Chapter 11.1 - Introduction

Hypothesis Testing Procedure

1. Formulate Ho and Ha

2. Design an Experiment
Take sample and find it's proportion
If Ho were true, what would you expect the distribution of samples to look like?

3. Formulate the Decision Rule
based on how willing you are to be wrong.
(α = 1% or α = 5%)

4. Do the Experiment – determine the test statistic

5. Find the p-value
That is the probability that you could have gotten this result by chance.

6. Determine the Conclusion

" Statistically Significant or
Not Statistically Significant,
That is the Question."

Chapter 11.2 – Population Proportion

Hypothesis Testing Procedure
for Population proportion

1. Formulate Ho and Ha

2. Determine if the distribution of sample proportions is normally distributed.

Are both np and n(1-p) > 5 ?

3. Calculate the Standard Error of the Proportion

$$SEP = \sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

4. From the sample calculate the test statistic (z)

5. From the test statistic calculate the p-value.

6. Determine the conclusion. (Reject Ho ???)

" Statistically Significant or
Not Statistically Significant,
That is the Question."

Chapter 11.2 – Population Proportion

According to a recent census report 50% of U.S. families earn more than $20,000 per year. A sociologist from a university believes this percent is too low.

He randomly selects 100 families and determines that 64 have incomes of more than $20,000. Use α = 1%.

a) State the null and alternative hypotheses.

Ho: $p = 0.50$
Ha: $p > 0.50$

Is the distribution of sample proportions normally distributed?
$np = (100)(0.5) = 50 > 5$ ✔
$n(1-p) = (100)(0.5) = 50 > 5$ ✔

Chapter 11.2 – Population Proportion

According to a recent census report 50% of U.S. families earn more than $20,000 per year. A sociologist from a university believes this percent is too low.

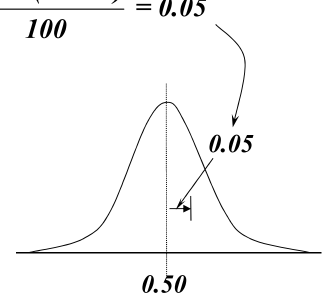
He randomly selects 100 families and determines that 64 have incomes of more than $20,000. Use α = 1%.

a) State the null and alternative hypotheses.

Ho: $p = 0.50$
Ha: $p > 0.50$

b) Sketch the sampling distribution of the mean (or proportion) . Label the diagram with the mean, std error, test statistic and the p-value.

$$SEP = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.50 \times (1-0.05)}{100}} = 0.05$$



Chapter 11.2 – Population Proportion

According to a recent census report 50% of U.S. families earn more than \$20,000 per year. A sociologist from a university believes this percent is too low.

He randomly selects 100 families and determines that 64 have incomes of more than \$20,000. Use $\alpha = 1\%$.

a) State the null and alternative hypotheses.

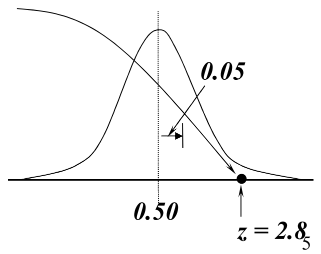
Ho: p = 0.50
Ha: p > 0.50

b) Sketch the sampling distribution of the mean (or proportion) . Label the diagram with the mean, std error, test statistic and the p-value.

c) Determine the test statistic from the data

$$\hat{p} = 64/100 = 0.64$$

$$z = \frac{0.64 - 0.50}{0.05} = 2.8$$



Chapter 11.2 – Population Proportion

According to a recent census report 50% of U.S. families earn more than \$20,000 per year. A sociologist from a university believes this percent is too low.

He randomly selects 100 families and determines that 64 have incomes of more than \$20,000. Use $\alpha = 1\%$.

a) State the null and alternative hypotheses.

