



# Machine Learning #10

## ▼ Define the Bayesian interpretation of probability.

***The Bayesian interpretation of probability is a mathematical framework used to update and quantify beliefs about the likelihood of an event occurring based on new evidence or data.*** In this interpretation, probability is seen as a measure of uncertainty, or the degree of belief, rather than the frequency of occurrence of an event.

***The Bayesian interpretation uses Bayes' theorem to calculate the probability of an event given some observed evidence.*** Bayes' theorem states that the probability of an event A given evidence B is proportional to the probability of evidence B given event A, multiplied by the prior probability of event A, divided by the prior probability of evidence B. Mathematically, this can be written as:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

where  $P(A|B)$  is the posterior probability of event A given evidence B,  $P(B|A)$  is the likelihood of evidence B given event A,  $P(A)$  is the prior probability of event A, and  $P(B)$  is the prior probability of evidence B.

The Bayesian interpretation is widely used in machine learning, particularly in the fields of probabilistic modeling, decision theory, and Bayesian networks. It allows for the incorporation of prior knowledge or assumptions into the analysis, and can be useful for making predictions or decisions based on uncertain or incomplete information.

## ▼ Define probability of a union of two events with equation.

***The probability of the union of two events A and B is defined as the probability that either event A or event B or both events occur.***

Mathematically, it can be represented as:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

where  $P(A \cup B)$  denotes the probability of the union of events A and B,  $P(A)$  denotes the probability of event A,  $P(B)$  denotes the probability of event B, and

$P(A \cap B)$  denotes the probability of the intersection of events A and B.

### ▼ What is joint probability? What is its formula?

**Joint probability is the probability of two events occurring simultaneously. It is the probability that event A and event B both happen together.** It is represented by  $P(A \cap B)$  and can be calculated using the formula:

$$P(A \cap B) = P(A) * P(B|A)$$

where  $P(A)$  is the probability of event A occurring and  $P(B|A)$  is the conditional probability of event B given that event A has occurred.

### ▼ What is chain rule of probability?

The chain rule of probability is a formula that **is used to calculate the probability of two or more events occurring together**. The rule is based on the concept that the probability of two events A and B occurring together can be calculated by multiplying the probability of A occurring by the conditional probability of B occurring given that A has occurred. The formula for the chain rule of probability is as follows:

$$P(A \text{ and } B) = P(A) * P(B|A)$$

This can be extended to calculate the joint probability of more than two events as follows:

$$P(A \text{ and } B \text{ and } C) = P(A) * P(B|A) * P(C|A \text{ and } B)$$

In general, for any sequence of events  $E_1, E_2, \dots, E_n$ , the chain rule of probability states that:

$$P(E_1 \text{ and } E_2 \text{ and } \dots \text{ and } E_n) = P(E_1) * P(E_2|E_1) * P(E_3|E_1 \text{ and } E_2) * \dots * P(E_n|E_1 \text{ and } E_2 \text{ and } \dots \text{ and } E_{n-1})$$

The chain rule of probability is a fundamental concept in probability theory and is used extensively in statistical modeling and machine learning.

### ▼ What is conditional probability means? What is the formula of it?

**Conditional probability is the probability of an event occurring given that another event has occurred.** It is denoted by  $P(A|B)$ , which reads as "the probability of A given B." The formula for conditional probability is:

$$P(A|B) = P(A \text{ and } B) / P(B)$$

where  $P(A \text{ and } B)$  is the probability of both A and B occurring, and  $P(B)$  is the probability of B occurring.

## ▼ What are continuous random variables?

*Continuous random variables are those which can take on any value within a specified range.* They are defined over a continuous range of values, *and the probability of any individual value occurring is zero. Instead, the probability is defined over intervals of values. The probability density function (PDF) is used to describe the probabilities of a continuous random variable.*

Examples of continuous random variables include height, weight, temperature, and time.

## ▼ What are Bernoulli distributions? What is the formula of it?

*Bernoulli distribution is a discrete probability distribution that models an experiment with only two possible outcomes: success and failure.* It is named after the Swiss mathematician Jacob Bernoulli. The outcome of the experiment is represented by a random variable  $X$ , where  $X = 1$  denotes a success and  $X = 0$  denotes a failure.

The probability mass function (PMF) of a Bernoulli distribution is given by:

$$P(X = x) = p^x * (1 - p)^{(1-x)}$$

where  $p$  is the probability of success and  $1-p$  is the probability of failure. The mean of a Bernoulli distribution is  $p$ , and the variance is  $p(1-p)$ .

