1. Introduction

Welcome to the series of E-learning modules on Likelihood Ratio Test Procedure. In this module we are going cover the basic terminology, concepts and principles of Likelihood Ratio Test Procedure, properties and the test criteria of Likelihood Ratio Test Procedure.

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

- Describe the basic terminologies of Likelihood Ratio Test Procedure(or LRTP)
- Explain the fundamental concept and principle of LRTP
- Describe the properties and test procedure of LRTP

In any testing problem, the following should be obvious from the description of the problem:

- The form of the population distribution
- The parameters of interest and
- The framing of a null and alternative hypothesis

The most crucial step is the choice of the best test that is the best statistic and the critical region C
By best we mean the one which in addition to controlling alpha at any desired low level has the minimum type two error beta, or maximum power one minus beta compared to beta of all other tests having this alpha.

This leads to Most Powerful test and Uniformly Most Powerful tests.
Most Powerful test is the test procedure used to test the simple null hypothesis against a simple alternative hypothesis.
However when we need to test a simple or composite null hypothesis against a simple or composite alternative hypothesis the test procedure adopted is known as a Uniformly Most Powerful test.

Neymann Pearson Lemma based on the magnitude of the ratio of the two probability density functions provides the best test for testing a simple null hypothesis against a simple alternative hypothesis.
The best test in any given situation depends upon the *nature of the population distribution* and the *form of the alternative hypothesis* being considered.
In this topic we shall discuss a general method of test construction called the Likelihood Ratio.

A test was introduced by Neyman and Pearson for testing a hypothesis simple or composite against a simple or composite alternative hypothesis. This test is related to the maximum likelihood estimates.
Before defining the test let us discuss about some notations and terminologies.

Parameter space:
Let us consider a random variable X with probability density function f of (x, theta).
In most common applications, though not always, the functional form of population distribution is assumed to be known except for the value of some unknown parameter theta which may take any value on a set omega.
This is expressed by writing a probability density function in the form f of (x coma theta) coma

theta belongs to omega.

The set omega which is the set of all possible values of theta is called the parameter space. Such a situation gives rise not to one probability distribution but a family of probability distributions which we write as {f of (x coma theta), theta belongs to omega}.
For example: X follows Normal distribution with mean mu and variance sigma square , then the parameter space Omega is equal to { (mu coma variance sigma square) colon minus infinity less than mu less than infinity, where sigma lies between zero and infinity}

In the following discussion we shall consider a general family of distributions
{f of (x , theta one, theta two, etc till theta k), theta i belongs to omega and i equal to one, two, etc till k}
Let x one, x two, etc till x n be a random sample of size n greater than one from a population with p. d. f,
F of (x, theta one, theta two, etc till theta k), where omega, the parameter space, is the totality of all the points that (theta one, theta two, etc till theta k), can assume.

Then we are interested to test the null hypothesis that (theta one, theta two, etc till theta k) belongs to the parameter space under the null hypothesis against alternative hypothesis of the type
H one: (theta one, theta two, etc till theta k) belongs to the parameter space under the alternative hypothesis where parameter space under the alternative hypothesis is equal to Omega minus parameter space under the null hypothesis.

Hence the entire parameter space Omega can be thought to be composed of two sub parameter spaces, namely parameter space under the null hypothesis and parameter space under the alternative hypothesis.
That is, Omega is equal to parameter space under the null hypothesis union parameter space under the alternative hypothesis.
Where the symbol "H circle naught" is the parameter space under the null hypothesis H naught which belongs to the parameter space omega and the symbol "H circle one" is the parameter space under the alternative hypothesis H one which is also a part of the parameter space omega.

2. Maximum Likelihood Estimates

A likelihood function of a number of sample observations is defined to be their joint density function.
To obtain a maximum likelihood estimate, one first specifies the joint density function for all observations.
For an independently identically distributed sample, this joint density function is, F of (x one, x two, etc till x n, theta) is equal to f of (x one , theta) into f of (x two , theta) into etc f of (x n , theta).

In case the sample observations are independent, the likelihood function happens to be the product of the density functions of the random observations. Then, the likelihood function of the random observations denoted by L is equal to L of ((x one, x two, etc till x n, theta) is given by product of f of (x i, theta); i runs from one to n, where f of (x i, theta) is the probability density function of the population.

The method of maximum likelihood is the one in which for a given set of values x one, x two, etc till x n) an estimator of theta is found that maximizes L.
Thus if there exists theta cap, a function of x one, x two, etc till x n) for which L is maximum for variations in theta .
Which implies theta cap is the m.l.e of theta, and then Delta L by delta theta is equal to zero and delta square L by delta theta square is less than zero.

