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

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

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 $A_1, A_2, \ldots, A_n$,

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

Probability Theory

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 is 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 $P(A) = \frac{m(A)}{n} = \frac{3}{6}$

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

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

Hence, $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{1}{6} / \frac{3}{6} = \frac{1}{3}$.

$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1}{6} / \frac{2}{6} = \frac{1}{2}$.

Probability Theory

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

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

Random Variable

Definition

Let S be an event space of a random experiment E and R be the set of real numbers. A mapping $X : S \rightarrow R$ 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.

Problem (Probability density function (p.d.f))

Show that a function $f(x)$ given by

$$f(x) = \begin{cases} x, & \textit{if } 0 < x < 1 \\ k - x, & \textit{if } 1 < x < 2 \\ 0, & \textit{elsewhere} \end{cases}$$

is a probability density function for a suitable value of the constant k.

Calculate the probability that the random variable lies between $\frac{1}{2}$ and $\frac{3}{2}$.

Problem (Probability distribution function):

If $F(x)$ denotes the distribution function of a random variable X, then show that

1. $P(a < X \leq b) = F(b-0) - F(a)$
2. $P(a \leq X \leq b) = F(b) - F(a-0)$

Important Discrete Distributions

Bernoulli Random Variables

- A Bernoulli trial is random experiment in which there are only two possible outcomes, usually called “success” or “‘failure”, with respective probabilities p and q, where $p + q = 1$, that is, $q = 1 - p$, and $0 < p < 1$.
- A sequence of such trials is a Bernoulli sequence if the trials are independent and the probability of success (s) or failure (f) is constant from trial to trail, where $P(s) = p$ and $P(f) = 1 - p = q$.
- Bernoulli random variable describes a Bernoulli trial and thus assumes only two values: 1 (for success) with probability p and 0 (for failure) with probability $q = 1 - p$.

Important Discrete Distributions

Example (Bernoulli Random Variables)

Let a sample space S have an event A such that $0 < P(A) < 1$, by identifying the occurrence of A with success and $\bar{A}$ (complement of A) with failure.

The corresponding Bernoulli random variable X is defined to be 1 for every point of A and 0 for every point of $\bar{A}$, that is, its probability mass function (p.m.f) is defined by

$$\begin{aligned} f_i &= P(X = x_i) = p^k q^{1-k}, k \in \{0, 1\} \\ f_i &= \begin{cases} p, & \textit{if } k = 1 \\ 1 - p = q, & \textit{if } k = 0 \\ 0, & \textit{otherwise.} \end{cases} \end{aligned}$$

where $P(X = 1) = p$ and $P(X = 0) = q = 1 - p$.

Binomial Random Variables

Consider a Bernoulli sequence of n trials where the probability of success on each trial is p.

The random variable X, that counts the number of successes in the n trials is called a binomial random variable with parameters n and p.

The random variable X (discrete random variable) whose spectrum consists of the points $0, 1, 2, 3, \ldots, n$, that is, $x_i = i$, $i = 0, 1, 2, 3, \ldots, n$ and the pmf is defined by

$$
\begin{aligned}
f_i &= P(X = x_i) = P(X = i) \\
&= {}^nC_i p^i q^{n-i}, \\
& \quad with\, q = 1 - p, 0 < p < 1.
\end{aligned}
$$

Problem (Binomial Random Variables)

The interactive computer system at GNU Glue has 20 communication lines to the central computer system. The lines operate independently and the probability that any particular line is in use is 0.6.

What is the probability that 10 or more lines are in use?

Important Discrete Distributions

Geometric Random Variables

Suppose a sequence of Bernoulli trials is continued until the first success occurs.

Let X be the random variable that counts the number of trials before the trial at which the first success occurs.

The pmf of X is given by

$$\begin{aligned} f_i &= P(X = x_i) = P(X = i) \\ &= q^i p, \\ & \quad \textit{with}\, q = 1 - p, 0 < p < 1, \\ & \quad p = P[\textit{success}], q = P[\textit{failure}]. \end{aligned}$$

Poisson Random Variables

Let E be a random experiment whose event space S consists of two points, one of which is called "success" and the other is "failure".