Ho: p = 0.50
Ha: p > 0.50

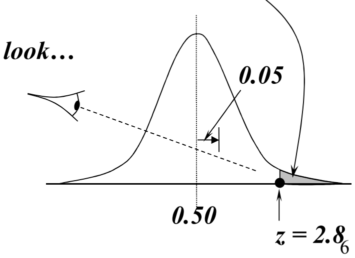
b) Sketch the sampling distribution of the mean (or proportion) . Label the diagram with the mean, std error, test statistic and the p-value.

c) Determine the test statistic from the data

d) Determine the p-value from the test statistic

p -value = **normalcdf (2.8, E99) = 0.00256**

Now let's take a closer look…



Chapter 11.2 – Population Proportion

According to a recent census report 50% of U.S. families earn more than \$20,000 per year. A sociologist from a university believes this percent is too low.

He randomly selects 100 families and determines that 64 have incomes of more than \$20,000. Use $\alpha = 1\%$.

a) State the null and alternative hypotheses.

Ho: p = 0.50
Ha: p > 0.50

b) Sketch the sampling distribution of the mean (or proportion) . Label the diagram with the mean, std error, test statistic and the p-value.

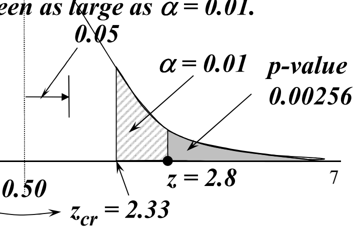
c) Determine the test statistic from the data

d) Determine the p-value from the test statistic

e) Recall $\alpha = 0.01$ and since the p-value < α,
Reject Ho.

Note: The p-value was far smaller than we needed to reject Ho. It could have been as large as $\alpha = 0.01$.

The test statistic could then have been invNormal(0.99)= 2.33



Chapter 11.2 – Population Proportion

According to a recent census report 50% of U.S. families earn more than \$20,000 per year. A sociologist from a university believes this percent is too low.

He randomly selects 100 families and determines that 64 have incomes of more than \$20,000. Use $\alpha = 1\%$.

a) State the null and alternative hypotheses.

Ho: p = 0.50
Ha: p > 0.50

b) Sketch the sampling distribution of the mean (or proportion) . Label the diagram with the mean, std error, test statistic and the p-value.

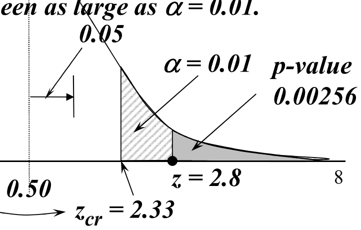
c) Determine the test statistic from the data

d) Determine the p-value from the test statistic

e) Recall $\alpha = 0.01$ and since the p-value < α,
Reject Ho.

Note: The p-value was far smaller than we needed to reject Ho. It could have been as large as $\alpha = 0.01$.

This is called the Critical Value and can also be used as the test criterion.



Chapter 11.2 – Population Proportion

Let's summarize what we did…

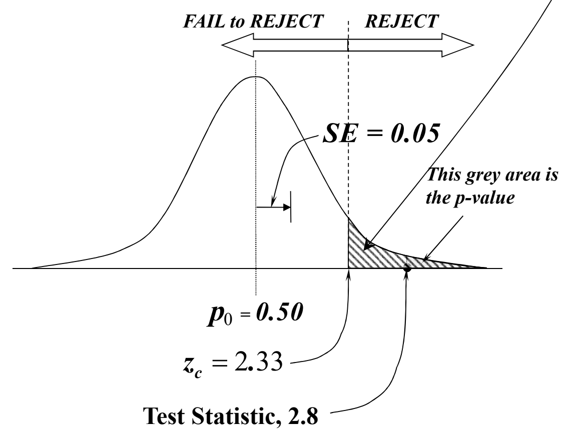
Key Data H_o: $p = 0.50$ H_a: $p > 0.50$

$n = 100$ $\alpha = 1\%$

$$\text{Std Error} = \sqrt{\frac{0.50(1-0.50)}{100}} = 0.05$$

$$\text{Test Statistic: } z = \frac{0.64-0.50}{0.05} = 2.8$$

p - value = normalcdf(2.8, E99) = 0.0026



Chapter 11.2 – Population Proportion

An independent research group is interested in showing that the percent of babies delivered by Cesarean Section is increasing. Last year, 20% of the babies born were delivered by Cesarean Section. The research group randomly inspects the medical records of 100 recent births and finds that 25 of the births were by Cesarean Section.

