HCAI5DS02 – Data Analytics and Visualization.

Lecture – 07

Statistical Modeling

Statistical Inference and Hypothesis Testing.

Siman Giri

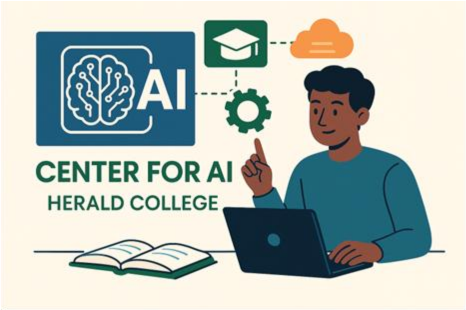


image generated via copilot.

1. Statistical Inference.

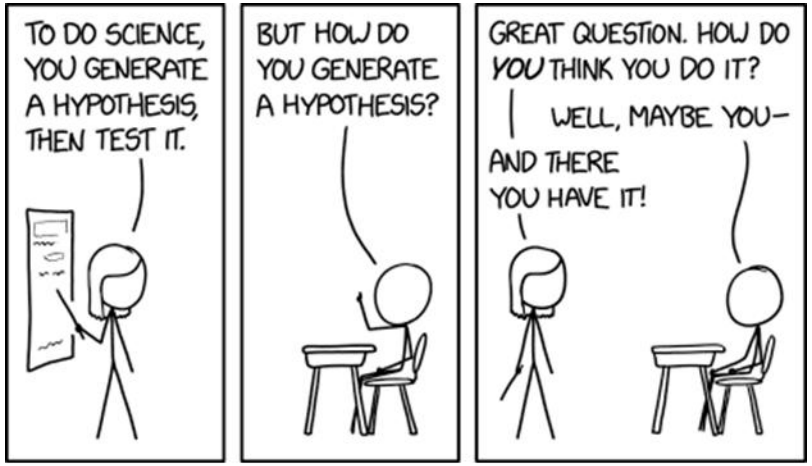
{"What can we say about the population?"}

1.1 Statistical Inference – Introduction.

- **Definition:**
  - **Statistical inference** is **the process** of **drawing conclusions** about a **population**
    - based on **information from a sample** supported by **probability theory**.
  - **Goals:**
    - Estimate unknown population parameters (e.g. mean, variance).
    - Test hypotheses (e.g. is a new product better?)
    - Quantify uncertainty (e.g. confidence intervals).
    - Make probabilistic statements.
  - **Key Tools:**
    - **Point estimation (e.g. sample mean).**
    - **Confidence intervals.**
    - **Hypothesis testing.**
    - **p – values and test statistics.**
    - **t – tests, z – tests, chi – square tests.**
- **Example:**
  - You conduct a **survey of 500 people** to **estimate the average income of an entire city.**
    - You use a **confidence interval** to report the **likely range** of **the true mean**.

1.2 Statistical Inference and Hypothesis Testing.

- **Statistical inference** allows us to **evaluate claims** about a **population parameter** using **observed data**.
  - This is done **through hypothesis testing**, where we:
    1. Formulate a null hypothesis (e.g. no effect, no difference)
    2. Collect sample data
    3. Calculate a test statistic (e.g. t – value, z – score)
    4. Evaluate the evidence against the null hypothesis using probability (p – values)
    5. **Make a decision: reject or fail to reject the null.**



©Explain xkcd.

1.2.1 Why Hypothesis Testing?

- One key goal of statistical inference is to **make informed decisions** or **judgements** about an **unknown population parameter using sample data**.
  - **For example:**
    - A company wants to know if a **new digital marketing campaign** has **increased** **average weekly sales**.
    - You collect post – campaign data and compute a **95% confidence interval** for **the new average sales**:
      - **[Rs: 9,800 – Rs: 10,200]**
  - What the **Confidence Interval** tells us:
    - We're **95% confident** that average weekly sales lie between **Rs: 9,800 and Rs: 10,200**.
    - It gives a **range** of plausible values, but not a **definitive yes or no** about **whether sales actually increased**.
- **Why is not a confidence interval enough?**
  - A confidence interval tells us the **likely range of a parameter (like average sales),**
    - but it **doesn't give us** a clear decision about whether a change is real or due to chance.
      - **Is the increase from the old average (Rs: 10,000) statistically significant?**
      - **Is this observed change real or just random variation?**
    - **Should we continue investing the campaign?**
- That's where **Hypothesis testing** comes in:
  - **Hypothesis testing** answers **yes/no** questions like:
    - *"Did the campaign cause a significant increase in average sales?"*

1.3 So, What is Hypothesis Testing?

- A statistical hypothesis is a **statement or assumption** about a **population parameter or a real – world phenomenon**
  - that can be tested using data.
  - ***"It is a claim, which we want to evaluate with evidence."***
- **For Example:**
  - "This new teaching method improves student performance."
  - "Customers prefer Brand A over Brand B."
  - These are hypotheses claims we can test with data.
- There are **usually two competing hypotheses**:
  - **The null hypothesis**, denoted $\mathbf{H_0}$ is the **hypothesis to be tested**.
    - This is the **default assumption**.
    - **In statistical notations -** $\mathbf{H_0: \mu = \mu_0}$
    - we call $\boldsymbol{\mu_0}$ **the null value**, and when we run a hypothesis test, $\boldsymbol{\mu_0}$ will be replaced by some number.
  - **The alternative hypothesis**, denoted $\mathbf{H_A}$ is **the alternative to null**.
    - **In statistical notations -** $\mathbf{H_A: \mu \neq \mu_0}$ **– The true mean is different from the null value.**

This section is framed in terms of questions about a population mean μ, but the same logic applies to p (and other population parameters).

1.3.1 Understanding What a Hypothesis Can Test?

1. **Single – Sample Hypothesis (One Group vs. a value):**
   - You are testing whether the population mean of one group is equal to some fixed number or benchmark.
   - *Example: Is the average delivery time less than 30 minutes?*
   - **Hypothesis:**
     - **null** $\rightarrow \mathbf{H_0: \mu = 30}$; **alternate** $\rightarrow \mathbf{H_a: \mu < 30}$
2. **Two – Sample Hypothesis (Between two Groups):**
   - You are testing whether two different groups have the same mean, proportion etc.
   - *Example: Do male and female employees earn the same average salary?*
   - **Hypothesis:**
     - **null** $\rightarrow \mathbf{H_0: \mu_1 = \mu_2}$; **alternate** $\rightarrow \mathbf{H_a: \mu_1 \neq \mu_2}$
3. **Hypothesis about relationships (Between Variables):**
   - You are testing whether two variables are related (not necessarily groups).
   - *Example: Is there a correlation between study time and exam score?*
   - **Hypothesis:**
     - **null** $\rightarrow \mathbf{H_0:\ \rho = 0}$ **(no correlation)** ; **alternate** $\rightarrow \mathbf{H_a:\ \rho \neq 0}$
   - This tests for **association**, not difference between groups.

