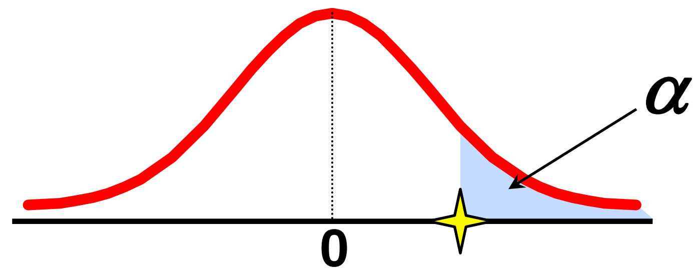


Two tailed test

Rejection region is shaded

H_0: $\mu = 3$

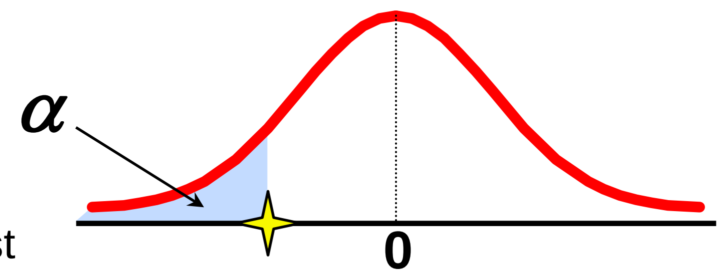
H_1: $\mu > 3$



Upper tail test

H_0: $\mu = 3$

H_1: $\mu < 3$



Lower tail test

Errors in Making Decisions

- **Type I Error**
  - Reject a true null hypothesis
  - Considered a serious type of error

The probability of Type I Error is α

- Called level of significance of the test
- Set by researcher in advance

Errors in Making Decisions

(continued)

- **Type II Error**
  - Fail to reject a false null hypothesis

The probability of Type II Error is β

Outcomes and Probabilities

Possible Hypothesis Test Outcomes

	Actual Situation	
Decision	H_0 True	H_0 False
Do Not Reject $\mathbf{H_0}$	**No error** **(1 - α)**	**Type II Error** **(β)**
Reject $\mathbf{H_0}$	**Type I Error** **(α)**	**No Error** **(1 - β)**

Key:
Outcome
(Probability)

Type I & II Error Relationship

- Type I and Type II errors can not happen at the same time
  - Type I error can only occur if H_0 is true
  - Type II error can only occur if H_0 is false

If Type I error probability (α) ⬆, then

Type II error probability (β) ⬇

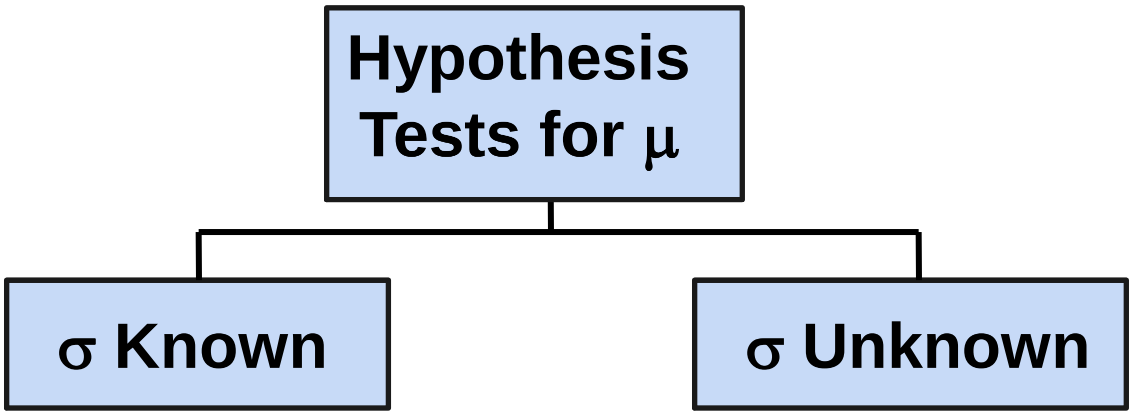
Factors Affecting Type II Error

- $\beta \uparrow$ when the difference between hypothesized parameter and its true value $\downarrow$
- $\beta \uparrow$ when $\alpha \downarrow$
- $\beta \uparrow$ when $\sigma \uparrow$
- $\beta \uparrow$ when $n \downarrow$

How to choose between Type I and Type II errors

- Choice depends on the cost of the error
- Choose little type I error when the cost of rejecting the maintained hypothesis is high
  - A criminal trial: convicting an innocent person
  - The Exxon Valdise: Causing an oil tanker to sink
- Choose large type I error when you have an interest in changing the status quo
  - A decision in a startup company about a new piece of software
  - A decision about unequal pay for a covered group.

Hypothesis Tests for the Mean



Z Test of Hypothesis for the Mean (σ Known)

- Convert sample statistic ($\bar{X}$) to a Z test statistic

Hypothesis Tests for μ

σ Known

σ Unknown

The test statistic is:

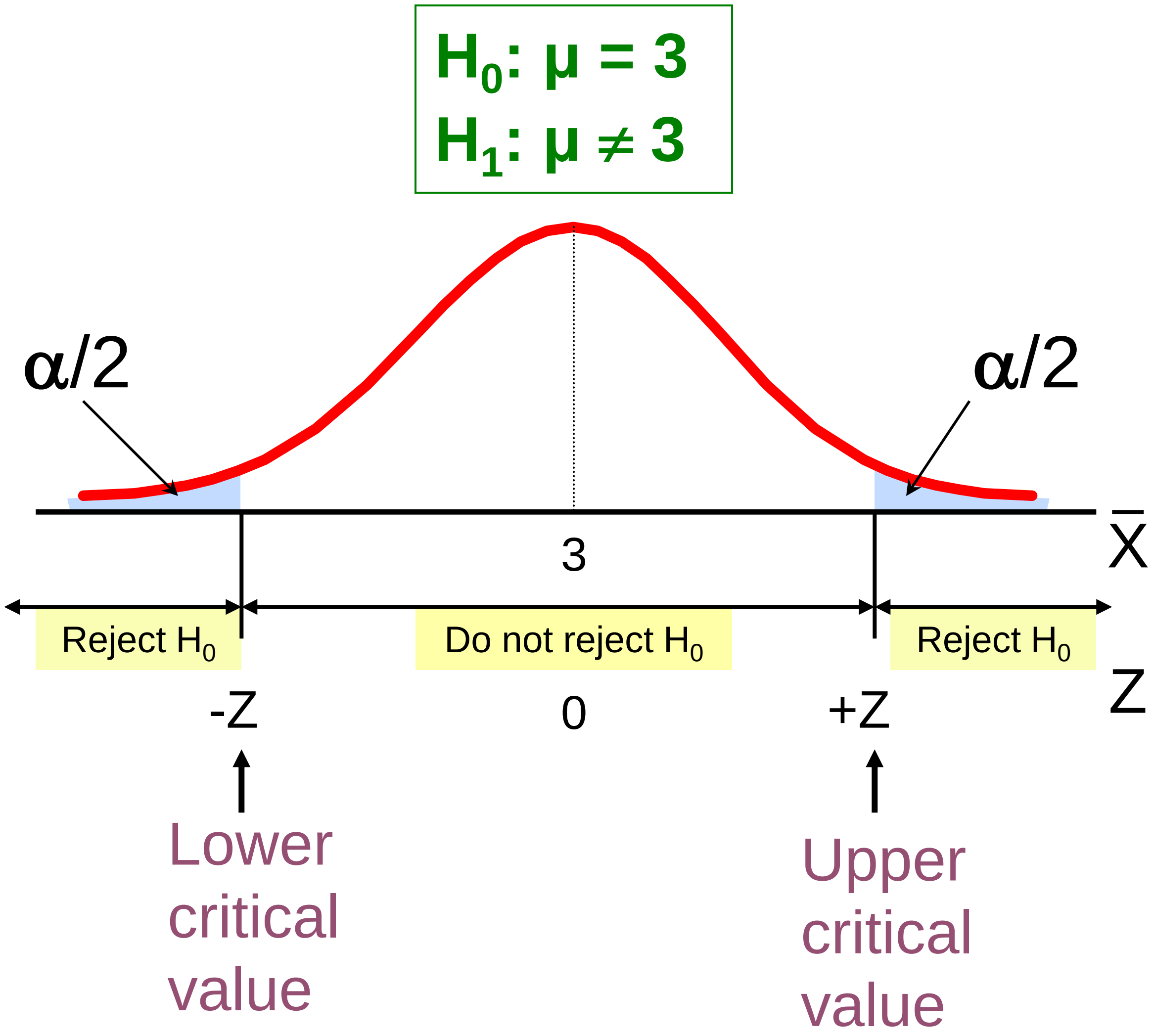
$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Critical Value Approach to Testing

- For two tailed test for the mean, σ known:
- Convert sample statistic ($\bar{X}$) to test statistic (Z statistic)
- Determine the critical Z values for a specified level of significance α from a table or computer
- Decision Rule: If the test statistic falls in the rejection region, reject H_0 ; otherwise do not reject H_0

Two Tailed Tests

- There are two cutoff values (critical values), defining the regions of rejection



Review: 10 Steps in Hypothesis Testing

1. State the null hypothesis, H_0
2. State the alternative hypotheses, H_1
3. Choose the level of significance, α
4. Choose the sample size, n
5. Determine the appropriate statistical technique and the test statistic to use
6. Find the critical values and determine the rejection region(s)

