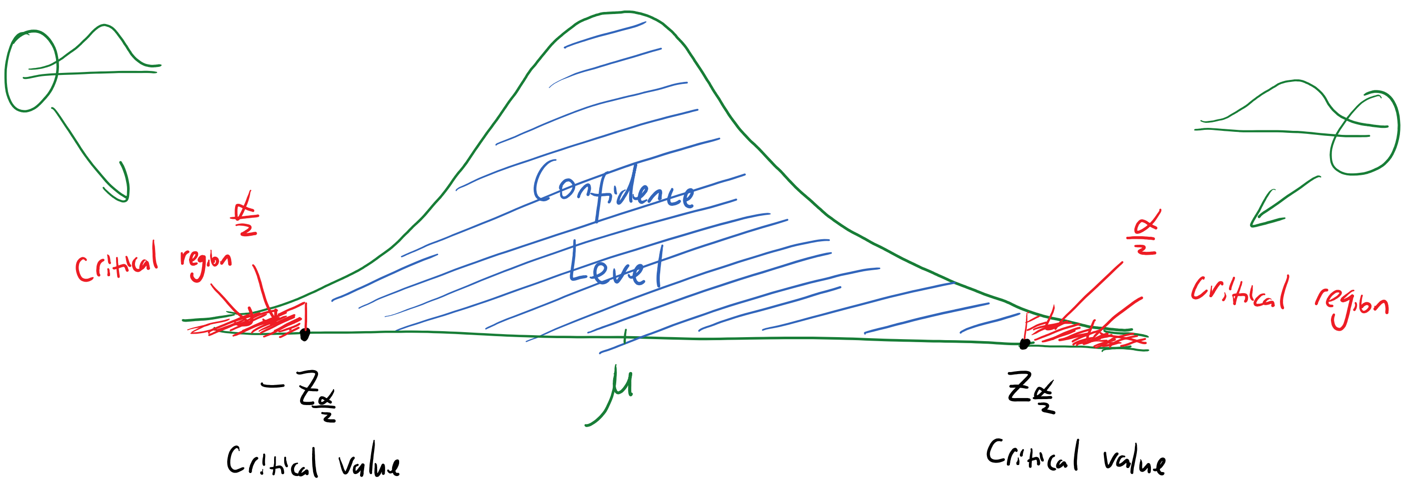
# Hypothesis Testing




Defn: $\alpha$ — Significance level, area of the tails
(Level of significance; signifance)

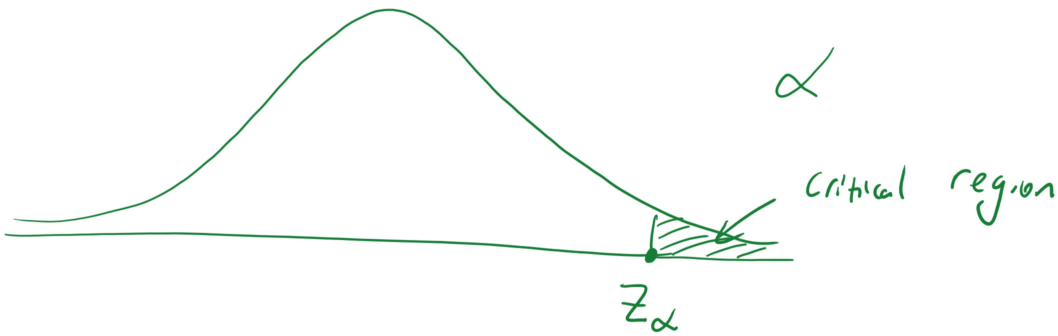
- usually given: 5%, 10%, 1%
- if it's <u>not</u> given, automatically 5%

Confidence Level — $1-\alpha$, the area (probability) in the center
C.L. of the N.D. density curve

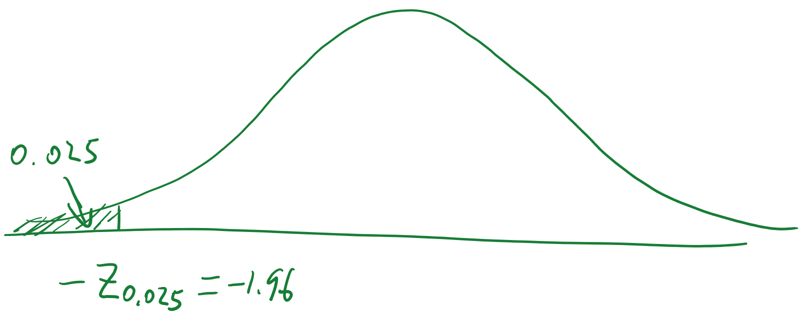
Critical Region — the area of the tails

Critical Value — $Z_{\frac{\alpha}{2}}$ and $-Z_{\frac{\alpha}{2}}$

eg.



