Lecture 02

Table of Contents

[Random Variables 3](#_Toc120900545)

[Probability Functions 3](#_Toc120900546)

[Discrete Random Variables 4](#_Toc120900547)

[Bernoulli Random Variables 4](#_Toc120900548)

[Binomial Random Variable 5](#_Toc120900549)

[Geometric Random Variables 6](#_Toc120900550)

[Poisson Random Variables 6](#_Toc120900551)

[Continuous Random Variables 7](#_Toc120900552)

[Uniform Random Variables 7](#_Toc120900553)

[Exponential Random Variables 8](#_Toc120900554)

[Gamma Random Variables 8](#_Toc120900555)

[Normal Distribution 8](#_Toc120900556)

[Expected Values 9](#_Toc120900557)

[Law of Large Numbers 9](#_Toc120900558)

[Variance 10](#_Toc120900559)

## Random Variables

A **Random Variable** takes on a defined set of values with different probabilities. For example, if we roll a die, the outcome is random. There are six possible values for , each of which occur with a probability of .

Random variables can be discrete or continuous. A **discrete random variable** has a countable number of outcomes, e.g., the outcome of a dice roll, while a **continuous random variable** does not, e.g., the weight of a person.

## Probability Functions

A **Probability Function** maps values of against their respective probabilities, . This value will always be between and . The area under a probability function is always equal to .

The probability function which gives the probability that a **discrete** random variable is exactly equal to some specified value is called the **probability mass function** (PMF). For example, the PMF of for a dice roll is given by .

The probability function which gives the probability random variable is equal to of less than some specified value is called the **cumulative distribution function** (CDF). It is denoted by .

For **continuous random variables**, we cannot find the probability at some specified value. Instead, we use integration to calculate the area under the curve for a range of values. This is called a **Probability Density Function** (PDF). For example, suppose . The probability that the value of is between and is given by:

## Discrete Random Variables

We will be studying 4 types of **discrete random variables**, Bernoulli random variables, Binomial random variables, Geometric random variables and Poisson random variables.

### Bernoulli Random Variables

A **Bernoulli Random Variable** has two possible outcomes, either a success or a failure. Thus, . The **PMF** is given by:

The **expected outcome** is given by .

The expected outcome for any discrete random variable is defined as

For example, if we are tossing a coin and success is defined as getting a head, the expected value is . This means that for a large number of trials, on average, the outcome tends to .

### Binomial Random Variable

A **Binomial Random Variable** is the combination of multiple independent Bernoulli trials. Here, is defined as the number of successes in trials. The binomial distribution is given by:

, where

The expected value is given by . This can be derived as follows:

Let .

### Geometric Random Variables

For **Geometric Random Variables**, is the number of failed trials before the first success.

### Poisson Random Variables

**Poisson Random Variables** are used when we have a large number of trials but is small. The random variable taking on one of the values from , , , , is said to be a Poisson random variable with the parameter if for some value of ,

## Continuous Random Variables

### Uniform Random Variables

A random variable is said to be a **uniform random variable** distributed over the interval if its PDF is given by

For any continuous random variable, we can find the probability of some value occurring as

, where is some small value.

The **expected value** for a uniform random variable is given by:

### Exponential Random Variables

A random variable is said to be an **exponential random variable** with the parameter if

### Gamma Random Variables

For , ,

### Normal Distribution

If and , the above distribution is said to be **normal**. For normal distributions, all values of are provided in a table, so we should convert the distribution to a normal one before using it. We can do this by **normalizing** it.

## Expected Values

The **Expected Value** ( or ) for a distribution can be calculated without using any of the above formulae.

For discrete distributions:

For continuous distributions:

## Law of Large Numbers

When dealing with **large numbers**, we can use the **frequency method** instead of the classical method. This will still give us reliable results.

Number of TVs per Household:

|  |  |
| --- | --- |
| No. of  TVs | No. of  Households |
| 0 | 1,218 |
| 1 | 32,379 |
| 2 | 37,961 |
| 3 | 19,387 |
| 4 | 7,714 |
| 5 | 2,842 |
| Total | 101,501 |

From this, we can say that .

The **Law of Large Numbers** states that the **expected value** can be interpreted as the **long-run average**.

## Variance

Sometimes, the expected value can be a misleading metric. Consider that we need to make a choice between two players in a game, both of whom have an average score of 50. The average score is basically the same as the expected value. Looking at just this value does not help us. Instead, we can also look at the **variance**. Suppose the first player consistently scores 50 in every match, while the second player scores 200 25% of the time and 0 75% of the time (which still puts their average at 50). In this case, we can pick the first player, since they have less variance.