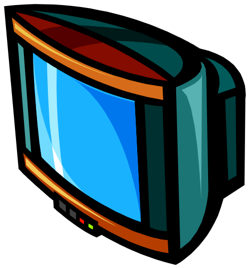
Review: 10 Steps in Hypothesis Testing

7. Collect data and compute the test statistic from the sample result
8. Compare the test statistic to the critical value to determine whether the test statistics falls in the region of rejection
9. Make the statistical decision: Reject H_0 if the test statistic falls in the rejection region
10. Express the decision in the context of the problem

Hypothesis Testing Example

Test the claim that the true mean # of TV sets in US homes is equal to 3.
(Assume σ = 0.8)

- 1-2. State the appropriate null and alternative hypotheses
  - H_0: μ = 3 H_1: μ ≠ 3 (This is a two tailed test)
- 3. Specify the desired level of significance
  - Suppose that α = .05 is chosen for this test
- 4. Choose a sample size
  - Suppose a sample of size n = 100 is selected



Hypothesis Testing Example *(continued)*

- 5. Determine the appropriate technique
  - σ is known so this is a Z test
- 6. Set up the critical values
  - For α = .05 the critical Z values are ±1.96
- 7. Collect the data and compute the test statistic
  - Suppose the sample results are

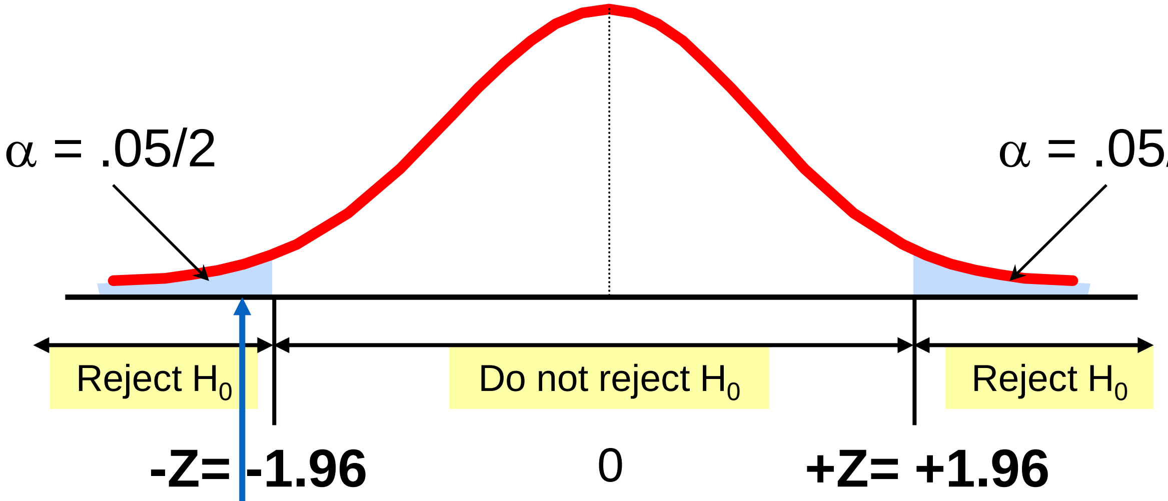
    $n = 100$, $\bar{X} = 2.84$ ($\sigma = 0.8$ is assumed known)

So the test statistic is:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{2.84 - 3}{\frac{0.8}{\sqrt{100}}} = \frac{-.16}{.08} = -2.0$$

Hypothesis Testing Example *(continued)*

- 8. Is the test statistic in the rejection region?

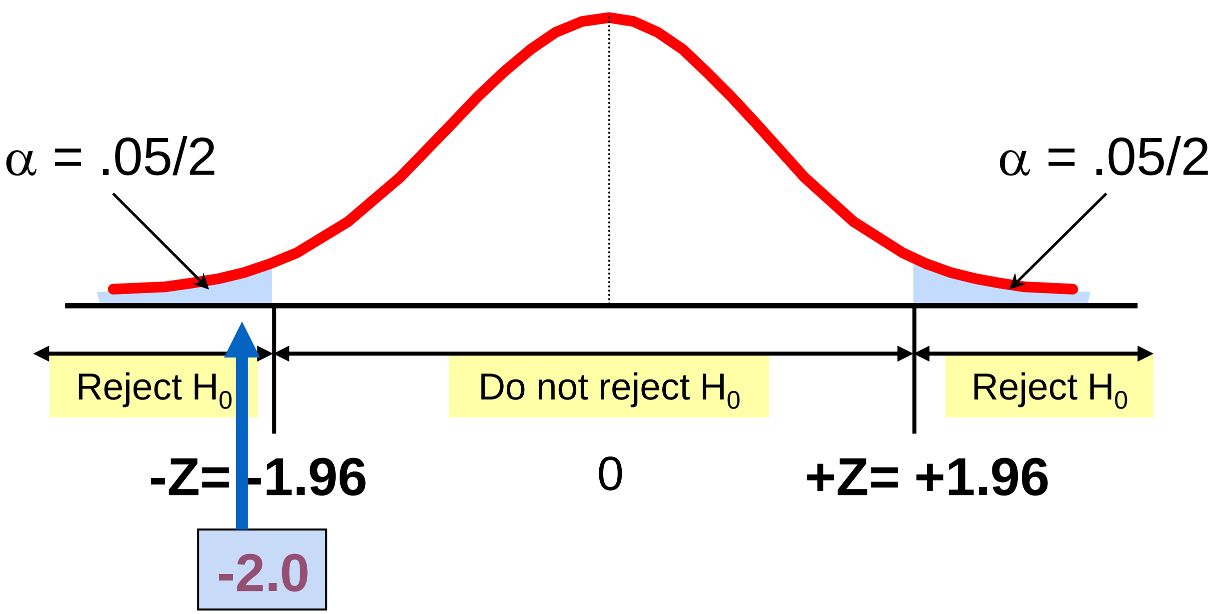


Reject H_0 if $Z < -1.96$ or $Z > 1.96$; otherwise do not reject H_0

Here, $Z = -2.0 < -1.96$, so the test statistic is in the rejection region

Hypothesis Testing Example *(continued)*

- 9-10. Reach a decision and interpret the result



Since Z = -2.0 < -1.96, we reject the null hypothesis and conclude that there is sufficient evidence that the mean number of TVs in US homes is not equal to 3

Pre-selecting α vs. using a P-Value

- In classical hypothesis testing, we typically pre-select α to be .05 or .01 and then determine the critical region.
  - We can then reject the hypothesis with that level of significance. (Remember that in a 2-sided test, with $\alpha = .05$, for example, the critical region would have .025 in either tail.)
- The alternative is calculating the P-value, or probability of obtaining the calculated result if H_0 is true. The P-value provides more information than just that the hypothesis was rejected or not.
  - If rejected, the P-value may be much less than .05 or .01, giving us additional confidence in our decision.
  - If not rejected, the P-value may be very close to .05 or .01, allowing us the option of rejecting at a slightly reduced level.
  - The judgment of the experimenter is used to interpret the calculated P-value result.

P-Value Approach to Testing

- p-value: Probability of obtaining a test statistic more extreme ($\leq$ or $\geq$) than the observed sample value given H_0 is true
  - Also called observed level of significance
  - Smallest value of α for which H_0 can be rejected

P-Value Approach to Testing *(continued)*

- Convert Sample Statistic (e.g. $\bar{X}$) to Test Statistic (e.g. Z statistic)
- Obtain the p-value from a table or computer
- Compare the p-value with α
  - If p-value $< \alpha$, reject H_0
  - If p-value $\geq \alpha$, do not reject H_0

P-Value Example

- Example: How likely is it to see a sample mean of 2.84 (or something further from the mean, in either direction) if the true mean is $\mu = 3.0$?

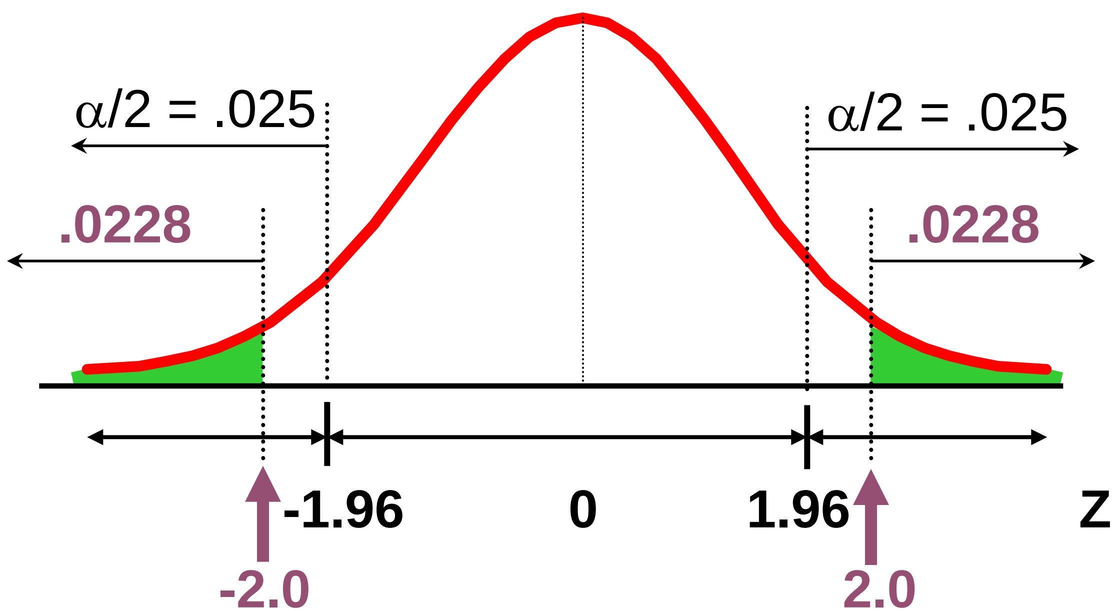
$\bar{X} = 2.84$ is translated to a Z score of Z = -2.0

