Lecture: 05

Hypothesis Testing:
Critical Region and P-Value

(PART02)

RECAP: The significance level (α)

α represents the **threshold below which you** would reject the null hypothesis (probability value used to define the (unlikely) sample outcomes) **if the null hypothesis is true.**

- It is the probability of making a Type I error in hypothesis testing.

Interpretation: If you set your significance level at α = 0.05 (5%), it means you are willing to accept a 5% chance of incorrectly rejecting a true null hypothesis (a false positive).

Common Values for α : are 0.05 (5%) and 0.01 (1%), with other levels possible depending on the desired balance between Type I and Type II errors.

Topics we will cover

1. Concept of Critical Region
2. Critical Value
3. One Tailed Test - Left and Right Tailed
4. Two Tailed Test
5. Concept of P-Value

Next: Set Criteria for the Decision

After determining the significance level, the next step is to calculate the test statistic.

- Use one of two approaches to make your decision about the null hypothesis:

  (1) compare the test statistic to the critical value, or

  (2) calculate the p-value

Based on this, decide whether to reject or fail to reject the null hypothesis.

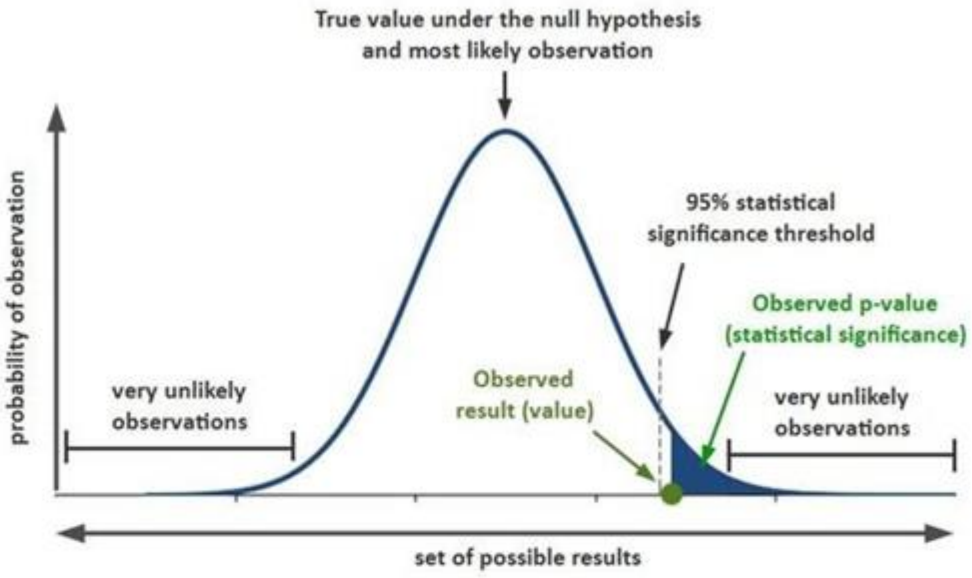
(1) Critical Region and Critical Value

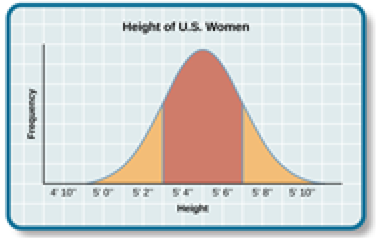
Critical Region & Critical Value

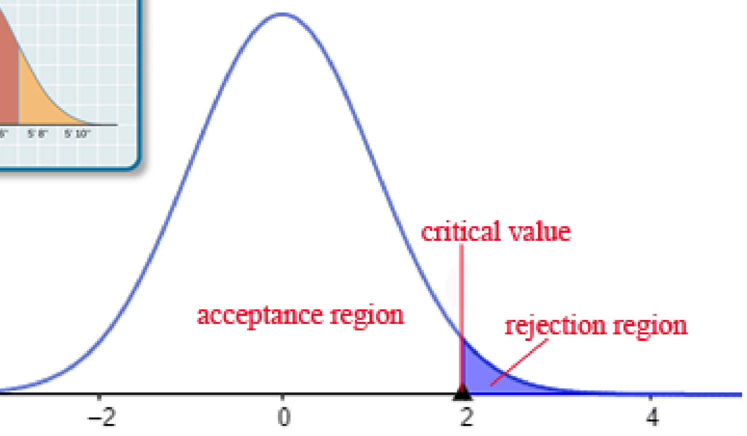
The **critical region (or region of rejection)** is the range of test statistic values where we reject the null hypothesis because the data is too extreme.

- Its size depends on the significance level (α), which sets how much evidence is needed to reject Ho.
  - The size of this region changes with different significance levels (e.g., 0.05 or 0.01).

The **critical value** is the boundary/cutoff point separating the critical region from other values—crossing it means rejecting the null hypothesis.

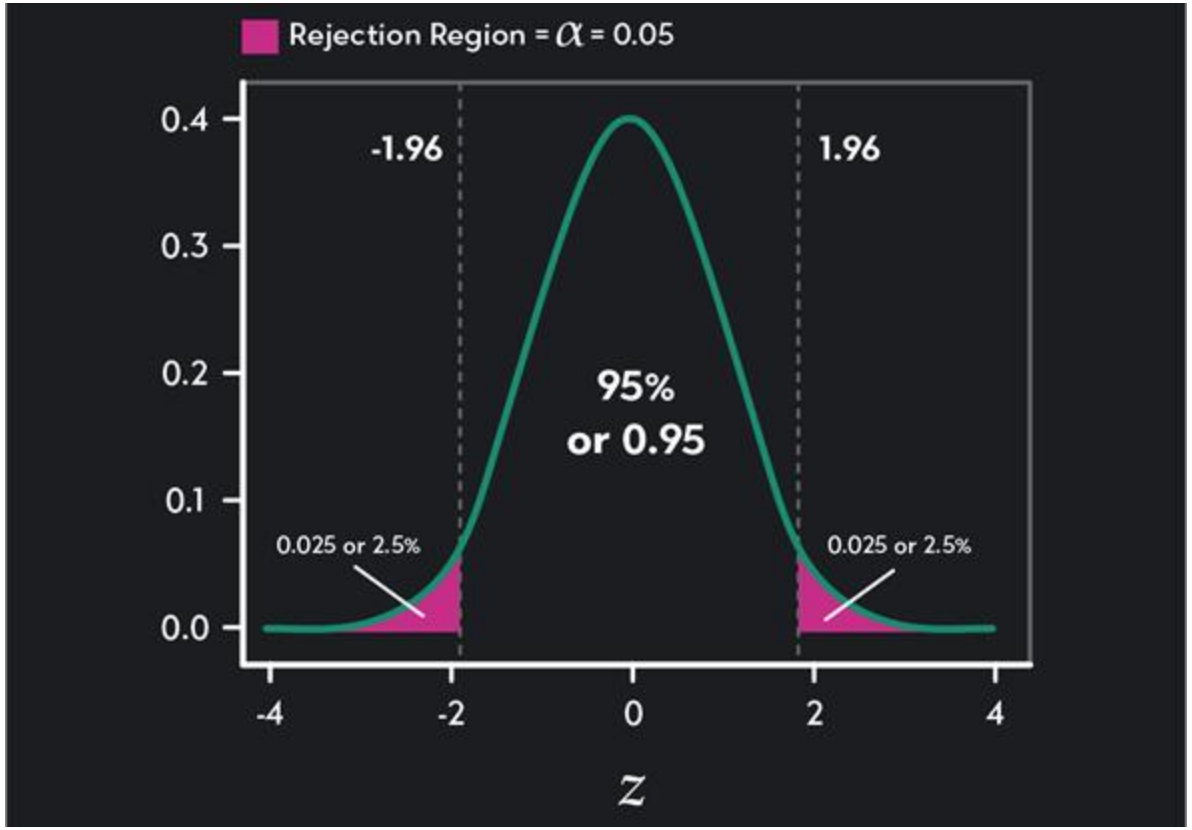






Critical Region (also known as the rejection region)

The figure below shows how the critical values mark the boundaries of two rejection regions (shaded in pink).

