Lecture Notes on Probability

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

RANDOM VARIABLS

* A random variable is a mapping from the event space of a random experiment to the set of real numbers.

* Consider a random experiment E having event space S. Then a map $X : S \to R$ is known as random variable, where R is the set of real numbers.

* **Example:** E = Experiment of tossing a coin

  $S = \{H, T\}$

  $H \to$ occurrence of Head

  $T \to$ occurrence of Tail

  Define $X : S \to R$ by

  $X(H) = 0$ & $X(T) = 1$

Here the event $(X = 0) \Leftrightarrow$ the occurrence of $\{H\}$

" " $(X = 1) \Leftrightarrow$ the occurrence of $\{T\}$

" " $(X = x) \Leftrightarrow$ the occurrence of Φ when $x \neq 0$ and $x \neq 1$

" " $(a \leq X \leq b) \Rightarrow$ the occurrence of at least one event point of S for which the random variable lies within the interval $[a, b]$

$(0 < X < 1) = \Phi$

$(X < 0) = \Phi$

$(0 \le X < 1) = (X = 0) + (0 < X < 1) = \{H\} + \Phi = \{H\}$

$(0 < X \le 1) = (0 < X < 1) + (X = 1) = \Phi + \{T\} = \{T\}$

$(0 \le X \le 1) = (X = 0) + (0 < X < 1) + (X = 1) = \{H\} + \Phi + \{T\} = \{H, T\}$

$(-\infty < X < 0) = \Phi$

$(-\infty < X \le x) = (-\infty < X < 0) + (0 \le X < x)$ for $0 < x < 1$

$= \Phi + (X = 0) + (0 < X < x)$

$= \Phi + \{H\} + \Phi = \{H\}$

$(-\infty < X \le 0) = (-\infty < X < 0) + (X = 0) = \Phi + \{H\} = \{H\}$

$(-\infty < X \le 1) = (-\infty < X < 0) + (X = 0) + (0 < X < 1) + (X = 1)$

$= \Phi + \{H\} + \Phi + \{T\} = \{H, T\}$

$(-\infty < X \le x) = (-\infty < X < 0) + (X = 0) + (0 < X < 1) + (X = 1) + (1 < X \le x)$

(for $x > 1$)

$= \Phi + \{H\} + \Phi + \{T\} + \Phi = \{H, T\}$

$(1 \le X < \infty) = (X = 1) + (1 < X < \infty) = \{T\} + \Phi = \{T\}$

$(-5 < X \le -3) = \Phi$

$(3 < X \le 5) = \Phi$

$(X > 2) = \Phi$

$(X \ge 2) = \Phi$

* Probability Distribution Function:

Let X be a random variable connected with a random experiment E having event space S. Then the probability distribution function of X is generally denoted by $F_X(x)$ or $F(x)$ and defined by

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

Example: Find the distribution (probability distribution) function of a random variable X connected with the random experiment of tossing a coin defined by

$X(H) = 0$ & $X(T) = 1$

where $H(T)$ denotes the occurrence of Head (Tail)

Answer: According to the definition, we have

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

i) For $x < 0$, $(-\infty < X \leq x) = \Phi$

$$\Rightarrow P(-\infty < X \leq x) = P(\Phi) = 0$$

$$\Rightarrow F(x) = 0$$

$\therefore$ $F(x) = 0$ for $x < 0$ ———— ①

ii) For $x=0$, $(-\infty < X \le x) = (-\infty < X \le 0)$

$$= (-\infty < X < 0) + (X=0)$$

$$= \Phi + \{H\} = \{H\}$$

$\Rightarrow P(-\infty < X \le x) = P(H) = \frac{1}{2} \Rightarrow F(x) = \frac{1}{2}$

$\Rightarrow F(x) = \frac{1}{2}$ for $x = 0$ ——————— ②

iii) For $0 < x < 1$,

$$(-\infty < X \le x) = (-\infty < X < 0) + (X=0) + (0 < X \le x)$$

$$= \Phi + \{H\} + \Phi = \{H\}$$

$\Rightarrow P(-\infty < X \le x) = P(H) = \frac{1}{2}$

$\Rightarrow F(x) = \frac{1}{2}$

$\therefore F(x) = \frac{1}{2}$ for $0 < x < 1$ ——————— ③

iv) For $x = 1$,

$$(-\infty < X \le x) = (-\infty < X \le 1)$$

$$= (-\infty < X < 0) + (X=0) + (0 < X < 1) + (X=1)$$

$$= \Phi + \{H\} + \Phi + \{T\} = \{H, T\} = S$$

$\Rightarrow P(-\infty < X \le x) = P(S) = 1$

$\Rightarrow F(x) = 1$

$\therefore F(x) = 1$ for $x = 1$ ——————— ④

v) For $x > 1$,

$$(-\infty < X \leq x) = (-\infty < X < 0) + (X = 0) + (0 < X < 1)$$
$$+ (X = 1) + (1 < X \leq x)$$
$$= \Phi + \{H\} + \Phi + \{T\} + \Phi$$
$$= \{H, T\} = S$$

$\Rightarrow P(-\infty < X \leq x) = P(S) = 1$

$\Rightarrow F(x) = 1$

$\therefore F(x) = 1$ for $x > 1$ ———— (5)

(1), (2), (3), (4) and (5) can be written as

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{2} & \text{for } x = 0 \\ \frac{1}{2} & \text{for } 0 < x < 1 \\ 1 & \text{for } x = 1 \\ 1 & \text{for } x > 1 \end{cases} = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{2} & \text{for } 0 \leq x < 1 \\ 1 & \text{for } x \geq 1 \end{cases}$$

Properties of Distribution function

P1 $P(a < X \leq b) = F(b) - F(a)$

P2 Distribution function is monotonically non-decreasing

P3 $F(\infty) = 1$ **P4** $P(-\infty) = 0$

P5 $F(a) - F(a-0) = P(X = a)$

P6 $F(a+0) = F(a)$

Proof of P1

$$F(b) = P(-\infty < x \le b)$$

$$= P\{(-\infty < x \le a) \cup (a < x \le b)\}$$

$$= P(-\infty < x \le a) + P(a < x \le b) \quad [A3]$$

$$= F(a) + P(a < x \le b)$$

$$\Rightarrow P(a < x \le b) = F(b) - F(a)$$

Proof of P2: Consider any two real numbers x_1 and x_2 such that $x_2 \ge x_1$

If $x_2 = x_1$ then $F(x_2) = F(x_1)$ ——— ①

If $x_2 > x_1$ then

$$F(x_2) = P(-\infty < x \le x_2)$$

$$= P\{(-\infty < x \le x_1) \cup (x_1 < x \le x_2)\}$$

$$= P(-\infty < x \le x_1) + P(x_1 < x \le x_2) \quad [A3]$$

$$= F(x_1) + P(x_1 < x \le x_2)$$

$$\Rightarrow F(x_2) - F(x_1) = P(x_1 < x \le x_2) \ge 0 \quad [A1]$$

$$\Rightarrow F(x_2) - F(x_1) \ge 0$$

$$\Rightarrow F(x_2) \ge F(x_1) \quad \text{———} \quad ②$$

Combining ① and ②, we get

$F(x_2) \ge F(x_1)$ when $x_2 \ge x_1$

i.e., $x_2 \ge x_1 \Longrightarrow F(x_2) \ge F(x_1)$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

$\Rightarrow F$ is monotonically non-decreasing

Proof of P3 : Consider the sequence of events $\{A_n\}$ defined by

$A_n = (-\infty < x \leq n) = (-\infty, n]$

Now $y \in A_n \Rightarrow y \in (-\infty < x \leq n)$

$\Rightarrow -\infty < y \leq n$

$\Rightarrow -\infty < y \leq n < n+1$

$\Rightarrow y \in (-\infty, n+1) \subseteq (-\infty, n+1]$

$\Rightarrow y \in A_{n+1} = (-\infty, n+1]$

$\therefore A_n \subseteq A_{n+1} \Rightarrow \{A_n\}$ is expanding

$\Rightarrow \lim A_n = \bigcup_{n=1}^{\infty} A_n$ ——— ①

& $P(\lim A_n) = \lim P(A_n)$ ——— ②

Now $y \in \lim A_n \Rightarrow y \in \bigcup_{n=1}^{\infty} A_n$

$\Rightarrow y \in A_r$ for some $r \in N$

$\Rightarrow y \in (-\infty, r] \subseteq (-\infty, \infty)$

$\Rightarrow \lim A_n \subseteq (-\infty, \infty)$ ——— ③

Let $y \in (-\infty, \infty)$. Then for the real numbers y and $x = 1 > 0$, there exists a positive integer m (from the Archimedean property

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

of real numbers) such that

$mx > y \Rightarrow m.1 > y \Rightarrow y < m$

$\Rightarrow -\infty < y < m \Rightarrow y \in (-\infty, m) \subseteq (-\infty, m] = A_m$

$\Rightarrow y \in A_m \Rightarrow y \in \bigcup_{n=1}^{\infty} A_n \Rightarrow y \in \lim A_n$

$\therefore y \in (-\infty, \infty) \Rightarrow y \in \lim A_n$

$\therefore (-\infty, \infty) \subseteq \lim A_n$ ——— (4)

$\therefore$ From (3) & (4), we get

$\lim A_n = (-\infty, \infty) = S$

$\Rightarrow P(\lim A_n) = P(S) = 1$ [A2]

$\Rightarrow \lim P(A_n) = 1$ [using (2)]

$\Rightarrow \lim P(-\infty < X \leq n) = 1$

$\Rightarrow \lim F(n) = 1$

$\Rightarrow \boxed{F(\infty) = 1}$

<u>Proof of P4</u>: Consider the sequence of events $\{A_n\}$ defined by

$A_n = (-\infty < X \leq -n) = (-\infty, -n]$ ——— (1)

For $n \in N$, $n+1 > n \Rightarrow -(n+1) < -n$ ——— (2)

$\therefore y \in A_{n+1} \Rightarrow y \in (-\infty, -(n+1)]$

$\Rightarrow -\infty < y \leq -(n+1) < -n$ [from (2)]

$\Rightarrow -\infty < y < -n$

$\Rightarrow y \in (-\infty, -n) \subseteq (-\infty, -n] = A_n$

$\Rightarrow y \in A_n$

$\therefore y \in A_{n+1} \Rightarrow y \in A_n$

$\therefore A_{n+1} \subseteq A_n \quad \forall n \in N$

$\Rightarrow \{A_n\}$ is contracting

$\Rightarrow \lim A_n = \bigcap_{n=1}^{\infty} A_n$ ——— (3)

& $P(\lim A_n) = \lim P(A_n)$ ——— (4)

Now we shall prove that $\lim A_n = \Phi$. If possible let $y \in \lim A_n$ for some $y \in (-\infty, \infty)$

$y \in \lim A_n \Rightarrow y \in \bigcap_{n=1}^{\infty} A_n$

$\Rightarrow y \in A_n \quad \forall n \in N$ ——— (5)

$\Rightarrow y \in (-\infty < x \leq -n) \quad \forall n \in N$

$\Rightarrow -\infty < y \leq -n \quad \forall n \in N$ ——— (6)

Again, from Archimedean property, for the real numbers $-y$ and $x = 1 > 0$, there exists a positive integer m such that

$m \cdot x > -y \Rightarrow m \cdot 1 > -y \Rightarrow -y < m$ — (7)

Putting $n = m$ in (6), we get

$-\infty < y \leq -m \Rightarrow y \leq -m \Rightarrow -y \geq m$ —— (8)

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

From (7) & (8), we get

$-y < m \leq -y \Rightarrow -y < -y$

which is untrue.