$$P(Z < -2.0) = .0228$$

$$P(Z > 2.0) = .0228$$

p-value

=.0228 + .0228 = .0456



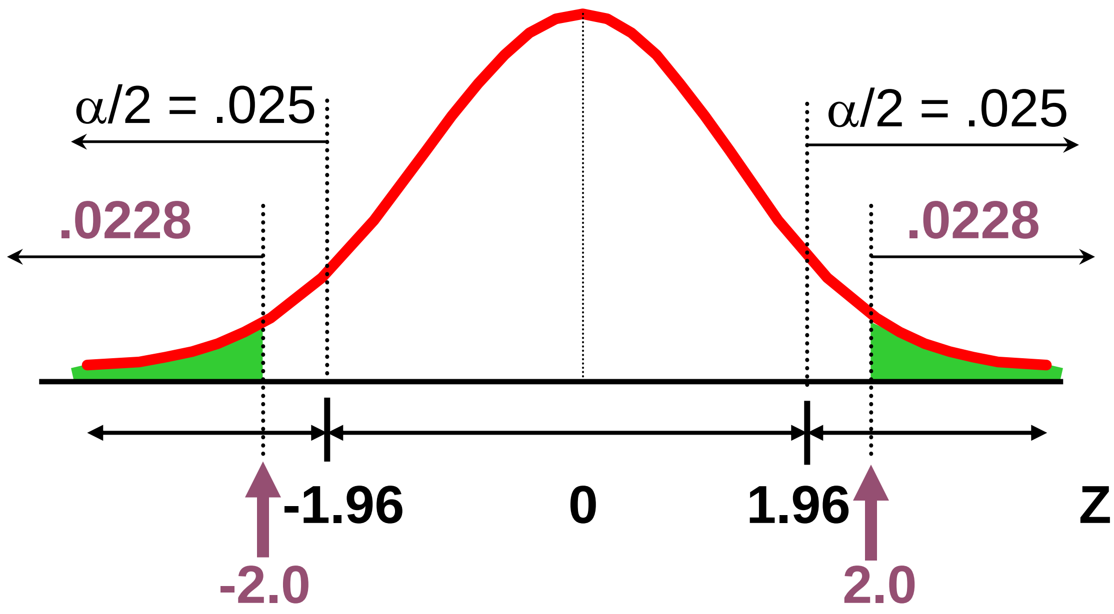
P-Value Example

(continued)

- Compare the p-value with α
  - If p-value < α, reject H_0
  - If p-value $\geq$ α, do not reject H_0

Here: p-value = .0456
α = .05

Since .0456 < .05, we reject the null hypothesis



Connection to Confidence Intervals

- For $\bar{X} = 2.84$, $\sigma = 0.8$ and $n = 100$, the 95% confidence interval is:

$$2.84 - (1.96)\frac{0.8}{\sqrt{100}} \quad \text{to} \quad 2.84 + (1.96)\frac{0.8}{\sqrt{100}}$$

$$\boxed{2.6832 \le \mu \le 2.9968}$$

- Since this interval does not contain the hypothesized mean (3.0), we reject the null hypothesis at $\alpha = .05$

One Tail Tests

- In many cases, the alternative hypothesis focuses on a particular direction

H_0: $\mu = 3$

H_1: $\mu < 3$

This is a lower tail test since the alternative hypothesis is focused on the lower tail below the mean of 3

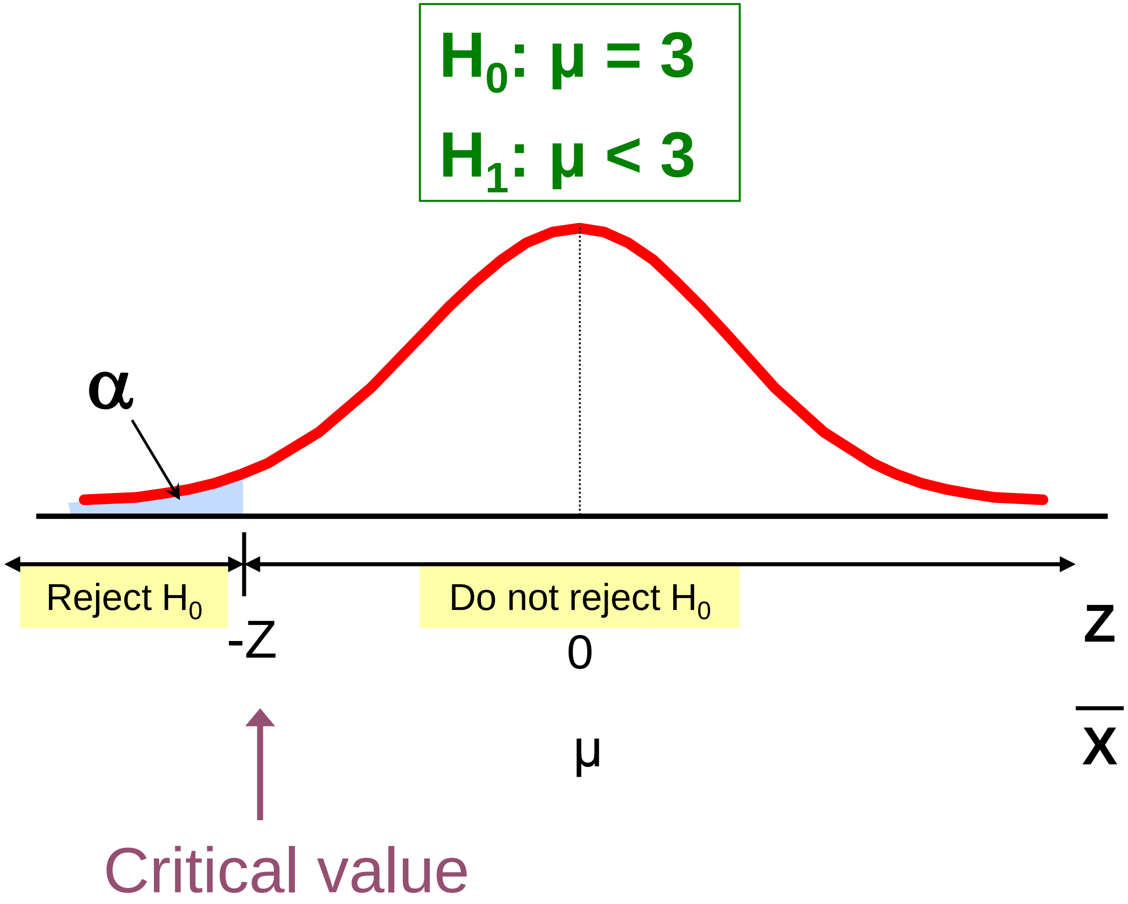
H_0: $\mu = 3$

H_1: $\mu > 3$

This is an upper tail test since the alternative hypothesis is focused on the upper tail above the mean of 3