Can the research group conclude that the percent of births by Cesarean section has increased?

Use $\alpha = 5\%$

Key Data

Ho: p = 0.20

Ha: p > 0.20 ⟶ *Directional test*

N = 100

$\alpha = 5\%$

$$SEP = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.20(1 - 0.20)}{100}} = 0.04$$

$$\hat{p} = \frac{25}{100} = 0.25$$

Chapter 11.2 – Population Proportion

Key Data

Ho: p = 0.20

Ha: p > 0.20

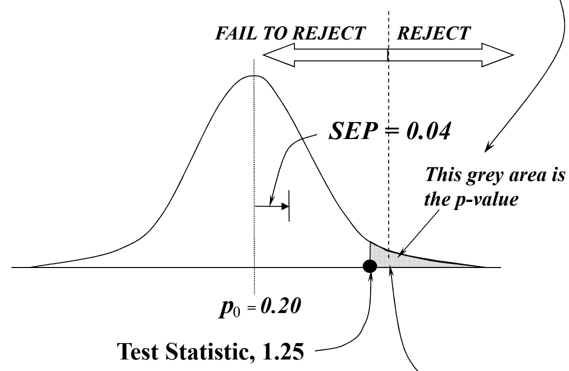
n = 100 $\alpha = 5\%$

Sample Data yields $\hat{p} = 25/100 = 0.25$

$$\text{Test Statistic} = z = \frac{0.25 - 0.20}{0.04} = 1.25$$

p-value = normalcdf(1.25, E99) = 0.106

p-value > α **→ Fail to Reject**



For an $\alpha = 5\%$**, the critical z score (**z_c**) is 1.645**

Chapter 11.2 – Population Proportion

College Degree

Last year, 75% of adults within the USA stated that they believed a college degree was very important. This year a national poll randomly selected 1020 American Adults and found 794 adults believed that a college degree is very important.

Can you conclude that the percent of adults within the USA who believe that a college degree is very important has significantly increased this year?

Perform a hypothesis test using an α **of 5%.**

Ho? p = 0.75

Ha? p > 0.75

SEP? $SEP = \sqrt{\frac{0.75 \times 0.25}{1020}} = 0.01356$

Data? → Test Statistic? $\hat{p} = \frac{794}{1020} = 0.7784$

$$z = \frac{0.7784 - 0.75}{0.01356} = 2.094$$

p-value? normalcdf(2.094, E99) = 0.0181

Conclusion? p-value < α **→ REJECT**

Chapter 11.2 – Population Proportion

Unemployment

According to the latest figures unemployment is at 8.7%. The Congressman from the third district believes that the unemployment rate is lower in his district.

To test his belief he interviews 200 residents of his district and finds 9 of them to be unemployed. Is the Congressman's belief correct?

Reject if $z < z_c = -1.645$

(What α is he using?)

Key Data

Ho: $\mu = 0.087$

Ha: $\mu < 0.087$

N = 200 $\alpha = ?$

Sample Data yields $\hat{p} = 9/200 = 0.045$

Chapter 11.2 – Population Proportion

Key Data

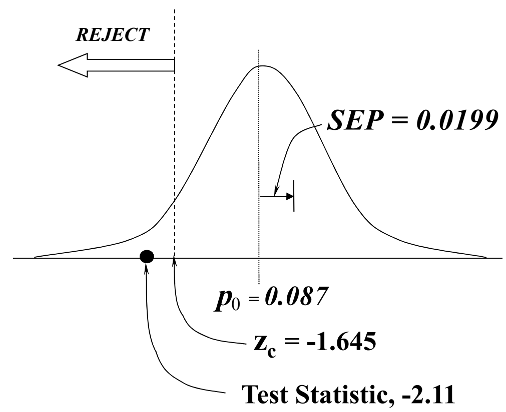
Ho: $\mu = 0.087$

Ha: $\mu < 0.087$

n = 200

Sample Data yields $\hat{p} = 9/200 = 0.045$

Decision: Reject Ho if $z < -1.645$



Any test statistic must be further from the mean than z_c in order to Reject. Therefore z_c marks the edge of the α area. Thus α = normalcdf(-E99, -1.645) = .05

Chapter 10.2 – Population Proportion

Divorce

According to NCHS (National Center for Health Statistics), the national divorce rate in 2004 was 3.6 per 1000 population (i.e. .0036). A sample of 20,000 people in New York State that year showed a divorce rate of 3.0 per 1000 population.

