

Unit-IV

Order Statistics and Limit Theorems

4.1

Order Statistics

Independent and identically distributed random variables:

We say that $X_1, X_2, \dots, X_n$ are *independent and identically distributed* random variables (i.i.d.r.vs) if

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = \prod_{i=1}^n F_{X_i}(x_i) \quad (\text{independent}) \quad \dots (1)$$

$$\text{and } F_{X_i}(x) = F(x) \quad \forall i = 1, 2, \dots, n \quad (\text{identically distributed}) \quad \dots (2)$$

where $F_{X_i}(x)$ is the c.d.f. of X_i for $i = 1, 2, \dots, n$ and $F_{X_1, \dots, X_n}(x_1, \dots, x_n)$ is the j.c.d.f. of $X_1, \dots, X_n$.

For continuous random variables, the c.d.fs are replaced with p.d.fs in equations (1) and (2) while for discrete random variables the c.d.fs are replaced with p.m.fs.

Definition: We say that $X_1, X_2, \dots, X_n$ is a random sample from a population with c.d.f. $F(x)$ (or p.d.f. $f(x)$ or p.m.f. $p(x)$) if $X_1, \dots, X_n$ are i.i.d.r.vs with common c.d.f. $F(x)$ (or p.d.f. $f(x)$ or p.m.f. $p(x)$).

Definition: Let $X_1, X_2, \dots, X_n$ be a random sample from a population with c.d.f. $F(x)$. Define

$$X_{(1)} = \text{smallest of } X_1, X_2, \dots, X_n$$

$$X_{(2)} = \text{second smallest of } X_1, X_2, \dots, X_n,$$

.

$$X_{(r)} = r^{\text{th}} \text{ smallest of } X_1, X_2, \dots, X_n,$$

.

$$X_{(n)} = \text{largest of } X_1, X_2, \dots, X_n.$$

The ordered values $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ are known as the **order statistics** (o.s) of the n r.vs $X_1, X_2, \dots, X_n$.

Note:

1. o.s are r.vs themselves (as functions of $X_1, \dots, X_n$)
2. o.s. satisfy $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$
3. $X_1, X_2, \dots, X_n$ are i.i.d.r.vs but $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ are neither independent nor identically distributed because of order restriction.

Distributions of o. s. in continuous case:

Let $X_1, X_2, \dots, X_n$ be a random sample from a continuous population with c.d.f. $F(x)$ and p.d.f. $f(x)$.

Marginal distributions:

- 1) The c.d.f. and p.d.f. of X_n , the n^{th} o.s. are given by

$$F_{X_{(n)}}(x) = [F(x)]^n \text{ and } f_{X_{(n)}}(x) = n[F(x)]^{n-1}f(x) \text{ respectively.}$$

- 2) The c.d.f. and p.d.f. of $X_{(1)}$, the first o.s. are given by

$$F_{X_{(1)}}(x) = 1 - [1 - F(x)]^n \text{ and } f_{X_{(1)}}(x) = n[1 - F(x)]^{n-1}f(x) \text{ respectively.}$$

- 3) The c.d.f. and p.d.f. of $X_{(j)}$, $1 \leq j \leq n$, the j^{th} o.s. are given by

$$F_{X_{(j)}}(x) = \sum_{i=j}^n \binom{n}{i} [F(x)]^i [1 - F(x)]^{n-i} \text{ and}$$

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} [F(x)]^{j-1} [1 - F(x)]^{n-j} f(x)$$

respectively.

Joint distributions

4) For $1 \leq i < j \leq n$, the j.p.d.f. of $X_{(i)}$ and $X_{(j)}$ is given by

$$f_{X_{(i)}, X_{(j)}}(u, v) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(u)]^{i-1} [F(v) - F(u)]^{j-i-1} [1 - F(v)]^{n-j} f(u) f(v)$$

for $-\infty < u < v < \infty$

5) The j.p.d.f. of k -order statistics $X_{(j_1)}, X_{(j_2)}, \dots, X_{(j_k)}$ where

$1 \leq r_1 < r_2 < \dots < r_k \leq n$ and $1 \leq k \leq n$ is for $x_1 \leq x_2 \leq \dots \leq x_k$ given by

$$f_{X_{(j_1)}, \dots, X_{(j_k)}}(x_1, \dots, x_k) = \frac{n!}{(j_1 - 1)!(j_2 - j_1 - 1)! \dots (j_k - j_{k-1} - 1)!(n - j_k)!} \times$$

$$F^{j_1-1}(x_1) [F(x_2) - F(x_1)]^{j_2-j_1-1} \dots [F(x_k) - F(x_{k-1})]^{j_k-j_{k-1}-1} \times$$

$$[1 - F(x_k)]^{n-j_k} f(x_1) f(x_2) \dots f(x_k)$$

6) The j.p.d.f. of $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ is given by

$$f_{X_{(1)}, \dots, X_{(n)}}(x_1, \dots, x_n) = \begin{cases} n! f(x_1) \dots f(x_n) & , \quad -\infty < x_1 < \dots < x_n < \infty \\ 0 & , \quad \text{otherwise} \end{cases}$$

Distribution of Range: Let us obtain the p.d.f. of the r.v. $R_{ij} = X_{(j)} - X_{(i)}$

for $i < j$. The j.p.d.f. of $X_{(i)}$ and $X_{(j)}$ is given by

$$f_{X_{(i)}, X_{(j)}}(u, v) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(u)]^{i-1} [F(v) - F(u)]^{j-i-1} [1 - F(v)]^{n-j} f(u) f(v) \dots (1)$$

Let $R_{ij} = X_{(j)} - X_{(i)}$ and $X = X_{(i)} \Rightarrow r_{ij} = v - u$ and $x = u$

$$\Rightarrow u = x \text{ and } v = r_{ij} + x$$

The Jacobian of transformation is given by

$$J = \frac{\partial(u,v)}{\partial(x,r_{ij})} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial r_{ij}} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial r_{ij}} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1 \text{ and } |J| = 1 \quad \dots(2)$$

From (1) and (2), the j.p.d.f. of $X_{(i)}$ and R_{ij} is given by

$$\begin{aligned} f_{X_{(i)}, R_{ij}}(x, r_{ij}) \\ = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(x)]^{i-1} [F(x + r_{ij}) - F(x)]^{j-i-1} \times \\ [1 - F(x + r_{ij})]^{n-j} f(x) f(x + r_{ij}) \quad \dots (3) \end{aligned}$$

From (2), the m.p.d.f. of R_{ij} is given by

$$f_{R_{ij}}(r_{ij}) = \int_{-\infty}^{\infty} f_{X_{(i)}, R_{ij}}(x, r_{ij}) dx \quad \dots (4)$$

Let $j = n$ and $i = 1$. Then the range is given by $W = X_{(n)} - X_{(1)}$. From (3) and (4), the p.d.f. of W is given by

$$g(w) = n(n-1) \int_{-\infty}^{\infty} [F(x+w) - F(x)]^{n-2} f(x+w) f(x) dx$$

The c.d.f. of w is given by

$$\begin{aligned} G(w) &= P(W \leq w) = \int_0^w g(u) du \\ &= \int_0^w \left(n(n-1) \int_{-\infty}^{\infty} [F(x+u) - F(x)]^{n-2} f(x+u) f(x) dx \right) du \\ &= n \int_{-\infty}^{\infty} f(x) \left[\int_0^w (n-1) f(x+u) [F(x+u) - F(x)]^{n-2} du \right] dx \\ \Rightarrow G(w) &= n \int_{-\infty}^{\infty} f(x) [F(x+w) - F(x)]^{n-1} dx \end{aligned}$$

Example 1: Let X_1, X_2, X_3, X_4 be a random sample of size 4 from uniform $[0, \theta]$ distribution. Find the p.d.f. of $X_{(1)}, X_{(3)}$ and $X_{(4)}$.

Solution: Since each $X \sim U(0, \theta)$, its p.d.f. is given by $f(x) = \frac{1}{\theta}, 0 < x < \theta$ and its c.d.f is given by

$$F(x) = P(X \leq x) = \int_0^x f(t)dt = \int_0^x \frac{1}{\theta} dt = \left[\frac{t}{\theta} \right]_0^x = \frac{x}{\theta}$$

The p.d.f. of $X_{(1)}$ is given by

$$f_{X_{(1)}}(x) = n[1 - F(x)]^{n-1}f(x) = 4\left(1 - \frac{x}{\theta}\right)^{4-1} \frac{1}{\theta}$$

$$\Rightarrow f_{X_{(1)}}(x) = \frac{4}{\theta}\left(1 - \frac{x}{\theta}\right)^3, 0 < x < \theta$$

The p.d.f of $X_{(3)}$ is given by

$$f_{X_{(3)}}(x) = \frac{4!}{2!1!} \left[\frac{x}{\theta} \right]^2 \left(1 - \frac{x}{\theta}\right)^1 \frac{1}{\theta} = \frac{12x^2(\theta - x)}{\theta^4}$$

$$\Rightarrow f_{X_{(3)}}(x) = \frac{12x^2(\theta - x)}{\theta^4}, 0 < x < \theta$$

The p.d.f of $X_{(4)}$ is given by

$$f_{X_{(4)}}(x) = n[F(x)]^{n-1}f(x) = 4\left(\frac{x}{\theta}\right)^3 \frac{1}{\theta} = \frac{4x^3}{\theta^4}$$

$$\Rightarrow f_{X_{(4)}}(x) = \frac{4x^3}{\theta^4}, 0 < x < \theta$$

Example 2: Let $X_1, X_2, \dots, X_n$ be i.i.d.r.v s with common p.d.f .

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

Find (i) p.d.f. of $X_{(j)}$, $1 \leq j \leq n$

(ii) j.p.d.f. of $X_{(j)}$ and $X_{(k)}$ for $1 \leq j < k \leq n$

(iii) p.d.f. of $R = X_{(n)} - X_{(1)}$

Solution: Given

$$\text{p.d.f: } f(x) = 1, 0 < x < 1$$

$$\text{c.g.f: } F(x) = \int_0^x f(t)dt = x \implies F(x) = x, 0 < x < 1$$

(i) The pdf of $X_{(j)}$ is given by

$$f_{X_{(j)}}(x_j) = \frac{n!}{(j-1)!(n-j)!} x_j^{j-1} (1-x_j)^{n-j} \text{ for } 0 < x_j < 1, 1 \leq j \leq n$$

(ii) The j.p.d.f. of $X_{(j)}$ and $X_{(k)}$ is given by

$$f_{X_{(j)}, X_{(k)}}(x_j, x_k) = \frac{n!}{(j-1)!(k-j-1)!(n-k)!} x_j^{j-1} (x_k - x_j)^{k-j-1} (1-x_k)^{n-k},$$

$$0 < x_j < x_k < 1 \text{ where } 1 \leq j < k \leq n$$

The j.p.d.f. of $X_{(1)}$ and $X_{(n)}$ is given by

$$f_{X_{(1)}, X_{(n)}}(x_1, x_n) = n(n-1)(x_n - x_1)^{n-2}, 0 < x_1 < x_n < 1$$

(iii) The p.d.f. of $R = X_{(n)} - X_{(1)}$ is given by

$$g(w) = n(n-1)w^{n-2}(1-w), 0 < w < 1$$

Example 3: Let $X_{(1)}, X_{(2)}, X_{(3)}$ be the o.s. of i.i.d.r.vs X_1, X_2, X_3 with common p.d.f.

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

Let $Y_1 = X_{(3)} - X_{(2)}$ and $Y_2 = X_{(2)}$. Show that Y_1 and Y_2 are independent.

Solution: The c.d.f. is given by $F(x) = \int_0^x f(t)dt = 1 - e^{-x\beta}, x > 0$.

Then the j.p.d.f. of $X_{(2)}$ and $X_{(3)}$ is given by

$$f_{X_{(2)}, X_{(3)}}(x, y) = \frac{3!}{1! 0! 0!} (1 - e^{-\beta x}) \beta e^{-\beta x} \beta e^{-\beta y}, 0 < x < y < \infty$$

Here $y_1 = y - x$ and $y_2 = x$

$$\Rightarrow x = y_2 \text{ and } y = y_1 + y_2$$

The Jacobian of transformation is given by

$$J = \frac{\partial(x, y)}{\partial(y_1, y_2)} = \begin{vmatrix} \frac{\partial x}{\partial y_1} & \frac{\partial x}{\partial y_2} \\ \frac{\partial y}{\partial y_1} & \frac{\partial y}{\partial y_2} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix} = -1 \text{ and } |J| = 1$$

The j.p.d.f. of Y_1 and Y_2 is given by

$$f(y_1, y_2) = 3! \beta^2 (1 - e^{-\beta y_2}) e^{-\beta y_2} e^{-\beta(y_1 + y_2)}, 0 < y_1 < \infty, 0 < y_2 < \infty \dots (1)$$

The m.p.d.f. of Y_2 is given by

$$f_2(y_2) = 3! \beta e^{-2\beta y_2} (1 - e^{-\beta y_2}), 0 < y_2 < \infty \dots (2)$$

and the m.p.d.f. of Y_1 is given by

$$f_1(y_1) = \beta e^{-\beta y_1}, 0 < y_1 < \infty \dots (3)$$

From (1), (2) and (3), Y_1 and Y_2 are independent.

Example 4: Let $X_1, X_2, \dots, X_n$ be a random samples from a population with continuous density. Show that $Y = \min(X_1, X_2, \dots, X_n)$ is exponential with parameter $n\lambda$ iff each X_i is exponential with parameter λ .

Solution: Let X_i be the i.i.d exponential variates with parameter λ and p.d.f.

$$f(x) = \lambda e^{-\lambda x}, x > 0, \lambda > 0$$

$$\text{and } F(x) = P(X \leq x) = \int_0^x f(u) du = \lambda \int_0^x e^{-\lambda u} du = 1 - e^{-\lambda x}$$

The distribution function of $Y_1 = \min(X_1, \dots, X_n)$ is given by

$$F_{Y_1}(y) = 1 - [1 - F(x)]^n = 1 - [1 - 1 + e^{-\lambda x}]^n = 1 - e^{-(n\lambda)x}$$

which is the distribution function of exponential distribution with parameter $n\lambda$.

Thus, $Y_1 \sim \exp(n\lambda)$.

Converse: Let $Y_1 = \min(X_1, \dots, X_n) \sim \exp(n\lambda)$

$$\text{Then } P(Y_1 \leq y) = 1 - e^{-n\lambda y}$$

$$\text{Now, } P(Y_1 \geq y) = 1 - P(Y_1 \leq y) = 1 - (1 - e^{-n\lambda y}) = e^{-n\lambda y}$$

$$\Rightarrow P[\min(X_1, \dots, X_n) \geq y] = e^{-n\lambda y}$$

$$\Rightarrow P[X_1 \geq y, X_2 \geq y, \dots, X_n \geq y] = e^{-n\lambda y}$$

$$\Rightarrow \prod_{i=1}^n P(X_i \geq y) = e^{-n\lambda y} \quad (\because \text{ } X\text{'s are i. d. d})$$

$$\Rightarrow P(X_i \geq y) = e^{-\lambda y} \Rightarrow P(X_i \leq y) = 1 - e^{-\lambda y}$$

which is $\exp(\lambda)$ distribution. Thus, X_i 's are i.d.d $\text{Exp}(\lambda)$.

Example 5: For exponential distribution $f(x) = e^{-x}, x \geq 0$, show that the c.d.f. of $X_{(n)}$ in a random sample of size n is $F_n(x) = (1 - e^{-x})^n$. Hence prove that as $n \rightarrow \infty$, the c.d.f. of $X_n - \ln n$ tends to the limiting form

$$\exp(-\exp(-x)), -\infty < x < \infty.$$

Solution: Here $f(x) = e^{-x}, x \geq 0 \Rightarrow F(x) = P(X \leq x) = 1 - e^{-x}$.

The c.d.f of $X_{(n)}$ is given by $F_{X_{(n)}} = [F(x)]^n = (1 - e^{-x})^n$

The c.d.f. of $X_{(n)} - \ln n$ is given by

$$\begin{aligned} G_n(x) &= P[X_{(n)} - \ln n \leq x] = P[X_{(n)} \leq x + \ln n] \\ &= [1 - e^{-(x + \ln n)}]^n = [1 - e^{-x} e^{-\ln n}]^n \end{aligned}$$

$$\Rightarrow G_n(x) = \left(1 - \frac{e^{-x}}{n}\right)^n$$

$$\Rightarrow \lim_{n \rightarrow \infty} G_n(x) = \lim_{n \rightarrow \infty} \left(1 - \frac{e^{-x}}{n}\right)^n = e^{-e^{-x}} \quad \left(\because \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x\right)$$

Distribution of O.S. in discrete case: In discrete case there is no magic formula to compute the distribution of any $X_{(j)}$ or any of the joint distributions. A direct computation is the best course of action.

Let $X_1, X_2, \dots, X_n$ be a random sample, from a population with p.m.f.

$p(x_i) = P(X = x_i)$ for $i = 1, 2, \dots$

Let $r_i = \sum_{k=1}^i p(x_k)$. Then $P(X_{(j)} \leq x_i) = \sum_{k=j}^n \binom{n}{k} r_i^k (1 - r_i)^{n-k}$

$$P[X_{(j)} = x_i] = \sum_{k=j}^n \binom{n}{k} \left[r_i^k (1 - r_i)^{n-k} - r_{i-1}^k (1 - r_{i-1})^{n-k} \right]$$

Example 6: Let $X_1, X_2, \dots, X_n$ are i.i.d.r.vs with common geometric p.m.f. given by

$$p_k = P(X = k) = pq^{k-1}, \quad k = 1, 2, \dots, \quad 0 < p < 1, \quad q = 1 - p$$

- (i) Find p.m.f. of $X_{(r)}, 1 \leq r \leq n$ and
- (ii) Show that X_1 and $X_{(2)} - X_{(1)}$ are independent random variables and $X_{(2)} - X_{(1)}$ has a geometric distribution.