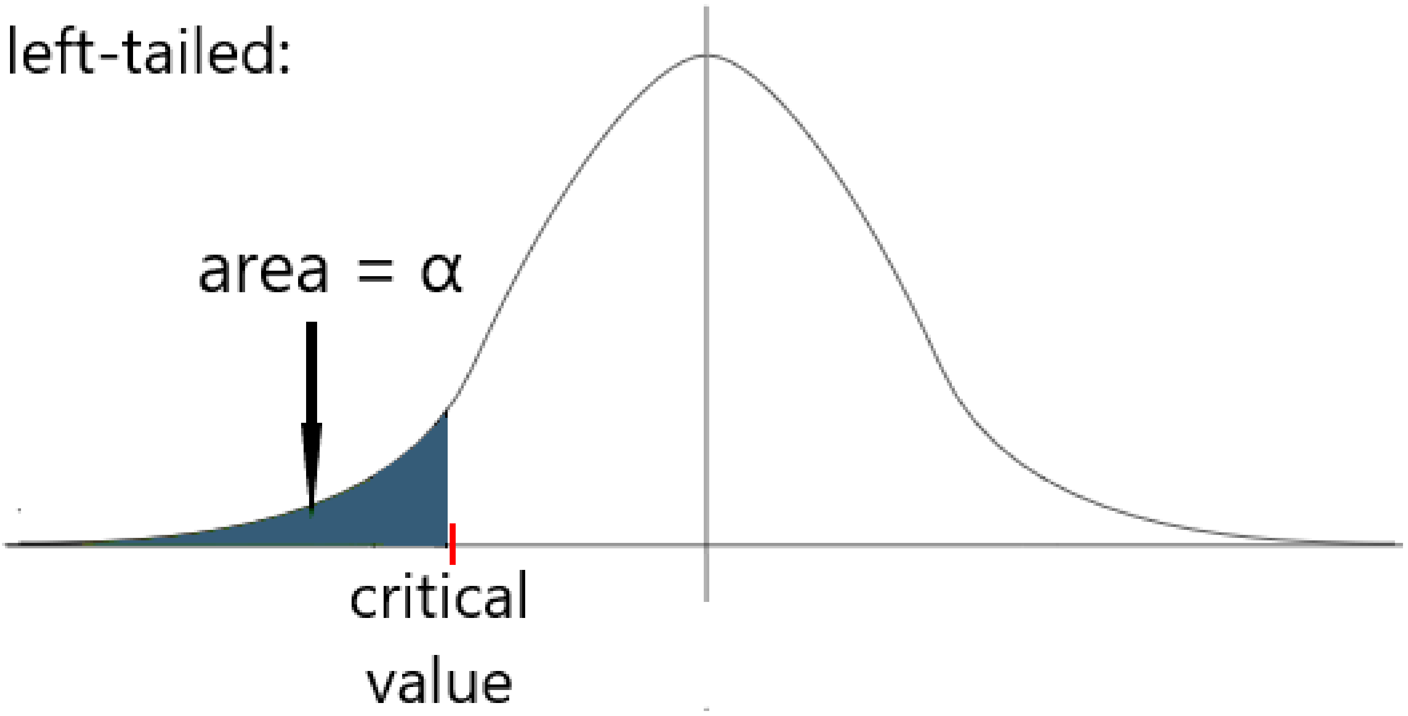


In this figure, any test result greater than 1.96 falls into the rejection region in the distribution's right tail, and any test result below -1.96 falls into the rejection region in the left tail of the distribution.

Note: we considered critical value 1.96 here. The critical value can be found using the appropriate table or statistical software for the distribution and degrees of freedom.

What Is a Critical Value C?

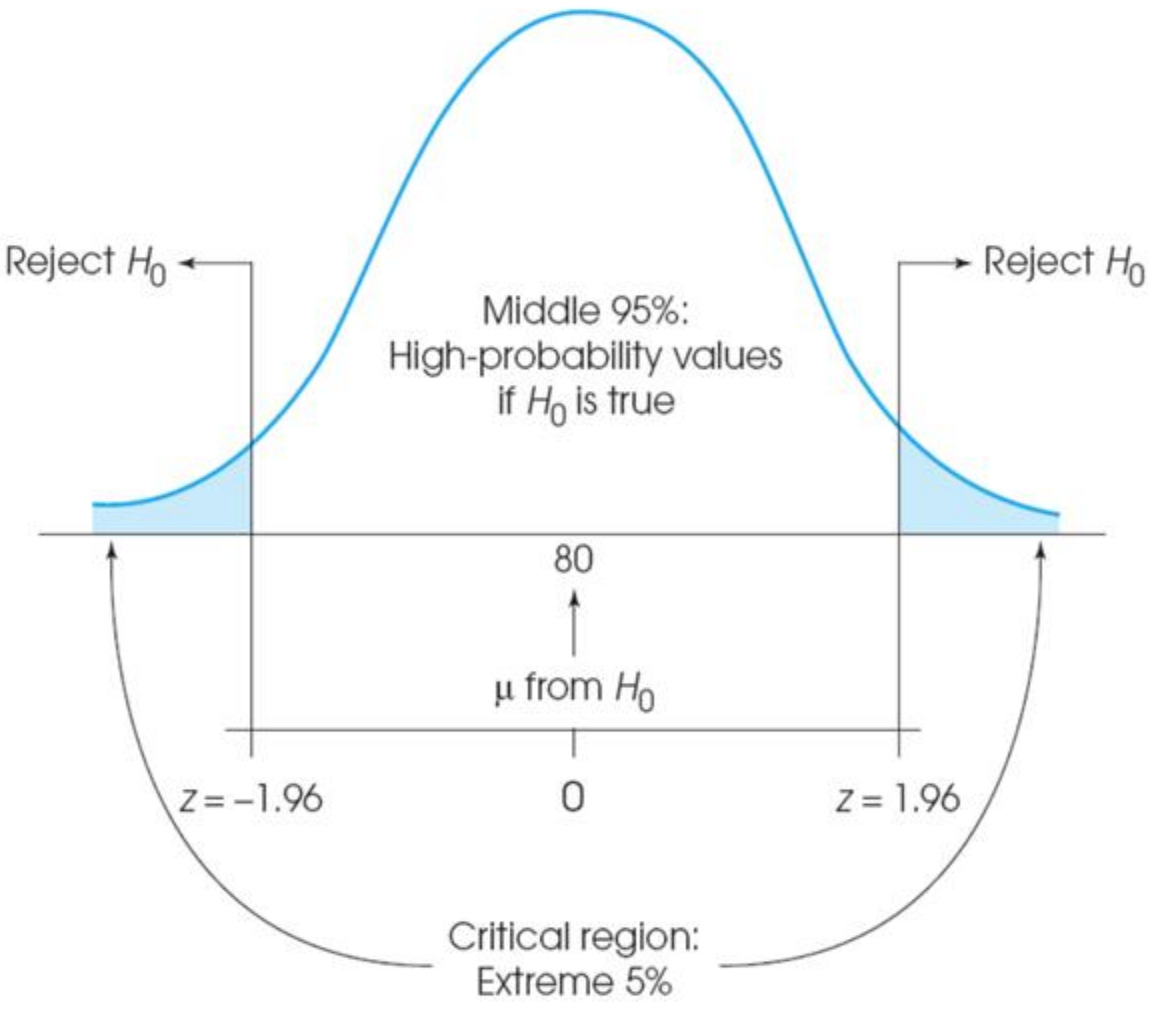
A number (cut-off point/ thresholds/boundary point) that **defines the rejection region** of a hypothesis test and is used to decide whether to reject the null hypothesis.



Critical Region Boundaries

In a hypothesis test, you compare the *calculated test statistic* to the *critical value(s)* to make a decision:

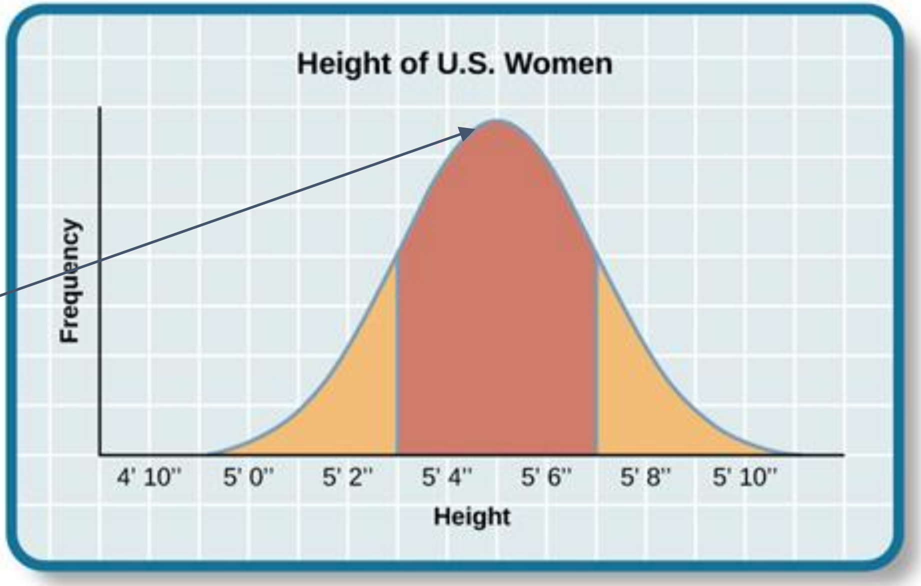
- If the test statistic falls in the critical region (beyond the critical value), you reject the null hypothesis.
- If the test statistic falls outside the critical region, you fail to reject the null hypothesis.



Example: Human Heights

Suppose we're investigating whether workers are selected based on their height.

Ques: If an employer selected someone, what is our most *likely* height?



Question: Based on the setup of the null hypothesis, where would we expect to observe rejection of the null hypothesis in our analysis?

If heights are *normally distributed*, the most likely worker we got is the mode.

Set up: Null Hypothesis (H0): The avg height of selected worker is just like the most common height in the human population.