Can New York State claim to have a divorce rate significantly lower than the national average at an α of 5%?

Mean or Proportion?,

Ho?

Ha?

Decision Rule?

Test Statistic?

Conclusion?

Chapter 11.2 – Population Proportion

Racing

The publisher of Hot Spark, a racing car magazine, claims that 70% its readership is composed of men between men between 21 and 35 years old. An Advertising agency manager disputes that claim feeling that it is too high.

The manager surveys 100 men who subscribe to the magazine and finds that only 64 of them are between 21 and 35 years old. Can the Ad agency manager justify his position with at an α of 5%.

Ho?

Ha?

Check np and n(1-p)

SEP?

Data? → Test Statistic?

p-value?

Conclusion?

Chapter 11.2 – Population Proportion

<u>Claim: 70% of readers are 21-35 years old</u>

- **Manager believes claim is <u>too high</u>.**
- **Decides on an α of 5%.**
- **Surveys 100 readers and finds 64 are in the 21-35 age group.**

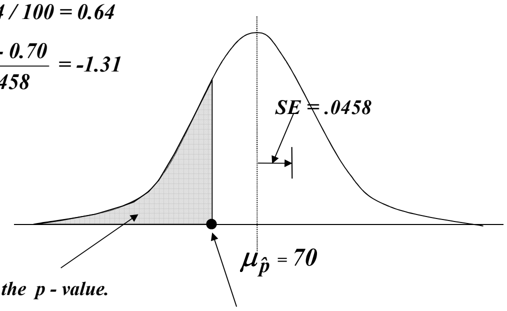
<u>Hypotheses:</u>
Ho: Prop of readers, $P = 70\%$
Ha: Prop of readers, $P < 70\%$

$\mu_p = p = 0.70$

$$SE = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{(0.70)(1 - 0.70)}{100}} = 0.0458$$

$\hat{p} = 64 / 100 = 0.64$

$$z = \frac{0.64 - 0.70}{0.0458} = -1.31$$



p = normalcdf (-E99, -1.31)
p = 0.095 = 9.5%
<u>CONCLUSION: FAIL to REJECT Ho</u>

Chapter 11.2 – Population Proportion

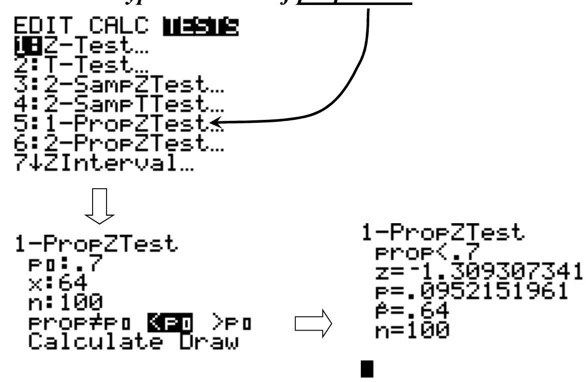
<u>Claim: 70% of readers are 21-35 years old</u>

- **Manager believes claim is <u>too high</u>.**
- **Decides on an α of 5%.**
- **Surveys 100 readers and finds 64 are in the 21-35 age group.**

<u>Hypotheses:</u>
Ho: Prop of readers, $P = 70\%$
Ha: Prop of readers, $P < 70\%$

Using the Calculator

This is a Hypothesis Test of <u>proportion</u> .



Since p = .095 which is greater than α = .05, we fail to reject.