Solution:

- (i) For any integer $x \geq 1$ and $r \geq 1$,

$$P[X_{(r)} = x] = P[X_{(r)} \leq x] - P[X_{(r)} \leq (x - 1)]$$

Now $P(X_{(r)} \leq x) = P[\text{at least } r \text{ of } X \text{ s are } \leq x]$

$$= \sum_{i=r}^n \binom{n}{i} [P(X_1 \leq x)]^i [P(X_1 > x)]^{n-i}$$

$$\text{and } P(X_1 \geq x) = \sum_{k=x}^{\infty} pq^{k-1} = (1-p)^{x-1} = q^{x-1}$$

$$\text{It follows that, } P[X_{(r)} = x] = \sum_{i=r}^n \binom{n}{i} q^{(x-1)(n-i)} [q^{n-i} (1-q^x)^i - (1-q^{x-1})^i]$$

$$x = 1, 2, \dots$$

- (ii) Let $n = r = 2$. Then, $P[X_{(2)} = x] = pq^{x-1} (pq^{x-1} + 2 - 2q^{x-1}), x \geq 1$

Also, for integers $x, y \geq 1$, we have $P[X_{(1)} = x, X_{(2)} - X_{(1)} = y]$

$$\begin{aligned} &= P[X_{(1)} = x, X_{(2)} = x + y] \\ &= P[X_1 = x, X_2 = x + y] + P[X_1 = x + y, X_2 = x] \\ &= 2pq^{x-1}pq^{x+y-1} = 2pq^{2x-2}pq^y \\ &= P(X_{(1)} = x)P(X_{(2)} = y) \end{aligned}$$

$$\text{and } P(X_{(1)} = 1, X_{(2)} - X_{(1)} = 0) = P(X_{(1)} = X_{(2)} = 1) = p^2$$

It follows that $X_{(1)}$ and $X_{(2)} - X_{(1)}$ are independent random variables and, moreover, that $X_{(2)} - X_{(1)}$ a geometric distribution.

Order Statistics

1. Definition: Order Statistics of a sample.

Let $X_1, X_2, \dots, X_n$ be a random sample from a population with p.d.f. $f(x)$. Then,

$$X_{(1)} = \min(X_1, X_2, \dots, X_n)$$

$$X_{(n)} = \max(X_1, X_2, \dots, X_n)$$

$$\text{and } X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(k)} \leq \dots \leq X_{(n)}$$

2. p.d.f.'s for $X_{(1)}$ and $X_{(n)}$

W.L.O.G.(Without Loss of Generality), let's assume X is continuous.

$$P(X_{(1)} > x) = P(X_1 > x, X_2 > x, \dots, X_n > x) = \prod_{i=1}^n P(X_i > x)$$

$$= \prod_{i=1}^n [1 - F_{X_i}(x)]$$

$$F_{X_{(1)}}(x) = 1 - \prod_{i=1}^n [1 - F_{X_i}(x)]$$

$$f_{X_{(1)}}(x) = -\frac{d}{dx} \prod_{i=1}^n [1 - F_{X_i}(x)] = -\frac{d}{dx} \prod_{i=1}^n [1 - F(x)]^n = n[1 - F(x)]^{n-1}f(x)$$

$$F_{X_{(n)}}(x) = P(X_{(n)} \leq x) = P(X_1 \leq x, X_2 \leq x, \dots, X_n \leq x) = \prod_{i=1}^n F_{X_i}(x) = [F(x)]^n$$

$$f_{X_{(n)}}(x) = n[F(x)]^{n-1}f(x)$$

Example 1. Let $X_n \stackrel{i.i.d}{\sim} \exp(\lambda), i=1, \dots, n$

Please (1). Derive the MLE of λ

(2). Derive the p.d.f. of $X_{(1)}$

(3). Derive the p.d.f. of $X_{(n)}$

Solutions.

(1).

$$L = \prod_{i=1}^n f(x_i) = \prod_{i=1}^n (\lambda e^{-\lambda x_i}) = \lambda^n e^{-\lambda \sum_{i=1}^n x_i}$$

$$l = \ln L = n \ln \lambda - \lambda \sum_{i=1}^n x_i$$

$$\frac{dl}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^n x_i = 0 \Rightarrow \hat{\lambda} = \frac{1}{\bar{X}}$$

Is $\hat{\lambda}$ an unbiased estimator of λ ? $E\left(\frac{1}{\bar{X}}\right) = ?$

$$M_{X_i}(t) = \frac{\lambda}{\lambda - t}$$

$$M_{\sum_{i=1}^n X_i}(t) = \left(\frac{\lambda}{\lambda - t}\right)^n$$

$$Y = \sum_{i=1}^n X_i \sim \text{gamma}(\lambda, n)$$

$$f_{Y - \sum_{i=1}^n X_i}(x) = \frac{\lambda}{\Gamma(n)} (\lambda y)^{n-1} e^{-\lambda y}$$

$$\text{Let } Y = \sum_{i=1}^n X_i$$

$$\begin{aligned} E\left(\frac{1}{Y}\right) &= \int_0^{\infty} \frac{1}{y} \frac{\lambda}{(n-1)!} (\lambda y)^{n-1} e^{-\lambda y} dy \\ &= \frac{\lambda}{n-1} \int_0^{\infty} \frac{\lambda}{(n-2)!} (\lambda y)^{n-2} e^{-\lambda y} dy \\ &= \frac{\lambda}{n-1} \end{aligned}$$

$$E\left(\frac{1}{\bar{X}}\right) = n \left(\frac{\lambda}{n-1}\right) = \frac{n\lambda}{n-1} \neq \lambda$$

$\hat{\lambda}$ is not unbiased

$$(2). X_{(1)} = \min(X_1, X_2, \dots, X_n)$$

$$P(X_{(1)} > x) = \prod_{i=1}^n P(X_i > x) = \prod_{i=1}^n [1 - F(x)] = [1 - F(x)]^n$$

$$F_{X_{(1)}}(x) = 1 - [1 - F(x)]^n$$

$$f_{X_{(1)}}(x) = n[1 - F(x)]^{n-1} f(x)$$

$$f(x) = \lambda e^{-\lambda x}, x > 0 \text{ (exponential distribution)}$$

$$F(x) = \int_0^x f(u) du = \int_0^x \lambda e^{-\lambda u} du = [-e^{-\lambda u}]_0^x = 1 - e^{-\lambda x}$$

$$\begin{aligned} f_{X_{(1)}}(x) &= n\lambda e^{-\lambda x} [1 - (1 - e^{-\lambda x})]^{n-1} = n\lambda e^{-\lambda x} (e^{-\lambda x})^{n-1} \\ &= n\lambda (e^{-\lambda x})^n, x > 0 \end{aligned}$$

$$(3). X_{(n)} = \max(X_1, X_2, \dots, X_n)$$

$$F_{X_{(n)}}(x) = P(X_{(n)} \leq x) = \prod_{i=1}^n P(X_i \leq x) = [F(x)]^n$$

$$f_{X_{(n)}}(x) = n[F(x)]^{n-1}f(x) = n[1 - e^{-\lambda x}]^{n-1}\lambda e^{-\lambda x}, x > 0$$

3. Order statistics are useful in deriving the MLE's.

Example 2. Let X be a random variable with pdf.

$$f(x) = \begin{cases} 1, & \text{if } x \in [\theta - \frac{1}{2}, \theta + 1/2] \\ 0, & \text{otherwise} \end{cases}$$

Derive the MLE of θ .

Solution.

Uniform Distribution $\Rightarrow$ important!!

$$L = \prod_{i=1}^n f(x_i) = \begin{cases} 1, & \text{if all } x_i \in [\theta - \frac{1}{2}, \theta + 1/2] \\ 0, & \text{otherwise} \end{cases}$$

MLE : $\max \ln L \rightarrow \max L$

$$\text{means} \Rightarrow \theta - \frac{1}{2} \leq X_1 \leq \theta + 1/2$$

$$\theta - \frac{1}{2} \leq X_2 \leq \theta + 1/2$$

...

$$\theta - \frac{1}{2} \leq X_n \leq \theta + 1/2$$

Now we re-express the domain in terms of the order statistics as follows:

$$\theta - \frac{1}{2} \leq X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)} \leq \theta + 1/2$$

$$\theta \leq X_{(1)} + \frac{1}{2}$$

$$\theta \geq X_{(n)} - \frac{1}{2}$$

Therefore,

$$\text{If } \theta \in \left[X_{(n)} - \frac{1}{2}, X_{(1)} + \frac{1}{2} \right], \text{ then } L = 1$$

Therefore, any $\hat{\theta} \in \left[X_{(n)} - \frac{1}{2}, X_{(1)} + \frac{1}{2} \right]$ is an MLE for θ .

4. The pdf of a general order statistic $X_{(j)}$

Let $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ denote the order statistics of a random sample, $X_1, X_2, \dots, X_n$, from a continuous population with cdf $F_X(x)$ and pdf $f_X(x)$. Then the pdf of $X_{(j)}$ is

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f_X(x) [F_X(x)]^{j-1} [1 - F_X(x)]^{n-j}$$

Proof: Let Y be a random variable that counts the number of $X_1, X_2, \dots, X_n$ less than or equal to x . Then we have $Y \sim B(n, F_X(x))$. Thus:

$$F_{X_{(j)}}(x) = P(Y \geq j) = \sum_{k=j}^n \binom{n}{k} [F_X(x)]^k [1 - F_X(x)]^{n-k}$$

5. The Joint Distribution of Two Order Statistics $X_{(i)}$ and $X_{(j)}$

Let $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ denote the order statistics of a random sample, $X_1, X_2, \dots, X_n$, from a continuous population with cdf $F_X(x)$ and pdf $f_X(x)$. Then the joint pdf of $X_{(i)}$ and $X_{(j)}$, $1 \leq i < j \leq n$, is

$$\begin{aligned} & f_{X_{(i)}, X_{(j)}}(u, v) \\ &= \frac{n!}{(i-1)!(j-1-i)!(n-j)!} f_X(u) f_X(v) [F_X(u)]^{i-1} [F_X(v) - F_X(u)]^{j-1-i} [1 - F_X(v)]^{n-j} \\ & \text{for } -\infty < u < v < \infty \end{aligned}$$

6. Special functions of order statistics

(1) Median (of the sample):

$$\begin{cases} X_{(k+1)}; & \text{if } n = 2k + 1 \\ \frac{X_{(k)} + X_{(k+1)}}{2}; & \text{if } n = 2k \end{cases}$$

(2) Range (of the sample): $X_{(n)} - X_{(1)}$

7. More examples of order statistics

Example 3. Let X_1, X_2, X_3 be a random sample from a distribution of the continuous type having pdf $f(x)=2x$, $0 < x < 1$, zero elsewhere.

(a) compute the probability that the smallest of X_1, X_2, X_3 exceeds the median of the distribution.

(b) If $Y_1 \leq Y_2 \leq Y_3$ are the order statistics, find the correlation between Y_2 and Y_3 .

Answer:

(a)

$$F(x) = P(X_i < x) = x^2;$$

$$\int_0^t 2x dx = \frac{1}{2}; t = \frac{\sqrt{2}}{2}$$

$$\begin{aligned} P(\min(X_1, X_2, X_3) > t) &= P(X_1 > t, X_2 > t, X_3 > t) = P(X_1 > t)P(X_2 > t)P(X_3 > t) \\ &= [1 - F(t)]^3 = (1 - t^2)^3 = \frac{1}{8} \end{aligned}$$

(b)

Please refer to the textbook/notes for the order statistics pdf and joint pdf formula. We have

$$f_{Y_3}(x) = 6 * x^5; 0 < x < 1$$

$$f_{Y_2}(x) = 12 * (x^3 - x^5); 0 < x < 1$$

$$E(Y_3) = 6/7,$$

$$E(Y_2) = 24/35;$$

$$f_{Y_2, Y_3}(y_2, y_3) = 24 * (y_2)^3 * y_3; 0 < y_2 \leq y_3 < 1$$

$$E(Y_2 Y_3) = \int_0^1 \left[\int_0^{y_3} y_2 * y_3 * 24 * (y_2)^3 * y_3 dy_2 \right] dy_3 = \frac{3}{5};$$

$$var(Y_3) = \frac{6}{8} - \left(\frac{6}{7}\right)^2 = \frac{6}{392};$$

$$var(Y_2) = \frac{1}{2} - \left(\frac{24}{35}\right)^2;$$

$$corr(Y_2, Y_3) = \frac{E(Y_2 Y_3) - E(Y_2)E(Y_3)}{\sqrt{var(Y_2)var(Y_3)}} = 0.57$$

Example 4. Let $Y_1 \leq Y_2 \leq Y_3$ denote the order statistics of a random sample of size 3 from a distribution with pdf $f(x) = 1$, $0 < x < 1$, zero elsewhere. Let $Z = (Y_1 + Y_3)/2$ be the midrange of the sample. Find the pdf of Z .

From the pdf, we can get the cdf : $F(x) = x, 0 < x < 1$
 Let

$$W = Y_1$$

$$Z = (Y_1 + Y_3)/2$$

The inverse transformation is:

$$Y_1 = W$$

$$Y_3 = 2Z - W$$

The joint pdf of Y_1 and Y_3 is:

$$f(y_1, y_3) = \begin{cases} 6(y_3 - y_1), & 0 < y_1 \leq y_3 < 1 \\ 0, & \text{o. w.} \end{cases}$$

We then find the Jacobian: $J = -2$

Now we can obtain the joint pdf of Z, W :

$$f(z, w) = \begin{cases} |2|6(2z - w - w) = 24(z - w), & 0 < w \leq 2z - w < 1 \\ 0, & \text{o. w.} \end{cases}$$

From $0 < w \leq 2z - w < 1$, we have:

$$w > 0;$$

$$w > 2z - 1;$$

$$w \leq z$$

Together they give us the domain of w as:

$$\max(0, 2z - 1) < w \leq z$$

Therefore the pdf of Z (non-zero portion) is:

$$f(z) = \begin{cases} \int_{2z-1}^z 24(z - w) dw = 12(z - 1)^2, & 2z - 1 > 0 \\ \int_0^z 24(z - w) dw = 12(z^2), & 2z - 1 \leq 0 \end{cases}$$

We also remind ourselves that:

$$0 < z < 1$$

Therefore the entire pdf of the midrange Z is:

$$f(z) = \begin{cases} \int_{2z-1}^z 24(z-w) dw = 12(z-1)^2, & 1/2 < z < 1 \\ \int_0^z 24(z-w) dw = 12(z^2), & 0 < z \leq 1/2 \\ 0, & \text{otherwise} \end{cases}$$

Example 5. Let $Y_1 \leq Y_2 \leq Y_3 \leq Y_4$ be the order statistics of a random sample of size $n = 4$ from a distribution with pdf $f(x) = 2x$, $0 < x < 1$, zero elsewhere.

(a) Find the joint pdf of Y_3 and Y_4 .

(b) Find the conditional pdf of Y_3 , given $Y_4 = y_4$.

(c) Evaluate $E[Y_3|y_4]$.

Solution:

(a)

$$f_{Y_1, Y_2, Y_3, Y_4}(y_1, y_2, y_3, y_4) = 4! (2y_1)(2y_2)(2y_3)(2y_4)$$

for $0 < y_1 \leq y_2 \leq y_3 \leq y_4 < 1$. We have:

$$\begin{aligned} f_{Y_3, Y_4}(y_3, y_4) &= \int_{y_2=0}^{y_3} \int_{y_1=0}^{y_2} f_{Y_1, Y_2, Y_3, Y_4}(y_1, y_2, y_3, y_4) dy_1 dy_2 \\ &= 48y_3^5 y_4 \end{aligned}$$

for $0 < y_3 \leq y_4 < 1$

(Note: You can also obtain the joint pdf of these two order statistics by using the general formula directly.)

(b)

$$f_{Y_3|Y_4}(y_3|y_4) = \frac{f_{Y_3, Y_4}(y_3, y_4)}{f_{Y_4}(y_4)} = \frac{48y_3^5 y_4}{8y_4^7} = \frac{6y_3^5}{y_4^6}$$

for $0 < y_3 \leq y_4$.

(c)

$$E[Y_3|Y_4 = y_4] = \int_0^{y_4} \frac{6y_3^5}{y_4^6} dy_3 = \frac{6y_4}{7}$$

Example 6. Suppose $X_1, \dots, X_n$ are iid with pdf $f(x; \theta) = 2x/\theta^2$, $0 < x \leq \theta$, zero elsewhere. Note this is a nonregular case. Find:

(a) The mle $\hat{\theta}$ for θ .

(b) The constant c so that $E(c\hat{\theta}) = \theta$.

(c) The mle for the median of the distribution.