Consider the extreme regions representing test statistic values that are unlikely to occur (far from the mean) if the null hypothesis is true.

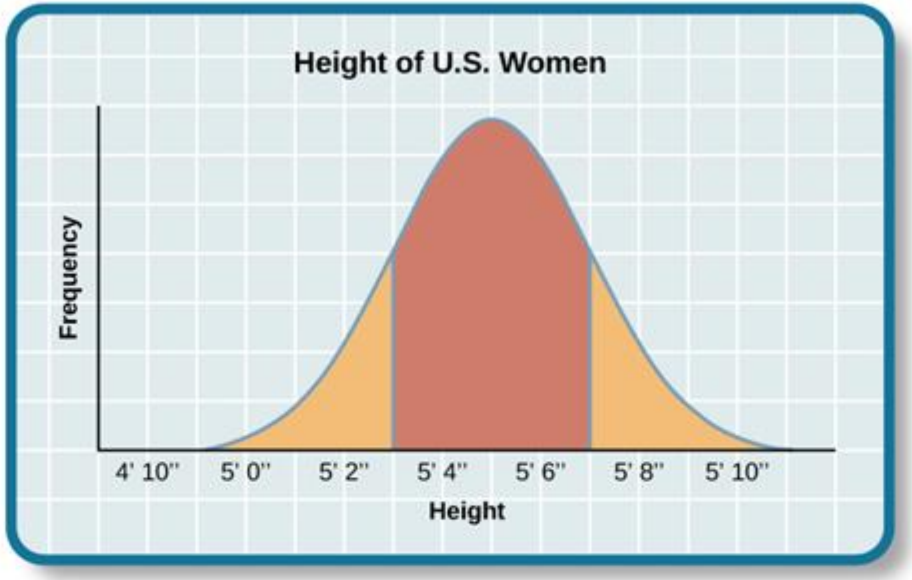
Location of the critical region:

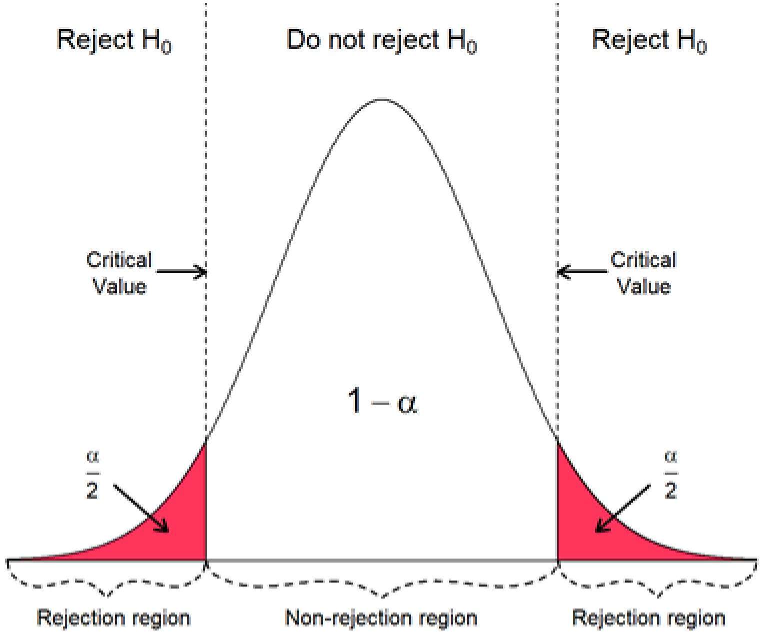
Depends on the significance level (α) and the type of test (one-tailed or two-tailed).

- **In a one-tailed test,** the critical region is on one side (tail) of the distribution.

- **In a two-tailed test,** the critical region is split between both ends (tails) of the distribution.

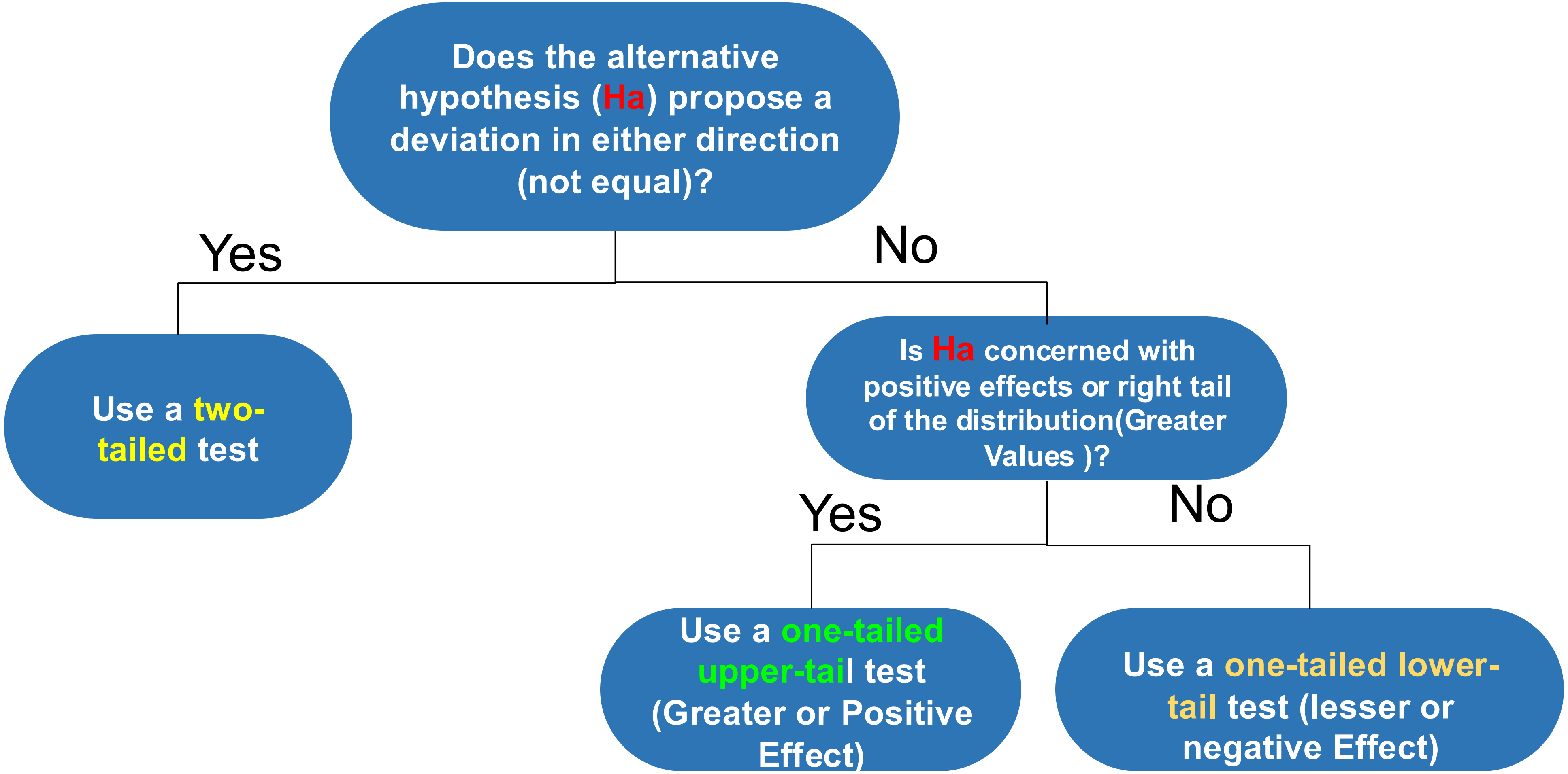
The choice depends on the direction of the research hypothesis.





Is it a one-tailed test or a two-tailed test?

