

Hypothesis Testing 1 sample

p. ①

Last chapter: Confidence Interval

- ① take sample
- ② Build Interval
- ③ We said 95% chance Population Mean (Proportion) is in our Interval.

Essence:

we took sample & made statement about population parameter with a 5% risk that our statement was not correct.

This chapter: Hypothesis Testing

statistical procedure that uses sample data to determine whether a statement about the value of a population parameter (μ or p) should be rejected or should not be rejected

Very

Similar

testing statements to see if they are reasonable or not

2 competing statements:

Null Hypothesis H_0

Alternative Hypothesis H_a

next 2 pages are
from last chapter →

p.5 & # p.6

Example 1 known

The Solid Construction Company

The Solid Construction Company constructs decks for residential homes. They send out two person teams to build decks. The company conducts a sample of 40 jobs and calculates a mean completion rate of 8 hours to build a typical deck. The standard deviation for the population is known to be equal to 3 hours (from past data).

Here we are given \bar{X} & expected to estimate a confidence interval for the population mean, M . P. 5

$$n = 40$$

$$\bar{X} = 8 \text{ hours}$$

$$\sigma = 3 \text{ hours}$$

$$\text{Standard error} = \frac{\sigma}{\sqrt{n}} = \frac{3}{\sqrt{40}} = .4743$$

what is a reasonable range of values for pop. mean

But what about sampling error? Does \bar{X} always = M ? We must add some "error room" to each side of our \bar{X} .

1. List variables:

2. What is the best estimation for the population mean?

Our best estimate for the population mean is our sample mean, $\bar{X} = 8$ hours. " $\bar{X} = 8$ hours" is a point estimate for our unknown population mean.

3. Determine a 90% Confidence Interval

- a. State the level of confidence, then divide by two and find probability in the body of the table then find z or use NORMSINV function

$$\text{Confidence Interval} = 90\%$$

$$\cdot 9/2 = .45$$

$$\alpha = \text{Risk that } M \text{ is not in interval} = .10$$

$$\alpha/2 = .10/2 = .05$$

$$z = \text{NORMSINV}(1 - .05) \approx 1.65$$

$$z = \text{NORMSINV}(.45 + .5) \approx 1.65$$

- b. Using the correct confidence interval formula, calculate the confidence limits

$$\bar{X} \pm z * \frac{\sigma}{\sqrt{n}}$$

$$= 8 \pm 1.644853627 * .474341649$$

$$= 8 \pm .780222582$$

$$\left. \begin{array}{l} 7.219777418 \\ 8.780222582 \end{array} \right\}$$

$$\left. \begin{array}{l} .22 * 60 \approx 13 \text{ mins.} \\ .78 * 60 \approx 47 \text{ mins.} \end{array} \right\}$$

4. Conclusions:

We are 90% sure that the population mean occurs between 8 h. 47 mins & 7 h. 13 mins. If we were to construct 100 similar intervals, we would expect to find pop. mean in 90 of them.

5. Would it be reasonable to conclude that the population mean is 9 hours? (If SCC claimed 9 hours in an ad, we would treat that as a population mean, $\mu = 9$ hours).

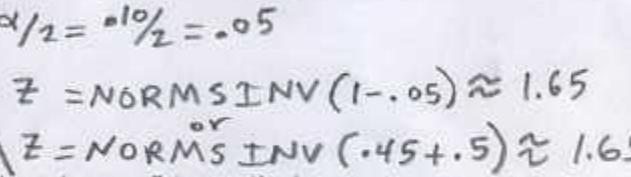
Because 9 hours is outside our interval, it would not be reasonable to conclude that the pop. mean is 9 hours. SCC's claim would not seem reasonable.

6. What if SCC claimed the population mean is 6 1/2 hours? (If SCC claimed 6.5 hours in an ad, we would treat that as a population mean, $\mu = 6.5$ hours).

Because 6.5 hours is outside our interval, it would not be reasonable to conclude that the population mean is 6.5 hours. SCC's claim would not seem reasonable.

7. What if SCC claimed 7 1/2 hours was their average job time? Is this reasonable? (If SCC claimed 7.5 hours in an ad, we would treat that as a population mean, $\mu = 7.5$ hours).

Because 7.5 hours is inside our interval, it would be reasonable to conclude that the pop. mean is 7.5 hours. SCC's claim seems reasonable.

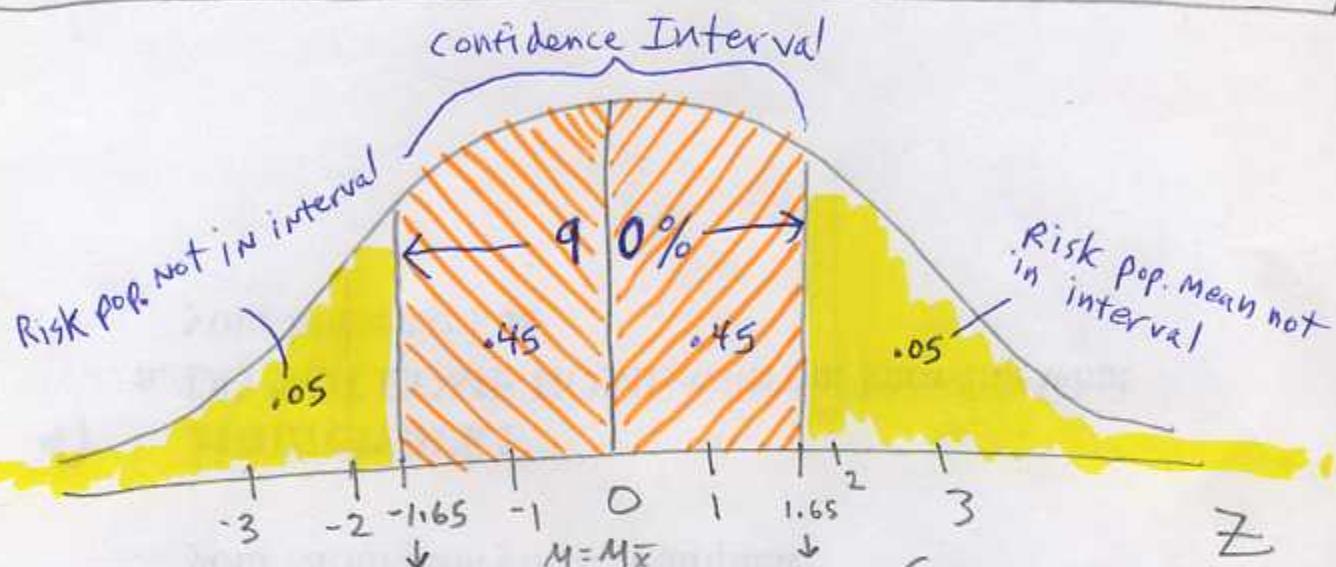


Picture

Visual summary for Example 1 or known

(6)

From $\bar{X} = 8$ hours, we created **Confidence Interval** so we were 90% sure that population mean was in the interval 7.22 hours to 8.78 hours.



{Confidence Limit}
on Low end}

$$8 - 1.65 * \frac{3}{\sqrt{40}} \quad \text{standard error}$$

$$8 - .78 \quad \text{Margin of error}$$

$$7.22 \text{ hours}$$

? {Confidence Limit
on upper end}

$$8 + 1.65 * \frac{3}{\sqrt{40}} \quad \text{standard error}$$

$$8 + .78 \quad \text{Margin of error}$$

$$8.78 \text{ hours}$$

Formulas:

$$Z = \text{NORMSINV}(1 - .05) = 1.65$$

$$\{\text{Confidence Limits}\} = \bar{X} \pm Z \frac{\sigma}{\sqrt{n}}$$

or

$$\text{Low value} = \text{NORMINV}(.05, \bar{X}, \frac{\sigma}{\sqrt{n}})$$

$$\text{Upper value} = \text{NORMINV}(1 - .05, \bar{X}, \frac{\sigma}{\sqrt{n}})$$

conclu.
z

We are 90% sure that the population mean occurs between 7.22 hours and 8.78 hours. A reasonable range of values for the population mean is 7.22 hours to 8.78 hours.



Let's see same example with Hypothesis T. →

Hypothesis Testing

statement from Solid Construction Company:

"It takes us 6.5 hours to build a typical deck"

A consumer group says:

"we don't think it takes 6.5 hours"

$$n = 40$$

$M_0 = 6.5$ = hypothesized population mean

$\sigma = 3$ hours (sigma known)

Confidence Interval (confidence coefficient) = .90

$\alpha = \alpha$ = risk that population parameter is not in interval = Risk that our test will reject original statement (H_0) when it is true. = $1 - .9 = .10$

$$\alpha/2 = .10/2 = .05$$

$$\bar{O}_{\bar{x}} = \text{standard error} = \frac{3}{\sqrt{40}} = .4743$$

$$z = \text{test statistic} = \frac{\bar{x} - M_0}{\bar{O}_{\bar{x}}} = \frac{6.5 - 8}{.4743} = -3.16228$$

critical value = Dividing line between whether statement is reasonable or not

$$= +/- 1 - \text{NORMSINV}(\alpha/2) = 1 - \text{NORMSINV}(.05) = \\ = 1.645$$

p-value = probability of getting test statistic or more (upper end)

$$p\text{-value} = 2 * \text{NORMSDIST}(-z \text{ test statistic}) = 2 * \text{NORMSDIST}(-3.16)$$

conclude:

Because -3.16 is not in

interval & because $p\text{-value} = .001565$

$.00157 \leq .10$ the original

statement is not reasonable, the

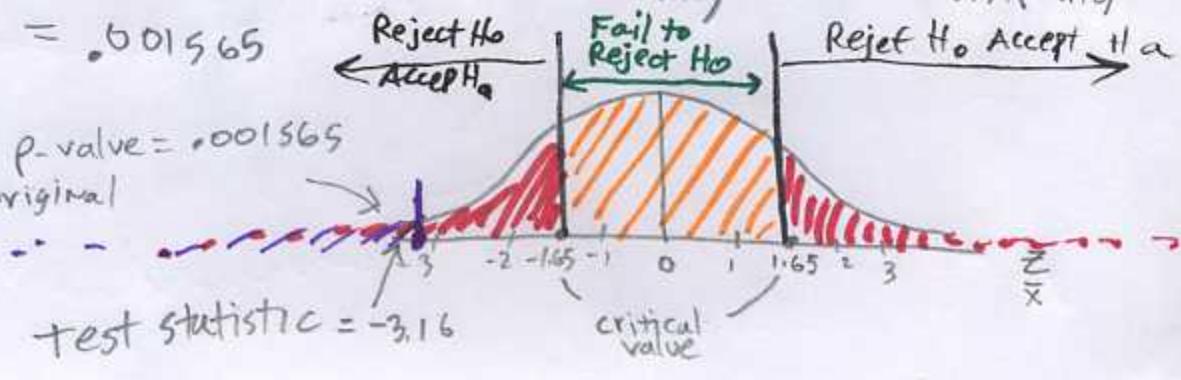
pop. mean is not 6.5 hours

$$= .001565$$

Reject H_0
Accept H_a

Fail to
Reject H_0

Reject H_0 Accept H_a



Step 1 List Null & Alternative Hypothesis

P. (4)

$$H_0: \bar{M} = 8 \text{ hours}$$

* $<>$ means \neq

$$H_a: \bar{M} <> 8 \text{ hours}$$

Step 2

Select Alpha

Risk that we reject H_0 when it is True

$$\alpha = .10$$

$$\alpha/2 = .1/2 = .05$$

Step 3 Collect sample Data, compute test statistic (z or t), Draw picture, calculate critical value & p-value, & state Rejection Rules

Reject H_0
Accept H_a

Fail to
Reject H_0

Reject H_0
Accept H_a

$$Z = -3.16$$

$$\text{critical value} = -1.65 \text{ & } 1.65$$

$$p\text{-value} = .001565$$



Rejection Rules:

If Z is between ± 1.65 , we fail to reject H_0 , otherwise we Reject H_0 & Accept H_a .

or
If $p\text{-value} \leq \alpha$, Reject H_0 & Accept H_a , otherwise fail to reject H_0

- ⑤ compare critical value to test statistic or compare p-value to α & conclude.

Because -3.16 is not between ± 1.65 & because $.001565$ is less than $.10$, we reject H_0 & accept H_a . From our sample evidence, it is more than reasonable to assume that the typical time to create deck is NOT 6.5 hours.

Examples of "statements" aims to see whether they are reasonable: P. 5

* Example 1: The new solicitation letter (asking for contributions to non-profit organization) is more effective than the old letter (old letter resulted in 15% contributing).

* Example 2: Is the manufacturer's claim that 16 oz. of catsup is in each bottle reasonable?

* Example 3: Is the mean monthly unpaid customer balance (customer owes money) more than \$400?

* Example 4: Is the new machine more efficient than the old one?

Hypothesis Testing (one - sample) ⑥

Hypothesis:

examples
page 2

- A statement about a population parameter subject to verification.
- A statement about the value of a population parameter developed for the purpose of testing.

Hypothesis Testing:

- A procedure based on sample evidence and Probability theory to determine whether the hypothesis is a reasonable statement.
- Hypothesis Testing does not prove that the hypothesis is true or false, but rather, it determines whether the hypothesis is a reasonable statement.
- Statistical procedure that uses sample data to determine whether a statement about the value of a population parameter (μ or p) should be rejected or should not be rejected

Steps of Hypothesis Testing

- ① Develop Null Hypothesis (H_0) & Alternative Hypothesis (H_1 or H_a)
- ② specify the level of significance (α)
- ③ collect sample Data & compute value of test statistic (Z or t), draw Picture.

P-value Approach

- ④ use value of test statistic to compute p-value
- ⑤ Reject H_0 if $p\text{-value} \leq \alpha$

Critical Value Approach

- ④ use level of significance to determine the critical value, and state rejection rule
- ⑤ use the value of the test statistic and the rejection rule to determine whether to reject H_0

Notes:

- ① If population data is Normally distributed, these methods are exact ($.99 = CI, \alpha = .01$, then 99 intervals contain μ , 1 does not)
- ② If pop. data is not normal, the bigger the n , the more exact.
 Pop Normal = any n can be used
 Approx. Normal $n \geq 15$
 Not Normal $n \geq 30$
 outliers $n \geq 50$

confidence Interval Hypothesis Testing

If:

$$H_0 : \mu = \mu_0$$

$$H_a \text{ or } H_1 : \mu < \mu_0$$

Then:

$\neq = \text{Not} = <>$

↑
Excel
Symbol
for
"Not" or
"Not Equal"

μ_0 = hypothesized population mean

- ① Select a simple random sample from the population and use the value of the sample mean \bar{X} to develop a confidence interval for the population mean μ .

$$\bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

\bar{X} = sample mean

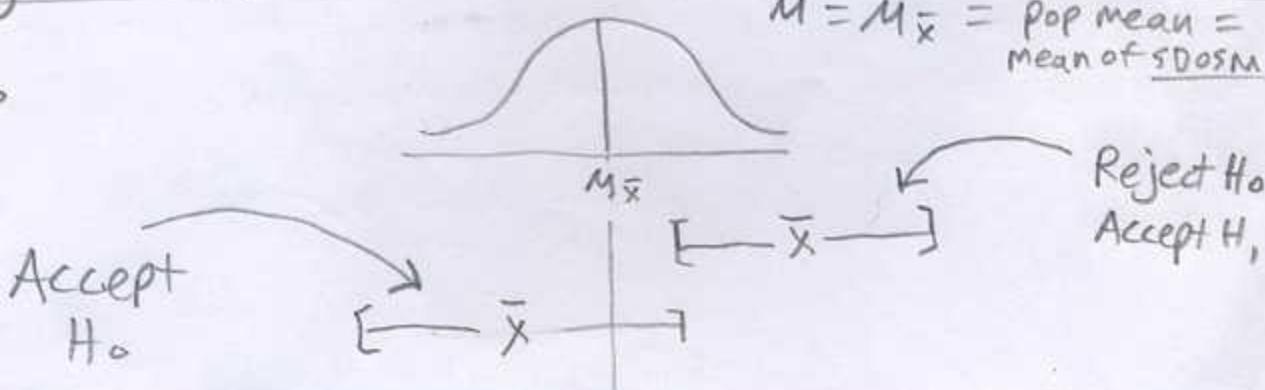
$Z_{\alpha/2}$ = upper Z

σ = pop S.D.

n = sample size

- ② If the confidence interval contains the hypothesized value μ_0 , do not reject H_0 , otherwise, Reject H_0 .

Example:



Test statistic (z or t) for Hypothesis Testing

About a population mean

σ Known

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

σ NOT Known

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

μ_0 = hypothesized mean

z & t = calculated test statistic, used to determine whether to reject the Null Hypothesis. Compare z or t to critical value to make decision, or used to calculate p-value.

\bar{x} = sample mean

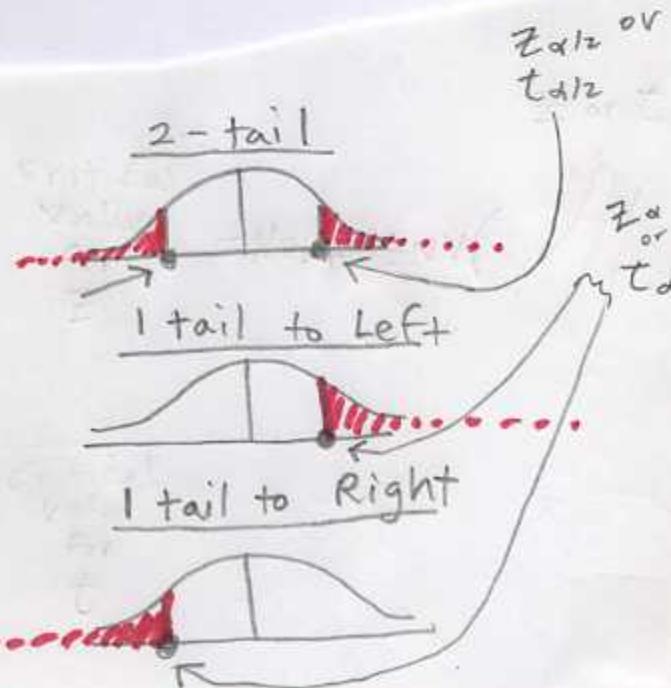
σ = population standard deviation

s = sample standard deviation

n = sample size

critical value (Z_α or T_α)

The dividing point between the region where the Null Hypothesis is rejected and the region where it is not rejected. Determined from alpha.



Test statistic for Hypothesis Tests About A Population Proportion

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

\bar{p} = sample proportion

p_0 = hypothesized pop. proportion

n = sample size

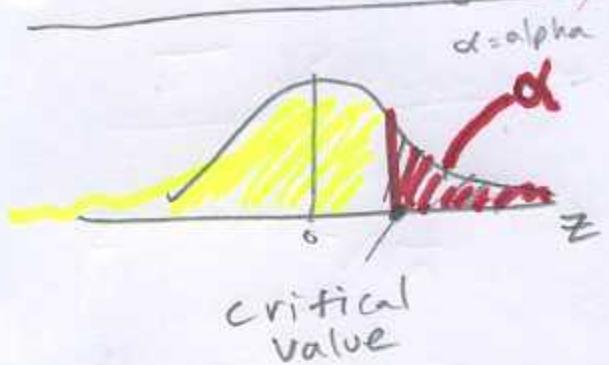
Must verify:

- ① Are there fixed # trials?
- ② Are results Independent?
- ③ Does each trial result in success or failure?
- ④ P stay same on each trial?
- ⑤ $n * p > 5$
 $n * (1-p) > 5$

Excel Functions

P.11

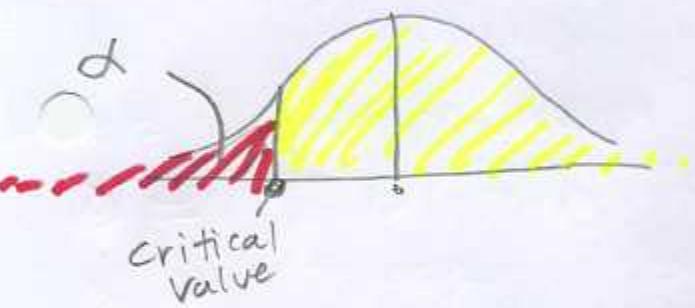
1-tail to right σ known for Z



$$\left. \begin{array}{l} \text{critical value} \\ \text{value} \end{array} \right\} = \text{NORMSINV}(1 - \underbrace{\alpha}_{\text{alpha}})$$

$$p\text{-value} = 1 - \text{NORMSDIST}(\underbrace{z}_{\text{calculated test statistic}})$$