Solution:

$$(a) L(\theta; X) = \prod_{i=1}^n \frac{2x_i}{\theta^2} = \frac{2^n \prod_{i=1}^n x_i}{\theta^{2n}} \leq \frac{2^n \prod_{i=1}^n x_i}{[\max(x_i)]^{2n}}$$

$$\text{So } \hat{\theta} = X_{(n)} \quad X_{(n)} = \max(X_1, \dots, X_n)$$

Dear students: note that this is no typo in the above – the truth is that $\theta \geq X_{(n)}$ – and so the smallest possible value for θ is $X_{(n)}$

$$(b) F_X(x) = \int_0^x \frac{2t}{\theta^2} dt = \frac{x^2}{\theta^2} \quad 0 < x \leq \theta$$

$$\text{So } F_{X_{(n)}}(x) = \left(\frac{x^2}{\theta^2}\right)^n = \frac{x^{2n}}{\theta^{2n}} \quad 0 < x \leq \theta$$

$$f_{X_{(n)}}(x) = \frac{2nx^{2n-1}}{\theta^{2n}} \quad 0 < x \leq \theta$$

$$E(c\hat{\theta}) = cE(\hat{\theta}) = c \int_0^\theta x \frac{2nx^{2n-1}}{\theta^{2n}} dx = \frac{2nc}{2n+1} \theta = \theta$$

$$\text{So } c = \frac{2n+1}{2n}$$

$$(c) \text{ Let } F(x) = \frac{x^2}{\theta^2} = \frac{1}{2}, \text{ then } x = \frac{\theta}{\sqrt{2}}$$

So the median of the distribution is $\frac{\theta}{\sqrt{2}}$

The mle for the median of the distribution is

$$\frac{\hat{\theta}}{\sqrt{2}} = \frac{X_{(n)}}{\sqrt{2}} = \frac{\sqrt{2}}{2} X_{(n)}$$

Mean Squared Error (M.S.E.)

How to evaluate an estimator?

For unbiased estimators, all we need to do is to compare their variances, the smaller the variance, the better is estimator.
Now, what if the estimators are not all unbiased? How do we compare them?

Definition: Mean Squared Error (MSE)

Let $T = t(X_1, X_2, \dots, X_n)$ be an estimator of $\tau(\theta)$, then the M.S.E. of the estimator T is defined as :

$$\begin{aligned} \text{MSE}_t(\tau(\theta)) &= E[(T - \tau(\theta))^2] : \text{average squared distance from } T \text{ to } \tau(\theta) \\ &= E[(T - E(T) + E(T) - \tau(\theta))^2] \\ &= E[(T - E(T))^2] + E[(E(T) - \tau(\theta))^2] + 2E[(T - E(T))(E(T) - \tau(\theta))] \\ &= E[(T - E(T))^2] + E[(E(T) - \tau(\theta))^2] + 0 \\ &= \text{Var}(T) + (E(T) - \tau(\theta))^2 \end{aligned}$$

Here $|E(T) - \tau(\theta)|$ is “the bias of T ”

If unbiased, $(E(T) - \tau(\theta))^2 = 0$.

The estimator has smaller mean-squared error is better.

Example 1. Let $X_1, X_2, \dots, X_n \stackrel{\text{i.i.d}}{\sim} N(\mu, \sigma^2)$

M.L.E. for μ is $\hat{\mu} = \bar{X}$; M.L.E. for σ^2 is $\hat{\sigma}^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$

1. M.S.E. of $\hat{\sigma}^2$?

2. M.S.E. of S^2 as an estimator of σ^2

Solution.

1.

$$\text{MSE}_{\hat{\sigma}^2}(\sigma^2) = E[(\hat{\sigma}^2 - \sigma^2)^2] = \text{Var}(\sigma^2) + (E(\hat{\sigma}^2) - \sigma^2)^2$$

To get $\text{Var}(\hat{\sigma}^2)$, there are 2 approaches.

a. By the first definition of the Chi-square distribution.

Note $X_i \sim N(\mu, \sigma^2)$; $W = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} \sim \chi_{n-1}^2$, Gamma($\lambda = \frac{1}{2}, S = \frac{n-1}{2}$)

$$E(W) = \frac{S}{\lambda} = n - 1; \text{Var}(W) = \frac{S}{\lambda^2} = 2(n - 1)$$

$$\text{Var}(\hat{\sigma}^2) = \text{Var}\left(\frac{W}{n} \sigma^2\right) = \frac{\sigma^4}{n^2} \text{Var}(W) = \frac{\sigma^4}{n^2} 2(n - 1)$$

b. By the second definition of the Chi-square distribution.

$$\text{For } Z \sim N(0,1), W = \sum_{i=1}^n Z_i^2$$

$$\begin{aligned} \text{Var}(Z^2) &= E\left[\left(Z^2 - E(Z^2)\right)^2\right] \\ &= E\left[\left(Z^2 - (\text{var}(Z^2) + E(Z))\right)^2\right] \\ &= E[(Z^2 - 1)^2] \end{aligned}$$

$$\text{Since } \text{Var}(Z) = E(Z^2) - E(Z)^2 =$$

$$\begin{aligned} 1 \text{ from } Z \sim N(0,1), E(Z^2) &= 1 = E[Z^4 - 2E(Z^2) + 1] \\ &= E(Z^4) - 1 \end{aligned}$$

Calculate the 4th moment of $Z \sim N(0,1)$ using the mgf of Z ;

$$M_Z(t) = e^{t^2/2}$$

$$M'_Z(t) = te^{t^2/2}$$

$$M''_Z(t) = te^{t^2/2} + t^2e^{t^2/2}$$

$$M^{(3)}_Z(t) = 3te^{t^2/2} + t^2e^{t^2/2}$$

$$M^{(4)}_Z(t) = 3e^{t^2/2} + 6t^2e^{t^2/2} + t^4e^{t^2/2}$$

$$\text{Set } t = 0, M^{(4)}_Z(0) = 3 = E(Z^4)$$

$$\text{Var}(Z^2) = 3 - 1 = 2$$

$$\text{Var}(W) = \sum_{i=1}^{n-1} \text{Var}(Z_i^2) = 2(n - 1)$$

$$\hat{\sigma}^2 = \frac{\sigma^2}{n} W, \text{Var}(\hat{\sigma}^2) = \frac{\sigma^4}{n^2} 2(n - 1)$$

$$\begin{aligned} \text{MSE}_{\hat{\sigma}^2}(\sigma^2) &= \text{Var}(\hat{\sigma}^2) + [E(\hat{\sigma}^2) - \sigma^2]^2 \\ &= \frac{2(n - 1)}{n^2} \sigma^4 + \left[E\left(\frac{n - 1}{n} S^2\right) - \sigma^2\right]^2 \\ &= \frac{2(n - 1)}{n^2} \sigma^4 \\ &\quad + \left[\frac{n - 1}{n} \sigma^2 - \sigma^2\right]^2 \text{ (we know } E(S^2) \\ &= \sigma^2) = \frac{2n - 1}{n^2} \sigma^4 \end{aligned}$$

The M.S.E. of $\hat{\sigma}^2$ is $\frac{2n-1}{n^2} \sigma^4$

We know S^2 is an unbiased estimator of σ^2

$$\begin{aligned} E[(S^2 - \sigma^2)^2] &= \text{Var}(S^2) + 0 = \text{Var}\left(\frac{\sigma^2 W}{n-1}\right) \\ &= \left(\frac{\sigma^2}{n-1}\right)^2 \text{var}(W) = \frac{2\sigma^4}{n-1} \end{aligned}$$

Exercise:

Compare the MSE of $\hat{\sigma}^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$ and $\hat{\sigma}^2 = S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$. Which one is a better estimator (in terms of the MSE)?

Solution.

$$\text{MSE}_{\hat{\sigma}^2}(\sigma^2) = E[(\hat{\sigma}^2 - \sigma^2)^2] = \text{Var}(\hat{\sigma}^2) + (E(\hat{\sigma}^2) - \sigma^2)^2$$

Then, we have M.S.E. of $\hat{\sigma}^2$ is $\frac{2n-1}{n^2} \sigma^4$.

$$\text{MSE}_{S^2}(\sigma^2) = E[(S^2 - \sigma^2)^2] = \text{Var}(S^2) + (E(S^2) - \sigma^2)^2$$

$$\begin{aligned} \text{MSE}_{S^2}(\sigma^2) &= \text{Var}(S^2) + 0 = \text{Var}\left(\frac{\sigma^2 W}{n-1}\right) = \\ &= \left(\frac{\sigma^2}{n-1}\right)^2 \text{var}(W) = \frac{2\sigma^4}{n-1} \end{aligned}$$

$$\therefore \frac{2n-1}{n^2} \sigma^4 - \frac{2\sigma^4}{n-1} = \frac{1-3n}{n^2(n-1)} \sigma^4 < 0$$

$$\therefore \text{MSE}_{\hat{\sigma}^2}(\sigma^2) = \frac{2n-1}{n^2} \sigma^4 < \frac{2\sigma^4}{n-1} = \text{MSE}_{S^2}(\sigma^2)$$

$$\therefore \hat{\sigma}^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n} \text{ is the better estimator}$$

Homework: Read the following chapter/sections from our textbook: Chapter 1, 2, 3 (3.1, 3.2, 3.3), 4 (4.1, 4.2, 4.3, 4.5, 4.6), 5 (5.1, 5.2, 5.3, 5.4), 7 (7.1, 7.2.1, 7.2.2, 7.3.1). These are materials covered so far in our class.

4.1. Order Statistics

Exercise

1. Let $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ be the set of order statistics of independent r.v.s with common p.d.f.

$$f(x) = \begin{cases} \beta e^{-x\beta} & , \quad x \geq 0 \\ 0 & , \quad \text{otherwise} \end{cases}$$

Show that $X_{(r)}$ and $X_{(s)} - X_{(r)}$ are independent for any $s > r$.

2. Let $X_1, X_2, \dots, X_n$ be i.i.d. r.v.s with p.d.f.

$$f(y) = \begin{cases} y^\alpha & \text{if } 0 < y < 1 \\ 0 & \text{otherwise, } \alpha > 0 \end{cases}$$

Show that $\frac{X_{(i)}}{X_{(n)}}$, $i = 1, 2, \dots, n-1$ and $X_{(n)}$ are independent.

3. Let $X_1, X_2, \dots, X_n$ be i.i.d. r.v.s with common p.d.f.

$$f(x) = \alpha \frac{\sigma^\alpha}{x^{\alpha+1}}, \quad x > \sigma \text{ where } \alpha > 0, \sigma > 0$$

Show that $X_{(1)}$ and $\left(\frac{X_{(2)}}{X_{(1)}}, \dots, \frac{X_{(n)}}{X_{(1)}}\right)$ are independent.

4. Let $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ be the order statistics of n independent r.v.s $X_1, X_2, \dots, X_n$ with common p.d.f.

$$f(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Show that $Y_1 = \frac{X_{(1)}}{X_{(2)}}$, $Y_2 = \frac{X_{(2)}}{X_{(3)}}$, $\dots$, $Y_{n-1} = \frac{X_{(n-1)}}{X_{(n)}}$ and $Y_n = X_{(n)}$ are independent.

5. An urn contains N identical marbles numbered 1 through N . From this urn n marbles are drawn, and let $X_{(n)}$ be the largest number drawn. Show that

$$P[X_{(n)} = k] = \frac{\binom{k-1}{n-1}}{\binom{N}{n}}, \quad k = n, n+1, \dots, N$$

$$\text{and } E(X_{(n)}) = \frac{n(N+1)}{(n+1)}.$$

P1:

Find the j.p.d.f of extremes (*i. e.*, $X_{(1)}$ and $X_{(n)}$).

P1:

Find the j.p.d.f of extremes (i. e., $X_{(1)}$ and $X_{(n)}$).

Solution:

The j.p.d.f. of $X_{(i)}$ and $X_{(j)}$ is given by

$$f_{X_{(i)}, X_{(j)}}(u, v) \\ = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(u)]^{i-1} [F(v) - F(u)]^{j-i-1} [1 - F(v)]^{n-j} f(u) f(v)$$

for $1 \leq i < j \leq n$ and $-\infty < u < v < \infty$.

Let $i = 1$ and $j = n$. Then the j.p.d.f. of $X_{(1)}$ and $X_{(n)}$ is given by

$$f_{X_{(1)}, X_{(n)}}(u, v) = n(n-1)[F(v) - F(u)]^{n-2} f(u) f(v), \quad -\infty < u < v < \infty$$

P2:

Find the j.p.d.f. of the mid range $M = \frac{1}{2}[X_{(1)} + X_{(2)}]$.

P2:

Find the j.p.d.f. of the mid range $M = \frac{1}{2}[X_{(1)} + X_{(2)}]$.

Solution:

Let $x = u$ and $y = \frac{1}{2}(u + v)$. Then $u = x$ and $v = 2y - x$ and the jacobian of transformation is given by

$$J = \frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -1 & 2 \end{vmatrix} = 2 \text{ and } |J| = 2$$

$X_{(1)}$ and $X_{(2)}$ given in P_1 , the j. p. d. f. of $X_{(1)}$ and M is given by

$$\begin{aligned} f_{X_{(1)},M}(x,y) &= f_{X_{(1)},X_{(n)}}(x,y)|J| \\ &= 2n(n-1)[F(2y-x) - F(x)]^{n-2}f(x)f(2y-x) \end{aligned}$$

The m.p.d.f. of M is given by

$$\begin{aligned} f_M(y) &= \int_{-\infty}^{\infty} f_{X_{(1)},M}(x,y)dx \\ &= 2n(n-1) \int_{-\infty}^{\infty} [F(2y-x) - F(x)]^{n-2} f(x) f(2y-x) dx \end{aligned}$$

P3:

Find the conditional probability density function of $X_{(i)}$ given $X_{(j)}$ for $1 \leq i < j \leq n$.

P3:

Find the conditional probability density function of $X_{(i)}$ given $X_{(j)}$ for $1 \leq i < j \leq n$.

Solution:

The c.p.d.f. of $X_{(i)}$ given $X_{(j)}$ is given by

$$\begin{aligned} f_{X_{(i)}|X_{(j)}}(x|y) &= \frac{f_{X_{(i)},X_{(j)}}(x,y)}{f_{X_{(j)}}(y)} \\ &= \frac{\frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(x)]^{i-1} [F(y)-F(x)]^{j-i-1} [1-F(y)]^{n-j} f(x)f(y)}{\frac{n!}{(j-1)!(n-j)!} [F(y)]^{j-1} [1-F(y)]^{n-j} f(y)} \\ &= \frac{(j-1)!}{(i-1)!(j-i-1)!} [F(x)]^{i-1} [F(y) - F(x)]^{j-i-1} [F(y)]^{1-j} f(x) \text{ for } y \geq x \end{aligned}$$

P4:

Find the conditional p.d.f. of $X_{(i)}$ given $X_{(i+1)}$

P4:

Find the conditional p.d.f. of $X_{(i)}$ given $X_{(i+1)}$

Solution:

Let $j = i + 1$ in P3. Then

$$f_{X_{(i)}|X_{(j+1)}}(x|y) = i[F(x)]^{i-1}[F(y)]^{-i}f(x)$$

4.2

Convergence of Sequence of Random Variables

In this module we investigate convergence properties of sequences of random variables. Throughout this module we assume that $\{X_1, X_2, \dots\}$ or $\{X_n\}$ is a sequence of r.v.s and X is a r.v. We consider **four different modes of convergence for random variables**.

1. **Almost sure convergence:** It is the **probabilistic version of pointwise convergence** known from elementary real analysis. It is also known as **convergence with probability one**.

The sequence of r.v.s $\{X_n\}$ is said to **converge almost surely** to a r.v. X if

$$P\left(\left\{w : \lim_{n \rightarrow \infty} X_n(w) = X(w)\right\}\right) = 1$$

In this case we write $X_n \xrightarrow{a.s.} X$ (or $X_n \rightarrow X$ with probability 1).

2. **Convergence in probability:** It is essentially mean that the probability that $|X_n - X|$ exceeds any prescribed strictly positive value, converges to zero. The basic idea behind this type of convergence is that the probability of an *unusual* outcome becomes smaller and smaller as the sequence progresses. The sequence of r.v.s $\{X_n\}$ is said to **converge in probability** to a r.v. X if

$$\lim_{n \rightarrow \infty} P(\{|X_n - X| > \epsilon\}) = 0$$

for every $\epsilon > 0$. It is denoted by $X_n \xrightarrow{P} X$.

3. **Convergence in r^{th} mean:** Let $\{X_n\}$ be a sequence of r.v.s such that $E(|X_n|^r) < \infty$ for some $r > 0$. We say that X_n **converges in the r^{th} mean** to a r.v. X if $E(|X|^r) < \infty$ and

$$E(|X_n - X|^r) \rightarrow 0 \text{ as } n \rightarrow \infty$$

and we write $X_n \xrightarrow{r} X$.

If $r = 2$, we call it as **convergence in quadratic mean** and it is denoted by

$$X_n \xrightarrow{q.m} X$$

4. **Convergence in distribution:** **Convergence in distribution** is very frequently used in practice, most often it arises from the application of the **central limit theorem** (to be discussed in module 4.5). Note that a cumulative distribution function (c.d.f) is briefly called as *distribution function (d.f)* also.

Let $\{F_n\}$ be a sequence of cumulative distribution functions (c.d.fs), if there exists a c.d.f. F such that as $n \rightarrow \infty$,

$$F_n(x) \rightarrow F(x)$$

for all x at which F is continuous, then we say that F_n **converges weakly** to F , and it is denoted by $F_n \xrightarrow{w} F$.

