Probability and Statistics

UG2, Core course, IIIT,H

Pawan Kumar

IIIT, Hyderabad

August 24, 2021

1 Why do we need Probability?

What is probability?

Probability is a measure that we use to measure how likely an event or outcome is. For example, today there is 10% chance of rain.

What is probability?

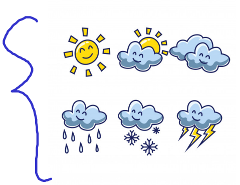
Probability is a measure that we use to measure how likely an event or outcome is. For example, today there is 10% chance of rain.

How likely is that one player or team will win the game?

How likely is that one will win the lottery?

How likely is that one will win the lottery?

How likely it is to rain or snow?



1 Study of Probability started with Games of Chance...

Figure: illustration of great war

- Mahabharat, happened around 900 BCE (Proof dates to 400BCE)

Figure: illustration of great war

- Mahabharat, happened around 900 BCE (Proof dates to 400BCE)
- Struggle between two major group of cousins: Adventure, Drama, Suspense, Thriller, etc

Figure: illustration of great war

- Mahabharat, happened around 900 BCE (Proof dates to 400BCE)
- Struggle between two major group of cousins: Adventure, Drama, Suspense, Thriller, etc
- ... and a game of chance!

1 What Does Game of Chance has to do with Mahabharat in India?

- It is a game of Pasha, similar to modern ludo

- It is a game of Pasha, similar to modern ludo
- In ludo, there is one dice with 6 sides
- There are two cuboidal dices with number of dots on them

1 What Does Game of Chance has to do with Mahabharat in India?

- It is a game of Pasha, similar to modern ludo
- In ludo, there is one dice with 6 sides
- There are two cuboidal dices with number of dots on them
- The move is determined by a randomly throwing the two dices

Figure: Cardino, Italy

Figure: Cardino, Italy

- Early theory of probability arose from games of chance played in Europe

Figure: Cardino, Italy

- Early theory of probability arose from games of chance played in Europe
- On the left is Cardano, an Italian mathematician, who studied game of chance

Figure: Cardino, Italy

- Early theory of probability arose from games of chance played in Europe
- On the left is Cardano, an Italian mathematician, who studied game of chance
- He gambled for about 25 years!

Figure: Cardino, Italy

- Early theory of probability arose from games of chance played in Europe
- On the left is Cardano, an Italian mathematician, who studied game of chance
- He gambled for about 25 years!
- His work on probability were published in famous 15 page "Liber de Ludo Aleae"

1 Game of Chance and Early History of Probability...

- Cardano's contributions have led some to consider Cardano as the real father of probability.

- Cardano's contributions have led some to consider Cardano as the real father of probability.
- Oystein Ore's biography of Cardano, titled Cardano: The Gambling Scholar:

> ... I have gained the conviction that this pioneer work on probability is so extensive and in certain questions so successful that it would seem much more just to date the beginnings of probability theory from Cardano's treatise rather than the customary reckoning from Pascal's discussions with his gambling friend de Méré and the ensuing correspondence with Fermat, all of which took place at least a century after Cardano began composing his *De Ludo Aleae*.

Random Experiment

A random experiment is a process by which we observe something uncertain. After the experiment, the result of the random experiment is known.

Random Experiment

A random experiment is a process by which we observe something uncertain. After the experiment, the result of the random experiment is known.

Outcome

An outcome is a result of a random experiment.

Random Experiment

A random experiment is a process by which we observe something uncertain. After the experiment, the result of the random experiment is known.

Outcome

An outcome is a result of a random experiment.

Sample Space

The set of all possible outcomes is called the sample space.

Random Experiment

A random experiment is a process by which we observe something uncertain. After the experiment, the result of the random experiment is known.

Outcome

An outcome is a result of a random experiment.

Sample Space

The set of all possible outcomes is called the sample space.

Event

An event is a subset of the sample space.

Random Experiment

A random experiment is a process by which we observe something uncertain. After the experiment, the result of the random experiment is known.

Outcome

An outcome is a result of a random experiment.

Sample Space

The set of all possible outcomes is called the sample space.

Event

An event is a subset of the sample space.

Trial

When we repeat a random experiment several times, we call each one of them a trial.

1 Examples of Experiments and Samples Spaces

- Random experiment: toss a coin;

1 Examples of Experiments and Samples Spaces

- Random experiment: toss a coin;
  Sample space: S={heads,tails} or as we usually write it, $\{H, T\}$.

