1. Define the Bayesian interpretation of probability.

**Bayes’ Theorem**

**Bayes’ Theorem**is used to determine the conditional probability of an event. It was named after an English statistician, **Thomas Bayes**who discovered this formula in 1763. Bayes Theorem is a very important theorem in mathematics, that laid the foundation of a unique statistical inference approach called the **Bayes’ inference. It is used to find the probability of an event, based on prior knowledge of conditions that might be related to that event.**

**For example,** if we want to find the probability that a white marble drawn at random came from the first bag, given that a white marble has already been drawn, and**there are three bags each containing some white and black marbles, then we can use Bayes’ Theorem.**

This article explores the Bayes theorem including its statement, proof, derivation, and formula of the theorem, as well as its applications with various examples.

**What is Bayes’ Theorem?**

**Bayes theorem (also known as the Bayes Rule or Bayes Law) is used to determine the conditional probability of event A when event B has already occurred.**

The general statement of Bayes’ theorem is “**The conditional probability of an event A, given the occurrence of another event B, is equal to the product of the event of B, given A and the probability of A divided by the probability of event B.”** i.e.

***P(A|B) = P(B|A)P(A) / P(B)***

*where,*

* ***P(A)****and****P(B)****are the probabilities of events A and B*
* ***P(A|B)****is the probability of event A when event B happens*
* ***P(B|A)****is the probability of event B when A happens*

**Check: [Bayes’s Theorem for Conditional Probability](https://www.geeksforgeeks.org/bayess-theorem-for-conditional-probability/" \t "_blank)**

**Bayes Theorem Statement**

**Bayes’ Theorem for n set of events is defined as,**

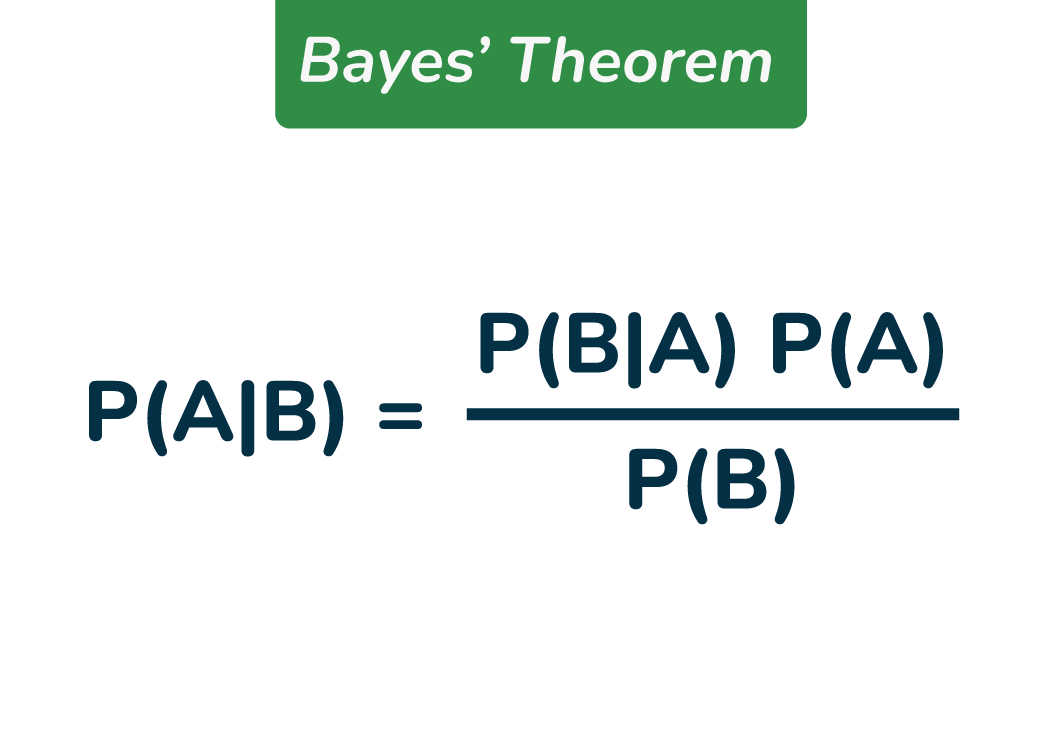
Let E1, E2,…, En be a set of events associated with the sample space S, in which all the events E1, E2,…, En have a non-zero probability of occurrence. All the events E1, E2,…, E form a partition of S. Let A be an event from space S for which we have to find probability, then according to Bayes’ theorem,

***P(Ei|A) = P(Ei)P(A|Ei) / ∑ P(Ek)P(A|Ek)***

***for k = 1, 2, 3, …., n***

**Bayes Theorem Formula**

For any two events A and B, then the formula for the Bayes theorem is given by: (the image given below gives the Bayes’ theorem formula)



*Bayes’ Theorem Formula*

where,

* **P(A)** and **P(B)** are the probabilities of events A and B also P(B) is never equal to zero.
* **P(A|B)** is the probability of event A when event B happens
* **P(B|A)** is the probability of event B when A happens

**Bayes Theorem Derivation**

The proof of Bayes’ Theorem is given as, according to the conditional probability formula,

**P(Ei|A) = P(Ei∩A) / P(A)…..(i)**

Then, by using the multiplication rule of probability, we get

**P(Ei∩A) = P(Ei)P(A|Ei)……(ii)**

Now, by the total probability theorem,

**P(A) =** **∑ P(Ek)P(A|Ek)…..(iii)**

Substituting the value of P(Ei∩A) and P(A) from eq (ii) and eq(iii) in eq(i) we get,

***P(Ei|A) = P(Ei)P(A|Ei) / ∑ P(Ek)P(A|Ek)***

Bayes’ theorem is also known as the formula for the **probability of “causes”**. **As we know, the Ei‘s are a partition of the sample space S, and at any given time only one of the events Ei occurs.** Thus we conclude that the Bayes’ theorem formula gives the probability of a particular Ei, given the event A has occurred.

**Terms Related to Bayes Theorem**

After learning about Bayes theorem in detail, let us understand some important terms related to the concepts we covered in formula and derivation.

* **Hypotheses:**Events happening in the sample space **E1, E2,… En is called the hypotheses**
* **Priori Probability:**Priori Probability is the initial probability of an event occurring before any new data is taken into account. P(Ei) is the priori probability of hypothesis Ei.
* **Posterior Probability:**Posterior Probability is the updated probability of an event after considering new information. Probability P(Ei|A) is considered as the posterior probability of hypothesis Ei.

**Conditional Probability**

* The probability of an event A based on the occurrence of another event B is termed [conditional Probability](https://www.geeksforgeeks.org/conditional-probability/).
* It is denoted as **P(A|B)** and represents the probability of A when event B has already happened.

**Joint Probability**

When the probability of two more events occurring together and at the same time is measured it is marked as Joint Probability. For two events A and B, it is denoted by joint probability is denoted as, **P(A∩B).**

**Random Variables**

Real-valued variables whose possible values are determined by random experiments are called random variables. The probability of finding such variables is the experimental probability.

**Bayes’ Theorem Applications**

Bayesian inference is very important and has found application in various activities, including medicine, science, philosophy, engineering, sports, law, etc., and Bayesian inference is directly derived from Bayes’ theorem.

**Example:** Bayes’ theorem defines the accuracy of the medical test by taking into account how likely a person is to have a disease and what is the overall accuracy of the test.

**Difference Between Conditional Probability and Bayes Theorem**

The difference between Conditional Probability and Bayes Theorem can be understood with the help of the table given below,

| **Bayes’ Theorem** | **Conditional Probability** |
| --- | --- |
| Bayes’ Theorem is derived using the definition of conditional probability. It is used to find the reverse probability. | Conditional Probability is the probability of event A when event B has already occurred. |
| **Formula:** P(A|B) = [P(B|A)P(A)] / P(B) | **Formula:** P(A|B) = P(A∩B) / P(B) |

**Theorem of Total Probability**

Let E1, E2, . . ., En is mutually exclusive and exhaustive events associated with a random experiment and lets E be an event that occurs with some Ei. Then, prove that

***P(E) = n∑i=1P(E/Ei) . P(Ej)***

**Proof:**

*Let S be the sample space. Then,*

*S = E1 ∪ E2 ∪ E3  ∪ . . . ∪ En and Ei ∩ Ej = ∅ for i ≠ j.*

*E = E ∩ S*

*⇒ E = E ∩ (E1 ∪ E2 ∪ E3 ∪ . . . ∪ En)*

*⇒ E = (E ∩ E1) ∪ (E ∩ E2) ∪ . . . ∪ (E ∩ En)*

*P(E) = P{(E ∩ E1) ∪ (E ∩ E2)∪ . . . ∪(E ∩ En)}*

*⇒ P(E) = P(E ∩ E1) + P(E ∩ E2) + . . . + P(E ∩ En)*

*{Therefore, (E ∩ E1), (E ∩ E2), . . . ,(E ∩ En)} are pairwise disjoint}*

*⇒ P(E) = P(E/E1) . P(E1) + P(E/E2) . P(E2) + . . . + P(E/En) . P(En)  [by multiplication theorem]*

*⇒ P(E) = n∑i=1P(E/Ei) . P(Ei)*

**Articles Related to Bayes’ Theorem**

* [*Probability Distribution*](https://www.geeksforgeeks.org/probability-distribution/)
* [*Bayes’ Theorem for Conditional Probability*](https://www.geeksforgeeks.org/bayess-theorem-for-conditional-probability/)
* [*Permutations and Combinations*](https://www.geeksforgeeks.org/bayess-theorem-for-conditional-probability/)
* [*Binomial Theorem*](https://www.geeksforgeeks.org/binomial-theorem/)

**Conclusion – Bayes’ Theorem**

Bayes’ Theorem offers a powerful framework for updating the probability of a hypothesis based on new evidence or information. By incorporating prior knowledge and updating it with observed data, Bayes’ Theorem allows for more accurate and informed decision-making in a wide range of fields, including statistics, machine learning, medicine, and finance. Its applications span from medical diagnosis and risk assessment to spam filtering and natural language processing.

Understanding and applying Bayes’ Theorem enables us to make better predictions, estimate uncertainties, and draw meaningful insights from data, ultimately enhancing our ability to make informed decisions in complex and uncertain situations.

