Hypothesis Testing 1 sample

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 (MU 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 σ KNOWN

Here we are given $\bar{X}$ & expected to estimate a confidence Interval for the population mean, μ.

P. 5

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).

1. List variables: $n = 40$
   $\bar{X} = 8$ hours
   $\sigma = 3$ hours
   Standard error $= \frac{3}{\sqrt{40}} = .4743$

what is a reasonable range of values for pop. mean

Another way to say

But what about sampling Error? Does $\bar{X}$ always $= \mu$? we must add some error "room" to each side of our $\bar{X}$

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 [redacted] then find z or use NORMSINV function

confidence Interval = 90%
$.9/2 = .45$
α = Risk that μ is not in interval = .10

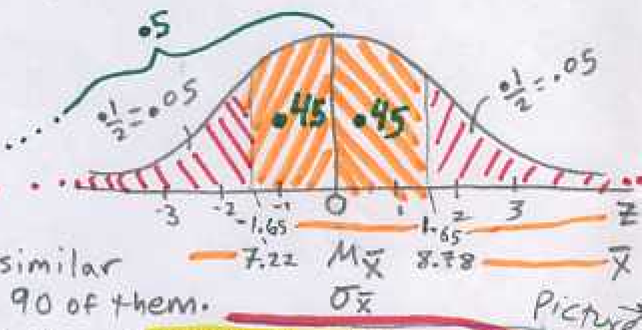
$\alpha/2 = .10/2 = .05$
$z = \text{NORMSINV}(1-.05) \approx 1.65$
or
$z = \text{NORMSINV}(.45+.5) \approx 1.65$

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

$\bar{X} \pm z * \frac{\sigma}{\sqrt{n}}$
$= 8 \pm 1.644853627 * .474341649$
$= 8 \pm .780222582$

→ 7.219777418 & 8.780222582

Minutes: $.22 * 60 \approx 13$ mins. $.78 * 60 \approx 47$ mins.



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 ½ hours? (If SCC claimed 6.5 hours in an ad, we would treat that as a population mean, mu = 6.5 hours).

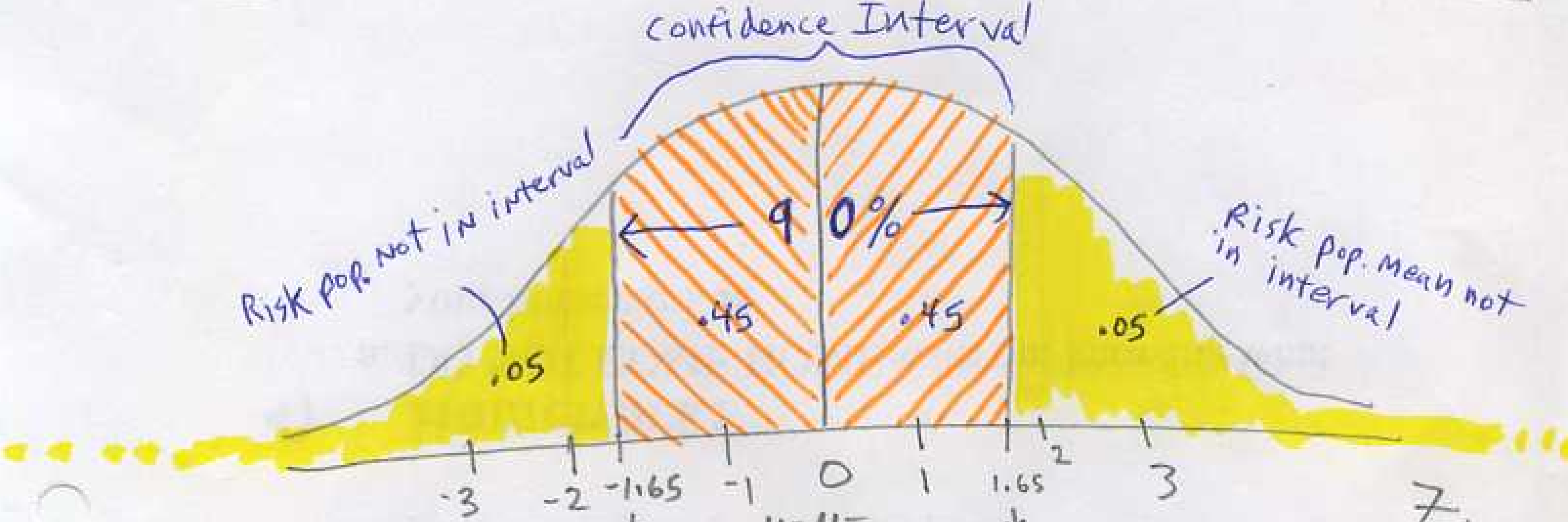
Because 6.5 hours is outside our interval, it would not be reasonable to conclude that the population mean is 6.5 hours. SCC's claim would not seem reasonable.

