4. Probability

Tasks you should be able to do after completing this module

- Estimate probabilities of specific events in many scenarios.
- Understand and apply Bayes' theorem to various computer science problems such as designing spam filters.
- Understand and apply random variables and random variable distributions to model a wide variety of problems.
- Apply Bernoulli, binomial, geometric, Poisson, and normal distributions to model and solve computer science problems.

Random Experiment (RE)

An experiment whose outcome is not known in advance, but the set of all possible outcomes is known.

Sample point, denoted ω or s, is an outcome of a RE.

Sample space, denoted $\Omega = \{\text{all possible outcomes of a RE}\}$.

Examples

1. A coin flip is a RE.

Sample points: Head and Tail

Sample space $\Omega = \{\text{Head, Tail}\}$

2. RE: Roll a six-faced die and observe the number of dots on the topmost face. Give its sample points and Sample space, Ω

3. A card is drawn from a standard deck of cards. Is it a RE?

4. Flip a coin until Head occurs. Is it a RE?

Notation: H for Head; T for Tail

Event: A subset of the sample space Ω, satisfying certain axioms. A can be ϕ (a null set), Ω, or a proper subset of Ω.

Examples

5. Roll a six-faced die and count the number of dots on the top face.

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

Let $A =$ event that the die roll results in a prime number. Then

$A = \{2, 3, 5\} \subseteq \Omega$

$B =$ event that an even number occurs $= \{2, 4, 6\}$

$C =$ event that a six occurs $= \{6\}$

$D =$ event that a seven occurs $= \{\,\} = \phi \subseteq \Omega$ An impossible event.

What is $A \cap B$?

A' or $\overline{A}$ is a complementary event of A.

$\therefore \overline{A}$ is an event that prime does not occur
since A is the event that prime occurs.

$\overline{A} = \Omega - A = \{1, 2, 3, 4, 5, 6\} - \{2, 3, 5\} = \{1, 4, 6\}$

A and $\overline{A}$ partition the sample space, Ω:

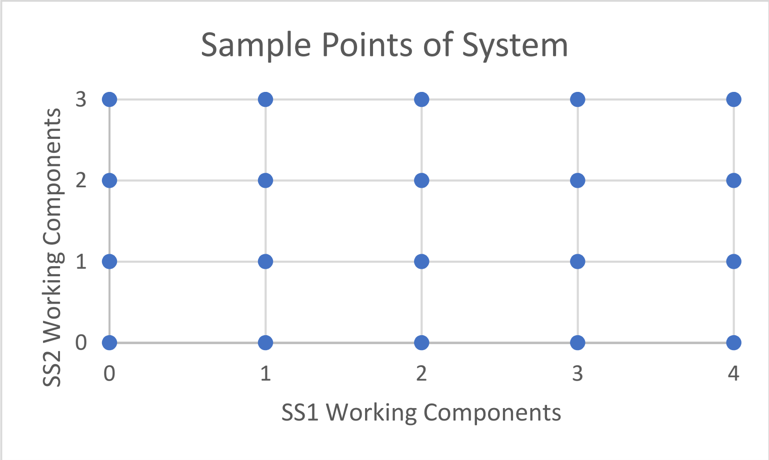
- A and $\overline{A}$ are not empty
- $A \cap \overline{A} = \phi$
- $A \cup \overline{A} = \Omega$

6. A system consists of two subsystems: SS1 with four components and SS2 with three components. We are interested in only the working components of the system. Components fail randomly.

The event space and events on the working condition of the system can be formulated as follows.

Samplel space, $\Omega = \{(x, y) | 0 \le x \le 4, 0 \le y \le 3\}$

where x and y are the number of working components in the system.



Give the following events.

$A =$ event that exactly one component is working

$B =$ exactly three components are working

$C =$ none of the SS2 components are working

Probability Measure

A probability measure assigns a numerical value in [0,1] for the probability or chance that an event occurs.

Probability of 0 ⇒ the event never occurs ⇒ an impossible event

Probability of 1 ⇒ the event always occurs ⇒ a definite event

Other events have probabilities between 0 and 1.

Classical probability

- Ω is finite
- All outcomes are likely

$P(A)$ = Probability that event A occurs

$$P(A) = \frac{|A|}{|\Omega|}.$$

Examples

7. Roll a fair die. $\Omega = \{1, 2, 3, 4, 5, 6\}$. $|\Omega| = 6$.

Each sample point occurs with probability $\frac{1}{6}$.

A = prime occurs = $\{2, 3, 5\}$.

$$\therefore P(A) = \frac{3}{6} = \frac{1}{2}.$$

D = a seven occurs = { } or ϕ. $\therefore$ $P(D) = 0$.

E = a seven does not occur. What is $P(E)$?

[Rosen 7.1.1] An urn contains four blue balls and five red balls. What is the probability that a blue ball is chosen if a ball is taken out of the urn at random (that is, any of the five balls is equally likely to be drawn).