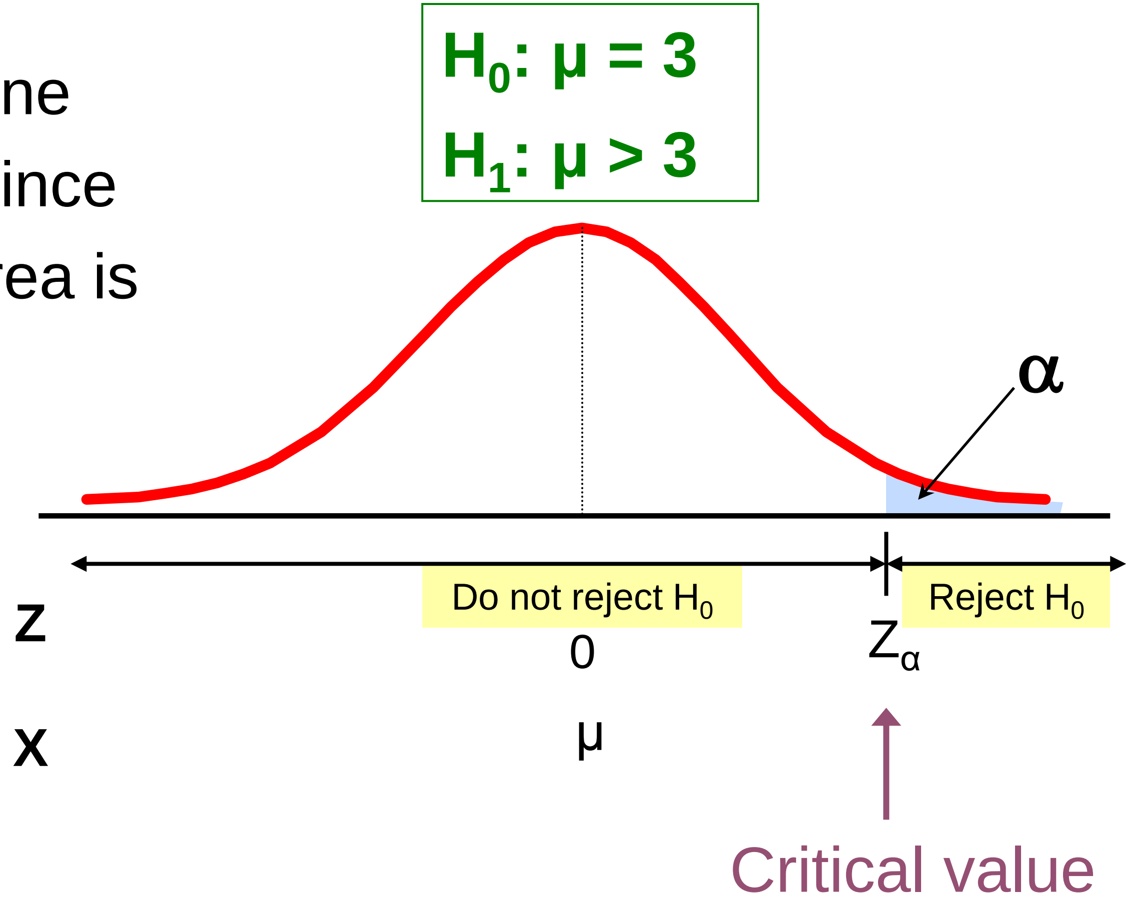
Lower Tail Tests

- There is only one critical value, since the rejection area is in only one tail



Upper Tail Tests

- There is only one critical value, since the rejection area is in only one tail



Example: Upper Tail Z Test for Mean (σ Known)

A phone industry manager thinks that customer monthly cell phone bill have increased, and now average over $52 per month. The company wishes to test this claim. (Assume $\sigma = 10$ is known)

Form hypothesis test:

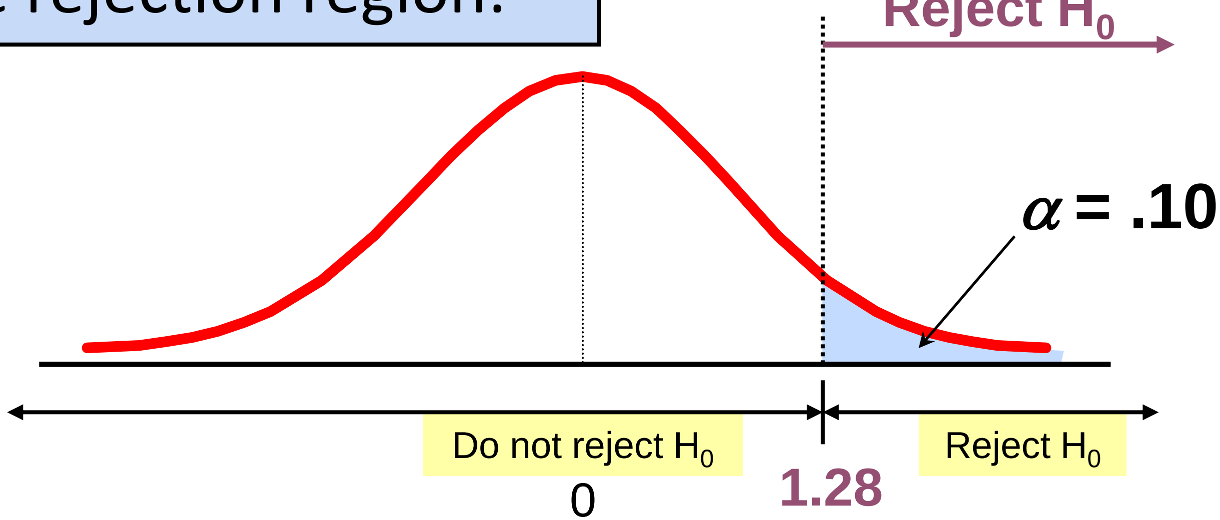
$H_0: \mu = 52$ the average is equivalent to $52 per month

$H_1: \mu > 52$ the average **is** greater than $52 per month (i.e., sufficient evidence exists to support the manager's claim)

Example: Find Rejection Region *(continued)*

- Suppose that α = .10 is chosen for this test

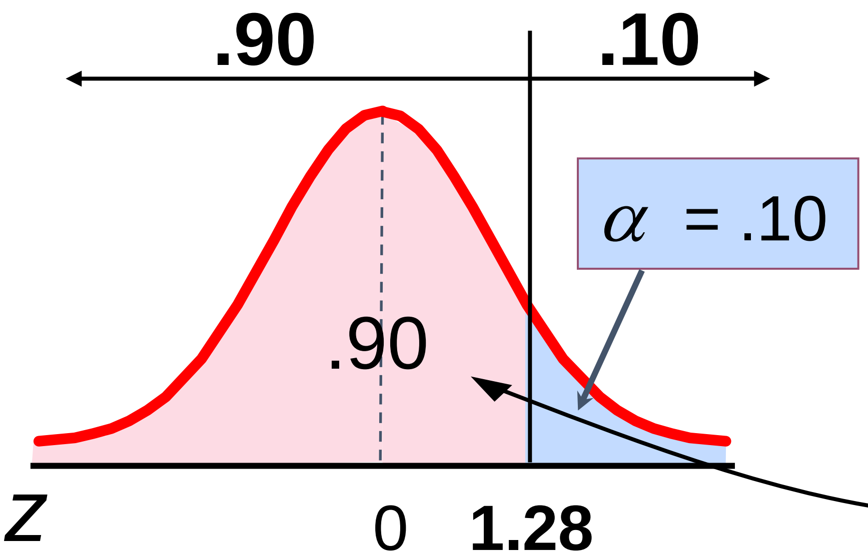
Find the rejection region:



Reject H_0 if $Z > 1.28$

Review:
Finding Critical Value - One Tail

What is Z given $\alpha = 0.10$?



Critical Value = 1.28

Standard Normal Distribution Table (Portion)

Z	.07	.08	.09
1.1	.8790	.8810	.8830
1.2	.8980	.8997	.9015
1.3	.9147	.9162	.9177

Example: Test Statistic

(continued)

Obtain sample evidence and compute the test statistic

Suppose a sample is taken with the following results:

$n = 64, \ \bar{X} = 53.1$ (σ=10 was assumed known)

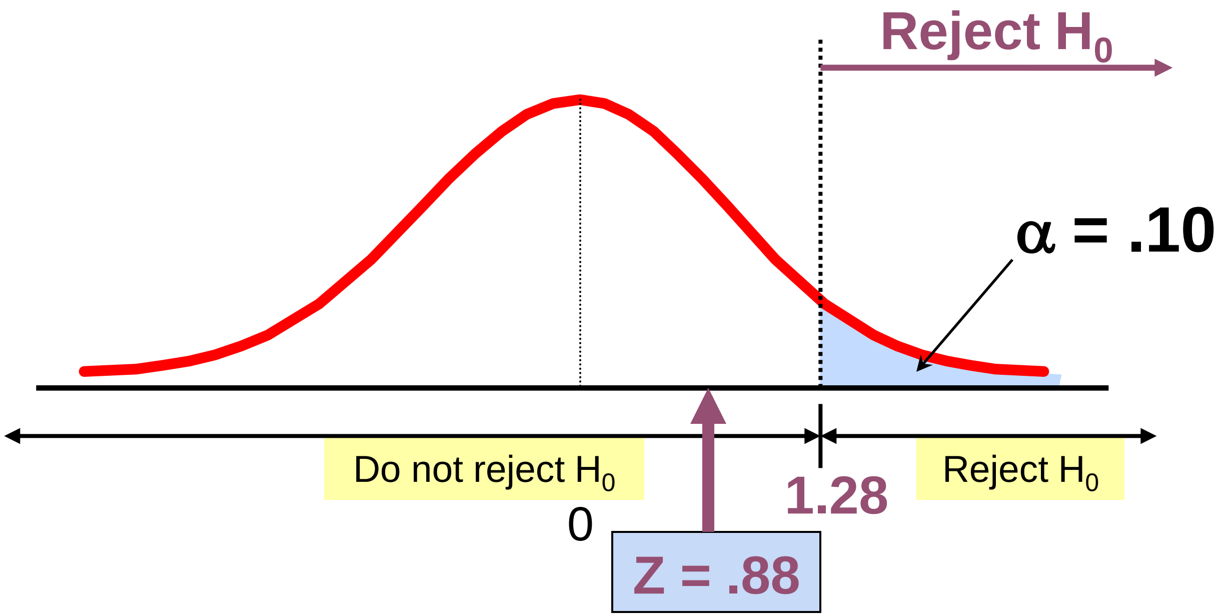
- Then the test statistic is:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{53.1 - 52}{\frac{10}{\sqrt{64}}} = 0.88$$

Example: Decision

(continued)

Reach a decision and interpret the result:



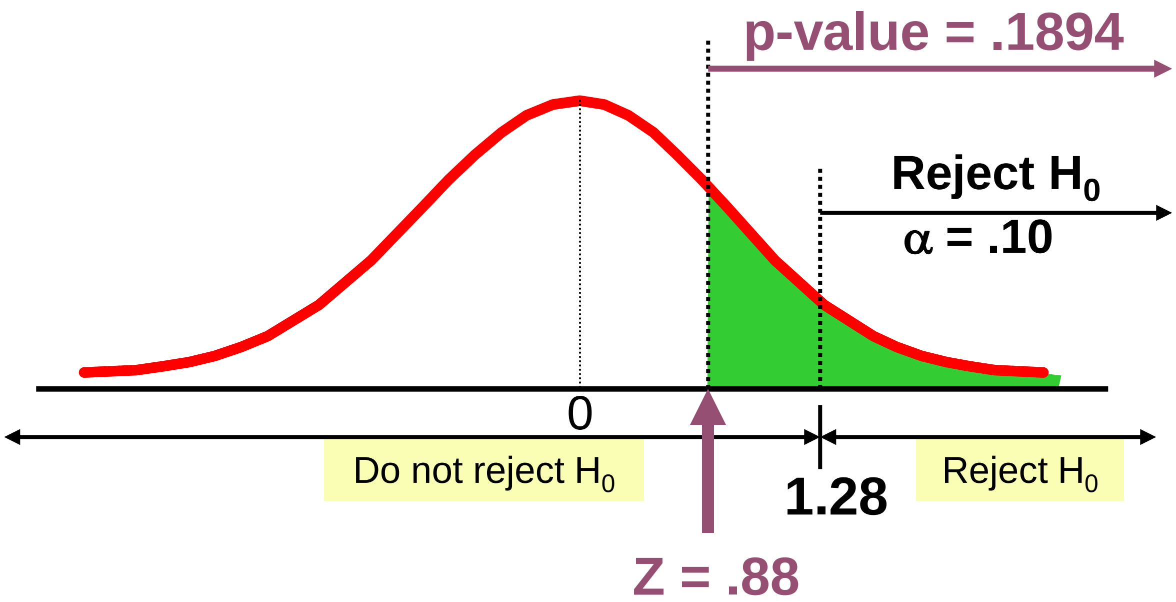
Do not reject H_0 since $Z = 0.88 \leq 1.28$

i.e.: there is not sufficient evidence that the mean bill is over \$52

p -Value Solution

(continued)

Calculate the p-value and compare to α



$$P(\bar{X} \geq 53.1 \mid \mu = 52.0)$$

$$= P\left(Z \geq \frac{53.1 - 52.0}{10/\sqrt{64}} \right)$$

$$= P(Z \geq 0.88) = 1 - .8106$$

$$= .1894$$

Do not reject H_0 since p-value = .1894 > α = .10

Z Test of Hypothesis for the Mean (σ UnKnown)

- Convert sample statistic ($\bar{X}$) to a t test statistic

Hypothesis Tests for μ

σ Known

σ Unknown

The test statistic is:

$$t_{n-1} = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$$

Example: Two-Tail Test (σ Unknown)

The average cost of a hotel room in New York is said to be \$168 per night. A random sample of 25 hotels resulted in $\bar{X}$ = \$172.50 and

S = \$15.40. Test at the

$\alpha = 0.05$ level.

(Assume the population distribution is normal)

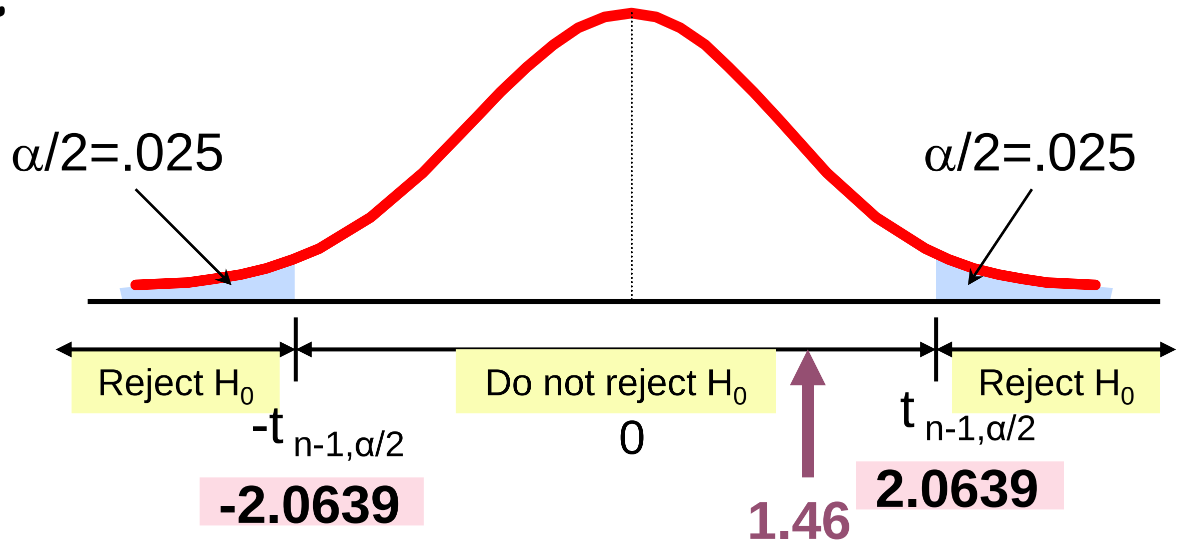
H_0: $\mu = 168$
H_1: $\mu \neq 168$

Example Solution: Two-Tail Test

H_0: $\mu = 168$
H_1: $\mu \neq 168$

- α = 0.05
- n = 25
- σ is unknown, so use a t statistic
- Critical Value:

$t_{24} = \pm\ 2.0639$



$$t_{n-1} = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} = \frac{172.50 - 168}{\frac{15.40}{\sqrt{25}}} = 1.46$$

Do not reject H_0: not sufficient evidence that true mean cost is different than \$168

Connection to Confidence Intervals

- For $\bar{X} = 172.5$, $S = 15.40$ and $n = 25$, the 95% confidence interval is:

$$\boxed{172.5 - (2.0639)\ 15.4/\sqrt{25}} \quad \text{to} \quad \boxed{172.5 + (2.0639)\ 15.4/\sqrt{25}}$$

$$\boxed{166.14 \le \mu \le 178.86}$$

- Since this interval contains the Hypothesized mean (168), we do not reject the null hypothesis at $\alpha = .05$

Hypothesis Tests for Proportions

- Involves categorical values
- Two possible outcomes
  - “Success” (possesses a certain characteristic)
  - “Failure” (does not possesses that characteristic)
- Fraction or proportion of the population in the “success” category is denoted by p

Proportions

(continued)

- Sample proportion in the success category is denoted by p_s

- $$p_s = \frac{X}{n} = \frac{\text{number of successes in sample}}{\text{sample size}}$$

- When both np and $n(1-p)$ are at least 5, p_s can be approximated by a normal distribution with mean and standard deviation

- $$\mu_{p_s} = p$$ $$\sigma_{p_s} = \sqrt{\frac{p(1-p)}{n}}$$

Hypothesis Tests for Proportions

- The sampling distribution of p_s is approximately normal, so the test statistic is a Z value:

$$Z = \frac{p_s - p}{\sqrt{\dfrac{p(1-p)}{n}}}$$

Hypothesis Tests for p

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

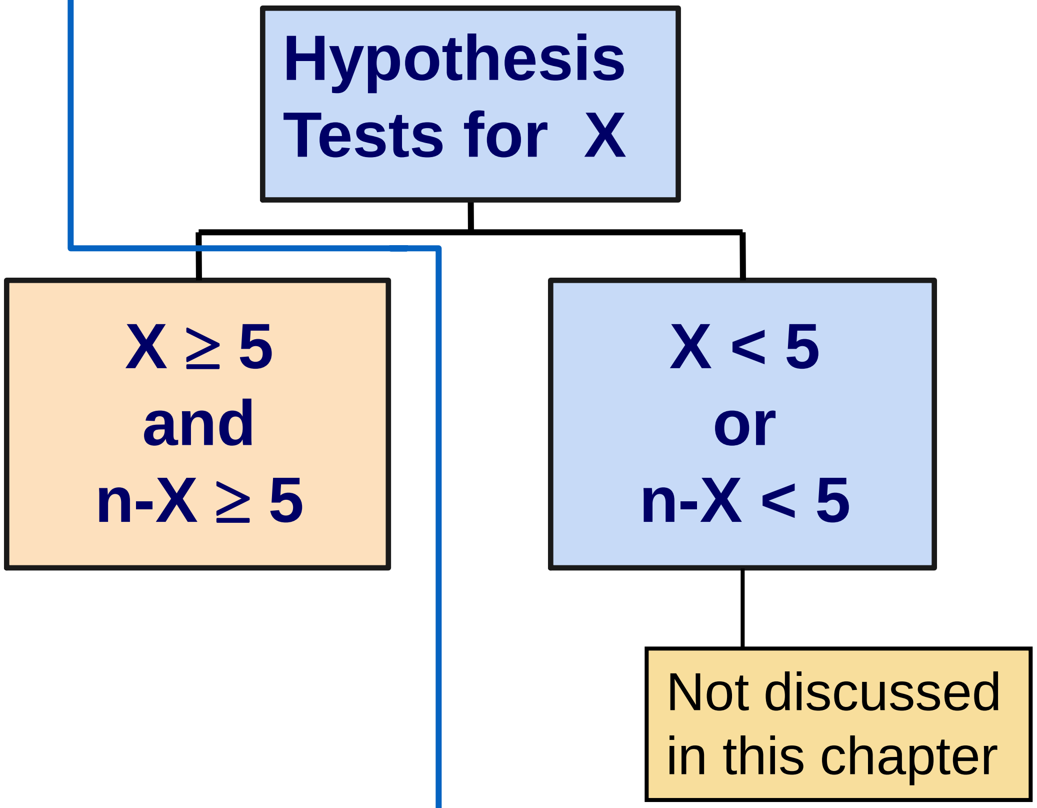
$np < 5$ or $n(1-p) < 5$

Not discussed in this chapter

Z Test for Proportion in Terms of Number of Successes

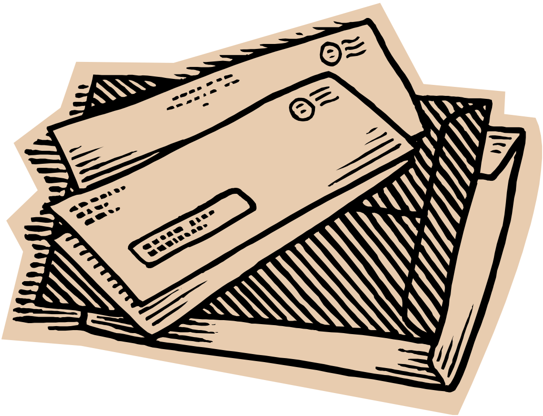
- An equivalent form to the last slide, but in terms of the number of successes, X:

$$Z = \frac{X - np}{\sqrt{np(1-p)}}$$



Example: Z Test for Proportion

A marketing company claims that it receives 8% responses from its mailing. To test this claim, a random sample of 500 were surveyed with 25 responses. Test at the α = .05 significance level.



Check:

$$n\,p = (500)(.08) = 40$$

$$n(1-p) = (500)(.92) = 460$$

✓

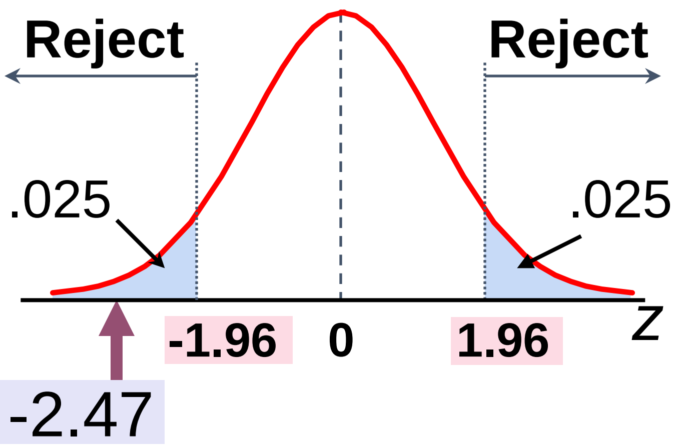
Z Test for Proportion: Solution

H_0: p = .08
H_1: p ≠ .08

α = .05

n = 500, p_s = .05

Critical Values: ± 1.96



Test Statistic:

$$Z = \frac{p_s - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{.05 - .08}{\sqrt{\frac{.08(1-.08)}{500}}} = -2.47$$

Decision:

Reject H_0 at α = .05

Conclusion:

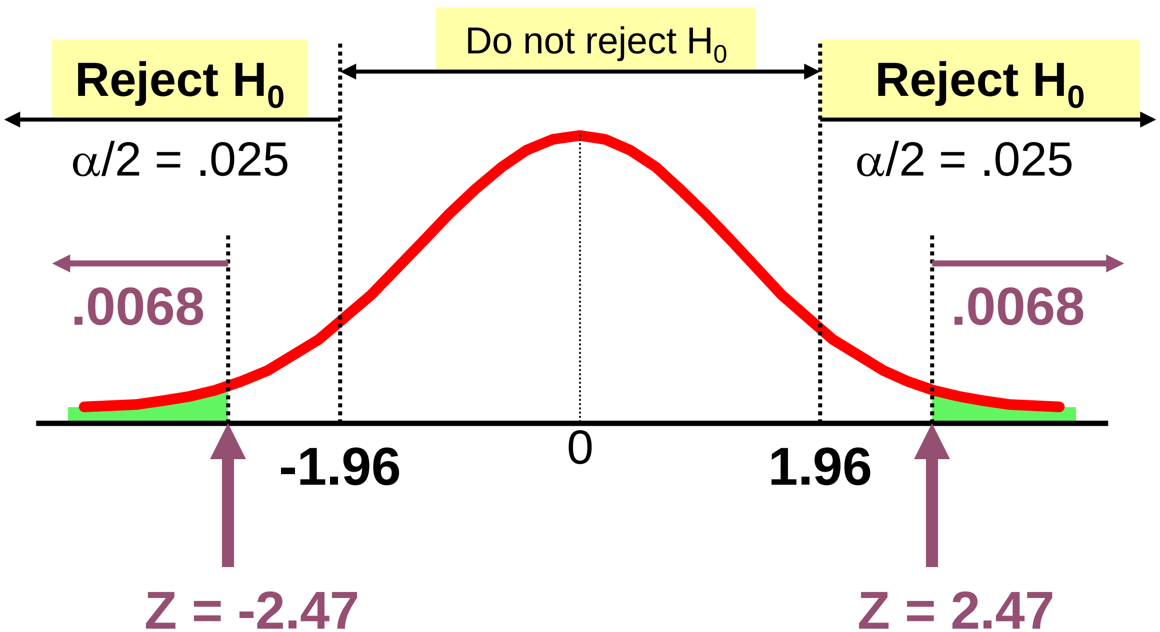
There is sufficient evidence to reject the company's claim of 8% response rate.

p-Value Solution

(continued)

Calculate the p-value and compare to α

(For a two sided test the p-value is always two sided)



p-value = .0136:

$$P(Z \le -2.47) + P(Z \ge 2.47)$$

$$= 2(.0068) = 0.0136$$

Reject H_0 since p-value = .0136 < α = .05

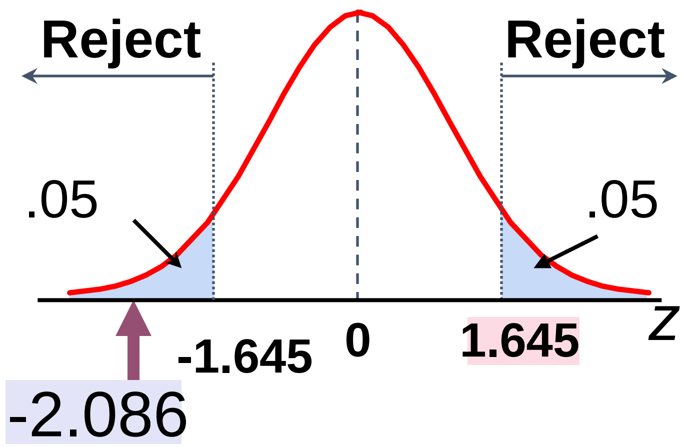
Other Example : Z Test for Proportion

H_0: p = .8
H_1: p ≠ .8

α = .1

n = 200, p_s = .741

Critical Values: ± 1.645



Test Statistic:

$$Z = \frac{p_s - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{.741 - .8}{\sqrt{\frac{.8(1-.8)}{200}}} = -2.086$$

Decision:

Reject H_0 at α = .1`

Conclusion:

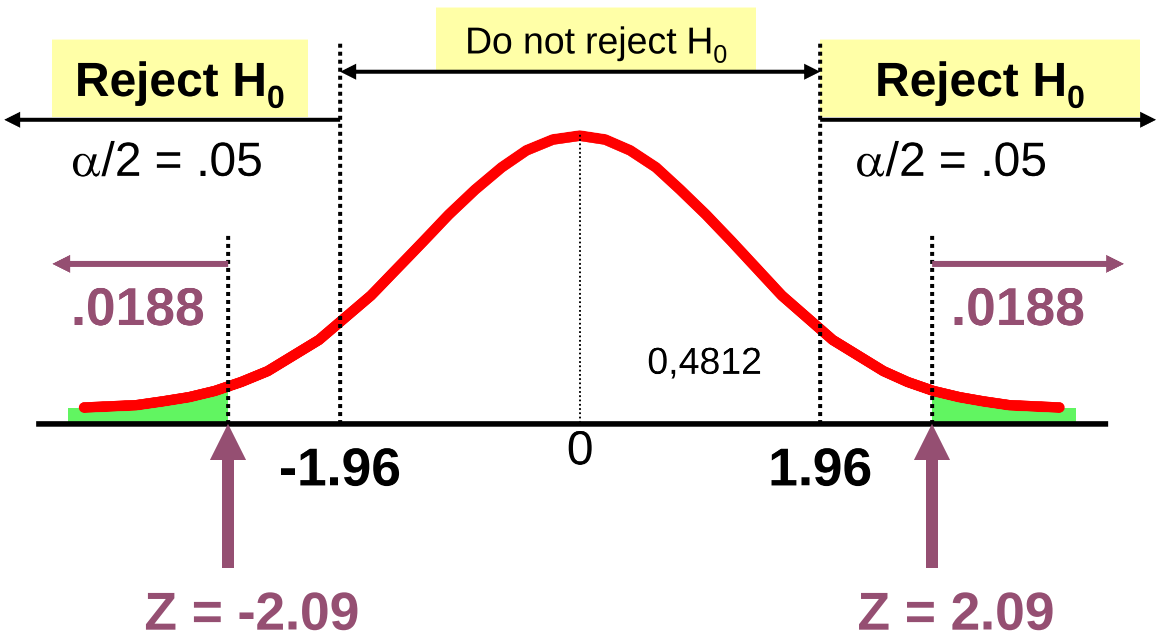
There is sufficient evidence to reject the company's claim of 80%

p-Value Solution

(continued)

Calculate the p-value and compare to α

(For a two sided test the p-value is always two sided)



p-value = .0376:

$$P(Z \le -2.09) + P(Z \ge 2.09)$$
$$= 2(.0188) = 0.0376$$

Reject H_0 since p-value = .0376 < α = .1

Chapter Summary

- Addressed hypothesis testing methodology
- Performed two tailed Z Test for the mean (σ known)
- Discussed critical value and p–value approaches to hypothesis testing
- Performed one-tail and two-tail tests . . .

