Practice Session | Hypothesis Testing

Tamojit Maiti
Masters in Applied Statistics and Operations Research, ISI Kolkata
Data Scientist at Sixt R&D, previously at Rapido & AB InBev

Critical Region

The critical region (or rejection region) is the set of all values of the test statistic that cause us to reject the null hypothesis

Critical Value

A critical value is any value that separates the critical region (where we reject the null hypothesis) from the values of the test statistic that do not lead to rejection of the null hypothesis. The critical values depend on the nature of the null hypothesis, the sampling distribution that applies, and the significance level $\boldsymbol{\alpha}$.

Two-tailed, Right-tailed, Left-tailed Tests

The tails in a distribution are the extreme regions bounded by critical values.

- Two-tailed test : H_0 : = , H_A : ≠
- Right tailed test : H_0 : = , H_A : >
- Left tailed test : H_0 : = , H_A : <

P-value

The P-value (or p-value or probability value) is the probability of getting a value of the test statistic that is at least as extreme as the one representing the sample data, assuming that the null hypothesis is true. The null hypothesis is rejected if the P-value is very small, such as 0.05 or less.

If a P-value is small enough, then we say the results are statistically significant

Conclusions in Hypothesis Testing based on P-value

We always test the null hypothesis. The initial conclusion will always be one of the following:

- Reject the null hypothesis - if the P-value ≤ $\boldsymbol{\alpha}$ (where $\boldsymbol{\alpha}$ is the significance level, such as 0.05).

- Fail to reject the null hypothesis - if the P-value > $\boldsymbol{\alpha}$

Example 1

An article distributed by the Associated Press included these results from a nationwide survey: Of 880 randomly selected drivers, 56% admitted that they run red lights. The claim is that the majority of all Americans run red lights. That is, $p > 0.5$. The sample data are $n = 880$, and $p = 0.56$.

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.56 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{880}}} = 3.56$$

H0 : $p = 0.5$, HA : $p > 0.5$, $\boldsymbol{\alpha} = 0.05$

We see that for values of $z = 3.50$ and higher, we use 0.9999 for the cumulative area to the left of the test statistic.

The P-value is 1 – 0.9999 = 0.0001. Since the P-value of 0.0001 is less than the significance level of = 0.05, we reject the null hypothesis.

There is sufficient evidence to support the claim.

Example 2

We have a sample of 106 body temperatures having a mean of 98.20°F. Assume that the sample is a simple random sample and that the population standard deviation is known to be 0.62°F. Use a 0.05 significance level to test the common belief that the mean body temperature of healthy adults is equal to 98.6°F..

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}} = \frac{98.2 - 98.6}{\frac{0.62}{\sqrt{106}}} = -6.64$$

H0 : μ = 98.6 , HA : μ ≠ 98.6, $\boldsymbol{\alpha}$ = 0.05 , $\bar{x}$ = 98.2 , σ = 0.62

This is a two-tailed test and the test statistic is to the left of the center, so the P-value is twice the area to the left of z = –6.64. Using the normal cumulative distribution function , the area to the left of z = –6.64 is 0.0001, so the P-value is 2(0.0001) = 0.0002.

Because the P-value of 0.0002 is less than the significance level of 0.05, we reject the null hypothesis. There is sufficient evidence to conclude that the mean body temperature of healthy adults differs from 98.6°F..

Type - I error

A Type I error is the mistake of rejecting the null hypothesis when it is true.

The symbol $\boldsymbol{\alpha}$ (alpha) is used to represent the probability of a type I error.

Type - II error

A Type II error is the mistake of failing to reject the null hypothesis when it is false.

The symbol β (beta) is used to represent the probability of a type II error

Type I and Type II errors

True State of Nature

		The null hypothesis is true	The null hypothesis is false
Decision	We decide to reject the null hypothesis	Type I error (rejecting a null hypothesis when it is true)	Correct Decision
	We fail to reject the null hypothesis	Correct Decision	Type II error (failing to reject a false null hypothesis)

Power of a hypothesis test

The power of a hypothesis test is the probability (1 - β) of rejecting a false null hypothesis, which is computed by using a particular significance level $\boldsymbol{\alpha}$ and a particular value of the population parameter that is an alternative to the value assumed true in the null hypothesis. That is, the power of the hypothesis test is the probability of supporting an alternative hypothesis that is true.