7. What if SCC claimed 7 ½ 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.

Visual summary for Example 1 σ KNOWN

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.



$\mu = \mu_{\bar{x}}$

$\sigma_{\bar{x}} = \frac{3}{\sqrt{40}}$

$\bar{X}$ (hours)

Confidence Limit on Low end	Confidence Limit on upper end
?	?
$8 - 1.65 * \frac{3}{\sqrt{40}}$ standard error	$8 + 1.65 * \frac{3}{\sqrt{40}}$ Standard error
$8 - .78$ Margin of error	$8 + .78$ Margin of error
7.22 hours	8.78 hours

Formulas:

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

$\left\{\begin{matrix}\text{confidence} \\ \text{Limits}\end{matrix}\right\} = \bar{X} \pm Z \frac{\sigma}{\sqrt{n}}$

or

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

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

Conclusion: 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"

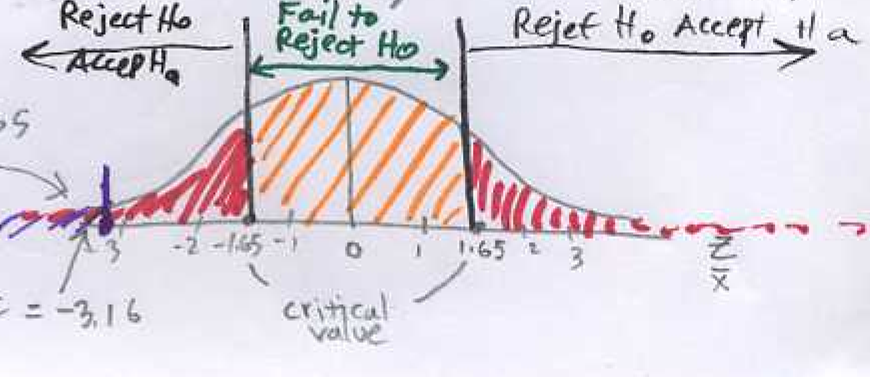
A consumer group says:
"we don't think it takes 8 hours"

$H_0: \mu = 8$ hour
$H_a: \mu \neq 8$ hour

$n = 40$
$\mu_0 = 6.5$ = hypothesized population mean
$\sigma = 3$ hours (sigma known)
Confidence Interval (confidence Coefficient) = .90
alpha = α = risk that population parameter is Not in interval = Risk that our test will reject original statement (H_0) when it is True. = $1 - .9 = .10$
$\alpha/2 = .10/2 = .05$
$\sigma_{\bar{x}}$ = Standard Error = $3/\sqrt{40} = .4743$
z = test statistic = $\dfrac{\bar{x} - \mu_0}{\sigma_{\bar{x}}} = \dfrac{6.5 - 8}{.4743} = -3.16228$
critical value = Dividing line between whether original statement is reasonable or not
= +/- 1 − NORMSINV($\alpha/2$) = 1 − NORSINV(.05) =
= 1.645

P-value = probability of getting Test statistic or more (upper end)
P-value = 2 * NORMSDIST(z test statistic) = 2*NORMSDIST(−3.16)
= .001565

