Conditions

Q.2 (a) How can we differentiate between continuous and discrete random variables. Explain with the help of examples.

Solution

Random variables can be classified as either discrete (i.e. it may assume any of a specified list of exact values) or as continuous (i.e. it may assume any numerical value in an interval or collection of intervals). The mathematical function describing the possible values of a random variable and their associated probabilities is known as a probability distribution. The realizations of a random variable, i.e. the results of randomly choosing values according to the variable's probability distribution are called random variates.

A discrete probability distribution shall be understood as a probability distribution characterized by a probability mass function. Thus, the distribution of a random variable X is discrete, and X is then called a discrete random variable, if

$$\sum_u \Pr(X = u) = 1$$

as u runs through the set of all possible values of X. It follows that such a random variable can assume only a finite or countably infinite number of values.

In cases more frequently considered, this set of possible values is a topologically discrete set in the sense that all its points are isolated points. But there are discrete random variables for which this countable set is dense on the real line (for example, a distribution over rational numbers).

Among the most well-known discrete probability distributions that are used for statistical modeling are the Poisson distribution, the Bernoulli distribution, the binomial distribution, the geometric distribution, and the negative binomial distribution. In addition, the discrete uniform distribution is commonly used in computer programs that make equal-probability random selections between a number of choices.

A continuous probability distribution is a probability distribution that has a probability density function. Mathematicians also call such a distribution absolutely continuous, since its cumulative distribution function is absolutely continuous with respect to the Lebesgue measure λ. If the distribution of X is continuous, then X is called a continuous random variable. There are many examples of continuous probability distributions: normal, uniform, chi-squared, and others.

Intuitively, a continuous random variable is the one which can take a continuous range of values — as opposed to a discrete distribution, where the set of possible values for the random variable is at most countable. While for a discrete distribution an event with probability zero is impossible (e.g. rolling 3½ on a standard die is impossible, and has probability zero), this is not so in the case of a continuous random variable. For example, if one measures the width of an oak leaf, the result of 3½ cm is possible, however it has probability zero because there are uncountably many other potential values even between 3 cm and 4 cm. Each of these individual outcomes has probability zero, yet the probability that the outcome will fall into the interval (3 cm, 4 cm) is nonzero. This apparent paradox is resolved by the fact that the probability that X attains some value within an infinite set, such as an interval, cannot be found by naively adding the probabilities for individual values. Formally, each value has an infinitesimally small probability, which statistically is equivalent to zero.

Example of a discrete R.V.:

Examples of discrete random variables include the number of children in a family, the Friday night attendance at a cinema, the number of patients in a doctor's surgery, the number of defective light bulbs in a box of ten.

Example of a continuous R.V.:

A random variable is called continuous if it can assume all possible values in the possible range of the random variable. Suppose the temperature in a certain city in the month of June in the past many years has always been between 35° to 45° centigrade. The temperature can take any value between the ranges 35° to 45°. The temperature on any day may be $40.15^\circ C$ or $40.16^\circ C$ or it may take any value between $40.15^\circ C$ and $40.16^\circ C$. When we say that the temperature is $40^\circ C$, it means that the temperature lies between somewhere between 39.5° to 40.5°. Any observation which is taken falls in the interval. There is nothing like an exact observation in the continuous variable. In discrete random variable the values of the variable are exact like 0, 1, 2 good bulbs. In continuous random variable the value of the variable is never an exact point. It is always in the form of an interval, the interval may be very small.