1 Examples of Experiments and Samples Spaces

- Random experiment: toss a coin;
  Sample space: S=$\{$heads,tails$\}$ or as we usually write it, $\{H, T\}$.
- Random experiment: roll a die;

1 Examples of Experiments and Samples Spaces

- Random experiment: toss a coin;
  Sample space: S=$\{$heads,tails$\}$ or as we usually write it, $\{H, T\}$.
- Random experiment: roll a die;
  Sample space: S=$\{1,2,3,4,5,6\}$.

1 Examples of Experiments and Samples Spaces

- Random experiment: toss a coin;
  Sample space: S={heads,tails} or as we usually write it, $\{H, T\}$.
- Random experiment: roll a die;
  Sample space: S={1,2,3,4,5,6}.
- Random experiment: observe the number of iPhones sold by an Apple store in Boston in 2015;

1 Examples of Experiments and Samples Spaces

- Random experiment: toss a coin;
  Sample space: S={heads,tails} or as we usually write it, $\{H, T\}$.
- Random experiment: roll a die;
  Sample space: S={1,2,3,4,5,6}.
- Random experiment: observe the number of iPhones sold by an Apple store in Boston in 2015;
  Sample space: S=$\{0,1,2,3, \cdots\}$.

- Random experiment: toss a coin;
  Sample space: S={heads,tails} or as we usually write it, $\{H, T\}$.
- Random experiment: roll a die;
  Sample space: S={1,2,3,4,5,6}.
- Random experiment: observe the number of iPhones sold by an Apple store in Boston in 2015;
  Sample space: S=$\{0,1,2,3, \cdots\}$.
- Random experiment: observe the number of goals in a basketball match;

1 Examples of Experiments and Samples Spaces

- Random experiment: toss a coin;
  Sample space: S={heads,tails} or as we usually write it, $\{H, T\}$.
- Random experiment: roll a die;
  Sample space: S={1,2,3,4,5,6}.
- Random experiment: observe the number of iPhones sold by an Apple store in Boston in 2015;
  Sample space: S={0,1,2,3, $\cdots$}.
- Random experiment: observe the number of goals in a basketball match;
  Sample space: S={0,1,2,3,$\cdots$ }.


262

LIBER
DE LVDO
ALEÆ.

CAPVT PRIMVM.
De Ludorum generibus.

CAPVT II.
De Ludorum conditionibus.

CAPVT III.
Quibus, & quando magis conveniat ludere.

262

LIBER
DE LVDO
ALEÆ.

CAPVT PRIMVM.

De Ludorum generibus.

CAPVT II.

De Ludorum conditionibus.

CAPVT III.

Quibus, & quando magis conveniat ludere.

- In Chapter 14 of the De Ludo Aleae, Cardano gives what some would consider the first definition of classical (or mathematical) probability:


262

LIBER
DE LVDO
ALEÆ.

CAPVT PRIMVM.
De Ludorum generibus.

CAPVT II.
De Ludorum conditionibus.

CAPVT III.
Quibus, & quando magis conveniat ludere.

- In Chapter 14 of the De Ludo Aleae, Cardano gives what some would consider the first definition of classical (or mathematical) probability:

> So there is one general rule, namely, that we should consider the whole circuit, and the number of those casts which represents in how many ways the favorable result can occur, and compare that number to the rest of the circuit, and according to that proportion should the mutual wagers be laid so that one may contend on equal terms.

Sample Space

$$= \frac{\#\ \text{favorable result occurs}}{\#\ \text{no. of possibilities or outcomes}}$$

2 Definition of Classical Probability...

Figure: Pascal, France

Figure: Pascal, France

- Pascal, French mathematician, 1749-1827

Figure: Pascal, France

- Pascal, French mathematician, 1749-1827
- Modern definition of probability is due to him

Figure: Pascal, France

- Pascal, French mathematician, 1749-1827
- Modern definition of probability is due to him
- Known for his work in
  - Engineering: Tidal Dynamics

2 Definition of Classical Probability...

Figure: Pascal, France

- Pascal, French mathematician, 1749-1827
- Modern definition of probability is due to him
- Known for his work in
  - Engineering: Tidal Dynamics
  - Mathematics: Laplace equation, Laplace Transform

2 Definition of Classical Probability...

Figure: Pascal, France

- Pascal, French mathematician, 1749-1827
- Modern definition of probability is due to him
- Known for his work in
  - Engineering: Tidal Dynamics
  - Mathematics: Laplace equation, Laplace Transform
  - Statistics: Bayesian Intrepretation