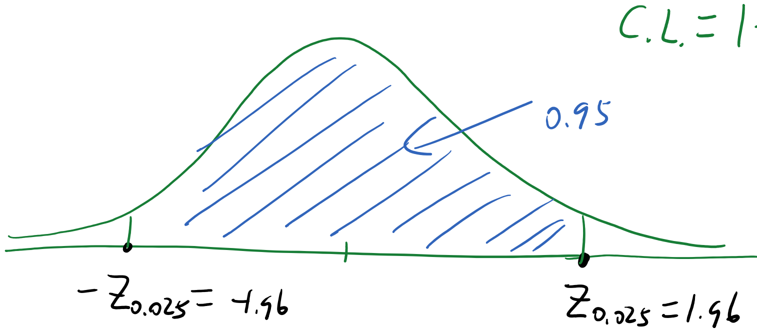

eg.




eg. $\alpha = 0.05$, two tails

Then

$C.L. = 1 - 0.05 = 0.95$

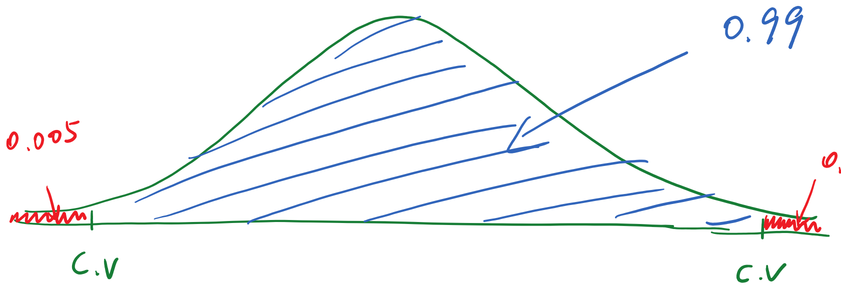



$\frac{0.05}{2} = 0.025$

eg. $\alpha = 0.01$, two tails


$\frac{0.01}{2} = 0.005$

$\frac{0.01}{2} = 0.005$




$-z_{0.005} = -2.58$ $\qquad$ $z_{0.005} = 2.58$

Steps: 1. Find $H_0$ and $H_1$, where $H_0$ is the null hypothesis and $H_1$ is the alternative hypothesis. $H_1$ is supposed to "override" the $H_0$.

old data

new testing data

2. Test $H_1$ by using Test Statistics.

3. Conclude: $H_1$ rejects $H_0$ or $H_1$ fails to reject $H_0$

4. Summarize. (Enough evidence / not enough evidence to support $H_1$ ...)

Claim: Base on the problem, we find $H_0$ first, then $H_1$. Keep in mind that $H_1$ is testing against $H_0$.

eg. $H_0: p = 0.2$ ⟵ general fact

$H_1: p > 0.2$ ⟵ testing against $H_0$

eg. $H_0: p = 0.001$

$H_1: p < 0.001$

eg. $H_0: p = 0.05$

$H_1: p \neq 0.05$

Eg. Assume that 100 babies are born to 100 couples treated with gender selection that is claimed to make girls more likely. We observe 58 girls in 100 babies. Write the hypotheses to test the claim the "with the XSORT method, the proportion of girls is greater than the 50% that occurs without any treatment".

old data

Sol:

$H_0: p = 0.5$

$H_1: p > 0.5$

* Eg. Now we consider the claim that the gender selection increases the likelihood of having a baby girl. Preliminary results from a test of gender selection involved 100 couples who gave birth to 58 girls and 42 boys. Use the given claim and the preliminary results to ~~calculate the~~ test statistic.

"tricky"

likelihood ??

$H_0: p = 0.5$

$H_1: p > 0.5$

increase means more than

(Similarly to majority)

gives $H_0: p=0.5$

Eg. Based on information from the National Cyber Security Alliance, 93% of computer owners believe they have antivirus programs installed on their computers. In a random sample of 400 scanned computers, it is found that 380 of them actually have antivirus software programs. Use the sample data from the scanned computers to test the claim that 93% of computers have antivirus software.

tells what?

if it's 93% or not

Sol:

$H_0: p = 0.93$

Sol:

$$H_0: p = 0.93$$

$$H_1: p \neq 0.93$$

tells what?

if it's 93% or not

not exactly

Eg. In recent years, there has been increasing concern about the health effects of computer terminals. It is known that the miscarriage rate under general conditions is about 20%. A random sample of 650 pregnant women working with a computer 1 to 20 hours per week was taken. For this sample, there were 155 miscarriages. Test the claim that computer terminals detrimentally affect pregnant women with a 0.02 significance level.