Substituting these estimates we obtain the maximum values of the likelihood function for the variation of the parameters in the parameter space.
Then the criteria for the likelihood ratio tests are defined as the quotient of these two maxima that is maximum under parameter space under the null hypothesis and a total parameter space.

In statistics, a likelihood ratio test is a statistical test used to compare the fit of two models, one of which (the null model*)* is a special case of the other (the alternative model).
The test is based on the likelihood ratio, which expresses how many times more likely the data are under one model than the other.

This likelihood ratio, or equivalently its logarithm, can then be used to compute a p-value, or compared to a critical value to decide whether to reject the null model in favour of the alternative model.

When the logarithm of the likelihood ratio is used, the statistic is known as a *log-likelihood ratio statistic*, and the probability distribution of this test statistic, assuming that the null model is true, can be approximated using Wilks' theorem.

In the case of distinguishing between two models, each of which has no unknown parameters, use of the likelihood ratio test can be justified by the Neyman–Pearson lemma, which demonstrates that such a test has the highest power among all competitors.

Suppose a composite null hypothesis H naught that theta belongs to the parameter space

under the null hypothesis H naught, is to be tested against a composite alternative hypothesis H one that theta belongs to parameter space under the alternative hypothesis H one.
For testing the above null hypothesis a test procedure called Likelihood ratio test procedure is followed which is explained as follows:

Let a random sample x one, x two etc till x n of size n be drawn from the given population with probability density function, f of (x coma theta).
Let lambda be equal to Supremum of L of (theta , x one, x two, etc till x n) when theta belongs to the parameter space under the null hypothesis H naught divided by Supremum of L of (theta , x one, x two, etc till x n) when theta belongs to the entire parameter space omega .

Since the supremum in the denominator is over a larger set of numbers, Lamda less than or equal to one.
Also likelihood functions are nonnegative and hence lambda greater than or equal to zero.
Thus zero less than or equal to lambda less than or equal to one.

That is, the test procedure is as follows.
If lambda is very large that is lambda greater than or equal to lambda alpha, the null hypothesis is accepted and if lambda is very small that is lambda less than lambda alpha the null hypothesis is to be rejected.
The constant lambda alpha is so chosen that the size of the test is alpha.

3. The Principle of LRTP

If the null hypothesis is true the sample drawn must have a larger likelihood under null hypothesis so that the numerator of lambda may be larger and consequently lambda may be larger.
So for higher values of lambda the null hypothesis may be accepted and if the value of lambda is small, the hypothesis may be rejected.

Background
The likelihood ratio, often denoted by the capital Greek letter lambda, is the ratio of the likelihood function varying the parameters over two different sets in the numerator and denominator.

A likelihood-ratio test is a statistical test for making a decision between two hypotheses based on the value of this ratio.
It is central to the Neyman–Pearson approach to statistical hypothesis testing, and, like statistical hypothesis testing generally, is both widely used and much criticized.

Interpretation
Being a function of the data xi, the Likelihood ratio lambda is therefore a statistic.
The likelihood-ratio test rejects the null hypothesis if the value of this statistic is too small.
How small is too small depends on the significance level of the test, that is, on what probability of Type one error is considered tolerable ("Type one" errors consist of the rejection of a null hypothesis that is true).

The numerator corresponds to the maximum likelihood of an observed outcome under the null hypothesis.
The denominator corresponds to the maximum likelihood of an observed outcome varying parameters over the whole parameter space.
The numerator of this ratio is less than the denominator.
The likelihood ratio hence is between zero and one.

Lower values of the likelihood ratio mean that the observed result was much less likely to occur under the null hypothesis as compared to the alternative.

Higher values of the statistic mean that the observed outcome was more than or equally likely or nearly as likely to occur under the null hypothesis as compared to the alternative and the null hypothesis cannot be rejected.

Hence the statistic lambda is bound to be between zero and one. (Likelihoods are non-negative and the likelihood of the smaller model can't exceed that of the larger model because it is *nested* on it).

Values close to zero indicate that the smaller model is not acceptable, compared to the larger model, because it would make the observed data very unlikely. Values close to one indicate that the smaller model is almost as good as the large model, making the data just as likely.

4. Assumptions of the Model

Assumptions of the model can be tested using LRTP

1. Calculate the maximum likelihood of the sample data based on an assumed distribution model (the maximum occurs when unknown parameters are replaced by their maximum likelihood estimates).
Repeat this calculation for other candidate distribution models that also appear to fit the data (based on probability plots).