1.3.1 The Logic of Hypothesis Testing.

- **Example: Hypothesis About the Mean:**
  - **Claim: "This new teaching method improves student performance."**
    - Suppose student **performance is measured by average test scores**.
  - **Step 1 – Define the parameter:**
    - **Let μ = the true mean test score of students using the new method**
    - **Let μ_0 = the known or historical mean score under the old method (e.g. 70)**
  - **Step 2 – Setup the null Hypothesis:**
    - The **null Hypothesis** always states that there is **no difference**, **no relationship** between **variables in a study**.
    - **In hypothesis testing,**
      - the null hypothesis serves **as the default or starting assumption** that researchers aim to test.
      - For example:
        1. In a **clinical trial**,
           - **the null hypothesis** might state that a new drug has **no effect** on a disease compared to a placebo.
        2. In an experiment **comparing two teaching methods**,
           - **the null hypothesis** might state that there is **no difference** in the **average test scores** of students taught by the two methods.
        3. In a study examining **the relationship** between two variables,
           - **the null hypothesis** might state that **there is no correlation** between them.
  - In our problem null hypothesis can be stated as:
    - **H_0: the mean of the population is $\rightarrow \mu = 70 (== \mu_0)$.**

This section is framed in terms of questions about a population mean μ, but the same logic applies to p (and other population parameters).

1.3.1 The Logic of Hypothesis Testing.

- **Step 3 – Setup Alternate Hypothesis:**
  - The **alternative hypothesis** (denoted as $\mathbf{H_1}$ **or** $\mathbf{H_a}$) is the statement that researchers
    - **aim to provide evidence for in a statistical hypothesis test.**
  - It is the opposite of the null hypothesis and represents
    - the **presence of an effect**, **difference, or relationship that the researcher expects or hopes to find**.
  - For Example:

Example 1:	**The new drug has no effect on the disease compared to a placebo.**		
Scenario	**Null Hypothesis**	**Alternate Hypothesis**	**In Practice**
Clinical Trial: New Drug vs. Placebo	• No difference in effect. • $\mathbf{H_0: \mu_{drug} = \mu_{placebo}}$	• $\mathbf{H_a: \mu_{drug} \neq \mu_{placebo}}$ • **Two-sided test.**	• $\mathbf{H_1: \mu_{drug} > \mu_{placebo}}$ • We are only interested in improvement. • **One – sided alternative.**

Example 2:	There is no difference in average test scores between methods A and methods B		
Scenario	**Null Hypothesis**	**Alternate Hypothesis**	**In Practice**
Teaching Methods Experiment Method A vs. Method B	$\mathbf{H_0}$: ?	$\mathbf{H_1}$: ?	$\mathbf{H_1}$: ?

1.3.1 The Logic of Hypothesis Testing.

- **Step 3 – Setup Alternate Hypothesis:**
  - The **alternative hypothesis** (denoted as H_1 **or** H_a) is the statement that researchers
    - **aim to provide evidence for in a statistical hypothesis test.**
  - It is the opposite of the null hypothesis and represents
    - the **presence of an effect**, **difference, or relationship that the researcher expects or hopes to find**.
  - For Example:

Example 1:	The new drug has no effect on the disease compared to a placebo.		
Scenario	**Null Hypothesis**	**Alternate Hypothesis**	**In Practice**
Clinical Trial: New Drug vs. Placebo	• No difference in effect. • $H_0: \mu_{drug} = \mu_{placebo}$	• $H_a: \mu_{drug} \neq \mu_{placebo}$ • Two-sided test.	• $H_1: \mu_{drug} > \mu_{placebo}$ • We are only interested in improvement. • One – sided alternative.

Example 2:	There is no difference in average test scores between methods A and methods B.		
Scenario	**Null Hypothesis**	**Alternate Hypothesis**	**In Practice**
Teaching Methods Experiment Method A vs. Method B	$H_0: \mu_A = \mu_B$	$H_1: \mu_A \neq \mu_B$	$H_1: \mu_A > \mu_B$

1.3.1 The Logic of Hypothesis Testing.

- **Step 4 – Test your Hypothesis:**
- **Example: Hypothesis About the Mean:**
  - **Claim: "This new teaching method improves student performance."**
    - Suppose student **performance is measured by average test scores**.
    - After implementing the new method in a sample of classrooms,
      - we observe an average score of 73 (which was 70 earlier).
    - We now test whether this observed increase is **statistically significant**, or just due to chance.
      - $\mathbf{H_0: \mu = 70}$
        $\mathbf{H_a: \mu > 70}$
  - **How:**
    I. Collect a sample test scores using the new method.
    II. Compute the **sample mean ($\bar{x}$)**
    III. Conduct a test (**e.g. one sample t – test**)
    IV. Calculate the **p – value**.
- **To understand:**
  - The null hypothesis serves as a **starting point** — a baseline assumption that any **observed pattern in the data is due to random chance**.
    - Nothing unusual is happening. Show me strong enough evidence to believe otherwise.

1.3.1 The Logic of Hypothesis Testing.

- **Step 5 – Making Decisions:**
  - The hypothesis we want to test is **whether H_a is likely true**.
  - So, there are two possible outcomes:
    - **Reject H_0 and accept H_a** because of **sufficient evidence** in the sample in favor of H_a.
    - Do not **reject H_0** because of **insufficient evidence** to **support H_a**.

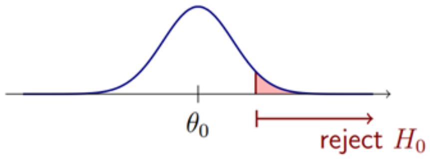
Very important!!!
Note that failure to **reject H_0** does not mean the null hypothesis is true. There is no formal outcome that says "**accept H_0**".It only means that we do not have sufficient evidence to **support H_a**.

2. Types of Hypothesis Testing.

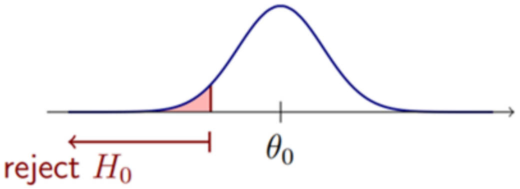
{One Vs. Two Tailed Test .}

2.1 One – Tailed Test.

- **One-Tailed Test:**
  - **Purpose:**
    - A **one-tailed test** is used when you are **interested in detecting a difference in a specific direction**.
    - You hypothesize that **the true parameter is either greater than or less than a certain value**, **but not both**.
  - **Examples:**
    - **Right-Tailed Test**: Testing if a new drug is more effective than the standard drug.
    - **Left-Tailed Test**: Testing if a machine produces fewer defective items than the industry standard.
  - **Null and Alternative Hypotheses:**