Therefore our supposition is wrong and consequently

$\lim A_n = \Phi$

$\Rightarrow P(\lim A_n) = P(\Phi) = 0$

$\Rightarrow \lim P(A_n) = 0$ [from (4)]

$\Rightarrow \lim P(-\infty < x \leq -n) = 0$

$\Rightarrow \lim F(-n) = 0$

$\Rightarrow F(-\infty) = 0$

<u>Proof of P5</u>: Consider the sequence of events defined by

$A_n = (a - \frac{1}{n} < x \leq a) = (a - \frac{1}{n}, a]$ ——— (1)

$n+1 > n \Rightarrow \frac{1}{n+1} < \frac{1}{n} \Rightarrow -\frac{1}{n+1} > -\frac{1}{n}$

$\Rightarrow a - \frac{1}{n+1} > a - \frac{1}{n}$ ——— (2)

Now $y \in A_{n+1} \Rightarrow a - \frac{1}{n+1} < y \leq a$

$\Rightarrow a - \frac{1}{n} < a - \frac{1}{n+1} < y \leq a$ [using (2)]

$\Rightarrow a - \frac{1}{n} < y \leq a \Rightarrow y \in A_n$

$\therefore y \in A_{n+1} \Rightarrow y \in A_n$

$\therefore\ A_{n+1} \subseteq A_n \quad \forall n \in \mathbb{N}$

$\Rightarrow \{A_n\}$ is contracting

$\Rightarrow \lim A_n = \bigcap_{n=1}^{\infty} A_n$ ——— (3)

& $P(\lim A_n) = \lim P(A_n)$ ——— (4)

For all $n \in \mathbb{N}$, $a - \frac{1}{n} < a \leq a \Rightarrow a \in A_n$

$\Rightarrow a \in A_n \quad \forall n \in \mathbb{N}$

$\Rightarrow a \in \bigcap_{n=1}^{\infty} A_n$

$\Rightarrow a \in \lim A_n$

If possible let $y \in \lim A_n$ such that $y \neq a$

$\Rightarrow y \in \bigcap_{n=1}^{\infty} A_n$ and $y \neq a$

$\Rightarrow y \in A_n$ for all $n \in \mathbb{N}$ and $y \neq a$

$\Rightarrow a - \frac{1}{n} < y \leq a \quad \forall n \in \mathbb{N}$ and $y \neq a$

$\Rightarrow -\frac{1}{n} < y - a \leq 0 \quad \forall n \in \mathbb{N}$ and $y \neq a$

$\Rightarrow +\frac{1}{n} > a - y \geq 0 \quad \forall n \in \mathbb{N}$ and $y \neq a$

$\Rightarrow \frac{1}{n} > |a-y| \quad \forall n \in \mathbb{N}$ and $|y-a| > 0$

$\Rightarrow |a-y| < \frac{1}{n} \quad \forall n \in \mathbb{N}$ and $|a-y| > 0$ ——— (5)

From the Archimedean property, for the real numbers $x = |a-y| > 0$ and $y = 1$, there exists a positive integer m such that

$m|a-y| > 1 \Rightarrow |a-y| > \frac{1}{m}$ ——— (6)

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

Putting $n=m$ in ⑤, we get

$|a-y| < \frac{1}{m}$ ——— ⑦

From ⑦ & ⑥, we get

$|a-y| < \frac{1}{m} < |a-y| \Rightarrow |a-y| < |a-y|$

which is untrue and consequently our supposition is wrong.

Therefore $\lim A_n = \{a\} = (X=a)$

$\Rightarrow P(\lim A_n) = P(X=a)$

$\Rightarrow \lim P(A_n) = P(X=a)$

$\Rightarrow \lim P(a-\frac{1}{n} < X \leq a) = P(X=a)$

$\Rightarrow \lim [F(a) - F(a-\frac{1}{n})] = P(X=a)$

$\Rightarrow F(a) - \lim F(a-\frac{1}{n}) = P(X=a)$

$\Rightarrow F(a) - \lim\limits_{h\to 0+} F(a-h) = P(X=a)$, $h=\frac{1}{n}$

$\Rightarrow F(a) - F(a-0) = P(X=a)$

<u>Proof of P6</u>: Consider the sequence of events $\{A_n\}$ defined by

$A_n = (a < X \leq a+\frac{1}{n}) = (a, a+\frac{1}{n}]$

Now $\forall n \in N$, $n+1 > n \Rightarrow \frac{1}{n+1} < \frac{1}{n}$

$\Rightarrow a+\frac{1}{n+1} < a+\frac{1}{n}$ ——— ①

$y \in A_{n+1} \Rightarrow y \in (a < X \leq a+\frac{1}{n+1})$

$\Rightarrow a < y \leq a + \frac{1}{n+1} < a + \frac{1}{n}$ [from ①]

$\Rightarrow a < y < a + \frac{1}{n}$

$\Rightarrow y \in (a, a+\frac{1}{n}) \subseteq (a, a+\frac{1}{n}] = A_n$

$\Rightarrow y \in A_n$

$\therefore y \in A_{n+1} \Rightarrow y \in A_n$

$\therefore A_{n+1} \subseteq A_n \quad \forall n \in N$

$\Rightarrow \{A_n\}$ is contracting

$\Rightarrow \lim A_n = \bigcap_{n=1}^{\infty} A_n$ ——— ②

& $P(\lim A_n) = \lim P(A_n)$ ——— ③

Now we shall prove that $\lim A_n = \Phi$

If possible let $y \in \lim A_n$

$\Rightarrow y \in \bigcap_{n=1}^{\infty} A_n$ [from ②]

$\Rightarrow y \in A_n \quad \forall n \in N$

$\Rightarrow y \in (a < x \leq a + \frac{1}{n}) \quad \forall n \in N$

$\Rightarrow a < y \leq a + \frac{1}{n} \quad \forall n \in N$

$\Rightarrow 0 < y - a \leq \frac{1}{n} \quad \forall n \in N$

$\Rightarrow |y-a| \leq \frac{1}{n} \quad \forall n \in N$ with $|y-a| > 0$ ——— ④

Again, from the Archimedean property

of two real numbers $x = |y-a| > 0$ and $y_0 = 1$, we can find a positive integer m such that

$$m x > y_0 \Rightarrow m|y-a| > 1 \Rightarrow |y-a| > \frac{1}{m} \quad \text{——— (5)}$$

Putting $n = m$ in (4), we get

$$|y-a| \le \frac{1}{m} \Rightarrow \frac{1}{m} \ge |y-a| \quad \text{——— (6)}$$

From (5) & (6), we get

$$|y-a| > \frac{1}{m} \ge |y-a|$$

$$\Rightarrow |y-a| > |y-a|$$

which is untrue and consequently our supposition is wrong

$\therefore \lim A_n = \Phi$

$\Rightarrow P(\lim A_n) = P(\Phi) = 0$

$\Rightarrow \lim P(A_n) = 0$ [using (3)]

$\Rightarrow \lim P(a < x \le a + \frac{1}{n}) = 0$

$\Rightarrow \lim [F(a + \frac{1}{n}) - F(a)] = 0$

$\Rightarrow \lim F(a + \frac{1}{n}) - F(a) = 0$

$\Rightarrow \lim\limits_{h \to 0+} F(a+h) - F(a) = 0$

$\Rightarrow F(a+0) - F(a) = 0$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

* Properties P5 and P6 indicate that

1. the distribution function is continuous from right
2. the distribution function is not continuous from left
3. $F(a+0)$, $F(a-0)$ both exist
4. $F(a+0) - F(a-0) = F(a) - F(a) + P(X=a) = P(X=a)$
5. the distribution function has jump type discontinuity only at $x=a$ having height of jump $P(X=a)$.

Probability Mass function and Probability density function:

Consider a straight line and suppose a mass is distributed along this straight line, the total mass being unity. Then we can compare the distribution of mass on this straight line with the probability distribution of a random variable X, where

total mass = total probability of X

$$= P(-\infty < X < \infty) = F(\infty) = 1$$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

The mass included within the interval $(x, x+\delta x]$

$= P(x < X \le x+\delta x) = F(x+\delta x) - F(x)$

$\therefore$ the line density of the mass

$$= \lim_{\delta x \to 0} \frac{F(x+\delta x) - F(x)}{\delta x}$$

If the limit exists finitely, then the line density or linear mass density is given by

$$f_X(x) = f(x) = F'(x) = F_X'(x) = \lim_{\delta x \to 0} \frac{F(x+\delta x) - F(x)}{\delta x}$$

This mass density is known as the probability density function

P1 Existence of probability density function

$\Rightarrow$ Existence of $F'(x)$

$\Rightarrow$ $F(x)$ is continuous

$\Rightarrow$ $F(x+0) = F(x-0) = F(x)$

$\Rightarrow$ $F(x+0) - F(x-0) = 0$

$\Rightarrow$ $P(X = x) = 0$

Continuous random variable: A random variable X is said to be continuous if its probability distribution function is differentiable.

Discrete random variable: A random variable X is said to be discrete if its probability distribution function is not differentiable.

Properties of probability density function (pdf):

P1: $f(x) \geq 0 \quad \forall x$

Proof: Let $F(x)$ is the probability distribution function of a continuous random variable X having pdf $f(x)$.

Now $F(x)$ is monotonically non-decreasing

$\Rightarrow F'(x) \geq 0 \quad \forall x$

$\Rightarrow f(x) \geq 0 \quad \forall x$

P2: $\int_a^b f(x)\,dx = F(b) - F(a) = P(a < x \leq b) = P(a \leq x \leq b)$

Proof: $\int_a^b f(x)\,dx = \int_a^b F'(x)\,dx = \int_a^b \frac{dF}{dx}\,dx$

$= \int_a^b dF = \int_a^b dF(x) = F(b) - F(a)$

P3 $F(x) = \int_{-\infty}^{x} f(t)\,dt$

Proof: From P2, we get

$$F(b) - F(a) = \int_a^b f(x)\,dx = \int_a^b f(t)\,dt \quad \text{——— (1)}$$

Putting $b = x$ in (1), we get

$$F(x) - F(a) = \int_a^x f(t)\,dt \quad \text{——— (2)}$$

Making limit as $a \longrightarrow -\infty$, we get, from (2),

$$F(x) - F(-\infty) = \int_{-\infty}^{x} f(t)\,dt$$

$$\Rightarrow F(x) = \int_{-\infty}^{x} f(t)\,dt \quad [\text{as } F(-\infty) = 0]$$

P4 $\int_{-\infty}^{\infty} f(x)\,dx = 1$

Proof: From P3, we get

$$F(x) = \int_{-\infty}^{x} f(t)\,dt \quad \text{——— (1)}$$

Making limit as $x \to \infty$, we get

$$F(\infty) = \int_{-\infty}^{\infty} f(t)\,dt$$

$$\Rightarrow \int_{-\infty}^{\infty} f(t)\,dt = F(\infty) = 1 \quad [\text{as } F(\infty) = 1]$$

$$\Rightarrow \int_{-\infty}^{\infty} f(t)\,dt = 1 \Rightarrow \int_{-\infty}^{\infty} f(x)\,dx = 1$$

* A function $f(x)$ is said to be a pdf of a continuous random variable X if P1 and P4 hold simultaneously.

Example: Let

$$f(x) = kx(1-x) \quad \text{for } 0 \le x \le 1$$
$$= 0 \quad \text{otherwise}$$

be a pdf of a continuous random variable X.