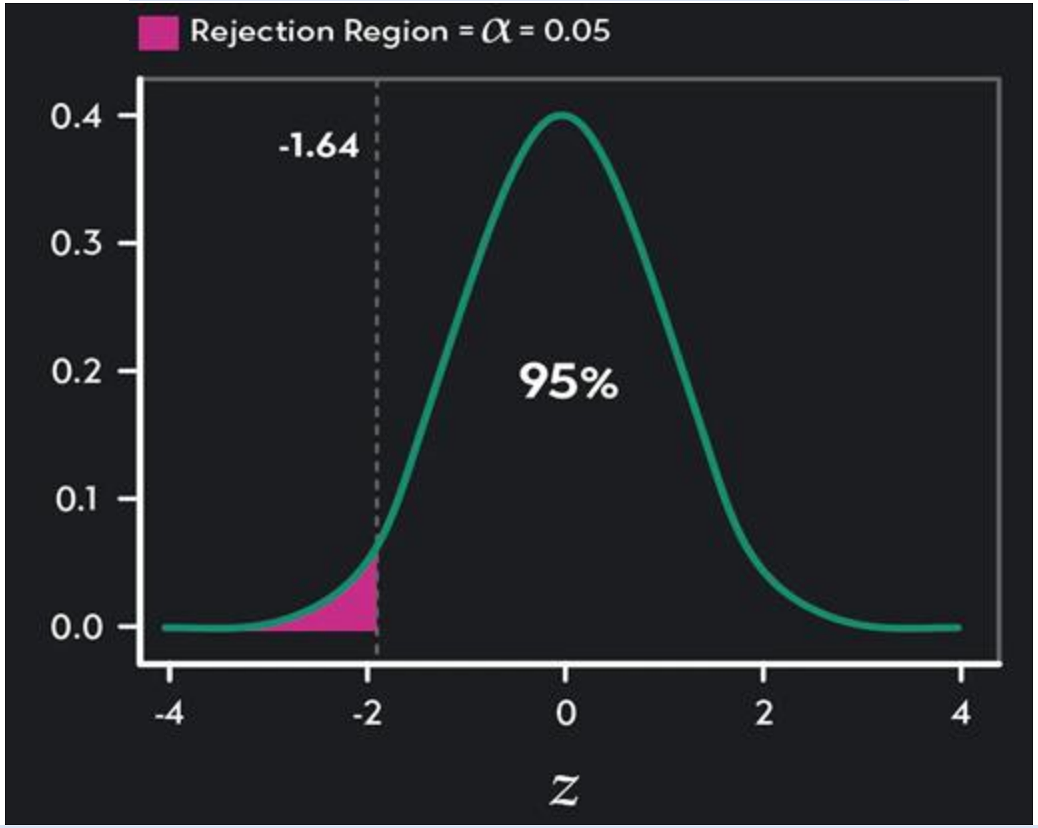
Hypothesis tests can be one-tailed or two-tailed, depending on the alternative hypothesis Ha.



One Tailed Test

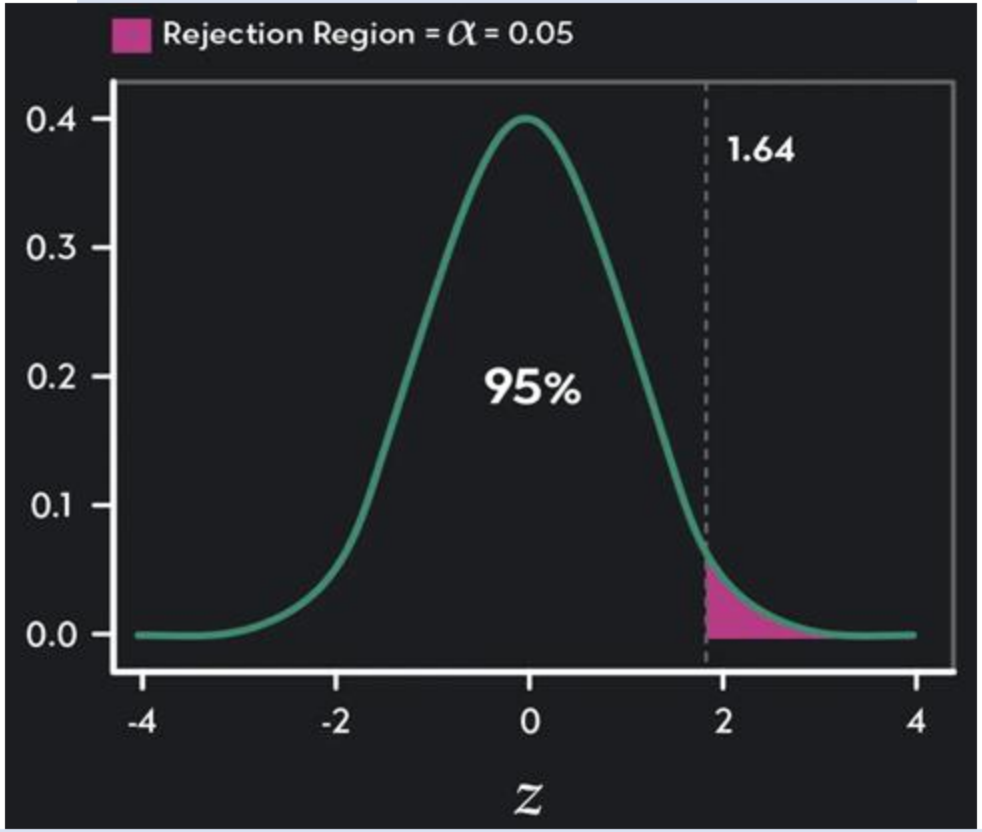
A one-tailed test has **one rejection region** (either in the **right** or **left tail** of the distribution) and **one critical value**. The total α probability goes into that <u>one side</u>.

H0: μ ≥ μ0 ; Ha: μ < μ0



Left-tailed/lower-tailed test: critical value and rejection region will be in the left tail of the distribution

H0: μ ≤ μ0 vs. Ha: μ > μ0



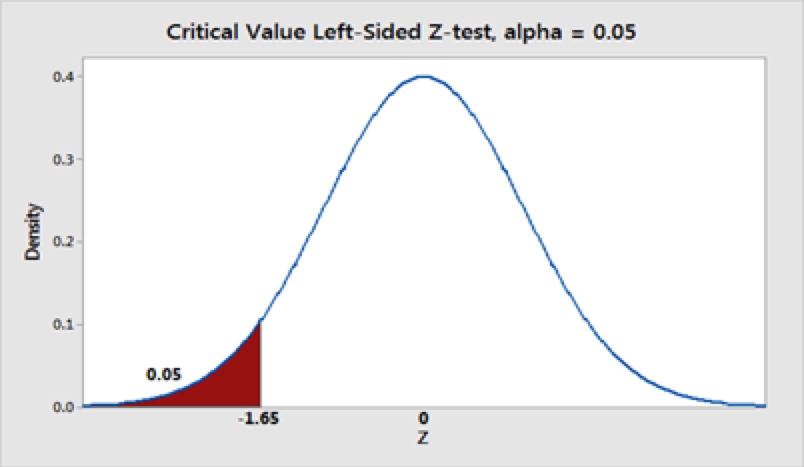
Right-tailed/upper-tailed test: critical value and rejection region will be in the right tail of the distribution

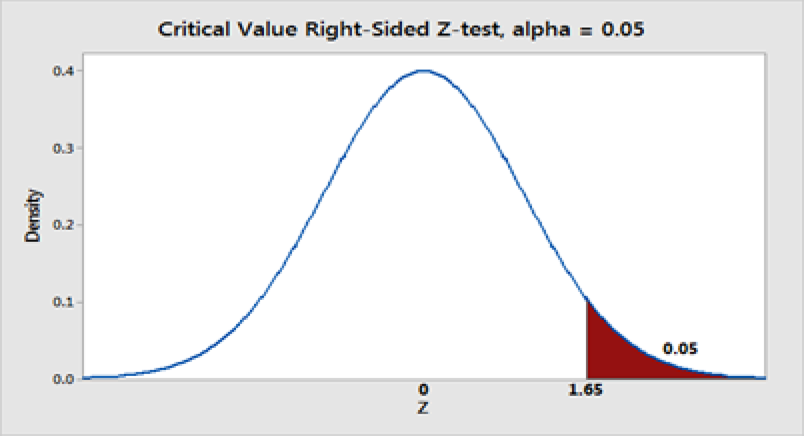
One Tailed Test is one sided!

The critical area of a distribution is one-sided so that it is either greater than or less than a certain value, but not both.

If the sample being tested falls into the one-sided critical area, the alternative hypothesis will be accepted instead of the null hypothesis.