**Also Check:**

* [**Bayes’ Theorem in Data Mining**](https://www.geeksforgeeks.org/bayes-theorem-in-data-mining/)
* [**Bayes Theorem in Artificial Intelligence**](https://www.geeksforgeeks.org/bayes-theorem/)
* [**Bayes Theorem in Machine learning**](https://www.geeksforgeeks.org/bayes-theorem-in-machine-learning/)

**Bayes Theorem Examples**

**Example 1:** **A person has undertaken a job. The probabilities of completion of the job on time with and without rain are 0.44 and 0.95 respectively. If the probability that it will rain is 0.45, then determine the probability that the job will be completed on time.**

**Solution:**

*Let E1 be the event that the mining job will be completed on time and E2 be the event that it rains. We have,*

*P(A) = 0.45,*

*P(no rain) = P(B) = 1 − P(A) = 1 − 0.45 = 0.55*

*By multiplication law of probability,*

*P(E1) = 0.44, and P(E2) = 0.95*

*Since, events A and B form partitions of the sample space S, by total probability theorem, we have*

*P(E) = P(A) P(E1) + P(B) P(E2)*

*⇒ P(E) = 0.45 × 0.44 + 0.55 × 0.95*

*⇒ P(E) = 0.198 + 0.5225 = 0.7205*

*So, the probability that the job will be completed on time is 0.7205*

**Example 2: There are three urns containing 3 white and 2 black balls; 2 white and 3 black balls; 1 black and 4 white balls respectively. There is an equal probability of each urn being chosen. One ball is equal probability chosen at random. what is the probability that a white ball is drawn?**

**Solution:**

*Let E1, E2, and E3 be the events of choosing the first, second, and third urn respectively. Then,*

*P(E1) = P(E2) = P(E3) =1/3*

*Let E be the event that a white ball is drawn. Then,*

*P(E/E1) = 3/5, P(E/E2) = 2/5, P(E/E3) = 4/5*

*By theorem of total probability, we have*

*P(E) = P(E/E1) . P(E1) + P(E/E2) . P(E2) + P(E/E3) . P(E3)*

*⇒ P(E) = (3/5 × 1/3) + (2/5 × 1/3) + (4/5 × 1/3)*

*⇒ P(E) = 9/15 = 3/5*

**Example 3:** **A card from a pack of 52 cards is lost. From the remaining cards of the pack, two cards are drawn and are found to be both hearts. find the probability of the lost card being a heart.**

**Solution:**

*Let E1, E2, E3, and E4 be the events of losing a card of hearts, clubs, spades, and diamonds respectively.*

*Then P(E1) = P(E2) = P(E3) = P(E4) = 13/52 = 1/4.*

*Let E be the event of drawing 2 hearts from the remaining 51 cards. Then,*

*P(E|E1) = probability of drawing 2 hearts, given that a card of hearts is missing*

*⇒ P(E|E1) = 12C2 / 51C2 = (12 × 11)/2! × 2!/(51 × 50) = 22/425*

*P(E|E2) = probability of drawing 2 clubs ,given that a card of clubs is missing*

*⇒ P(E|E2) = 13C2 / 51C2 = (13 × 12)/2! × 2!/(51 × 50) = 26/425*

*P(E|E3) = probability of drawing 2 spades ,given that a card of hearts is missing*

*⇒ P(E|E3) = 13C2 / 51C2 = 26/425*

*P(E|E4) = probability of drawing 2 diamonds ,given that a card of diamonds is missing*

*⇒ P(E|E4) = 13C2 / 51C2 = 26/425*

*Therefore,*

*P(E1|E) = probability of the lost card is being a heart, given the 2 hearts are drawn from the remaining 51 cards*

*⇒ P(E1|E) = P(E1) . P(E|E1)/P(E1) . P(E|E1) + P(E2) . P(E|E2) + P(E3) . P(E|E3) + P(E4) . P(E|E4)*

*⇒ P(E1|E) = (1/4 × 22/425) / {(1/4 × 22/425) + (1/4 × 26/425) + (1/4 × 26/425) + (1/4 × 26/425)}*

*⇒ P(E1|E) = 22/100 = 0.22*

*Hence, The required probability is 0.22.*

**Example 4:** **Suppose 15 men out of 300 men and 25 women out of 1000 are good orators. An orator is chosen at random. Find the probability that a male person is selected. Assume that there are equal numbers of men and women.**

**Solution:**

*Gievn,*

* *Total Men = 300*
* *Total Women = 1000*
* *Good orators among Men = 15*
* *Good orators among Women = 25*

*Total number of good orators = 15 (from men) + 25 (from women) = 40*

*Probability of selecting a male orator:*

*P(Male Orator) = Numbers of male orators / total no of orators = 15/40*

**Example 5: A man is known to speak the lies 1 out of 4 times. He throws a die and reports that it is a six. Find the probability that is actually a six.**

**Solution:**

*In a throw of a die, let*

*E1 = event of getting a six,*

*E2 = event of not getting a six and*

*E = event that the man reports that it is a six.*

*Then, P(E1) = 1/6, and P(E2) = (1 – 1/6) = 5/6*

*P(E|E1) = probability that the man reports that six occurs when six has actually occurred*

*⇒ P(E|E1) = probability that the man speaks the truth*

*⇒ P(E|E1) = 3/4*

*P(E|E2) = probability that the man reports that six occurs when six has not actually occurred*

*⇒ P(E|E2) = probability that the man does not speak the truth*

*⇒ P(E|E2) = (1 – 3/4) = 1/4*

*Probability of getting a six ,given that the man reports it to be six*

*P(E1|E) = P(E|E1) × P(E1)/P(E|E1) × P(E1) + P(E|E2) × P(E2)     [by Bayes’ theorem]*

*⇒ P(E1|E) = (3/4 × 1/6)/{(3/4 × 1/6) + (1/4 × 5/6)}*

*⇒ P(E1|E) = (1/8 × 3) = 3/8*

*Hence the probability required is 3/8.*

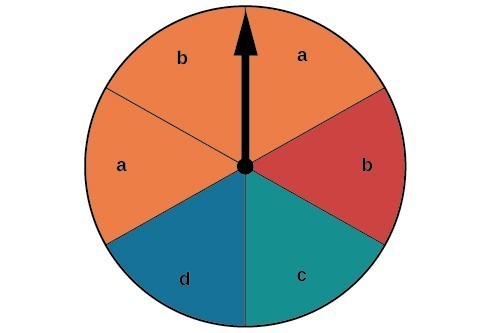
1. Define probability of a union of two events with equation.

**Computing the Probability of the Union of Two Events**

We are often interested in finding the probability that one of multiple events occurs. Suppose we are playing a card game, and we will win if the next card drawn is either a heart or a king. We would be interested in finding the probability of the next card being a heart or a king. The **union of two events** 𝐸 and 𝐹,written 𝐸∪𝐹*E* and *F*,written *E*∪*F*, is the event that occurs if either or both events occur.

𝑃(𝐸∪𝐹)=𝑃(𝐸)+𝑃(𝐹)−𝑃(𝐸∩𝐹)*P*(*E*∪*F*)=*P*(*E*)+*P*(*F*)−*P*(*E*∩*F*)

Suppose the spinner in Figure 2 is spun. We want to find the probability of spinning orange or spinning a 𝑏*b*.



**Figure 2**

There are a total of 6 sections, and 3 of them are orange. So the probability of spinning orange is 36=12​6​​3​​=​2​​1​​. There are a total of 6 sections, and 2 of them have a 𝑏*b*. So the probability of spinning a 𝑏*b* is 26=13​6​​2​​=​3​​1​​. If we added these two probabilities, we would be counting the sector that is both orange and a 𝑏*b* twice. To find the probability of spinning an orange or a 𝑏*b*, we need to subtract the probability that the sector is both orange and has a 𝑏*b*.

12+13−16=23​2​​1​​+​3​​1​​−​6​​1​​=​3​​2​​

The probability of spinning orange or a 𝑏*b* is 23​3​​2​​.

**A GENERAL NOTE: PROBABILITY OF THE UNION OF TWO EVENTS**

The probability of the union of two events 𝐸*E* and 𝐹*F* (written 𝐸∪𝐹*E*∪*F* ) equals the sum of the probability of 𝐸*E* and the probability of 𝐹*F* minus the probability of 𝐸*E* and 𝐹*F* occurring together (( which is called the **intersection** of 𝐸*E* and 𝐹*F* and is written as 𝐸∩𝐹*E*∩*F* ).

𝑃(𝐸∪𝐹)=𝑃(𝐸)+𝑃(𝐹)−𝑃(𝐸∩𝐹)*P*(*E*∪*F*)=*P*(*E*)+*P*(*F*)−*P*(*E*∩*F*)

**EXAMPLE 3: COMPUTING THE PROBABILITY OF THE UNION OF TWO EVENTS**

A card is drawn from a standard deck. Find the probability of drawing a heart or a 7.

**SOLUTION**

A standard deck contains an equal number of hearts, diamonds, clubs, and spades. So the probability of drawing a heart is 14​4​​1​​. There are four 7s in a standard deck, and there are a total of 52 cards. So the probability of drawing a 7 is 113​13​​1​​.

The only card in the deck that is both a heart and a 7 is the 7 of hearts, so the probability of drawing both a heart and a 7 is 152​52​​1​​. Substitute 𝑃(𝐻)=14,𝑃(7)=113,and𝑃(𝐻∩7)=152*P*(*H*)=​4​​1​​,*P*(7)=​13​​1​​,and*P*(*H*∩7)=​52​​1​​ into the formula.

P(E∪ F)=P(E)+P(F)−P(E∩ F) =14+113−152 =413𝑃(𝐸∪ 𝐹)=𝑃(𝐸)+𝑃(𝐹)−𝑃(𝐸∩ 𝐹) =14+113−152 =413

The probability of drawing a heart or a 7 is 413​13​​4​​.

1. What is joint probability? What is its formula?

# Joint Probability | Concept, Formula and Examples

**Last Updated :**19 Sep, 2023

Probability theory is a cornerstone of statistics, offering a powerful tool for navigating uncertainty and randomness in various fields, including business. One key concept within probability theory is Joint Probability, which enables us to analyse the likelihood of multiple events occurring simultaneously.

