

why we having different formula of p value  
The p-value is:

Two-tailed  
Right-tailed  
Left-tailed

We have different formulas for p-value (two-tailed, right-tailed, left-tailed) because the meaning of “extreme evidence against the null hypothesis” depends on the direction of the alternative hypothesis ( $H_1$ ).

Let’s explain this intuitively, mathematically, and with visual diagrams (in words) so that it becomes permanently clear.

## ✓ 1. Why do we have different p-value formulas?

Because p-value always measures:

“How extreme is my sample result assuming  $H_0$  is true?”

But what counts as extreme depends on your chosen direction in  $H_1$ :

### ◆ Case 1: Right-tailed test

Alternative hypothesis:

$$H_1 : \mu > \mu_0$$

What is considered “extreme”?

☰  > ChatGPT 5.2 ▾

  ...

So the p-value is:

$$p = P(Z \geq z_{\text{obs}})$$

## Why?

You're checking how likely you would get a **value equal or higher** than the observed one.

### Visual intuition:

bash



Copy code

### ◆ Case 2: Left-tailed test

Alternative hypothesis:

$$H_1 : \mu < \mu_0$$

"Extreme" means:

👉 Values much smaller than the null value.

So the p-value is:

$$p = P(Z \leq z_{\text{obs}})$$

### Visual intuition:

bash



Copy code

### ◆ Case 3: Two-tailed test

Alternative hypothesis:

$$H_1 : \mu \neq \mu_0$$

"Extreme" values are those far from  $\mu_0$  in *both* directions.

So the p-value is:

$$p = 2 \times P(Z \geq |z_{\text{obs}}|)$$


Why multiply by 2?

Because extreme evidence can appear on either side:

markdown

```

Left tail      Right tail
  /\           /\
 _/  \_____/  \_____/
 -|z|      |z|
  
```

 Copy code

## 2. The real reason behind different formulas

The p-value formula depends solely on  $H_1$ , not the data.

Alternative Hypothesis	Direction	Formula for p-value
$H_1 : \mu > \mu_0$	Right	$P(Z \geq z_{\text{obs}})$
$H_1 : \mu < \mu_0$	Left	$P(Z \leq z_{\text{obs}})$
$H_1 : \mu \neq \mu_0$	Two-tail	$(2P(Z \geq  z_{\text{obs}} ))$

Think of hypothesis testing like **court trials**:

- **Right-tailed test** = proving someone is **too high** (e.g., pollution exceeds limit).
- **Left-tailed test** = proving someone is **too low** (e.g., battery life is worse than promised).
- **Two-tailed test** = proving something is **simply different**, higher or lower.

Different claims → different "extreme regions".

## 3. Mathematical Example (Simple)

Suppose you compute a test statistic:

$$z_{\text{obs}} = 1.8$$

 **Right-tailed test:**

$$p = P(Z \geq 1.8) = 0.0359$$

▶ Left-tailed test:

$$p = P(Z \leq 1.8) = 0.9640$$

(This is obviously not evidence for a left-tail claim.)

▶ Two-tailed test:

$$p = 2P(Z \geq 1.8) = 2 \times 0.0359 = 0.0718$$

## ★ 4. Why the formulas must differ — Intuition summary

Different p-value formulas exist because:

Test Type	What counts as "extreme"?	Why formula changes
Right-tailed	Big values	Only upper tail matters
Left-tailed	Small values	Only lower tail matters
Two-tailed	Very small or very big values	Both tails matter → multiply by 2

If you used the **wrong p-value formula**, you would be measuring the **wrong kind of extremeness**, leading to:

- wrong conclusions
- invalid significance levels
- false-positive or false-negative results

## 🔥 If you want next:

I can explain how p-value is mathematically derived from CDFs, PDFs, and integrals → with diagrams.

Do you want:

- ✓ Derivation of p-value using integrals
- ✓ Real-world ML examples (feature selection, A/B testing)
- ✓ A flowchart: "Which tail test should I use?"

## ✓ Step-by-step numeric example for all 3 cases

Just say next.



first tell me what variable defines what

Perfect question 👍

Before touching **p-values**, we must be **100% clear about the variables and symbols**.

I'll go from **data** → **statistic** → **distribution** → **p-value**, and every symbol will be defined clearly.

### 1 Raw data variables (what you actually observe)

◆  $X_1, X_2, \dots, X_n$

- **Meaning:** Individual data points (random variables)
- **Example:** Heights of students, daily returns, model errors
- **Type:** Random variables

◆  $n$

- **Meaning:** Sample size
- **Example:** Number of students measured
- **Role:** Controls variability and test statistic scale

### 2 Population parameters (usually unknown)

These describe the **true population**, not the sample.

