Z-Test, P-Value, and Critical Value: Explained Intuitively

# 🔍 Introduction

This document walks you through the concepts of Z-test, p-value, and critical value in a simple, intuitive way. It includes a real-world example using OLS (Ordinary Least Squares) regression where you're testing whether an independent variable (X) significantly impacts the dependent variable (Y).

# 📊 Real-World Scenario: Does X Affect Y?

Imagine you're running a regression to understand if a marketing variable X (say, advertisement budget) has an impact on Y (sales).  
  
You formulate hypotheses:  
H₀: X has no effect on Y (coefficient = 0)  
H₁: X has an effect on Y (coefficient ≠ 0)  
  
You run an OLS regression and get a t-statistic of 2.2 and a p-value of 0.03.

**Significance level:** It represent chances of committing Type 1 Error, means rejecting null hypothesis when its true.

If confidence level 95% then alpha – 5%

If confidence level 90%, then alpha – 10%

So, when am setting significance level to be 5% in any test, I am accepting 5% Type 1 error, means rejecting it when it’s true.

Think of it as when am saying I can take risk of 5% of rejecting my null hypothesis when it’s true, that means test values will lie in the outer side of distribution(rejection region), that means 95% of the time, it should fall in acceptance region, which is the middle region of sampling distribution.

**What does a test result or test statistic mean?**

A t-statistic of 2.2 means sample result is 2.2 standard error away from hypothesized population parameter under the null hypothesis.

Let’s say you are doing a one-sample t-test where:

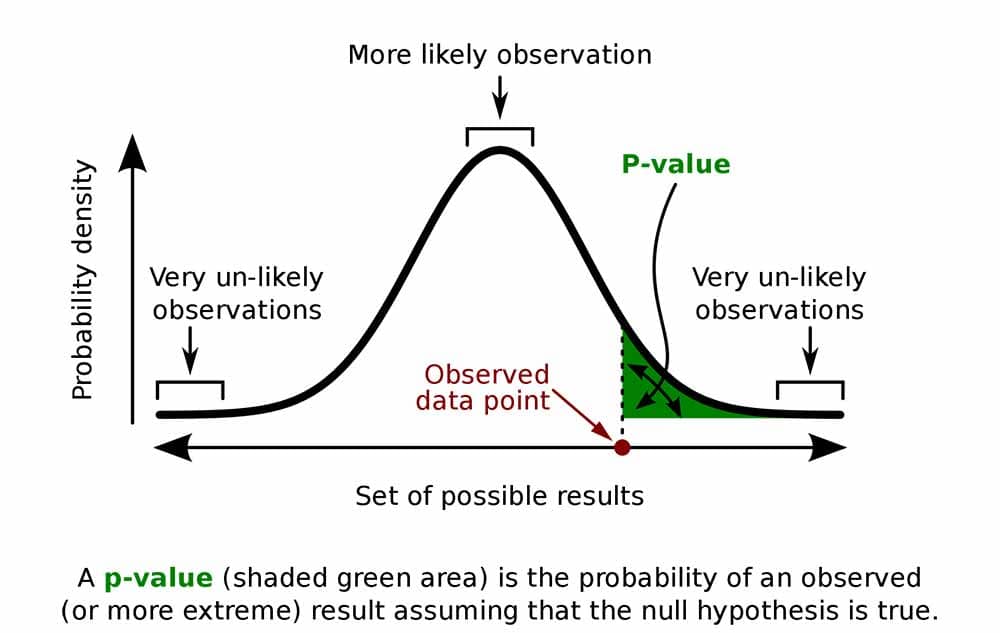
* Null hypothesis H0:μ=100
* Your sample mean xˉ=107
* Standard error (SE) = 3.18

Then:

t=107−1003.18≈2.2

➡️ This tells us that your sample mean (107) is **2.2 SEs above the hypothesized mean** of 100.

# 📉 What Does P-Value Mean?



Assume H₀ is true — i.e., X has no effect on Y. Then the p-value answers:  
“What’s the probability of observing a result this extreme (t = 2.2 or higher) just by chance?”  
  
A p-value of 0.03 means there's only a 3% chance you'd see this extreme result if the null hypothesis were true just by random chance or sampling variability.