Chapter 11.3 – Population Mean

<u>Hypothesis Testing Procedure</u>
for a Population Mean

1. ***Formulate Ho and Ha***
2. ***Is the population Normally distributed?***
   If not, is n > 30?
3. ***Calculate the Standard Error of the Mean***

$$SEM = \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

4. ***From the sample, calculate the <u>test statistic</u> (z)***
5. ***From the test statistic, calculate the p-value***
6. ***Determine the conclusion***

If the p-value < α : REJECT Ho

Else: Fail to REJECT Ho

Chapter 11.3 – Population Mean

An unemployment agency claims that the mean age of recipients of unemployment benefits is 37 years with a standard deviation of 5 years. A trade union association believes this claim is too large.

The association randomly interviews 400 recipients of unemployment benefits and obtains a mean age of 36 years.

Can the association reject the unemployment agency's claim at an α of 5%?

Use the following model to perform this hypothesis test, and:

a) State the null and alternative hypotheses.

b) Draw the sampling distribution of the mean labeling the mean, Standard Error and test statistic (z).

c) Determine the p-value for the test statistic.

d) State your conclusion.

Chapter 11.3 – Population Mean

An unemployment agency claims that the mean age of recipients of unemployment benefits is 37 years with a standard deviation of 5 years. A trade union association believes this claim too large.

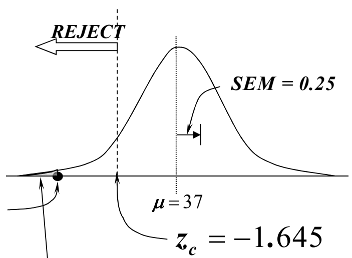
The association randomly interviews 400 recipients of unemployment benefits and obtains a mean age of 36 years.

Can the association reject the unemployment agency's claim at an α of 5%?

Ho: $\mu = 37$
Ha: $\mu < 37$
Distribution of Sample Means Normal?
Yes. (n > 30) n = 400

$SEM = 5 / \sqrt{400} = 0.25$

$z = \dfrac{36 - 37}{0.25} = -4$



Test Statistic = z = -4

p = normalcdf(- E99, - 4) = 3.17E-5

p - value < α, REJECT Ho

Chapter 11.3 – Population Mean

A manufacturer of light bulbs claims that his bulbs have a mean lifetime of 1800 hrs. with a standard deviation of 100 hrs. A consumer agency feels his bulbs do not last that long so they purchase a random sample of 400 bulbs and run them until they burn out.

The mean lifetime of the sample was 1788 hours.

Can the consumer agency reject the manufacturer's claim at an α of 5%?

Mean or Proportion?,
Ho? Ha?
Decision Rule?
Test Statistic?
Conclusion?

Using the Calculator for Hypothesis Testing

For the Normal Distribution:

STAT $\Rightarrow$ TESTS $\Rightarrow$ Z-TEST

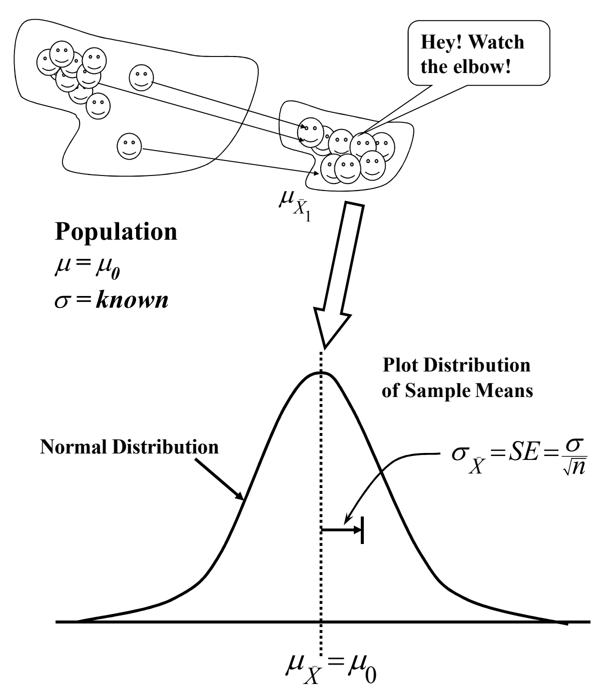
When you know the stats

```
Z-Test
 Inpt:Data Stats
 μ0:1800
 σ:100
 x̄:1788
 n:400
 μ:≠μ0 <μ0 >μ0
 Calculate Draw
```

When you have the data

```
Z-Test
 Inpt:Data Stats
 μ0:1800
 σ:100
 List:L1
 Freq:1
 μ:≠μ0 <μ0 >μ0
 Calculate Draw
```


Thus far the population standard deviation has been KNOWN!