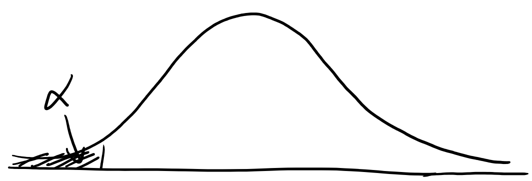
Sol:

$$H_0: p = 0.2 \quad \leftarrow \text{old data}$$

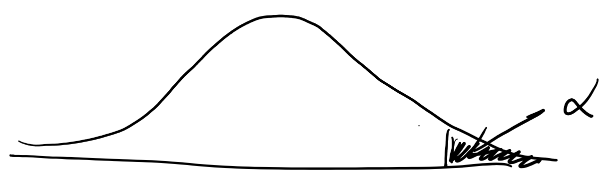
$$H_1: p > 0.2 \quad \leftarrow \text{increasing}$$

Tails

- $<$ means left tail



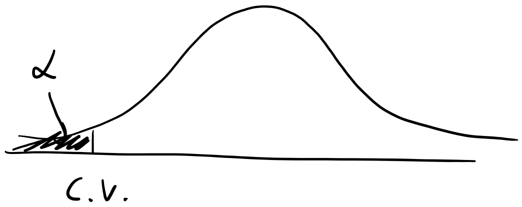

- $>$ means right tail




- $\neq$ means two tails

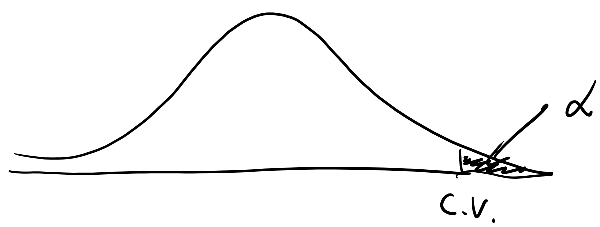

eg. $H_0: p = 0.6$

$H_1: p < 0.6$

⇒ 

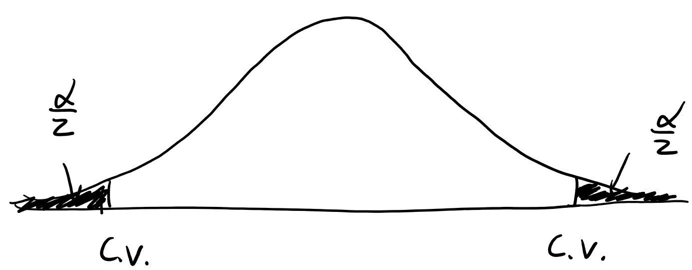

eg. $H_0: p = 0.08$

$H_1: p > 0.08$

⇒ 


eg. $H_0: p = 0.86$

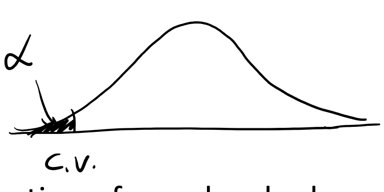
$H_1: p \neq 0.86$

⇒ 


Eg. a. Claim: The proportion of people who have smoked once is less than 0.6. $<$

$H_0$: $p = 0.6$

$H_1$: $p < 0.6$

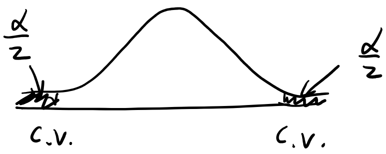



b. Claim: The proportion of people who have smoked once is exactly 0.6. Yes or no

$\neq$

$H_0$: $p = 0.6$

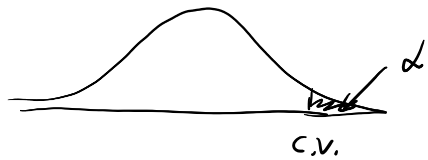
$H_1$: $p \neq 0.6$




c. Claim: The proportion of people who have smoked once is at least 0.6.

$H_0$: $p = 0.6$

$H_1$: $p > 0.6$




more than

$>$