

Essentials of Data Science With R Software - 1

Probability and Statistical Inference

Probability Theory

:::

Lecture 8

Introduction to Probability

Shalabh

Department of Mathematics and Statistics

Indian Institute of Technology Kanpur

Introduction to Statistics:

It has become a rule rather than exception that if we want to learn and know about any phenomenon or a process, one must collect data and learn from there.

Statistics is the science of learning from data.

It is related to the collection of data and then extracting the hidden information by its descriptive analysis and drawing of conclusions.

Introduction to Statistics:

Sometimes a statistical analysis begins with a given set of data

Statistics describes, summarizes, and analyse the data.

In case, the data is not available, in such cases, the statistical design of experiment is appropriately used to generate data.

At the end of the experiment, the data is described and summarized using the tools of descriptive statistics.

Inferential Statistics:

After completing the experiment, data is described and summarized with an aim to draw a statistical conclusion using the tools of inferential statistics.

The concept of chance is considered and utilized to draw a conclusion from the data.

Some assumptions are made about the chances/probabilities of obtaining the different data values. These assumptions are referred to as probability model for the data.

Inferential Statistics:

A careful description and presentation of the data enable us to infer an appropriate probability model for a given data set which can be verified by using the additional data.

The tools of statistical inference lay the foundation of the formulation of a probability model to describe the data.

Thus an understanding of statistical inference data to make valid inferences requires knowledge of the theory of probability.

Elements of Probability:

The probability of an event is subjected to various meanings or interpretations depending upon how one says it.

For example, if a medical doctor says that “there are 70% percent chances that the patient will be cured,” gives some intuitive idea about the success of treatment.

One can also say that the doctor feels that, over the long run, in 70% percent of such ailments, the patients have recovered

Interpretation of Probability:

Probability: Measure of uncertainty

Broadly, there are two types of interpretation of probability

- 1. Frequency interpretation**
- 2. Subjective interpretation of probability.**

Frequency Interpretation of Probability:

The probability of a given outcome of an experiment indicates a “property” of that outcome.

Such a property can be determined by continual repetition of the experiment.

A popular interpretation of probability is as follows:

The probability of the outcome is observed as the proportion of the experiments that result in the outcome.

Subjective Interpretation of Probability:

In the subjective interpretation, the probability of an outcome is not thought of as being a property of the outcome.

It is considered as a statement about the beliefs of the person who is quoting the probability.

The probability is about the chances of occurrence of the outcome.

Subjective Interpretation of Probability:

So probability becomes a subjective or personal concept and has no meaning outside of expressing one's degree of belief.

This interpretation of probability is often favored by decision makers.

Interpretation of Probability:

Irrespective of frequency or subjective interpretation of probability, practically there is a consensus that the mathematics of probability are the same in either case.

For example, if we say that the probability of raining tomorrow is 0.8, then we feel that

- 80% chances are there for the rain tomorrow and the expected weather will be cloudy.**
- 20% chances are there for the not raining tomorrow and the expected weather will not be cloudy.**