1. Introduction

Welcome to the series of E-learning modules on Critical region and critical function. In this module we are going cover the concept of critical region, critical function and critical values. Its interpretation and role in testing of hypothesis-one tailed and two tailed tests.

By the end of this session, you will be able to:

- Explain critical region
- Describe critical function and critical values
- Interpret the values and understand its importance
- Understand one and two tailed tests

In testing of hypothesis suppose we take a sample of size n from a population X based on which a hypothesis must be tested. Then the set of all possible samples (x one, x two till x n) of size n from the population constitute a sample space.

Sample space is written in symbols as,
Capital Omega equal is to { x one, x two till x n, colon x i belongs to X}
Based on the samples drawn from the population a decision is taken about the acceptance and rejection of null hypothesis.

Let the sample space omega be divided into two groups, omega and omega complement.
Omega complement is equal to capital omega minus small letter omega.
Depending on the null hypothesis and the test, so that when and only when a sample point sample point falls in omega, a null hypothesis shall be rejected. The set or region omega is known as a critical region.
Thus critical region or rejection region is that part of the sample space, where if a sample point falls, a null hypothesis is rejected.

The set of all those sample observations which suggest the acceptance of a null hypothesis is called the acceptance region and is denoted by A.
The set of those sample observations which leads to the rejection of the null hypothesis is called a critical region or rejection region and is denoted by omega or C

Giving a test procedure is nothing but giving a critical region or an acceptance region.
We represent any test procedure as follows:
Phi of x is equal to {greater than zero if (x one, x two, till, x n) belong to A. Less than or equal to zero if (x one, x two till xn) belong to C}.

A function Phi of x is known as test function.

Probability of a selected sample belonging to the critical region is called the size of the critical region.
Size of the critical region is nothing but the size of the test or level of significance that is probability of rejecting a null hypothesis when it is true.

A statistical hypothesis test specifies a critical region that is, a set of numbers. If the observed data is in the critical region, then reject the maintained hypothesis.
If the observed data is NOT in the critical region, then accept the maintained hypothesis.

2. Example

Example: Let the critical region be zero, one, two, nine, ten, and eleven heads. If you flip the coin eleven times, reject the maintained hypothesis that the number of heads is a binomial distribution with probability zero point five of heads and n equal to eleven if you observe zero, one, two, nine, ten, and eleven heads.

Accepting the maintained hypothesis does not prove it to be true and rejecting the maintained hypothesis does not prove it to be false.
Similarly, accepting the alternate hypothesis does not prove it to be true and rejecting the alternative does not prove it to be false.

Many authors use '*fail to reject*' so that students will not think that something was proved with a statistical hypothesis. But the only meaning that *'fail to reject'* can have in statistical hypothesis testing is accept.

There are only two outcomes of a statistical test. The data is either in the critical region or it is not in the critical region. If the data is NOT in the critical region, you accept the maintained and reject the alternative.
Reject the alternative must mean accept the maintained.
Fail to reject the maintained means accept the maintained.

If *'fail to reject'* had any real meaning other than accept, then *'fail to accept'* would also have a different meaning.
Now you would have four outcomes: accept the maintained, reject the maintained, fail to reject the maintained, fail to accept the maintained.
However there are only two outcomes of a statistical test: the data is either in the critical region or it is not in the critical region.

The only outcomes are to accept the maintained (that is, reject the alternative) or accept the alternative (or reject the maintained).
Fail to reject must mean accept and fail to accept must mean reject.

For some statistical hypotheses, there are many tests.
Homoscedasticity, non serially correlated errors and even normality of a random variable are all a hypothesis with many different tests.
In some sense, saying 'fail to reject' means that the current test accepted the maintained (or null) but some other test might reject the maintained.
Then, 'fail to reject' is not about a hypothesis test, but about many hypotheses tests.

However, in such a case there are many critical regions (as many as there are tests). ACCEPT or REJECT is about one single critical region. We will discuss distinguishing among hypothesis tests, but we will not use 'fail to reject'.

If accepting a hypothesis does not prove the hypothesis, then what does it do? Acceptance allows one to proceed as if the hypothesis were true.
But there are two outcomes: Accepting a true hypothesis and accepting a false hypothesis.

Accepting a true hypothesis would be a 'correct' decision while rejecting a true hypothesis would be an 'incorrect' decision.
Rather than call it incorrect, we call it an error.