## What is Joint Probability in Business Statistics?

In the realm of business statistics, Joint Probability refers to the likelihood of two or more events happening together or in conjunction with each other. It helps answer questions such as, “What is the probability of both event A and event B occurring in a business context?”

## What does Joint Probability tell us?

Joint probability offers valuable insights into the likelihood of multiple events happening together. This helps us in several ways:

**1. Co-occurrence:** Joint probability helps us understand how likely it is for two or more events to happen at the same time. This is important for seeing how events are connected and the probability of them occurring together.

**2. Risk Evaluation:** In areas like finance and insurance, joint probability helps us assess the risk when multiple events overlap. **For instance,** it can estimate the chance of multiple financial instruments facing losses simultaneously.

**3. Quality Check:** Businesses can use joint probability to gauge the reliability and quality of their products or processes. It shows the likelihood of multiple defects or issues occurring at once, which allows for proactive quality improvement efforts.

**4. Event Relationships:** Joint probability can indicate if events are related or not. If joint probability significantly differs from the product of individual probabilities, it suggests events are connected, and the occurrence of one affects the likelihood of the other.

**5. Decision Support:** When businesses need to make choices involving multiple factors or events, joint probability provides a numerical foundation for decision-making. It helps assess how different variables together impact the desired outcome.

**6. Resource Management:** In situations with limited resources, understanding joint probability helps optimise resource allocation. **For example,** in supply chain management, it can estimate the chance of multiple supply chain disruptions happening at the same time, enabling better risk management strategies.

## Formula for Joint Probability

The formula for calculating joint probability hinges on whether the events are independent or dependent:

#### 1. For Independent Events

When events A and B are independent, meaning that the occurrence of one event does not impact the other, we use the multiplication rule:

*P(A∩B) = P(A) x P(B)*

Here, P(A) is the probability of occurrence of event A, P(B) is the probability of occurrence of event B, and P(A∩B) is the joint probability of events A and B.

#### 2. For Dependent Events

Events are often dependent on each other, meaning that one event’s occurrence influences the likelihood of the other. Here, we employ a modified formula:

*P(A∩B) = P(A) x P(B|A)*

Here, P(A) is the probability of occurrence of event A, P(B|A) is the conditional probability of occurrence of event B when event A has already occurred, and P(A∩B) is the joint probability of events A and B.

## Examples of Joint Probability

#### Example 1: Independent Events

Suppose you are running an e-commerce platform, and you want to find the probability of a customer purchasing a red shirt (event A) and a blue hat (event B) independently. Find out the Joint Probability where

P(A): The probability of a customer buying a red shirt is 0.3.

P(B): The probability of a customer purchasing a blue hat is 0.2.

#### Solution:

P(A∩B) = P(A) x P(B)

P(A∩B) = P(customer buying a red shirt) x P(customer buying a blue hat)

P(A∩B) = 0.3 x 0.2

**P(A∩B) = 0.6**

#### Example 2: Dependent Events

Imagine you are in the insurance business, and you want to determine the probability of a customer filing a claim (event A) and receiving a payout (event B), given that a claim was filed. Find out the Joint Probability where

P(A): The probability of a customer filing a claim is 0.1.

The probability of a customer receiving a payout given that a claim was filed is 0.8.

#### Solution:

P(A∩B) = P(A) x P(B|A)

P(A∩B) = P(customer filing a claim) x P(customer receiving a payout given that a claim was filed)

P(A∩B) = 0.1 x 0.8

**P(A∩B) = 0.08**

## Difference between Joint Probability and Conditional Probability

#### Joint Probability (P(A∩B))

Joint Probability addresses the simultaneous occurrence of events A and B without considering any specific order or sequence. It quantifies the combined probability of events occurring together, providing insights into their co-occurrence in a business context.

#### Conditional Probability (P(B|A))

Conditional Probability focuses on the probability of event B happening, given that event A has already occurred. This kind of probability is utilised when the occurrence of one event influences the likelihood of another event, making it a valuable tool for understanding cause-and-effect relationships in business statistics.

| **Basis** | **Joint Probability** | **Conditional Probability** |
| --- | --- | --- |
| **Definition** | Probability of multiple events occurring together. | Probability of an event occurring given another event has occurred. |
| **Application** | Provides insights into the combined occurrence of events, often used in risk assessment, quality control, and event co-occurrence analysis. | Useful for understanding cause-and-effect relationships; i.e., helps predict outcomes based on known information. |
| **Focus** | Focuses on events occurring together, regardless of order. | Focuses on events that depend on or are influenced by the occurrence of another event. |
| **Example** | Probability of a customer buying both a red shirt (A) and a blue hat (B) independently. | Probability of a customer buying a blue hat (B) given that he has already bought a red shirt (A). |

In conclusion, Joint probability plays a pivotal role in business statistics, offering a framework to assess the likelihood of multiple events occurring concurrently. By harnessing joint probability, businesses gain valuable insights into the combined outcomes of different events, aiding in decision-making, risk assessment, quality control, and various other applications in the corporate world.

Summer-time is here and so is the time to skill-up! More than 5,000 learners have now completed their journey from **basics of DSA to advanced level development programs** such as Full-Stack, Backend Development, Data Science.

1. What is chain rule of probability?

**Chain rule**  
Generalizing the product rule leads to the **chain rule**. Let 𝐸1,2,....𝐸𝑛 be n events. The joint probability of all the n events is given by,

𝑃(⋂𝑖=1,..,𝑛𝐸𝑖)=𝑃(𝐸𝑛|⋂𝑖=1,..,𝑛−1𝐸𝑖)∗𝑃(⋂𝑖=1,..,𝑛−1𝐸𝑖)

The chain rule can be used iteratively to calculate the joint probability of any no.of events.

1. What is conditional probability means? What is the formula of it?

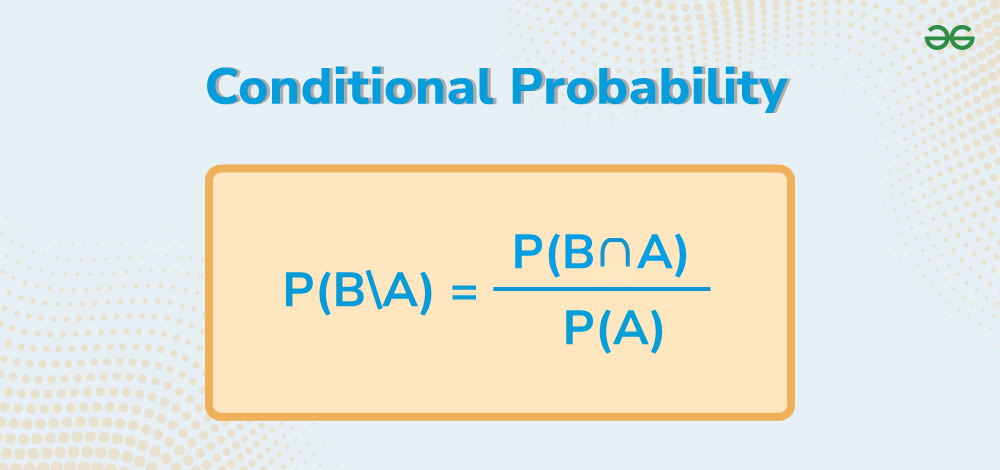
**Conditional Probability**

Last Updated : 18 Jun, 2024

**Conditional probability** is one type of probability in which the possibility of an event depends upon the existence of a previous event. As this type of event is very common in real life, conditional probability is often used to determine the probability of such cases.

Conditional probability is **the likelihood of an outcome occurring based on a previous outcome in similar circumstances. In probability notation, this is denoted as A given B, expressed as P(A|B), indicating that the probability of event A is dependent on the occurrence of event B.**

To know about conditional probability, we need to be familiar with independent events and dependent events. Let’s understand **conditional probability, and its formula with solved examples in this article.**



*Conditional Probability*

**Table of Content**

* [What is Conditional Probability?](https://www.geeksforgeeks.org/conditional-probability/#what-is-conditional-probability)
  + [Conditional Probability Definition](https://www.geeksforgeeks.org/conditional-probability/#conditional-probability-definition)
* [Conditional Probability Formula](https://www.geeksforgeeks.org/conditional-probability/#conditional-probability-formula)
* [How to Calculate Conditional Probability?](https://www.geeksforgeeks.org/conditional-probability/#how-to-calculate-conditional-probability)
  + [Conditional Probability of Independent Events](https://www.geeksforgeeks.org/conditional-probability/#conditional-probability-of-independent-events)
* [Conditional Probability vs Joint Probability vs Marginal Probability](https://www.geeksforgeeks.org/conditional-probability/#conditional-probability-vs-joint-probability-vs-marginal-probability)
* [Conditional Probability and Bayes’ Theorem](https://www.geeksforgeeks.org/conditional-probability/#bayes-theorem)
* [Conditional Probability Examples](https://www.geeksforgeeks.org/conditional-probability/#conditional-probability-examples)
  + [Tossing a Coin](https://www.geeksforgeeks.org/conditional-probability/#tossing-a-coin)
  + [Drawing Cards](https://www.geeksforgeeks.org/conditional-probability/#drawing-cards)
* [Properties of Conditional Probability](https://www.geeksforgeeks.org/conditional-probability/#properties-of-conditional-probability)
* [Multiplication Rule of Probability](https://www.geeksforgeeks.org/conditional-probability/#multiplication-rule-of-probability)
  + [How to Apply the Multiplication Rule?](https://www.geeksforgeeks.org/conditional-probability/#how-to-apply-the-multiplication-rule)
* [Applications of Conditional Probability](https://www.geeksforgeeks.org/conditional-probability/#applications-of-conditional-probability)
* [Conditional Probability Questions](https://www.geeksforgeeks.org/conditional-probability/#conditional-probability-questions)
  + [Resources related to Conditional Probability](https://www.geeksforgeeks.org/conditional-probability/#resources-related-to-conditional-probability)

**What is Conditional Probability?**

**Conditional probability** is the[probability](https://www.geeksforgeeks.org/probability-in-maths/)that depends on a previous result or [event](https://www.geeksforgeeks.org/events-in-probability/). Due to this fact, they help us understand how events are related to each other. Simply put, conditional probability tells us the likelihood of the occurrence of an event based on the occurrence of some previous outcome.