Chapter Summary

(continued)

- Performed t test for the mean (σ unknown)
- Performed Z test for the proportion
- Discussed pitfalls and ethical issues

EXERCISES

Problem 1

The Daytona Beach Tourism Commission recently claimed that the average amount of money a typical college student spends per day during spring break is over $70. Based upon previous research, the population standard deviation is estimated to be $17.32. The Commission surveys 35 students and find that the mean spending is $67.57. Is there evidence that the average amount spent by students is less than $70 at α 5% ?

Problem 2

The manufacturer of a new product claims that his product will increase output per machine by 29 units per hour. A line manager adopts the product on 15 of his machines, and finds that the average increase was only 26 with a standard deviation of 6.2. Is there evidence to doubt the manufacturer's claim at α 5%?

EXERCISES

Problem 3

The supervisor of a production line believes that the average time to assemble an electronic component is 14 minutes. Assume that assembly time is normally distributed with a standard deviation of 3.4 minutes. The supervisor times the assembly of 14 components, and finds that the average time for completion was 11.6 minutes. Is there evidence that the average amount of time required to assemble a component is something other than 14 minutes at α 10% ?

Problem 4

The manufacturer of a certain chewing gum claims that four out of five dentists surveyed prefer their type of gum. You decide to test their claim. You find that in a sample of 200 doctors, 74.1% do actually prefer their gum. Is this evidence sufficient to doubt the manufacturer's claim α 10% ?

Hypothesis Testing for Differences/Changes in the Variance of a Population

The Chi-Square (χ^2) Test of Variance

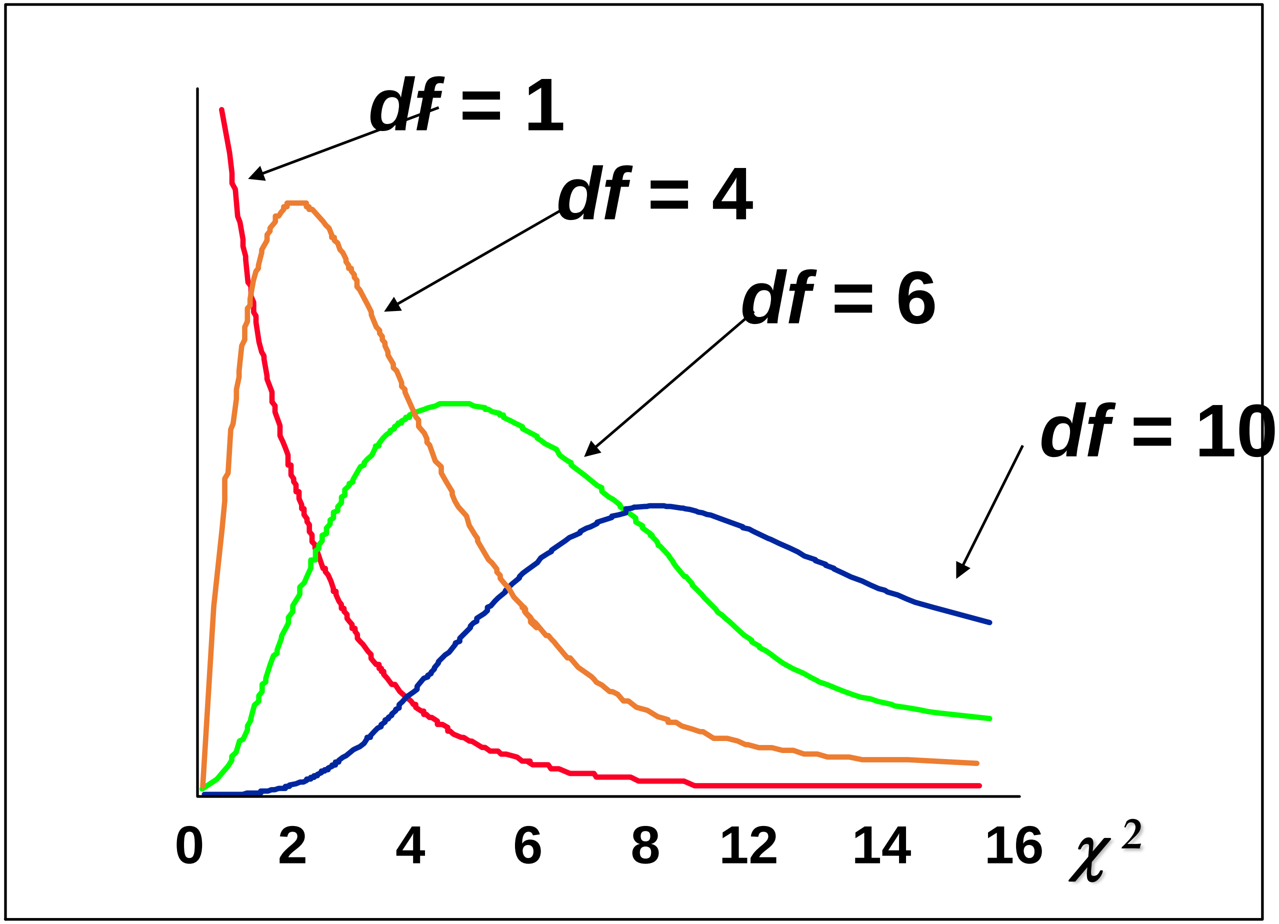
- Intended to determine whether a change has occurred in the dispersion of a population, as measured by the variance of the data; or
- Whether it is reasonable that a sample with a particular s^2 value could have been randomly drawn from a population with a hypothesized value of X

Hypothesis Testing for Differences/Changes in the Variance of a Population

The Chi-Square (χ^2) Test of Variance

- Employs the χ^2 family of distributions
- The specific distribution to be employed depends on the degrees of freedom (*df*) for the test
- The χ^2 distribution originates at 0, and is positively skewed, but becomes normal as *df* increases

The Chi-Square Family of Distributions



The Chi-Square (χ^2) Test of Variance

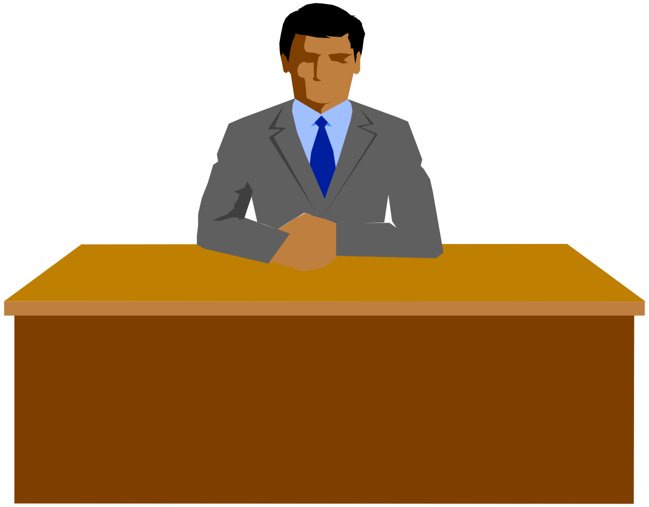
Underlying Assumptions

- The population is normally distributed with a mean of μ and a standard deviation of σ
- The sample was randomly drawn from the population associated with the hypothesis test / experiment / study

The Chi-Square (χ^2) Test of Variance

Sample Problem # 1

- A management professor has given careful thought to the design of examinations. In order for him to be reasonably certain that an exam does a good job of distinguishing the differences in achievement shown by the students, the standard deviation *cannot be too small*. On the other hand, if the standard deviation is too large, there will tend to be a lot of very low scores, which is bad for student morale. Past experience has led the professor to believe that a standard deviation of 13 points on a 100-pt exam indicates that the exam does a good job of balancing these two objectives.



The professor just gave an exam to his class of 31 students. The mean score was 72.7 and the standard deviation was 15.9. Does this exam meet his criterion At α 0.1 ?