If $\{X_n\}$ is a sequence of r.vs and $\{F_n\}$ is the corresponding sequence of c.d.fs, then we say that X_n **converges in distribution** (or **law**) to X if there exists an r.v X with c.d.f. F such that $F_n \xrightarrow{w} F$. We write $X_n \xrightarrow{d} X$ or $X_n \xrightarrow{L} X$.

Note: It is quite possible for a given sequence of c.d.fs to converge to a function that is not a c.d.f.

Example: Let $F_n(x) = \begin{cases} 0, & x < n \\ 1, & x \geq n \end{cases}$

As $n \rightarrow \infty$, $F_n(x) \rightarrow F(x) = 0$ which is not a c.d.f.

Example 1: Let $X_1, X_2, \dots, X_n$ be i.i.d.r.vs with common p.d.f

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

Let $X_{(n)} = \max(X_1, \dots, X_n)$. Then show that $X_{(n)} \xrightarrow{L} X$, where X is degenerate at $x = \theta$.

(Note: We say that a r.v. X is **degenerate at $x = \theta$** if $P(X = \theta) = 1$)

Solution: Corresponding to p.d.f. $f(x) = \frac{1}{\theta}$, the c.d.f. is given by

$$F(x) = \int_0^x f(t)dt = \frac{1}{\theta} \int_0^x dt = \frac{x}{\theta}$$

$$\Rightarrow F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{x}{\theta} & , \quad 0 \leq x < \theta \\ 1 & , \quad x \geq \theta \end{cases}$$

Then the c.d.f. of $X_{(n)}$ is given by

$$F_n(x) = [F(x)]^n = \begin{cases} 0 & , \quad x < 0 \\ \left(\frac{x}{\theta}\right)^n & , \quad 0 \leq x < \theta \\ 1 & , \quad x \geq \theta \end{cases}$$

We see that as $n \rightarrow \infty$

$$F_n(x) = F(x) = \begin{cases} 0 & \text{if } x < \theta \\ 1 & \text{if } x \geq \theta \end{cases}$$

which is the d.f. of $P(X = \theta) = 1$. i. e., X is degenerate at $x = \theta$.

Thus $F_n \xrightarrow{w} F$ and hence $X_n \xrightarrow{L} X$.

The following example shows that convergence in distribution does not imply convergence of moments.

Example 2: Let F_n be a sequence of c.d.fs defined by

$$F_n(x) = \begin{cases} 0 & , \quad x < 0 \\ 1 - \frac{1}{n} & , \quad 0 \leq x < n \\ 1 & , \quad x \geq n \end{cases}$$

Show that $X_n \xrightarrow{L} X$ does not imply $E(X_n^k) \rightarrow E(X^k)$.

Solution: We see that as $n \rightarrow \infty$

$$F(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

Note that F_n is the c.d.f. of the r.v. X_n with p.m.f.

$$P(X_n = 0) = 1 - \frac{1}{n}, P(X_n = n) = \frac{1}{n}$$

and F is the c.d.f. of the r.v. degenerate at 0 i.e., $P(X = 0) = 1$.

Thus, $F_n \xrightarrow{w} F$ and hence $X_n \xrightarrow{L} X$. We have

$E(X_n^k) = 0^k \left(1 - \frac{1}{n}\right) + n^k \left(\frac{1}{n}\right) = n^{k-1}$, where k is a positive integer. Also, $E(X^k) = 0^k 1 = 0$. Hence $E(X_n^k) \not\rightarrow E(X^k)$ as $n \rightarrow \infty$

Therefore, $X_n \xrightarrow{L} X$ does not imply $E(X_n^k) \rightarrow E(X^k)$.

The next example shows that weak convergence of distribution of function does not imply the convergence of corresponding p.m.fs or p.d.fs.

Example 3: Let $\{X_n\}$ be a sequence of r.vs with p.m.f.

$$f_n(x) = P(X_n = x) = \begin{cases} 1, & \text{if } x = 2 + \frac{1}{n} \\ 0, & \text{otherwise} \end{cases}$$

Show that $F_n \xrightarrow{w} F$ does not imply $f_n \rightarrow f$.

Solution: Note that $f_n(x) \rightarrow f(x)$ as $n \rightarrow \infty$, where $f(x) = 0$ for all x .

The c.d.f. of X_n is given by

$$F_n(x) = P(X_n \leq x) = \begin{cases} 0, & x < 2 + \frac{1}{n} \\ 1, & x \geq 2 + \frac{1}{n} \end{cases}$$

which converges to

$$F(x) = \begin{cases} 0 & , x < 2 \\ 1 & , x \geq 2 \end{cases}$$

at all continuity points of F . Since F is the c.d.f. of a r.v. degenerate at $x = 2$
i.e., $P(X = 2) = 1$

$$i.e., f(x) = \begin{cases} 1, & x = 2 \\ 0, & otherwise \end{cases}$$

Thus, convergence of distribution functions does not imply the convergence of corresponding p.m.fs.

Example 4: Let $\{X_n\}$ be a sequence of r.vs with p.m.f $P(X_n = 1) = \frac{1}{n}$ and $P(X_n = 0) = 1 - \frac{1}{n}$. Then show that $X_n \xrightarrow{P} 0$.

Solution: We have $P(|X_n| > \epsilon) = \begin{cases} P(X_n = 1) = \frac{1}{n}, & 0 < \epsilon < 1 \\ 0 & , \epsilon \geq 1 \end{cases}$

It follows that $P(|X_n| > \epsilon) \rightarrow 0$ as $n \rightarrow \infty$, and we conclude that $X_n \xrightarrow{P} 0$

Example 5: Let $\{X_n\}$ be a sequence of r.vs defined by

$$P(X_n = 0) = 1 - \frac{1}{n}, P(X_n = 1) = \frac{1}{n}, n = 1, 2, \dots$$

Show that $X_n \xrightarrow{q.m} X$, where $P(X = 0) = 1$.

Solution: Consider $E(|X_n - 0|^2) = E(|X_n|^2) = E(X_n^2) = 0^2 \left(1 - \frac{1}{n}\right) + 1^2 \left(\frac{1}{n}\right)$

$$= \frac{1}{n} \rightarrow 0 \text{ as } n \rightarrow \infty$$

Thus, $X_n \xrightarrow{q.m} X$, where X is degenerate at 0.

Example 6: Let $\{X_n\}$ be a sequence of independent r.vs defined by

$$P(X_n = 0) = 1 - \frac{1}{n} \text{ and } P(X_n = 1) = \frac{1}{n}, \quad n = 1, 2, \dots$$

Show that $X_n \xrightarrow{q.m} 0$ but $X_n \not\xrightarrow{a.s} 0$

Solution: $E(|X_n - 0|^2) = E(|X_n|^2) = 0^2 \left(1 - \frac{1}{n}\right) + 1^2 \left(\frac{1}{n}\right) = \frac{1}{n} \rightarrow 0$ as $n \rightarrow \infty$

Hence $X_n \xrightarrow{q.m} 0$.

Also, $P(X_n = 0 \text{ for every } m \leq n \leq n_0) = \prod_{n=m}^{n_0} \left(1 - \frac{1}{n}\right) = \frac{m-1}{n_0}$ which converges to

zero as $n \rightarrow \infty$ for all values of m . Thus, $X_n \not\xrightarrow{a.s} 0$

Example 7: Let $\{X_n\}$ be a sequence of independent r.vs defined by

$$P(X_n = 0) = 1 - \frac{1}{n^r} \text{ and } P(X_n = n) = \frac{1}{n^r}, \quad r \geq 2, \quad n = 1, 2, \dots$$

Show that $X_n \xrightarrow{a.s} 0$ but $X_n \not\xrightarrow{r} 0$.

Solution: We have $P(X_n = 0 \text{ for } m \leq n \leq n_0) = \prod_{n=m}^{n_0} \left(1 - \frac{1}{n^r}\right)$

As $n_0 \rightarrow \infty$, the infinite product converges to some nonzero quantity, which itself converges to 1 as $m \rightarrow \infty$.

That is, $P\left[\lim_{n \rightarrow \infty} X_n = 0\right] = 1$. Therefore $X_n \xrightarrow{a.s} 0$

However, $E(|X - 0|^r) = E(|X|^r) = 0^r \left(1 - \frac{1}{n^r}\right) + n^r \times \frac{1}{n^r} = 1$

and hence $E(|X|^r) = 1$ as $n \rightarrow \infty$. Therefore, $X_n \not\xrightarrow{r} 0$

Thus, $X_n \xrightarrow{a.s} 0$ but $X_n \not\xrightarrow{r} 0$

A sufficient condition for *a. s.* convergence:

We state a sufficient condition for the *a. s.* convergence without proof which is sometimes to verify.

$$X_n \xrightarrow{a.s.} X \Leftrightarrow \lim_{n \rightarrow \infty} P \left[\bigcup_{m=n}^{\infty} |X_m - X| > \epsilon \right] = 0, \quad \forall \epsilon > 0$$

Example 8: Let $\{X_n\}$ be a sequence of r.vs with $P \left(X_n = \pm \frac{1}{n} \right) = \frac{1}{2}$. Show that $X_n \xrightarrow{r} 0$ and $X_n \xrightarrow{a.s.} 0$.

Solution: We have $E(|X_n - 0|^r) = E(|X_n|^r) = \frac{1}{n^r} \left(\frac{1}{2} \right) + \frac{1}{n^r} \left(\frac{1}{2} \right) = \frac{1}{n^r} \rightarrow 0$ as $n \rightarrow \infty$ and hence $X_n \xrightarrow{r} 0$. It follows that

$$\bigcup_{j=n}^{\infty} \{ |X_j| > \epsilon \} = \{ |X_n| > \epsilon \}$$

Choosing $n > \frac{1}{\epsilon}$, we see that

$$\begin{aligned} P \left[\bigcup_{j=n}^{\infty} \{ |X_j| > \epsilon \} \right] &= P \left(\{ |X_n| > \epsilon \} \right) \leq P \left(|X_n| > \frac{1}{n} \right) = 0 \text{ as } n \rightarrow \infty \\ \Rightarrow \lim_{n \rightarrow \infty} P \left[\bigcup_{j=n}^{\infty} \{ |X_j| > \epsilon \} \right] &= 0 \Rightarrow X_n \xrightarrow{a.s.} 0 \end{aligned}$$

Implications always valid between modes of convergence

We state the following implications always valid between modes of convergence without proof.

- 1) $X_n \xrightarrow{r} X \Rightarrow X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$
- 2) $X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$

Counter examples to implications among the modes of convergence

$$1) X_n \xrightarrow{d} X \not\Rightarrow X_n \xrightarrow{P} X \quad (\text{See P1})$$

$$2) X_n \xrightarrow{P} X \not\Rightarrow X_n \xrightarrow{r} X \quad (\text{See P2})$$

$$3) X_n \xrightarrow{P} X \not\Rightarrow X_n \xrightarrow{a.s} X \quad (\text{See P3})$$

$$4) X_n \xrightarrow{r} X \not\Rightarrow X_n \xrightarrow{a.s} X$$

$$5) X_n \xrightarrow{a.s} X \not\Rightarrow X_n \xrightarrow{r} X$$

The following theorem is known as **Slutsky's Theorem** and is very useful in finding the limiting distribution of certain r.vs. This theorem is stated without proof.

Theorem 1: Slutsky's Theorem: Let $\{X_n, Y_n\}, n = 1, 2, \dots$ be a sequence of pairs of random variables and let c be a constant. If $X_n \xrightarrow{L} X$ and $Y_n \xrightarrow{P} c$, then

$$(i) \quad X_n + Y_n \xrightarrow{L} X + c$$

$$(ii) \quad X_n Y_n \xrightarrow{L} cX$$

$$(iii) \quad \frac{X_n}{Y_n} \xrightarrow{L} \frac{X}{c} \quad \text{if } c \neq 0$$

An example presented in **P4** as an application of **Slutsky's theorem**.

4.2. Convergence of sequences of random variables

Exercise:

1. Let $X_1, X_2, \dots$ be a sequence of r.v.s with corresponding d.f.s given by

$$F_n(x) = \begin{cases} 0 & , \quad x < -n \\ \frac{x+n}{2n} & , \quad -n \leq x < n \\ 0 & , \quad x \geq n \end{cases}$$

Does F_n converge to some d.f.

2. Let $X_1, X_2, \dots$ be a i.i.d $U(0, \theta)$ r.v.s. Let $X_{(1)} = \min(X_1, X_2, \dots, X_n)$ and consider the sequence $Y_n = nX_{(1)}$. Does Y_n converge in distribution to some r.v. Y ? If so, find the d.f. of r.v. Y .

3. Let $X_1, X_2, \dots$ be i.i.d. r. vs with continuous d.f. F . Let $X_{(n)} = \max(X_1, X_2, \dots, X_n)$ and consider the sequence of r.v.s $Y_n = n[1 - F(X_{(n)})]$. Find the limiting d.f. of Y_n .

4. Let $X_1, X_2, \dots$ be a sequence of i.i.d r.v.s with common p.d.f

$$f(x, \theta) = \begin{cases} e^{-x+\theta}, & \text{if } x \geq \theta \\ 0 & , \quad \text{if } x < \theta \end{cases}$$

write $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$

(a) Show that $\bar{X}_n \xrightarrow{P} 1 + \theta$

(b) Show that $\min\{X_1, X_2, \dots, X_n\} \xrightarrow{P} \theta$

5. Let $X_1, X_2, \dots$ be i.i.d $U[0, \theta]$ r.v.s. Show that $\max\{X_1, X_2, \dots, X_n\} \xrightarrow{P} \theta$

Answers:

1. No

2. Yes. $Y_n \rightarrow Y$ where $F(y) = \begin{cases} 0 & \text{if } y < 0 \\ 1 - e^{-y/\theta} & \text{if } y \geq 0 \end{cases}$

3. $F(y) = \begin{cases} 0 & \text{if } y \leq 0 \\ 1 - e^{-y} & \text{if } y > 0 \end{cases}$

P1:

The j. p. m. f of X and X_n is given by

$X \backslash X_n$	0	0	$P(X = x)$
0	0	$\frac{1}{2}$	$\frac{1}{2}$
1	$\frac{1}{2}$	0	$\frac{1}{2}$
$P(X_n = x)$	$\frac{1}{2}$	$\frac{1}{2}$	1

Show that $X_n \xrightarrow{d} X$ but $X_n \not\xrightarrow{P} 0$

P1:

The j. p. m. f of X and X_n is given by

$X_n \backslash X$	0	1	$P(X = x)$
0	0	$\frac{1}{2}$	$\frac{1}{2}$
1	$\frac{1}{2}$	0	$\frac{1}{2}$
$P(X_n = x)$	$\frac{1}{2}$	$\frac{1}{2}$	1

Show that $X_n \xrightarrow{d} X$ but $X_n \not\xrightarrow{P} 0$

Solution:

The p.m.f of X_n is $P(X_n = 0) = \frac{1}{2} = P(X_n = 1)$ and its c.d.f is given by

$$F_n(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{1}{2} & , \quad 0 \leq x < 1 \\ 1 & , \quad x \geq 1 \end{cases}$$

The p.m.f of X is $P(X = 0) = \frac{1}{2} = P(X = 1)$ and its c.d.f is given by

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

Thus, $F_n(x) \xrightarrow{w} F(x)$ as $n \rightarrow \infty$. That is $X_n \xrightarrow{d} X$

$$\begin{aligned} \text{But } P\left[|X_n - X| > \frac{1}{2}\right] &= P[|X_n - X| = 1] \\ &= P[X_n = 0, X = 1] + P[X_n = 1, X = 0] \\ &= \frac{1}{2} + \frac{1}{2} = 1 \not\rightarrow 0 \text{ as } n \rightarrow \infty \end{aligned}$$

and hence $X_n \not\xrightarrow{P} 0$. Therefore, $X_n \xrightarrow{d} X$ but $X_n \not\xrightarrow{P} 0$

P2:

Let $\{X_n\}$ be a sequence of r.vs defined by $P(X_n = 0) = 1 - \frac{1}{n^r}$ and $P(X_n = n) = \frac{1}{n^r}$, $r > 0, n = 1, 2, \dots$.

Show that (i) $X_n \xrightarrow{P} 0$ and (ii) $X_n \not\xrightarrow{r} 0$

P2:

Let $\{X_n\}$ be a sequence of r.v.s defined by $P(X_n = 0) = 1 - \frac{1}{n^r}$ and $P(X_n = n) = \frac{1}{n^r}$, $r > 0, n = 1, 2, \dots$.

Show that (i) $X_n \xrightarrow{P} 0$ and (ii) $X_n \not\xrightarrow{r} 0$

Solution:

(i) We have

$$P(|X_n - 0| > \epsilon) = P(X_n = n) = \frac{1}{n^r}$$

Therefore $P(|X_n - 0| > \epsilon) \rightarrow 0$ as $n \rightarrow \infty$. Thus, $X_n \xrightarrow{P} 0$

(ii) We have $E(|X_n - 0|^r) = E(|X_n|^r) = 0^r \left(1 - \frac{1}{n^r}\right) + n^r \left(\frac{1}{n^r}\right) = 1$

$$\Rightarrow E(|X_n - 0|^r) = 1$$

$$\Rightarrow \lim_{n \rightarrow \infty} E(|X_n - 0|^r) = 1$$

$$\Rightarrow X_n \not\xrightarrow{r} 0$$

Thus $X_n \xrightarrow{P} 0$ but $X_n \not\xrightarrow{r} 0$

P3:

For each positive integer n there exist integers m and k (uniquely determined) such that

$$n = 2^k + m, 0 \leq m < 2^k, k = 0, 1, 2, \dots$$

Thus, for $n = 1$, we have $k = 0$ and $m = 0$; for $n = 5$, we have $k = 2$ and $m = 1$; and so on. Define r.v.s X_n for $n = 1, 2, \dots$ on $\Omega = [0, 1]$ by

$$X_n(w) = \begin{cases} 2^k, & \frac{m}{2^k} \leq w < \frac{m+1}{2^k} \\ 0, & \text{otherwise} \end{cases}$$

Show that $X_n \xrightarrow{P} 0$ but $X_n \not\xrightarrow{a.s.} 0$.