Conclude:
Because −3.16 is not in Interval & because p-value = .001565
.00157 ≤ .10 the original statement is Not reasonable - the pop. mean is not 8 hours



Step 1 List Null & Alternative Hypothesis

$H_0 : \mu = 8$ hours

$H_a : \mu <> 8$ hours

* <> means $\neq$

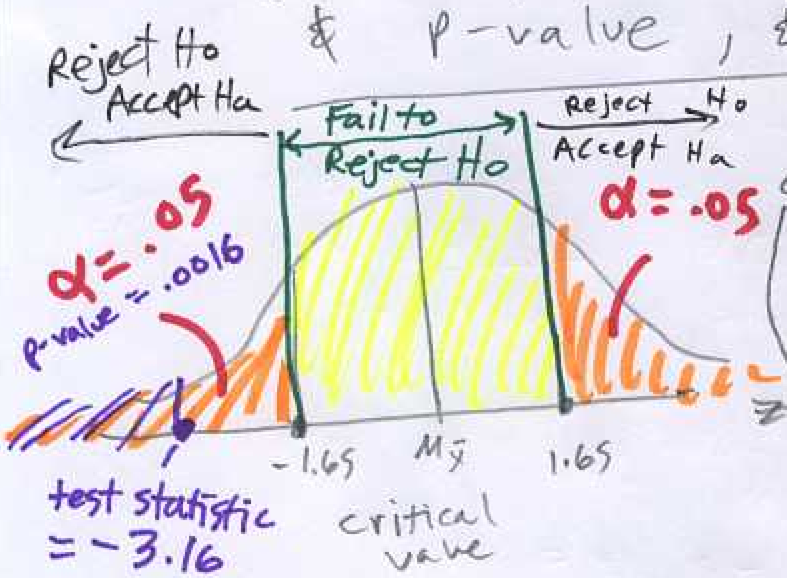
Step 2 Select Alpha

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



Z = -3.16

critical value = -1.65 & 1.65

p-value = .001565

Rejection Rules:

If Z is between +/- 1.65, we fail to reject Ho, otherwise we Reject Ho & Accept Ha.

or

If p-value ≤ α, Reject Ho & Accept Ha, otherwise fail to reject Ho

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

Because -3.16 is not between +/- 1.65 & because .0016 is less than .10, we reject Ho & accept Ho. 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" to see whether they are reasonable:

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 16oz. 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)

Hypothesis:

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

examples page 2

Hypothesis Testing:

- A procedure based on sample evidence and Probability theory to determine whether the hypothesis is a reasonable statement.
- Hypothesis Testing <u>does not prove</u> that the <u>hypothesis is true</u> 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 (MU 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-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, α = .01, then 99 intervals contain mu, 1 does not)

② If pop. data is not Normal, the bigger the n, the more exact.

Pop Normal = any n can be used
Appox. 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 or $H_1 : \mu <> \mu_0$

Then:

$\neq$ = 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.

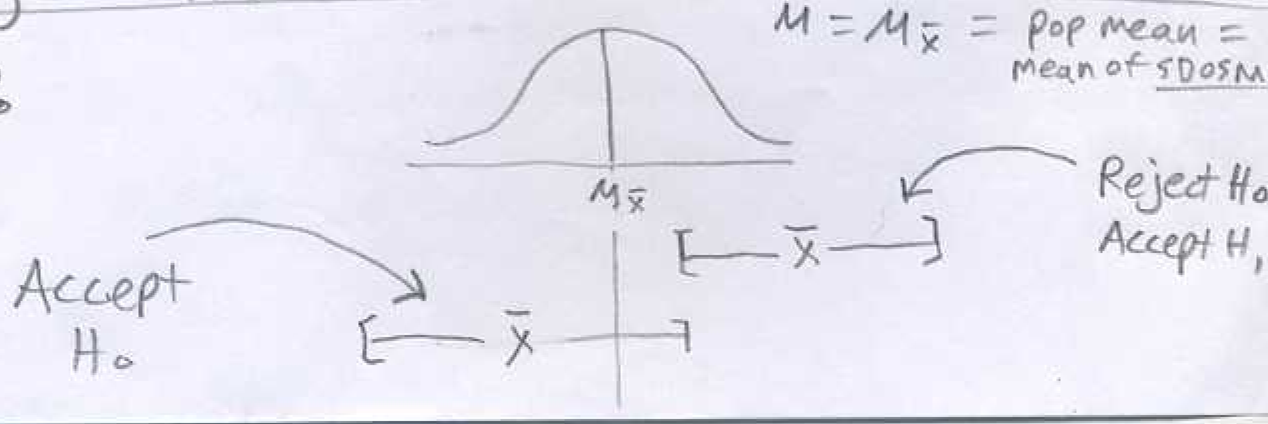
$$\bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$\bar{x}$ = 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