This result is rare enough (3%) **under the null**, and I was ready to reject the null for anything rarer than 5%. So I reject it — **knowing there is up to 5% chance of making a Type I error**, not exactly 3%.

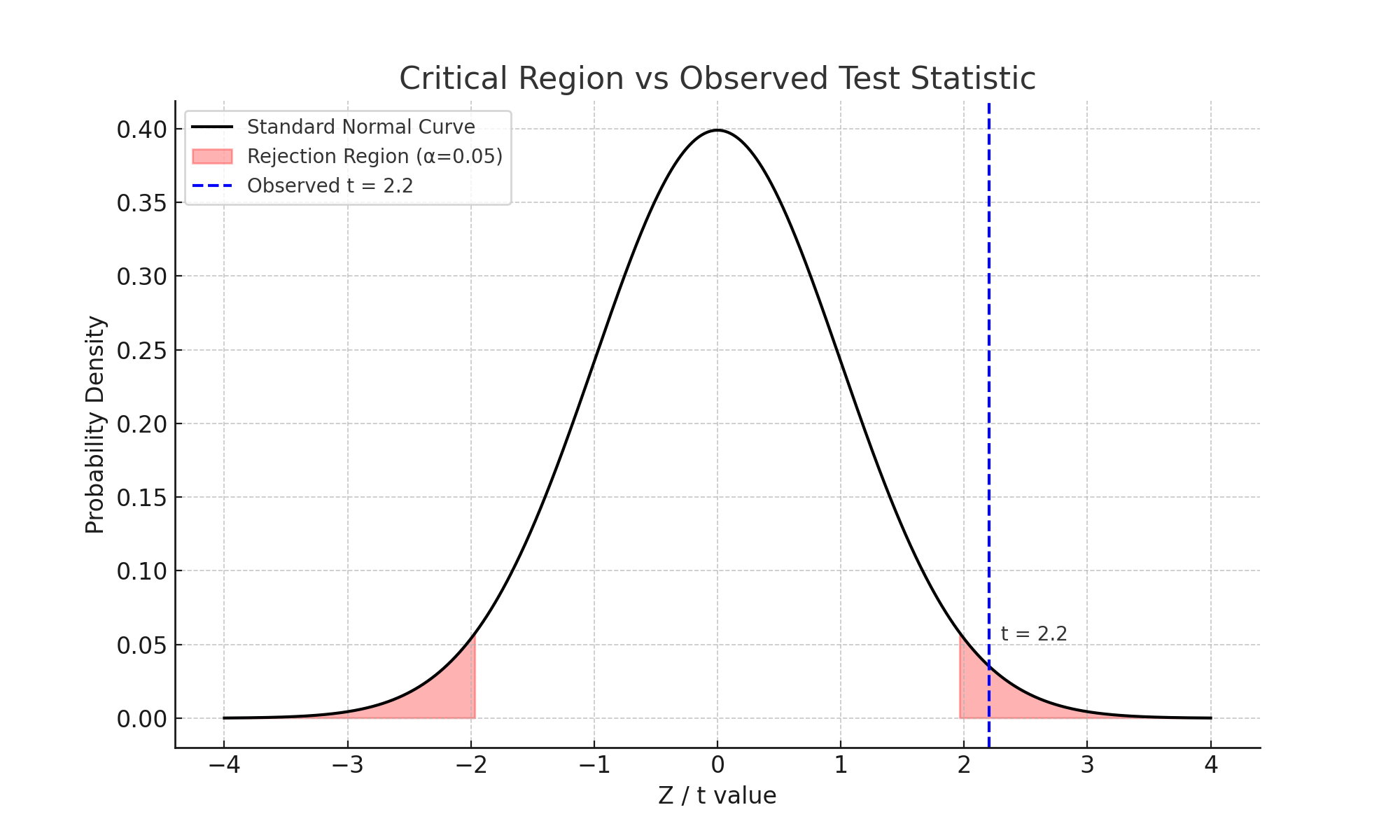
# 📏 What is Critical Value?

The critical value defines the threshold beyond which results are considered rare/extreme enough to reject H₀.  
For a 95% confidence level (α = 0.05), the critical z-values are ±1.96 for a two-tailed test.  
  
If your test statistic is greater than +1.96 or less than -1.96, it falls in the rejection region.

