# Probability Questions

## Devesh Nath

# 1 Questions

## 1.1 What is the probability of getting at least one head when flipping two coins?

The probability of getting at least one head when flipping two coins can be calculated by finding the complement of the probability of getting no heads (i.e., getting tails on both coins). The probability of getting tails on one coin is $\frac{1}{2}$, so the probability of getting tails on both coins is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. Therefore, the probability of getting at least one head is $1 - \frac{1}{4} = \frac{3}{4}$.

## 1.2 Define and differentiate between discrete and continuous probability distributions.

A discrete probability distribution is one where the set of possible outcomes is discrete, or countable. Examples include the roll of a die or the number of heads in a series of coin flips. Each outcome has a specific probability associated with it.

A continuous probability distribution, on the other hand, is one where the set of possible outcomes is continuous, meaning it can take on any value within a given range. Examples include the height of a person or the time it takes to run a race. Probabilities are described using a probability density function (PDF), and the probability of any single exact outcome is zero; instead, probabilities are calculated over intervals.

## 1.3 How do you calculate the probability of independent events occurring together?

To calculate the probability of independent events occurring together, you multiply the probabilities of each individual event. If $A$ and $B$ are two independent events, then the probability of both $A$ and $B$ occurring is given by:

$$P(A \cap B) = P(A) \times P(B)$$

For example, if the probability of event $A$ is 0.5 and the probability of event $B$ is 0.3, then the probability of both events occurring together is:

$$P(A \cap B) = 0.5 \times 0.3 = 0.15$$

## 1.4 What is a joint probability, and how is it different from conditional probability?

Joint probability refers to the probability of two events occurring together. If $A$ and $B$ are two events, the joint probability is denoted as $P(A \cap B)$ and represents the likelihood of both events happening at the same time.

Conditional probability, on the other hand, is the probability of one event occurring given that another event has already occurred. It is denoted as $P(A|B)$, which represents the probability of event $A$ occurring given that event $B$ has occurred. The relationship between joint probability and conditional probability is given by:

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

In summary, joint probability measures the likelihood of two events happening together, while conditional probability measures the likelihood of one event happening given that another event has already happened.

## 1.5 Explain the law of total probability

The law of total probability is a fundamental rule relating marginal probabilities to conditional probabilities. It states that if $\{B_1, B_2, \ldots, B_n\}$ is a partition of the sample space $S$, then for any event $A$ within $S$, the probability of $A$ can be found by:

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i)$$

This can also be expressed in terms of conditional probabilities:

$$P(A) = \sum_{i=1}^{n} P(A|B_i) \times P(B_i)$$

Here, $B_i$ are mutually exclusive events that cover the entire sample space, and $P(A|B_i)$ is the conditional probability of $A$ given $B_i$. The law of total probability allows us to break down complex probabilities into simpler, more manageable parts. If you have a tree of probabilities, you can sum the probabilities of the branches that lead to the event of interest to find the total probability of that event.

## 1.6 Explain the difference between mutually exclusive and independent events.

Mutually exclusive events are events that cannot occur at the same time. If $A$ and $B$ are mutually exclusive, then the occurrence of $A$ means $B$ cannot occur, and vice versa. Mathematically, this is represented as:

$$P(A \cap B) = 0$$

For example, when rolling a single die, the events of rolling a 3 and rolling a 5 are mutually exclusive because you cannot roll both a 3 and a 5 at the same time.

Independent events, on the other hand, are events where the occurrence of one event does not affect the probability of the other event occurring. If $A$ and $B$ are independent, then:

$$P(A \cap B) = P(A) \times P(B)$$

For example, the outcome of rolling a die and the outcome of flipping a coin are independent events because the result of the die roll does not affect the result of the coin flip.

In summary, mutually exclusive events cannot happen simultaneously, while independent events do not influence each other's occurrence.

## 1.7 What is the difference between a permutation and a combination?

A permutation is an arrangement of objects in a specific order. The order of the objects matters in permutations. The number of permutations of $n$ objects taken $r$ at a time is given by:

$$P(n, r) = \frac{n!}{(n - r)!}$$

where $n!$ (n factorial) is the product of all positive integers up to $n$.

A combination, on the other hand, is a selection of objects without regard to the order. The order of the objects does not matter in combinations. The number of combinations of $n$ objects taken $r$ at a time is given by:

$$C(n, r) = \binom{n}{r} = \frac{n!}{r!(n - r)!}$$

In summary, permutations consider the order of objects, while combinations do not.