**With** **the help of conditional probability, we can tell apart**[**dependent and independent events**](https://www.geeksforgeeks.org/dependent-and-independent-events-probability/)**.**When the probability of one event happening doesn’t influence the probability of any other event, then events are called independent, otherwise dependent events.

**Conditional Probability Definition**

***Conditional Probability****is defined as the probability of any event occurring when another event has already occurred. In other words, it calculates the probability of one event happening given that a certain condition is satisfied****.****It is represented as P (A | B) which means the probability of A when B has already happened.*

**For Example,**let’s consider the case of rolling two dice, sample space of this event is as follows:

**{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6),**  
**(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6),**  
**(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6),**  
**(4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6),**  
**(5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6),**  
**(6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)}**

Now, consider an event A = getting 3 on the first die and B = getting a sum of 9.

Then the probability of getting 9 when on the first die it’s already 3 is P(B | A),

which can be calculated as follows:

All the cases for the first die as 3 are (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6).

In all of these cases, only one case has a sum of 9.

Thus, P (B | A) = 1/36.

In case, we have to find P (A | B),

All cases where the sum is 9 are (3, 6), (4, 5), (5, 4), and (6, 3).

In all of these cases, only one case has 3 on the first die i.e., (3, 6)

Thus, P(A | B) = 1/36.

**Conditional Probability Formula**

As we can **calculate**the **Conditional Probability** of simple cases without any formula, just as we have seen in the above heading but for complex cases, we need a **Conditional Probability Equation** as we can’t possibly count all the cases for those. Let’s consider two events A and B, then the **formula for conditional probability** of A when B has already occurred is given by:

***P(A|B) = P (A ∩ B) / P(B)***

*Where,*

* ***P (A ∩ B)****represents the probability of both events A and B occurring simultaneously.*
* ***P(B)****represents the probability of event B occurring.*

In other words, the conditional probability of A given B has already occurred is equal to the probability of the intersection of A and B divided by the probability of event B.

**How to Calculate Conditional Probability?**

To **calculate**the conditional probability, we can use the following step-by-step method:

***Step 1:****I****dentify the Events. Let’s call them Event A and Event B.***

***Step 2: Determine the Probability of Event A i.e., P(A)***

***Step 3: Determine the Probability of Event B i.e., P(B)***

***Step 4: Determine the Probability of Event A and B i.e., P(A∩B).***

***Step 5: Apply the Conditional Probability Formula and calculate the required probability.***

**Conditional Probability of Independent Events**

When two events are independent, those conditional probability is the same as the probability of the event individually i.e., P (A | B) is the same as P(A) as there is no effect of event B on the probability of event A. For independent events, A and B, the conditional probability of A and B with respect to each other is given as follows:

* **P(B|A) = P(B)**
* **P(A|B) = P(A)**

**Check,**[**Probability Formulas**](https://www.geeksforgeeks.org/probability-formulas/)

**Conditional Probability vs Joint Probability vs Marginal Probability**

The difference between Conditional Probability, **Joint Probability**, and **Marginal Probability** is given in the following table:

| **Parameter** | **Conditional Probability** | **Joint Probability** | **Marginal Probability** |
| --- | --- | --- | --- |
| **Definition** | The probability of an event occurring given. that another event has already occurred. | The probability of two or more  events occurring simultaneously. | The probability of an event occurring  without considering any other events. |
| **Calculation** | P (A | B) | P (A ∩ B) | P(A) |
| **Variables involved** | Two or more events | Two or more events | Single event. |

**Conditional Probability and Bayes’ Theorem**

**Bayes’ Theorem is a fundamental concept in probability theory named after the Reverend Thomas Bayes**. It provides a mathematical framework for updating beliefs or hypotheses in light of new evidence or information. This theorem is extensively used in various fields, including statistics, machine learning, and artificial intelligence.

At its core, [Bayes’ Theorem](https://www.geeksforgeeks.org/bayes-theorem/" \t "_blank) enables us to calculate the probability of a hypothesis being true given observed evidence. **The theorem is expressed mathematically as follows:**

***P(A∣B)******=******(P(B∣A) × P(A))​ / P(B)***

Where:

* **P(A∣B)** is the posterior probability of hypothesis **A** given evidence **B**.
* **P(B∣A)** is the likelihood of observing evidence **B** given that hypothesis **A** is true.
* **P(A)** is the prior probability of hypothesis **A** before observing any evidence.
* **P(B)** is the probability of observing evidence **B** regardless of the truth of hypothesis **A**.

Here’s a breakdown of how Bayes’ Theorem works:

* ***Prior Probability P(A):****This represents our initial belief in the likelihood of hypothesis****A****being true before considering any new evidence.*
* ***Likelihood P(B∣A):****This indicates the probability of observing the evidence****B****given that hypothesis****A****is true. It quantifies how well the evidence supports the hypothesis.*
* ***Evidence P(B):****This term serves as a normalization factor and represents the total probability of observing the evidence****B****across all possible hypotheses.*
* ***Posterior Probability P(A∣B)****: This is the updated probability of hypothesis****A****being true after taking into account the observed evidence****B****. It’s what we’re ultimately interested in determining.*

**Bayes’ Theorem is particularly powerful because it allows us to incorporate new evidence incrementally, refining our beliefs as more data becomes available**. This iterative process of updating beliefs with new evidence forms the basis of Bayesian inference, which is widely used in fields such as medical diagnosis, spam filtering, weather forecasting, and many others.

Bayes’ Theorem provides a principled approach for reasoning under uncertainty, making it a cornerstone of probabilistic reasoning and decision-making in diverse domains.

**Read in Detail: [Bayes’s Theorem for Conditional Probability](https://www.geeksforgeeks.org/bayess-theorem-for-conditional-probability/" \t "_self)**

**Conditional Probability Examples**

**There are various examples of conditional probability as in real life all the events are related to each other and happening any event affects the probability of another event.**For **example**, if it rains, the probability of road accidents increases as roads have less friction. Let’s consider some problem-based examples here:

**Tossing a Coin**

Let’s consider two events in tossing two coins be,

* **A: Getting a head on the first coin.**
* **B: Getting a head on the second coin.**

Sample space for tossing two coins is:

**S = {HH, HT, TH, TT}**

Conditional probability of getting a head on the second coin (B) given that we got a head on the first coin (A) is = P(B|A)

Since the coins are independent (one coin’s outcome does not affect the other), P(B|A) = P(B) = 0.5 (50%), which is the probability of getting a head on a single coin toss.

**Drawing Cards**

In a deck of 52 cards where two cards are being drawn, then let’s consider the events be.

* **A: Drawing a red card on the first draw, and**
* **B: Drawing a red card on the second draw.**

Conditional probability of drawing a red card on the second draw (B) given that we drew a red card on the first draw (A) is = P(B|A)

After drawing a red card on the first draw, there are 25 red cards and 51 cards remaining in the deck. So, P(B|A) = 25/51 ≈ 0.49 (approximately 49%).

**Properties of Conditional Probability**

**Some of the common properties of conditional property are:**

**Property 1:**Let’s consider an event A in any sample space S of an experiment.

***P(S|A) = P(A|A) = 1***

**Property 2:**For any two events A and B of a sample space S, and an event X such that P(X) ≠ 0,

***P((A ∪ B)|X) = P(A|X) + P(B|X) – P((A ∩ B)|X)***

**Property 3:** The order of set or events is important in conditional probability, i.e.,

***P(A|B) ≠ P(B|A)***

**Property 4:**The complement formula for probability only holds conditional probability if it is given in the context of the first argument in conditional probability i.e.,

***P(A’|B)=1-P(A|B)***

***P(A|B’) ≠ 1-P(A|B)***

**Property 5:**For any two or three independent events, the intersection of events can be calculated using the following formula:

* For the intersection of two events A and B,

***P(A ⋂ B) = P(A) P(B)***

* For the intersection of three events A, B, and C,

***P (A ⋂ B ⋂ C) = P(A) P(B) P(C)***

**Multiplication Rule of Probability**

[**Multiplication Rule of Probability**](https://www.geeksforgeeks.org/multiplication-theorem/), when applied in the context of conditional probability, helps us calculate the probability of the intersection of two events when the probability of one event depends on the occurrence of the other event. **This rule is crucial in understanding the joint probability of events under specific conditions.**

**In the context of conditional probability, the Multiplication Rule is often stated as follows:**

***P(A∩B) = P(A) × P(B∣A)***

**Here’s what each term represents**:

* **P(A∩B)**: **This denotes the probability that both events A and B occur simultaneously.**
* **P(A)**:**This represents the probability of event A happening.**
* **P(B∣A)**: **This is the conditional probability of event B occurring given that event A has already occurred.**

**How to Apply the Multiplication Rule?**

To apply the Multiplication Rule in the context of conditional probability, we can use the following steps:

* First we calculate the probability of event A occurring.
* Then, we compute the probability of event B occurring given that event A has occurred.
* Multiplying these probabilities together gives us the joint probability of both events happening under the specified conditions.
* This rule is particularly useful when dealing with events that are not independent,meaning that the occurrence of one event affects the probability of the other event.

**Applications of Conditional Probability**

Various applications of conditional probability are,

**Finance and Risk Management**

* **Example:** Assessing the probability of default for a borrower given certain financial indicators.
* **Application:** Banks and financial institutions use conditional probability to evaluate the risk associated with loans and investments.

**Healthcare and Diagnostics**

* **Example:**Determining the probability of a patient having a specific disease given the results of diagnostic tests.
* **Application:**Conditional probability is crucial in medical diagnoses and decision-making, helping healthcare professionals make informed decisions based on test results.

**Marketing and Customer Relationship Management (CRM)**

* **Example:** Predicting the probability of a customer making a purchase based on their past buying behavior.
* **Application:**Businesses use conditional probability to tailor marketing strategies, optimize customer experiences, and personalize product recommendations.