# 🔗 Connection Between P-Value and Critical Value

Both are two sides of the same coin:  
- If test statistic > critical value → reject H₀  
- If p-value < α → reject H₀  
  
In our case: p-value (0.03) < α (0.05) → reject H₀. The result is statistically significant.

# 📊 Visual: Critical Region vs P-Value



# ✅ Summary

• A p-value tells you the probability of seeing such an extreme result assuming H₀ is true.  
• A critical value is a threshold based on your α level that determines whether to reject H₀.  
• In both methods, you're checking whether the result is rare/extreme enough to disbelieve the null.  
• In regression, you use these values to check if a feature (X) significantly impacts your target (Y).

**Final Revised Summary**

When I conduct a hypothesis test, I’m essentially checking whether there is a real, statistically significant difference between groups — in other words, is the observed difference due to something meaningful, or could it just be due to random chance? The null hypothesis (H₀) assumes that there is no real difference, while the alternate hypothesis (H₁) assumes that a true difference exists.

The p-value is the probability of observing a test statistic as extreme as mine — or more extreme — purely by random chance, assuming the null hypothesis is true. It does not tell me whether the null is true or false. Instead, it asks: "If the null hypothesis were true, how unusual would my observed result be?"

For example, a p-value of 0.03 means there’s a 3% chance I’d see a result this extreme just by random variation, assuming H₀ is correct. If this probability is low enough — typically below a significance level (α) of 0.05 — I reject the null, accepting that the result is likely not due to chance. Becasue when p value is less, you will find test result value fall right or left to the critical value, or I can say at the tail of distribution, so this much distance can't happen purely due to chance, so we can say there is a statistical difference.p-value of 0.03 means there’s a 3% chance of observing a test statistic this extreme — or even more extreme — purely by random variation, assuming the null hypothesis (H₀) is true. Now, if this probability is less than the chosen significance level (α = 0.05), we say the result is statistically significant and we reject H₀.

Why? Because low p-values correspond to results that fall far out in the tails of the sampling distribution — that is, they lie far from the expected value under H₀. This much deviation is unlikely to occur just by chance, which makes us question the assumption that H₀ is true.

In simpler terms: When a test result is so far from the null expectation that it lands in the tail of the distribution, we doubt it happened by chance — and thus, we reject the null hypothesis, concluding there’s likely a real difference.

This threshold, α (significance level), represents the maximum risk I’m willing to take for making a Type I error — i.e., rejecting a true null hypothesis. So, when p-value < α, I say the result is statistically significant and I reject H₀, knowing there's still a small (e.g., 5%) chance that I could be wrong.

In summary: Statistical significance means the result is likely not due to random noise. P-value quantifies how surprising my data is, assuming H₀ is true. Significance level (α) is the risk I’m okay with for wrongly rejecting H₀. If p-value < α → I reject H₀ (strong evidence against H₀). If p-value ≥ α → I fail to reject H₀ (not enough evidence against it).

# 🧪 What is a Z-Test?

A Z-test is a type of hypothesis test used to determine whether there is a significant difference between sample data and a known population value (mean or proportion), assuming the population standard deviation is known.  
  
It is most commonly used when:  
• The population standard deviation is known.  
• The sample size is large (n > 30).  
• The sampling distribution is approximately normal.  
  
Z-Test Formula for Mean:  
z = (x̄ - μ) / (σ / √n)  
  
Where:  
• x̄ = sample mean  
• μ = population mean  
• σ = population standard deviation  
• n = sample size  
  
We compare the calculated z-statistic with the critical z-value or use the corresponding p-value to make decisions.

# 📉 How is P-Value Calculated Visually?

To calculate the p-value from the test statistic (e.g., z = 2.2), we calculate the probability of observing a value at least this extreme — i.e., in the tails of the standard normal distribution.  
  
This is done by integrating the area under the normal curve beyond the observed statistic.  
• For a two-tailed test, we calculate the area to the right of +z and to the left of -z and sum them.  
• For a one-tailed test, we only consider one side.  
  
For z = 2.2:  
p-value = P(Z ≥ 2.2) + P(Z ≤ -2.2)  
This area represents how likely we are to see a value this extreme under the null hypothesis.