Find k and $P(X > \frac{1}{2})$

Answer: $f(x)$ is a pdf of X

$$\Rightarrow \int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\Rightarrow \underbrace{\int_{-\infty}^{0} f(x)\,dx}_{0} + \int_{0}^{1} f(x)\,dx + \underbrace{\int_{1}^{\infty} f(x)\,dx}_{0} = 1$$

$$\Rightarrow \int_{0}^{1} f(x)\,dx = 1$$

$$\Rightarrow \int_{0}^{1} kx(1-x)\,dx = 1$$

$$\Rightarrow k\int_{0}^{1} (x - x^2)\,dx = 1$$

$$\Rightarrow k\left(\frac{x^2}{2} - \frac{x^3}{3}\right)\Big|_0^1 = 1$$

$$\Rightarrow k\left(\frac{1}{2} - \frac{1}{3}\right) = 1 \Rightarrow \boxed{k = 6}$$

$$P(X > \tfrac{1}{2}) = 1 - P(X \le \tfrac{1}{2}) = 1 - P(-\infty < X \le \tfrac{1}{2})$$

$$= 1 - \int_{-\infty}^{1/2} f(x)\,dx$$

$$= 1 - \Big(\underbrace{\int_{-\infty}^{0} f(x)\,dx}_{0} + \int_{0}^{1/2} f(x)\,dx \Big)$$

$$= 1 - \int_{0}^{1/2} f(x)\,dx$$

$$= 1 - \int_{0}^{1/2} k\,x(1-x)\,dx$$

$$= 1 - k\left(\frac{x^2}{2} - \frac{x^3}{3}\right)\Big|_0^{1/2}$$

$$= 1 - k\left(\frac{1}{2}\cdot\frac{1}{2^2} - \frac{1}{3}\cdot\frac{1}{2^3}\right)$$

$$= 1 - k\,\frac{1}{2^3}\left(1 - \frac{1}{3}\right)$$

$$= 1 - k\cdot\frac{1}{2^3}\cdot\frac{2}{3}$$

$$= 1 - k\cdot\frac{1}{2^2\cdot 3} = 1 - \frac{6}{2^2\cdot 3} = 1 - \frac{1}{2} = \frac{1}{2}$$

<u>Problem</u>: Let the spectrum of a random variable X consists of points $\{1, 2, \dots, n\}$. If

$$P(X = i) \propto \frac{1}{i(i+1)} \quad \forall\, i = 1(1)n$$

then find $P(m < X \le r)$, where m, r are both positive integers and $m, r \le n$.

Answer: $P(-\infty < x < \infty) = F(\infty) = 1$

$$\Rightarrow \sum_{i=1}^{n} P(x=i) = 1 \quad \text{——————} \quad (1)$$

Now $P(x=i) \propto \frac{1}{i(i+1)}$

$\Rightarrow P(x=i) = \frac{k}{i(i+1)}$ [$k \to$ proportionality constant]

$$\Rightarrow P(x=i) = k\left[\frac{1}{i} - \frac{1}{i+1}\right] \quad \text{——————} \quad (2)$$

Putting $i = 1, 2, \dots, s$ in (2), we get

$P(x=1) = k\left[\frac{1}{1} - \frac{1}{2}\right]$

$P(x=2) = k\left[\frac{1}{2} - \frac{1}{3}\right]$

$P(x=3) = k\left[\frac{1}{3} - \frac{1}{4}\right]$

$\dots \quad \dots \quad \dots$

$P(x=s) = k\left[\frac{1}{s} - \frac{1}{s+1}\right]$

$P(x=1) + P(x=2) + \dots + P(x=s) = k\left[1 - \frac{1}{s+1}\right]$ (adding columnwise)

$$\Rightarrow \sum_{i=1}^{s} P(x=i) = k\frac{s}{s+1} \quad \text{——————} \quad (3)$$

Putting $s = n$ in (3), we get

$$\sum_{i=1}^{n} P(x=i) = k\frac{n}{n+1} \quad \text{——————} \quad (4)$$

From (1) and (4), we get

$k\frac{n}{n+1} = 1 \Rightarrow \boxed{k = \frac{n+1}{n}}$

Now $F(s) = P(-\infty < X \le s)$

$$= P(X=1) + P(X=2) + \cdots + P(X=s)$$

$$= \sum_{i=1}^{s} P(X=i) = k\frac{s}{s+1}$$

$\therefore\ P(m < X \le r) = F(r) - F(m)$

$$= k\frac{r}{r+1} - k\frac{m}{m+1}$$

$$= k\left[\frac{r}{r+1} - \frac{m}{m+1}\right]$$

$$= k\,\frac{r(m+1) - m(r+1)}{(r+1)(m+1)}$$

$$= k\,\frac{r-m}{(r+1)(m+1)}$$

$$= \frac{n+1}{n}\cdot\frac{r-m}{(r+1)(m+1)}$$

<u>Theorem</u>: Let X be a continuous random variable having pdf $f_X(x)$. Let $y = g(x)$ be a continuously differentiable strictly monotonic function, i.e., either $g'(x) > 0 \quad \forall x$

or $g'(x) < 0 \quad \forall x$.

Then the pdf of the random variable $Y = g(X)$ is given by

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

Case 1: Let $y = g(x)$ be strictly monotonic increasing

$\therefore \frac{dy}{dx} = g'(x) > 0 \quad \forall x$

$\Rightarrow g^{-1}$ exists finitely and it is also strictly monotonic increasing

$$\therefore F_Y(y) = P(-\infty < Y \leq y) = P(Y \leq y)$$
$$= P(Y \leq y)$$
$$= P(g(X) \leq g(x))$$
$$= P(g^{-1}(g(X)) \leq g^{-1}(g(x)))$$
$$= P(I(X) \leq I(x))$$
$$= P(X \leq x)$$
$$= F_X(x)$$

$$\Rightarrow F_Y'(y) = \frac{d}{dy}(F_Y(y)) = \frac{d}{dy}(F_X(x)) = \frac{d}{dx}(F_X(x))\frac{dx}{dy}$$

$$\Rightarrow F_Y'(y) = F_X'(x)\frac{dx}{dy}$$

$$\Rightarrow f_Y(y) = f_X(x)\frac{dx}{dy}$$

$$\Rightarrow f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right| \quad \left[\text{As } \frac{dy}{dx} > 0 \Rightarrow \frac{dx}{dy} > 0 \Rightarrow \frac{dx}{dy} = \left|\frac{dx}{dy}\right|\right] \qquad (1)$$

Case-2: Let $y = g(x)$ be strictly decreasing

$\therefore \frac{dy}{dx} = g'(x) < 0 \quad \forall x$

$\Rightarrow g'(x)$ exists finitely and it is also strictly monotonically decreasing.

$$F_Y(y) = P(-\infty < Y \le y) = P(Y \le y)$$

$$= P(g(X) \le g(x))$$

$$= P(g^{-1}(g(X)) \ge g^{-1}(g(x))) = P(I(X) \ge I(x))$$

$$= P(X \ge x)$$

$$= 1 - P(X < x)$$

$$= 1 - P(X \le x) \quad [\text{as } P(X = x) = 0]$$

$$= 1 - F_X(x)$$

$$\Rightarrow F_Y'(y) = \frac{d}{dy}(F_Y(y)) = \frac{d}{dy}\{1 - F_X(x)\}$$

$$= -\frac{d}{dy}(F_X(x))$$

$$= -\frac{d}{dx}(F_X(x))\frac{dx}{dy}$$

$$= -F_X'(x)\frac{dx}{dy}$$

$$\Rightarrow f_Y(y) = -f_X(x)\frac{dx}{dy}$$

$$= f_X(x)\left|\frac{dx}{dy}\right| \quad \text{As } \frac{dy}{dx} < 0 \Rightarrow \frac{dx}{dy} < 0$$

$$\Rightarrow \left|\frac{dx}{dy}\right| = -\frac{dx}{dy}$$

$$\Rightarrow f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right| \quad \text{———} \quad (2)$$

Combining (1) and (2), we can write

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

Discrete Distributions:

(1) Bernoulli Trials and Binomial distribution:

Consider a random experiment E having event space S. Consider any event A connected with E. Instead of the events A or $\bar{A}$, we shall use two symbols or abbreviations "success" or "s" and "failure" or "f" defined as follows:

$s \equiv$ occurrence of A

$f \equiv$ occurrence of $\bar{A}$

Therefore the random experiment E can be regarded as a random experiment whose event space contains only two event points namely success or s and failure or f where the 'success' ('s') denotes the occurrence of the event A whereas the 'failure' ('f') denotes the occurrence of $\bar{A}$.

$P(s) = P(\text{success}) = P(A) = p$ (say)

$P(f) = P(\text{failure}) = P(\bar{A}) = 1 - P(A) = 1-p$

If $1-p = q$ then $p+q=1$ with $0<p<1$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

If $P(A) = P(S) = p$ remains unchanged at each trial of the random experiment E, then a sequence of trials of E is known as Bernoulli trials.

Consider n repetitions of the random experiment E. Then our aim is to find the probability of exactly i success out of n independent repetitions of E.

Let A_i be the event of exactly i success connected with the compound random experiment E_n, which consists of n independent repetitions of E.

Let n trials be represented by n rooms.

Therefore, each room is filled by either 's' or 'f' and consequently an event point in favour of A_i is of the following form:

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

$$\boxed{s \,|\, f \,|\, f \,|\, f \,|\, \cdots \,|\, s \,|\, f}$$

where exactly i rooms are filled by 's's and exactly $(n-i)$ rooms are filled by 'f's, i.e.,

$$\underbrace{s\,f\,f\,f\,\cdots\,s\,f}$$

$i \rightarrow$ success and $n-i \rightarrow$ failure

$$P(sff\cdots sf) = P(s)\,P(f)\,P(f)\cdots P(s)\,P(f)$$

$$= \underbrace{P(s)\,P(s)\cdots P(s)}_{i}\ \underbrace{P(f)\,P(f)\cdots P(f)}_{n-i}$$

$$= (P(s))^{i}\,(P(f))^{n-i}$$

$$= p^{i} q^{n-i}$$

Now i rooms can be chosen from n rooms in ${}^{n}c_{i}$ ways and consequently A_i contains ${}^{n}c_{i}$ event points each having the probability $p^{i} q^{n-i}$

$$\therefore\ P(A_i) = {}^{n}c_{i}\, p^{i} q^{n-i}, \quad i = 0, 1, 2, \ldots, n,$$
$$p+q=1, \quad 0<p<1$$

This law is known as Binomial law and we can define the following random variable

$$P(x=i) = {}^{n}c_{i}\, p^{i} q^{n-i}, \quad i = 0(1)n, \quad p+q=1, \quad 0<p<1$$

This probability distribution is known as Binomial distribution with parameters n and p, and is generally denoted by $B(n,p)$, i.e., $X = B(n,p)$.

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

Poisson Distribution: Poisson distribution is defined by

$$P(x=i) = \frac{e^{-\mu}\mu^{i}}{i!}, \quad i = 0, 1, 2, \ldots \quad \text{with } \mu > 0$$

Therefore the spectrum of the random variable X contains all non-negative integers, i.e.,

$x = i, \quad i = 0, 1, 2, 3, \ldots$

Theorem: Poisson distribution is the limiting case of Binomial distribution.