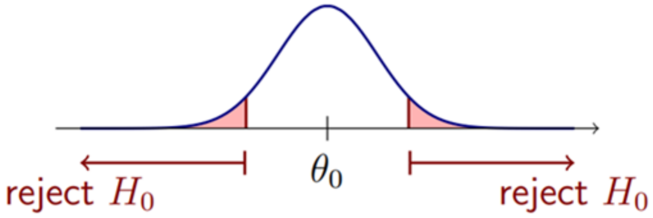
- **Right Tailed:**
  - $H_0: \mu \leq \mu_0$
    $H_a: \mu > \mu_0$



- **Left Tailed:**
  - $H_0: \mu \geq \mu_0$
    $H_a: \mu < \mu_0$

2.2 Two – Tailed Test.

- **Purpose**:
  - A two-tailed test is used when you are interested in detecting any significant difference
    - from the null hypothesis, **regardless of the direction**.
  - You hypothesize that the true parameter is different from a specific value, but **you don't specify a direction**.
- **Examples**:
  - Testing if the average height of a group is different from a known average (it could be either greater or less).
  - Testing if a new teaching method leads to a different average test score compared to the traditional method (it could be higher or lower).
- **Null and Alternative Hypotheses**:



- **Two Tailed:**
  - $H_0: \mu = \mu_0$
    $H_a: \mu \neq \mu_0$

A. Before we Perform {Hypothesis}Test …

{Let’s define some of the terminologies used …

A.1 Remember – Statistic …

- **Sample Statistic:**
  - This is a statistic specifically used
    - to estimate a corresponding population parameter.
  - Sample Statistic are used to infer values
    - about the population.

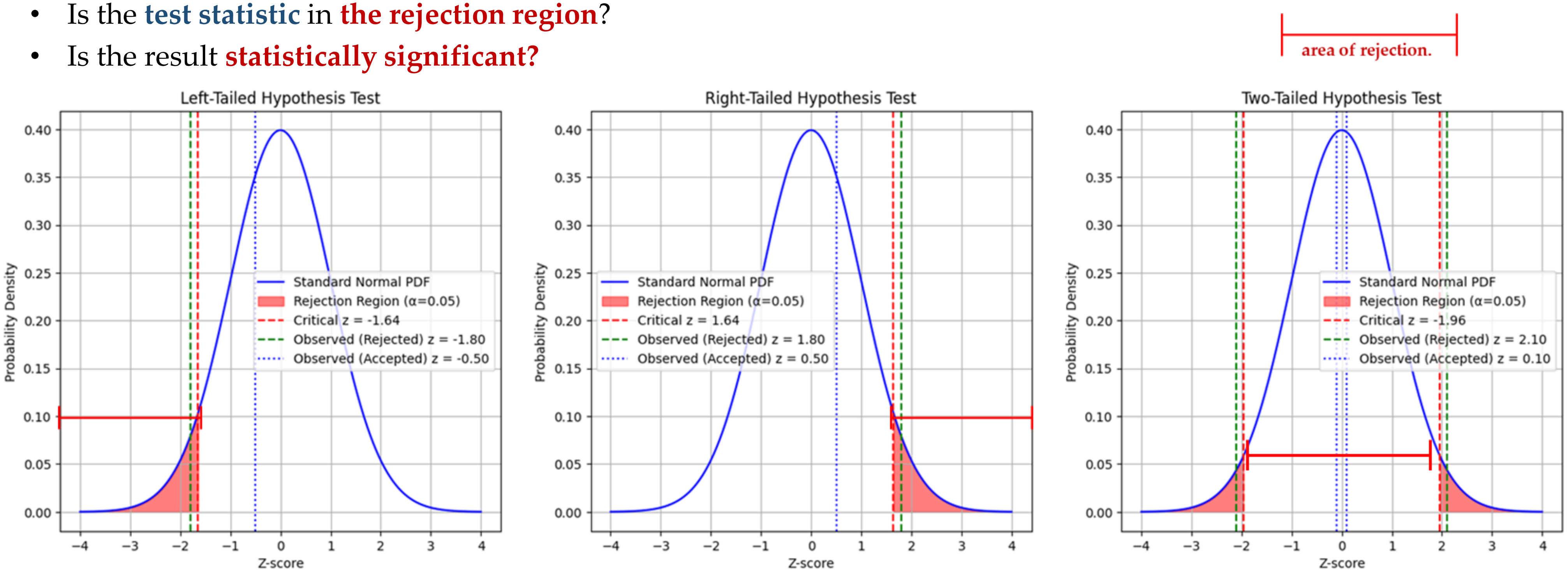
Example: Sample Statistics.	
Population Parameter	**Sample Statistic**
Population mean μ	Sample mean $\bar{x}$
Population proportion p	Sample proportion $\hat{p}$
Population variance σ^2	Sample variance s^2

Common Forms of Test Statistics		
Test Type	**Formula**	**Description**
One – sample t - test	$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$	Sample mean vs. hypothesized mean
Z - test	$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$	Sample proportion vs. population proportion
Two – sample t - test	$t = \frac{\bar{x}_1 - \bar{x}_2}{SE_{diff}}$	Comparing two sample means.

- **Test Statistic:**
  - A **test statistic** is a specific kind of statistic used to
    - determine **how far** the sample data **deviate** from
    - what is **expected** under the **null hypothesis**.
    - *"A **test statistic** measures the distance (in standard error units) between your **sample estimate** and the **null hypothesis value**."*
  - It converts the **difference between**
    - your **observed value** and **the null value** into a **standardized value**
      - (e.g., a z-score or t-score)
      - that can be compared to a **critical value**.
  - It answers the question:
    - ***"Is the observed result unusual enough to reject the null hypothesis?"***

A.1.1 Test Statistic – How it is used?

- Calculate the **test statistic** from your **sample data.**
- Compare it to **the critical value** or use it to **compute the p – value**.
- Make a decision:
  - Is the **test statistic** in **the rejection region**?
  - Is the result **statistically significant?**



A.1.2 How to Compute test statistics.

- A **test statistic** tells us **how many standard errors** our observed sample result is from the value specified in the **null hypothesis**.
  - $\textbf{Test statistics} = \dfrac{\textbf{sample estimate} - \textbf{Null value}}{\textbf{Standard error}}$
- First determine **one sample or two sample test**:
  - **One sample test:**
    - You have **one sample** and
      - **want to test** whether **its mean (or proportion)** is **equal to a known or hypothesized population value**.
    - **Example: Is the average weight of packaged rice equal to 1 kg?**
      - $\begin{cases} H_0: \mu = 1 \\ H_a: \mu \neq 1 \end{cases}$
  - **Two sample test:**
    - You want to compare **the means or proportions** of **two independent groups**.
    - **Example: Do customers who received a discount spend more than those who didn't?**
      - $\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$

A.1.3 Test Statistic for Z test.