3. Examples and Critical Function

The probability of making a Type two error is the probability that the data is NOT in the critical region conditional upon assuming the alternative hypothesis.
Example: The critical region is zero, one, two, nine, ten and eleven.
The probability of zero, one, two, nine, ten, or eleven heads occurring given the number of heads is a Binomial (zero point five, eleven) is equal to zero point zero six five four.
The probability of a Type one error for this critical region is zero point zero six five four. It is the probability that we observe zero, one, two, nine, ten, or eleven heads in eleven flips assuming the flips are a binomial distribution with n equal to eleven and p equal to zero point five.

If we observe zero, one, two, nine, ten, or eleven heads, we reject the null hypothesis and accept the alternative hypothesis.
If we observe three, four, five, six, seven, or eight heads we accept the null hypothesis and reject the alternative hypothesis.

What is the alternative hypothesis?
In this example, and in most real situations, the alternative hypothesis is the negation of the null hypothesis.

In this example, the alternative hypothesis is that the number of heads was not generated according to a binomial distribution with probability zero point five and n equal to eleven.
We know that there were eleven trials.
We do not know whether it was a binomial distribution and we do not know whether the probability was zero point five.

One alternative is the data was generated by a different distribution.
For example, the data could have been generated by flipping the coin until two heads were observed and it took eleven trials.
That distribution (that is, flipping until a certain number of successes are observed) is called the negative binomial.
Another more common alternative hypothesis in the example is that the distribution is binomial but the probability is not zero point five.

Unless the alternative hypothesis is specified, we cannot know what the probability of a Type two error (that is rejecting a true alternative) is.
In most real life cases, the alternative is complex and hence the probability of Type error error is unknown.

The size (or significance level) of a statistical test is the probability of a Type one error.
The power of a test is one minus the probability of a Type two error. For a given size, we want a statistical test with the highest power.

For most tests we encounter, we specify a size, we obtain a critical region and theoretical results indicate what alternatives have high power and what alternatives do not.
For most tests we encounter, we never know the probability of a Type two error. We rely on

prior research to tell us what tests are powerful against what alternatives.

Both the size and the power are probabilities of the critical region.
The difference is the assumption made to compute the probability.
For the size, the probability is computed assuming the null hypothesis.
For the power, the probability is computed assuming some alternative hypothesis.

Therefore, let's return finally to the question of whether we
(a) reject or accept the null hypothesis; and
(b) If we reject the null hypothesis, do we accept the alternative hypothesis?

If our statistical analysis shows that the two distributions are the same at the significance level (either zero point zero five or zero point zero one) that we have set, we simply accept the null hypothesis.
Alternatively, if the two distributions are different, we need to either accept or reject the alternative hypothesis.
This will depend on whether we made a one- or two-tailed prediction.

Critical function:
With each test for testing a null hypothesis against an alternative hypothesis a function is associated called the critical function of a test.
The critical function of a test is that function which gives the probability of rejecting the null hypothesis.

Suppose we have a population with an unknown parameter theta, then the critical function is a function of theta which is given as probability of reject H naught given theta.

Critical Value(s)
The critical value(s) for a hypothesis test is a threshold to which the value of the test statistic in a sample is compared, to determine whether or not the null hypothesis is rejected.

The critical value for any hypothesis test depends upon a test statistic, which is specific to the type of the test and the significance level alpha which defines the sensitivity of the test at which the test is carried out, and whether the test is one-sided or two-sided.

A value of alpha is equal to zero point zero five implies that the null hypothesis is rejected five percent of the time when it is in fact true.
Critical values are essentially cut-off values that define regions where the test statistic is unlikely to lie.
For example, a region where the critical value is exceeds the probability alpha if the null hypothesis is true. The null hypothesis is rejected if the test statistic lies within this region which is often referred to as the rejection region(s).