$\mu = \mu_{\bar{x}}$ = pop mean = mean of SDOSM

Example:



Test statistic (z or t) for Hypothesis Testing About a population mean

σ Known	σ Not Known
$z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$	$t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$

μ_0 = hypothesized mean

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

$\bar{x}$ = sample mean

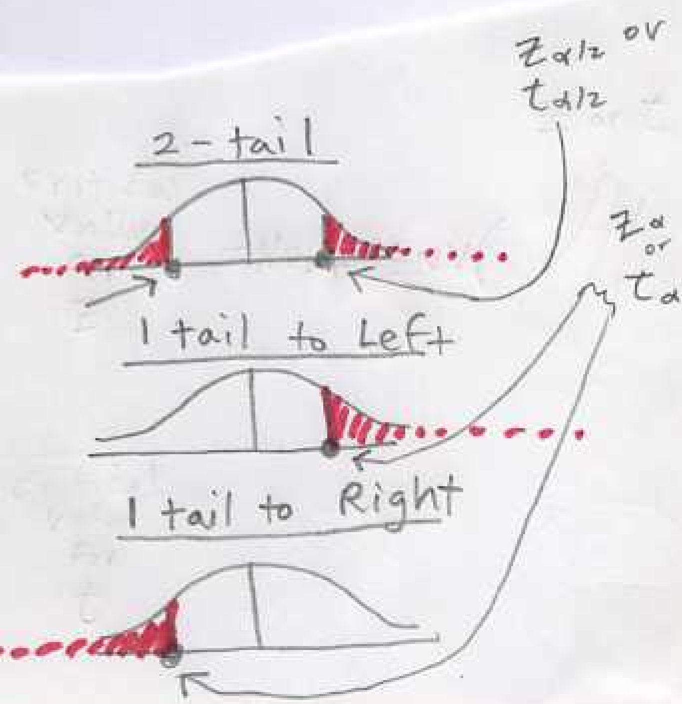
σ = population standard deviation

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

<u>must verify:</u>

(1) Are there fixed # Trials?

(2) Are results Independent?

(3) Does each Trial result in success or Failure?

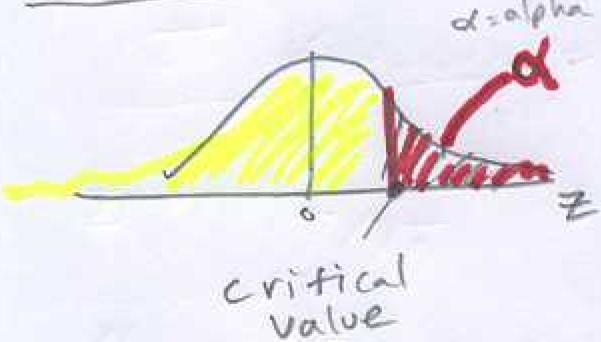
(4) p stay same on each trial?