## 1.8 What is a Markov chain, and where is it used?

A Markov chain is a stochastic process that undergoes transitions from one state to another on a state space. It is characterized by the Markov property, which states that the future state depends only on the current state and not on the sequence of events that preceded it. This property is also known as "memorylessness."

Mathematically, if $X_n$ represents the state at step $n$, then the Markov property is given by:

$$P(X_{n+1} = x | X_n = x_n, X_{n-1} = x_{n-1}, \ldots, X_0 = x_0) = P(X_{n+1} = x | X_n = x_n)$$

## 1.9 How does probability differ from likelihood?

**Probability** refers to the measure of the chance that a particular event will occur. It is a forward-looking concept, meaning it is used to predict future outcomes based on a known

model. For example, if we know the probability distribution of a random variable, we can calculate the probability of observing a specific outcome.

**Likelihood**, on the other hand, is a measure of how well a particular model explains the observed data. It is a backward-looking concept, meaning it is used to infer the parameters of a model based on observed data. In other words, likelihood is used in the context of parameter estimation. Given a set of observed data, the likelihood function evaluates how likely it is to observe that data under different parameter values of the model.

Mathematically, if $X$ represents the observed data and $\theta$ represents the parameters of the model, the probability of observing $X$ given $\theta$ is denoted as $P(X|\theta)$. The likelihood of the parameters $\theta$ given the observed data $X$ is denoted as $L(\theta|X)$ and is proportional to $P(X|\theta)$. In summary, probability is used to predict future outcomes based on a known model, while likelihood is used to estimate the parameters of a model based on observed data.

## 1.10   What is conditional independence?

Conditional independence refers to a situation where two events are independent given the occurrence of a third event. In other words, two events $A$ and $B$ are conditionally independent given an event $C$ if the occurrence of $A$ and $B$ are independent when $C$ is known to have occurred.

Mathematically, $A$ and $B$ are conditionally independent given $C$ if:

$$P(A \cap B|C) = P(A|C) \times P(B|C)$$

This can also be expressed as:

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

Conditional independence is an important concept in probability theory and statistics, particularly in the context of Bayesian networks and graphical models, where it simplifies the representation and computation of joint probabilities.

Example: Sum of two dice rolls is C, and the individual dice rolls are A and B. A and B are independent, but not conditionally independent given C.

## 1.11   Describe the difference between prior, likelihood, and posterior in Bayesian analysis

In Bayesian analysis, the concepts of prior, likelihood, and posterior are fundamental to updating beliefs based on observed data.

**Prior:** The prior distribution represents the initial beliefs about the parameters before observing any data. It encapsulates any existing knowledge or assumptions about the parameters. The prior is denoted as $P(\theta)$, where $\theta$ represents the parameters of interest.

**Likelihood:** The likelihood function measures how well the observed data is explained by different parameter values. It is the probability of the observed data given the parameters and is denoted as $P(X|\theta)$, where $X$ represents the observed data.

**Posterior:** The posterior distribution represents the updated beliefs about the parameters after observing the data. It combines the prior distribution and the likelihood function using Bayes' theorem. The posterior is denoted as $P(\theta|X)$ and is calculated as:

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)}$$

Here, $P(X)$ is the marginal likelihood or evidence, which ensures that the posterior distribution is a valid probability distribution.

In summary, the prior represents initial beliefs, the likelihood measures the fit of the data to the parameters, and the posterior represents the updated beliefs after considering the data.

## 1.12 Explain the Monty Hall problem and its solution

**Problem Statement:** You are a contestant on a game show. There are three doors: behind one door is a car (the prize), and behind the other two doors are goats. You choose one of the doors, say Door 1. The host, Monty Hall, who knows what is behind each door, then opens another door, say Door 3, which has a goat behind it. Monty then gives you the option to switch your choice to the remaining unopened door (Door 2) or stick with your original choice (Door 1). What should you do to maximize your chances of winning the car?

**Solution:** To maximize your chances of winning the car, you should always switch doors. Here is the reasoning:

- When you first choose a door, there is a 1/3 chance that the car is behind the door you chose and a 2/3 chance that the car is behind one of the other two doors.

- Monty then opens a door with a goat, which does not change the initial probabilities.

- If you stick with your original choice, your probability of winning the car remains 1/3.

- If you switch, your probability of winning the car is 2/3 because Monty's action of opening a door with a goat effectively transfers the 2/3 probability to the remaining unopened door.

Another way to think about this is that in the first round, you are likely to select a goat with (2/3) probability. The host opens another door with a goat so witching gives you a higher probability of winning the car.

Thus, switching doors gives you a higher probability of winning the car.