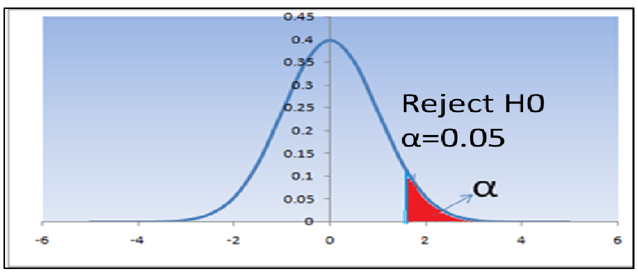
4. One Tailed Test

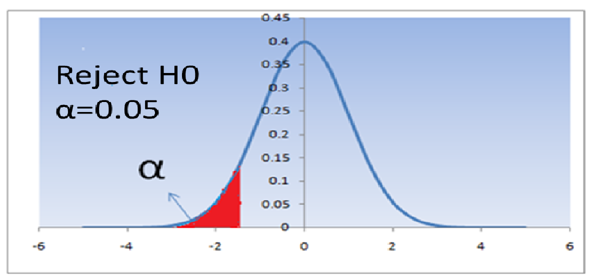
One tailed or one sided tests
A test of a statistical hypothesis, where the region of rejection is on only one side of the sampling distribution, is called a **one-tailed test**.

For example, suppose the null hypothesis states that the mean is less than or equal to twenty. The alternative hypothesis would be that the mean is greater than twenty. The region of rejection would consist of a range of numbers located on the right side of sampling distribution; that is, a set of numbers greater than twenty.

A one-sided test is a statistical hypothesis test, in which the values for which we can reject the null hypothesis, H naught are located entirely in one tail of the probability distribution.
In other words, the critical region for a one-sided test is the set of values less than the critical value of the test, or the set of values greater than the critical value of the test.

Figure 1





Just have a look at this figure, one tailed tests concentrating on the right and left side tail is formatted here. Accordingly we keep the level of significance completely either on the right tail or on the left tail depending upon the situation.
If the alternative hypothesis is of the greater than type, then the critical region is towards the right side of the curve.
On the other hand if our alternative hypothesis is less than type, then the critical region is towards the left side of the curve.
A one-sided test is also referred to as a one-tailed test of significance. The choice between a one-sided and a two-sided test is determined by the purpose of the investigation or prior reasons for using a one-sided test.

Example
Suppose we wanted to test a producer's claim that there are, on average, fifty five matches in

a box. We could set up the following hypotheses .H naught is that mu is equal to fifty five , against the alternative hypothesis that mu is less than fifty five or that mu is greater than fifty five.
Either of these two alternative hypotheses would lead to a one-sided test.

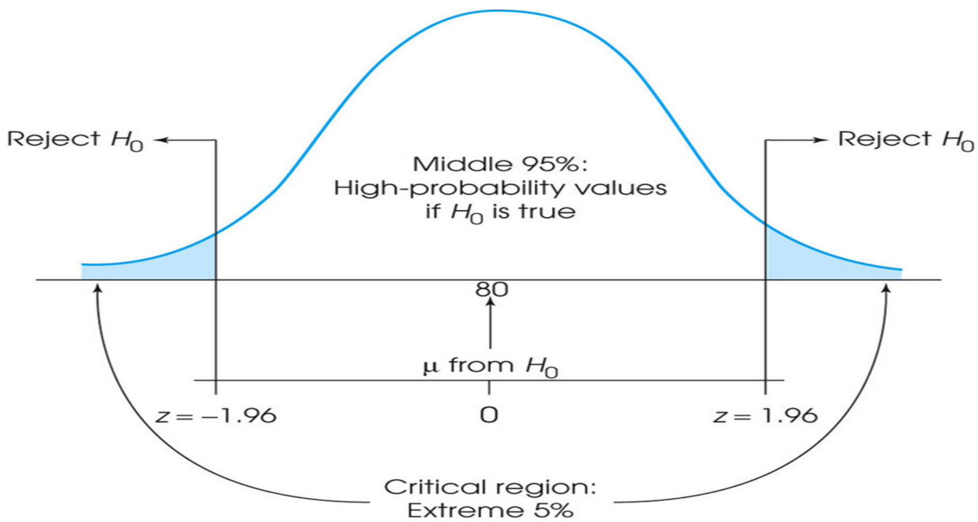
Presumably, we would want to test the null hypothesis against the first alternative hypothesis since it would be useful to know if there is likely to be less than fifty five matches, on average, in a box (no one would complain if they get the correct number of matches in a box or more).

5. Two Tailed Tests and Summary

Two tailed tests
A test of a statistical hypothesis, where the region of rejection is on both sides of the sampling distribution, is called a **two-tailed test**.
For example, suppose the null hypothesis states that the mean is equal to twenty.
The alternative hypothesis would be that the mean is less than twenty or greater than twenty.

The region of rejection would consist of a range of numbers located on both sides of sampling distribution
That is, the region of rejection would consist partly of numbers that were less than twenty and partly of numbers that were greater than twenty.