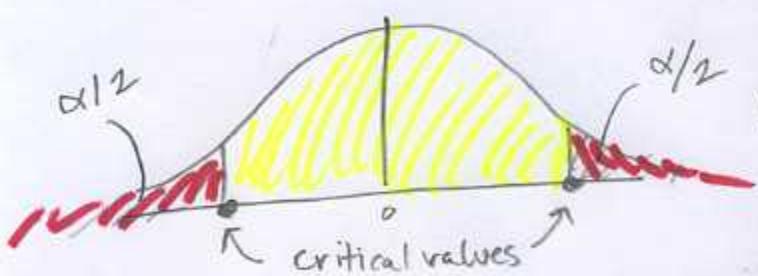
1 tail to left σ known for Z



$$\left. \begin{array}{l} \text{critical value} \\ \text{value} \end{array} \right\} = \text{NORMSINV}(\underbrace{\alpha}_{\text{alpha}})$$

$$p\text{-value} = \text{NORMSDIST}(\underbrace{z}_{\text{calculated test statistic}})$$

2 tail σ known for Z



$$\left. \begin{array}{l} \text{critical value} \\ \text{value} \end{array} \right\} = \pm \text{NORMSINV}(\alpha/2)$$

$$p\text{-value} = 2 * \text{NORMDIST}(z \text{ on low end}, z \text{ on high end})$$

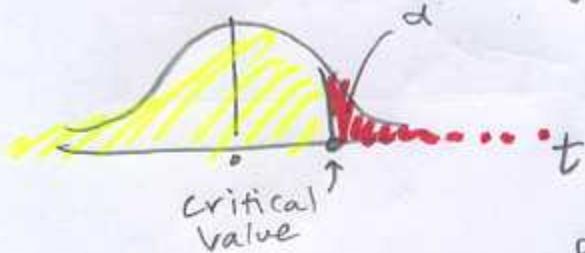
Notes: Z can be used for:

- ① Mean with σ known
- ② Proportions (Binomial Tests met)

Excel Functions

1 tail to right σ Not known for t

$$\text{critical value} = TINV(2*\alpha, \text{degrees of freedom})$$



ABOUT TINV: ① TINV always calculates value for a 2 tail test - that is why we multiply by 2. TINV always gives value on upper end.

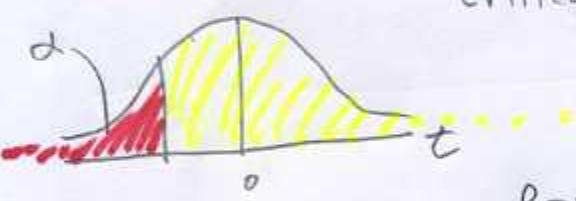
$$P\text{-value} = TDIST(t, \text{degree of freedom}, 1)$$

Always must be positive

t →
 $\text{df} \uparrow$
 $\# \text{ of tails in test}$

1 tail to left σ NOT KNOWN for t

$$\text{critical value} = -TINV(2*\alpha, \text{degree of freedom})$$



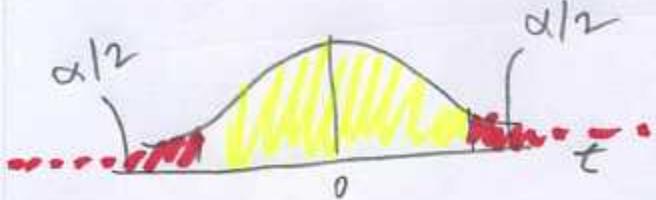
Because TINV always calculates for 2-tail test, Because TINV always calculates for upper value.

$$\text{more common } \rightarrow P\text{-value when } t \text{ is negative} = TDIST(-t, \text{degree of freedom}, 1)$$

$$\text{less common } \rightarrow P\text{-value when } t \text{ is positive} = 1 - TDIST(t, \text{degree of freedom}, 1)$$

2 tail σ NOT KNOWN for t

$$\text{critical value} = +/- TINV(\alpha, \text{degree of freedom})$$



$$P\text{-value} = TDIST(t, \text{degree of freedom}, 2)$$

t →
 $\text{df} \uparrow$
 $\# \text{ of tails}$

Example of Hypothesis Testing

(13)

Statement from Official Report:

"The yearly salary earned by full-time
realtors is \$85,000 (with $\sigma = \$12,549$)

Researcher Believes:

"Realtors make more than \$85,000"

① If we take a sample & get an $\bar{x} = \$88,595$,
we must decide about the Sample Error
 $\$88,595 - \$85,000 = \$3,595$.

② Is the difference Sample Error (^{original statement reasonable})

③ Is the difference a True difference (^{original statement not reasonable})

True difference = "statistically significant"

Small sample Error = "statistically insignificant"

(2)

① Null Hypothesis symbol : H_0
 ↑
 "H sub zero"
 "no difference."

② Alternative Hypothesis symbol :

H_a or H_1
 ↑
 "H sub a"
 ↑
 textbook uses
 ↑
 "H sub 1"
 ↑
 MICHAEL GIRVIN uses

Both Mean same thing:
 "Alternative Hypothesis."

Step
1

14

State the Null Hypothesis (H_0) and the Alternative Hypothesis (H_1)

Null Hypothesis = H_0 = "H sub zero"

A statement about the value of a population parameter developed for the purpose testing numerical evidence.

No difference from the assumed population parameter assumed true

Alternative Hypothesis = H_1 = "H sub one"

A statement that is accepted if the sample data provides sufficient evidence that the null hypothesis is not reasonable.

accepted if H_0 rejected

Example of Step 1 From Example 5:

$$H_0: M \leq \$85,000$$

$$\bar{X} = \$88,595$$

$$H_a \text{ also } H_1: M > \$85,000$$

$$\text{Difference} = \\ (88595 - 85000) = 3595$$

IF the difference of \$3,595 turns out to be due to sampling error, we will fail to reject $H_0: M = 85,000$. The hypothesis is reasonable.

IF the difference of \$3,595 turns out to be a true difference (statistically significant), we will reject H_0 and accept $H_1: M \neq 85,000$. (Hypothesis Not reasonable).

Q: How did we come up with
 $H_0: \mu = 85000$, $H_1: \mu \neq 85000$?

A: Because our sample mean salary was \$88,595 and the original claim was $\mu = 85,000$, we had to design a test to see if the population means is actually larger than \$85,000. So, we wrote the alternative hypothesis first. Then we can write our Null hypothesis.

once you know the comparative symbol for H_1 , just put the opposite for the H_0

H_0 always get equal sign

colon says: "Here is the hypothesis"

$$H_0: \mu \leq 85,000$$

Always write H_0 second

H_0 or

$$H_1: \mu > 85,000$$

Always write H_1 first

The direction of arrow tells us this is a 1-tail test to Right →

Null & Alternative Hypotheses

P. 16

- ① Research Hypotheses → usually it is $\rightarrow H_0$ or H_a

Example:

Does process take less than 30 mins?

$$H_0: M \geq 30 \text{ mins}$$

$$H_a \text{ or } H_1: M < 30 \text{ mins}$$

- ② Validity of claim → claim is usually given $\rightarrow H_0$

Example:

Is catsup bottle filled with 16 oz?

(A)

Ad point of view
(consumers')

$$H_0: M \geq 16 \text{ oz.}$$

$$H_a \text{ or } H_1: M < 16 \text{ oz.}$$

Consumer only cares if they get 16 oz. or more

(B)

Manufacturer's
Filling Machine
point of view

$$H_0: M = 16 \text{ oz}$$

$$H_a \text{ or } H_1: M \neq 16 \text{ oz.}$$

Manufacturer cares if it is filled too much or too little

similar but from different points of view

MN
OR

- ③ Decision Making → choose between 2 things $\rightarrow H_0$ or H_1

Example:

Should we accept box of 20 meter boomerangs?

(if they fly too short, can't be used in competition)
(if they fly too long, times will not be fast enough.)

Choice #1

$$H_0: M = 20 \text{ meters} \rightarrow \text{accept Box}$$

Choice #2

$$H_a \text{ or } H_1: M \neq 20 \text{ meters} \rightarrow \text{reject Box}$$

Step
2

Select the Level of Significance

$$\text{Alpha} = \alpha$$

$$\text{Level of Significance} = \text{alpha} = \alpha = 1 - \{\text{Level of Confidence}\}$$

- The probability of rejecting the null hypothesis when it is true.
- Risk of rejecting H_0 when it is actually true
- "innocent but found guilty"

Standards for Level of Significance

$\alpha = .10$ often used in politics

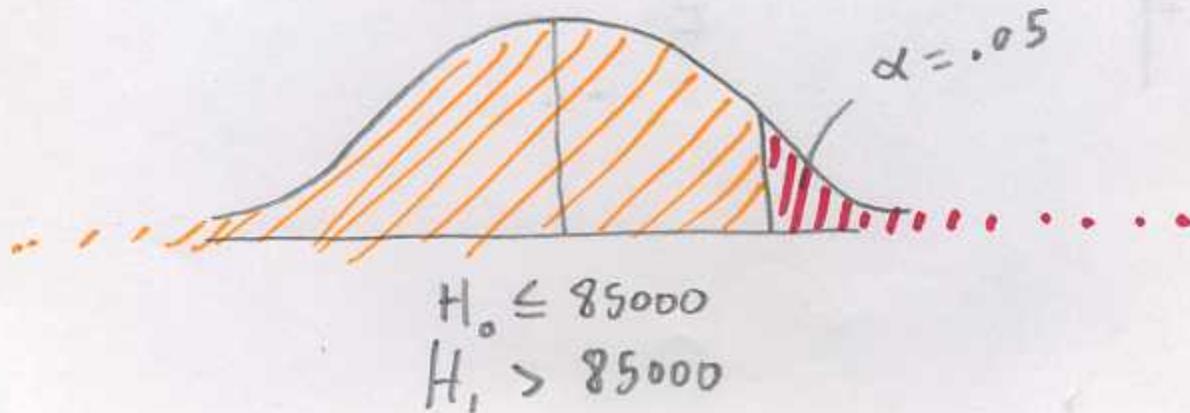
$\alpha = .05$ often used in consumer research projects