H0: μ ≥ μ0 ; Ha: μ < μ0

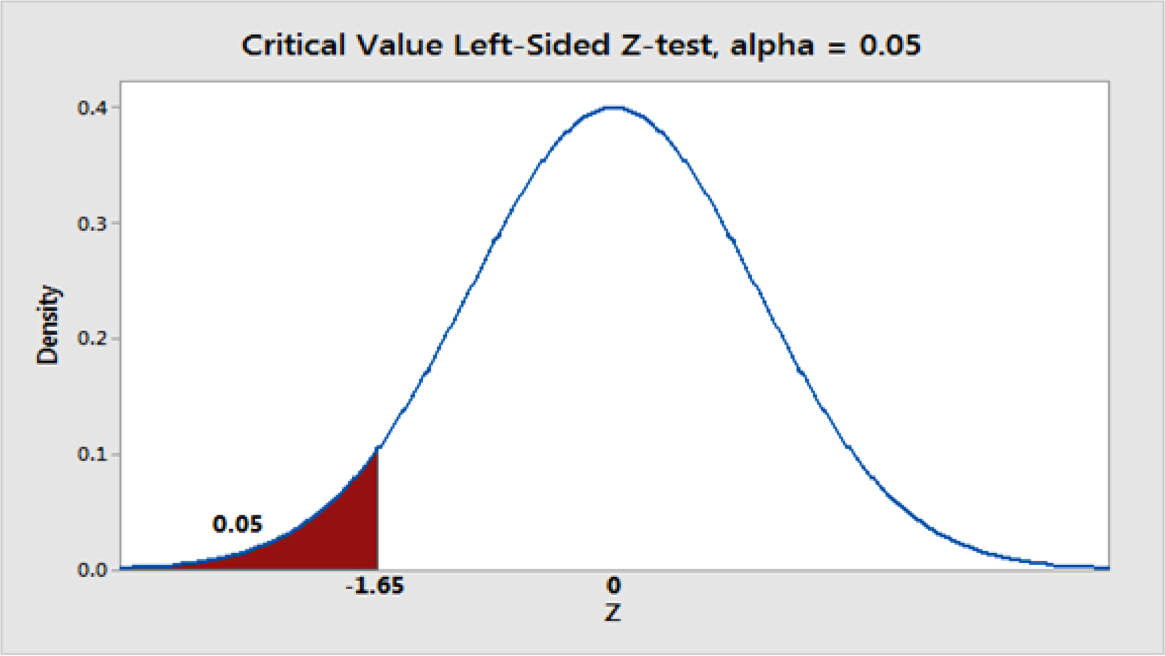




H0: μ ≤ μ0 vs. Ha: μ > μ0

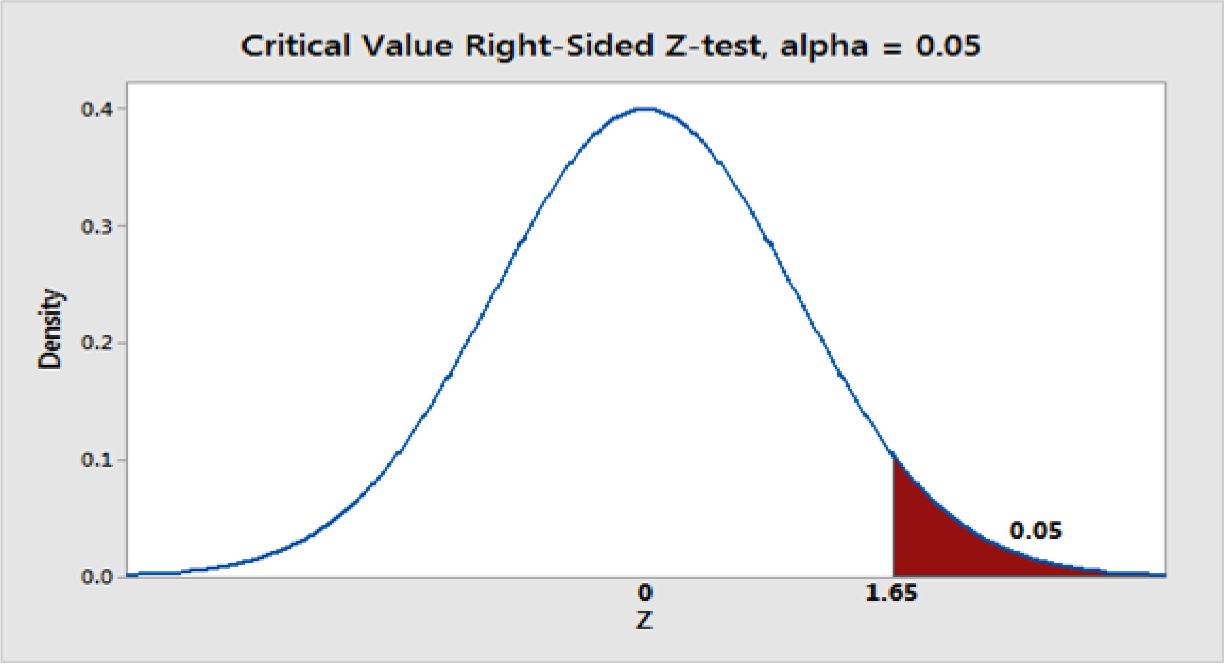
Two types of One Tailed Tests

Left-Tailed Test: rejection region is located to the extreme left of the distribution. Ha contains the condition: **$\mu < \mu 0$**.



H0: $\mu \geq \mu 0$; Ha: $\mu < \mu 0$

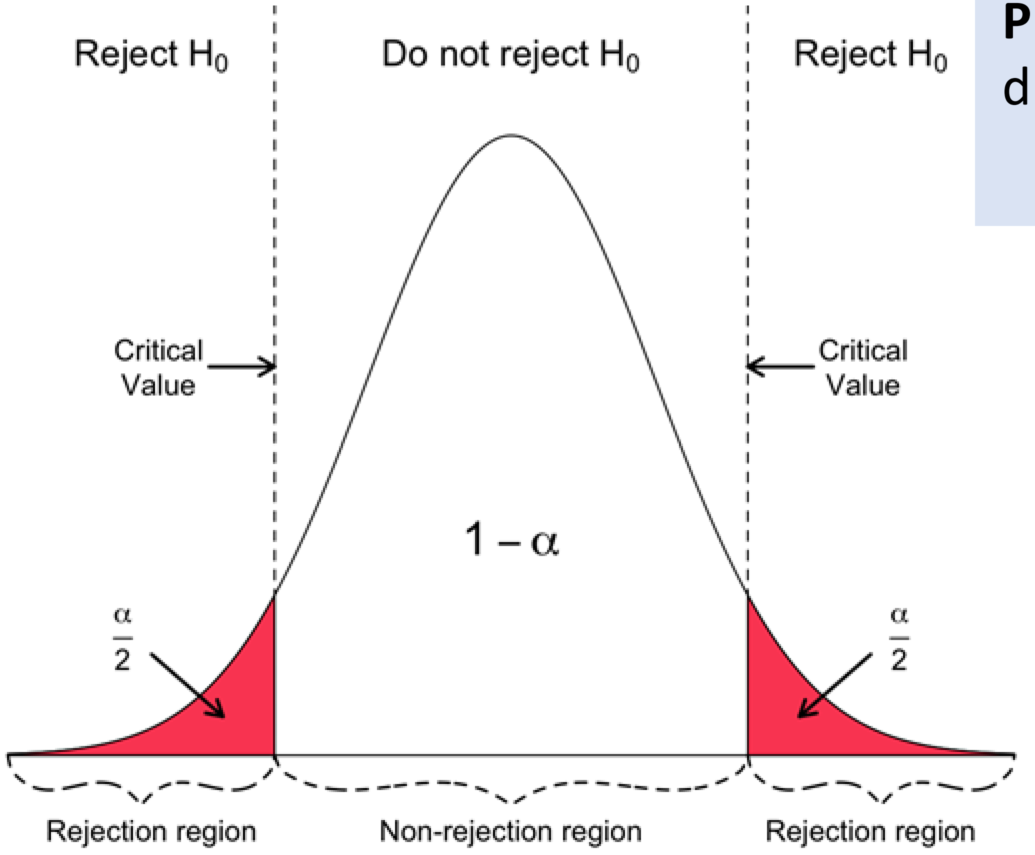
Right-Tailed Test: rejection region is located to the extreme right of the distribution. Ha contains the condition: **$\mu > \mu 0$**.



H0: $\mu \leq \mu 0$ vs. Ha: $\mu > \mu 0$

Example: Testing if a new drug increases blood pressure (right-tailed) or decreases blood pressure (left-tailed) compared to a baseline value

Two-Tailed Test



Used when the direction of the effect is not specified in the hypothesis.

- Checks for **significance in both directions**, unlike a one-tailed test.
- Useful when you are interested in detecting any significant difference, regardless of whether it is positive or negative.

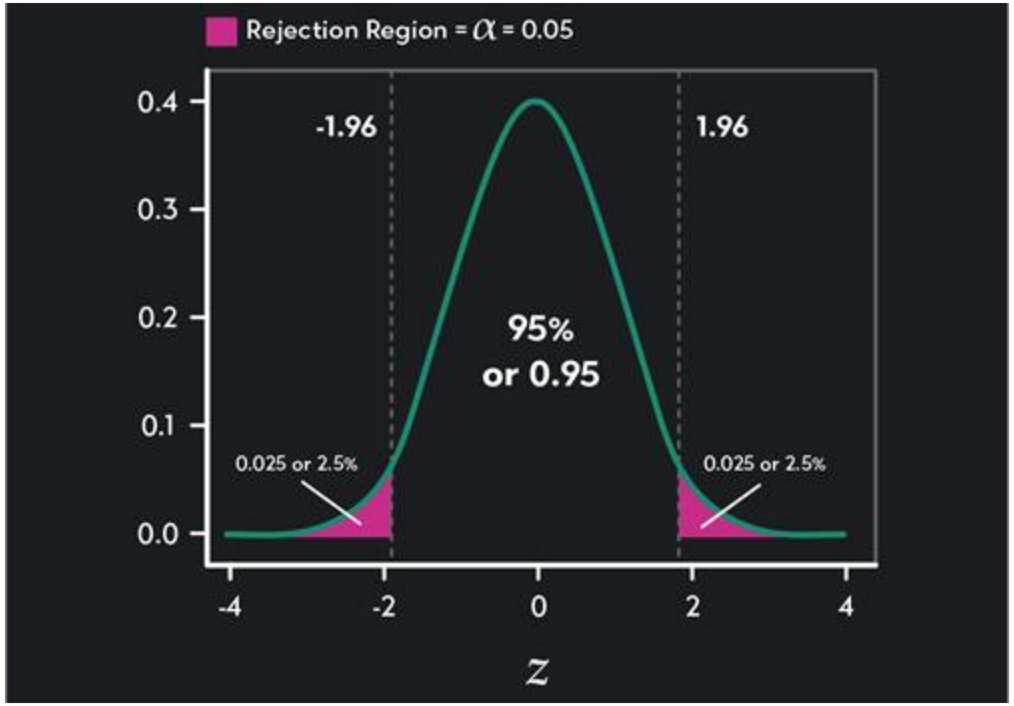
Purpose: determine if there is a significant difference in either direction from the specified value

Example: Testing if a new drug has an effect on blood pressure, without specifying whether the effect is an increase or a decrease.