Figure: Pascal, France

- Pascal, French mathematician, 1749-1827
- Modern definition of probability is due to him
- Known for his work in
  - Engineering: Tidal Dynamics
  - Mathematics: Laplace equation, Laplace Transform
  - Statistics: Bayesian Intrepretation
  - Physics: Existence of black holes, Gravitational collapse, stability of solar, Speed of sound, Surface tension, etc

equally likely

La théorie des hasards consiste à réduire tous les événemens du même genre, à un certain nombre de cas également possibles, c'est-à-dire, tels que nous soyons également indécis sur leur existence; et à déterminer le nombre des cas favorables à l'événement dont on cherche la probabilité. Le rapport de ce nombre à celui de tous les cas possibles, est la mesure de cette probabilité qui n'est ainsi qu'une fraction dont le numérateur est le nombre des cas favorables, et dont le dénominateur est le nombre de tous les cas possibles.

Figure: Book: "Théorie analytique des probabilités," Laplace, 1812

2 Definition of Classical Probability...the way it evolved

La théorie des hasards consiste à réduire tous les événemens du même genre, à un certain nombre de cas également possibles, c'est-à-dire, tels que nous soyons également indécis sur leur existence; et à déterminer le nombre des cas favorables à l'événement dont on cherche la probabilité. Le rapport de ce nombre à celui de tous les cas possibles, est la mesure de cette probabilité qui n'est ainsi qu'une fraction dont le numérateur est le nombre des cas favorables, et dont le dénominateur est le nombre de tous les cas possibles.

Figure: Book: Théorie analytique des probabilités, Laplace, 1812

Translation of last line: The probability of an event is a fraction whose numerator is the number of favourable cases and whose denominator is total number of cases. In the first line, he also mentions that all the events are equally possible (or equally likely).

Definition

- The number of outcomes is finite.

Definition

- The number of outcomes is finite.
- Each outcome i has probability $p_i \geq 0$

Definition

- The number of outcomes is finite.
- Each outcome i has probability $p_i \geq 0$
- Probabilities sum to 1: $\sum_i p_i = 1$

Definition

- The number of outcomes is finite.
- Each outcome i has probability $p_i \geq 0$
- Probabilities sum to 1: $\sum_i p_i = 1$
- Event: a set of outcomes

Definition

- The number of outcomes is finite.
- Each outcome i has probability $p_i \geq 0$
- Probabilities sum to 1: $\sum_i p_i = 1$
- Event: a set of outcomes
- Probability of the event: sum of outcome probabilities

Definition

- The number of outcomes is finite.
- Each outcome i has probability $p_i \geq 0$
- Probabilities sum to 1: $\sum_i p_i = 1$
- Event: a set of outcomes
- Probability of the event: sum of outcome probabilities
- Is $p_i = 0$ possible?

2 Finite Probability Space with Equally Likely Outcomes

Definition

A finite sample space where each outcome is equally likely

2 Finite Probability Space with Equally Likely Outcomes

Definition

A finite sample space where each outcome is equally likely

- Let $S = \{s_1, s_2, \cdots, s_N\}, \quad P(s_i) = P(s_j), \quad \forall i, j$

Definition

A finite sample space where each outcome is equally likely

- Let $S = \{s_1, s_2, \cdots, s_N\}, \quad P(s_i) = P(s_j), \quad \forall i, j$
- Since all outcomes are equally likely,

$$P(s_i) = \frac{1}{N}, \quad \forall i \in \{1, 2, \cdots, N\}$$

Definition

A finite sample space where each outcome is equally likely

- Let $S = \{s_1, s_2, \cdots, s_N\}, \quad P(s_i) = P(s_j), \quad \forall i, j$
- Since all outcomes are equally likely,

$$P(s_i) = \frac{1}{N}, \quad \forall i \in \{1, 2, \cdots, N\}$$

- If A is any event with cardinality $|A| = M$, then

$$P(A) = \frac{|A|}{|S|} = \frac{M}{N}$$

Probability

We assign a probability measure $P(A)$ or $Pr(A)$ to an event A.

Probability

We assign a probability measure $P(A)$ or $Pr(A)$ to an event A. This is a value between 0 and 1 that shows how likely the event is.

Probability

We assign a probability measure $P(A)$ or $Pr(A)$ to an event A. This is a value between 0 and 1 that shows how likely the event is. If $P(A)$ is close to 0, it is very unlikely that the event A occurs.

Probability

We assign a probability measure P(A) or Pr(A) to an event A. This is a value between 0 and 1 that shows how likely the event is. If P(A) is close to 0, it is very unlikely that the event A occurs. On the other hand, if $P(A)$ is close to 1, A is very likely to occur.