**Machine Learning and Artificial Intelligence**

* **Example:**Predicting the likelihood of a user clicking on a particular ad based on their online behavior.
* **Application:** Conditional probability is fundamental in machine learning algorithms for tasks such as classification, recommendation systems, and natural language processing.

**Weather Forecasting**

* **Example:**Estimating the probability of rain tomorrow given today’s weather conditions.
* **Application:** Meteorologists use conditional probability to make weather predictions based on historical data and current atmospheric conditions.

| **Articles Related to Conditional Probability:** | |
| --- | --- |
| [**Probability Theory**](https://www.geeksforgeeks.org/probability-theory/) | [**Addition Rule for Probability**](https://www.geeksforgeeks.org/addition-rule-for-probability/) |
| [**Complement of a Set**](https://www.geeksforgeeks.org/complement-of-a-set/) | [**Types of Events in Probability**](https://www.geeksforgeeks.org/types-of-events-in-probability/#article-meta-div) |
| [**Coin Toss Probability Formula**](https://www.geeksforgeeks.org/coin-toss-probability-formula/?ref=) | [**Mutually Exclusive Events**](https://www.geeksforgeeks.org/mutually-exclusive-events/?ref=) |

1. What are continuous random variables?

**Continuous Random Variable**

Continuous random variable is a random variable that can take on a continuum of values. In other words, a random variable is said to be continuous if it assumes a value that falls between a particular interval.

Continuous random variables are used to denote measurements such as height, weight, time, etc. The area under a density curve is used to represent a continuous random variable. In this article, we will learn about the definition of a continuous random variable, its mean, variance, types, and associated examples.

**What is a Continuous Random Variable?**

A continuous random variable and a discrete [random variable](https://www.cuemath.com/data/random-variable/) are the two types of random variables. A random variable is a variable whose value depends on all the possible outcomes of an experiment. A continuous random variable is defined over a range of values while a discrete random variable is defined at an exact value.

**Continuous Random Variable Definition**

A continuous random variable can be defined as a random variable that can take on an infinite number of possible values. Due to this, the [probability](https://www.cuemath.com/data/probability/) that a continuous random variable will take on an exact value is 0. The cumulative distribution function and the probability density function are used to describe the characteristics of a continuous random variable.

**Continuous Random Variable Example**

Suppose the probability density function of a continuous random variable, X, is given by 4x3, where x ∈ [0, 1]. The probability that X takes on a value between 1/2 and 1 needs to be determined. This can be done by integrating 4x3 between 1/2 and 1. Thus, the required probability is 15/16.

**Continuous Random Variable Formulas**

The probability density function (pdf) and the cumulative distribution function (CDF) are used to describe the probabilities associated with a continuous random variable. The continuous random variable formulas for these functions are given below.

**PDF of Continuous Random Variable**

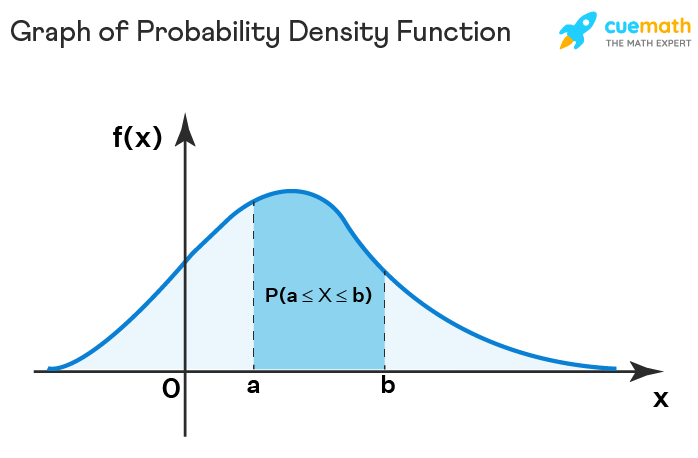
The [probability density function](https://www.cuemath.com/data/probability-density-function/) of a continuous random variable can be defined as a function that gives the probability that the value of the random variable will fall between a range of values. Let X be the continuous random variable, then the formula for the pdf, f(x), is given as follows:

f(x) = dF(x)dxd𝐹(𝑥)d𝑥 = F'(x)

where, F(x) is the cumulative distribution function.

For the pdf of a continuous random variable to be valid, it must satisfy the following conditions:

* ∫∞−∞f(x)dx=1∫−∞∞𝑓(𝑥)𝑑𝑥=1. This means that the total area under the graph of the pdf must be equal to 1.
* f(x) > 0. This implies that the probability density function of a continuous random variable cannot be negative.



**CDF of Continuous Random Variable**

The [cumulative distribution function](https://www.cuemath.com/data/probability-distribution/) of a continuous random variable can be determined by integrating the probability density function. It can be defined as the probability that the random variable, X, will take on a value that is lesser than or equal to a particular value, x. The formula for the cdf of a continuous random variable, evaluated between two points a and b, is given below:

P(a < X ≤ b) = F(b) - F(a) = ∫baf(x)dx∫𝑎𝑏𝑓(𝑥)𝑑𝑥

**Mean and Variance of Continuous Random Variable**

The mean and variance of a continuous random variable can be determined with the help of the probability density function, f(x).

**Mean of Continuous Random Variable**

The [mean](https://www.cuemath.com/data/mean/) of a continuous random variable can be defined as the [weighted average](https://www.cuemath.com/data/weighted-average/) value of the random variable, X. It is also known as the expectation of the continuous random variable. The formula is given as follows:

E[X] = μ=∫∞−∞xf(x)dx𝜇=∫−∞∞𝑥𝑓(𝑥)𝑑𝑥

**Variance of Continuous Random Variable**

The [variance](https://www.cuemath.com/data/variance/) of a continuous random variable can be defined as the expectation of the squared differences from the mean. It helps to determine the dispersion in the distribution of the continuous random variable with respect to the mean. The formula is given as follows:

Var(X) = σ2=∫∞−∞(x−μ)2f(x)dx𝜎2=∫−∞∞(𝑥−𝜇)2𝑓(𝑥)𝑑𝑥

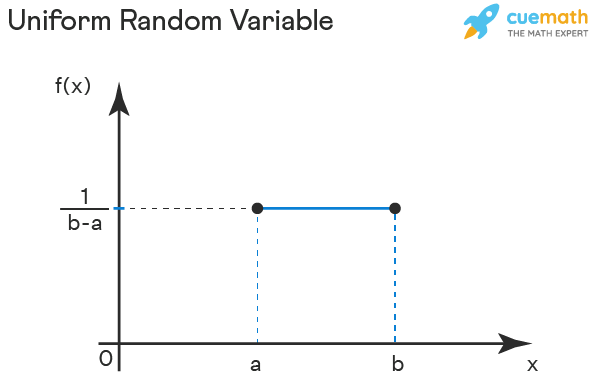
**Continuous Random Variable Types**

A continuous random variable is usually used to model situations that involve measurements. For example, the possible values of the temperature on any given day. As the temperature could be any real number in a given interval thus, a continuous random variable is required to describe it. Some important continuous random variables associated with certain probability distributions are given below.

**Uniform Random Variable**

A continuous random variable that is used to describe a [uniform distribution](https://www.cuemath.com/uniform-distribution-formula/) is known as a uniform random variable. Such a distribution describes events that are equally likely to occur. The pdf of a uniform random variable is as follows:

f(x)={1b−aa≤x≤b0otherwise𝑓(𝑥)={1𝑏−𝑎𝑎≤𝑥≤𝑏0𝑜𝑡ℎ𝑒𝑟𝑤𝑖𝑠𝑒



**Normal Random Variable**

A continuous random variable that is used to model a [normal distribution](https://www.cuemath.com/normal-distribution-formula/) is known as a normal random variable. If the parameters of a normal distribution are given as X∼N(μ,σ2)𝑋∼𝑁(𝜇,𝜎2) then the formula for the pdf is given as follows:

f(x) = 1σ√2Πe−12(x−μσ)21𝜎2Π𝑒−12(𝑥−𝜇𝜎)2

where,

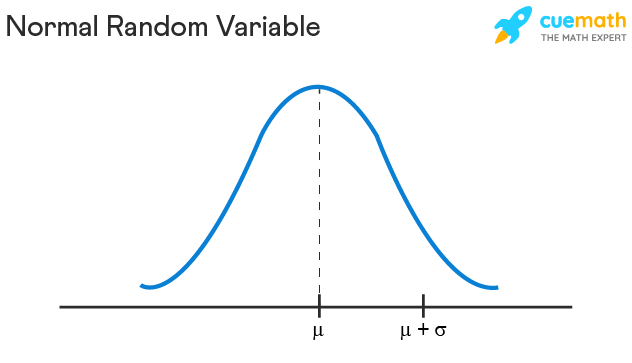
μ𝜇 = mean

σ𝜎 = standard deviation

σ𝜎2 = variance.