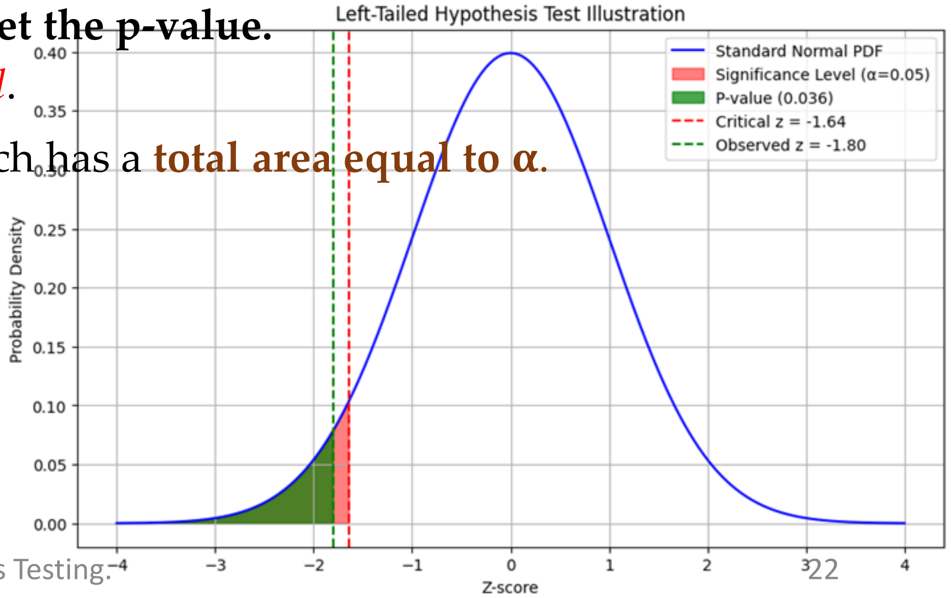
- **Used when:**
  - Population standard deviation σ is known or
  - Sample size $n \geq 30$
  - **One – sample Z – test:**
    - **for Mean:**
      - $z_{statistic} = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$
    - **for proportion:**
      - $z_{statistic} = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$
  - **Two sample Z – test:**
    - **for mean:**
      - $z_{statistic} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}$
    - **for proportion:**
      - $z_{statistic} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$
      - $\hat{p} \rightarrow \frac{x_1 + x_2}{n_1 + n_2}$ → called **pooled proportion** and x → **number of sucess** in **group 1**(x_1) **and 2**(x_2).

A.1.4 Test Statistic for t – test.

- **Used when:**
  - **σ** unknown and
  - Typically, small sample sizes $n < 30$
  - **One sample t – test for mean:**
    - $\mathbf{t_{statistic} = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}}$ **; for df** $\mathbf{= n - 1}$
  - **Two sample t – test for mean:**
    - $\mathbf{t_{statistic} = \frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}}$
    - This is called **Welch's t-test** (default in most statistical software)
      - and used when we have unequal variances.
    - **Degrees of freedom – welch – Satterthwaite Approximation**:
      - $\mathbf{df = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1^2)\text{\^{}}2}{n_1-1}+\frac{(s_2^2/n_2^2)\text{\^{}}2}{n_2-1}}}$

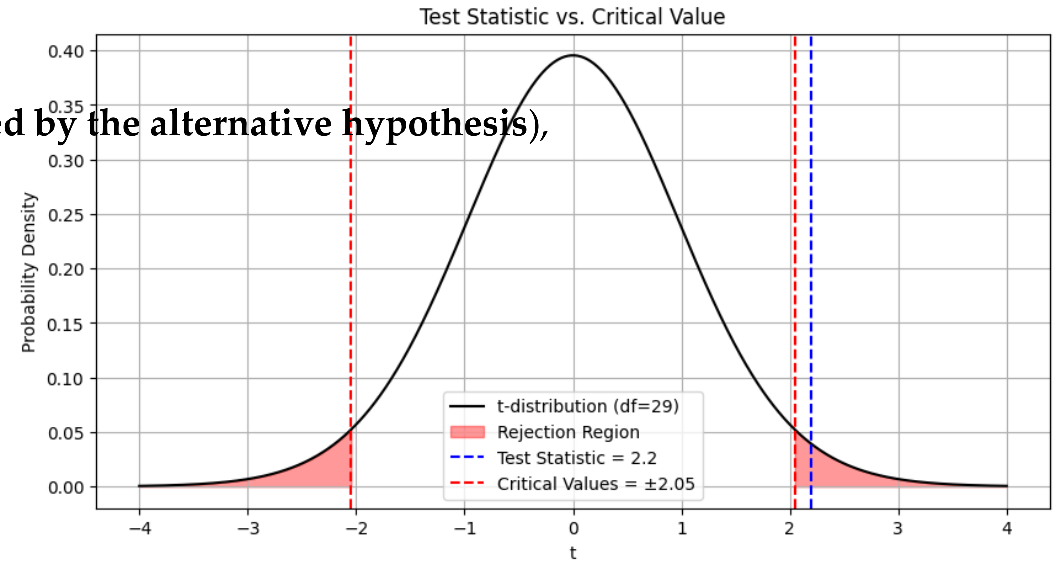
A.2 Some Key Concepts – Significance Level.

- **Significance level (denoted α):**
    - The **significance level α** is the **threshold** we set **before** doing the test to decide
        - how much evidence we need to **reject the null hypothesis**.
    - It is the **maximum probability** of making a **Type I error** (**rejecting a true null hypothesis**)
        - that we are **willing to tolerate**.
    - **Typical Values**: Common significance levels are **0.05, 0.01, and 0.10**.
    - **Usage**:
        - It is used to **determine the critical value(s)** and to **interpret the p-value.**
        - If the ***p-value is less than α***, the **null hypothesis** is *rejected*.
- The **shaded red area** in the **plots represents the critical region**, which has a **total area equal to α.**



A.3 Some Key Concepts – Critical Value.

- **Critical Value:**
  - A critical value is the cutoff point on **the distribution curve** that defines the **boundary** of the rejection region in a **hypothesis test.**
  - It is the **value beyond** which we consider a result to be **statistically significant** at a given **significance level α**.
  - It is **fixed before the test** and acts like a **benchmark**.
- **What Does It Do?**
  - The critical value helps you decide:
    - **When your test statistic is far enough from the null hypothesis value**
    - **Whether to reject the null hypothesis.**
- **General Rule:**
  - If the **test statistic** exceeds the critical value (**in the direction specified by the alternative hypothesis**),
    - you **reject the null hypothesis**.
- **Depends On:**
  - **Type of test** (z, t, etc.)
  - **Significance level** ($\alpha \rightarrow 0.05, 0.01$)
  - **One-tailed or two-tailed test**
  - **Degrees of freedom** (for t-tests)