What if we do NOT KNOW the population standard deviation?

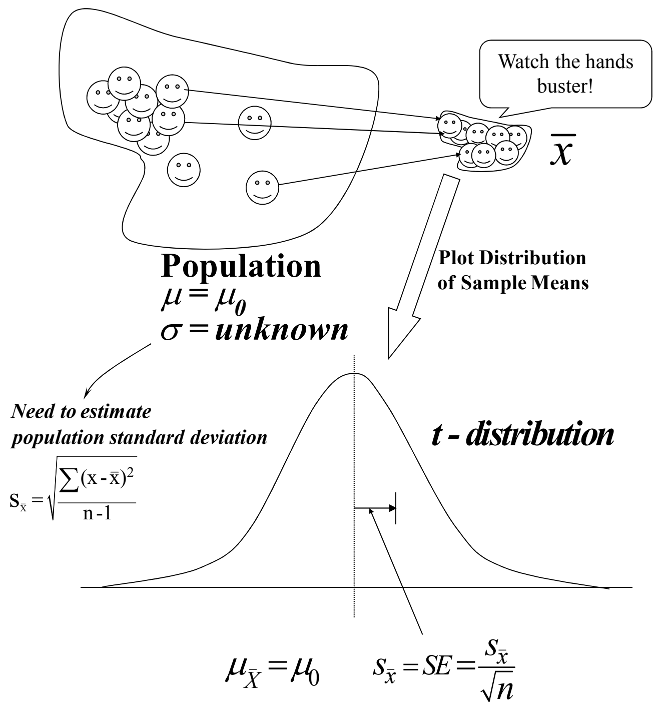
The sample means will NOT be Normally Distributed.

They will be distributed as a t-distribution

Courtesy of William Gosset of the Guinness Brewery.

Chapter 11.4 – the *t* - Distribution

For σ unknown…



Chapter 11.4 – the *t* - Distribution

Properties of the t-distribution

…bell shaped

…symmetric about the mean

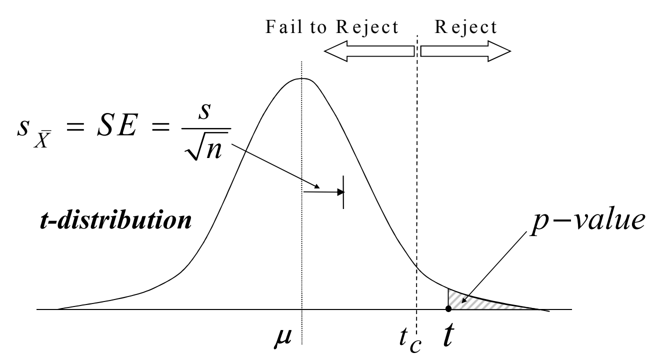
…is different for each sample size (n)

(use degrees of freedom, df = n-1)

…approximates the normal distribution for df > 30

Chapter 11.4 – the *t* - Distribution

Decide Ho and Ha
Draw the t-distribution of sample means
Determine the Standard Error of the Mean
Determine the sample mean, $\bar{x}$
Determine the test statistic, t → $t = \frac{\bar{x} - \mu}{SE}$
Determine the p-value from the test statistic, t
… you will need the degrees of freedom (n-1)
Reject Ho if the p-value is less than < a
α also determines the critical value, t_c



Chapter 11 – Page 604 – Prob 94

Presently, the mean life expectancy of a rare strain of bacteria is 12 hours. A scientist claims she has developed a medium which will increase the mean life of the bacteria.