A normal distribution where μ𝜇 = 0 and σ𝜎2 = 1 is known as a standard normal distribution. Thus, a standard normal random variable is a continuous random variable that is used to model a standard normal distribution. The pdf formula is as follows:

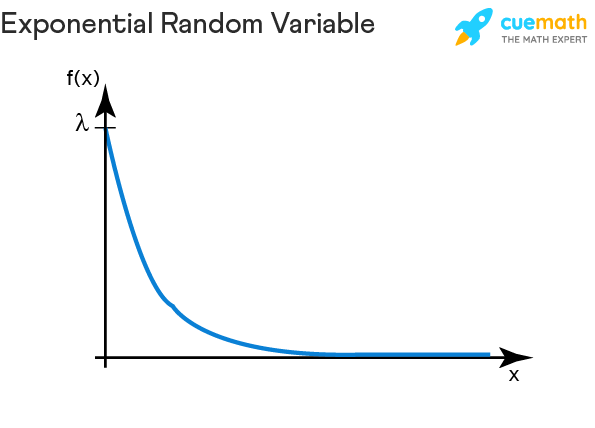
f(x) = 1√2Πe−x2212Π𝑒−𝑥22



**Exponential Random Variable**

[Exponential distributions](https://www.cuemath.com/exponential-distribution-formula/) are continuous probability distributions that model processes where a certain number of events occur continuously at a constant average rate, λ≥0𝜆≥0. Thus, a continuous random variable used to describe such a distribution is called an exponential random variable. The pdf is given as follows:

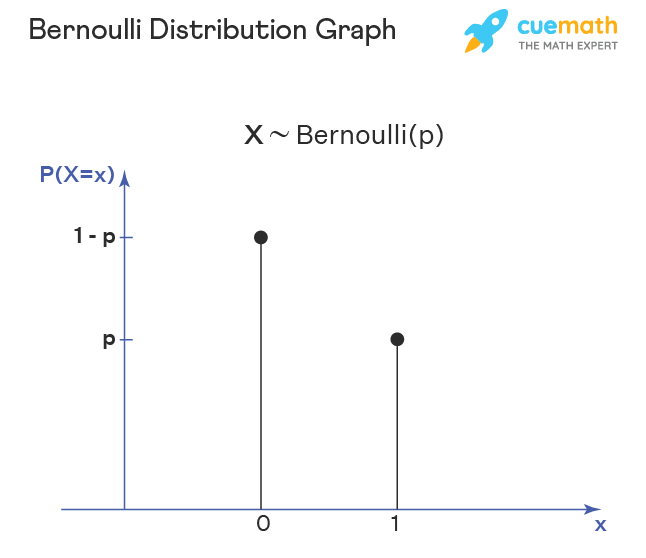
f(x) = λe−λx



**Continuous Random Variable vs Discrete Random Variable**

Both discrete and continuous random variables are used to model a random phenomenon. The differences between a continuous random variable and [discrete random variable](https://www.cuemath.com/algebra/discrete-random-variable/) are given in the table below:

| **Continuous Random Variable** | **Discrete Random Variable** |
| --- | --- |
| The value of a continuous random variable falls between a range of values. | The value of a discrete random variable is an exact value. |
| The probability density function is associated with a continuous random variable. | The probability mass function is used to describe a discrete random variable |
| A continuous random variable can take on an infinite number of values. | Such a variable can take on a finite number of distinct values. |
| Mean of a continuous random variable is E[X] = ∫∞−∞xf(x)dx∫−∞∞𝑥𝑓(𝑥)𝑑𝑥 | The mean of a discrete random variable is E[X] = ∑ x P(X = x), where P(X = x) is the probability mass function. |
| The variance of a continuous random variable is Var(X) = ∫∞−∞(x−μ)2f(x)dx∫−∞∞(𝑥−𝜇)2𝑓(𝑥)𝑑𝑥 | The variance of a discrete random variable is Var[X] = ∑(x − μ)2P(X = x) |
| The examples of a continuous random variable are uniform random variable, exponential random variable, normal random variable, and standard normal random variable. | The examples of a discrete random variable are binomial random variable, geometric random variable, Bernoulli random variable, and Poisson random variable. |

1. What are Bernoulli distributions? What is the formula of it?
2. **Bernoulli Distribution**
3. Bernoulli Distribution is a type of discrete probability distribution where every experiment conducted asks a question that can be answered only in yes or no. In other words, the random variable can be 1 with a probability p or it can be 0 with a probability (1 - p). Such an experiment is called a Bernoulli trial. A pass or fail exam can be modeled by a Bernoulli Distribution.
4. If we have a Binomial Distribution where n = 1 then it becomes a Bernoulli Distribution. As this distribution is very easy to understand, it is used as a basis for deriving more complex distributions. Bernoulli Distribution can be used to describe events that can only have two outcomes, that is, success or failure. In this article, we will learn about the formula, pmf, CDF, and other aspects of the Bernoulli Distribution.
5. **What is Bernoulli Distribution?**
6. Bernoulli Distribution is a special kind of distribution that is used to model real-life examples and can be used in many different types of applications. A random experiment that can only have an outcome of either 1 or 0 is known as a Bernoulli trial. Such an experiment is used in a Bernoulli distribution.
7. **Bernoulli Distribution Definition**
8. A discrete probability distribution wherein the random variable can only have 2 possible outcomes is known as a Bernoulli Distribution. If in a Bernoulli trial the random variable takes on the value of 1, it means that this is a success. The probability of success is given by p. Similarly, if the value of the random variable is 0, it indicates failure. The probability of failure is q or 1 - p. Bernoulli distribution can be used to derive a binomial distribution, [geometric distribution](https://www.cuemath.com/geometric-distribution-formula/), and negative binomial distribution.
9. **Bernoulli Distribution Example**
10. Suppose there is an experiment where you flip a coin that is fair. If the outcome of the flip is heads then you will win. This means that the probability of getting heads is p = 1/2. If X is the random variable following a Bernoulli Distribution, we get P(X = 1) = p = 1/2.
11. **Bernoulli Distribution Formula**
12. A binomial random variable, X, is also known as an indicator variable. This is because if an event results in success then X = 1 and if the outcome is a failure then X = 0. X can be written as X ∼∼ Bernoulli (p), where p is the parameter. The formulas for Bernoulli distribution are given by the probability mass function (pmf) and the cumulative distribution function (CDF).
13. **Probability Mass Function for Bernoulli Distribution**
14. We calculate the probability mass function for a Bernoulli distribution. The probability that a discrete random variable will be exactly equal to some value is given by the probability mass function. The formula for pmf, f, associated with a Bernoulli random variable over possible outcomes 'x' is given as follows:
15. PMF = f(x, p) = {pifx=1q=1−pifx=0{𝑝𝑖𝑓𝑥=1𝑞=1−𝑝𝑖𝑓𝑥=0
16. We can also express this formula as,
17. f(x, p) = px (1 - p)1 - x, x ϵ𝜖 {0, 1}
18. **Cumulative Distribution Function for Bernoulli Distribution**
19. The cumulative distribution function of a Bernoulli random variable X when evaluated at x is defined as the probability that X will take a value lesser than or equal to x. The formula is given as follows:
20. CDF = F(x, p) = ⎧⎪⎨⎪⎩0ifx<01−pif0≤x<11x≥1{0𝑖𝑓𝑥<01−𝑝𝑖𝑓0≤𝑥<11𝑥≥1
21. **Mean and Variance of Bernoulli Distribution**
22. The [arithmetic mean](https://www.cuemath.com/arithmetic-mean-formula/) of a large number of independent realizations of the random variable X gives us the expected value or mean. The expected value can also be thought of as the weighted average. Given below is the proof and formula for the mean of a Bernoulli distribution.
23. **Mean of Bernoulli Distribution Proof:**
24. We know that for X,
25. P(X = 1) = p
26. P(X = 0) = q
27. E[X] = P(X = 1) . 1 + P(X = 0) . 0
28. E[X] = p . 1 + q . 0
29. E[X] = p
30. Thus, the mean or expected value of a Bernoulli distribution is given by E[X] = p.
31. **Variance of Bernoulli Distribution Proof:**
32. The [variance](https://www.cuemath.com/data/variance-and-standard-deviation/) can be defined as the difference of the mean of X2 and the square of the mean of X. Mathematically this statement can be written as follows:
33. Var[X] = E[X2] - (E[X])2
34. Using the properties of E[X2], we get,
35. E[X2] = ∑x2P(X=x)∑𝑥2𝑃(𝑋=𝑥)
36. E[X2] = 12 . p + 02 . q = p
37. Substituting this value in Var[X] = E[X2] - (E[X])2 we have
38. Var[X] = p - p2
39. = p(1 - p)
40. = p . q
41. Hence, the variance of a Bernoulli distribution is Var[X] = p(1 - p) = p . q
42. **Bernoulli Distribution Graph**
43. The graph of a Bernoulli distribution helps to get a visual understanding of the probability density function of the Bernoulli random variable.
44. 
45. The graph shows that the probability of success is p when X = 1 and the probability of failure of X is (1 - p) or q if X = 0.
46. **Bernoulli Distribution and Binomial Distribution**
47. Bernoulli distribution is a special case of the [Binomial distribution](https://www.cuemath.com/algebra/binomial-distribution/) when the number of trials = 1. The difference between Bernoulli distribution and binomial distribution is given below:

|  |  |
| --- | --- |
| **Bernoulli Distribution** | **Binomial Distribtuion** |
| Bernoulli distribution is used when we want to model the outcome of a single trial of an event. | If we want to model the outcome of multiple trials of an event, Binomial distribution is used. |
| It is represented as X ∼∼ Bernoulli (p). Here, p is the [probability](https://www.cuemath.com/data/probability/) of success. | It is denoted as X ∼∼ Binomial (n, p). Where n is the number of trials. |
| Mean, E[X] = p | Mean, E[X] = np |
| Variance, Var[X] = p(1-p) | Variance, Var[X]= np(1-p) |
| Example: Suppose the probability of passing an exam is 80% and failing is 20%. Then the Bernoulli distribution can be used to model the passing or failing in such an exam. | Example: Suppose the probability of passing an exam is 80% and failing is 20%. Then if we want to find the probability that a student will pass in exactly 4 out of 5 exams, we use the Binomial Distribution. |

1. What is binomial distribution? What is the formula?

**Binomial Distribution Formula**

The binomial distribution is a commonly used discrete distribution in statistics. The normal distribution as opposed to a binomial distribution is a continuous distribution. The binomial distribution represents the probability for 'x' successes of an experiment in 'n' trials, given a success probability 'p' for each trial at the experiment.

**Binomial Distribution in Statistics:**The binomial distribution forms the base for the famous binomial test of statistical importance. A test that has a single outcome such as success/failure is also called a Bernoulli trial or Bernoulli experiment, and a series of outcomes is called a Bernoulli process. Consider an experiment where each time a question is asked for a yes/no with a series of n experiments. Then in the binomial probability distribution, the boolean-valued outcome the success/yes/true/one is represented with probability p and the failure/no/false/zero with probability q (q = 1 − p). In a single experiment when n = 1, the binomial distribution is called a Bernoulli distribution.