H0: μ = μ0 — The mean blood pressure with the new drug is equal to the baseline mean.

Ha: μ ≠ μ0 — The mean blood pressure with the new drug is different from the baseline mean (either higher or lower).

Two Tailed Test



We divide the rejection region into **two equal parts**:

1. One in the right tail and
2. One in the left tail of the distribution.

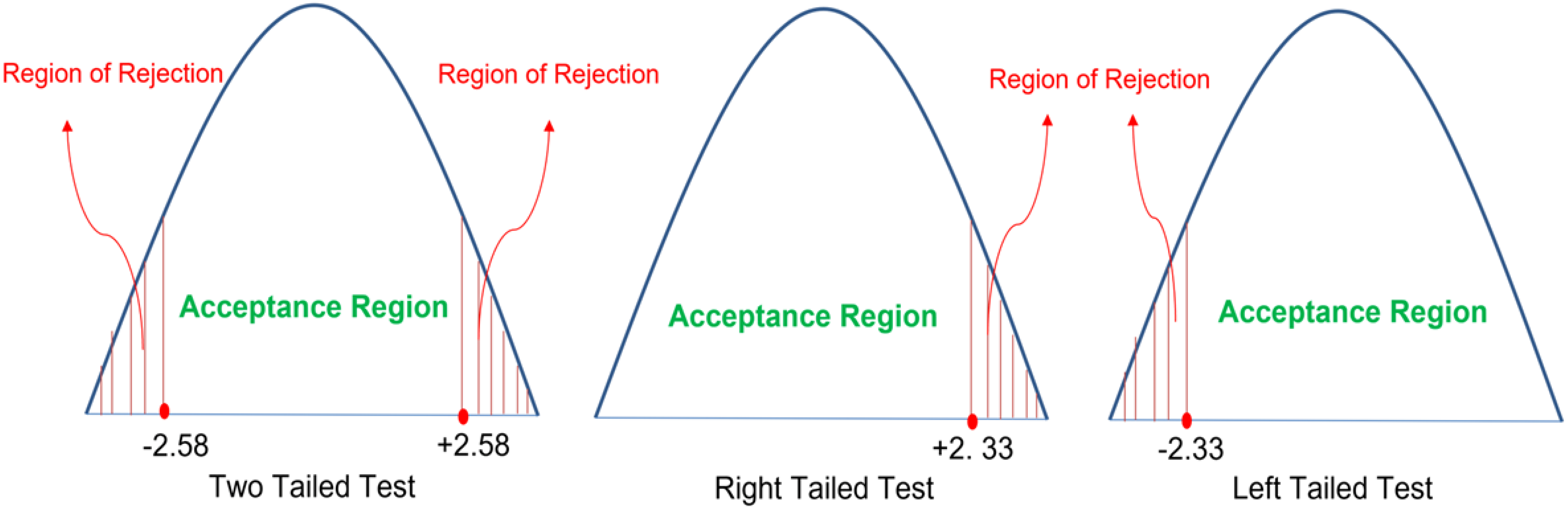
Each part contains an area of α/2.

For example, in a two-tailed test with a significance level of 0.05, each rejection region will contain 0.05/2 = 0.025 = 2.5% of the area under the distribution.

- Because we split the rejection region, a two-tailed test has two critical values.

Critical Regions in Hypothesis Testing

Level of Significance	1%	5%	10%
Two Tailed Test	-2.58, +2.58	-1.96, +1.96	-0.645, +0.645
Right Tailed Test	+2.33	+1.645	+1.28
Left Tailed Test	-2.33	-1.645	-1.28



With Level of Significance = 1%

Example: Critical Value

You are testing whether the **average weight of a batch of apples is <u>different</u> from 150 grams (μ)**. You will use a **1% significance level (α = 0.01)**, and you assume the **population standard deviation (σ)** is 10 grams. You collect a **sample of 30 apples** and find that the **sample mean (x') is 152 grams**.

Ho: $\mu = 150$
H1: $\mu \neq 150$

$\alpha = 0.01$
$\sigma = 10$

Choose the Appropriate Test and Test Statistics?

Calculate the Test Statistic (Z): The formula for the Z-statistic is:

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

$$Z = \frac{152 - 150}{\frac{10}{\sqrt{30}}} = \frac{2}{\frac{10}{5.477}} = \frac{2}{1.826} = 1.095$$

Determine the Critical Value: $\mathbf{Z_1 = -2.58}$ and $\mathbf{Z_2 = +2.58}$ (check previous slide- decided using Z table)

Since **1.095** is **within the range** of **-2.576** to **+2.576**, we **fail to reject** the null hypothesis.

How to Find a Critical Value

Unfortunately, the formulas for finding critical values are very complex. Typically, you don’t calculate them by hand. You can use some statistical software or statistical tables (Z-table, T distribution table, Chi-square table, F-table) to find them.

ALTERNATIVE: P-Value (Another approach to evaluating the significance of test results.)

(2) P-Value

Recap

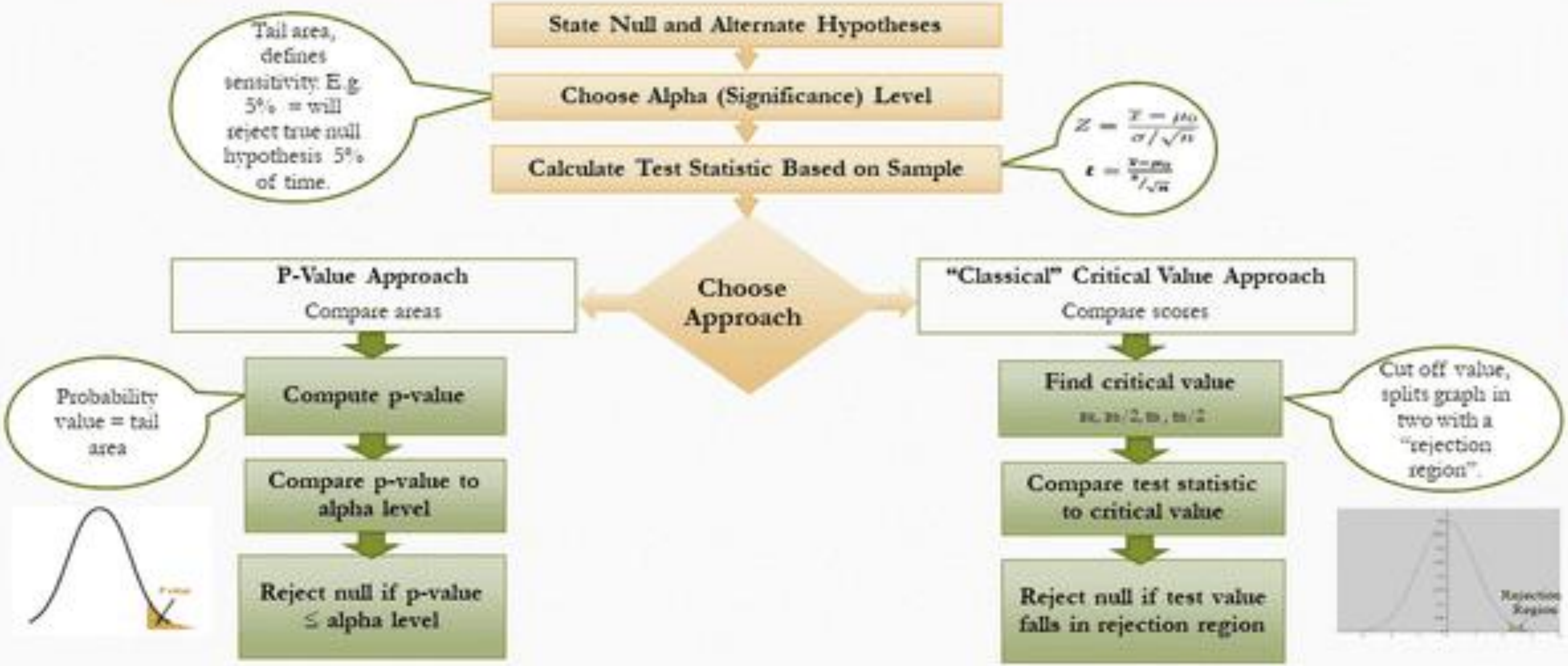
Recall: The **Critical Value Method**:

- We compare our test Stastics and the critical value(s)
- If the test statistic is more extreme than the critical value, it was significantly high or low, we reject the null hypothesis.