P3:

For each positive integer n there exist integers m and k (uniquely determined) such that

$$n = 2^k + m, 0 \leq m < 2^k, k = 0, 1, 2, \dots$$

Thus, for $n = 1$, we have $k = 0$ and $m = 0$; for $n = 5$, we have $k = 2$ and $m = 1$; and so on. Define r.v.s X_n for $n = 1, 2, \dots$ on $\Omega = [0, 1]$ by

$$X_n(w) = \begin{cases} 2^k, & \frac{m}{2^k} \leq w < \frac{m+1}{2^k} \\ 0, & \text{otherwise} \end{cases}$$

Show that $X_n \xrightarrow{P} 0$ but $X_n \not\xrightarrow{a.s.} 0$.

Solution:

Let the probability distribution of X_n be given by

$$P(I) = \text{length of the interval } I \subseteq \Omega.$$

Thus, $P(X_n = 2^k) = \frac{1}{2^k}$ and $P(X_n = 0) = 1 - \frac{1}{2^k}$

The limit $\lim_{n \rightarrow \infty} X_n(w) = \lim_{k \rightarrow \infty} X_n(w)$ does not exist for any $w \in \Omega$, so that X_n does not converge almost surely. But

$$P(|X_n - 0| > \epsilon) = P(|X_n| > \epsilon) = P(X_n > \epsilon)$$

$$= \begin{cases} 0, & \epsilon \geq 2^k \\ \frac{1}{2^k}, & 0 < \epsilon < 2^k \end{cases}$$

and we see that

$$P(|X_n - 0| > \epsilon) \rightarrow 0 \text{ as } n \text{ (and hence } k) \rightarrow \infty.$$

Thus, $X_n \xrightarrow{P} 0$ but $X_n \not\xrightarrow{a.s.} 0$.

P4:

Let $X_1, X_2, \dots$ be i.i.d.r.vs with common p.d.f. $N(0, 1)$. Determine the limiting distribution of the r.v.

$$W_n = \sqrt{n} \left(\frac{X_1 + \dots + X_n}{X_1^2 + \dots + X_n^2} \right)$$

P4:

Let $X_1, X_2, \dots$ be i.i.d.r.vs with common p.d.f. $N(0, 1)$. Determine the limiting distribution of the r.v.

$$W_n = \sqrt{n} \left(\frac{X_1 + \dots + X_n}{X_1^2 + \dots + X_n^2} \right)$$

Solution:

Let $U_n = \frac{1}{\sqrt{n}}(X_1 + \dots + X_n)$ and $V_n = \frac{X_1^2 + X_2^2 + \dots + X_n^2}{n}$. Then $W_n = \frac{U_n}{V_n}$

Since each $X \sim N(0, 1)$, the m.g.f. of X is given by $M_X(t) = e^{\frac{t^2}{2}}$

Then the m.g.f. of U_n is given by

$$\begin{aligned} M_{U_n}(t) &= E[e^{tU_n}] = E\left[e^{\left(\frac{t}{\sqrt{n}}\right) \sum_{i=1}^n X_i}\right] \\ &= \prod_{i=1}^n M_{X_i}\left(\frac{t}{\sqrt{n}}\right) \quad (\because X \text{ s are independent}) \\ &= \prod_{i=1}^n e^{t^2/2n} \quad (\because X \text{ s are identically distributed}) \\ &= e^{\frac{t^2}{2n}} \cdot e^{\frac{t^2}{2n}} \cdot e^{\frac{t^2}{2n}} \cdot e^{\frac{t^2}{2n}} \dots e^{\frac{t^2}{2n}} \quad (n \text{ times}) = e^{\frac{t^2}{2}} \end{aligned}$$

$\Rightarrow M_{U_n}(t) = e^{\frac{t^2}{2}}$ which is the m.g.f. of $N(0, 1)$ r.v. By uniqueness of m.g.f. ,

$U_n \sim N(0, 1)$ i.e., $U_n \xrightarrow{L} Z$, where $Z \sim N(0, 1)$.

Next, we find the m.g.f. of V_n and identify its probability distribution.

First, the m.g.f. of X^2 is given by

$$M_{X^2}(t) = E[e^{tx^2}] = \int_{-\infty}^{\infty} e^{tx^2} f(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx^2} e^{-\frac{1}{2}x^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2(1-2t)} dx$$

Let $(\sqrt{1-2t})x = y \Rightarrow dx = (1-2t)^{-\frac{1}{2}} dy$. Then

$$M_{X^2}(t) = \left\{ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2} dy \right\} (1-2t)^{-1/2} = (1-2t)^{-1/2}$$

($\because$ total probability of $N(0, 1)$ is 1)

$$\Rightarrow M_{X^2}(t) = (1-2t)^{-\frac{1}{2}}$$

$$\text{Next, } M_{V_n}(t) = E[e^{tV_n}] = E\left[e^{\left(\frac{t}{n}\right)(X_1^2 + \dots + X_n^2)}\right]$$

$$= \prod_{i=1}^n E\left[e^{\left(\frac{t}{n}\right)X_i^2}\right] \quad (\because X \text{ s are independent})$$

$$= \prod_{i=1}^n M_{X^2}\left(\frac{t}{n}\right)$$

$$= \prod_{i=1}^n \left(1 - \frac{2t}{n}\right)^{-\frac{1}{2}} = \left(1 - \frac{2t}{n}\right)^{-\frac{n}{2}}$$

$\Rightarrow M_{V_n}(t) = \left(1 - \frac{2t}{n}\right)^{-\frac{n}{2}}, t < \frac{n}{2}$, which is the m.g.f. of a Gamma distribution with two parameters $\alpha = \frac{n}{2}$ and $\beta = \frac{2}{n}$

The *p.d.f.* of Gamma variate with two parameters (α, β) is defined by

$$f(x) = \frac{1}{\sqrt{\alpha}} \frac{1}{\beta^\alpha} e^{-\frac{1}{\beta}x} x^{\alpha-1} \text{ for } x > 0, \alpha > 0, \beta > 0$$

Notation: $X \sim G(\alpha, \beta)$

If $X \sim G(\alpha, \beta)$ then its m.g.f. is given by $M_X(t) = (1 - pt)^{-\alpha}$

Mean = $E(X) = \alpha\beta$ and variance = $V(X) = \beta^2\alpha$

The variance of V_n is given by $V(V_n) = \beta^2 \alpha = \left(\frac{2}{n}\right)^2 \frac{n}{2} = \frac{2}{n}$

We have for any $\epsilon > 0$,

$$\begin{aligned} P\{|V_n - 1| > \epsilon\} &\leq \frac{V(V_n)}{\epsilon^2} \quad (\text{By chebychev's inequality}) \\ &= \frac{2}{n\epsilon^2} \rightarrow 0 \text{ as } n \rightarrow \infty \end{aligned}$$

Thus, $V_n \xrightarrow{P} 1$. We have thus shown that $U_n \xrightarrow{L} Z$ and $V_n \xrightarrow{P} 1$

It follows by Slutsky's theorem (iii)

$$W_n = \frac{U_n}{V_n} \xrightarrow{L} \frac{Z}{1} = Z, \text{ where } Z \text{ is } N(0, 1)$$

Hence, $W_n \sim N(0, 1)$.

4.3

Weak Law of Large Numbers

Let $\{X_n\}$ be a sequence of r.vs and let $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ be the mean of first n r.vs. The

weak laws deal with *limits of probabilities involving* $\bar{X}_n$. The strong laws deal with *probabilities involving limits of* $\bar{X}_n$.

Definition of Weak Law of Large Numbers

A sequence $\{X_n\}$ of r.vs is said to satisfy the **Weak Law of Large Numbers (WLLN)** if

$$\lim_{n \rightarrow \infty} P \left[\left| \frac{S_n}{n} - E \left(\frac{S_n}{n} \right) \right| < \epsilon \right] = 1$$

for any $\epsilon > 0$, where $S_n = \sum_{i=1}^n X_i$, i. e., $\frac{S_n}{n} \xrightarrow{P} E \left(\frac{S_n}{n} \right)$

Theorem1: Let $\{X_n\}$ be a sequence of r.vs and let $S_n = X_1 + \dots + X_n$ with $B_n = V(S_n) < \infty$. If $\frac{B_n}{n^2} \rightarrow 0$ as $n \rightarrow \infty$, then for any $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} P \left[\left| \frac{S_n}{n} - E \left(\frac{S_n}{n} \right) \right| < \epsilon \right] = 1$$

i. e., $\{X_n\}$ satisfies WLLN.

Proof: On applying Chebychev's inequality to the variable $\frac{S_n}{n}$, we have

$$P \left[\left| \frac{S_n}{n} - E \left(\frac{S_n}{n} \right) \right| \geq \epsilon \right] \leq \frac{V \left(\frac{S_n}{n} \right)}{\epsilon^2} = \frac{V(S_n)}{n^2 \epsilon^2} = \frac{B_n}{n^2 \epsilon^2} \rightarrow 0$$

as $n \rightarrow \infty$. Thus,

$$\lim_{n \rightarrow \infty} P \left[\left| \frac{S_n}{n} - E \left(\frac{S_n}{n} \right) \right| \geq \epsilon \right] = 0 \Rightarrow \lim_{n \rightarrow \infty} P \left[\left| \frac{S_n}{n} - E \left(\frac{S_n}{n} \right) \right| < \epsilon \right] = 1$$

$\Rightarrow \{X_n\}$ satisfies WLLN.

Corollary 1: Let $\{X_n\}$ be a sequence of r.v.s, $\bar{X}_n = \frac{S_n}{n}$ and $\mu = E\left(\frac{S_n}{n}\right)$.

If $\frac{B_n}{n^2} \rightarrow 0$ as $n \rightarrow \infty$, then

$$\lim_{n \rightarrow \infty} P[\bar{X}_n \leq k] = \begin{cases} 0 & , \text{if } k < \mu \\ 1 & , \text{if } k > \mu \end{cases}$$

Proof: Since WLLN holds for $\{X_n\}$, we have

$$\lim_{n \rightarrow \infty} P[|\bar{X}_n - \mu| < \epsilon] = 1 \Rightarrow \lim_{n \rightarrow \infty} P[|\bar{X}_n - \mu| \geq \epsilon] = 0 \quad \dots (1)$$

Since $\{\bar{X}_n \leq \mu - \epsilon\} \subset \{|\bar{X}_n - \mu| \geq \epsilon\}$, we have

$$\begin{aligned} P(\bar{X}_n \leq \mu - \epsilon) &\leq P(|\bar{X}_n - \mu| \geq \epsilon) \\ \Rightarrow \lim_{n \rightarrow \infty} P(\bar{X}_n \leq \mu - \epsilon) &\leq \lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| \geq \epsilon) \\ \Rightarrow \lim_{n \rightarrow \infty} P(\bar{X}_n \leq \mu - \epsilon) &= 0 \\ \Rightarrow \lim_{n \rightarrow \infty} P(\bar{X}_n \leq k) &= 0, \text{ where } k = \mu - \epsilon, \text{ i.e., } k < \mu \text{ since } \epsilon > 0 \\ \Rightarrow \lim_{n \rightarrow \infty} P(\bar{X}_n \leq k) &= 0 \text{ if } k < \mu \end{aligned}$$

Further, $P(\bar{X}_n \leq \mu + \epsilon) + P(|\bar{X}_n - \mu| > \epsilon) \geq 1$, since the region is larger than sample space covered.

$$\begin{aligned} \Rightarrow \lim_{n \rightarrow \infty} P(\bar{X}_n \leq \mu + \epsilon) &\geq 1 \quad \left(\because \lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| > \epsilon) = 0 \right) \\ \Rightarrow \lim_{n \rightarrow \infty} P(\bar{X}_n \leq \mu + \epsilon) &= 1 \\ \Rightarrow \lim_{n \rightarrow \infty} P(\bar{X}_n \leq k) &= 1 \text{ where } k = \mu + \epsilon \text{ i.e., } k > \mu \text{ since } \epsilon > 0 \\ \Rightarrow \lim_{n \rightarrow \infty} P(\bar{X}_n \leq k) &= 1 \text{ if } k > \mu \end{aligned}$$

Thus, $\lim_{n \rightarrow \infty} P(\bar{X}_n \leq k) = \begin{cases} 0, & k < \mu \\ 1, & k > \mu \end{cases}$

Variations of the WLLN

The following are some special cases of Theorem 1 which are stated without proof.

Theorem 2: (Bernoulli's WLLN)

Let $\{X_n\}$ be a sequence of Bernoulli trials with probability of success equal to p . If S_n is the number of successes in n trials, then

$$\lim_{n \rightarrow \infty} P \left[\left| \frac{S_n - np}{n} \right| < \epsilon \right] = 1, \quad \forall \epsilon > 0$$

Theorem 3: (Khinchine's WLLN)

Let $\{X_n\}$ be a sequence of i.i.d.r.vs with $E(X_i) = \mu < \infty, i = 1, 2, \dots$, then the WLLN holds i.e.,

$$\lim_{n \rightarrow \infty} P \left[\left| \frac{S_n}{n} - \mu \right| > \epsilon \right] = 0$$

Theorem 4: (Bernstein's WLLN)

Let $\{X_n\}$ be a sequence of random variables for which $\text{var}(X_n) = \sigma_n^2 < k, \forall i$, where k is independent of n . If $\sigma_{ij} = \text{cov}(X_i, X_j) \rightarrow 0$ as $|i - j| \rightarrow \infty$ (Asymptotic uncorrelatedness) then the WLLN holds.

Example 1: Let $\{X_n\}$ be i.i.d.r.vs with mean μ and variance σ^2 , if

$$\frac{X_1^2 + X_2^2 + \dots + X_n^2}{n} \xrightarrow{P} c$$

as $n \rightarrow \infty$ for some constant $c(0 \leq c < \infty)$, then find c .

Solution: Here $E(X_i) = \mu$ and $V(X_i) = \sigma^2 \forall i$.

Let $S_n = X_1^2 + X_2^2 + \dots + X_n^2$. Then

$$E(S_n) = nE(X_1^2) \quad (\because X\text{s are i.i.d.r.vs})$$

$$= n \left[V(X_1) + (E(X_1))^2 \right]$$

$$\Rightarrow E(S_n) = n(\sigma^2 + \mu^2)$$

Since $E(X^2) = V(X) + (E(X))^2 = \sigma^2 + \mu^2$ exists for each X^2 in S_n , by Khinchine's WLLN, we have

$$\frac{S_n}{n} = \frac{X_1^2 + X_2^2 + \dots + X_n^2}{n} \quad E(X_1^2) = \mu^2 + \sigma^2$$

Thus, $c = \mu^2 + \sigma^2$.

Example 2: If the i.i.d. r.vs $X_k (k = 1, 2, \dots)$ assume the value $2^{r-2 \ln r}$ with probability $\frac{1}{2^r}$, examine if the WLLN holds for the sequence $\{X_k\}$.

Solution:

$$\begin{aligned} E(X_k) &= \sum_{r=1}^{\infty} 2^{r-2 \ln r} \cdot \frac{1}{2^r} = \sum_{r=1}^{\infty} (2^{-2})^{\ln r} = \sum_{r=1}^{\infty} \left(\frac{1}{4}\right)^{\ln r} \\ &= \sum_{r=1}^{\infty} (r)^{\ln\left(\frac{1}{4}\right)} \left(\because a^{\ln n} = n^{\ln a}\right) \\ &= \sum_{r=1}^{\infty} \left(\frac{1}{r}\right)^{\ln 4} \text{ converges since } \ln 4 = 1.39 > 1 \quad \left(\sum_{n=1}^{\infty} \frac{1}{n^p} \text{ converges if } p > 1 \right) \end{aligned}$$

Thus $E(X_k) < \infty$

Since $\{X_k\}$ are i.i.d.r.vs with $E(X_k) < \infty$, the WLLN holds for the sequence, by Khinchine's theorem.

Example 3: Let $\{X_n\}$ be a sequence of i.i.d $U(0, 1)$ r.vs. For the geometric mean $G_n = (X_1 \cdot X_2 \cdot \dots \cdot X_n)^{\frac{1}{n}}$, show that $G_n \xrightarrow{P} c$ where c is some constant. Find c .

Solution: Let $Y = -\ln X$ where $X \sim U(0,1)$. The c.d.f. of Y is given by

$$F_Y(y) = P(Y \leq y) = P(-\ln X \leq y) = P(X \geq e^{-y}) = \int_{e^{-y}}^1 1 \, dx = 1 - e^{-y}$$

$\Rightarrow F_Y(y) = 1 - e^{-y}$ and the p.d.f of Y is given by

$$f_Y(y) = \frac{d}{dx}(F_Y(y)) = e^{-y} \text{ for } y > 0.$$

Then $E(Y) = V(Y) = 1$.