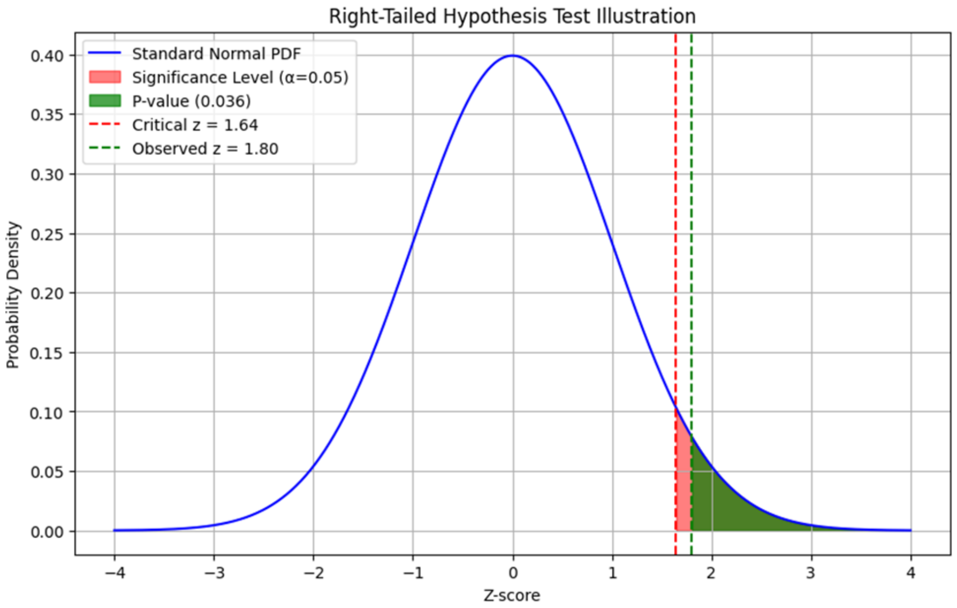
A.3.1 How to Compute the Critical Value?

- For a **z – test** – when population standard deviation is known or large n:
  - **Two tailed test:**
    - $z_{critical} = \pm z_{\frac{\alpha}{2}}$
  - **One tailed test:**
    - $z_{critical} = z_{\alpha}(\text{right tail}) \text{ or } z_{critical} = -z_{\alpha}(\text{left tail})$

```
from scipy.stats import norm

alpha = 0.05

z_critical_two_tail = norm.ppf(1 - alpha/2)  # For two-tailed
z_critical_right = norm.ppf(1 - alpha)       # For right-tailed
z_critical_left = norm.ppf(alpha)            # For left-tailed
```



- **Interpretation:**
  - Since the **observed z-score (1.80)** is greater than the **critical z-value (1.64)**, it falls into **the critical region.**
  - Alternatively, since the **p-value (0.036)** is less than the **significance level ($\alpha = 0.05$),**
  - we would **reject the null hypothesis** at the **0.05 significance level.**
- This **suggests** there is **enough evidence** to support **the alternative hypothesis** in a **right-tailed test.**

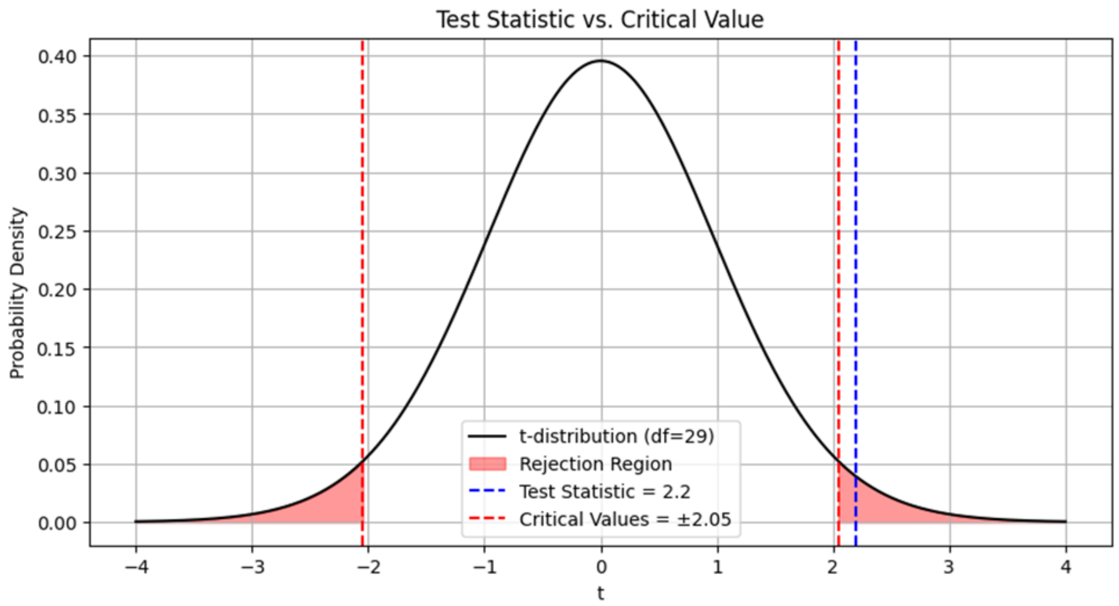
A.3.1 How to Compute the Critical Value?

- For a **t – test** when population std is unknown and small sample size:
- Use the t – distribution with calculated degree of freedom:
  - **Two tailed test:**
    - $\mathbf{t_{critical} = \pm t_{\frac{\alpha}{2}}}$
  - **One tailed test:**
    - $\mathbf{t_{critical} = t_{\alpha}(right\ tail)\ or\ t_{critical} = -t_{\alpha}(left\ tail)}$

```
from scipy.stats import t

df = 29
alpha = 0.05

t_critical_two_tail = t.ppf(1 - alpha/2, df)
t_critical_right = t.ppf(1 - alpha, df)
t_critical_left = t.ppf(alpha, df)
```



Interpret - ?

A.4 Some Key Concepts – p –Value.

- The **p-value** is the **probability of obtaining a result** as extreme (or more extreme)
  - than the observed test statistic, **assuming the null hypothesis is true**.
- It answers:
  - *"If H_0 is true, how rare is this result?"*
- Decision Rule:
  - If **p – value < α** → Reject the null hypothesis.
  - If **p – value ≥ α** → Do not reject.
- **Important – p – value is calculated from the data and *α* is chosen before the test.**
- How to compute the **p – value**?
  - The **p-value** is the **area under the curve** beyond the observed test statistic (**depends on test direction**).

```
from scipy.stats import norm

z = 2.1  # your test statistic

# Two-tailed p-value
p_two_tailed = 2 * (1 - norm.cdf(abs(z)))

# Right-tailed
p_right = 1 - norm.cdf(z)

# Left-tailed
p_left = norm.cdf(z)
```