$\alpha = .01$ often used in Quality Assurance

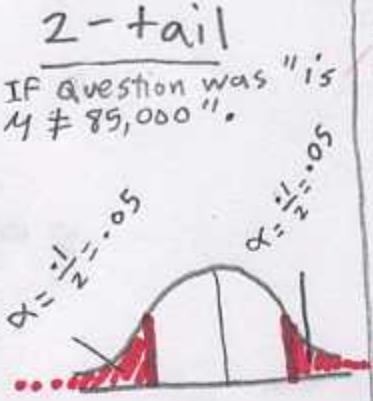
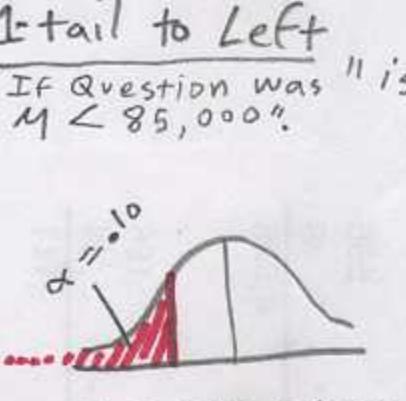
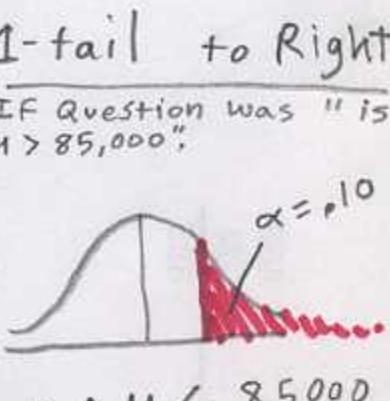
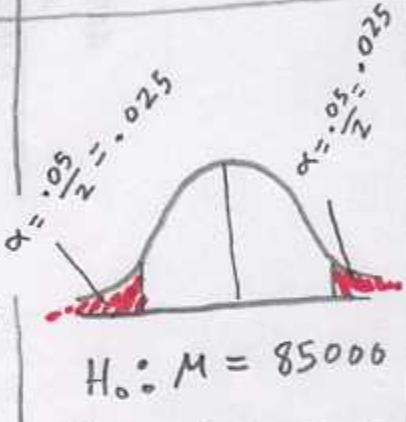
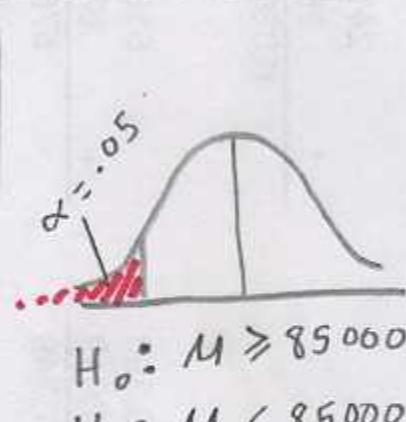
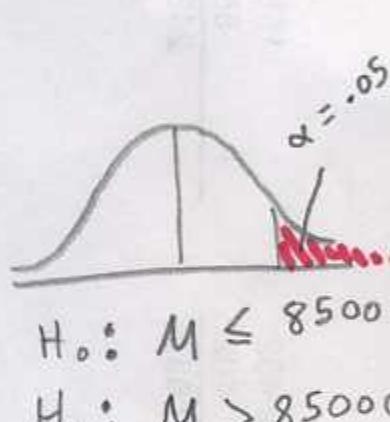
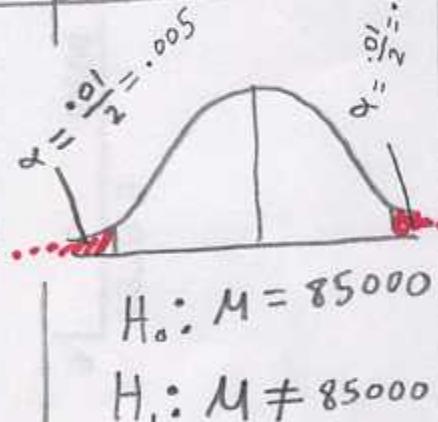
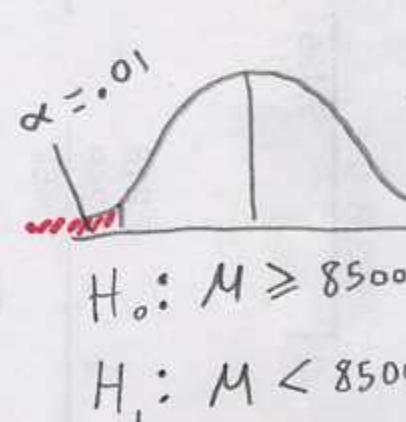
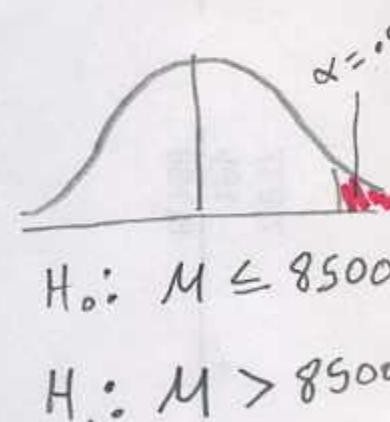
Example of step 2 from Example 5:

If we choose $\alpha = .05$, we are taking a 5% risk of rejecting H_0 even though it is true.

A picture would look like this:



Picture Examples for Level of significance where $M = 85,000$

<u>Alpha</u>	2-tail IF question was "is $M \neq 85,000$ ". 	1-tail to Left IF Question was "is $M < 85,000$ ". 	1-tail to Right IF Question was "is $M > 85,000$ ". 
$\alpha = .10$	$H_0: M = 85000$ $H_1: M \neq 85000$	$H_0: M \geq 85000$ $H_1: M < 85000$	$H_0: M \leq 85000$ $H_1: M > 85000$
$\alpha = .05$	 $H_0: M = 85000$ $H_1: M \neq 85000$	 $H_0: M \geq 85000$ $H_1: M < 85000$	 $H_0: M \leq 85000$ $H_1: M > 85000$
$\alpha = .01$	 $H_0: M = 85000$ $H_1: M \neq 85000$	 $H_0: M \geq 85000$ $H_1: M < 85000$	 $H_0: M \leq 85000$ $H_1: M > 85000$
These picture examples show the 3 possibilities at 3 different alpha values.			

Errors & Correct Conclusions in Hypothesis Testing

		Population Condition
		H ₀ True (H _a)
Conclusion	H ₀ True	correct conclusion
	Reject H ₀	Type I. α
Accept H ₀		Type II. β
		correct conclusion

Type I Error = α = alpha

H₀ True, But we reject H₀

"Innocent, but found guilty"

reject H₀ when true
Accept H₀ when H₀ is true

Type II Error = β = beta

H₀ False, But we accept H₀

Accept H₀ when false

Because we control for α , we can say "Accept H₀"

Because we don't control for β , we can't say "Accept H₀"

Accept H₀ when H₀ true

Level of significance (^{more comprehensive} definition)

P.20

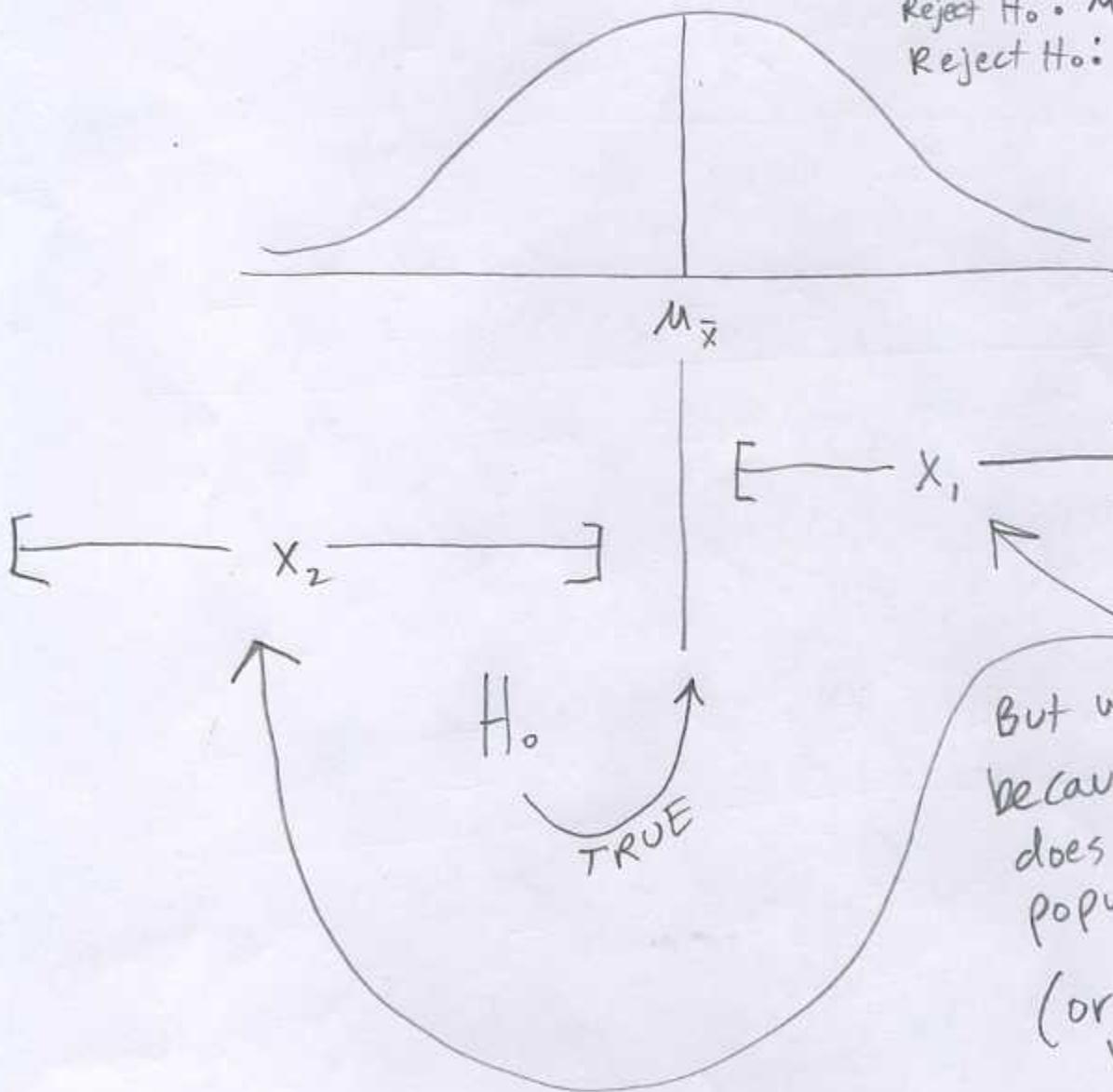
The level of significance is the probability of making a Type I error when the null hypothesis is true as an equality.