◆  $\mu$

- **Meaning:** True population mean
- **Example:** True average height of all students
- **Unknown** → that's why we test hypotheses

◆  $\mu_0$ 

- **Meaning:** Mean assumed under the null hypothesis
  - **Example:** "The average height is 170 cm"
  - **Fixed reference value**
- 

◆  $\sigma$ 

- **Meaning:** Population standard deviation
  - **Known** → Z-test
  - **Unknown** → t-test
- 

**3 Sample statistics (computed from data)**

These summarize the observed data.

◆  $\bar{X}$ 

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

- **Meaning:** Sample mean
  - **Estimator of  $\mu$**
- 

◆  $s$ 

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$$

- **Meaning:** Sample standard deviation
  - **Estimator of  $\sigma$**
- 

**4 Test statistic (this is the KEY variable)**

This converts your data into a **standardized scale**.

---

◆  $Z$  (Z-statistic)

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

- **Meaning:** How many standard errors away from  $\mu_0$
- **Distribution:** Standard Normal  $N(0, 1)$
- **Used when**  $\sigma$  is known

#### ◆ $T$ (t-statistic)

$$T = \frac{\bar{X} - \mu_0}{s / \sqrt{n}}$$

- **Meaning:** Same idea as  $Z$ , but uncertainty in  $\sigma$
- **Distribution:** t-distribution with  $n - 1$  degrees of freedom

#### ◆ $z_{\text{obs}}, t_{\text{obs}}$

- **Meaning:** Observed numerical value of the test statistic
- **Example:**  $z_{\text{obs}} = 1.85$
- **Used directly to compute p-value**

## 5 Distribution variables (probability layer)

These describe the **sampling distribution**.

#### ◆ $\phi(z)$

- **Meaning:** PDF of standard normal distribution
- **Role:** Shape of bell curve

#### ◆ $\Phi(z)$

$$\Phi(z) = P(Z \leq z)$$

- **Meaning:** CDF of standard normal
- **Used to compute probabilities (p-values)**

## 6 p-value variables (final decision layer)

### ◆ p-value

- Meaning:

$$\text{p-value} = P(\text{Test statistic as extreme as observed} \mid H_0 \text{ true})$$

- Not the probability that  $H_0$  is true ❌
- Is a conditional probability under  $H_0$

### ◆ Tail direction (defined by $H_1$ )

Symbol	Meaning
$H_0$	Null hypothesis
$H_1$	Alternative hypothesis
$\geq, \leq, \neq$	Decide tail type
(	$z_{\{\text{text}\{\text{obs}\}\}}$

## 7 Significance level

### ◆ $\alpha$

- Meaning: Maximum allowed probability of false rejection
- Common values: 0.05, 0.01
- Decision rule:

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

## 🧠 Big mental picture (VERY important)

SCSS

Data (X)





Sample statistic ( $\bar{X}$ ,  $s$ )



Test statistic ( $Z$  or  $T$ )



Sampling distribution (Normal /  $t$ )



p-value (tail probability)



Decision (compare with  $\alpha$ )

## 🔥 One-line intuition summary

📄 Copy code

Variable type	What it represents
$X_i$	Raw data
$\mu$	True unknown reality
$\mu_0$	Claim being tested
$\bar{X}$	Evidence from data
$Z, T$	Standardized evidence
p-value	Strength of evidence

## Next logical step 🚀

Now that variables are clear, the **next step** is:

👉 How the p-value formula is derived from these variables (integrals + geometry of tails)

Just say **next** and we'll derive it properly.



use above metal picture and curve a real world example that uses all variable so that i can understand when they appear

Great — now we'll attach every variable to **ONE** concrete real-world story, following the **mental picture pipeline** you already saw.

I'll reuse the same symbols and show exactly when and why each one appears.

## Real-world example: Battery life of smartphones

### Situation (business + statistics combined)

A smartphone company claims:

"Average battery life = 10 hours"

You suspect the battery life might be **different** (could be higher or lower).

## Step 0: Mental picture (what's happening conceptually)

SCSS

Reality ( $\mu$ ) —unknown

↓ sample

Data ( $X_1, X_2, \dots, X_n$ )

↓ summarize

Sample mean ( $\bar{X}$ ), variability ( $s$ )

↓ standardize

Test statistic (T or Z)

↓ probability

p-value (tail area)

↓ decision

Reject or not reject  $H_0$

Keep this pipeline in mind.



## Step 1: Define the variables in the real world

### ◆ Raw data


You randomly test 25 phones.

Each phone gives battery life in hours:

$$X_1, X_2, \dots, X_{25}$$

Example:

```
9.4, 10.2, 9.8, 10.1, 9.6, ...
```

 Copy code

→ These are **random variables** because if you test again, values will change.

### ◆ Sample size

$$n = 25$$

Controls uncertainty (larger  $n \rightarrow$  more confidence).