Figure 2



A two-sided test is a statistical hypothesis test in which the values for which we can reject the null hypothesis, H naught are located in both tails of the probability distribution.
In this case we describe the level of significance on either side of the tail equally.
For example if the level of significance alpha equal to zero point zero five , then we distribute this amount equally at the rate zero point zero two five on either side of the tail. The value of standard normal variate for alpha equal to zero point zero two five is one point nine six as shown in the figure. In other words, the critical region for a two-sided test is the set of values less than a first critical value of the test and the set of values greater than a second critical value of the test.
A two-sided test is also referred to as a two-tailed test of significance.
The choice between a one-sided test and a two-sided test is determined by the purpose of the investigation or prior reasons for using a one-sided test.

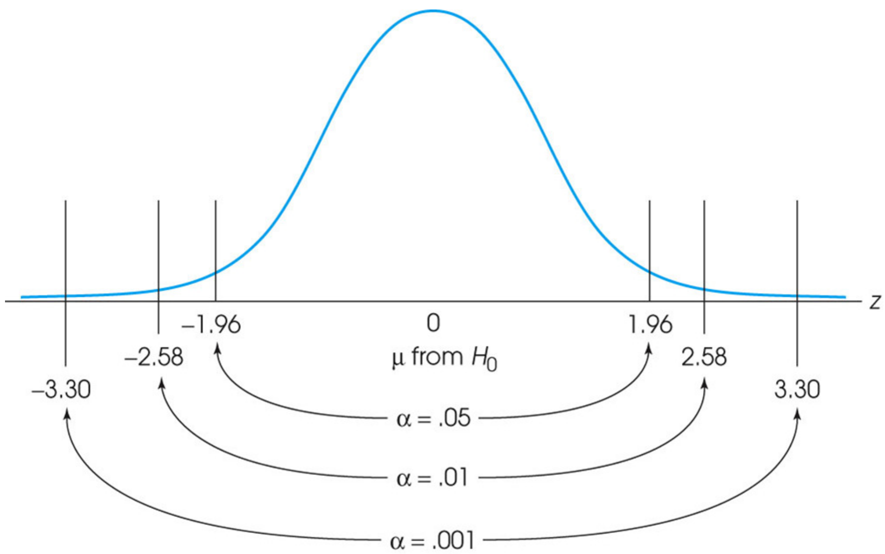
Example
Suppose we wanted to test a manufacturer’s claim that there are, on average, fifty five matches in a box. We could set up the following hypotheses. The null hypothesis is that mu is equal to fifty five.

An alternative hypothesis could be tested against the above null, leading this time to a two-sided test: H one is that mu is not equal to fifty five.

Here, nothing specific can be said about the average number of matches in a box; only that, if we could reject the null hypothesis in our test, we would know that the average number of matches in a box is likely to be less than or greater than fifty five.

The relationship between the probabilities of taking two decisions in terms of null hypothesis whether to accept or to reject the null hypothesis can be illustrated using the sampling distribution when the null hypothesis is true.

Figure 3



The decision point is set by alpha, the area in the tail or tails of the distribution. Setting alpha smaller moves the decision point further into the tails of the distribution.

Summary:
State the hypotheses and select an alpha level.
The null hypothesis, always states that the treatment has no effect (no change, no difference).
According to the null hypothesis, the population mean after treatment is the same is it was before treatment.
The alpha level establishes a criterion, or "cut-off", for making a decision about the null hypothesis.

Locate the critical region. The critical region consists of outcomes that are very unlikely to occur if the null hypothesis is true.
That is, the critical region is defined by sample means that are almost impossible to obtain if the treatment has no effect.
The phrase “almost impossible” means that these samples have a probability (p) that is less than the alpha level.

Hence the critical region CR, or rejection region RR, is a set of values of the test statistic for which the null hypothesis is rejected in a hypothesis test. That is, the sample space for the test statistic is partitioned into two regions; one region (the critical region) will lead us to reject the null hypothesis H 0, the other will not. So, if the observed value of the test statistic is a member of the critical region, we conclude "Reject H naught"; if it is not a member of the critical region then we conclude "Do not reject H naught".

Here’s a summary of our learning in this session where have understood the following concepts:

- Critical region
- Critical values and critical function
- Interpretation and importance of critical region
- One tailed and Two tailed tests