Examples:

Reject $H_0: \bar{M} = 16$ oz. when $\bar{M} = 16$

Reject $H_0: \bar{M} \leq 16$ when $\bar{M} = 16$

Reject $H_0: \bar{M} \geq 16$ when $\bar{M} = 16$



But we Reject H_0
because our interval
does not contain
population mean

(or above some
value or below
some value).

- ① By selecting α , the person is controlling the probability of making a Type I error $P(2)$
- ② Cost making Type I error high, make α small. Example $CI = .99 \quad \alpha = .01$
- ③ Cost making Type I error low, make α bigger. Example $CI = .90 \quad \alpha = .01$
- ④ Applications of hypothesis testing that only control for the Type I error are called "significance tests."
- ⑤ only 2 conclusions are possible:
- ① "fail to reject H_0 " or "Do Not Reject H_0 "
 - ② Reject H_0 , accept H_1
- ⑥ Because we will not control for Type II error, we don't say "accept H_0 !"

Hypothesized mean = M_0 .

\bar{X} = point estimate of M

But what about Sampling Error?

How big can Sampling Error
be before we say it is

"statistically significant"?

When you declare statistically
significant without testing the
whole population, you risk Type I, α ,
of rejecting H_0 when it is True,
that is $M = M_0$.

Example of Step 3 (collect Data, calculate test statistic, Draw Picture) 0.23

Because we know the population standard deviation $\sigma = 12,549$, we can use the test statistic, Z .

$$n = 36$$

$$\bar{X} = 88,595$$

$$\sigma = 12,549$$

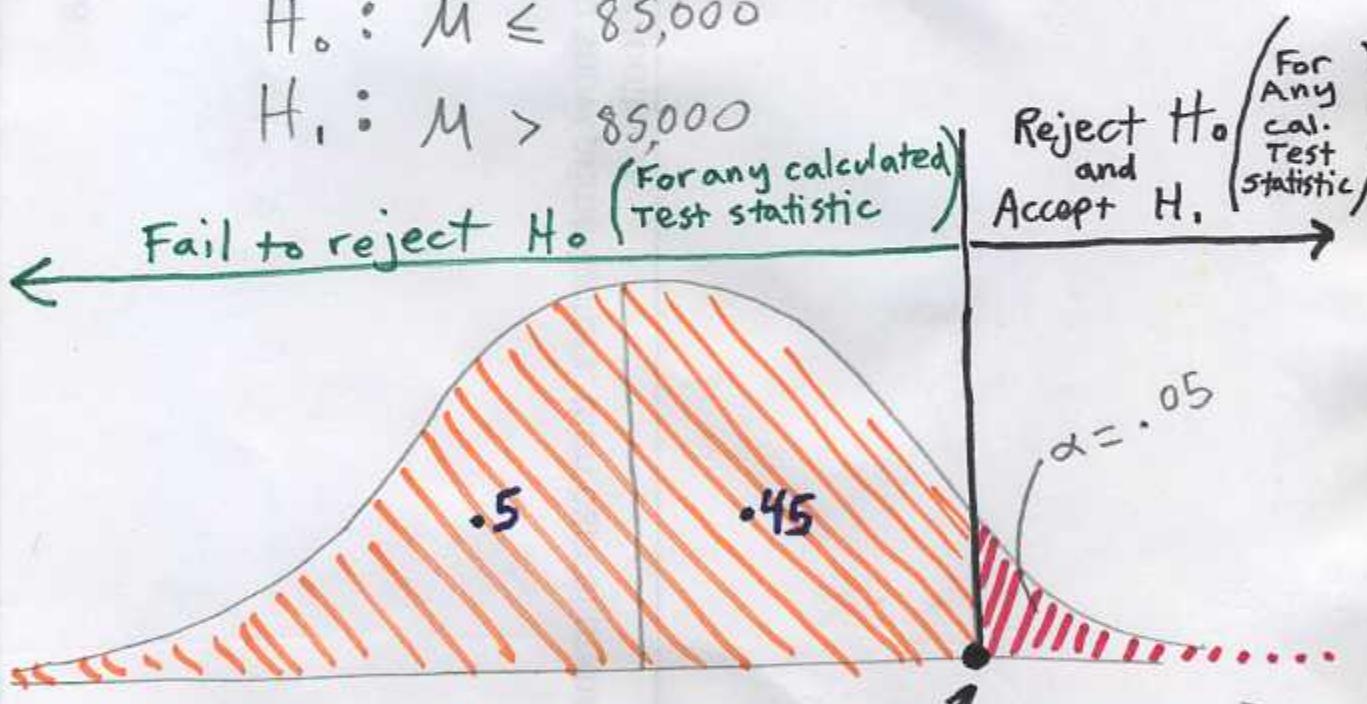
$$M = M_{\bar{X}} = 85,000$$

$$\sigma_{\bar{X}} = \frac{12549}{\sqrt{36}} = 2,091.50$$

$$\alpha = .05$$

$$H_0: M \leq 85,000$$

$$H_1: M > 85,000$$



$$\left\{ \begin{array}{l} \text{critical} \\ \text{value} \end{array} \right\} = \text{NORMSINV}(1 - .05) = 1.6448$$

$$M = M_{\bar{X}} = 85,000$$

$$\sigma_{\bar{X}} = 2,091.50$$

Step 4

Formulate the decision rule
for critical value

P.
24

① From the picture, create your rule:

fail to reject H_0

$$H_0: M \leq 85000$$

$$H_1: M > 85000$$

reject H_0 and

$$\alpha = .05$$

Accept H_1

$$= \text{NORMSINV}(1 - .05) = 1.645$$

critical value

② Decision Rule or Rejection Rule

If our calculated test statistic is greater than 1.645, we reject H_0 and accept H_1 , otherwise we fail to reject H_0 .

Step 5

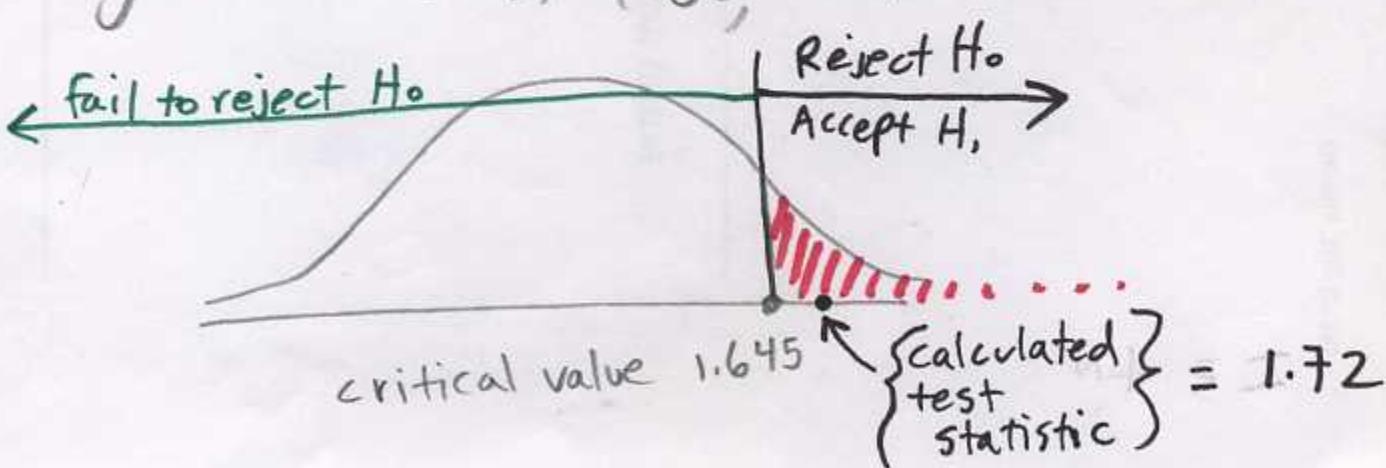
Conclude w/ critical value
& Rejection rule

P
25

$$\left\{ \text{Calculated test statistic} \right\} = \frac{88595 - 85000}{\frac{12549}{\sqrt{36}}} = 1.72$$