Proof: For $B(n, p)$ we have

$$P(x=i) = \binom{n}{i} p^{i} q^{n-i}, \quad i = 0, 1, 2, \ldots, n, \quad p+q=1, \quad 0<p<1$$

$$P(x=i) = \frac{n!}{i!(n-i)!} p^{i} q^{n-i}$$

$$= \frac{n(n-1)\cdots(n-i+1)(n-i)!}{i!\,(n-i)!} p^{i} q^{n-i}$$

$$= \frac{n(n-1)\cdots(n-i+1)}{i!} p^{i} q^{n-i} \quad \text{——} \quad (1)$$

Putting $p = \frac{\mu}{n} \Leftrightarrow \mu = np$ in (1), we get

$$P(x=i) = \frac{n(n-1)\cdots(n-\overline{i-1})}{i!} \left(\frac{\mu}{n}\right)^{i} \left(1-\frac{\mu}{n}\right)^{n-i}$$

$$= \frac{\mu^{i}}{i!} \cdot \frac{n(n-1)\cdots(n-\overline{i-1})}{n^{i}} \frac{\left(1-\frac{\mu}{n}\right)^{n}}{\left(1-\frac{\mu}{n}\right)^{i}}$$

$$P(X=i) = \frac{\mu^i}{i!}\,\frac{n(n-1)\cdots(n-\overline{i-1})}{n^i}\cdot\frac{\left(1-\frac{\mu}{n}\right)^n}{\left(1-\frac{\mu}{n}\right)^i}$$

$$= \frac{\mu^i}{i!}\cdot\frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{i-1}{n}\right)}{\left(1-\frac{\mu}{n}\right)^i}\cdot\left(1-\frac{\mu}{n}\right)^n$$

$$\longrightarrow \frac{\mu^i}{i!}e^{-\mu} \quad \text{as } n\to\infty$$

because $\left(1-\frac{\mu}{n}\right) = \left[\left(1-\frac{\mu}{n}\right)^{-\left(\frac{\mu}{n}\right)^{-1}}\right]^{-\mu}$

$$= \left\{\left(1-\frac{\mu}{n}\right)^{-\frac{n}{\mu}}\right\}^{-\mu} \to e^{-\mu} \quad \text{as } n\to\infty$$

In this process, we have the following assumption:

$p = \frac{\mu}{n}$ and $\mu = np$

① If p is constant, then $\mu\to\infty$ as $n\to\infty$, i.e., μ is not finite for large value of n.

② If μ is constant, then $p\to 0$ as $n\to\infty$, this value of p is also not possible.

✓Therefore we take p in such a way that $\mu = np$ is of moderate value.

<u>Stochastic process</u>: A family of random variable $X(t)$, which depends parametrically on t, is usually called a stochastic process.

Poisson Process: Poisson process is a particular example of stochastic process. Here we count the number of changes (a general name) in a given interval of t. Here the parameter t is known as time (a general name).

Hypothesis for any Poisson Process:

① The number of changes during the time interval $(t, t+h)$ is independent of the number of changes in $(0, t)$ $\forall t > 0$ & $\forall h > 0$.

② The probability of exactly one change in $(t, t+h)$ is $\lambda h + o(h)$, where $\lambda > 0$ and $\frac{o(h)}{h} \to 0$ as $h \to 0$

③ The probability of more than one change in $(t, t+h)$ is $o(h)$.

Let the event $(X(t) = i)$ denotes the number of exactly i changes in the interval $(0, t)$

Let $A_i = (X(t) = i)$

Let B_i be the event of exactly i changes in $(t, t+h)$.

$\therefore$ $(X(t+h)=i)$ = the event of exactly i changes in $(0, t+h)$

$= A_i B_0 + A_{i-1} B_1 + A_{i-2} B_2 + \cdots + A_0 B_i$

$$\Rightarrow P(X(t+h)=i) = P(A_i B_0 + A_{i-1} B_1 + A_{i-2} B_2 + \cdots + A_0 B_i)$$

$$= P(A_i B_0) + P(A_{i-1} B_1) + \cdots + P(A_0 B_i)$$

$$= P(A_i B_0) + P(A_{i-1} B_1) + \sum_{r=2}^{i} P(A_{i-r} B_r)$$

$$= P(A_i) P(B_0) + P(A_{i-1}) P(B_1) + \sum_{r=2}^{i} P(A_{i-r}) P(B_r)$$

$$= P(A_i) P(B_0) + P(A_{i-1}) [\lambda h + o(h)] + \sum_{r=2}^{i} P(A_{i-r})\, o(h)$$

$$= P(A_i) P(B_0) + \lambda h P(A_{i-1}) + P(A_{i-1})\, o(h) + \sum_{r=2}^{i} P(A_{i-r})\, o(h)$$

$$= P(A_i) P(B_0) + \lambda h P(A_{i-1}) + \sum_{r=1}^{i} P(A_{i-r})\, o(h)$$

$$= P(A_i) P(B_0) + \lambda h P(A_{i-1}) + \sum_{r=1}^{i} P(A_{i-r})\, o(h)$$

$$= P(A_i) P(B_0) + \lambda h P(A_{i-1}) + o(h) \sum_{r=1}^{i} P(A_{i-r})$$

$$P(B_0) = 1 - P(\overline{B_0})$$
$$= 1 - P(B_1 + B_2 + \cdots + B_i)$$
$$= 1 - [P(B_1) + P(B_2) + \cdots + P(B_i)]$$
$$= 1 - [\lambda h + o(h) + o(h) + \cdots + o(h)]$$
$$= 1 - [\lambda h + i\, o(h)] = 1 - \lambda h - i\, o(h)$$

$$\therefore P(X(t+h) = i) = P(A_i)\{1 - \lambda h - i\, o(h)\} + \lambda h\, P(A_{i-1}) + o(h) \sum_{r=1}^{i} P(A_{i-r}) \quad \text{——— (1)}$$

Let $P(X(t) = i) = P_i(t)$. Therefore, from (1), we get

$$P(X(t+h) = i) = P(X(t) = i)\{1 - \lambda h - i\, o(h)\} + \lambda h\, P(X(t) = i-1) + o(h) \sum_{r=1}^{i} P(A_{i-r})$$

$$\Rightarrow P_i(t+h) = P_i(t)\{1 - \lambda h - i\, O(h)\} + \lambda h\, P_{i-1}(t) + O(h) \sum_{r=1}^{i} P(A_{i-r})$$

$$\Rightarrow \frac{P_i(t+h) - P_i(t)}{h} = -\lambda P_i(t) + \lambda P_{i-1}(t) - i\, P_i(t) \frac{o(h)}{h} + \frac{o(h)}{h} \sum_{r=1}^{i} P(A_{i-r})$$

Making $h \to 0$, we get

$$P_i'(t) = -\lambda P_i(t) + \lambda P_{i-1}(t) \quad \text{——— (2)}$$

Now $(X(t+h) = 0) = A_0 B_0$

$$\Rightarrow P(X(t+h) = 0) = P(A_0 B_0) = P(A_0)P(B_0)$$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

$\Rightarrow P_0(t+h) = P_0(t)\{1-\lambda h - i\,o(h)\}$

$\Rightarrow \dfrac{P_0(t+h)-P_0(t)}{h} = -\lambda P_0(t) - i\,P_0(t)\dfrac{o(h)}{h}$

Making $h \longrightarrow 0$, we get

$P_0'(t) = -\lambda P_0(t)$ ——— ③

Again $P_0(0) = 1$ ——— ④

$P_i(0) = 0 \quad \forall i = 1, 2, \ldots$ ——— ⑤

The general solution of ③ can be written as

$P_0(t) = Ae^{-\lambda t}$ ($A \rightarrow$ integration constant)

Using ④, we get $P_0(0) = Ae^{-\lambda \cdot 0} = A = 1$ [as $P_0(0)=1$]

$\therefore P_0(t) = e^{-\lambda t}$

Set $P_i(t) = e^{-\lambda t} Q_i(t)$ ——— ⑥ $i = 0, 1, 2, \ldots$

$\Rightarrow P_i'(t) = -\lambda e^{-\lambda t} Q_i(t) + e^{-\lambda t} Q_i'(t)$

$= -\lambda e^{-\lambda t} Q_i(t) + \lambda e^{-\lambda t} Q_{i-1}(t)$ [from ②]
$(i = 1, 2, \ldots)$

$\Rightarrow Q_i'(t) = \lambda Q_{i-1}(t)$ ——— ⑦ $i = 1, 2, \ldots$

Putting $t=0$ in ⑥, we get

$P_i(0) = Q_i(0) \Rightarrow Q_i(0) = 0 \quad \forall i = 1, 2, \ldots$ [as $P_i(0)=0$]

Putting $i=1$ in ⑦,

$Q_1'(t) = \lambda Q_0(t) = \lambda e^{\lambda t} P_0(t) = \lambda e^{\lambda t} e^{-\lambda t} = \lambda$

$\Rightarrow Q_1(t) = \lambda t + Q_1(0) = \lambda t$ [as $Q_1(0) = 0$]

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

Putting $i=2$ in (7), we get

$$Q_2'(t) = \lambda Q_1(t) = \lambda \lambda t = \lambda^2 t$$

$$\Rightarrow Q_2(t) = \frac{1}{2}\lambda^2 t^2 + Q_2(0) = \frac{1}{2}\lambda^2 t^2 = \frac{(\lambda t)^2}{2!}$$

Putting $i=3$ in (7), we get

$$Q_3'(t) = \lambda Q_2(t) = \frac{1}{2}\lambda^3 t^2$$

$$\Rightarrow Q_3(t) = \frac{1}{3!}(\lambda t)^3$$

By mathematical induction, we get

$$Q_n(t) = \frac{1}{n!}(\lambda t)^n$$

$$P_i(t) = e^{-\lambda t} Q_i(t) = e^{-\lambda t} \cdot \frac{(\lambda t)^i}{i!}$$

$$= \frac{e^{-\mu}\mu^i}{i!}, \quad \mu = \lambda t$$

$$P_i(t) = \frac{e^{-\mu}\mu^i}{i!}, \quad \mu = \lambda t$$

Example of Poisson Process:

(1) No of road accidents in a particular place in a definite time interval

(2) No of telephone calls in a given interval of time.

Geometric Distribution:

Consider an infinite sequence of Bernoulli trials. Let X be the random variable which denotes the number of failures preceding the first success with probability of success p, $0<p<1$ and probability of failure $q=1-p$.

Therefore the spectrum of $X = \{0, 1, 2, \ldots\}$

$$\underbrace{\boxed{f}\,\boxed{f}\,\boxed{f}\;\cdots\;\boxed{f}}_{i \to \text{failures}}\,\boxed{s} \longrightarrow \text{Probability} = P(\underbrace{f\,f\cdots f}_{i}\,s)$$

$$= P(f)P(f)\cdots P(f)P(s) = [P(f)]^i P(s)$$

$$P(X=i) = q^i p \quad \forall\, i = 0, 1, 2, \ldots, \quad q = 1-p, \; 0<p<1$$

Negative Binomial Distribution:

Consider an infinite sequence of Bernoulli trials. Let X be the random variable denoting the number of failures that precede the rth success, where $r \geq 1$ is a fixed positive integer.