The scientist tests 16 cultures of the newly treated bacteria and finds they have a mean life of 13 hours with s = 1 hour. Do these results show that the medium is effective in increasing the bacteria's life expectancy?
(Use α = 1%)

Chapter 11.4 – the *t* - Distribution

Hypothesis Testing Procedure
For a population Mean with σ unknown

1. *Formulate Ho and Ha*
2. *Is the population normally distributed?*
   Continue if YES or n > 30
3. *Calculate the Standard Error of the Mean*

$$SEM = S_{\bar{x}} = \frac{S}{\sqrt{n}}$$

4. *From the sample, calculate the test statistic (t).*
5. *From the test statistic, calculate the p-value.*
6. *Determine the Conclusion.*

Chapter 11.4 – the *t* - Distribution

Hypothesis Testing Procedure
For a population Mean with σ unknown

1. *Formulate Ho and Ha*
   $Ho: \mu = 12$ *hrs*
   $Ha: \mu > 12$ *hrs*
2. *Is the population Normally distributed?*
   YES, but σ is unknown.
   So, sample means are t-distributed.
3. *Calculate the Standard Error of the Mean*

$$SEM = S_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{1}{\sqrt{16}} = 0.25$$

4. *From the sample, calculate the test statistic (t)*

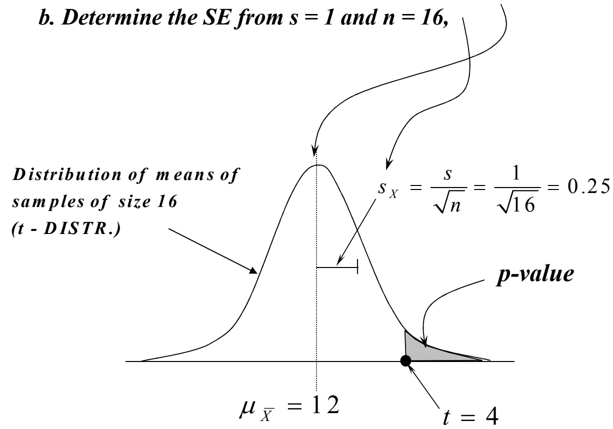
$$t = \frac{x - \mu}{s_{\bar{x}}} = \frac{13 - 12}{0.25} = 4.0$$

5. *From the test statistic, calculate the p-value*
   p-value = tcdf(4.0, E99, 15) = 5.79E-4
6. *Determine the Conclusion*
   Since p-value < α = 1%, Reject Ho

Chapter 11 Page 537/Prob 94

a. Decide Ho and Ha
Ho: μ = 12 hours
Ha: μ > 12 hours
Since we are not told σ, the distribution is t-distributed

b. Determine the SE from s = 1 and n = 16,



Determine the test statistic, t, from the sample mean

$$\bar{x} = 13 \quad \text{then } t = \frac{\bar{x} - \mu}{SE} = \frac{13 - 12}{0.25} = 4$$

Determine the p-value

p-value = tcdf (t1, t2, df) = tcdf (4, E99, 15) = 0.000580)

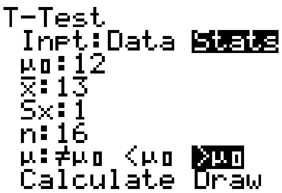
Thus, the p-value < α and we REJECT Ho

Using the Calculator for Hypothesis Testing

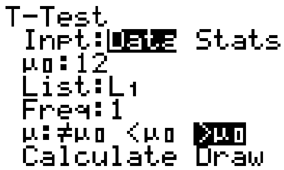
For the T- Distribution:

STAT ⇒ TESTS ⇒ T-TEST

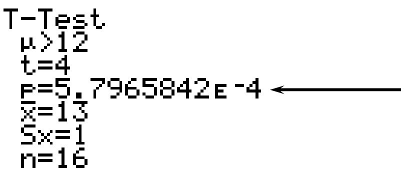
When you know the stats



When you have the data



And the results…



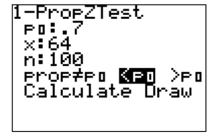
p-value

SUMMARY

Testing a Proportion?

- ***Both np and n(1- p) > 5 ?***

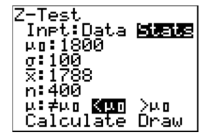
1-PropZTest



Testing a Mean?

- ***σ is known***
- ***Population is Normally distributed or n > 30***

Z -Test



Testing a Mean?

- ***σ is unknown***

T -Test