Thus, the sequence $\{Y_n\}$ is i.i.d with finite mean $E(Y_n) = 1$. Hence, by Khinchine's WLLN

$$\sum_{i=1}^n \frac{Y_i}{n} \xrightarrow{P} E(Y_1) = 1 \quad \dots (1)$$

$$\text{But } \ln G_n = \sum_{i=1}^n \ln \frac{X_i}{n} = - \sum_{i=1}^n \frac{Y_i}{n}$$

$$\Rightarrow \sum_{i=1}^n \frac{Y_i}{n} = - \ln G_n \quad \dots (2)$$

From (1) and (2), we have

$$- \ln G_n \xrightarrow{P} 1 \text{ i.e., } G_n \xrightarrow{P} e^{-1}$$

Thus, $c = \frac{1}{e}$.

Example 4: Let X_i can have only two values i^α and $-i^\alpha$ with equal probabilities.

If $\{X_i\}$ is a sequence of independent r.vs, then show that WLLN holds if $\alpha < \frac{1}{2}$.

Solution: Here $E(X_i) = i^\alpha \frac{1}{2} - i^\alpha \frac{1}{2} = 0$ and

$$V(X_i) = E(X_i^2) = i^{2\alpha} \frac{1}{2} + i^{2\alpha} \frac{1}{2} = i^{2\alpha}$$

Let $S_n = \sum_{k=1}^n X_k$. Then

$$B_n = V(S_n) = \sum_{i=1}^n V(X_i) \quad (\because X_i \text{ s are independent})$$

$$= \sum_{i=1}^n i^{2\alpha} = 1^{2\alpha} + 2^{2\alpha} + \dots + n^{2\alpha}$$

$$= \int_0^n x^{2\alpha} dx \quad (\text{Euler - Maclaurion formula})$$

$$\Rightarrow B_n = \frac{n^{2\alpha+1}}{2\alpha+1} \Rightarrow \frac{B_n}{n^2} = \frac{n^{2\alpha+1}}{2\alpha+1} \rightarrow 0 \text{ as } n \rightarrow \infty \text{ if } \alpha < \frac{1}{2}$$

Thus, $\frac{B_n}{n^2} \rightarrow 0$ as $n \rightarrow \infty$ when $\alpha < \frac{1}{2}$

Therefore, $\{X_n\}$ holds WLLN when $\alpha < \frac{1}{2}$.

4.3. Weak Law of Large Numbers

Exercise:

1) For the following sequence of independent r.v.s, does the *WLLN* hold?

a) $P(X_k = \pm k) = \frac{1}{2\sqrt{k}}, P(X_k = 0) = 1 - \frac{1}{\sqrt{k}}$

b) $P(X_k = \pm 2^k) = \frac{1}{2^{2k+1}}, P(X_k = 0) = 1 - \frac{1}{2^{2k}}$

c) $P\left(X_k = \pm \frac{1}{k}\right) = \frac{1}{2}$

d) $P(X_k = \pm \sqrt{k}) = \frac{1}{2}$

2) Examine if *WLLN* holds for the sequence $\{X_i\}$ of i.i.d.r.v.s with

$$P[x_i = (-1)^{k-1} \cdot k] = \frac{6}{\pi^2 k^2}, \quad k = 1, 2, 3, \dots, \quad i = 1, 2, \dots$$

Hint: $\log_e 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \dots$

3) Let $X_1, X_2, \dots, X_n$ be jointly normal with $E(X_i) = 0, E(X_i)^2 = 1$ for all i and

$$\text{cov}(X_i, X_j) = \begin{cases} \rho, & |j - i| = 1 \\ 0, & \text{otherwise} \end{cases}$$

Examine if *WLLN* holds for the sequence $\{X_n\}$.

ANSWERS

1)

a) No

b) Yes

c) Yes

d) No

e) No

2) Yes

3) Yes

P1:

A variate X_k has the distribution

$$P(X_k = 0) = 1 - \left(\frac{2}{3^{2k+2}} \right), \quad P(X_k = \pm 3^k) = 3^{-(2k+2)}$$

If $\{X_k\}$ is a sequence of independent r.vs, then show that $\{X_k\}$ obeys *WLLN*

P1:

A variate X_k has the distribution

$$P(X_k = 0) = 1 - \left(\frac{2}{3^{2k+2}}\right), \quad P(X_k = \pm 3^k) = 3^{-(2k+2)}$$

If $\{X_k\}$ is a sequence of independent r.v.s, then show that $\{X_k\}$ obeys *WLLN*

Solution:

Here $E(X_k) = 3^k \cdot 3^{-(2k+2)} - 3^k \cdot 3^{-(2k+2)} = 0$ and

$$\begin{aligned} V(X_k) &= E(X_k^2) = 3^{2k} 3^{-(2k+2)} + 3^{2k} 3^{-(2k+2)} \\ &= 2 \cdot 3^{2k} 3^{-(2k+1)} = \frac{2}{9} \end{aligned}$$

Let $S_n = \sum_{k=1}^n X_k$. Then

$$B_n = V(S_n) = \sum_{k=1}^n V(X_k) \quad (\because X_k \text{ s are independent})$$

$$= \sum_{k=1}^n \frac{2}{9} = \frac{2n}{9}$$

and hence $\frac{B_n}{n^2} = \frac{2n}{9n^2} = \frac{2}{9n} \rightarrow 0$ as $n \rightarrow \infty$

Since $\frac{B_n}{n^2} \rightarrow 0$ as $n \rightarrow \infty$, $\{X_k\}$ obeys the *WLLN*.

P2:

Let $\{X_n\}$ be a sequence of independent r.vs such that

$$P(X_n = \pm 1) = \frac{1}{2}(1 - 2^{-n}), \quad P(X_n = \pm 2^{-n}) = 2^{-n-1}$$

Does the *WLLN* hold for this sequence?

P2:

Let $\{X_n\}$ be a sequence of independent r.vs such that

$$P(X_n = \pm 1) = \frac{1}{2}(1 - 2^{-n}), \quad P(X_n = \pm 2^{-n}) = 2^{-n-1}$$

Does the *WLLN* hold for this sequence?

Solution:

$$\text{Here } E(X_n) = \frac{1}{2}(1 - 2^{-n}) - \frac{1}{2}(1 - 2^{-n}) + 2^{-n} \cdot 2^{-n-1} - 2^{-n} \cdot 2^{-n-1} = 0$$

$$\begin{aligned} \text{and } V(X_n) &= E(X_n^2) = \frac{1}{2}(1 - 2^{-n}) + \frac{1}{2}(1 - 2^{-n}) + 2^{-2n} \cdot 2^{-n-1} + 2^{-2n} \cdot 2^{-n-1} \\ &= (1 - 2^{-n}) + 2^{-3n} \end{aligned}$$

$$\Rightarrow V(X_n) = 2^{-3n} - 2^{-n} + 1$$

Let $S_n = \sum_{i=1}^n X_i$. Then

$$V(S_n) = \sum_{i=1}^n V(X_i) \quad (\because X_i \text{ s are independent})$$

$$= \sum_{i=1}^n (2^{-3i} - 2^{-i} + 1) = \sum_{i=1}^n 2^{-3i} - \sum_{i=1}^n 2^{-i} + n$$

$$\Rightarrow B_n = V(S_n) = \frac{1}{7}(1 - 8^{-n}) - (1 - 2^{-n}) + n = -\frac{1}{7 \cdot 8^n} + \frac{1}{2^n} - \frac{6}{7} + n$$

$$\Rightarrow \frac{B_n}{n^2} = -\frac{1}{7 \cdot n^2 \cdot 8^n} + \frac{1}{n^2 \cdot 2^n} - \frac{6}{7 \cdot n^2} + \frac{1}{n} \rightarrow 0 \text{ as } n \rightarrow \infty$$

Since $\frac{B_n}{n^2} \rightarrow 0$ as $n \rightarrow \infty$, *WLLN* holds for the given sequence.

P3:

Show that *WLLN* holds to the mean of a sequence of independent r.v s X_k specified by

$$P(X_k = \pm\sqrt{\ln k}) = \frac{1}{2}$$

P3:

Show that *WLLN* holds to the mean of a sequence of independent r.v s X_k specified by

$$P(X_k = \pm\sqrt{\ln k}) = \frac{1}{2}$$

Solution:

Here $E(X_k) = \sqrt{\ln k} \cdot \frac{1}{2} - \sqrt{\ln k} \cdot \frac{1}{2} = 0$ and

$$V(X_k) = E(X_k^2) = \frac{1}{2} \ln k + \frac{1}{2} \ln k = \ln k$$

Let $S_n = \sum_{k=1}^n X_k$. Then

$$V(S_n) = \sum_{k=1}^n V(X_k) \quad (\because X_k \text{'s are independent})$$

$$= \sum_{k=1}^n \ln k = \ln 1 + \ln 2 + \dots + \ln n = \ln(n!)$$

$$\Rightarrow V(S_n) = B_n = \ln(n!) \quad \dots (1)$$

By Stirling's approximation

$$n! = \sqrt{2\pi} e^{-n} n^{n+\frac{1}{2}} \quad \dots (2)$$

$$\begin{aligned} \text{We have } \frac{B_n}{n^2} &= \frac{\ln(\sqrt{2\pi} e^{-n} n^{n+\frac{1}{2}})}{n^2} = \frac{-n + (n+\frac{1}{2}) \ln n + \ln \sqrt{2\pi}}{n^2} \\ &= \frac{(\ln \sqrt{2\pi} - n)}{n^2} + \frac{1}{n} \left(1 + \frac{1}{2n}\right) \ln n \\ &= \frac{\ln \sqrt{2\pi} - n}{n^2} + \left(1 + \frac{1}{2n}\right) \ln(n^{1/n}) \rightarrow 0 \text{ as } n \rightarrow \infty \quad \left(\because \lim_{n \rightarrow \infty} n^{1/n} = 1\right) \end{aligned}$$

Hence $\lim_{n \rightarrow \infty} \frac{B_n}{n^2} = 0$. Thus, the sequence $\{X_n\}$ holds *WLLN*.

P4:

The variates $X_1, X_2, \dots$ have equal expectations and finite variations. Is *WLLN* applicable to this sequence if all covariances σ_{ij} are negative?

P4:

The variates $X_1, X_2, \dots$ have equal expectations and finite variations. Is *WLLN* applicable to this sequence if all covariances σ_{ij} are negative?

Solution:

We have

$$\begin{aligned} V(aX + bY) &= a^2V(X) + b^2V(Y) + 2ab \operatorname{cov}(X, Y) \\ &= a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_{XY} \\ &\leq a^2\sigma_X^2 + b^2\sigma_Y^2 \quad (\because \sigma_{XY} < 0) \\ \Rightarrow 0 \leq V(aX + bY) &\leq a^2\sigma_X^2 + b^2\sigma_Y^2 \\ \therefore 0 \leq V\left(\frac{X_1 + \dots + X_n}{n}\right) &\leq \frac{1}{n^2} \sum_{i=1}^n V(X_i) < \frac{A}{n} \quad \dots(1) \end{aligned}$$

where A is the upper bound of $V(X_i) \forall i = 1, 2, \dots, n$

$$\text{Now, } \frac{B_n}{n^2} = V\left(\frac{X_1 + \dots + X_n}{n}\right) < \frac{A}{n} \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (\text{from (1)})$$

Thus, *WLLN* is applicable to this sequence.

4.4

Strong Law of Large Numbers

Definition: A sequence of r.vs $\{X_n\}$ is said to satisfy the **strong law of large numbers (SLLN)** if

$$\left[\frac{S_n - E(S_n)}{n} \right] \xrightarrow{a.s} 0 \text{ as } n \rightarrow \infty$$

We state the following theorems without proof which are useful in checking whether a given *sequence satisfies SLLN or not*.

Theorem1: (Kolmogorov's SLLN)

This theorem is helpful when the r.vs in the sequence are *independent but not identically distributed*.

Statement: Let $\{X_n\}$ be a sequence of independent r.vs with $E(X_i) = \mu$ and $V(X_i) = \sigma_i^2 < \infty$ for $i = 1, 2, \dots$. If $\sum_{k=1}^{\infty} \frac{\sigma_k^2}{k^2} < \infty$, then the SLLN holds for the sequence $\{X_n\}$.

Theorem 2:

This theorem is helpful when the r.vs in the sequence are independent and identically distributed (i.i.d).

Statement: The sequence $\{X_n\}$ of i.i.d.r.vs holds SLLN iff $E(X_n)$ exists.

Theorem 3: (Borel's SLLN):

This theorem is helpful when the sequence consists of *Bernoulli trials*.

Statement: For a sequence of Bernoulli trials with constant probability of success, the SLLN holds.

Example 1: Let $\{X_n\}$ be a sequence of independent random variables with p.m.f. given by

$$P(X_n = \pm 2^n) = \frac{1}{2^{(2n+1)}}, P(X_n = 0) = 1 - \frac{1}{2^{2n}}$$

Does the SLLN hold for $\{X_n\}$?

Solution: We have $E(X_n) = 2^n \frac{1}{2^{2n+1}} - 2^n \frac{1}{2^{2n+1}} = 0$ and

$$\sigma_n^2 = V(X_n) = E(X_n^2) = 2^{2n} \frac{1}{2^{2n+1}} + 2^{2n} \frac{1}{2^{2n+1}} = 1$$

Further, $\sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2}$ converges ($\because \sum_{n=1}^{\infty} \frac{1}{n^p}$ converges if $p > 1$).

Hence, the SLLN holds for $\{X_n\}$.

Example 2: For what value of α does the SLLN hold for the sequence

$$P(X_k = \pm k^\alpha) = \frac{1}{2}$$

Solution: We have $E(X_k) = k^\alpha \frac{1}{2} - k^\alpha \frac{1}{2} = 0$ and

$$\sigma_k^2 = V(X_k) = E(X_k^2) = k^{2\alpha} \frac{1}{2} + k^{2\alpha} \frac{1}{2} = k^{2\alpha}$$

Further, $\sum_{k=1}^{\infty} \frac{\sigma_k^2}{k^2} = \sum_{k=1}^{\infty} \frac{k^{2\alpha}}{k^2} = \sum_{k=1}^{\infty} \frac{1}{k^{2-2\alpha}}$ converges if $2 - 2\alpha > 1$

($\because \sum_{k=1}^{\infty} \frac{1}{k^p}$ converges if $p > 1$).

$$\Rightarrow 2\alpha < 1 \Rightarrow \alpha < \frac{1}{2}$$

Thus, SLLN holds if $\alpha < \frac{1}{2}$.

Example 3: Let $\{X_n\}$ be a sequence of independent r.vs with p.m.f. given by

$$P\left(X_n = \pm \frac{1}{n}\right) = \frac{1}{2}$$

Check whether SLLN holds for $\{X_n\}$ or not.

Solution: We have $E(X_n) = \frac{1}{n} \frac{1}{2} - \frac{1}{n} \frac{1}{2} = 0$ and

$$\sigma_n^2 = V(X_n) = E(X_n^2) = \frac{1}{n^2} \frac{1}{2} + \frac{1}{n^2} \frac{1}{2} = \frac{1}{n^2}$$

Further, $\sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^4}$ converges ($\because \sum_{k=1}^{\infty} \frac{1}{k^p}$ converges if $p > 1$).

Therefore, $\{X_n\}$ obeys SLLN.

Example 4: Let $\{X_n\}$ be a sequence of independent r.vs with p.m.f. given by

$$P(X_k = \pm 2^{-k}) = \frac{1}{2}$$

Check whether SLLN holds or not.

Solution: Here $E(X_k) = 2^{-k} \frac{1}{2} - 2^{-k} \frac{1}{2} = 0$ and

$$\sigma_k^2 = V(X_k) = E(X_k^2) = 2^{-2k} \frac{1}{2} + 2^{-2k} \frac{1}{2} = 2^{-2k}$$

Further $\sum_{k=1}^{\infty} \frac{\sigma_k^2}{k^2} = \sum_{k=1}^{\infty} 2^{-2k} \frac{1}{k^2} = \sum_{k=1}^{\infty} \frac{1}{k^2 2^{2k}}$ converges. Therefore, $\{X_n\}$ obeys the SLLN.

Example 5: Let $\{X_n\}$ be i.i.d.r.vs with mean μ and variance σ^2 and as $n \rightarrow \infty$,

$$\frac{X_1^2 + \dots + X_n^2}{n} \xrightarrow{a.s} c$$

for some constant c ($0 \leq c < \infty$), then find c .

Solution: Here $E(X_i) = \mu$ and $V(X_i) = \sigma^2 \forall i$.

Let $S_n = X_1^2 + \dots + X_n^2$. Then

$$E(S_n) = nE(X_1^2) = n[V(X_1) + (E(X_1))^2] = n(\sigma^2 + \mu^2)$$

$$\Rightarrow E(S_n) = n(\sigma^2 + \mu^2)$$

$$\Rightarrow E\left(\frac{S_n}{n}\right) = \sigma^2 + \mu^2$$

By Theorem 2,

$$\frac{S_n}{n} \xrightarrow{a.s} E\left(\frac{S_n}{n}\right) = (\sigma^2 + \mu^2)$$

$$\Rightarrow \frac{X_1^2 + \dots + X_n^2}{n} \xrightarrow{a.s} c, \text{ where } c = \sigma^2 + \mu^2.$$

Example 6: If the i.i.d.r.vs $\{X_n\}$ assume the value $2^{r-2 \ln r}$ with probability $\frac{1}{2^r}$ for $r = 1, 2, \dots$, examine if the SLLN holds for the sequence $\{X_n\}$.