Therefore the spectrum of $X = \{0, 1, 2, \ldots\}$. Let $X=x$.

$$\underbrace{\boxed{\;}\,\boxed{\;}\,\boxed{\;}\,\boxed{\;}\;\cdots\;\boxed{\;}\,\boxed{\;}\,\boxed{\;}}_{x+r-1 \to \text{rooms}}\,\underset{r\text{th success}}{\boxed{s}}$$

- x rooms are filled by f
- $r-1$ rooms are filled by s

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

Now x rooms can be chosen from $x+r-1$ rooms in

$$\binom{x+r-1}{x} = \binom{x+r-1}{r-1} \text{ ways}$$

$$\therefore\ P(X=x) = \binom{x+r-1}{x} (P(f))^x (P(s))^r, \quad x = 0, 1, 2, \ldots$$

$$\Rightarrow P(X=x) = \binom{x+r-1}{x} (1-p)^x p^r, \quad x = 0, 1, 2, \ldots \qquad (1)$$

Geometric Distribution is a particular case of a Negative Binomial distribution:

For $r=1$, (1) assumes the following form:

$$P(X=x) = \binom{x}{x} (1-p)^x p, \quad x = 0, 1, 2, \ldots$$

$$= (1-p)^x p = q^x p, \quad x = 0, 1, 2, \ldots$$

This is the probability mass function of the Geometric Distribution.

Therefore from the Bernoulli trials, we have the following Distributions:

(1) Binomial Distribution $:= B(n, p)$

$$f(x) = P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n$$

$$= 0 \quad \text{elsewhere}$$

② Poisson Distribution $:= P(\mu)$

$$f(x) = P(X=x) = \frac{e^{-\mu}\mu^x}{x!}, \quad x = 0, 1, 2, \ldots$$
$$= 0 \quad \text{elsewhere}$$

③ Geometric Distribution $:= G(p)$

$$f(x) = P(X=x) = p(1-p)^x, \quad x = 0, 1, 2, \ldots$$
$$= 0 \quad \text{elsewhere}$$

④ Negative Binomial Distribution $:= NB(r, p)$

$$f(x) = P(X=x) = \binom{x+r-1}{x}(1-p)^x p^r, \quad x = 0, 1, 2, \ldots$$
$$= 0 \quad \text{elsewhere}$$

Show that

ⓐ For Binomial (n, p) distribution

$$\sum_{x=0}^{n} P(X=x) = 1$$

ⓑ For Poisson (μ) distribution

$$\sum_{x=0}^{\infty} P(X=x) = 1$$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

(c) For Geometric $G(p)$ distribution

$$\sum_{x=0}^{\infty} p(1-p)^{x} = 1$$

(d) For Negative Binomial $NB(r, p)$ distribution

$$\sum_{x=0}^{\infty} P(X=x) = 1$$

Some continuous distributions

(1) Rectangular or uniform distribution:

The pdf of this distribution is given by

$$f(x) = k \quad \text{for } a < x < b$$
$$= 0 \quad \text{otherwise}$$

where k is a constant.

As $f(x)$ is a pdf,

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\Rightarrow \underbrace{\int_{-\infty}^{a} f(x)\,dx}_{0} + \int_{a}^{b} f(x)\,dx + \underbrace{\int_{b}^{\infty} f(x)\,dx}_{0} = 1$$

$$\Rightarrow \int_{a}^{b} f(x)\,dx = 1 \Rightarrow \int_{a}^{b} k\,dx = 1 \Rightarrow k(b-a) = 1$$

$$\Rightarrow k = \frac{1}{b-a}$$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

② Normal distribution

The pdf of this distribution is given by

$$f(x) = A e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}, \quad -\infty < x < \infty$$

where A, m, σ are constants with $\sigma > 0$

As $f(x)$ is a pdf, we have

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\Rightarrow \int_{-\infty}^{\infty} A e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2} dx = 1$$

$$\Rightarrow \int_{-\infty}^{\infty} A e^{-\frac{1}{2}y^2} \sigma\, dy = 1$$

Put $y = \frac{x-m}{\sigma}$
$\Downarrow$
$dy = \frac{1}{\sigma} dx$
$\Downarrow$
$dx = \sigma\, dy$

$$\Rightarrow A\sigma \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2} dy = 1$$

$$\Rightarrow 2A\sigma \int_{0}^{\infty} e^{-\frac{1}{2}y^2} dy = 1$$

Put $\frac{1}{2}y^2 = p$
$\Downarrow$
$y^2 = 2p$
$\Downarrow$
$y = \sqrt{2p}$
$\Downarrow$
$dy = \sqrt{2} \cdot \frac{1}{2} p^{\frac{1}{2}-1} dp$
$= \frac{1}{\sqrt{2}\sqrt{p}} dp$

$$\Rightarrow 2A\sigma \int_{0}^{\infty} e^{-p} \cdot \frac{1}{\sqrt{2}\sqrt{p}} dp = 1$$

$$\Rightarrow \sqrt{2} A\sigma \int_{0}^{\infty} e^{-p} p^{\frac{1}{2}-1} dp = 1$$

$$\Rightarrow \sqrt{2} A\sigma\, \Gamma\left(\tfrac{1}{2}\right) = 1 \Rightarrow \sqrt{2} A\sigma \sqrt{\pi} = 1$$

$$\Rightarrow A = \frac{1}{\sqrt{2\pi}\,\sigma}$$

Therefore for Normal distribution, pdf is given by

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

where m and σ are two real constants with $\sigma > 0$ and these two constants are known as the parameters of the distribution.

The random variable corresponding to the normal distribution is known as normal variate and this is generally denoted by $N(m, \sigma)$.

$N(0,1)$ is said to be Standard Normal variate.

③ Cauchy distribution:

The pdf for this distribution is given by

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

where $\lambda > 0$ and μ are two parameters of the distribution.

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

4 Gamma distribution:

The pdf of this distribution is given by

$$f(x) = \frac{e^{x} x^{l-1}}{\Gamma(l)} \quad \text{for } 0 < x < \infty$$

$$= 0 \qquad \text{otherwise}$$

where $l > 0$ is the only parameter of the distribution and this distribution is known as Gamma distribution.

The random variable corresponding this distribution is known as $\Gamma(l)$-variate.

5 β-distribution of 1st kind:

The pdf for this distribution is given by

$$f(x) = \frac{x^{l-1}(1-x)^{m-1}}{\beta(l, m)}, \quad 0 < x < 1$$

$$= 0 \qquad \text{otherwise}$$

where $l > 0$ and $m > 0$ are two parameters of this distribution.

⑥ β- distribution of 2nd kind:

The pdf of this distribution is given by

$$f(x) = \frac{x^{l-1}}{\beta(l,m)(1+x)^{l+m}} \quad , \quad 0<x<\infty$$

$$= 0 \quad , \quad \text{Otherwise}$$

where $l>0$ and $m>0$ are the parameters of this distribution.

⑦ χ^2- distribution:

The pdf of this distribution is given by

$$f(x) = \frac{e^{-\frac{x}{2}} \left(\frac{x}{2}\right)^{\frac{n}{2}-1}}{2\,\Gamma\left(\frac{n}{2}\right)} \quad , \quad x>0$$

$$= 0 \quad \text{Otherwise}$$

where n is a positive integer and this is the only parameter of the distribution. This parameter of this distribution is known as the number of degrees of freedom.

A χ^2- distribution with n degrees of freedom is generally known as $\chi^2(n)$ distribution.

Sometimes, the variable x in χ^2-distribution is replaced by χ^2 and consequently pdf can be written as

$$f(\chi^2) = \frac{e^{-\frac{1}{2}\chi^2}\left(\frac{1}{2}\chi^2\right)^{\frac{n}{2}-1}}{2\Gamma(\frac{n}{2})} \quad \text{for } \chi^2 > 0$$

$$= 0 \quad \text{for } \chi^2 < 0$$

(8) t-distribution or Student's distribution:

The pdf of this distribution is given by

$$f(t) = \frac{1}{\sqrt{n}\, B(\frac{1}{2}, \frac{n}{2}) \left(1+\frac{t^2}{n}\right)^{(n+1)/2}}, \quad -\infty < t < \infty.$$

Here n is a positive integer and this positive integer is known as the degrees of freedom of the distribution.

(9) F-distribution:

The pdf of the distribution

$$f(F) = \frac{m^{m/2}\, n^{n/2}\, F^{\frac{m}{2}-1}}{B(\frac{m}{2}, \frac{n}{2})\,(mF+n)^{(m+n)/2}} \quad \text{for } F > 0$$

$$= 0 \quad \text{for } F < 0$$

Here m, n both are positive integers and these positive integers are known as the parameters of the distribution. This distribution is referred as $F(m,n)$.

Problem 1: If X is uniformly distributed in $(-1, 1)$, then find the pdf of $|X|$

Answer: The pdf of X is given by

$$f(x) = \frac{1}{2} \quad \text{for } -1 < x < 1$$
$$= 0 \quad \text{otherwise}$$

Let $Y = |X|$

$$y = |x| = \begin{cases} x & \text{for } x > 0 \\ 0 & \text{for } x = 0 \\ -x & \text{for } x < 0 \end{cases}$$

$y = |x|$ is not differentiable at $x = 0$ and consequently we cannot apply the theorem for transformation.

Now $-1 < x < 1 \Rightarrow 0 \leq |x| < 1 \Rightarrow 0 \leq y < 1$

case-1

For $y < 0$,

$$F_Y(y) = P(Y \leq y) = P(|X| \leq y < 0)$$
$$= P(|X| < 0) = P(\Phi) = 0$$

$\therefore F_Y(y) = 0$ for $y < 0$ ——— ①

case 2

For $y = 0$

$$F_Y(y) = P(Y \leq y) = P(Y \leq 0) = P(Y < 0) + P(Y = 0)$$
$$= P(|X| < 0) + P(|X| = 0)$$
$$= P(\Phi) + P(X = 0) = P(\Phi) + 0 = 0 + 0 = 0$$

$\therefore \; F_Y(y) = 0 \quad \text{for} \quad y = 0$ ——————— (2)

<u>Case 3</u>: $y > 0$

<u>Subcase 3A</u>: $0 < y < 1$

$$F_Y(y) = P(Y \le y) = P\{(Y<0) \cup (Y=0) \cup (0<Y\le y)\}$$

$$= P(Y<0) + P(Y=0) + P(0<Y\le y)$$

$$= P(|X|<0) + P(|X|=0) + P(0<|X|\le y)$$

$$= P(\Phi) + P(X=0) + P(|X|\le y)$$

$$= 0 + 0 + P(-y \le X \le y)$$

$$= P(-y \le X \le y)$$

$$= \int_{-y}^{y} f(x)\,dx = \int_{-y}^{y} \frac{1}{2}\,dx = y$$

$F_Y(y) = y \quad \text{for} \quad 0 < y < 1$ ——————— (3)

<u>subcase 3B</u>: $y = 1$

$$F_Y(y) = P(Y \le y) = P(Y \le 1) = P(|X| \le 1)$$

$$= P(-1 \le X \le 1) = \int_{-1}^{1} f(x)\,dx = \int_{-1}^{1} \frac{1}{2}\,dx$$

$$= 1$$

$\therefore \; F_Y(y) = 1 \quad \text{for} \quad y = 1$ ——————— (4)

<u>subcase 3c:</u> $y > 1$

$$F_Y(y) = P(Y \le y) = P(|X| \le y)$$

$$= P\{(|X| \le 1) \cup (1 < |X| \le y)\}$$

$$= P(|X| \le 1) + P(1 < |X| \le y)$$

$$= P(|X| \le 1) + P(1 < Y \le y)$$

$$= P(|X| \le 1) + P(\Phi) = P(|X| \le 1) + 0$$

$$= \int_{-1}^{1} f(x)\,dx = \int_{-1}^{1} \tfrac{1}{2}\,dx = 1$$

$\therefore F_Y(y) = 1$ for $y > 1$ ——— (5)

Combining (1), (2), (3), (4) and (5), we get

$$F_Y(y) = 0 \quad \text{for } y < 0$$
$$= 0 \quad \text{for } y = 0$$
$$= y \quad \text{for } 0 < y < 1$$
$$= 1 \quad \text{for } y = 1$$
$$= 1 \quad \text{for } y > 1$$

$$\Rightarrow F_Y(y) = 0 \quad \text{for } y \le 0$$
$$= y \quad \text{for } 0 < y < 1$$
$$= 1 \quad \text{for } y \ge 1$$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

$\Rightarrow F_Y'(y) = 1$ for $0<y<1$

$= 0$ otherwise

$\Rightarrow f_Y(y) = 1$ for $0<y<1$

$= 0$ otherwise

$\Rightarrow Y = |X|$ is uniformly distributed in $(0,1)$

Problem 2: The pdf of X is given by

$$f_X(x) = kx^2, \quad -3<x<6$$
$$= 0 \quad \text{otherwise}$$

Find the pdf of $U = \frac{1}{3}(12-X)$.