Formal Definition of Probability Measure

Let $\mathbb{F} =$ family of events in Ω.

$\Omega \in \mathbb{F},\ \ \phi \in \mathbb{F}$

Probability measure P is a real-valued function on $\mathbb{F}$.

$$P\colon \mathbb{F} \to [0,1]$$

Axioms of Probability Measure

1. $P(A) \geq 0$ for every event A.

2. $P(\Omega) = 1$.

3. $P(A \cup B) = P(A) + P(B)$ if A and B are mutually exclusive events.

Note:
1. This probability measure includes experiments with outcomes not equally likely and infinite outcomes.

2. We deal with mostly discrete probabilities of REs with finite or countabliy infinite outcomes.

Probability of an Event A is the sum of the probabilities of the outcomes in A.

$$\therefore P(A) = \sum_{\omega \in A} P(\omega)$$

where ω is a sample point with probability $P(\omega)$.

Uniform Distribution

Suppose Ω is a finite set with n elements. The uniform distribution assigns the probability of $1/n$ to each element of Ω.

[This corresponds to the classical probability measure.]

[Rosen 7.1.3′] In a lottery, players win a large prize when they pick four digits (a digit is $0, 1, \ldots, 9$) that match, in the correct order, four digits selected by a random mechanical process. Repetitions allowed. A smaller prize is won if only the first digit is not matched. Compute the probability of winning each prize. Also, compute the probability of winning either the grand prize or the smaller prize.

Experiments with Outcomes not Equally Likely

Random experiments often have outcomes that are not equally likely.

Example [Rosen 7.2.1]

Assign the probabilities to the sample points of a RE of flipping a coin. Consider two cases:
(a) the coin is fair,
(b) the coin is biased so that heads comes up twice as often as a tails.

[Rosen 7.2.2] Suppose a die is biased (or loaded) so that 3 appears twice as often as any other number and all other numbers are equally likely. What is the probability that an odd number occurs when we roll the die?

Theorem

1. $P(\phi) = 0$

2. $P(\overline{A}) = 1 - P(A)$

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

Example

[Rosen 7.1.8] A sequence of 10 bits are randomly generated with the probabilities of generating 0 or 1 being equal. What is the probability that one of the bits is a zero.

Example

What is the probability that a randomly selected positive integer not exceeding 100 is divisible by 2 or 5?

Conditional Probability

Ω is the sample space of a RE.

A, B are events based on the RE, $A, B \subseteq \Omega$.

Let $P(B) > 0$.

$P(A|B) =$ Probability that A occurs given that B occurred.

- $P(A|B)$ revises the probability of A based on the additional information [B occurred.]
- $P(A|B)$ revises the probability of A using B as the new sample space and $A \cap B$ as the new event space.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0$$

Examples

8. There are two urns. The first urn has one blue marble and four red marbles. The second urn has two blue marbles and three red marbles.

a. If an urn is picked at random and a marble is drawn at random, what is the probability that a blue marble is drawn?

b. If the first urn is slelected and a marble is drawn, what is the probability that it is a blue marble?

9. An urn contains two blue and three red marbles. Two marbles are drawn at random one after another without replacement (that is, the first marble is not put back in the urn). What is the probability that a blue marble is drawn in the second attempt if the first marble drawn is also a blue marble?

10. A random experiment consists of rolling a six-faced fair die. Let A be the event prime occurs. Let B be the event an even number occurs.

In one instance of the RE, it is observed that B occurred. What is the probability that A also occurred?

Multiplication rule

$$P(A \cap B) = \begin{cases} P(A|B) \cdot P(B), & P(B) > 0 \\ P(B|A) \cdot P(A), & P(A) > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} P(A \cap B \cap C) &= P(A|B \cap C) \cdot P(B \cap C) \\ &= P(A|B \cap C) \cdot P(B|C) \cdot P(C) \end{aligned}$$

In general:

If $A_1, A_2, \ldots, A_n$ are events and $A_1 \cap A_2 \cap \cdots \cap A_{n-1} \neq \phi$, then

$$\begin{aligned} &P(A_1 \cap \cdots \cap A_n) \\ &\quad = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap \cdots A_{n-1}) \end{aligned}$$

Example

There are 100 computers, 75 of which are of brand X. If three computers are selected at random without replacement, what is the probability tht each of the three selected computers is of brand X.

Independent Events

Two events are independent if and only if

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

In other words,

$P(A|B) = P(A)$ provided B is not a null event, and

$P(B|A) = P(B)$ provided A is not a null event.

If A and B are independent events,
so are $\overline{A}$ and B, A and $\overline{B}$, and $\overline{A}$ and $\overline{B}$.

In general, nonnull events $A_1, \dots, A_n$ are independent if

$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \cdot P(A_2) \cdots P(A_n)$

Example

A fair coin is flipped repeatedly until a head occurs. Each coin flip is independent of other flips. If the sequence of coin flips are of interest, then it is a RE.

$$\Omega = \{H, TH, TTH, TTTH, \dots\}$$

Sample points: H = head in the first flip
TH = tail in the first flip and head in the second
TTH = tail in the first two flips and head in the third

…

Caclculate the probabilities the sample points.