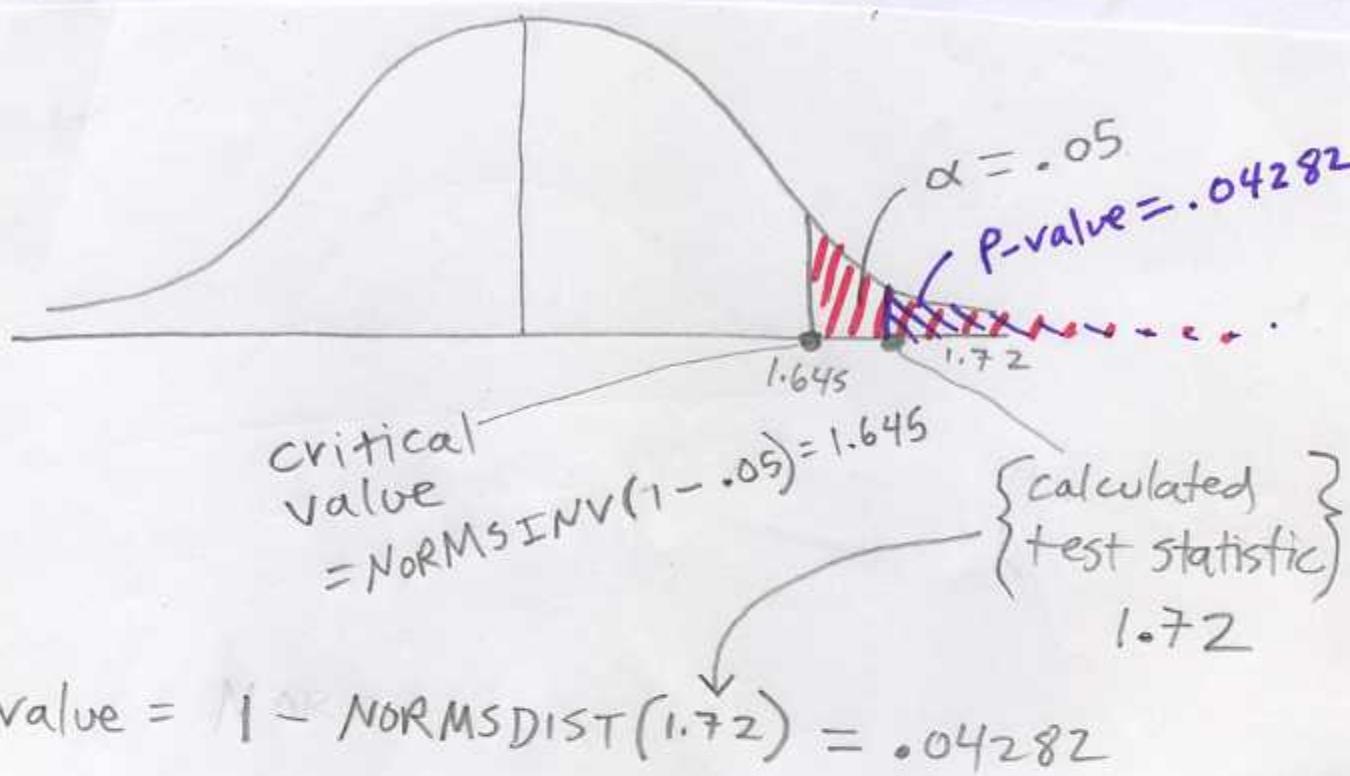
Make Decision:

Because our calculated test statistic is greater than 1.645, we reject H_0 and accept H_1 . It is reasonable to assume that the mean salary for real estate agents is greater than \$85000.



Step 4 & 5 for p-value

P.26



Because the p-value is less than alpha ($.04282 \leq .05$), we reject H_0 & accept H_1 . It is reasonable to assume that the mean salary for real estate agents is greater than \$85,000.

Interpreting P-value

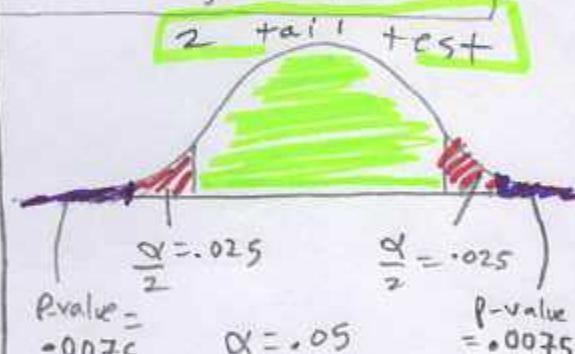
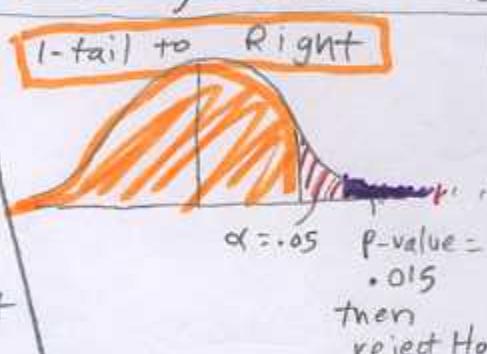
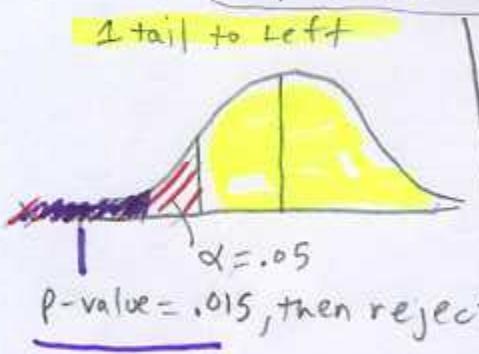
- $p\text{-value} > .10$ / Insufficient evidence to say H_0 is TRUE
- $.05 < p\text{-value} \leq .10$ weak evidence to say H_0 is TRUE
- $.01 \leq p\text{-value} \leq .05$ strong evidence to say H_0 is TRUE
- $p\text{-value} < .01$ very strong evidence to say H_0 is TRUE

p-value (also called "observed p-value")

A p-value is a probability that provides a measure of evidence against the null hypothesis provided by the sample.

Rule:

$p\text{-value} \leq \alpha$, then: Reject H_0 , Accept H_1



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

p-value is probability of getting test statistic or less

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

p-value is probability of getting a value greater than or equal to test statistic

2-tail
 $2 * p\text{-value} \leq \alpha$, Reject H_0

probability of getting a value less than or equal to the lower end test statistic or a value greater than or equal to the upper end test statistic.

$$M = M_{\bar{x}} = M_0 = 85000$$

$$\bar{x} = 88,595$$

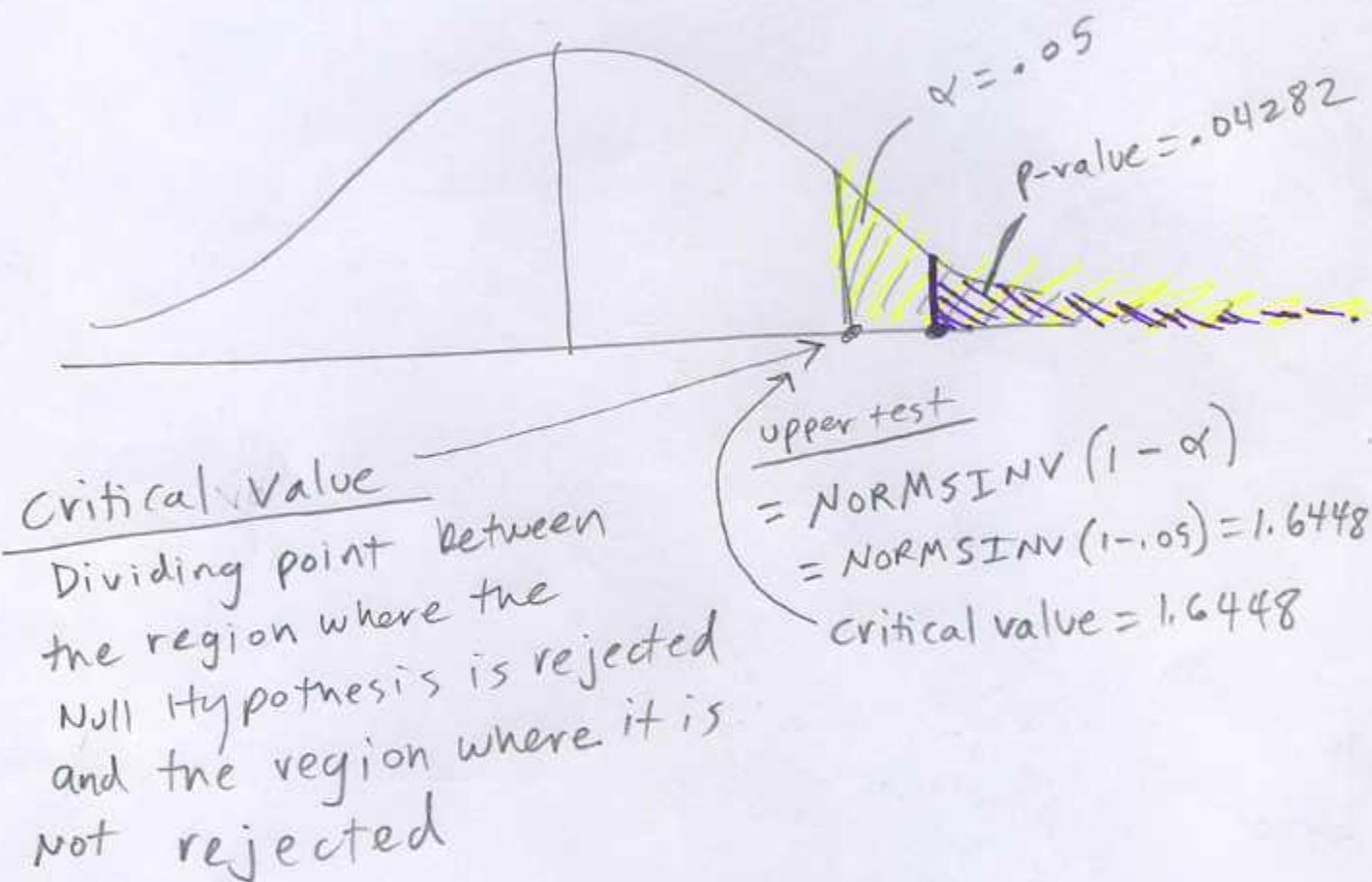
$$n = 36$$

$$\sigma = 12,549$$

$$\sigma_{\bar{x}} = 12549 / \sqrt{36} = 2,091.50$$

$$\alpha = .05$$

test statistic = $\frac{88,595 - 85000}{2091.5} = 1.72$



p-value

probability of getting test statistic

$$\text{or More} = 1 - \text{NORMSDIST}(z)$$

$$\text{p-value} = 1 - \text{NORMDIST}(1.72) = .04282$$

Summary for
Real Estate Example:

(P.29)

concluding :

use Z or t to compare to critical value

use p-value to compare to alpha α

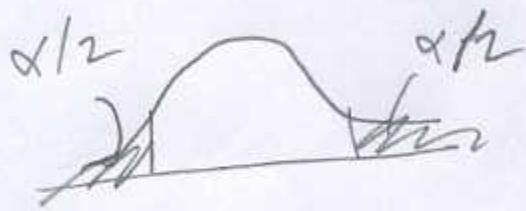
Critical value



t or $Z \geq$ critical value,
Reject H_0 , Accept H_a



t or $Z \leq$ critical value, Reject
 H_0 , Accept H_a



$-t$
 $-Z \leq$ critical value $\leq Z$,
Fail to Reject H_0

p-value

p-value $\leq \alpha$, Reject H_0 ,
Accept H_a

Hypothesis Testing Example #1

- In the past, 15% of the mail order solicitations for a certain charity resulted in a financial contribution
- A new solicitation letter that has been drafted is sent to a sample of 200 people and 35 responded with a contribution
- Assume Experiment passes all the binomial tests
- At the .05 significance level can it be concluded that the new letter is more effective?
-

Conduct A Test Of Hypothesis About A Population Proportion

- Test statistic for testing a single population proportion:

$$z = \frac{p - \pi}{\sqrt{\frac{\pi(1 - \pi)}{n}}}$$

32

example 1 $\pi = .15$ proportion of mail order solicitations that result in financial contribution

P 34

$$x = 35$$

$$n = 200$$

$$p = \frac{x}{n} = \frac{35}{200} = .175$$

- Step 1: State null and alternate hypotheses

$$H_0: \pi \leq .15$$

$$H_1: \pi > .15$$

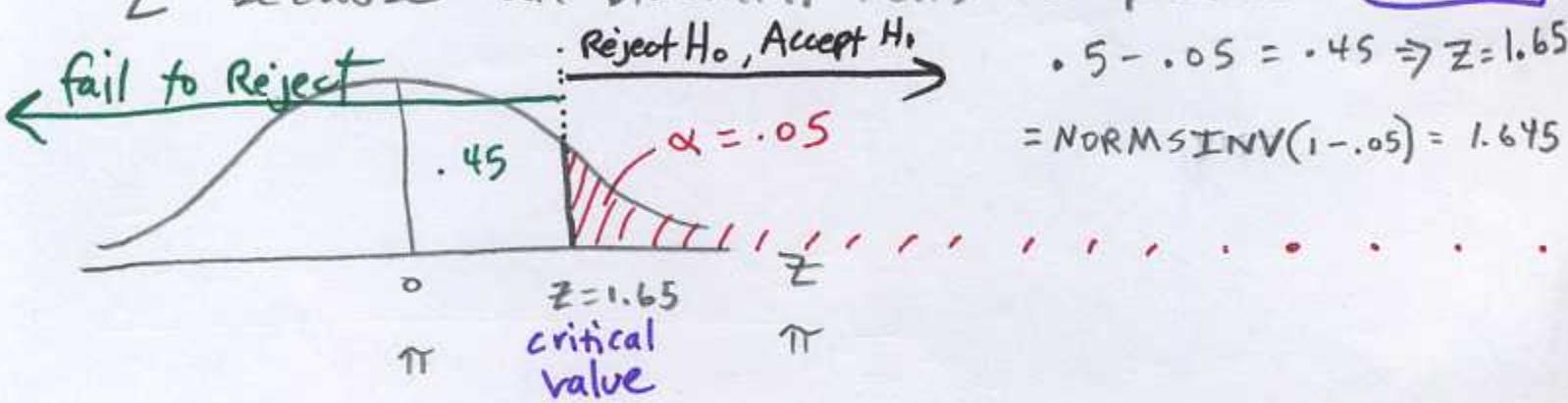
- Step 2: Select a level of significance

$$\alpha = .05$$

- Step 3: Identify the test statistic (z or t) and draw curve with critical value

Z because all binomial tests are passed

Critical value



- Step 4: Formulate a decision rule

If the calculated test statistic is greater than 1.65, we reject H_0 and accept H_1 , otherwise we fail to reject H_0 .

sample
 $p = \text{proportion}$
 $\pi = \text{pop proportion}$

- Step 5: Take a random sample, compute the test statistic, compare it to critical value, and make decision to reject or not reject null and hypotheses

calculated test statistic

$$= \frac{.175 - .15}{\sqrt{\frac{.15(.85)}{200}}} = .990147543$$

$$Z = \frac{p - \pi}{\sqrt{\frac{\pi * (1 - \pi)}{n}}}$$

Because .99 is less than 1.65, we fail to reject H_0 . The evidence suggests that the new letter is not more effective. The difference between .15 and .175 seems to be due to sample error.

Can we conclude that the new letter is more efficient?
Is the proportion more than .15?

Conduct A Test Of Hypothesis About A Population Mean

Hypothesis Testing Example # 2

- The processors of Fries' Catsup indicate on the label that the bottle contains 16 ounces of catsup
- The standard deviation of the process is 0.5 ounces
- A sample of 36 bottles from last hour's production revealed a mean weight of 16.12 ounces per bottle
- At the .05 significance level is the process out of control?
- Can we conclude that the mean amount per bottle is different from 16 ounces?

1

example 2

$M = 16.02$, $\sigma = .502$, $n = 36$, $\bar{X} = 16.12 \text{ oz}$.

$\alpha = .05$, $SE = \frac{\sigma}{\sqrt{n}} = \frac{.5}{\sqrt{36}} = .125$

- Is the production machine accurately filling the bottles?
- Can we conclude that the mean amount/bottle is different from 16 oz.?

• Step 1: State null and alternate hypotheses

$$H_0: M = 16 \text{ oz}$$

$$H_1: M \neq 16 \text{ oz.} \quad \text{2 tail test}$$

• Step 2: Select a level of significance

$$\alpha = .05 \quad \alpha/2 = \frac{.05}{2} = .025$$

critical value

• Step 3: Identify the test statistic (z or t) and draw curve with critical value

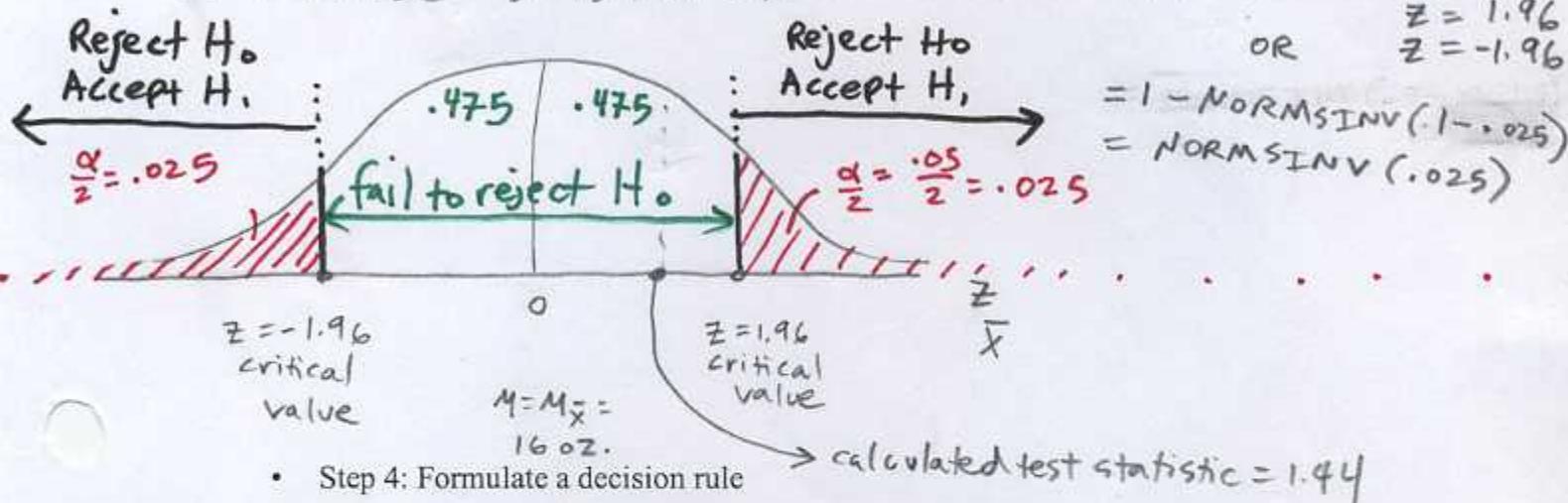
Z because σ is known

$$Z = .5 - .025 = .475$$

$$Z = 1.96$$

$$OR \quad Z = -1.96$$

$$Z = -1.96$$



• Step 4: Formulate a decision rule

If our calculated test statistic is greater than 1.96 or less than -1.96, we reject H_0 and accept H_1 , otherwise we fail to reject H_0 .

computed test statistic

• Step 5: Take a random sample, compute the test statistic, compare it to critical value, and make decision to reject or not reject null and hypotheses

$$Z = \frac{\bar{X} - M}{\frac{\sigma}{\sqrt{n}}} = \frac{(16.12 - 16)}{\frac{.5}{\sqrt{36}}} = 1.44$$

$$p\text{-value} = (1 - NORMSDIST(1.44)) * 2 = .1499$$

Because our calculated test statistic is between 1.96 and -1.96, we fail to reject H_0 . It seems reasonable that 16 oz. is the population mean. It is reasonable to assume that the $(16.12 - 16) = .12$ is due to sampling error. We cannot conclude that the mean is different from 16 oz.. It is reasonable to assume that the production machine is filling accurately.

Example 3

$$\begin{aligned} M &= 3 \text{ mins wait time} \\ \sigma &= 1 \text{ min} \\ n &= 50 \\ \bar{X} &= 2.75 \text{ mins} \end{aligned}$$

can we conclude that the mean wait time is less than 3 mins?