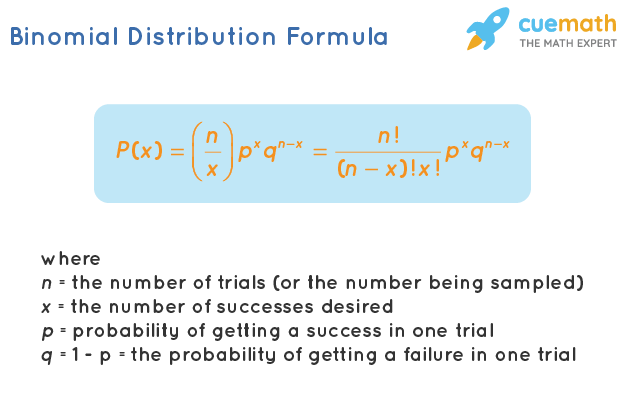
**What Is the Binomial Distribution Formula?**

The binomial distribution formula is for any random variable X, given by;  P(x:n,p) = nCx𝑥 px(1-p)n-x **Or** P(x:n,p) = nCx px (q)n-x

where,

* n = the number of experiments
* x = 0, 1, 2, 3, 4, …
* p = Probability of success in a single experiment
* q = Probability of failure in a single experiment (= 1 – p)

The binomial distribution formula is also written in the form of n-Bernoulli trials, where nCx = n!/x!(n-x)!. Hence, P(x:n,p) = n!/[x!(n-x)!].px.(q)n-x



1. What is Poisson distribution? What is the formula?

Poisson distribution is one example of a discrete probability distribution. Its formula is given by: P(X=k)=(lambda^(k)e^(-lambda))/k! where lambda is the expected value of the random variable X, and k is the number of occurrences.

1. Define covariance.

**Covariance Matrix**

Last Updated : 13 Jun, 2024

Covariance Matrix is a type of matrix used to describe the covariance values between two items in a random vector. It is also known as the variance-covariance matrix because the variance of each element is represented along the matrix’s major diagonal and the covariance is represented among the non-diagonal elements.

It’s particularly important in fields like data science, machine learning, and finance, where understanding relationships between multiple variables is crucial and comes in handy when it comes to stochastic modeling and principal component analysis.

In this article, we will discuss about various things related to Covariance Matrix such as it’s definition, example, and formula as well.

**Table of Content**

* [What is Covariance Matrix?](https://www.geeksforgeeks.org/covariance-matrix/#what-is-covariance-matrix)
* [Covariance Matrix Example](https://www.geeksforgeeks.org/covariance-matrix/#covariance-matrix-example)
* [Covariance Matrix Formula](https://www.geeksforgeeks.org/covariance-matrix/#covariance-matrix-formula)
  + [2 ⨯ 2 Covariance Matrix](https://www.geeksforgeeks.org/covariance-matrix/#2-2-covariance-matrix)
  + [3 ⨯ 3 Covariance Matrix](https://www.geeksforgeeks.org/covariance-matrix/#3-3-covariance-matrix)
* [How to Find Covariance Matrix?](https://www.geeksforgeeks.org/covariance-matrix/#how-to-find-covariance-matrix)
* [Properties of Covariance Matrix](https://www.geeksforgeeks.org/covariance-matrix/#properties-of-covariance-matrix)
* [Solved Examples](https://www.geeksforgeeks.org/covariance-matrix/#solved-examples-on-covariance-matrix)
* [Practice Problems](https://www.geeksforgeeks.org/covariance-matrix/#practice-problems-on-covariance-matrix)
* [FAQs](https://www.geeksforgeeks.org/covariance-matrix/#faqs-on-covariance-matrix)

**What is Covariance Matrix?**

The [variance](https://www.geeksforgeeks.org/variance)-covariance matrix is a [square matrix](https://www.geeksforgeeks.org/square-matrix) with diagonal elements that represent the variance and the non-diagonal components that express covariance. The covariance of a variable can take any real value- positive, negative, or zero. A positive covariance suggests that the two variables have a positive relationship, whereas a negative covariance indicates that they do not. If two elements do not vary together, they have a zero covariance.

**Learn More,** [**Diagonal Matrix**](https://www.geeksforgeeks.org/diagonal-matrix)

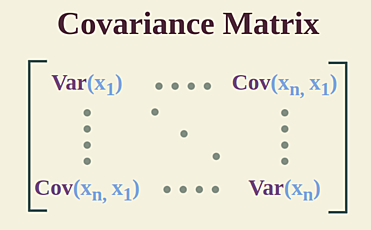
**Covariance Matrix Example**

Let’s say there are 2 data sets X = [10, 5] and Y = [3, 9]. The variance of Set X = 12.5 and the variance of set Y = 18. The covariance between both variables is -15. The covariance matrix is as follows:

[𝑉𝑎𝑟𝑖𝑎𝑛𝑐𝑒 𝑜𝑓 𝑆𝑒𝑡 𝑋𝐶𝑜𝑜𝑟𝑒𝑙𝑎𝑡𝑖𝑜𝑛 𝑜𝑓 𝐵𝑜𝑡ℎ 𝑆𝑒𝑡𝑠𝐶𝑜𝑜𝑟𝑒𝑙𝑎𝑡𝑖𝑜𝑛 𝑜𝑓 𝐵𝑜𝑡ℎ 𝑆𝑒𝑡𝑠𝑉𝑎𝑟𝑖𝑎𝑛𝑐𝑒 𝑜𝑓 𝑆𝑒𝑡 𝑌]=[12.5−15−1518][*Variance* *of* *Set* *XCoorelation* *of* *Both* *Sets*​*Coorelation* *of* *Both* *SetsVariance* *of* *Set* *Y*​]=[12.5−15​−1518​]

**Covariance Matrix Formula**

The general form of a covariance matrix is given as follows:



Where,

* **Sample Variance:**var(x1) = ∑1𝑛(𝑥𝑖−𝑥‾)2𝑛−1*n*−1∑1*n*​(*xi*​−*x*)2​
* **Sample Covarinace:** cov(x1, y1) = ∑1𝑛(𝑥𝑖−𝑥‾)(𝑦𝑖−𝑦‾)𝑛−1*n*−1∑1*n*​(*xi*​−*x*)(*yi*​−*y*​)​
* **Population Variance:** var(xn) = ∑1𝑛(𝑥𝑖−𝜇)2𝑛*n*∑1*n*​(*xi*​−*μ*)2​
* **Population Covariance:** cov(xn, yn) = ∑1𝑛(𝑥𝑖−𝜇𝑥)(𝑦𝑖−𝜇𝑦)𝑛*n*∑1*n*​(*xi*​−*μx*​)(*yi*​−*μy*​)​

Here, **μ** is Mean of Population

𝑥‾*x* is Mean of Sample

**n** is Number of Observation

**xi**is the Observation in Dataset x

Let’s see the format of Covariance Matrix of 2 ⨯ 2 and 3 ⨯ 3.

**2 ⨯ 2 Covariance Matrix**

We know that in a 2 ⨯ 2 [matrix](https://www.geeksforgeeks.org/matrices) there are two rows and two columns. Hence, the 2 ⨯ 2 Covariance Matrix can be expressed as [var(x)cov(x,y)cov(x,y)var(y)][var(x)cov(x,y)​cov(x,y)var(y)​]

**3 ⨯ 3 Covariance Matrix**

In a 3⨯3 Matrix there are 3 rows and 3 columns. We know that in a Covariance Matrix the diagonal elements are variance and non-diagonal elements are covariance. Hence, a 3⨯3 Covariance Matrix can be given as [var(x)cov(x,y)cov(x,z)cov(x,y)var(y)cov(y,z)cov(x,z)cov(y,z)var(z)]​var(x)cov(x,y)cov(x,z)​cov(x,y)var(y)cov(y,z)​cov(x,z)cov(y,z)var(z)​​

**How to Find Covariance Matrix?**

The dimensions of a covariance matrix are determined by the number of variables in a given data set. If there are only two variables in a set, then the covariance matrix would have two rows and two columns. Similarly, if a data set has three variables, then its covariance matrix would have three rows and three columns.

The data pertains to marks scored by Anna, Caroline, and Laura in Psychology and History. Make a covariance matrix.

| **Student** | **Psychology(X)** | **History(Y)** |
| --- | --- | --- |
| **Anna** | **80** | **70** |
| **Caroline** | **63** | **20** |
| **Laura** | **100** | **50** |

The following steps have to be followed:

***Step 1:****Find the mean of variable X. Sum up all the observations in variable X and divide the sum obtained with the number of terms. Thus, (80 + 63 + 100)/3 = 81.*

***Step 2:****Subtract the mean from all observations. (80 – 81), (63 – 81), (100 – 81).*

***Step 3:****Take the squares of the differences obtained above and then add them up. Thus, (80 – 81)2 + (63 – 81)2 + (100 – 81)2.*

***Step 4:****Find the variance of X by dividing the value obtained in Step 3 by 1 less than the total number of observations. var(X) = [(80 – 81)2 + (63 – 81)2 + (100 – 81)2] / (3 – 1) = 343.*

***Step 5:****Similarly, repeat steps 1 to 4 to calculate the variance of Y. Var(Y) = 633.333*

***Step 6:****Choose a pair of variables.*

***Step 7:****Subtract the mean of the first variable (X) from all observations; (80 – 81), (63 – 81), (100 – 81).*

***Step 8:****Repeat the same for variable Y; (70 – 47), (20 – 47), (50 – 47).*

***Step 9:****Multiply the corresponding terms: (80 – 81)(70 – 47), (63 – 81)(20 – 47), (100 – 81)(50 – 47).*

***Step 10:****Find the covariance by adding these values and dividing them by (n – 1). Cov(X, Y) = [(80 – 81)(70 – 47) + (63 – 81)(20 – 47) + (100 – 81)(50 – 47)]/(3-1) = 260.*

***Step 11:****Use the general formula for the covariance matrix to arrange the terms. The matrix becomes: [343260260633.333][343260​260633.333​]*

**Properties of Covariance Matrix**

The Properties of Covariance Matrix are mentioned below:

* A covariance matrix is always square, implying that the number of rows in a covariance matrix is always equal to the number of columns in it.
* A covariance matrix is always symmetric, implying that the [transpose](https://www.geeksforgeeks.org/transpose-of-a-matrix) of a covariance matrix is always equal to the original matrix.
* A covariance matrix is always positive and semi-definite.
* The [eigenvalues](https://www.geeksforgeeks.org/eigen-values) of a covariance matrix are always real and non-negative.