The Chi-Square (χ^2) Test of Variance

Sample Problem # 1

Summary of Data:

$\sigma_{H_0} =$ 13 <---- hypothesized value of standard deviation

s = 15.9 <--- sample standard deviation

n = 31 <----- sample size

1. H_0: $\sigma^2 = 169$

   H_1: $\sigma^2 \neq 169$

2. $\alpha = 0.10$

The Chi-Square (χ^2) Test of Variance

Sample Problem # 1

- Step 3: State the Associated Test Statistic (χ^2)

$$\chi^2 = \frac{(n-1)\,s^2}{\sigma^2}$$

- Step 4: $\chi^2 = \chi^2$ (n-1) *df* if H_0 is True

The Chi-Square (χ^2) Test of Variance

Sample Problem # 1

Step 5: Determine the critical values of χ^2

$$\chi^2_{1-\alpha/2,\, df} = \chi^2_{1-0.10/2,\, n-1}$$

$$= \chi^2_{0.95,\, 31-1}$$

$$= \chi^2_{0.95,\, 30}$$

$$= 18.493$$

$$\chi^2_{\alpha/2,\, df} = \chi^2_{0.10/2,\, n-1}$$

$$= \chi^2_{0.05,\, 31-1}$$

$$= \chi^2_{0.05,\, 30}$$

$$= 43.773$$

The Chi-Square (χ^2) Test of Variance

Sample Problem # 1

Step 5: Determine the critical values of χ^2

$$\chi^2_{1-\alpha/2,\ df} = \chi^2_{1-0.10/2,\ n-1}$$

$$= \chi^2_{0.95,\ 31-1}$$

$$= \chi^2_{0.95,\ 30}$$

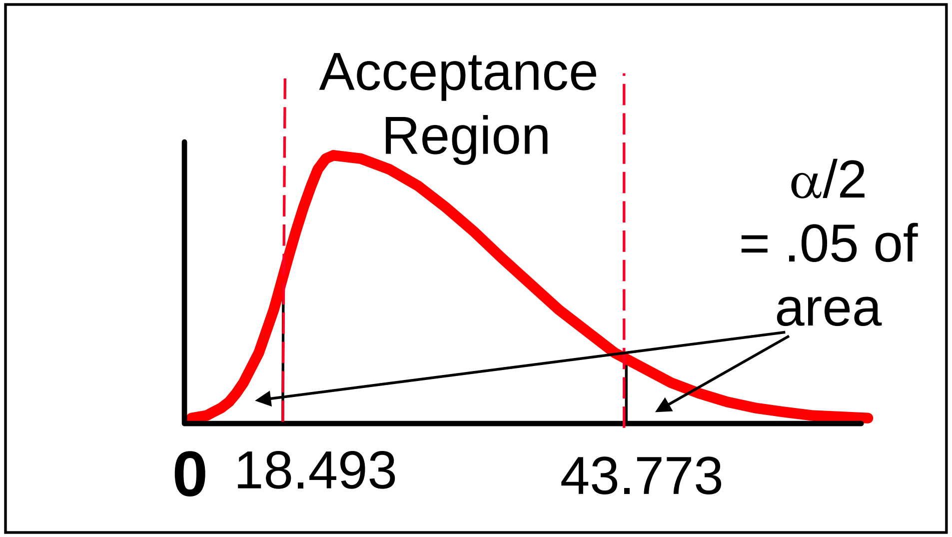
$$= 18.493$$

$$\chi^2_{\alpha/2,\ df} = \chi^2_{0.10/2,\ n-1}$$

$$= \chi^2_{0.05,\ 31-1}$$

$$= \chi^2_{0.05,\ 30}$$

$$= 43.773$$



The Chi-Square (χ^2) Test of Variance

Sample Problem # 1

- Step 6: Calculate the Value of the Test Statistic (χ^2)

$$\boxed{\chi^2 = \frac{(n-1)\, s^2}{\sigma^2}}$$

$$= \frac{(31-1)\, 15.9^2}{13^2}$$

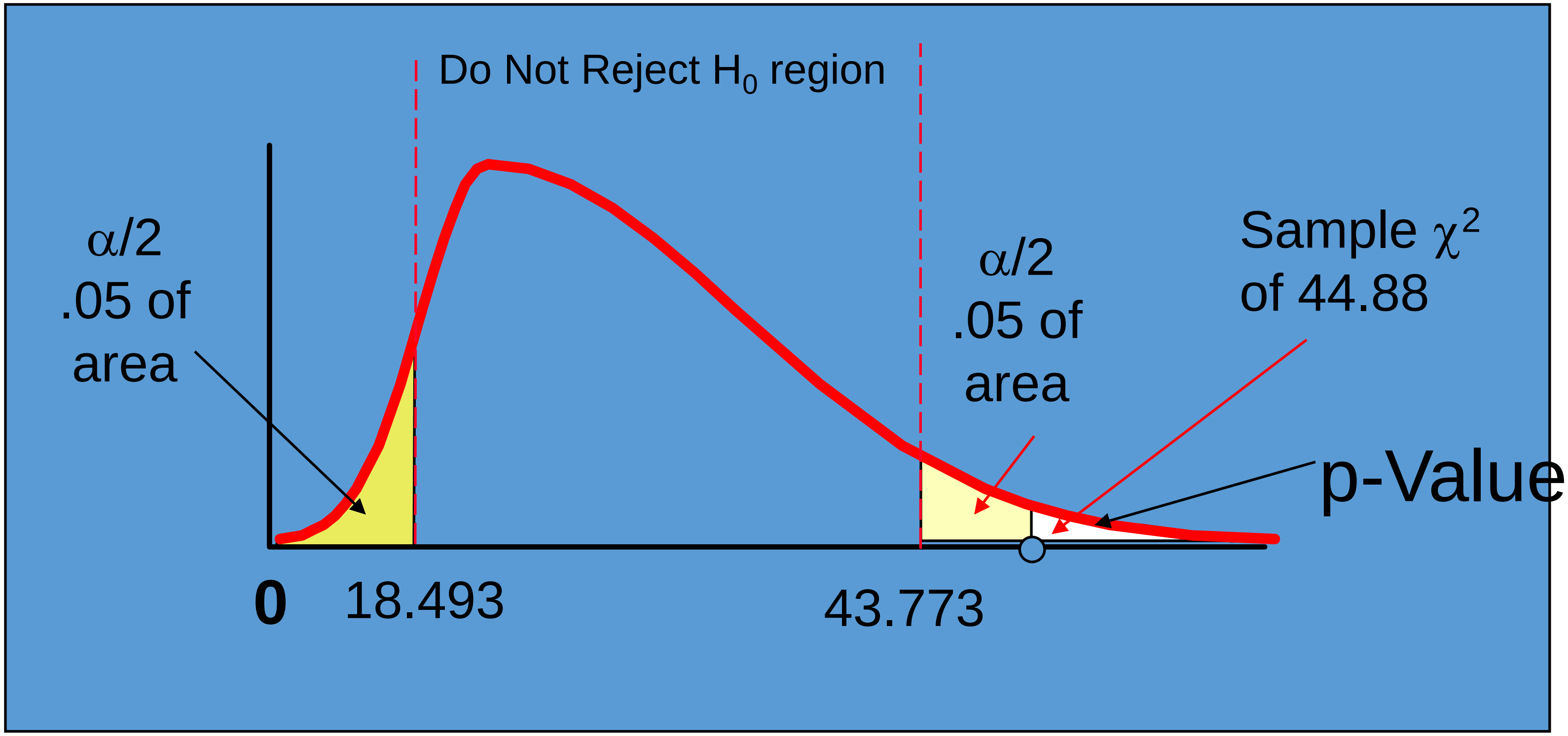
$$= \frac{(30)(15.9)^2}{(13)^2}$$

$$= 44.88$$

The Chi-Square (χ^2) Test of Variance

- Step 7 :

Sample Problem # 1



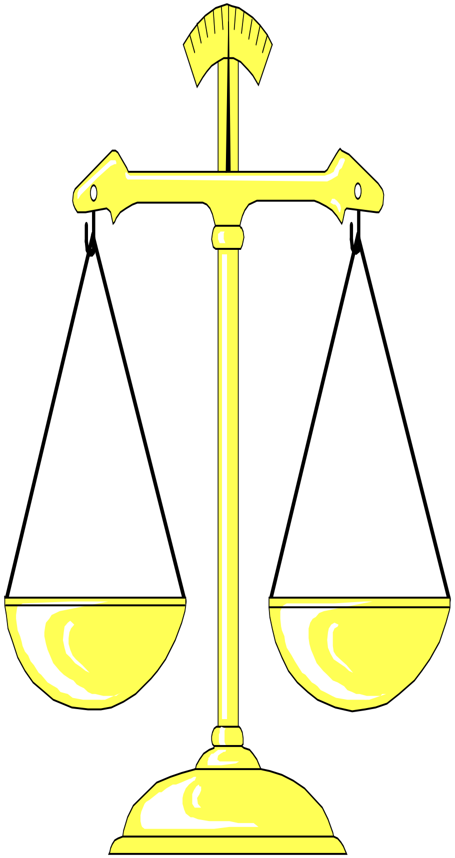
DECISION : reject Ho

Conclusion : The exam didn’t meet the professor’ criterion

The Chi-Square (χ^2) Test of Variance
Sample Problem # 2 : Self-Review

- Precision Analytics manufactures a wide line of precision instruments and has a fine reputation in the field for quality of its instruments. It will not release an analytic balance for sale unless that balance shows a variability significantly below one microgram when weighing quantities of about 500 grams.

- A new balance has just been delivered to the quality control division from the production line. It is tested to weigh the same 500-gm standard weight 30 different times. The sample standard deviation turns out to be 0.73 microgram. Should the balance be sold at α 0.1 ?



Thank You

"We trust in GOD, all others must bring data"