For a fair coin, $P(H) = P(T) = \frac{1}{2}$.

$P(TH) = P(T) \cdot P(H)$

$P(TTH) = P(T) \cdot P(T) \cdot P(H)$

$P(TTTH) =$

Is $P(\Omega) = 1$?

Law of Total Probability

Let $B_1, B_2, \ldots, B_n$ be n events that partition Ω.

1. $B_i \neq \phi, \ 1 \leq i \leq n$
2. $B_i \cap B_j = \phi, \ 1 \leq i \leq n$
3. $B_1 \cup B_2 \cup \cdots \cup B_n = \Omega$

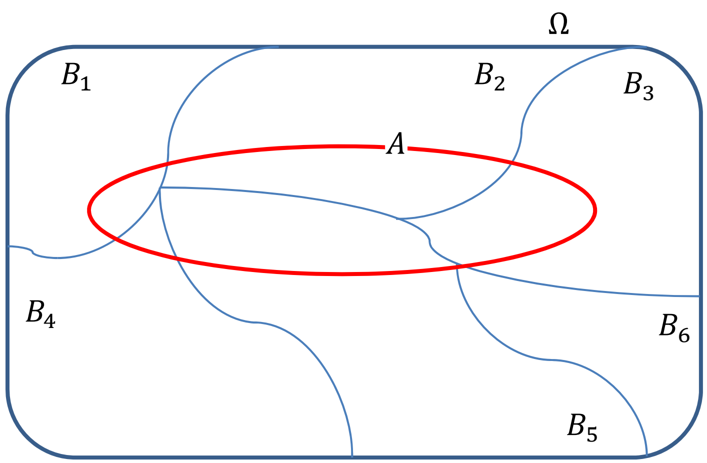
Note: $P(B_1) + P(B_2) + \cdots + P(B_n) = 1$.

Let A be another event.

$P(A|B_1), P(A|B_2), \ldots, P(A|B_n)$ are known

and $P(B_1), \ldots, P(B_n)$ are known.

Calculate $P(A)$ using the known probabilities.



Examples

9. An urn contains two blue and three red marbles. Two marbles are drawn at random one after another without replacement. What is the probability that a blue marble is drawn in the second attempt?

10. A supercomputer center receives jobs from three different sources.

Source	% of Jobs	% of Jobs that require graphics processing
1	15%	1%
2	35%	5%
3	2%	2%

If a job is picked at random, what is the probability that it requires graphics processing?

11. In the previous example on supercomputer center:

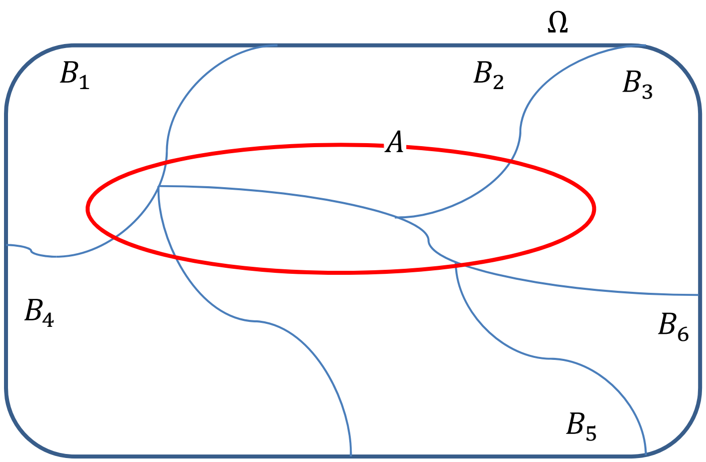
If a job picked at random requires graphics processing, what is the probability that it is from source 2?

Bayes' Theorem

Let $B_1, B_2, \dots, B_n$ be n events that partition Ω.

1. $B_i \neq \phi, \ \ 1 \leq i \leq n$
2. $B_i \cap B_j = \phi, \ \ 1 \leq i \leq n$
3. $B_1 \cup B_2 \cup \cdots \cup B_n = \Omega$

Let A be another nonnull event.
$P(A|B_1), P(A|B_2), \dots, P(A|B_n)$ and $P(B_1), \dots, P(B_n)$ are known.



By Law of Total Probability,

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_n)P(B_n)$$

Bayes' Theorem allows us to calculate new probabilities based on the occurrence of event A.

$$\boldsymbol{P(B_i|A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(B_i) \cdot P(A|B_i)}{P(A)}}$$

Examples

12. About one in thousand messages seen on a computer network contains malware. An anti-virus scanner with 99% accuracy is used to scan and flag messages suspected of containing malware: that is, the scanner flags 99% of the messages containg malware and does not flag 99% of the messages free of malware. If a message is flagged by the scanner, what is the probability that it contains malware?