Now Come Alternative Approach to help us assessing the evidence against the null hypothesis (H_0): **P-Value Method**

- Rather than, comparing the values in the horizontal axis, **we compare the probabilities** (areas in the tails)

P-Value vs Critical Value



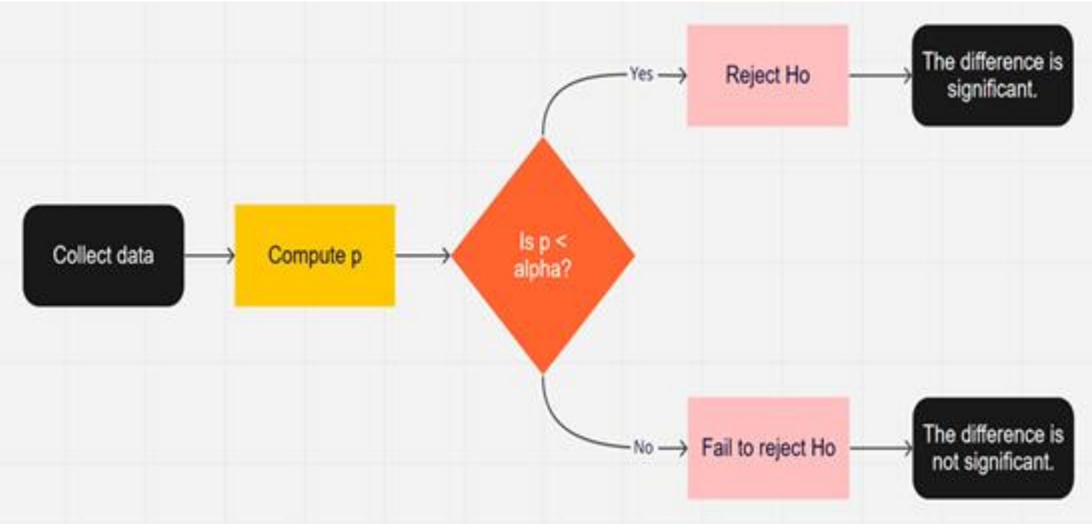
The Use of P –Values in Decision Definition Making

p-value: the probability of observing your test statistic (or something more extreme) if the null hypothesis (H_0) is true.

Smaller p-values indicate stronger evidence against the null hypothesis.

Compare the test statistic with the critical region using p-value:

- If p-value ≤ α, reject the null hypothesis.
- If p-value > α, fail to reject the null hypothesis.



P-Value Procedure

- Classify the test as "two tailed, left tailed or right tailed"
- Find the P-Value using the area associated with the test statistic
  - If test is two tailed, P is double the area in the nearest tail
  - If test is left tailed, P is the area in the left tail
  - If test is right tailed, P is the area in the right tail
  - We can use associated table (Z-table, T-table etc.) of test statistic to find the P-Value.

Make a Decision:

- If P-Value <= α; the value of the test statistic is significantly high or low, so we **reject the null hypothesis.**
- If P-Value >α; the value of the test statistic is likely given our assumption, so we f**ail to reject the null hypothesis**.

P-Value - Example

Suppose you are testing the following hypothesis at a significance level (α) of 5% and you got the p-value as 3%, and your sample statistic (sample mean) is $\bar{x} = 25$

H_0: $\mu = 20$ (The population mean is 20)

H_1: $\mu > 20$ (The population mean greater than 20)

Given, α=5%, we are fine to reject our null hypothesis 5 out of 100 times even though it is true.

P-value is 3% which is less than α
$0.03 < 0.05 \rightarrow$ We reject the null Hypothesis (extremely strong evidence against the null hypothesis.)

What about P-Value is 6%?

Example: Now use P Value Approach

You are testing whether the **average weight of a batch of apples is <u>different</u> from 150 grams (μ)**. You will use a **1% significance level (α = 0.01)**, and you assume the **population standard deviation (σ)** is 10 grams. You collect a **sample of 30 apples** and find that the **sample mean (x') is 152 grams**.

Ho: $\mu = 150$
H1: $\mu \neq 150$
$\alpha = 0.01$
$\sigma = 10$

Calculate the Test Statistic (Z): The formula for the Z-statistic is:

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

$$Z = \frac{152 - 150}{\frac{10}{\sqrt{30}}} = \frac{2}{\frac{10}{5.477}} = \frac{2}{1.826} = 1.095$$

Find P-Value: What is the p-value that corresponds to this z-score?

From the Z-table or calculator, the area to the right of Z = 1.095 is about 0.1368 (NO NEED TO FIND THIS DURING EXAM, THIS WILL BE GIVEN).

Since it's a two-tailed test, **P = 2 * 0.1368 =0.274**

Since **P-Value (0.2736) > α (0.01), we fail to reject Ho**

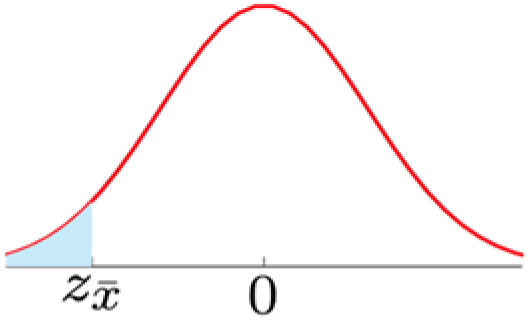
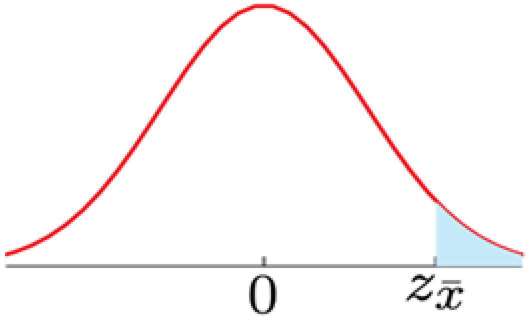
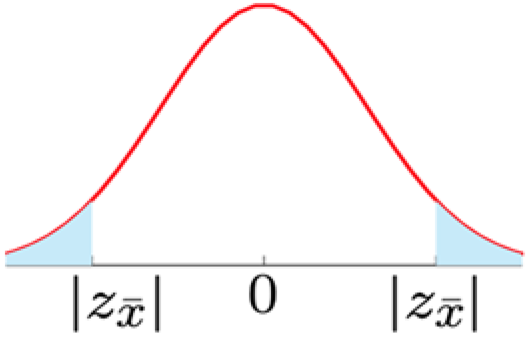
PRACTICE

WRITING NULL AND ALTERNATIVE HYPOTHESIS

The school board claims that atleast 60% of students bring a phone at school. A teacher believe that this number is too high and randomly samples 25 students to test at at level of significance of .02

Write down **H0**, **H, C, alpha.**

P-Values and Types of Tests:

Graph	Test	Conclusion
	1. Left-tailed Test $H_0 : \mu = k \quad H_1 : \mu < k$ P-value $= P(z < z_{\bar{x}})$ This is the probability of getting a test statistic as low as or lower than $z_{\bar{x}}$	If P-value $\leq \alpha$, we reject H_0 and say the data are statistically significant at the level α. If P-value $> \alpha$, we do not reject H_0.
	2. Right-tailed Test $H_0 : \mu = k \quad H_1 : \mu > k$ P-value $= P(z > z_{\bar{x}})$ This is the probability of getting a test statistic as high as or higher than $z_{\bar{x}}$	If P-value $\leq \alpha$, we reject H_0 and say the data are statistically significant at the level α. If P-value $> \alpha$, we do not reject H_0.
	3. Two-tailed Test $H_0 : \mu = k \quad H_1 : \mu \neq k$ P-value $= 2P(z > \lvert z_{\bar{x}} \rvert)$ This is the probability of getting a test statistic either lower than $\lvert z_{\bar{x}} \rvert$ or higher than $\lvert z_{\bar{x}} \rvert$	If P-value $\leq \alpha$, we reject H_0 and say the data are statistically significant at the level α. If P-value $> \alpha$, we do not reject H_0.