**Read More,**

* [**Types of Matrices**](https://www.geeksforgeeks.org/types-of-matrices)
* [**Matrix Multiplication**](https://www.geeksforgeeks.org/matrix-multiplication)
* [**Variance and Standard Deviation**](https://www.geeksforgeeks.org/variance-and-standard-deviation)

**Solved Examples on Covariance Matrix**

**Example 1: The marks scored by 3 students in Physics and Biology are given below:**

| **Student** | **Physics(X)** | **Biology(Y)** |
| --- | --- | --- |
| **A** | **92** | **80** |
| **B** | **60** | **30** |
| **C** | **100** | **70** |

**Calculate Covariance Matrix from the above data.**

**Solution:**

*Sample covariance matrix is given by ∑1(𝑥𝑖−𝑥‾)2𝑛−1        n−1∑1n​(xi​−x)2​        .*

*Here, μx = 84, n = 3*

*var(x) = [(92 – 84)2 + (60 – 84)2 + (100 – 84)2] / (3 – 1) = 448*

*Also, μy = 60, n = 3*

*var(y) = [(80 – 60)2 + (30 – 60)2 + (70 – 60)2] / (3 – 1) = 700*

*Now, cov(x, y) = cov(y, x) = [(92 – 84)(80 – 60) + (60 – 84)(30 – 60) + (100 – 84)(70 – 60)] / (3 – 1) = 520.*

*The population covariance matrix is given as: [448520520700][448520​520700​]*

**Example 2. Prepare the population covariance matrix from the following table:**

| **Age** | **Number of People** |
| --- | --- |
| **29** | **68** |
| **26** | **60** |
| **30** | **58** |
| **35** | **40** |

**Solution:**

*Population variance is given by ∑1(𝑥𝑖−𝜇)2𝑛        n∑1n​(xi​−μ)2​        .*

*Here, μx = 56.5, n = 4*

*var(x) = [(68 – 56.5)2 + (60 – 56.5)2 + (58 – 56.5)2 + (40 – 56.5)2 ] / 4 = 104.75*

*Also, μy = 30, n = 4*

*var(y) = [(29 – 30)2 + (26 – 30)2 + (30 – 30)2 + (35 – 30)2] / 4 = 10. 5*

*Now, cov(x, y) = ∑14(𝑥𝑖−𝜇𝑥)(𝑦𝑖−𝜇𝑦)44∑14​(xi​−μx​)(yi​−μy​)​*

*cov(x, y) = -27*

***The population covariance matrix is given as:****[104.7−27−2710.5][104.7−27​−2710.5​]*

**Example 3. Interpret the following covariance matrix:**

[𝑋𝑌𝑍𝑋6032−4𝑌32300𝑍−4080]​*XYZ*​*X*6032−4​*Y*32300​*Z*−4080​​

**Solution:**

1. *The diagonal elements 60, 30, and 80 indicate the variance in data sets X, Y, and Z respectively. Y shows the lowest variance whereas Z displays the highest variance.*
2. *The covariance for X and Y is 32. As this is a positive number it means that when X increases (or decreases) Y also increases (or decreases)*
3. *The covariance for X and Z is -4. As it is a negative number it implies that when X increases Z decreases and vice-versa.*
4. *The covariance for Y and Z is 0. This means that there is no predictable relationship between the two data sets.*

**Example 4. Find the sample covariance matrix for the following data:**

| **X** | **Y** | **Z** |
| --- | --- | --- |
| **75** | **10.5** | **45** |
| **65** | **12.8** | **65** |
| **22** | **7.3** | **74** |
| **15** | **2.1** | **76** |
| **18** | **9.2** | **56** |

**Solution:**

*Sample covariance matrix is given by ∑1(𝑥𝑖−𝑥‾)2𝑛−1        n−1∑1n​(xi​−x)2​        .*

*n = 5,*

* *μx = 22.4, var(X) = 321.2 / (5 – 1) = 80.3*
* *μy = 12.58, var(Y) = 132.148 / 4 = 33.037*
* *μz = 64, var(Z) = 570 / 4 = 142.5*

*Now, cov(X, Y) = ∑15(𝑥𝑖−22.4)(𝑦𝑖−12.58)5−1=−11.765−1∑15​(xi​−22.4)(yi​−12.58)​=−11.76*

*⇒ cov(X, Z) = ∑15(𝑥𝑖−22.4)(𝑧𝑖−64)5−1=34.975−1∑15​(xi​−22.4)(zi​−64)​=34.97*

*⇒ cov(Y, Z) = ∑15(𝑦𝑖−12.58)(𝑧𝑖−64)5−1=−40.87 5−1∑15​(yi​−12.58)(zi​−64)​=−40.87*

*The covariance matrix is  given as:*

*[80.3−13.86514.25−13.86533.037−39.525014.25−39.5250142.5]​80.3−13.86514.25​−13.86533.037−39.5250​14.25−39.5250142.5​​*

**Practice Problems on Covariance Matrix**

**Problem 1:**Given two sets of data points: X = [2, 4, 6, 8, 10] and Y = [1, 3, 5, 7, 9], calculate the covariance between X and Y.

**Problem 2:**Calculate the covariance matrix for the following dataset:

| **X1** | **X2** | **X3** |
| --- | --- | --- |
| 4 | 2 | 0 |
| 4 | 5 | 6 |
| 8 | 10 | 12 |
| 12` | 9 | 6 |

**Problem 3:**Given the covariance matrix:

Σ=[4−20−231015]Σ=​4−20​−231​015​​

Identify the variances and covariances between the variables.

**FAQs on Covariance Matrix**

**Define Covariance Matrix**

*A covariance matrix is a type of matrix used to describe the covariance values between two items in a random vector.*

**What is the Formula for Covariance Matrix?**

*The Formula for Covariance Matrix is given as*

*[Var⁡(𝑥1)……Cov⁡(𝑥𝑛,𝑥1)⋮…⋮⋮…⋮Cov⁡(𝑥𝑛,𝑥1)……Var⁡(𝑥𝑛)]​*Var*(x1​)⋮⋮*Cov*(xn​,x1​)​………………​*Cov*(xn​,x1​)⋮⋮*Var*(xn​)​​*

*Where,****Sample Variance:****var(x1) = ∑1(𝑥𝑖−𝑥‾)2𝑛−1n−1∑1n​(xi​−x)2​*

* ***Sample Covarinace:****cov(x1, y1) = ∑1𝑛(𝑥𝑖−𝑥‾)(𝑦𝑖−𝑦‾)𝑛−1n−1∑1n​(xi​−x)(yi​−y​)​*
* ***Population Variance:****var(xn) = ∑1𝑛(𝑥𝑖−𝜇)2𝑛n∑1n​(xi​−μ)2​*
* ***Population Covariance:****cov(xn, yn) = ∑1𝑛(𝑥𝑖−𝜇𝑥)(𝑦𝑖−𝜇𝑦)𝑛n∑1n​(xi​−μx​)(yi​−μy​)​*

**What is the General Form of a 3 ⨯ 3 Covariance Matrix?**

*The general form of a 3 ⨯ 3 covariance matrix is given as follows:*

*[var(x)cov(x,y)cov(x,z)cov(x,y)var(y)cov(y,z)cov(x,z)cov(y,z)var(z)]​*var*(*x*)*cov*(*x*,*y*)*cov*(*x*,*z*)​*cov*(*x*,*y*)*var*(*y*)*cov*(*y*,*z*)​*cov*(*x*,*z*)*cov*(*y*,*z*)*var*(*z*)​​*

**What are the Properties of Covariance Matrix?**

*Covariance Matrix is a square matrix and is also symmetric in nature i.e. the transpose of the original matrix gives the original matrix itself*

**What are the sectors where Covariance Matrix can be used?**

*Covariance Matrix is used in the field of Mathematics, Machine Learning, Finance and Economics. Covariance Matrix is used in [Cholskey Decomposition](https://www.geeksforgeeks.org/cholesky-factorization/" \t "_blank) to perfom Monte Carlo Simulation which is used to create Mathematical Models.*

1. Define correlation

Correlation is a statistical term describing the degree to which two variables move in coordination with one another. If the two variables move in the same direction, then those variables are said to have a positive correlation. If they move in opposite directions, then they have a negative correlation.

1. Define sampling with replacement. Give example.

Sampling with replacement is used to find **probability with replacement**. In other words, you want to find the [probability](https://www.statisticshowto.com/probability-and-statistics/probability-main-index/)of some event where there’s a number of balls, cards or other objects, and you replace the item each time you choose one.

Let’s say you had a population of 7 people, and you wanted to sample 2. Their names are:

* John
* Jack
* Qiu
* Tina
* Hatty
* Jacques
* Des

1. What is sampling without replacement? Give example.

Sampling Without Replacement

Sampling without Replacement is a way to figure out **probability without replacement**. In other words, you don’t replace the first item you choose before you choose a second. This dramatically changes the odds of choosing sample items. Taking the above example, you would have the same list of names to choose two people from. And your list of results would similar, except you couldn’t choose the same person twice:

* John, Jack
* John, Qui
* Jack, Qui
* Jack Tina…

But now, your two items are **dependent**, or linked to each other. When you choose the first item, you have a 1/7 probability of picking a name. But then, assuming you don’t replace the name, you only have six names to pick from. That gives you a 1/6 chance of choosing a second name. The odds become:

* P(John, Jack) = (1/7) \* (1/6) = .024.
* P(John, Qui) = (1/7) \* (1/6) = .024.
* P(Jack, Qui) = (1/7) \* (1/6) = .024.
* P(Jack Tina) = (1/7) \* (1/6) = .024…

As you can probably figure out, I’ve only used a few items here, so the odds only change a little. But larger samples taken from small populations can have more dramatic results.

You can tell *how*dramatic these results are by calculating the [covariance](https://www.statisticshowto.com/probability-and-statistics/statistics-definitions/covariance/). That’s a measure of how much probabilities of two items are linked together; the higher the covariance, the more dramatic the results. A covariance of zero would mean there’s no difference between sampling with replacement or sampling without

1. What is hypothesis? Give example.

A hypothesis is an assumption that is made based on some evidence. This is the initial point of any investigation that translates the research questions into predictions. It includes components like variables, population and the relation between the variables. A research hypothesis is a hypothesis that is used to test the relationship between two or more variables.