BasicHT_P-Value & Critical Value

Tuesday, 3 December 2024 11:15 PM

In hypothesis testing, the p-value and critical value are used to make decision whether to reject the null hypothesis.

Probability (p)-Value:–

It is the probability that the test statistics equals the obsered value or a value even more extreme, assuming that the null hypothesis is true.
It helps determine the strength of the evidence against the null hypothesis.

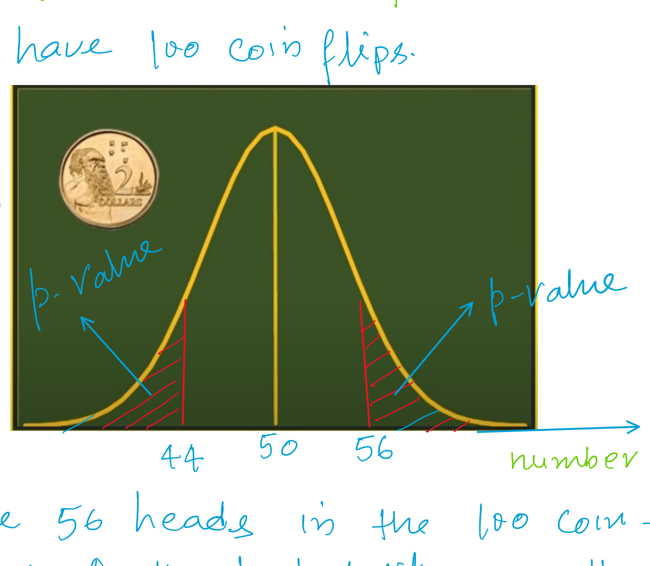
$$\boxed{p\text{-value} = P\left(\text{test statistics takes observed value or beyond in tails} \mid H_0 \text{ true}\right)}$$

Example:- Suppose that we want to test wether a coin is fair. So, we have

H_0: the coin is fair ($P(H) = P(T) = 0.5$)
H_1: The coin is not fair

For this let we have 100 coin flips.

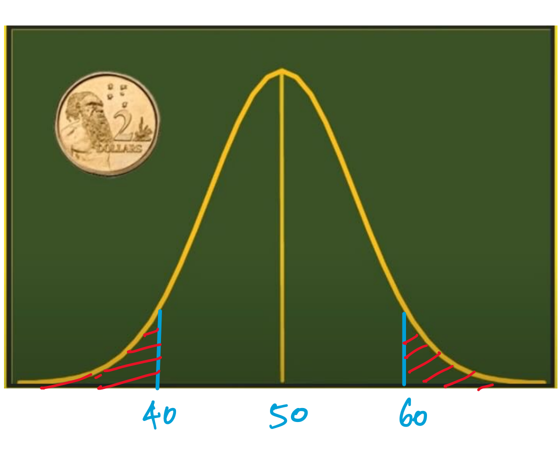
Fig shows the prob distribution of the number of heads.



Let we observe 56 heads in the 100 coin flips. Then, the p-values is the probability of getting a sample (test statistics) equals the observed value (56) or a value even more extreme.

The p-value in the above example is 0.193 (see later how to calculate it). It means that if the coin is fair (Ho is true), the prob of getting 56 heads (or a sample more extreme) is 0.193.

Let us now assume that we get 60 heads in 100 coin flips.



p-value now is 0.032.

The value 0.032 of p-value indicates that if the coin is fair, the prob of getting 60 heads (or a sample more extreme) is 0.032.
We now start thinking that the null hypothesis may not be true because we are getting 60 heads out of 100 coin flips.

p-value (90 heads in 100 coin flips) = 0.0001 (almost zero)

Thus, smaller the p-value, stronger is the evidence against the null hypothesis.

To put it in other words, larger is the difference between sample statistics and hypothesized population parameter, smaller is the p-value and stronger is the evidence against the null hypothesis.

The obvious question now is what should be the p-value to reject a null hypothesis?
The answer is the significance level (α).

Significance level (α):- The significance level α is the threshold for rejecting the null hypothesis when it is actualy true.