Extra Reading Slides

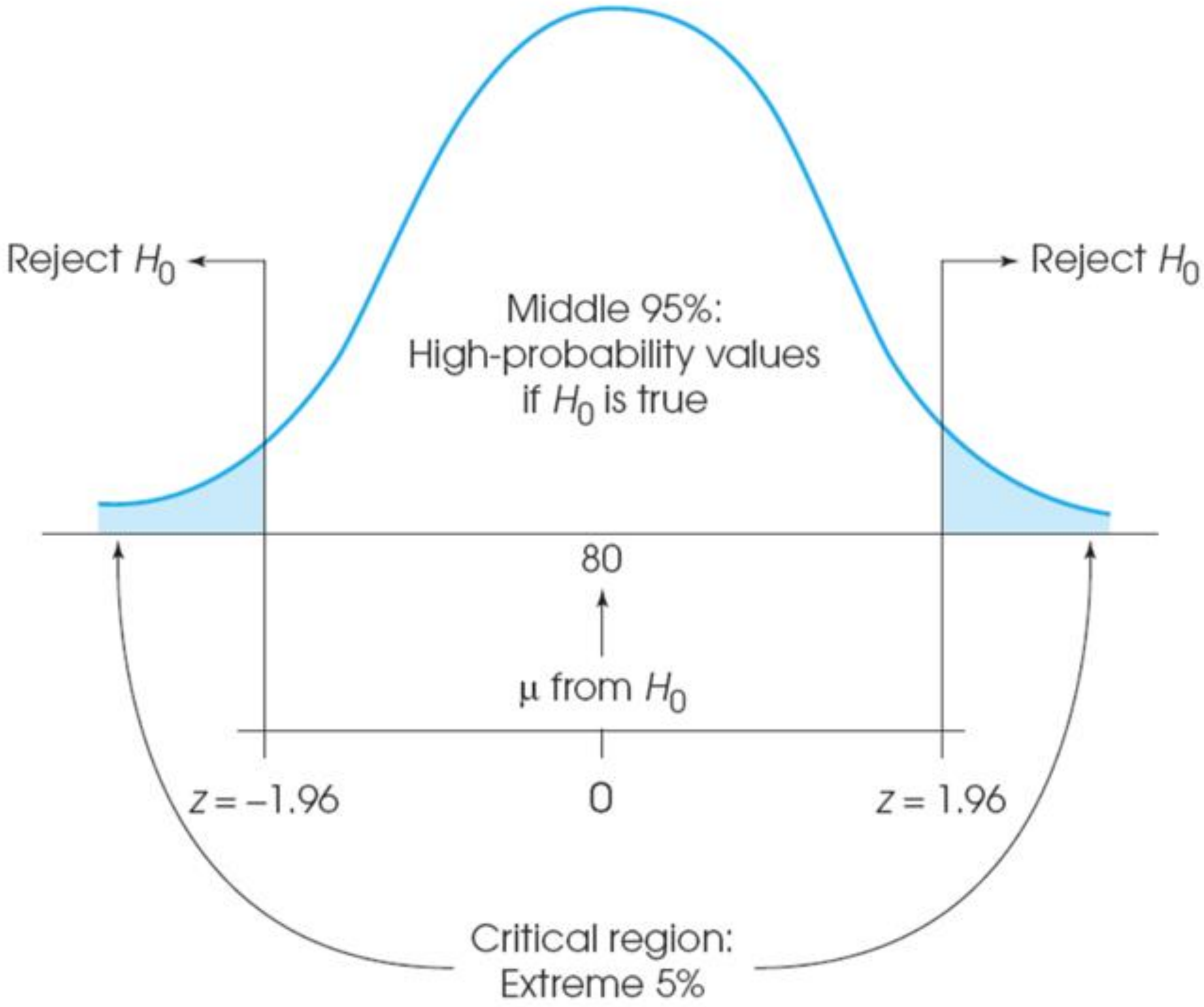
Critical Region Boundaries

Assume normal distribution

Example: if significance level of 5%:

α = .05, boundaries of critical region divide middle 95% from extreme 5%

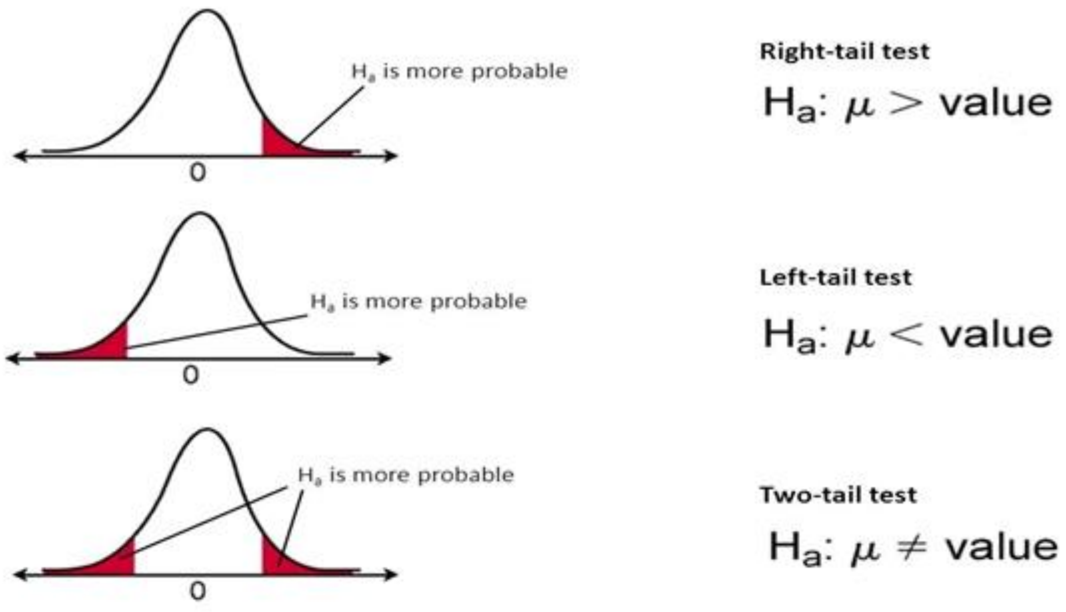
- 2.5% in each tail (2-tailed test which tests for differences in both directions)



One tailed (one sided test), Two tailed (left, right sided test)

In hypothesis testing, two-tailed tests, left-tailed tests, and right-tailed tests are different ways to specify the directionality of the test and the critical region.

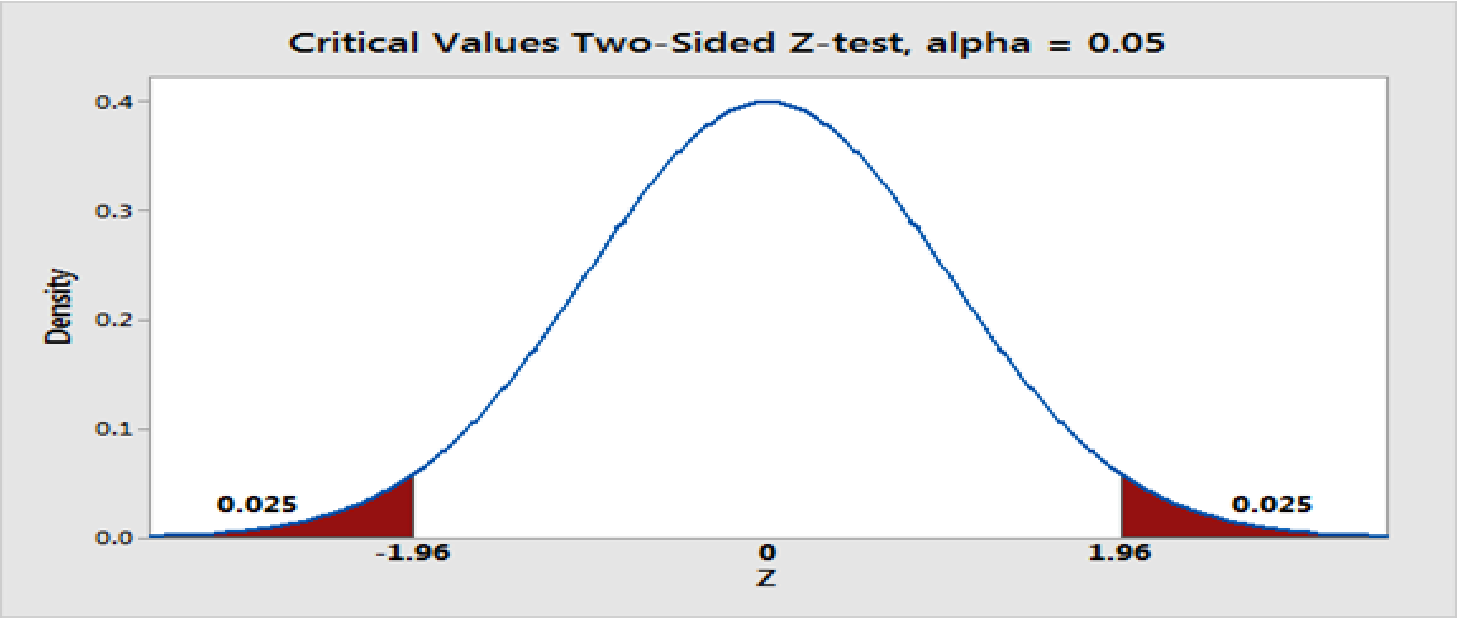
Connection to the critical region in how they determine which values of the test statistic lead to the rejection of the null hypothesis.



Two-sided hypothesis tests

Two-sided hypothesis tests have two rejection regions.

- you'll need two critical values that define them.
- Because there are **two rejection regions**, we must split our significance level in half →Each rejection region has a probability of α / 2, making the total likelihood for both areas equal the significance level.



H0: $\mu = \mu 0$
Ha: $\mu \neq \mu 0$