Always gets =

- Step 1: State null and alternate hypotheses

Null hypothesis $H_0: M \geq 3 \text{ mins}$ (statement about population)

P.34

Alternative hypothesis $H_1: M < 3 \text{ mins}$ (statement accepted if sample data rejects H_0)

H_1 is taken if the sample data provides us with enough evidence that H_0 is not reasonable

- Step 2: Select a level of significance

$\alpha = .05$ (Risk of making error) (Risk of Rejecting H_0)

when it is true) (If confidence level is how sure we are - α = risk)

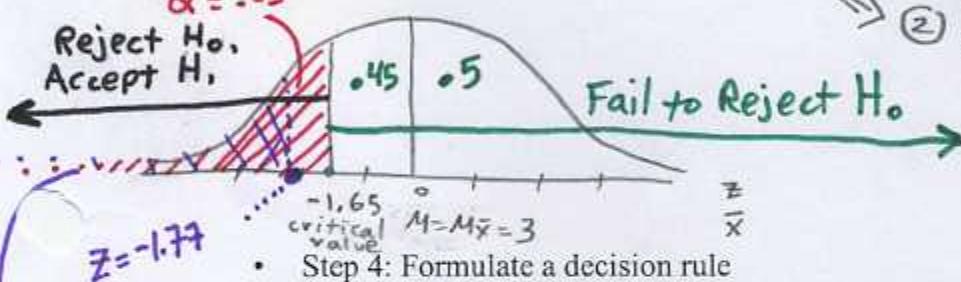
- Step 3: Identify the test statistic (z or t) and draw curve with critical value

or known → so use z

$Z = 1.65$

Indicates that this is a one tail test, pointing to left

Critical value = "hurdle" = $Z \Rightarrow$
 $\begin{cases} ① 0.5 - 0.05 = 0.45 \Rightarrow \text{look up } P(x) \\ ② \text{Excel in table \& find } Z \\ = \text{NORMSINV}(0.05) = -1.645 = Z \end{cases}$



- Step 4: Formulate a decision rule

test statistic IF our computed test statistic is less than -1.65 , we reject H_0 and accept H_1 , otherwise we fail to reject H_0

(Notice we "fail to reject H_0 " instead of saying H_0 is false)

- Step 5: Take a random sample, compute the test statistic, compare it to critical value, and make decision to reject or not reject null and hypotheses

{Computed test Statistic}

$$= Z \frac{\bar{X} - M}{\sigma / \sqrt{n}} = \frac{2.75 - 3}{1 / \sqrt{50}} = -1.76776695296637$$

P-value = Area associated w/ computed test statistic
 $= .0385$ Excel = NORMSDIST(-1.77)

- ① $P\text{-Val.} > \alpha$
fail to Reject H_0
- ② $P\text{-Val.} < \alpha$
Reject H_0

① Because -1.767 is less than -1.65 , we reject H_0 and accept H_1 . The evidence suggests that the mean wait time is less than 3 mins. We can be reasonably sure that the mean wait time is less than 3 minutes.

- ② Because $.0385$ is less than $.05$, reject H_0 and accept H_1

Hypothesis Testing Example 4

- The current rate for producing 5 amp fuses at Neary Electric Co. is 250 per hour
- A new machine has been purchased and installed that, according to the supplier, will increase the production rate
- A sample of 10 randomly selected hours from last month revealed:
 - Mean hourly production for new machine = 256 units
 - Sample standard deviation = 6 per hour
- At the .05 significance level can Neary conclude that the new machine is faster?

31

Example 4

$$\begin{aligned} M &= 250 \text{ Amp-Fuses/hour} \\ n &= 10 \\ \bar{x} &= 256 \text{ Amp Fuses/hour} \end{aligned}$$

$$\begin{aligned} s &= 6 \\ \alpha &= .05 \end{aligned}$$

- Step 1: State null and alternate hypotheses

$$H_0: M \leq 250 \text{ AF/hour}$$

$$H_1: M > 250 \text{ AF/hour}$$

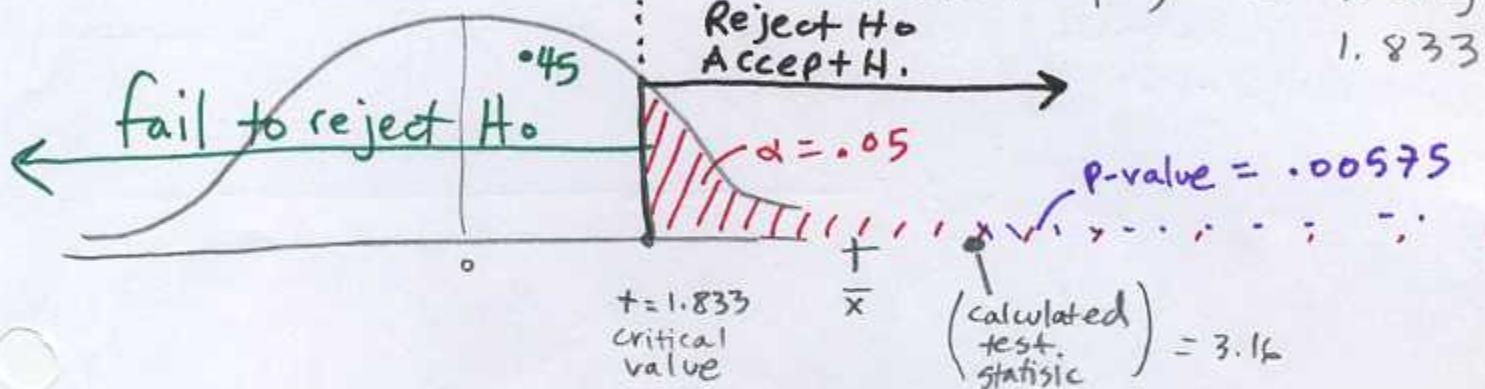
- Step 2: Select a level of significance

$$\alpha = .05$$

can nearly conclude that the new machine is faster? Are more than 250 amp-fuses produced per hour?

- Step 3: Identify the test statistic (z or t) and draw curve with critical value

+ because σ not known one tail } tables $\Rightarrow 1.833$
 $\alpha = .05$
 $df = 10 - 1 = 9$ } $= TINV(.05*2, 9) =$
 1.833113



- Step 4: Formulate a decision rule

IF our calculated test statistic is greater than 1.833, we reject H_0 and accept H_1 , otherwise we fail to reject H_0 .

- Step 5: Take a random sample, compute the test statistic, compare it to critical value, and make decision to reject or not reject null and hypotheses

calculated
test
statistic

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{256 - 250}{6/\sqrt{10}} = 3.16$$

$$p\text{-value} = TDIST(3.16, 9, 1) = .00575 \text{ (extremely unlikely)}$$

Because 3.16 is greater than 1.833, we reject H_0 & accept H_1 . The evidence suggests that the new machine is more productive (more than 250 AF/hour).

Example
5

$$M = \$85,000 \quad \bar{X} = 88595$$

$$\sigma = \$12,549 \quad \left\{ \begin{array}{l} \text{standard} \\ \text{error} \end{array} \right\} = \frac{12549}{\sqrt{36}} = \$2091.50$$

$$n = 36$$

- Step 1: State null and alternate hypotheses

$$H_0: M \leq 85000$$

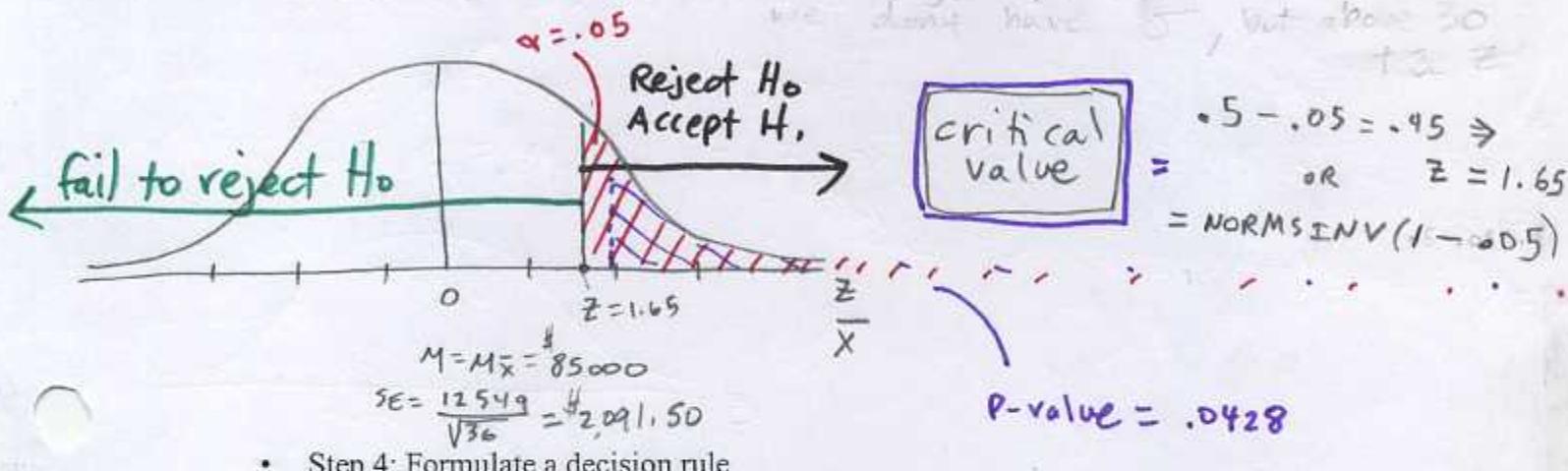
$$H_1: M > 85000 \quad (\text{one tail to right})$$

- Step 2: Select a level of significance

$\alpha = .05 \Rightarrow$ risk of rejecting H_0 even though true

- Step 3: Identify the test statistic (z or t) and draw curve with critical value

Z because σ known



- Step 4: Formulate a decision rule

If our computed test statistic is greater than 1.65 we reject H_0 & accept H_1 , otherwise we fail to reject H_0 .

- Step 5: Take a random sample, compute the test statistic, compare it to critical value, and make decision to reject or not reject null and hypotheses

$$\boxed{\text{Computed test statistic}} = \frac{(\bar{X} - M)}{\sigma/\sqrt{n}} = \frac{(88595 - 85000)}{12549/\sqrt{36}} = 1.72$$

$$P\text{-value} = 1 - \text{NORMSDIST}(1.72) = .0428$$

$P < \alpha \rightarrow$ reject
 $P > \alpha \rightarrow$ Don't

Because $1.72 > 1.65$ and because our p-value (.0428) is less than .05, we reject H_0 and accept H_1 . The evidence suggests that the mean yearly salary is greater than \$85,000.

Can we conclude that the mean yearly salary for RETA is greater than \$85,000?

p 37