Solution: $u = \frac{1}{3}(12-x) \Rightarrow 3u = 12-x$

$\Rightarrow x = 12-3u = 3(4-u)$

$\Rightarrow x = 3(4-u)$ ———— (1)

Now $-3<x<6 \Rightarrow -3<3(4-u)<6$

$\Rightarrow -1<4-u<2 \Rightarrow 1>u-4>-2$

$\Rightarrow -2<u-4<1 \Rightarrow 4-2<u<4+1$

$\Rightarrow 2<u<5$

$\therefore -3<x<6 \Rightarrow 2<u<5$

Again $u = \frac{1}{3}(12-x) \Rightarrow \frac{du}{dx} = -\frac{1}{3} < 0 \quad \forall x$

Therefore we can apply the transformation rule

$$f_U(u) = f_X(x)\left|\frac{dx}{du}\right|$$

$$= \begin{cases} kx^2\left|\frac{dx}{du}\right| & \text{for } -3<x<6 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} k\{3(4-u)\}^2 \cdot 3 & \text{for } 2<u<5 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} k \cdot 27(4-u)^2 & \text{for } 2<u<5 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} 27k(4-u)^2 & \text{for } 2<u<5 \\ 0 & \text{otherwise} \end{cases}$$

To find the value of k, we have

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \Rightarrow \int_{-\infty}^{-3} f(x)\,dx + \int_{-3}^{6} f(x)\,dx + \int_{6}^{\infty} f(x)\,dx = 1$$

$$\Rightarrow \int_{-3}^{6} f(x)\,dx = 1 \Rightarrow \int_{-3}^{6} kx^2\,dx = 1$$

$$\Rightarrow k\left.\frac{x^3}{3}\right|_{-3}^{6} = 1 \Rightarrow \frac{k}{3}(6^3+3^3) = 1 \Rightarrow k = \frac{1}{81}$$

$$\therefore f_U(u) = \begin{cases} \frac{1}{3}(4-u)^2 & \text{for } 2<u<5 \\ 0 & \text{otherwise} \end{cases}$$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

Problem-3: If X is a standard normal variate then prove that $Y = \frac{X^2}{2}$ is a $\Gamma(\frac{1}{2})$ variate.

Solution: $X = N(0,1) \Rightarrow$ the pdf of X is given by

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

Now $y = \frac{x^2}{2} \Rightarrow \frac{dy}{dx} = x \Rightarrow \frac{dy}{dx} > 0$ for $x > 0$

and $\frac{dy}{dx} < 0$ for $x < 0$

Therefore, we cannot apply the tranformation rule.

$$Y = \frac{X^2}{2} \Rightarrow 0 \leq Y < \infty$$

Case 1: For $y < 0$,

$$F_Y(y) = P(Y \leq y) = P(Y \leq y < 0) = P(Y < 0)$$
$$= P\left(\frac{X^2}{2} < 0\right) = P(X^2 < 0) = P(\Phi) = 0$$

$\therefore F_Y(y) = 0$ for $y < 0$ ——— (1)

Case 2: For $y = 0$

$$F_Y(y) = P(Y \leq y) = P(Y \leq 0) = P\{(Y<0) \cup (Y=0)\}$$
$$= P(Y<0) + P(Y=0) = P\left(\frac{X^2}{2} < 0\right) + P\left(\frac{X^2}{2} = 0\right)$$
$$= P(X^2 < 0) + P(X^2 = 0) = P(\Phi) + P(X = 0)$$
$$= 0 + 0 = 0$$

$\therefore\ F_Y(y) = 0 \quad \text{for} \quad y = 0$ ——— (2)

Case 3: $y > 0$

$$F_Y(y) = P(Y \le y) = P\left(\frac{X^2}{2} \le y\right) = P(X^2 \le 2y)$$

$$= P\left(|X|^2 \le (\sqrt{2y})^2\right) = P\left((|X| + \sqrt{2y})(|X| - \sqrt{2y}) \le 0\right)$$

$$= P\left(\underbrace{(|X| + \sqrt{2y})}(|X| - \sqrt{2y}) \le 0\right)$$

$$= P(|X| - \sqrt{2y} \le 0)$$

$$= P(|X| \le \sqrt{2y}) = P(-\sqrt{2y} < X < \sqrt{2y})$$

$$= \int_{-\sqrt{2y}}^{\sqrt{2y}} f_X(x)\,dx = \int_{-\sqrt{2y}}^{\sqrt{2y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx$$

$$= \frac{1}{\sqrt{2\pi}} \cdot 2 \int_0^{\sqrt{2y}} e^{-\frac{x^2}{2}}\,dx$$

Put $z = \frac{x^2}{2}$

$\Rightarrow x = \sqrt{2z}$

$\Rightarrow dx = \sqrt{2} \cdot \frac{1}{2} \cdot \frac{1}{\sqrt{z}}\,dz$

$\Rightarrow dx = \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{z}}\,dz$

when $x = 0$, $z = 0$

when $x = \sqrt{2y}$, $z = y$

$$= \sqrt{\frac{2}{\pi}} \int_0^y e^{-z} \cdot \frac{dz}{\sqrt{2} \cdot \sqrt{z}}$$

$$= \frac{1}{\sqrt{\pi}} \int_0^y e^{-z} z^{\frac{1}{2}-1}\,dz$$

$$\therefore\ F_Y(y) = \frac{1}{\sqrt{\pi}} \int_0^y e^{-z} z^{\frac{1}{2}-1}\,dz \quad \text{——— (3)}$$

Combining (1), (2) and (3), we get

$$F_Y(y) = \begin{cases} 0 & \text{for } y < 0 \\ 0 & \text{for } y = 0 \\ \frac{1}{\sqrt{\pi}} \int_0^y e^{-z} z^{\frac{1}{2}-1}\, dz & \text{for } y > 0 \end{cases}$$

$$= \begin{cases} 0 & \text{for } y \leq 0 \\ \frac{1}{\sqrt{\pi}} \int_0^y e^{-z} z^{\frac{1}{2}-1}\, dz & \text{for } y > 0 \end{cases}$$

$$\therefore \; f_Y(y) = \frac{1}{\sqrt{\pi}} e^{-y} y^{\frac{1}{2}-1} \quad \text{for } y > 0$$

$$= 0 \quad \text{otherwise}$$

or $$f_Y(y) = \frac{e^{-y} y^{\frac{1}{2}-1}}{\Gamma(\frac{1}{2})} \quad \text{for } y > 0$$

$$= 0 \quad \text{otherwise}$$

which shows that Y is $\Gamma(\frac{1}{2})$ variate.

<u>Problem-4</u>: Let X be a continuous random variable with pdf $f(x)$. Let $Y = aX^2$ with $a > 0$ and $g(y)$ be the pdf of Y. Show that

$$g(y) = \begin{cases} \frac{1}{2\sqrt{a}\sqrt{y}} \left\{ f(\sqrt{y/a}) + f(-\sqrt{y/a}) \right\}, & y > 0 \\ 0, & \text{otherwise} \end{cases}$$

$F_Y(y) = P(Y \le y)$

we have the following three cases depending on the values of y.

<u>Case 1:</u> $y < 0$ and for this case, we have

$$F_Y(y) = P(Y \le y) = P(aX^2 \le y < 0) = P(aX^2 < 0)$$
$$= P(X^2 < 0) \text{ (as } a > 0)$$
$$= P(\Phi) = 0$$

$\therefore F_Y(y) = 0$ for $y < 0$ ——————— ①

<u>Case 2:</u> $y = 0$ and for this case we have

$$F_Y(y) = P(Y \le y) = P(aX^2 \le 0) = P(X^2 \le 0) \text{ (as } a > 0)$$
$$= P\{(X^2 < 0) \cup (X^2 = 0)\}$$
$$= P(X^2 < 0) + P(X^2 = 0)$$
$$= P(\Phi) + P(X = 0)$$ (Probability mass at a point is equal to zero for continuous random variable)
$$= 0 + 0 = 0$$

$\therefore F_Y(y) = 0$ for $y = 0$ ——————— ②

<u>Case 3:</u> $y > 0$ and for this case we have

$$F_Y(y) = P(Y \le y) = P(aX^2 \le y) = P\left(X^2 \le \frac{y}{a}\right) \text{ (as } a > 0)$$
$$= P\left(|X|^2 \le \left(\sqrt{\frac{y}{a}}\right)^2\right) = P\left\{\underbrace{\left(|X| + \sqrt{\frac{y}{a}}\right)}\left(|X| - \sqrt{\frac{y}{a}}\right) \le 0\right\}$$
$$= P\left(|X| - \sqrt{\frac{y}{a}} \le 0\right) = P\left(|X| \le \sqrt{\frac{y}{a}}\right)$$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

$$F_Y(y) = P\left(|X| \le \sqrt{\frac{y}{a}}\right) = P\left(-\sqrt{\frac{y}{a}} \le X \le \sqrt{\frac{y}{a}}\right)$$

$$= \int_{-\sqrt{\frac{y}{a}}}^{\sqrt{y/a}} f(x)\,dx = \int_0^{\sqrt{y/a}} f(x)\,dx + \int_{-\sqrt{y/a}}^{0} f(x)\,dx$$

$$= \int_0^{\sqrt{y/a}} f(x)\,dx - \int_0^{-\sqrt{y/a}} f(x)\,dx$$

Put $x = -z$ in the 2nd integral.

$dx = -dz$

when $x = 0$, then $z = 0$

when $x = -\sqrt{y/a}$ then $z = \sqrt{y/a}$

$$= \int_0^{\sqrt{y/a}} f(x)\,dx - \int_0^{\sqrt{y/a}} f(-z)\,(-dz)$$

$$= \int_0^{\sqrt{y/a}} f(x)\,dx + \int_0^{\sqrt{y/a}} f(-z)\,dz$$

$$= \int_0^{\sqrt{y/a}} f(x)\,dx + \int_0^{\sqrt{y/a}} f(-x)\,dx$$

Put $ax^2 = p$ in both the integrals.