Solution: By Theorem 2, SLLN holds for i.i.d.r.vs $\{X_n\}$ if $E(X_k)$ exists $\forall k$.

Here we have to verify whether $E(X_k)$ is finite or not.

We have

$$\begin{aligned} E(X_k) &= \sum_{r=1}^{\infty} 2^{r-2 \ln r} \frac{1}{2^r} = \sum_{r=1}^{\infty} 2^{-2 \ln r} = \sum_{r=1}^{\infty} \left(\frac{1}{4}\right)^{\ln r} \\ &= \sum_{r=1}^{\infty} r^{\ln\left(\frac{1}{4}\right)} \quad (\because a^{\ln n} = n^{\ln a}) \end{aligned}$$

$$= \sum_{r=1}^{\infty} \left(\frac{1}{r}\right)^{\ln 4} = \sum_{r=1}^{\infty} \frac{1}{r^{\ln 4}} \text{ where } \ln 4 = 1.39 > 1$$

which converges.

Thus, $E(X)$ is finite and hence the SLLN holds for $\{X_n\}$.

4.4. Strong Law of Large Numbers

Exercise:

For the following sequences of independent r.v.s, does the SLLN hold or not?

1. $P(X_k = \pm 2^{k+1}) = 2^{-(k+3)}, P(X_k = 0) = 1 - 2^{-(k+2)}$

2. $P(X_k = \pm 2^{-k}) = \frac{1}{2^{k+1}}, P(X_k = \pm 1) = \frac{1}{2}(1 - 2^{-k})$

3. Examine if SLLN holds for the sequence $\{X_i\}$ of i.i.d.r.vs with

$$P[X_i = (-1)^{k-1}k] = \frac{6}{\pi^2 k^2}, k = 1, 2, 3, \dots, i = 1, 2, \dots$$

4. Let $X_1, X_2, \dots, X_n$ be jointly normal with $E(X_i) = 0, E(X_i^2) = 1$ for all i and

$$\text{cov}(X_i, X_j) = \begin{cases} \rho, & |j - i| = 1 \\ 0, & \text{otherwise} \end{cases}$$

Examine if SLLN holds for the sequence $\{X_i\}$.

Answer

1. No
2. Yes
3. Yes
4. Yes

P1:

Investigate the a.s. convergence of $\{\overline{X_n}\}$ to 0 where X_n s are independent and

$$P(X_n = 0) = 1 - \frac{1}{n}, P(X_n = \pm 1) = \frac{1}{2n}$$

P1:

Investigate the a.s. convergence of $\{\bar{X}_n\}$ to 0 where X_n s are independent and

$$P(X_n = 0) = 1 - \frac{1}{n}, P(X_n = \pm 1) = \frac{1}{2n}$$

Solution:

Here $E(X_n) = 0 \left(1 - \frac{1}{n}\right) + \frac{1}{2n} - \frac{1}{2n} = 0$ and

$$\sigma_n^2 = V(X_n) = E(X_n^2) = 0^2 \left(1 - \frac{1}{n}\right) + 1^2 \cdot \frac{1}{2n} + 1^2 \cdot \frac{1}{2n} = \frac{1}{n}$$

Thus, we have

$$\sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n} \cdot \frac{1}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^3} \text{ is convergent and by SLLN } \frac{S_n}{n} \xrightarrow{a.s.} E\left(\frac{S_n}{n}\right)$$

$$\text{But } E\left(\frac{S_n}{n}\right) = \frac{1}{n} E\left(\sum_{k=1}^n X_k\right) = \frac{1}{n} \sum_{k=1}^n E(X_k) = 0 \text{ and } \frac{S_n}{n} = \bar{X}_n$$

Therefore, $\bar{X}_n \xrightarrow{a.s.} 0$

P2:

A variate X_k has the distribution

$$P(X_k = 0) = 1 - \left(\frac{2}{3^{2k+2}}\right), P(X_k = \pm 3^k) = 3^{-(2k+2)}$$

If $\{X_k\}$ is a sequence of independent r.vs, then show that $\{X_k\}$ obeys *SLLN*.

P2:

A variate X_k has the distribution

$$P(X_k = 0) = 1 - \left(\frac{2}{3^{2k+2}}\right), P(X_k = \pm 3^k) = 3^{-(2k+2)}$$

If $\{X_k\}$ is a sequence of independent r.vs, then show that $\{X_k\}$ obeys *SLLN*.

Solution:

Here $E(X_k) = 3^k \cdot 3^{-(2k+2)} - 3^k \cdot 3^{-(2k+2)} = 0$ and

$$\sigma_k^2 = V(X_k) = E(X_k^2) = 3^{2k} 3^{-(2k+2)} + 3^{2k} 3^{-(2k+2)} = \frac{2}{9}$$

Then we have

$$\sum_{k=1}^{\infty} \frac{\sigma_k^2}{k^2} = \frac{2}{9} \sum_{k=1}^{\infty} \frac{1}{k^2} \quad \text{Converges.}$$

Thus $\{X_n\}$ obeys the *SLLN*.

P3:

Let $\{X_n\}$ be a sequence of independent r.vs such that

$$P(X_n = \pm 1) = \frac{1}{2}(1 - 2^{-n}), P(X_n = \pm 2^{-n}) = 2^{-n-1}$$

Does the *SLLN* hold for this sequence?

P3:

Let $\{X_n\}$ be a sequence of independent r.vs such that

$$P(X_n = \pm 1) = \frac{1}{2}(1 - 2^{-n}), P(X_n = \pm 2^{-n}) = 2^{-n-1}$$

Does the *SLLN* hold for this sequence?

Solution:

$$\text{Here } E(X_n) = 1 \cdot \frac{1}{2}(1 - 2^{-n}) - 1 \cdot \frac{1}{2}(1 - 2^{-n}) + 2^{-n}2^{-n-1} - 2^{-n}2^{-n-1} = 0$$

$$\begin{aligned} \text{and } \sigma_n^2 = V(X_n) = E(X_n^2) &= 1^2 \frac{1}{2}(1 - 2^{-n}) + 1^2 \frac{1}{2}(1 - 2^{-n}) + 2^{-2n}2^{-2n-1} \\ &\quad + 2^{-2n}2^{-n-1} \end{aligned}$$

$$= 2 \cdot \frac{1}{2}(1 - 2^{-n}) + 2 \cdot 2^{-2n}2^{-n-1}$$

$$\Rightarrow \sigma_n^2 = 1 - 2^{-n} + 2^{-3n}$$

$$\text{Further, we have } \sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} - \sum_{n=1}^{\infty} \frac{1}{2^n n^2} + \sum_{n=1}^{\infty} \frac{1}{2^{3n} n^2} \text{ converges.}$$

Thus $\{X_n\}$ obeys *SLLN*.

4.5

Central Limit Theorem

Let $\{X_n\}$ be a sequence of independent random variables. Let $S_n = \sum_{i=1}^n X_i$. In laws of large numbers we considered convergence of $\frac{S_n}{n}$ to $E\left(\frac{S_n}{n}\right)$ which is a constant either in *probability* (in case of WLLN) or *almost surely* (in case of SLLN). Here we consider some different situations, namely, $\frac{S_n}{n} \xrightarrow{d} Z$, where Z is a normal variate. If the sequence $\frac{S_n}{n} \xrightarrow{d} Z$, $\frac{S_n}{n}$ is said to follow the **central limit theorem** (CLT) or **normal convergence**. In this module we consider different Central Limit Theorems.

Definition: A sequence of independent r.vs $\{X_i\}$ with mean $E(X_i) = \mu_i$ and $V(X_i) = \sigma_i^2 \forall i$ is said to follow **Central Limit Theorem** (CLT) under certain conditions, if the random variable $S_n = X_1 + X_2 + \dots + X_n$ is **asymptotically normal (AN)** with mean μ and variance σ^2 where $\mu = \sum_{i=1}^n \mu_i$ and $\sigma^2 = \sum_{i=1}^n \sigma_i^2$.

Notation: $S_n \sim AN(\mu, \sigma^2)$. Read as S_n follows **asymptotically normal** with mean μ and variance σ^2 .

Note:

1. S_n is asymptotically normal means S_n follows normal distribution as $n \rightarrow \infty$.
2. If $Z_n = \frac{(S_n - \mu)}{\sigma}$, then Z_n follows asymptotically standard normal with mean 0 and variance 1 and we write $Z_n \sim AN(0, 1)$.

Variations of the CLT

The following are some variations of the CLT which are stated without proof.

Theorem 1 (De Moivre-Laplace CLT) : If $\{X_n\}$ is a sequence of Bernoulli trials with constant probability of success equal to p , then the distribution of the r.v. $S_n = X_1 + \cdots + X_n$ where X_i 's are independent, is asymptotically normal (i. e., S_n is $AN(np, np(1 - p))$)

Theorem 2 (Lindeberg-Levy CLT) : This CLT theorem is for i.i.d.r.vs.

If $\{X_i\}$ is a sequence of i.i.d.r.vs with mean $E(X_i) = \mu_1$ and variance $V(X_i) = \sigma_1^2$ for all i , then the sum $S_n = X_1 + \cdots + X_n$ is asymptotically normal with mean $\mu = n\mu_1$ and variance $\sigma^2 = n\sigma_1^2$.

Theorem 3 (Liapounoff's CLT): This CLT theorem is for independent but not identically distributed random variables.

Let $\{X_i\}$ be a sequence of independent random variables with mean $E(X_i) = \mu_i$ and variance $V(X_i) = \sigma_i^2 \forall i$. Let us assume that third absolute moment, say ρ_i^3 of X_i about its mean exists i. e., $\rho_i^3 = E\{|X_i - \mu_i|^3\}$ for $i = 1, 2, \dots, n$ is finite. Let $\rho^3 = \sum_{i=1}^n \rho_i^3$, $\mu = \sum_{i=1}^n \mu_i$ and $\sigma^2 = \sum_{i=1}^n \sigma_i^2$. If $\lim_{n \rightarrow \infty} \frac{\rho}{\sigma} = 0$, then the sum $S_n = \sum_{i=1}^n X_i$ is $AN(\mu, \sigma^2)$.

Example 1: If $\{X_i\}$ are i.i.d.r.vs with p.m.f $(X_i = \pm 1) = \frac{1}{2}$, find the asymptotic distribution of $S_n = \sum_{i=1}^n X_i$.

Solution: Here $E(X_i) = 1 \cdot \frac{1}{2} - 1 \cdot \frac{1}{2} = 0$ and

$$V(X_i) = E(X_i^2) = 1^2 \cdot \frac{1}{2} + (-1)^2 \cdot \frac{1}{2} = 1$$

Let $S_n = X_1 + \dots + X_n$. Then $E(S_n) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) = 0$ and

$$V(S_n) = \sum_{i=1}^n V(X_i) = \sum_{i=1}^n 1 = n.$$

Since mean and variance exist for $\{X_i\}$, by Lindeberg-Levy CLT, $S_n \sim AN(0, n)$ or $\frac{S_n}{\sqrt{n}} \sim AN(0, 1)$.

Example 2: If $\{X_i\}$ are i.i.d. with $E(X_i) = 0$, $V(X_i) = \sigma^2$, $0 < \sigma^2 < \infty$ and $\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$, then show that for any $\epsilon > 0$

$$P(\overline{X}_n \geq \epsilon) = \frac{\sigma}{\epsilon\sqrt{n}} \frac{1}{\sqrt{2\pi}} e^{-\frac{n\epsilon^2}{2\sigma^2}} \text{ as } n \rightarrow \infty.$$

Solution: Let $S_n = X_1 + \dots + X_n$. Then $E(S_n) = \sum_{i=1}^n E(X_i) = 0$ and

$$V(S_n) = \sum_{i=1}^n V(X_i) = n\sigma^2. \text{ Since } \{X_i\} \text{ are i.i.d with finite mean and variance, then}$$

we have $S_n \sim AN[0, n\sigma^2]$ (by Lindeberg-Levy CLT).

Let $\overline{X}_n = \frac{S_n}{n}$. Then $E(\overline{X}_n) = \frac{1}{n} E(S_n) = \frac{0}{n} = 0$ and

$$V(\overline{X}_n) = \frac{1}{n^2} V(S_n) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

Thus, $\overline{X}_n \sim AN\left[0, \frac{\sigma^2}{n}\right]$.

We have $P(\overline{X}_n \geq \epsilon) = P\left(\frac{\overline{X}_n - 0}{\frac{\sigma}{\sqrt{n}}} \geq \frac{\epsilon - 0}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(Z \geq \frac{\sqrt{n}\epsilon}{\sigma}\right)$ where $Z \sim N(0, 1)$

$$= 1 - P\left(Z \leq \frac{\sqrt{n}\epsilon}{\sigma}\right)$$

$$= 1 - \Phi\left(\frac{\sqrt{n}\epsilon}{\sigma}\right), \text{ where } \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{1}{2}t^2} dt$$

$$\Rightarrow P(\bar{X}_n \geq \epsilon) = 1 - \Phi\left(\frac{\sqrt{n}\epsilon}{\sigma}\right) \text{ for } \epsilon > 0 \quad \dots(1)$$

$$\text{But } 1 - \Phi(z) = \frac{1}{z\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad \dots (2)$$

(A result in normal distribution)

From (1) and (2), we have

$$P(\bar{X}_n \geq \epsilon) = 1 - \Phi\left(\frac{\sqrt{n}\epsilon}{\sigma}\right) \Rightarrow P(\bar{X}_n \geq \epsilon) = \frac{\sigma}{\epsilon\sqrt{n}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{n\epsilon^2}{2\sigma^2}}$$

Example 3: Examine if CLT holds for the sequence $\{X_k\}$ with p.m.f

$$P(X_k = \pm 2^k) = 2^{-(2k+1)}, (X_k = 0) = 1 - 2^{-2k}.$$

Solution: Since it is a non identically distributed sequence of r.vs, for CLT to hold, we have to verify the Liapounoff's condition.

$$\text{We have } \mu_k = E(X_k) = 2^k \cdot 2^{-(2k+1)} - 2^k \cdot 2^{-(2k+1)} = 0,$$

$$\sigma_k^2 = V(X_k) = E(X_k^2) = 2^{2k} \cdot 2^{-(2k+1)} + 2^{2k} \cdot 2^{-(2k+1)} = 1 \text{ and}$$

$$\begin{aligned} \rho_k^3 &= E\{|X_k - 0|^3\} = E(|X_k|^3) = 2^{3k} \cdot 2^{-(2k+1)} + 2^{3k} \cdot 2^{-(2k+1)} \\ &= 2 \cdot 2^{3k} \cdot 2^{-(2k+1)} = 2^k \end{aligned}$$

Further, we have

$$\mu = \sum_{k=1}^n \mu_k = 0,$$

$$\sigma^2 = \sum_{k=1}^n \sigma_k^2 = \sum_{i=1}^n 1 = n,$$

$$\rho^3 = \sum_{k=1}^n \rho_k^3 = \sum_{k=1}^n 2^k = 2 + 2^2 + \dots + 2^n = 2(2^n - 1) \text{ and}$$

$$\frac{\rho^3}{(\sigma^2)^{\frac{3}{2}}} = \frac{2(2^n - 1)}{n^{\frac{3}{2}}}$$

Thus, $\lim_{n \rightarrow \infty} \frac{\rho^3}{(\sigma^2)^{\frac{3}{2}}} = \lim_{n \rightarrow \infty} \frac{2(2^n - 1)}{n^{\frac{3}{2}}} = \infty$. Thus, the Liapounoff's condition is not satisfied and hence we cannot say that CLT holds for $\{X_k\}$.

Example 4: Examine if CLT holds for the sequence $\{X_k\}$ with p.m.f

$$P(X_k = \pm k^\alpha) = \frac{1}{2} \cdot k^{-2\alpha}, P(X_k = 0) = 1 - k^{1-2\alpha}, \alpha < \frac{1}{2}.$$

Solution: Since it is a non identically distributed of r.vs, for CLT to hold, we have to verify the Liapounov's condition.

$$\text{We have } \mu_K = E(X_k) = k^\alpha \cdot \frac{1}{2} \cdot k^{-2\alpha} - k^\alpha \cdot \frac{1}{2} \cdot k^{-2\alpha} = 0,$$

$$\sigma_k^2 = V(X_k) = E(X_k^2) = k^{2\alpha} \cdot \frac{1}{2} \cdot k^{-2\alpha} + k^{2\alpha} \cdot \frac{1}{2} \cdot k^{-2\alpha} = \frac{1}{2} + \frac{1}{2} = 1 \text{ and}$$

$$\begin{aligned} \rho_k^3 &= E\{|X_k - 0|^3\} = E\{|X_k|^3\} = k^{3\alpha} \cdot \frac{1}{2} \cdot k^{-2\alpha} + k^{3\alpha} \cdot \frac{1}{2} \cdot k^{-2\alpha} \\ &= \frac{1}{2} \cdot k^\alpha + \frac{1}{2} k^\alpha = k^\alpha \end{aligned}$$

Further, we have

$$\mu = \sum_{k=1}^n \mu_k = \sum_{k=1}^n 0 = 0,$$

$$\sigma^2 = \sum_{k=1}^n \sigma_k^2 = \sum_{k=1}^n 1 = n \text{ and } \rho^3 = \sum_{k=1}^n \rho_k^3 = \sum_{k=1}^n k^\alpha = 1^\alpha + 2^\alpha + \dots + n^\alpha$$

$$\text{Note that } \rho^3 \leq n \cdot n^\alpha = n^{\alpha+1}$$

$$\text{Now } \lim_{n \rightarrow \infty} \frac{\rho^3}{(\sigma^2)^{3/2}} \leq \lim_{n \rightarrow \infty} \frac{n^{\alpha+1}}{n^{3/2}} = \lim_{n \rightarrow \infty} n^{\alpha - \frac{1}{2}} = 0, \text{ if } \alpha < \frac{1}{2}$$

$$\text{Thus } \lim_{n \rightarrow \infty} \frac{\rho^3}{(\sigma^2)^{3/2}} = 0 \text{ if } \alpha < \frac{1}{2}$$

Therefore, CLT holds for the sequence $\{X_k\}$.