Probability

We assign a probability measure P(A) or Pr(A) to an event A. This is a value between 0 and 1 that shows how likely the event is. If P(A) is close to 0, it is very unlikely that the event A occurs. On the other hand, if $P(A)$ is close to 1, A is very likely to occur. The main subject of probability theory is to develop tools and techniques to calculate probabilities of different events.

Probability

We assign a probability measure P(A) or Pr(A) to an event A. This is a value between 0 and 1 that shows how likely the event is. If P(A) is close to 0, it is very unlikely that the event A occurs. On the other hand, if $P(A)$ is close to 1, A is very likely to occur. The main subject of probability theory is to develop tools and techniques to calculate probabilities of different events.

Probability Axioms...

- Axiom 1: For any event A, $\quad P(A) \geq 0$

Probability

We assign a probability measure P(A) or Pr(A) to an event A. This is a value between 0 and 1 that shows how likely the event is. If P(A) is close to 0, it is very unlikely that the event A occurs. On the other hand, if $P(A)$ is close to 1, A is very likely to occur. The main subject of probability theory is to develop tools and techniques to calculate probabilities of different events.

Probability Axioms...

- Axiom 1: For any event A, $\quad P(A) \geq 0$
- Axiom 2: Probability of the sample space S is $P(S) = 1$

Probability

We assign a probability measure P(A) or Pr(A) to an event A. This is a value between 0 and 1 that shows how likely the event is. If P(A) is close to 0, it is very unlikely that the event A occurs. On the other hand, if $P(A)$ is close to 1, A is very likely to occur. The main subject of probability theory is to develop tools and techniques to calculate probabilities of different events.

Probability Axioms...

- Axiom 1: For any event A, $\quad P(A) \geq 0$
- Axiom 2: Probability of the sample space S is $P(S) = 1$
- Axiom 3: If $A1, A2, A3, \cdots$ are disjoint events, then $P(A_1 \cup A_2 \cup A_3 \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$

Quiz

Prove the following:

- For any event A, $P(A^c) = 1 - P(A)$

$S = A \cup A^c$

Axiom-2: $P(S) = P(A \cup A^c) \rightarrow (*)$

$= 1$

Axiom-3: Since A and A^c are disjoint,

$P(A \cup A^c) = P(A) + P(A^c)$

From $(*)$ $P(A) + P(A^c) = 1$

$\Rightarrow P(A^c) = 1 - P(A)$

Take $A_1 = S$, $A_2 = \phi$

Here $S \cap \phi = \phi \Rightarrow$

S and ϕ are mutually disjoint.

Axiom-3

$\Rightarrow P(S \cup \phi) = P(S) + P(\phi)$

Quiz

Prove the following:

- For any event A, $P(A^c) = 1 - P(A)$
- The probability of the empty set is zero, i.e., $P(\phi) = 0$

Axiom-1: $P(S) = 1$

From above

$1 = P(S) = P(S \cup \phi) = P(S) + P(\phi)$

$\Rightarrow P(\phi) = 1 - P(S) \overset{\text{Axiom-1}}{=} 1 - 1$

$= 0$

$P(A^c) \geq 0$, because A^c is some event.

Quiz

Prove the following:

- For any event A, $P(A^c) = 1 - P(A)$
- The probability of the empty set is zero, i.e., $P(\phi) = 0$
- For any event $A, P(A) \leq 1$

From before,

$P(A^c) = 1 - P(A) \geq 0$

$\Rightarrow P(A) \leq 1$

Vamsi

$A \cup X = B$

$X = B \setminus A$

Here A & X are mutually disjoint.

Quiz

Prove the following:

- For any event A, $P(A^c) = 1 - P(A)$
- The probability of the empty set is zero, i.e., $P(\phi) = 0$
- For any event A, $P(A) \leq 1$ $\leftarrow$
- If $A \subseteq B$, then $P(A) \leq P(B)$

(Axiom-3)

Hence $P(A \cup X) = P(A) + P(X) = P(B)$

Here $P(A) \geq 0$, $P(B) \geq 0$, $P(X) \geq 0$

because A, B, X are events (Axiom 1)

Since $P(X) \geq 0$

$\Rightarrow \boxed{P(A) \leq P(B)}$ //

① Are $A \setminus B$ & $A \cap B$ mutually exclusive ? yes

Quiz

Prove the following:

- For any event A, $P(A^c) = 1 - P(A)$
- The probability of the empty set is zero, i.e., $P(\phi) = 0$
- For any event $A, P(A) \leq 1$
- If $A \subseteq B$, then $P(A) \leq P(B)$
- $P(A - B) = P(A) - P(A \cap B)$

② Also $(A \setminus B) \cup (A \cap B) = A$?