$ax^2 = p \Rightarrow x = \sqrt{p/a}$

$\Rightarrow x = \frac{1}{\sqrt{a}} \cdot \sqrt{p}$

$\Rightarrow dx = \frac{1}{\sqrt{a}} \cdot \frac{1}{2} \cdot p^{\frac{1}{2}-1}\,dp$

when $x = 0$, $p = 0$

when $x = \sqrt{\frac{y}{a}}$, $p = a \cdot \frac{y}{a} = y$

$$= \int_0^{y} f\left(\sqrt{\frac{p}{a}}\right) \cdot \frac{1}{2\sqrt{a}}\, p^{\frac{1}{2}-1}\,dp + \int_0^{y} f\left(-\sqrt{\frac{p}{a}}\right) \cdot \frac{1}{2\sqrt{a}}\, p^{\frac{1}{2}-1}\,dp$$

$$= \int_0^{y} \frac{1}{2\sqrt{a}}\, p^{\frac{1}{2}-1} \left\{ f\left(\sqrt{\frac{p}{a}}\right) + f\left(-\sqrt{\frac{p}{a}}\right) \right\} dp \quad \text{——— } (3)$$

Combining (1), (2) and (3), we get

$$F_Y(y) = \begin{cases} \int_0^{y} \frac{1}{2\sqrt{a}}\, p^{\frac{1}{2}-1} \left\{ f\left(\sqrt{p/a}\right) + f\left(-\sqrt{p/a}\right) \right\} dp, & y > 0 \\ 0 & \text{otherwise} \end{cases}$$

As $g(y) = F_Y'(y)$, we have

$$g(y) = \begin{cases} \frac{1}{2\sqrt{a}} y^{\frac{1}{2}-1}\{f(\sqrt{Y/a}) + f(-\sqrt{Y/a})\}, & y > 0 \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{1}{2\sqrt{a}\sqrt{y}}\{f(\sqrt{Y/a}) + f(-\sqrt{Y/a})\}, & y > 0 \\ 0, & \text{otherwise} \end{cases}$$

<u>Problem-5</u>: Let $F(x)$ be a distribution function. Then show that $G(x)$ is also a distribution function for any $h \neq 0$, where

$$G(x) = \frac{1}{2h}\int_{x-h}^{x+h} F(t)\,dt$$

<u>Answer</u>: To prove that $G(x)$ is a probability distribution function, we shall prove the following properties of $G(x)$.

① G is monotonically non-decreasing

② $0 \leq G(x) \leq 1$

③ $G(\infty) = 1$

④ $G(-\infty) = 0$

To prove the above mentioned results of $G(x)$, we take $h > 0$.

$$G(x) = \frac{1}{2h}\int_{x-h}^{x+h} F(t)\,dt$$

Put $x+p=t$
$\Rightarrow dp = dt$
when $t = x-h$, $p=-h$
when $t = x+h$, $p = h$

$$= \frac{1}{2h}\int_{-h}^{h} F(x+p)\,dp$$

$$= \frac{1}{2h}\int_{-h}^{h} F(x+t)\,dt$$

① Let $x_1 \geq x_2$.

Now $x_1 \geq x_2 \Rightarrow x_1 + t \geq x_2 + t$

$\Rightarrow F(x_1+t) \geq F(x_2+t)$ [As F is monotonically non-decreasing]

$$\Rightarrow \int_{-h}^{h} F(x_1+t)\,dt \geq \int_{-h}^{h} F(x_2+t)\,dt$$

$$\Rightarrow \frac{1}{2h}\int_{-h}^{h} F(x_1+t)\,dt \geq \frac{1}{2h}\int_{-h}^{h} F(x_2+t)\,dt$$

$$\Rightarrow G(x_1) \geq G(x_2)$$

$\Rightarrow G$ is monotonically non-decreasing ——— Ⓐ

② As $F(x)$ is a probability distribution function, we have

$$0 \leq F(x) \leq 1 \quad \forall x$$

$$\Rightarrow 0 \leq F(x+t) \leq 1 \quad \forall x \text{ and } \forall t$$

$$\Rightarrow \int_{-h}^{h} 0 \cdot dt \leq \int_{-h}^{h} F(x+t)\,dt \leq \int_{-h}^{h} 1 \cdot dt$$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

$$\Rightarrow \int_{-h}^{h} 0 \cdot dt \leq \int_{-h}^{h} F(x+t)\,dt \leq \int_{-h}^{h} 1 \cdot dt$$

$$\Rightarrow 0 \leq \int_{-h}^{h} F(x+t)\,dt \leq 2h$$

$$\Rightarrow 0 \leq \frac{1}{2h}\int_{-h}^{h} f(x+t)\,dt \leq 1$$

$$\Rightarrow 0 \leq G(x) \leq 1 \text{ ———————— } (B)$$

(2) For $-h \leq t \leq h$, we have

$$-h \leq t \leq h \Rightarrow x-h \leq x+t \leq x+h$$

$$\Rightarrow F(x-h) \leq F(x+t) \leq F(x+h)$$ [As F is monotonically non-decreasing]

$$\Rightarrow \int_{-h}^{h} F(x-h)\,dt \leq \int_{-h}^{h} F(x+t)\,dt \leq \int_{-h}^{h} F(x+h)\,dt$$

$$\Rightarrow 2h\,F(x-h) \leq \int_{-h}^{h} F(x+t)\,dt \leq 2h\,F(x+h)$$

$$\Rightarrow F(x-h) \leq \frac{1}{2h}\int_{-h}^{h} F(x+t)\,dt \leq F(x+h)$$

$$\Rightarrow F(x-h) \leq G(x) \leq F(x+h) \text{ ———————— } (100)$$

Making $x \longrightarrow -\infty$ in (100), we get

$$F(-\infty) \leq G(-\infty) \leq F(-\infty) \Rightarrow G(-\infty) = F(-\infty) = 0$$

$$\Rightarrow G(-\infty) = 0 \text{ ———————— } (C)$$

Making $x \longrightarrow +\infty$ in (100), we get

$$F(\infty) \leq G(\infty) \leq F(\infty) \Rightarrow G(\infty) = F(\infty) = 1$$

$$\Rightarrow G(\infty) = 1 \text{ ———————— } (D)$$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

Combining Ⓐ, Ⓑ, Ⓒ and Ⓓ, we can conclude that $G(x)$ is a probability distribution function for $h>0$

For $h<0$, Let $h_1 = -h > 0$.

$h_1 = -h \Rightarrow h = -h_1$

$$G(x) = \frac{1}{2h}\int_{x-h}^{x+h} F(t)\,dt = \frac{1}{2(-h_1)}\int_{x-(-h_1)}^{x+(-h_1)} F(t)\,dt$$

$$= -\frac{1}{2h_1}\int_{x+h_1}^{x-h_1} F(t)\,dt = \frac{1}{2h_1}\int_{x-h_1}^{x+h_1} F(t)\,dt$$

$\Rightarrow G(x) = \frac{1}{2h_1}\int_{x-h_1}^{x+h_1} F(t)\,dt$ where $h_1 > 0$ and consequently from the earlier part we can conclude that $G(x)$ is a probability distribution function for $h_1 > 0 \Rightarrow -h > 0 \Rightarrow h < 0$

$\therefore$ $G(x)$ is a probability distribution function for all $h \neq 0$

<u>Problem-6:</u> If X is a Poisson variate with prameter μ, then show that

$$P(X \leq n) = \frac{1}{n!}\int_{\mu}^{\infty} e^{-x} x^n\,dx$$

Answer: As X is a Poisson variate with parameter μ,

$$P(X=r) = \frac{e^{-\mu}\mu^r}{r!}, \quad r = 0, 1, 2, \ldots$$

$$P(X \le n) = P(X=0) + P(X=1) + \cdots + P(X=n)$$

$$= \sum_{r=0}^{n} P(X=r)$$

Let $I_r = \frac{1}{r!}\int_{\mu}^{\infty} e^{-x} x^r \, dx$

$$= \frac{1}{r!}\left[x^r \frac{e^{-x}}{(-1)} \Big|_{\mu}^{\infty} - \int_{\mu}^{\infty} r x^{r-1} \frac{e^{-x}}{(-1)} dx \right]$$

$$= \frac{\mu^r e^{-\mu}}{r!} + \frac{1}{(r-1)!}\int_{\mu}^{\infty} x^{r-1} e^{-x} dx$$

$$= \frac{\mu^r e^{-\mu}}{r!} + I_{r-1}$$

$$= P(X=r) + I_{r-1}$$

$$\Rightarrow I_r - I_{r-1} = P(X=r) \quad \text{——————} \quad (1)$$

Putting $r = 1, 2, \ldots, n$ we get the following equations

$$I_1 - I_0 = P(X=1)$$
$$I_2 - I_1 = P(X=2)$$
$$I_3 - I_2 = P(X=3)$$
$$\cdots$$
$$I_n - I_{n-1} = P(X=n)$$

$$I_n - I_0 = P(X=1) + P(X=2) + \cdots + P(X=n)$$

$$\Rightarrow I_n - I_0 = P(X=1) + P(X=2) + \cdots + P(X=n)$$

$$\Rightarrow I_n = I_0 + \sum_{r=1}^{n} P(X=r) \quad \text{——— ②}$$

$$I_0 = \frac{1}{0!}\int_{\mu}^{\infty} e^{-x} x^0 dx = \int_{\mu}^{\infty} e^{-x} dx = \left.\frac{e^{-x}}{(-1)}\right|_{\mu}^{\infty}$$

$$= \frac{e^{\mu}}{1} = \frac{e^{-\mu}\mu^0}{0!} = P(X=0) \quad \text{——— ③}$$

From ② and ③, we get

$$I_n = P(X=0) + \sum_{r=1}^{n} P(X=r) = \sum_{r=0}^{n} P(X=r) = P(X \le n)$$

$$\Rightarrow P(X \le n) = I_n = \frac{1}{n!}\int_{\mu}^{\infty} e^{-x} x^n dx$$

<u>Problem-7:</u> If $X = B(n,p)$, then show that

$$P(X \le k) = \frac{\int_0^q x^{n-k-1}(1-x)^k dx}{\int_0^1 x^{n-k-1}(1-x)^k dx}$$

<u>Answer:</u> As X is a $B(n,p)$,

$$P(X=r) = \binom{n}{r} p^r q^{n-r}, \quad r = 0, 1, 2, \ldots, n$$

where $0 < p < 1$ and $q = 1-p$

Now $P(X \le k) = P(X=0) + P(X=1) + \cdots + P(X=k)$

$$= \sum_{r=0}^{k} P(X=r)$$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

Let $I_r = \dfrac{\int_0^q x^{n-r-1}(1-x)^r\,dx}{\int_0^1 x^{n-r-1}(1-x)^r\,dx}$

Now $\int_0^1 x^{n-r-1}(1-x)^r\,dx = \beta(n-r, r+1)$

$$= \frac{\Gamma(n-r)\Gamma(r+1)}{\Gamma(n+1)} = \frac{\Gamma(n-r)\, r!}{n!}$$

$$= \frac{\Gamma(n-r)}{\frac{n!}{r!\,(n-r)!}(n-r)!} = \frac{\Gamma(n-r)}{(n-r)!}\cdot\frac{1}{\binom{n}{r}}$$

$$= \frac{1}{(n-r)\binom{n}{r}}$$

$$I_r = (n-r)\binom{n}{r}\int_0^q x^{n-r-1}(1-x)^r\,dx \qquad \text{——— (1)}$$

$$= (n-r)\binom{n}{r}\left[(1-x)^r\frac{x^{n-r}}{n-r}\Big|_0^q - \int_0^q r(1-x)^{r-1}(-1)\cdot\frac{x^{n-r}}{n-r}\,dx\right]$$

$$= (n-r)\binom{n}{r}\left[\frac{(1-q)^r q^{n-r}}{n-r} + \frac{r}{n-r}\int_0^q x^{n-r}(1-x)^{r-1}\,dx\right]$$

$$= \binom{n}{r}p^r q^{n-r} + r\binom{n}{r}\int_0^q x^{n-r}(1-x)^{r-1}\,dx$$

$$= P(X=r) + r\,\frac{n!}{(n-r)!\,r!}\int_0^q x^{n-r}(1-x)^{r-1}\,dx$$

$$= P(X=r) + \frac{n!}{(n-r)!\,(r-1)!}\int_0^q x^{n-r}(1-x)^{r-1}\,dx$$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

