Discrete Structures (Monsoon 2021)

Ashok Kumar Das

Associate Professor
IEEE Senior Member
Center for Security, Theory and Algorithmic Research
International Institute of Information Technology, Hyderabad (IIIT Hyderabad)

E-mail: *ashok.das@iiit.ac.in*
URL: `http://www.iiit.ac.in/people/faculty/ashokkdas`
`https://sites.google.com/view/iitkgpakdas/`

Topic: **Probability Theory**

Probability Theory

Definition

If the sample space S of an experiment E consists of finitely many outcomes (points) that are equally likely, the probability of an event A connected with the experiment E is

$$\begin{aligned} P(A) \text{ or } Pr(A) &= \frac{\text{no. of points in } A}{\text{no. of points in } S} \\ &= \frac{m(A)}{n}, \; say \end{aligned}$$

Probability Theory

Axioms of Probability

1. For any event A in S, $P(A) \geq 0$.
2. The probability of a certain event S is $P(S) = 1$.
3. For mutually exclusive events A and B, that is $A \cap B = \emptyset$:
   $P(A \cup B) = P(A) + P(B)$.
   If S is infinite (has infinitely many points), then for mutually exclusive events $A_1, A_2, \ldots$:

   $$P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \ldots$$

   $[A_i \cap A_j = \emptyset, i \neq j; i, j = 1, 2, 3, \ldots]$

Probability Theory

- **Problem:** If A and B be two events, then show that $P(\bar{A} \cup \bar{B}) = 1 - P(A \cap B)$, where $\bar{X}$ denotes the complement of an event X.
- **Problem (Boole's Inequality):** For any n events $A_1, A_2, \ldots, A_n$, $P(A_1 \cup A_2 \cup \cdots \cup A_n) \leq P(A_1) + P(A_2) + \cdots + P(A_n)$.
- **Problem:** Find the probability of occurrence of only one of the events A and B.

Conditional Probability

Definition

The conditional probability of an event B on the hypothesis that another event A has occurred will be denoted by $P(B|A)$ and defined by

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

provided that $P(A) \neq 0$.

In a similar way,

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

provided that $P(B) \neq 0$.

Conditional Probability

Theorem

For any two events A and B, $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$, if $P(A) \neq 0$ and $P(B) \neq 0$.

Proof.

We have, $P(B|A) = \frac{P(A \cap B)}{P(A)} \Rightarrow P(A \cap B) = P(A)P(B|A)$

Again, $P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(B)P(A|B)$.

Thus, $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$.

$\square$

Probability Theory

Conditional Probability

Problem: For n events $A_1, A_2, \ldots, A_n$,

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

Conditional Probability

Problem: A die is rolled. Let A be the event that the result is an even face and B the event that the result if multiple of 3. Then, compute $P(B|A)$ and $P(A|B)$.

Solution: Here, n = points in the sample space $S = 6$ since $S = \{1, 2, 3, 4, 5, 6\}$.

$A = \{2, 4, 6\}$, $m(A) = 3$ and $P(A) = \frac{m(A)}{n} = 36$

$B = \{3, 6\}$, $m(B) = 2$ and $P(B) = \frac{m(B)}{n} = 26$

$A \cap B = \{6\}$, $m(A \cap B) = 1$ and $P(A \cap B) = \frac{m(A \cap B)}{n} = 16$.

Hence, $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{1}{3}$.

$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1}{2}$.

Probability Theory

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Conditional Probability

Theorem (Baye's Theorem)

If $A_1, A_2, \ldots, A_n$ be a given set of n pairwise mutually exclusive events, one of which certainly occurs, and if X be an arbitrary event such that $P(X) \neq 0$, then

$$P(A_i|X) = \frac{P(A_i)P(X|A_i)}{\sum_{i=1}^{n} P(A_i)P(X|A_i)}$$

$(i = 1, 2, \ldots, n)$.

Stochastic Independence

Let A and B be events connected with a random experiment E.
The event B is said to be *independent* of A or *stochastically independent* of A, if the probability of B is noway depended on the occurrence of A, that is, if

$$P(B|A) = P(B).$$

We have,

$$\begin{aligned} P(B|A) &= \frac{P(A \cap B)}{P(A)}, P(A) \neq 0 \\ P(A \cap B) &= P(A)P(B|A) \\ &= P(A).P(B). \end{aligned}$$

In other words, two events A and B are said to be independent if

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

Probability Theory

Random Variable

Definition

Let S be an event space of a random experiment E and R be the set of real numbers. A mapping $X : S \to R$ is called a random variable or a stochastic variable or simply a variate.

- The range of the mapping X is called the spectrum of X.
- If this spectrum is a discrete set, X is called a *discrete* random variable.
- If this spectrum is a continuous set, X is called a *continuous* random variable.

Distribution Function

Definition

Let X be a random variable defined on S, the event space of a random experiment E, and $x \in R$ be a real number. The distribution function of X is denoted by $F_x(x)$ or $F(x)$ defined in $R = (-\infty, \infty)$ by

$$F(x) = P(X \leq x) = P(-\infty < X \leq x).$$

Note that $F(x)$ is a function of the real variable x.

Probability Theory

Properties of Distribution Function $F(x)$

1. $F(x)$ is a monotonic non-decreasing function of x. That is, if $b > a$, $F(b) \geq F(a)$.
2. $F(-\infty) = 0$.
3. $F(\infty) = 1$.
4. $F(a) - F(a-0) = P(X = a)$, where $F(a-0)$ denotes the left-handed limit of $F(x)$ at $x = a$.
5. $F(a+0) = F(a)$, where $F(a+0)$ denotes the right-handed limit of $F(x)$ at $x = a$.

Application of Distribution Function $F(x)$

If the distribution function $F(x)$ of a random variable X is given, we can find the probability that X lies in any arbitrary interval $(a, b]$.
In other words, $P(a < X \leq b) = F(b) - F(a)$.

Probability mass function (p.m.f)

- If X is a discrete random variable, its distribution function can be calculated from its probability mass function (p.m.f) f_i defined for all the reals x_i by
$$f_i = P(X = x_i).$$
- If X is a discrete random variable, then

$$\begin{aligned} F(x) &= P(-\infty < X \leq x) = P[\sum_{a=-\infty}^{i} (X = x_a)] \\ &= \sum_{a=-\infty}^{i} P(X = x_a) = \sum_{a=-\infty}^{i} f_a \end{aligned}$$

Again, $F(\infty) = 1 \Rightarrow \sum_{i=-\infty}^{\infty} f_i = 1.$

Probability density function (p.d.f)

- If X is a continuous random variable, its distribution function can be calculated from its probability density function (p.d.f) $f(\cdot)$ which is characterized by the following properties:
  - (i) $f(x) \geq 0$ for all real x
  - (ii) $\int_{-\infty}^{\infty} f(x)dx = 1$
  - $P(a \leq X \leq b) = \int_{a}^{b} f(x)dx$, for all reals a, b with $a < b$

  where $f(x) = F'(x) = \frac{d}{dx}F(x)$.
- Then, $dF(x) = f(x)dx \Rightarrow F(x) = \int_{-\infty}^{x} f(t)dt = P(-\infty < X \leq x)$.

Probability Differential

- We have,

$$\begin{aligned} P(x < X \leq x + dx) &= F(x + dx) - F(x) \\ &= dF(x) \\ &= \frac{d}{dx} F(x).dx \\ &= f(x)dx. \end{aligned}$$

- $P(x < X \leq x + dx) = f(x)dx.$
- $dF(x)$ is called the probability differential.