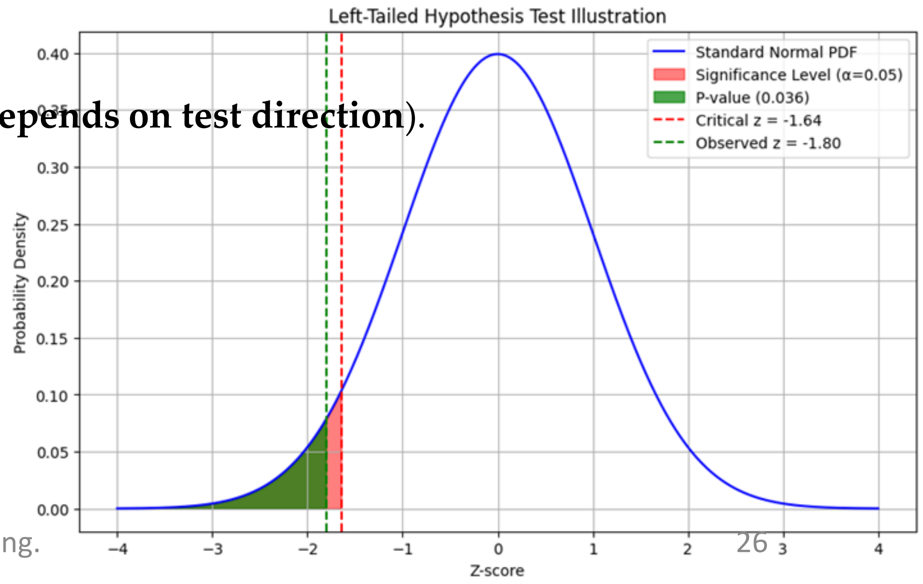
```
from scipy.stats import t

t_stat = 2.2; df = 29 # your test statistic

# Two-tailed
p_two_tailed = 2 * (1 - t.cdf(abs(t_stat), df))

# Right-tailed
p_right = 1 - t.cdf(t_stat, df)

# Left-tailed
p_left = t.cdf(t_stat, df)
```



A.5 Hypothesis Testing: Error.

- When we perform a hypothesis test, we make a **decision**
  - either to **reject** or **not reject** the null hypothesis $\mathbf{H_0}$
  - But since we're using a sample (not the full population),
    - there's always a chance of making the **wrong decision**.
- There are **two main types of errors**:
  - **Type I Error (False Positive)**
  - **Type II Error (False Negative)**

	H_0 is true	H_1 is true
Do not reject H_0	Correct decision	Type II error
Reject H_0	Type I error	Correct decision

Fig: Types of error in Hypothesis Testing.

A.5.1 Hypothesis Testing: Error.

Type – I Error

- **Definition**:
  - Rejecting the null hypothesis when it is actually **true**.
  - Controlled by the **significance level** $\boldsymbol{\alpha}$.
  - The **probability** of committing a **Type I error**
    - is called the **level of significance.**
- ***"I'm willing to accept a α = 5% chance of incorrectly thinking the new page is better when it's not."***

Type – II Error

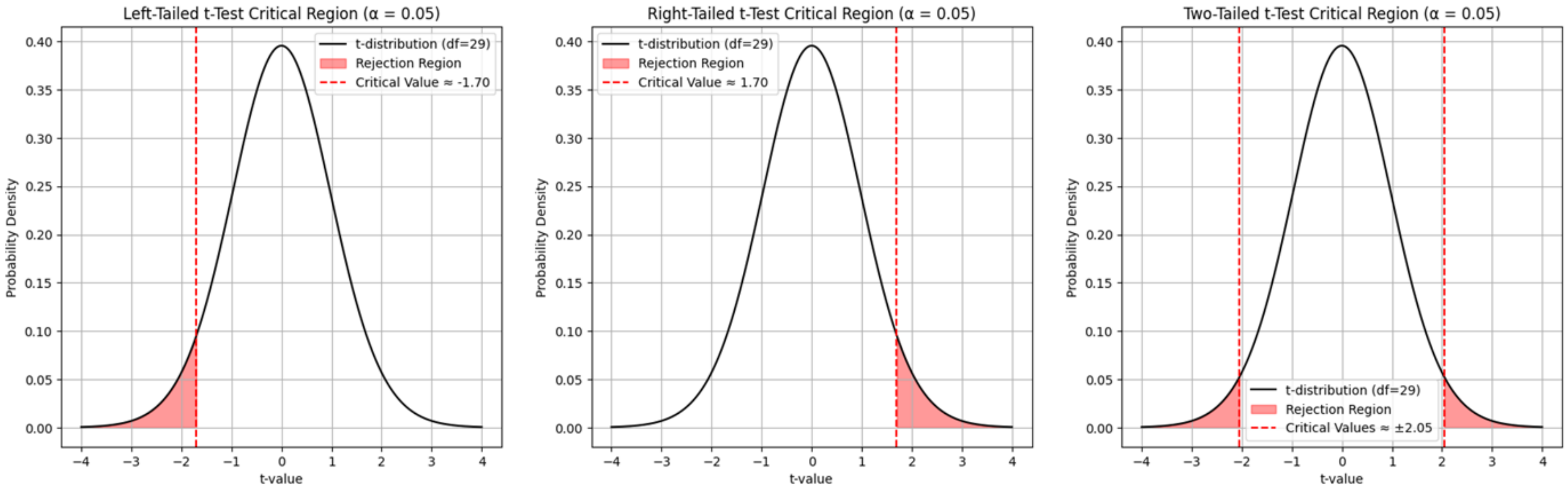
- Type II Error (False Negative)
  - **Definition**:
    - **Failing to reject** the null hypothesis when it is actually **false**.
  - Probability is denoted by β.
  - **Power** $= \mathbf{1 - \beta}$:
    - probability of correctly rejecting a false $\mathbf{H_0}$

Getting Back to Hypothesis testing …

{3. Critical Value and Decision Rule.}

3.1 Step 1: Determine the Critical Value.

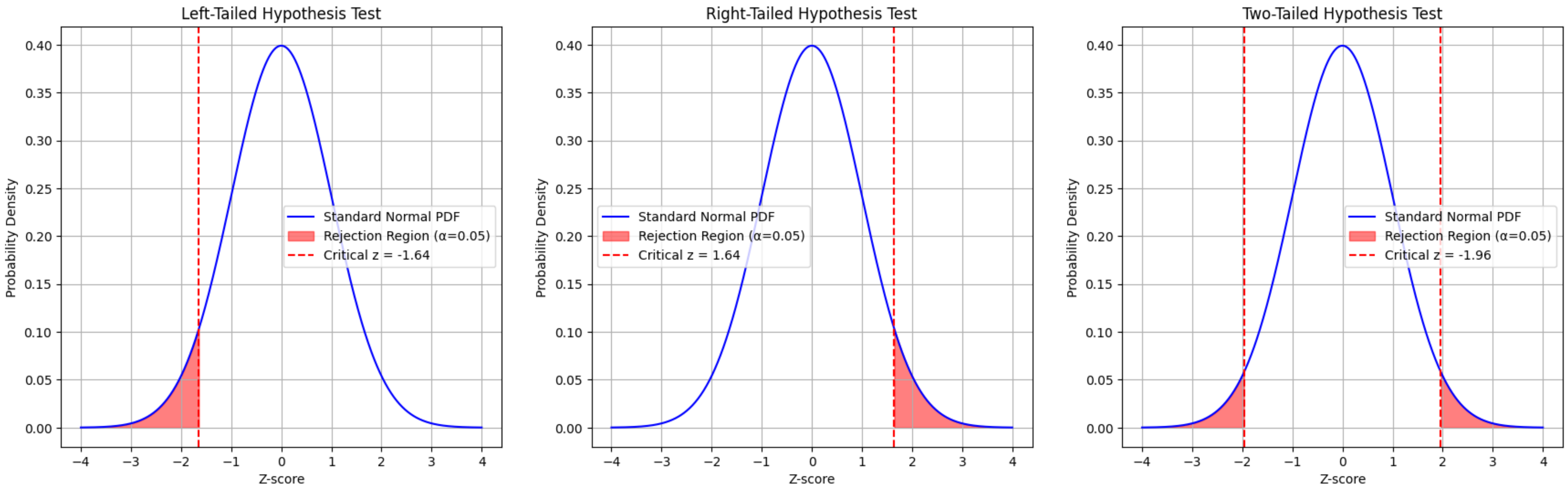
- Find the **critical value** from the appropriate distribution (e.g., t-distribution or z-distribution), based on:
  - The chosen significance level α.
  - The **degrees of freedom** (if applicable)
  - The **type of test** (one-tailed or two-tailed)