If all the models have the same number of unknown parameters, and there is no convincing reason to choose one particular model over another based on the failure mechanism or previous successful analyses, then pick the model with the largest likelihood value.

2. Many model assumptions can be viewed as putting restrictions on the parameters in a likelihood expression that effectively reduce the total number of unknown parameters.

Example where assumptions can be tested by the Likelihood Ratio Test
It is suspected that a type of data, typically modelled by a Weibull distribution, can be fit adequately by an exponential model.
The exponential distribution is a special case of the Weibull, with the shape parameter gamma set to one.

If we write the Weibull likelihood function for the data, the exponential model likelihood function is obtained by setting gamma to one, and the number of unknown parameters has been reduced from two to one.

Clearly, we could come up with many more examples like this, for which an important assumption can be restated as a reduction or restriction on the number of parameters used to formulate the likelihood function of the data.
In all these cases, there is a simple and very useful way to test whether the assumption is consistent with the data.

Alternate Method for The Likelihood Ratio Test Procedure
Let L one be the maximum value of the likelihood of the data without the additional assumption.
In other words, L one is the likelihood of the data with all the parameters unrestricted and maximum likelihood estimates substituted for these parameters.

Let L naught be the maximum value of the likelihood when the parameters are restricted (and reduced in number) based on the assumption.
Assume *k* parameters were lost (i.e., L naught has *k* less parameters than L one).

Form the ratio lambda is equal to L naught by L one.
This ratio is always between zero and one and the less likely the assumption is, the smaller lambda will be.
Under certain regularity conditions, minus twice the log of the likelihood ratio has

approximately in large samples a chi-square distribution with degrees of freedom equal to the difference in the number of parameters between the two models.

Since zero less than or equal to lambda less than or equal to one, minus 2 Ln lambda is an increasing function of lambda and approaches infinity when lambda tends to zero.
The critical region for minus two Ln lambda being the right hand tail of the Chi square distribution.

This can be quantified at a given confidence level as follows:

- Calculate chi square is equal to minus two Ln lambda. The smaller lambda is, the larger chi square will be.
- We can tell when chi square is significantly large by comparing it to the hundred into (one minus alpha) percentile point of a Chi Square distribution with *k* degrees of freedom.
- The likelihood ratio test computes chi square and rejects the assumption if chi square is larger than a Chi-Square percentile with *k* degrees of freedom, where the percentile corresponds to the confidence level chosen by the analyst.

5. Demerits, Applications and Properties

Demerits:
While Likelihood Ratio test procedures are very useful and widely applicable, the computations are difficult to perform by hand, especially for censored data, and appropriate software is necessary.

The tests discussed in this paper are asymptotically equivalent, and are therefore expected to give similar results in large samples.
Their small-sample properties are not known, but some simulation studies suggest that the likelihood ratio test may be better than its competitors in small samples.

Applications:
Likelihood ratio may refer to:
- The ratio of two likelihood functions
- Likelihood-ratio test, a statistical test for comparing two models
- Likelihood ratios in diagnostic testing, ratios based on sensitivity and specificity, used to assess diagnostic tests
- Bayes factor, ratio of likelihoods used to update prior probabilities to posterior in Bayes' theorem and Bayesian inference

Properties of LRTP
Likelihood Ratio Test principle is an intuitive one.
If we are testing a simple null hypothesis against a simple alternative hypothesis, then the Likelihood Ratio Principle leads to the same test as given by Neyman Pearson Lemma.
This suggests that Likelihood Ratio Test has some desirable properties, especially large sample properties.

In Likelihood ratio test the probability of Type one error is controlled by suitably choosing the cut off point lambda alpha.
Likelihood ratio test is generally Uniformly Most Powerful if an UMP test at all exists.

The important asymptotic properties of Likelihood ratio test are:
- Under certain conditions minus two Ln lambda has an asymptotic Chi square distribution that is minus two Ln lambda is distributed as central Chi square when null hypothesis holds
- Under certain conditions Likelihood ratio test is consistent
- LR test procedure is at least asymptotically unbiased

This is because as 'n' tends to infinity P T of theta tends to one for all theta belongs to the parameter space under alternative hypothesis by consistency and therefore infimum of P T of theta when theta belongs to the parameter space under alternative hypothesis is equal to one is greater than or equal to Supremum of P T of theta when theta belongs to the parameter space under null hypothesis.

Here’s a summary of our learning in this session:

- Basic concept of terminologies used in Likelihood ratio teat
- Principle and Properties of the test
- Procedure to construct likelihood ratio test and its application