(5) $n * p > 5$

$n * (1-p) > 5$

Excel Functions

1-tail to right σ known for Z



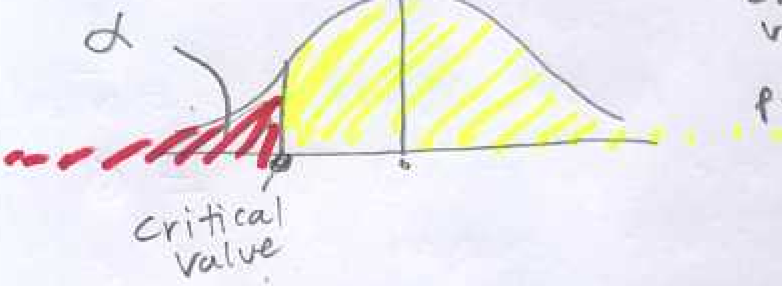
critical value } = 1 – NORMSINV(α)

alpha

p-value = 1 – NORMSDIST(z)

calculated test statistic

1 tail to left σ known for z



critical value } = NORMSINV(α) alpha

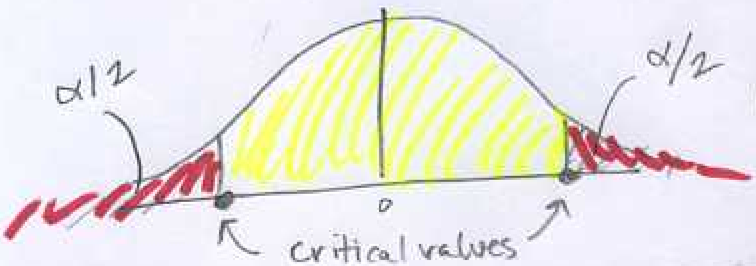
p-value = NORSDIST(z)

calculated test statistic

2 tail σ known for Z

alpha/2

critical value } = +/– NORMSINV(α/2)



p-value = 2 * NORMDIST(z on low end)

z on low end

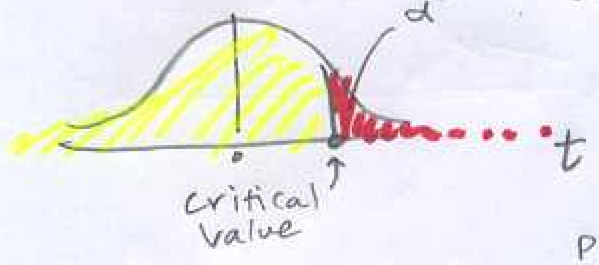
Notes: z can be used for:

1. Mean with σ known
2. Proportions (Binomial Tests Met)

Excel Functions

1 tail to right σ Not Known for t

critical value} = TINV(2*α, 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-value = TDIST(positive t, degree of freedom, 1)

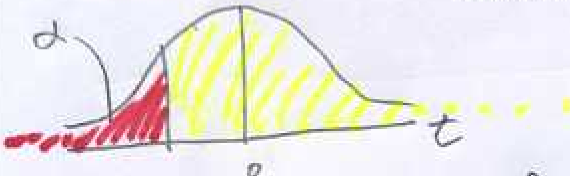
Always must be positive

df

# of tails in test

1 tail to left σ Not Known for t

critical value = − TINV(2*α, degree of freedom)



Because TINV only calculates for upper value.

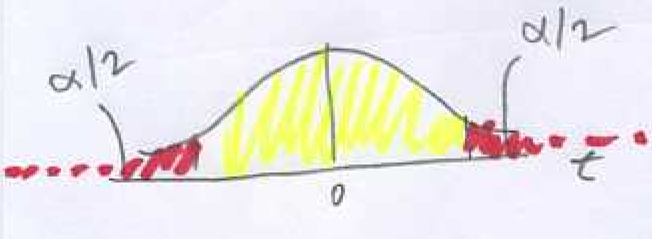
Because TINV always calculates for 2-tail test.

more common → P-value when t is negative} = TDIST(− negative t, Degree of freedom, 1)

less common → P-value when t is postive} = 1 − TDIST(positive t, Degree of freedom, 1)

2 tail σ Not Known for t

critical value} = +/− TINV(α, degree of freedom)



p-value = TDIST(positive t, degree of freedom, 2)

tail