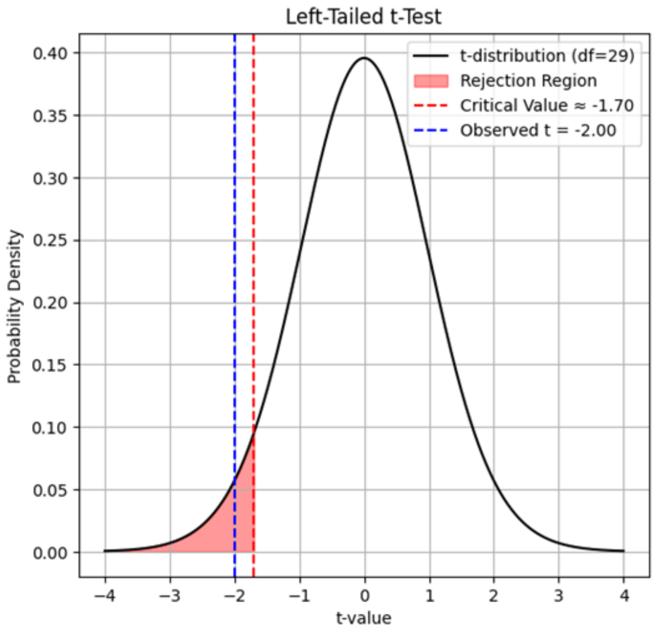
T – test.

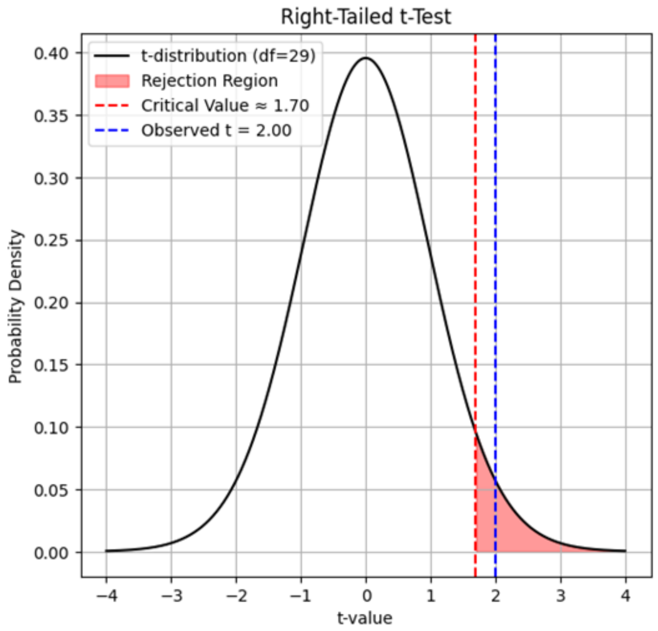
3.1 Step 1: Determine the Critical Value.

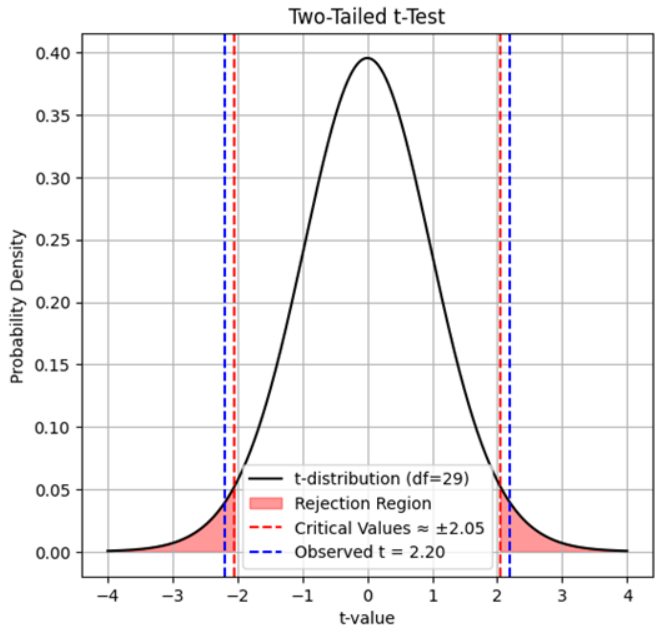
- Find the **critical value** from the appropriate distribution (e.g., t-distribution or z-distribution), based on:
  - The chosen significance level α.
  - The **degrees of freedom** (if applicable)
  - The **type of test** (one-tailed or two-tailed)