Axiom - 3

Since $A \setminus B$ and $A \cap B$ are mutually exclusive

$\Rightarrow P((A \setminus B) \cup (A \cap B)) =$

$P(A \setminus B) + P(A \cap B) = P(A)$

$\Rightarrow P(A \setminus B) = P(A) - P(A \cap B)$

Coin Toss example

A coin is considered fair if the likelihood of getting heads or tails is same. Calculate the probability of obtaining head, when the coin is tossed once.

$A = \{H\}$ ← favourable event.

$S = \{H, T\}$ ← Sample space

$P(A) = \frac{1}{2} = 50\%$

Coin Toss example

A coin is considered fair if the likelihood of getting heads or tails is same. Calculate the probability of obtaining head, when the coin is tossed once.

Answer

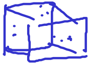
There are two favourable cases, the set of all possible cases is $\{H, T\}$. The number of head possible in one toss of the coin is 1. Hence, the probability of obtaining head is the fraction: $\frac{1}{2}$

Equally likely.

favourable event

Quiz

We roll a fair dice. What is the probability of the event $E = \{1, 6\}$?



$$P(E) = \frac{2}{6} = \frac{1}{3}$$

$$= 33.33\ldots$$

$$\approx 33\%$$

$$E = \{1, 6\}$$

$$S = \{1, 2, 3, 4, 5, 6\}$$

Coin Toss example

A coin is considered fair if the likelihood of getting heads or tails is same.

equilly likely

H H
H T
T H
T T

Coin Toss example

A coin is considered fair if the likelihood of getting heads or tails is same. Calculate the probability of obtaining at least one head, when two fair coins are tossed simultaneously.

Random Experiment.

Sample Space, $S = \{(H,H), (H,T), (T,H), (T,T)\}$

favourable event: $A = \{(H,H), (H,T), (T,H)\}$

$P(A) = \frac{3}{4} = 75\%$

Coin Toss example

A coin is considered fair if the likelihood of getting heads or tails is same. Calculate the probability of obtaining at least one head, when two fair coins are tossed simultaneously.

Answer

- Total number possibilities are $\{H, H\}, \{H, T\}, \{T, T\}, \{T, H\}$

Coin Toss example

A coin is considered fair if the likelihood of getting heads or tails is same. Calculate the probability of obtaining at least one head, when two fair coins are tossed simultaneously.

Answer

- Total number possibilities are $\{H, H\}, \{H, T\}, \{T, T\}, \{T, H\}$
- Favourable cases are $\{H, H\}, \{H, T\}, \{T, H\}$

Coin Toss example

A coin is considered fair if the likelihood of getting heads or tails is same. Calculate the probability of obtaining at least one head, when two fair coins are tossed simultaneously.

Answer

- Total number possibilities are $\{H, H\}, \{H, T\}, \{T, T\}, \{T, H\}$
- Favourable cases are $\{H, H\}, \{H, T\}, \{T, H\}$
- Hence the probability of atleast one head is $= \dfrac{3}{4}$

$P(A^c) = 1 - P(A)$

$\underline{2}\ \underline{2}\ \underline{2}\ \_\ \_\ \_\ \_\ \_\ \_\ \underline{2} = 2^{10}$

Coin Toss example

Ten fair coins are tossed simultaneously. What is the probability of getting atleast one head?

Define: A = event of getting atleast one head.

A^c = event of getting no head, i.e.,

getting $\{(T, T, T, T, T, T, T, T, T, T)\}$

Sample Space: S

& $|S| = 2^{10}$

We know

$P(A^c) = 1 - P(A)$

$\Rightarrow P(A) = 1 - P(A^c)$

$= 1 - 2^{-10}$

$$\boxed{P(A^c) = \frac{1}{2^{10}} = 2^{-10}}$$

Coin Toss example

Ten fair coins are tossed simultaneously. What is the probability of getting atleast one head?

Answer

- Total number possibilities are 2^{10}

Coin Toss example

Ten fair coins are tossed simultaneously. What is the probability of getting atleast one head?

Answer

- Total number possibilities are 2^{10}
- Only case with no head: $\{T, T, T, T, T, T, T, T, T, T\}$. Favourable cases are: $2^{10} - 1$

Coin Toss example

Ten fair coins are tossed simultaneously. What is the probability of getting atleast one head?

Answer

- Total number possibilities are 2^{10}
- Only case with no head: $\{T, T, T, T, T, T, T, T, T, T\}$. Favourable cases are: $2^{10} - 1$
- Hence the probability of atleast one head is $= \dfrac{2^{10} - 1}{2^{10}}$