## ▼ What is binomial distribution? What is the formula?

*Binomial distribution is a discrete probability distribution that describes the number of successes in a fixed number of independent Bernoulli trials,* where each trial has only two possible outcomes: success or failure. The formula for binomial distribution is:

$$P(X = k) = (n \text{ choose } k) * p^k * (1 - p)^{(n-k)}$$

where:

- $P(X=k)$  is the probability of  $k$  successes in  $n$  trials
- $(n \text{ choose } k)$  is the binomial coefficient, also known as "n choose k", which represents the number of ways to choose  $k$  items from a set of  $n$  items
- $p$  is the probability of success on each trial
- $(1-p)$  is the probability of failure on each trial
- $k$  is the number of successes to be observed in  $n$  trials.

▼ [Reference Link](#)

### Binomial Distribution Formula

Formula to Calculate Binomial Distribution The binomial distribution formula calculates the probability of getting x successes in the n trials of the

<https://www.wallstreetmojo.com/binomial-distribution-formula/>

### Binomial Distribution Formula

$$P(X) = {}^n C_x p^x (1 - p)^{n-x}$$



## ▼ What is Poisson distribution? What is the formula?

**Poisson distribution is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time or space, given that these events occur with a known average rate and independently of the time since the last event.** It is used to model rare events that occur randomly over time or space.

The formula for Poisson distribution is:

$$P(X = k) = \frac{(e^{-\lambda} * \lambda^k)}{k!}$$

where:

- $P(X = k)$  is the probability of k events occurring in the given time or space interval
- e is the base of the natural logarithm
- $\lambda$  (lambda) is the expected number of events occurring in the given interval
- k is the number of events that actually occur
- k! is the factorial of k (i.e.,  $k! = k * (k-1) * (k-2) * \dots * 2 * 1$ )

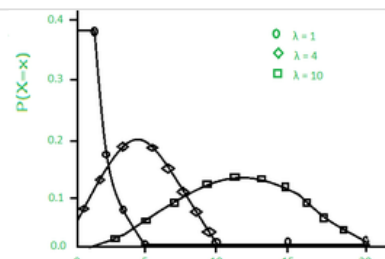
The mean (or expected value) and variance of the Poisson distribution are both equal to  $\lambda$ .

### ▼ [Reference Link](#)

#### Poisson Distribution Formula - GeeksforGeeks

A Computer Science portal for geeks. It contains well written, well thought and well explained computer science and programming articles, quizzes and

<https://www.geeksforgeeks.org/poisson-distribution-formula/>



## ▼ Define covariance.

Covariance is a statistical measure that quantifies the relationship between two variables. *It measures how much two variables change together and provides information about the direction of their relationship.* A positive covariance indicates that the two variables move in the same direction, while a negative covariance indicates that they move in opposite directions. The covariance is calculated as the average of the product of the deviations of each variable from their respective means. It is represented mathematically as:

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$$

### ▼ Define correlation

Correlation is a statistical measure that indicates the degree of association between two variables. *It measures the strength and direction of the linear relationship between two continuous variables.* Correlation coefficient, denoted by  $r$ , is a value between -1 and 1 that indicates the extent to which the two variables are linearly related. A positive value of  $r$  indicates a positive linear relationship, a negative value of  $r$  indicates a negative linear relationship, and a value of 0 indicates no linear relationship.

### ▼ Define sampling with replacement. Give example.

*Sampling with replacement is a sampling method in which each item in a population has an equal probability of being selected for a sample, and after each selection, the item is returned to the population before the next selection. This means that the same item can be selected more than once in the sample.*

For example, suppose we have a bag of 5 balls numbered 1 to 5. If we sample 3 balls with replacement, we select one ball at a time, record its number, and then return it to the bag before selecting the next ball. Thus, on the first selection, we might select ball 4, on the second selection we might select ball 1 again, and on the third selection, we might select ball 5. With replacement, each ball has an equal probability of being selected on each draw, so the probability of drawing any specific ball on each selection is always  $1/5$ .

### ▼ What is sampling without replacement? Give example.

*Sampling without replacement is a method of selecting a subset of items from a larger set without replacing the selected items back into the larger set. This means that once an item is selected, it cannot be selected again for subsequent draws.*

For example, suppose you have a bag with 10 marbles of different colors. If you draw a marble from the bag, record its color, and put it back before drawing the next marble, you are sampling with replacement. However, if you draw a marble from the bag, record its color, and do not put it back before drawing the next marble, you are sampling without replacement. In this case, the probability of drawing a particular color marble will change with each draw.

### ▼ What is hypothesis? Give example.

In statistics, ***a hypothesis refers to an assumption about a population parameter or distribution.*** It is a statement that we want to test using statistical methods and evidence from a sample. A hypothesis typically consists of two parts: the null hypothesis and the alternative hypothesis.

For example, let's say we want to test whether a new drug is effective in reducing blood pressure. Our null hypothesis would be that there is no significant difference in blood pressure between the group that received the drug and the group that received a placebo. The alternative hypothesis would be that there is a significant difference in blood pressure between the two groups. We would then collect data from a sample and use statistical methods to determine whether we can reject the null hypothesis and accept the alternative hypothesis.