$$I_r = P(X=r) + (n-r+1)\frac{n!}{(n-r+1)!(r-1)!}\int_0^q x^{n-r}(1-x)^{r-1}dx$$

$$= P(X=r) + (n-\overline{r-1})\binom{n}{r-1}\int_0^q x^{n-r}(1-x)^{r-1}$$

$$= P(X=r) + I_{r-1} \quad [\text{using } ①]$$

$$\Rightarrow I_r - I_{r-1} = P(X=r) \qquad ——— ②$$

Putting $r = 1, 2, \ldots, k$ in ②, we get

$$I_1 - I_0 = P(X=1)$$

$$I_2 - I_1 = P(X=2)$$

$$I_3 - I_2 = P(X=3)$$

$$\ldots$$

$$I_k - I_{k-1} = P(X=k)$$

$$I_k - I_0 = P(X=1) + P(X=2) + \cdots + P(X=k)$$

$$\Rightarrow I_k = I_0 + \sum_{r=1}^{k} P(X=r)$$

$$= n\binom{n}{0}\int_0^q x^{n-1}\,dx + \sum_{r=1}^{k} P(X=r) \quad [\text{using } ①]$$

$$= n\frac{q^n}{n} + \sum_{r=1}^{k} P(X=r) = P(X=0) + \sum_{r=1}^{k} P(X=r)$$

$$= \sum_{r=0}^{k} P(X=r) = P(X \le k) \Rightarrow P(X \le k) = I_k$$

Problem-8: If r is the most probable number of success for $B(n,p)$, then show that

$$(n+1)p-1 \leq r \leq (n+1)p$$

Answer: Let $X = B(n,p)$

$$\therefore \; P(X=r) = \binom{n}{r} p^r (1-p)^{n-r}, \quad r = 0, 1, 2, \ldots, n$$

where $0<p<1$.

Let A_r be the event of exactly r success out of n trials.

$$\therefore \; P(A_r) = P(X=r)$$

Let r is the most probable number of success.

Therefore, we have

$$\text{① } P(A_0) \leq P(A_1) \leq \cdots \leq P(A_{r-1}) \leq P(A_r)$$

$$\text{② } P(A_r) \geq P(A_{r+1}) \geq \cdots \geq P(A_n)$$

$$\frac{P(A_r)}{P(A_{r-1})} \geq 1 \quad \text{———} \quad ①$$

$$\frac{P(A_r)}{P(A_{r+1})} \geq 1 \quad \text{———} \quad ②$$

From ①, we have

$$\frac{P(A_r)}{P(A_{r-1})} = \frac{\binom{n}{r} p^r q^{n-r}}{\binom{n}{r-1} p^{r-1} q^{n-r+1}} \geq 1$$

$$\Rightarrow \frac{n!\,(r-1)!\,(n-r+1)!}{n!\,r!\,(n-r)!} \, \frac{p}{q} \geq 1$$

$$\Rightarrow \frac{n-r+1}{r}\,\frac{p}{1-p} \geq 1$$

$$\Rightarrow (n-r+1)p \geq r(1-p)$$

$$\Rightarrow (n+1)p - rp \geq r - rp$$

$$\Rightarrow (n+1)p \geq r \Rightarrow \boxed{r \leq (n+1)p} \text{ ——— } (3)$$

From (2), we get

$$\frac{P(A_r)}{P(A_{r+1})} = \frac{\binom{n}{r} p^r q^{n-r}}{\binom{n}{r+1} p^{r+1} q^{n-r-1}} \geq 1$$

$$\Rightarrow \frac{(n!\,(r+1)!\,(n-r-1)!)\ q}{(n!\ r!\ (n-r)!)\ \ p} \geq 1$$

$$\Rightarrow \frac{r+1}{n-r}\,\frac{1-p}{p} \geq 1$$

$$\Rightarrow (r+1)(1-p) \geq (n-r)p$$

$$\Rightarrow r+1-rp-p \geq np-rp$$

$$\Rightarrow r+1-p \geq np$$

$$\Rightarrow r \geq np+p-1 = (n+1)p-1$$

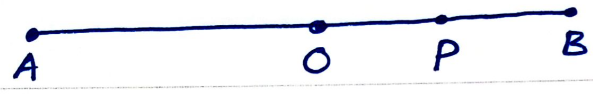
$$\Rightarrow r \geq (n+1)p-1 \text{ ——— } (4)$$

Combining (3) and (4), we get

$$\checkmark\ (n+1)p-1 \leq r \leq (n+1)p$$

Problem-9: Consider a line segment AB and a point P on AB. Let O be the mid-point of AB. Find the probability that AP, PB, BO form the sides of a triangle.

Answer:



Let $AB = 2a$

$\therefore AO = BO = a$

Let $OP = x$

$\therefore AP = a + x, \quad BP = a - x$

Let X be the random variable corresponding to x.

$\therefore X$ has uniform probability distribution on AB because the points on AB are uniformly distributed.

Therefore, the pdf of X is given by

$$f(x) = \begin{cases} \frac{1}{a-(-a)}, & -a < x < a \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{1}{2a}, & -a < x < a \\ 0, & \text{otherwise} \end{cases}$$

Now AP, PB, BO form the sides of a triangle iff

$AP + PB > BO$, $AP + BO > PB$ and $BP + BO > AP$

i.e., iff $a + x + a - x > a$, $a + x + a > a - x$, and

$a - x + a > a + x$

i.e., iff $\underbrace{2a > a}_{\text{holds good for any } a}$, $2a + x > a - x$ and $2a - x > a + x$

i.e., iff $2a + x > a - x$ and $2a - x > a + x$

i.e., iff $2x > -a$ and $a > 2x$

i.e., iff $x > -\frac{a}{2}$ and $\frac{a}{2} > x$

i.e., iff $-\frac{a}{2} < x < \frac{a}{2}$

$\therefore$ P(AP, BP, BO form the sides of a triangle)

$= P\left(-\frac{a}{2} < X < \frac{a}{2}\right)$

$$= \int_{-\frac{a}{2}}^{a/2} f_X(x)\,dx = \int_{-a/2}^{a/2} \frac{1}{2a}\,dx = \frac{1}{2a}\left\{\frac{a}{2} - \left(-\frac{a}{2}\right)\right\} = \frac{1}{2}$$

$= \frac{1}{2}$

<u>Problem-10</u>: The point P is taken at random on a line segment of length $2a$. Find the probability that the area of the rectangle having adjacent sides AP and PB will exceed $\frac{a^2}{2}$.

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

Let $AB = 2a$, $AP = X$

$\therefore$ $BP = 2a - X$

Therefore X has uniform probability distribution on $(0, 2a)$ because the points on AB are uniformly distributed. Consequently the pdf of X is given by

$$f(x) = \begin{cases} \dfrac{1}{2a-0} & , 0 < x < 2a \\ 0 & , \text{otherwise} \end{cases}$$

$$= \begin{cases} \dfrac{1}{2a} & , 0 < x < 2a \\ 0 & , \text{otherwise} \end{cases}$$

$P\left(AP \cdot PB > \frac{a^2}{2}\right)$

$= P\left(X(2a - X) > \frac{a^2}{2}\right)$

$= P\left(2aX - X^2 > \frac{a^2}{2}\right)$

$= P\left(X^2 - 2aX + \frac{a^2}{2} \quad 0\right)$

$= P\left(X^2 - 2aX + a^2 - a^2 + \frac{a^2}{2} < 0\right)$

$= P\left((X-a)^2 - \frac{a^2}{2} < 0\right)$

$= P\left(|X-a|^2 - \left(\frac{a}{\sqrt{2}}\right)^2 < 0\right)$

$= P\left\{\underbrace{\left(|X-a| + \frac{a}{\sqrt{2}}\right)}\left(|X-a| - \frac{a}{\sqrt{2}}\right) < 0\right\}$

$= P\left(|X-a| - \frac{a}{\sqrt{2}} < 0\right)$

Anup Bandyopadhyay, Department of Mathematics, Jadavpur University

$$= P\left(|x-a| < \frac{a}{\sqrt{2}}\right)$$

$$= P\left(-\frac{a}{\sqrt{2}} < x-a < \frac{a}{\sqrt{2}}\right)$$

$$= P\left(a-\frac{a}{\sqrt{2}} < x < a+\frac{a}{\sqrt{2}}\right)$$

$$= \int_{a-\frac{a}{\sqrt{2}}}^{a+\frac{a}{\sqrt{2}}} f(x)\,dx$$

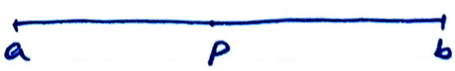
$$= \int_{a-\frac{a}{\sqrt{2}}}^{a+\frac{a}{\sqrt{2}}} \frac{1}{2a}\,dx = \frac{1}{2a}\left(a+\frac{a}{\sqrt{2}}-a+\frac{a}{\sqrt{2}}\right)$$

$$= \frac{1}{2a}\cdot\frac{2a}{\sqrt{2}} = \frac{1}{\sqrt{2}}$$

Problem-11: A point P is chosen at random in a given interval which divides the interval into two sub-intervals. Find the probability that the ratio of the length of the left sub-interval to that of the right sub-interval is less than a constant k.

Answer: Consider a given interval $[a,b]$ with $b>a$.

Let $P \in [a,b)$ and



x be the random variable which measures the length of the lef sub-interval. Therefore x is uniformly distributed over $(0, b-a)$ and consequently the pdf of X is given by

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

$$= \begin{cases} \frac{1}{b-a}, & 0 < x < b-a \\ 0, & \text{elsewhere} \end{cases}$$

Required probability

$$= P\left(\frac{X}{b-a-X} < k\right)$$

$$= P\left(X < k(b-a-X)\right)$$

$$= P\left(X < k(b-a) - kX\right)$$

$$= P\left(X + kX < k(b-a)\right)$$

$$= P\left((1+k)X < k(b-a)\right)$$

$$= P\left(X < \frac{k(b-a)}{1+k}\right)$$

$$= P\left(0 < X < \frac{k(b-a)}{1+k}\right)$$

$$= \int_0^{\frac{k(b-a)}{1+k}} f(x)\,dx$$

$$= \frac{1}{b-a}\left[\frac{k(b-a)}{1+k} - 0\right]$$

$$= \frac{k}{1+k}.$$

Problem-12: A point P is chosen at random on a circle of radius a and a point A is fixed on the circle. Show that the probability that the chord AP will exceed the length of the side of an equilateral triangle inscribed in the circle is $1/3$.

Problem-13: The random variable X is uniformly distributed over $(0, 2)$. Find the distribution function of the larger root of the quadratic equation $t^2 + 2t - x = 0$

Problem-14: A point P is chosen at random on a semicircle of radius unity and it is projected on the diameter. Prove that the distribution of the point of projection from the centre has the pdf

$$\begin{cases} \dfrac{1}{\pi\sqrt{1-x^2}}, & |x| < 1 \\ 0, & \text{otherwise} \end{cases}$$

Problem – 15: Three concentric circles of radii $\frac{1}{\sqrt{3}}$, 1, $\sqrt{3}$ units are drawn on a target board. If a shot falls within the innermost circle 3 points are scored. If it falls within the next two rings, the scores are 2 and 1 respectively and the score is zero if the shot falls outside the uppermost circle. If the probability density function of the hit at a distance r from the centre of the target is $\frac{2}{\pi}\cdot\frac{1}{1+r^2}$, find the probability distribution function of the scores.