3.2 Step 2: Decision Rule (Based on Tail Type)





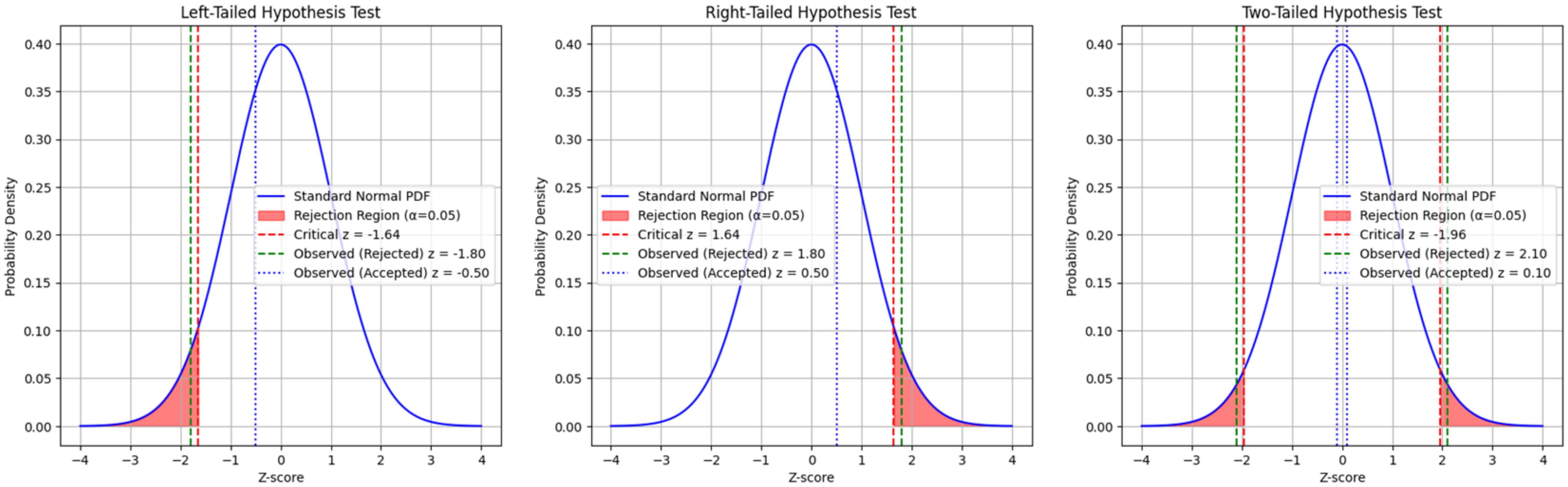


- **Left – Tailed Test:**
  - Reject the null hypothesis if:
    - **test statistic ≤ crtical value**

- **Right – Tailed Test:**
  - Reject the null hypothesis if:
    - **test statistic ≥ crtical value**

- **Two – Tailed Test:**
  - Reject the null hypothesis if:
    - **|test statisti| ≥ |crtical value| or**
    - **test statistic ≤ −crtical value or test statistic ≥ critical value**

3.2 Step 2: Decision Rule (Based on Tail Type)



- **Left – Tailed Test:**
  - Reject the null hypothesis if:
    - **test statistic ≤ crtical value**

- **Right – Tailed Test:**
  - Reject the null hypothesis if:
    - **test statistic ≥ crtical value**

- **Two – Tailed Test:**
  - Reject the null hypothesis if:
    - **|test statisti| ≥ |crtical value| or**
    - **test statistic ≤ −crtical value or test statistic ≥ critical value**

Let's do an Example …

Evaluating Sales Training …

- **Background:**
  - Your company recently conducted a **sales training program** for a group of sales employees, aiming to improve their monthly sales performance. Now, management wants to know whether the training actually led to a **statistically significant improvement** in average sales.
- **Objective:**
  - Use a **two-sample t-test** to determine whether the **trained group** has a significantly **higher average monthly sales** than the **untrained group**.

Data Summary:			
Group	Sample Size (n)	Mean Monthly Sales (Rs.)	Sample Std. Deviation (Rs.)
Trained Group	30	9,800	1,100
Untrained Group	35	9,200	1,000

Evaluating Sales Training …

- **Questions:**
  1. **State the null and alternative hypotheses.**
     - Be clear about whether it is one-tailed or two-tailed.
  2. **Check if the assumptions** for using a two-sample t-test are satisfied:
     - Are the samples independent?
     - Is it reasonable to assume normality or use the Central Limit Theorem?
  3. **Calculate the test statistic** and **degrees of freedom** using the Welch's t-test formula (unequal variances assumed).
  4. **Determine the p-value** and **interpret** the result at $\boldsymbol{\alpha = 0.05}$.
  5. **Make a business recommendation**:
     - Should the company roll out the training program company wide?
     - Discuss the risk of **Type I** and **Type II errors** in your conclusion.

1. State the Hypotheses.

- Goal:
  - Test if the training improved sales (i.e., trained group's mean is greater than untrained group).
  - Let:
    - μ_1: mean sales of trained group
    - μ_2: mean sales of untrained group
  - Hypotheses Statement:
    - $\begin{cases} H_0: \mu_1 = \mu_2 \text{ (no improvement)} \\ H_a: \mu_1 > \mu_2 \text{ (training improved sales)} \end{cases}$
  - Right Tailed Test.

2. Assumptions Check …

- Two **independent** samples.
- Sample Sizes > 30 → CLT applies i.e. normality assumption can be made reasonably.
- **Unequal variances** → we will use **Welch's t – test**.

Data Summary:			
Group	**Sample Size (n)**	**Mean Monthly Sales (Rs.)**	**Sample Std. Deviation (Rs.)**
Trained Group	30	9,800	1,100
Untrained Group	35	9,200	1,000

3. Compute the Test Statistic

- **Formula (Welch's t – test):**
  - $t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$
- **Given:**
  - $\bar{x}_1 = 9800, s_1 = 1100, n_1 = 30$
  - $\bar{x}_2 = 9200, s_2 = 1000, n_2 = 35$
    - $t = \frac{9800 - 9200}{\sqrt{\frac{1100^2}{30} + \frac{1000^2}{35}}} = \frac{600}{\sqrt{40333.33 + 28571.43}} \approx 2.287$
- **Compute Degree of Freedom:**
  - Using welch – Satterthwaite approximation:
    - $df = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}} \approx 57.3$

4. Find a p – value.

- **Using scipy.stats.t:**

```
from scipy.stats import t
import numpy as np

# Given values
x1_bar = 9800
x2_bar = 9200
s1 = 1100
s2 = 1000
n1 = 30
n2 = 35

# Compute t-statistic
numerator = x1_bar - x2_bar
denominator = np.sqrt((s1**2)/n1 + (s2**2)/n2)
t_stat = numerator / denominator

# Degrees of freedom
df_num = ((s1**2)/n1 + (s2**2)/n2)**2
df_denom = (((s1**2)/n1)**2)/(n1-1) + (((s2**2)/n2)**2)/(n2-1)
df = df_num / df_denom

# One-tailed p-value
p_val = 1 - t.cdf(t_stat, df)

print(f"t-statistic: {t_stat:.3f}")
print(f"Degrees of freedom: {df:.2f}")
print(f"p-value: {p_val:.4f}")
```

→

Cautions !!!		
Test Type	**Condition**	**Python Code**
Right - tailed	**t > 0**	p_val= 1 - t.cdf(t_stat, df)
Left - tailed	**t < 0**	p_val= t.cdf(t_stat, df)
Two - tailed	**any**	p_val= 2 *(1 – t.cdf(abs(t_stat), df))

- **Expected Output:**
  - **t-statistic: 2.287**
  - **Degrees of freedom: 57.31**
  - **p-value: 0.0129**

5. Decision and Interpretation.

- $\mathbf{p = 0.0129 < \alpha = 0.05}$
- **Decision: Reject H_0**
- Interpretation:
  - There is **statistically significant** evidence at the **5% level** that the sales training program
    - **increased average monthly sales.**
  - The company should consider rolling out the program more broadly.

Thank You.