Commonly used value of α is 0.05, meaning there is 5% chance of rejecting a null hypotheses when it is actually true. It is set by the researchers.

Interpretation:

1. If $p \leq \alpha$ (significance level), we reject the null hypothesis, suggesting that the observed data is statistically significant.
2. If $p > \alpha$, we fail to reject the null hypo-thesis, meaning that there is not enough evidence to support the alternative hypothesis.

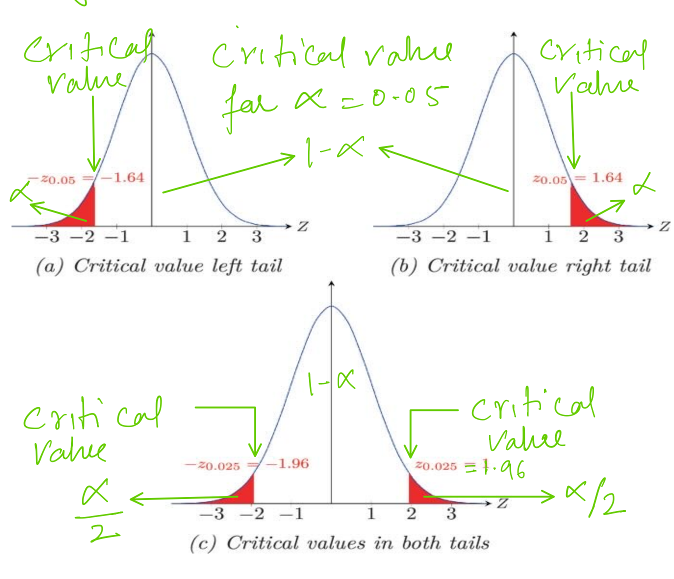
Critical value:– A critical value is a threshold or cut off point that helps in deciding whether to reject the null hypothesis. It is determine by the chosen significance level (α) and the distribution of the test statistics.

The critical value marks the boundary for the region of rejection (also called the critical region). If the test statistics falls beyond this critical value, the null hypothesis is rejected.

Critical value for a two-tailed test:–

Critical value for a two-tailed (two-sided) test for a given α is the Z-score at the value $\alpha/2$.
For example, for $\alpha = 0.05$, the critical value is the Z-score for at $\alpha/2 = 0.025$, which is ± 1.96, as show in Fig.

The Z-values in the red portion give the region of rejection for null hypothesis



y-axis is Z-score.

Critical value for a one-tailed test:-

The critical value for a one-sided (one-tailed) test for a given α is the Z-score at the value α.
For the above example with $\alpha = 0.05$, the critical value for one-side test is the Z-score at $\alpha = 0.05$, which is 1.64 for right tail and −1.64 for the left tail.

Example 1:

Scenario: We want to test if the average score on test for a sample of 15 students is significantly different from 80, with the population standard deviation of 12 and sample mean of 85.

Solution:

Null hypothesis (H_0): $\mu = 80$
Alternative hypothesis (H_1): $\mu \neq 80$ (two-tailed test)

Recall, the Z-statistics for sample is

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{85 - 80}{12/\sqrt{15}}$$

$$\boxed{Z = 1.61}$$

We next need to find p-value from the Z-score
From the Z-score table, $Z = 1.61$ corresponds to the cummulative prob of 0.9463.
The area under the upper tail beyond the Z-value
$1 - 0.9463 = 0.0537$
Since it is two-tailed test, the total p-value
$p\text{-value} = 2 \times 0.0537 = 0.1074$