E is repeated under identical conditions and independently.

If the probability of success is not the same in every trial, the sequence of trials is called a Poisson sequence of trials.

We say that a random variable X is a Poisson random variable with parameter $\alpha > 0$, if X has the mass points $0, 1, 2, 3, \ldots$, and if its pmf is given by

$$\begin{aligned} f_i &= P(X = x_i) = P(X = i) \\ &= e^{-\alpha} \frac{\alpha^i}{i!} \\ &\quad \textit{for } i = 0, 1, 2, 3, \ldots \end{aligned}$$

Important Discrete Distributions

Problem (Poisson Random Variables)

Suppose it has determined that the number of inquiries that arrive per second at the central computer installation of the Varoom Broom online computer system can be described by a Poisson random variable with an average rate of 10 messages per second.

What is the probability that 15 or fewer inquiries arrive in a one-second period?

Important Continuous Distributions

Uniform Random Variables

A continuous random variable X is said to be a uniform or rectangular distribution on the interval a to b, if its pdf is given by

$$f(x) = \begin{cases} \frac{1}{b-a}, & \textit{if } a < x < b \\ 0, & \textit{elsewhere} \end{cases}$$

The corresponding distribution function is easily calculated by integration to give

$$F(x) = \begin{cases} 0, & \textit{if } x \leq a \\ \frac{x-a}{b-a}, & \textit{if } a < x < b \\ 1, & \textit{if } x \geq b \end{cases}$$

Exponential Random Variables

A continuous random variable X has an exponential distribution with parameter $\theta > 0$, if its pdf f is given by

$$f(x) = \begin{cases} \theta e^{-\theta x}, & \text{if } x > 0 \\ 0, & \text{if } x \leq 0 \end{cases}$$

The corresponding distribution function F is given by

$$F(x) = \begin{cases} 1 - e^{-\theta x}, & \text{if } x > 0 \\ 0, & \text{if } x \leq 0 \end{cases}$$

Important Continuous Distributions

Normal Random Variables

A continuous random variable X is said to be a normal (Gaussian) random variable with parameter μ and $\sigma > 0$, if it has the pdf (f)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

The corresponding distribution function F is given by

$$\begin{aligned}
F(x) &= P(-\infty < X \leq x) \\
&= \int_{-\infty}^{x} f(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{y}{2\sigma^2}} dy, \textit{with } \frac{x-\mu}{\sigma} = y \\
&= \Phi(\frac{x-\mu}{\sigma})
\end{aligned}$$

Important Continuous Distributions

Normal Random Variables)

A standard normal variable is one with the parameters $\mu = 0$ and $\sigma = 1$.

The standard normal density $\psi(x)$ is defined by

$$\psi(x) \quad = \quad \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty$$

Important Continuous Distributions

Cauchy Random Variables

A continuous random variable X is said to have a Cauchy distribution with parameters $\lambda > 0$ and μ if its density function is given by

$$f(x) = \frac{1}{\pi} \frac{\lambda}{\lambda^2 + (x-\mu)^2}, \quad -\infty < x < \infty$$

X is also called Cauchy (λ, μ)-variate.

Gamma Distribution/Random Variables

A continuous random variable X is said to have a Gamma distribution with parameters $l > 0$ if its density function is given by

$$f(x) = \begin{cases} \frac{e^{-x}x^{l-1}}{\Gamma(l)}, & x > 0 \\ 0, & otherwise \end{cases}$$

where $\Gamma(n)$ is defined by

$$\begin{aligned}
\Gamma(n) &= \int_0^{-\infty} e^{-x}x^{n-1}dx, with\, 0 < n < \infty \\
\Gamma(n+1) &= n\Gamma(n), \\
\Gamma(n+1) &= n!, n = 1, 2, 3, \ldots \\
\Gamma(\frac{1}{2}) &= \sqrt{\pi}, \\
\Gamma(1) &= 1.
\end{aligned}$$

Problem (Probability density function (p.d.f))

A point X is chosen at random on a line segment AB whose middle point is O (origin).

Find the probability that AX, BX and AO can form the sides of a triangle.

End of this lecture