Essentials of Data Science With R Software - 1

Probability and Statistical Inference

Probability Theory

:::

Lecture 13

Axiomatic Definition of Probability

Shalab h

Department of Mathematics and Statistics

Indian Institute of Technology Kanpur

Relative Frequency and Probability of an Event:

If we assume that

- **the experiment is repeated a large number of times (mathematically, this would mean that *n* tends to infinity) and**
- **the experimental conditions remain the same (at least approximately) over all the repetitions,**

then the relative frequency f(*A*) converges to a limiting value for *A*.

This limiting value is interpreted as the probability of *A* and denoted by

$$P(A) = \lim_{n \to \infty} \frac{n(A)}{n}$$

where *n*(*A*) denotes the number of times an event *A* occurs out of *n* times.

Relative Frequency and Probability of an Event:

Although this definition is certainly intuitively pleasing and but it possesses a serious drawback:

How do we know that $n(A)/n$ will converge to some constant limiting value that will be the same for each possible sequence of repetitions of the experiment?

Relative Frequency and Probability of an Event:

For example, a coin is continuously tossed repeatedly.

1. **How do we know that the proportion of heads obtained in the first *n* tosses will converge to some value as *n* gets large?**
2. **Even if it converges to some value, how do we know that, if the experiment is repeatedly performed a second time, we will again obtain the same limiting proportion of heads?**

Relative Frequency and Probability of an Event:

This issue is answered by stating the convergence of $n(A)/n$ to a constant limiting value as an assumption, or an axiom, of the system.

However, to assume that $n(A)/n$ will necessarily converge to some constant value is a complex assumption.

We hope that such a constant limiting frequency exists, it is difficult to believe a priori that this will happen.

Relative Frequency and Probability of an Event:

In fact, it would be better to assume a set of simpler axioms about probability and then attempt to prove that such a constant limiting frequency does in some sense exist.

This approach is the modern axiomatic approach to probability theory.

Relative Frequency and Probability of an Event:

We assume that for each event *A* in the sample space Ω there exists a value *P*(*A*), referred to as the probability of *A*.

We then assume that the probabilities satisfy a certain set of axioms which will be more agreeable with our intuitive notion of probability.

Axiomatic Definition of Probability:

From a purely mathematical viewpoint, we suppose that for each event A of an experiment having a sample space Ω there is a number, denoted by $P(A)$ which satisfies the following three axioms:

Axiom 1: **Every random event A has a probability in the (closed) interval [0, 1], i.e., $0 \leq P(A) \leq 1$**

Axiom 2: **The sure event has probability 1, i.e., $P(\Omega) = 1$**

Axiom 3: **For any sequence of disjoint or mutually exclusive events $A_1, A_2, \ldots, A_n, \ldots$, (that is, events for which $A_i \cup A_j = \phi$ when $i \neq j$),**

$$P(A_1 \cap A_2 \cap \ldots \cap A_n \cap \ldots) = P(A_1) + P(A_2) + \ldots + P(A_n) + \ldots, \; n = 1, 2, \ldots, \infty$$

We call $P(A)$ the probability of the event A.

Axiomatic Definition of Probability:

Axiom 1 states that the probability that the outcome of the experiment is contained in *A* is some number between 0 and 1.

Axiom 2 states that, with probability 1, the outcome will be a member of the sample space Ω.

Axiom 3 states that for any set of mutually exclusive events the probability that at least one of these events occurs is equal to the sum of their respective probabilities.

Axiom 3 is called the theorem of additivity of disjoint events.

Axiomatic Definition of Probability:

It is to be noted that if we interpret *P*(*A*) as the relative frequency of the event *A* when a large number of repetitions of the experiment are performed, then *P*(*A*) would indeed satisfy the above axioms.

For instance,

- **the proportion (or frequency) of time that the outcome is in *A* is clearly between 0 and 1, and**
- **the proportion of time that it is in Ω is 1 (since all outcomes are in Ω).**
- **Also, if *A* and *B* have no outcomes in common, then the proportion of time that the outcome is in either *A* or *B* is the sum of their respective frequencies.**

Axiomatic Definition of Probability:

Example: Suppose a pair of dice is rolled and sum of the points on upper faces is obtained.

Suppose event *A* : sum is 4, 6, or 12 and

event *B* is that the sum is 7 or 9.

Then if outcome *A* occurs 10% time and outcome *B* occurs 20% time, then 30% of the time the outcome will be either 4, 6, 12, 7, or 9.