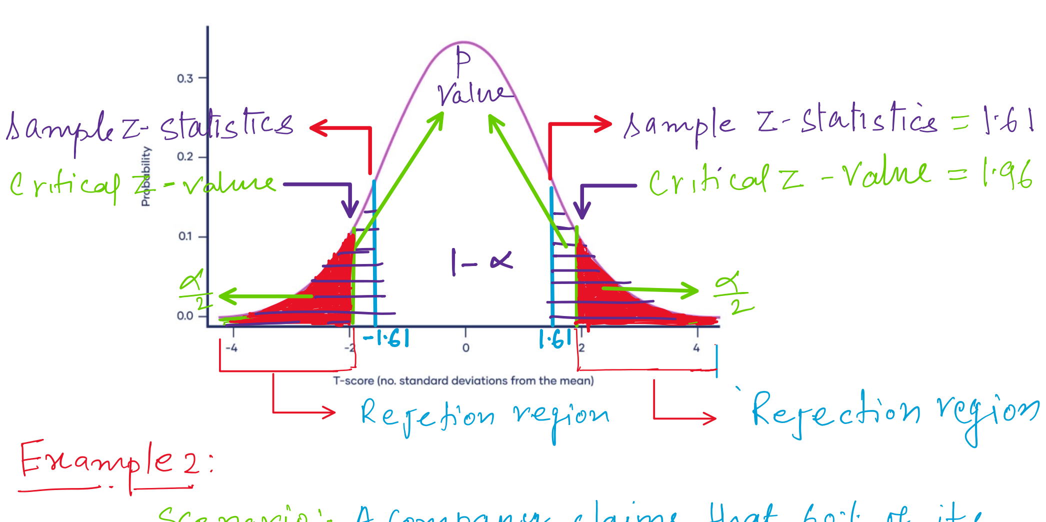
The critical value with $\alpha = 0.05$ for two tail test is 0.025, which leads to critical Z-value of (using the Z-table)

Critical Z-value $= \pm 1.96$

We see that p-value (0.1) > α (0.05), so we fail to reject the null hypothesis.
In other words, there is not enough evidence to conclude that the average score is significantly different from 80.

or

The critical Z-value (1.96) is higher than 1.61, meaning that the Z-statistics does not fall in the rejection region. Thus, there is not enough evidence to reject the null hypothesis.



Example 2:

Scenario:- A company claims that 60% of its customers are satisfied with their services. A survey of 200 customers shows that 130 are satisfied. We want to test if the proportion of satisfied customers is different from the claimed 60% at $\alpha = 0.05$.

Solution:

Null hypothesis (H_0): $P = 0.60$
Alternative hypothesis (H_1): $P \neq 0.60$

The sample size $n = 200$
Recall that the Z-statistics for proportion is

$$Z = \frac{\hat{p}_0 - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}, \quad \hat{p}_0 = \frac{130}{200} = 0.65, \quad p_0 = 0.60, \quad n = 200$$

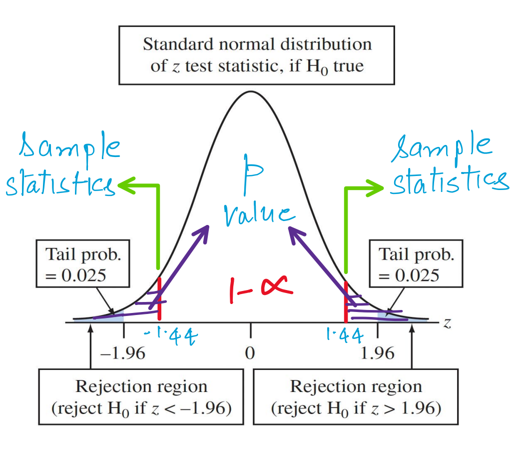
$$Z = \frac{0.65 - 0.60}{\sqrt{\frac{0.6(1-0.6)}{200}}} = \frac{0.05}{0.03464} \approx 1.44$$

From the Z-table, the cummulative prob for $Z = 1.44$ is approximately 0.9251.
The area in the upper tail beyond this Z-value is $1 - 0.9251 = 0.0749$.
Since its a two-tailed test, the total p-value

$$p\text{-value} = 2 \times 0.0749 = 0.1498$$

The critical value with $\alpha = 0.05$ for two tail test is 0.025, which leads to critical Z-value of (using the Z-table)

Critical Z-value $= \pm 1.96$



Since p-value (0.1498) > α (0.05), there is not enough evidence to conclude that the proportion of satisfied customers is different 60%.

or

Since the critical Z-value (1.96) is greater than the Z-statistics (1.44), there is not enough evidence to reject the null hypothesis.