## Step 2: Population parameters (unknown truth)

### ◆ $\mu$

- True average battery life of ALL phones
- Unknown
- This is what we care about

### ◆ $\mu_0$

From the company's claim:

$$\mu_0 = 10 \text{ hours}$$

This is the **reference value** under the null hypothesis.



### Step 3: Hypotheses (this decides tail type!)

You want to check **any difference**, not direction:

$$H_0 : \mu = 10$$

$$H_1 : \mu \neq 10$$



This choice will later force a two-tailed p-value.

---



### Step 4: Compute sample statistics

From your 25 phones:

#### ◆ Sample mean

$$\bar{X} = 9.6 \text{ hours}$$

This is your **observed evidence**.

---

#### ◆ Sample standard deviation

$$s = 1.2 \text{ hours}$$

Measures **spread of battery life**.

---



### Step 5: Build the test statistic (bridge to probability)

Because population  $\sigma$  is **unknown**, we use **t-statistic**:

$$T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

---

Substitute real values:

$$T = \frac{9.6 - 10}{1.2/\sqrt{25}} = \frac{-0.4}{0.24} = -1.67$$

---

#### ◆ $t_{\text{obs}} = -1.67$

This number means:

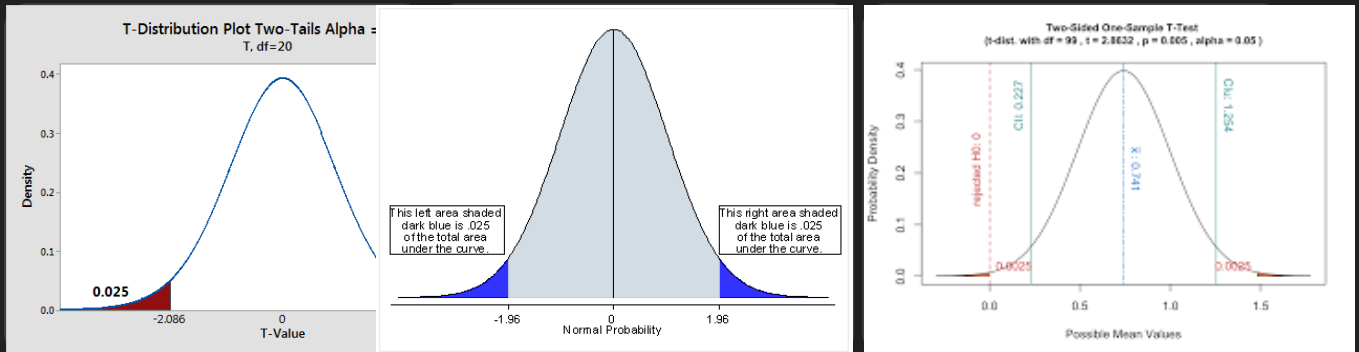
"My sample mean is 1.67 standard errors below the claimed mean."

## Step 6: Sampling distribution (curve appears here)

Under  $H_0$ , the test statistic follows:

$$T \sim t_{(n-1)} = t_{24}$$

So now we place  $-1.67$  on the t-curve:



## Step 7: p-value (this is where tail logic matters)

Because:

$$H_1 : \mu \neq 10$$

We look for extreme values on BOTH sides.

### Mathematical definition

$$\text{p-value} = P(|T| \geq |t_{\text{obs}}| \mid H_0)$$

### Numerically

$$\text{p-value} = 2 \times P(T \leq -1.67) \approx 2 \times 0.054 = 0.108$$

## Step 8: Decision variable

### Significance level

$$\alpha = 0.05$$

### Decision rule

Reject  $H_0$  if p-value  $\leq \alpha$

Here:

$$0.108 > 0.05$$

→ Fail to reject  $H_0$

### Where EACH variable appeared (summary table)

Variable	Appears when	Meaning in story
$X_i$	Data collection	Battery life of each phone
$n$	Experiment design	Number of phones tested
$\mu$	Reality	True average battery life
$\mu_0$	Claim	Company's 10-hour claim
$\bar{X}$	Evidence	Observed average (9.6 hrs)
$s$	Uncertainty	Variability in phones
$T$	Standardization	Distance from claim
$t_{\text{obs}}$	Observation	-1.67
Distribution	Assumption	$t_{(24)}$ curve
p-value	Probability	How extreme this is
$\alpha$	Risk tolerance	5% false alarm rate

### One-line intuition (lock this in)

Every variable exists to answer one question:

"Is my observed deviation from the claim too large to be explained by random sampling?"

## Next step (natural continuation)

- 👉 Same example but as right-tailed & left-tailed
- 👉 Same pipeline applied to ML A/B testing
- 👉 Derive p-value as an integral under the curve

Just say next.