Applications of central Limit Theorem:

In case of Bernoulli, Binomial and Poisson distributions, evaluation of probabilities using p.m.f. are tedious. Using normal approximation for large samples to these distributions, the probabilities can be easily evaluated.

(a) Let $\{X_n\}$ be a sequence of i.i.d Bernoulli variate *i. e.*, $B(1, p)$.

$$\text{Let } S_n = X_1 + \cdots + X_n$$

Then $S_n \sim B(n, p)$, where $E(S_n) = np$ and $V(S_n) = np(1 - p) = npq$

By Lindeberg Levy CLT for large n , $S_n \sim AN(E(S_n), V(S_n))$

$$\Rightarrow S_n \sim AN(np, np(1 - p)) \quad \dots (1)$$

$$\text{Let } Z_n = \frac{S_n - np}{\sqrt{np(1-p)}}$$

Then from (1), $Z_n \xrightarrow{d} Z$ where Z is $N(0, 1)$

$$\text{Thus, } \lim_{n \rightarrow \infty} P\left(a \leq \frac{S_n - np}{\sqrt{npq}} \leq b\right) = P(a \leq Z \leq b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}z^2} dz$$

and the *RHS* can be evaluated using standard normal tables for given real numbers a and b .

(b) Let $\{X_n\}$ be a sequence of i.i.d Binomial variates *. e.*, $B(r, p)$.

$$\text{Let } S_n = X_1 + \cdots + X_n$$

$$\text{Then } E(S_n) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) = \sum_{i=1}^n rp = nrp$$

$$(\because E(X_i) = rp \forall i)$$

$$\text{and } V(S_n) = V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) \quad (\because X_i \text{ s are independent})$$

$$= \sum_{i=1}^n rp(1-p) = nrp(1-p)$$

$$(\because V(X_i) = rp(1-p))$$

Thus $E(S_n) = nrp$ and $V(S_n) = nrp(1-p)$

By Lindberg – Levy CLT, for large n , we have

$$S_n \sim AN(nrp, nrp(1-p))$$

$$\text{Let } Z_n = \frac{S_n - nrp}{\sqrt{nrp(1-p)}}$$

Then $Z_n \xrightarrow{d} Z$ where $Z \sim N(0, 1)$

$$\text{Thus, } \lim_{n \rightarrow \infty} P\left(a \leq \frac{S_n - nrp}{\sqrt{nrp(1-p)}} \leq b\right) = P(a \leq Z \leq b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}z^2} dz$$

and the RHS can be evaluated using standard normal tables for given real numbers a and b .

(c) Let $\{X_n\}$ be a sequence of i.i.d Poisson variates . e., $P(\lambda)$. Let $S_n = \sum_{i=1}^n X_i$

Here $E(X_i) = V(X_i) = \lambda \forall i$. Then $E(S_n) = \sum_{i=1}^n E(X_i) = n\lambda$

$$\text{and } V(S_n) = V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) = n\lambda \quad (\because X_i \text{ s are independent})$$

Thus, by Lindeberg – Levy CLT, for large n , $S_n \sim AN(n\lambda, n\lambda)$

Let $Z_n = \frac{S_n - n\lambda}{\sqrt{n\lambda}}$, Then $Z_n \xrightarrow{d} Z$, where $Z \sim N(0, 1)$

$$\text{The probabilities } \lim_{n \rightarrow \infty} P\left(a \leq \frac{S_n - n\lambda}{\sqrt{n\lambda}} \leq b\right) = P(a \leq Z \leq b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}z^2} dz$$

and the RHS can be evaluated using standard normal for given real numbers a and b .

Example 5: A sample of 100 items is taken at random from a batch known to contain 40% defectives. What is the probability that the sample contains

(i) at least 44 defectives,

(ii) exactly 44 defectives?

Solution:

$$\text{Let } X_i = \begin{cases} 1 & , \text{ if the } i^{\text{th}} \text{ item is defective} \\ 0 & , \text{ if the } i^{\text{th}} \text{ item is nondefective} \end{cases}, \text{ for } i = 1, 2, \dots$$

It is given that $P(\text{defective}) = P(X_i = 1) = 40\% = 0.4$

Then X_i follows Bernoulli distribution *i. e.*, $B(1, 0.4)$

Let $S_n = X_1 + \dots + X_n$. Then $S_n \sim B(n, p)$

Since $n = 100$ and $p = 0.4$, $S_n \sim B(100, 0.4)$

Since n is large, computation of probabilities using binomial formula is difficult. Hence, by CLT, we use normal approximation to compute the probabilities of S_n instead of binomial distribution.

Here $E(S_n) = np = 100 \cdot (0.4) = 40$ and

$$V(S_n) = np(1 - p) = 100 \times 0.4 \times 0.6 = 24$$

$$\text{Let } Z_n = \frac{S_n - E(S_n)}{\sqrt{V(S_n)}} = \frac{S_n - 40}{\sqrt{24}} = \frac{S_n - 40}{4.9}$$

Then by Lindeberg Levy CLT, we have $Z_n \xrightarrow{d} Z$, where Z is $N(0, 1)$.

- (i) It should be noted that the continuous normal distribution is approximating the discrete binomial distribution so that the continuity correction has to be taken into account in determining the various probabilities. So finding the probability of at least 44 defectives in a sample of 100 items requires finding the area under the normal curve from **43.5 to 100.5**

Therefore, the probability of at least 44 defectives is given by

$$\begin{aligned}
 P(43.5 < S_n < 100.5) &= P\left(\frac{43.5-40}{4.9} < Z < \frac{100.5-40}{4.9}\right) \\
 &= P(0.7143 < Z < 12.347) \\
 &= P(0 < Z < 12.347) - P(0 < Z < 0.7143) \\
 &= 0.5 - 0.2624 \\
 &\quad \text{(See the standard normal distribution table)} \\
 &= 0.2376
 \end{aligned}$$

- (ii) The probability of exactly 44 defectives is

$$\begin{aligned}
 P(S_n = 44) &= P(43.5 < S_n < 44.5) \\
 &= P\left(\frac{43.5-40}{4.9} < Z < \frac{44.5-40}{4.9}\right) \\
 &= P(0.7143 < Z < 0.9184) \\
 &= P(0 < Z < 0.9184) - P(0 < Z < 0.7143) \\
 &= 0.3208 - 0.2624 \quad \text{(See table)} \\
 &= 0.0584
 \end{aligned}$$

Note: Using the binomial distribution, $P(S_n \geq 44) = \sum_{k=44}^{100} \binom{100}{k} (0.4)^k (0.6)^{100-k}$

and $P(S_n = 44) = \binom{100}{44} (0.4)^{44} (0.6)^{56} = 0.0576$ (Using **Binomial tables**).

As can be seen by comparing the answers, both sets of answers are remarkably close.

Example 6: Let $X_1, X_2, \dots$ be i.i.d. Poisson variables with parameter λ . Use CLT to estimate $P(120 \leq S_n \leq 160)$, where $S_n = X_1 + \dots + X_n$, $\lambda = 2$ and $n = 75$

Solution: Since X_i s are i.i.d $P(\lambda)$, $E(X_i) = \lambda = V(X_i)$ for $i = 1, 2, \dots, n$

$$\therefore E(S_n) = \sum_{i=1}^n E(X_i) = \sum_{i=1}^n \lambda = n\lambda \quad \text{and}$$

$$V(S_n) = V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) = n\lambda$$

Hence by Lindberg – Levy CLT, for large n , we have

$S_n \sim AN(n\lambda, n\lambda) = AN(150, 150)$. After applying the continuity correction, the required probability is

$$p = P(119.5 \leq S_n \leq 160.5) = P\left(\frac{119.5-150}{\sqrt{150}} \leq Z \leq \frac{160.5-150}{\sqrt{150}}\right),$$

where $Z \sim N(0, 1)$

$$= P(-2.45 \leq Z \leq 0.82)$$

$$= P(-2.45 \leq Z \leq 0) + P(0 \leq Z \leq 0.82)$$

$$= P(0 \leq Z \leq 2.45) + P(0 \leq Z \leq 0.82)$$

$$= 0.4929 + 0.2938 \quad (\text{From standard normal table})$$

$$= 0.7868$$

4.5. Central Limit Theorem

Exercise:

1. The lifetime of a certain brand of an electric bulb may be considered as a r.v. with mean 1200 *hrs* and standard deviation 250 *hrs*. Find the probability, using CLT that the average lifetime of 60 bulbs exceeds 1250 *hrs*.
2. A distribution has unknown mean μ and variance σ^2 equal to 1.5. Use CLT to find how large a sample should be taken from the distribution in order that the probability will be at least 0.95 that the sample mean will be within 0.5 of the population mean.
3. A random sample of size 100 is taken from a population whose mean is 60 and variance is 400. Use CLT, with what probability can we assert that the mean of the sample will not differ from $\mu = 60$ by more than 47.
4. The guaranteed average life of a certain type of electric light bulb is 1000 *hrs* with a standard deviation of 125*hrs*. It is decided to sample the output so as to ensure that 90% of the bulbs do not fall short of the guaranteed average by more than 2.5%. Use CLT to find the minimum sample size.
5. If X_i , $i = 1, 2, \dots, 50$ are independent r.vs, each having a Poisson distribution with parameter $\lambda = 0.03$ and $S_n = X_1 + \dots + X_n$, evaluate $P(S_n \geq 3)$, using CLT.

Answers:

1. 0.0606
2. at least 24
3. 0.9544
4. 41
5. 0.1112

P1:

Show that WLLN follows from CLT for the sequence of i.i.d.r.vs .

P1:

Show that WLLN follows from CLT for the sequence of i.i.d.r.vs .

Solution:

Let $\{X_n\}$ be a sequence of i.i.d.r.vs with $E(X_i) = \mu$ and $V(X_i) = \sigma^2 \forall i$.

Then $\{X_n\}$ obeys WLLN if

$$\lim_{n \rightarrow \infty} P \left[\left| \frac{S_n}{n} - \mu \right| > \epsilon \right] = 0 \text{ (By Khinchine's WLLN)}$$

for $\epsilon > 0$ and $S_n = \sum_{i=1}^n X_i$.

$$\Rightarrow \lim_{n \rightarrow \infty} P \left[\left| \frac{S_n}{n} - \mu \right| < \epsilon \right] = 1 \quad \dots (1)$$

Since $\epsilon > 0$ is fixed, for any positive constant k , we have

$$k\sigma\sqrt{n} < \epsilon n \quad \dots (2)$$

for all sufficiently large values of n (because any large multiple of $\sqrt{n}$ is negligible in comparison with any small multiple of n)

From (2), we have $|S_n - n\mu| < k\sigma\sqrt{n} \Rightarrow \{|S_n - n\mu| < \epsilon n\}$

$$i.e., \frac{|S_n - n\mu|}{\sigma\sqrt{n}} < k \Rightarrow \left| \frac{S_n - n\mu}{n} \right| < \epsilon$$

Therefore, $P \left\{ \frac{|S_n - n\mu|}{\sigma\sqrt{n}} < k \right\} \leq P \left\{ \left| \frac{S_n - n\mu}{n} \right| < \epsilon \right\}$

$$i.e., \lim_{n \rightarrow \infty} P \left\{ \frac{|S_n - n\mu|}{\sigma\sqrt{n}} < k \right\} \leq \lim_{n \rightarrow \infty} P \left\{ \left| \frac{S_n - n\mu}{n} \right| < \epsilon \right\} \quad \dots (3)$$

As $n \rightarrow \infty$, by Lindeberg-Levy CLT, $Z = \frac{S_n - n\mu}{\sigma\sqrt{n}} \sim N(0,1)$ i.e., standard normal variate. Then, we can choose large k such that

$$\lim_{n \rightarrow \infty} P \left\{ \frac{|S_n - n\mu|}{\sigma\sqrt{n}} < k \right\} = P(|Z| < k) = 1 \quad \dots(4)$$

From (3) and (4), we have

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{S_n}{n} - \mu \right| < \epsilon \right\} \geq 1$$

$$\Rightarrow \lim_{n \rightarrow \infty} P \left\{ \left| \frac{S_n}{n} - \mu \right| < \epsilon \right\} = 1 \ (\because \text{Probability can't be more than 1}) \text{ which is 1.}$$

Therefore, WLLN holds.

P2:

Examine if CLT holds for the sequence $\{X_k\}$ with p.m.f. $P(X_k = \pm k^\lambda) = \frac{1}{2}$.

P2:

Examine if CLT holds for the sequence $\{X_k\}$ with p.m.f. $P(X_k = \pm k^\lambda) = \frac{1}{2}$.

Solution:

We have $\mu_k = E(X_k) = k^\lambda \frac{1}{2} - k^\lambda \frac{1}{2} = 0$,

$\sigma_k^2 = V(X_k) = k^{2\lambda} \frac{1}{2} + k^{2\lambda} \frac{1}{2} = k^{2\lambda}$ and

$\rho_k^3 = E\{|X_k - 0|^3\} = E(|X_k|^3) = k^{3\lambda} \cdot \frac{1}{2} + k^{3\lambda} \cdot \frac{1}{2} = k^{3\lambda}$

Let $S_n = \sum_{k=1}^n X_k$. Then we have $\mu = \sum_{k=1}^n \mu_k = 0$, $\sigma^2 = \sum_{k=1}^n \sigma_k^2 = \sum_{k=1}^n k^{2\lambda}$

$\rho^3 = \sum_{k=1}^n \rho_k^3 = \sum_{k=1}^n k^{3\lambda}$ and

$$\begin{aligned} \frac{\rho^3}{(\sigma^2)^{\frac{3}{2}}} &= \frac{\sum_{k=1}^n k^{3\lambda}}{\left(\sum_{k=1}^n k^{2\lambda}\right)^{\frac{3}{2}}} = \frac{n^{3\lambda+1}}{3\lambda+1} \times \left(\frac{2\lambda+1}{n^{2\lambda+1}}\right)^{\frac{3}{2}} \quad \left(\because \sum_{k=1}^n k^\alpha = \int_0^n x^\alpha dx = \frac{n^{\alpha+1}}{\alpha+1} \right. \\ &\quad \left. \text{Euler - maclaurian formula} \right) \\ &= \frac{(2\lambda+1)^{\frac{3}{2}}}{(3\lambda+1)} n^{(3\lambda+1)-(2\lambda+1)\frac{3}{2}} \\ &= \frac{(2\lambda+1)^{\frac{3}{2}}}{(3\lambda+1)} n^{-\frac{1}{2}} \rightarrow 0 \text{ as } n \rightarrow \infty \end{aligned}$$

Since Liapounoff's condition is satisfied, CLT holds.

P3:

Examine if CLT holds for the sequence $\{X_k\}$ with p.m.f. $P(X_k = \pm 2^{-k}) = \frac{1}{2}$.

P3:

Examine if CLT holds for the sequence $\{X_k\}$ with p.m.f. $P(X_k = \pm 2^{-k}) = \frac{1}{2}$.

Solution:

$$\text{We have } \mu_k = E(X_k) = 2^{-k} \cdot \frac{1}{2} - 2^{-k} \cdot \frac{1}{2} = 0 ,$$

$$\sigma_k^2 = V(X_k) = E(X_k^2) = 2^{-2k} \cdot \frac{1}{2} + 2^{-2k} \cdot \frac{1}{2} = 2^{-2k} \text{ and}$$

$$\rho_k^3 = E\{|X_k - 0|^3\} = E(|X_k|^3) = 2^{-3k} \cdot \frac{1}{2} + 2^{-3k} \cdot \frac{1}{2} = 2^{-3k}$$

$$\text{Let } S_n = \sum_{k=1}^n X_k . \text{ Then we have } \mu = \sum_{k=1}^n \mu_k = 0$$

$$\sigma^2 = \sum_{k=1}^n \sigma_k^2 = \sum_{k=1}^n 2^{-2k} = \sum_{k=1}^n \frac{1}{2^{2k}} = \sum_{k=1}^n \frac{1}{4^k} = \frac{1}{4} + \frac{1}{4^2} + \dots + \frac{1}{4^n} = \frac{1}{3}(1 - 4^{-n}) \text{ and}$$

$$\rho^3 = \sum_{k=1}^n \rho_k^3 = \sum_{k=1}^n 2^{-3k} = \sum_{k=1}^n \frac{1}{8^k} = \frac{1}{8} + \frac{1}{8^2} + \dots + \frac{1}{8^n} = \frac{1}{7}(1 - 8^{-n}) \text{ and}$$

$$\frac{\rho^3}{(\sigma^2)^{\frac{3}{2}}} = \frac{\frac{1}{7}(1 - 8^{-n})}{\left[\frac{1}{3}(1 - 4^{-n})\right]^{\frac{3}{2}}} \rightarrow \frac{1}{7} \left(\frac{1}{3}\right)^{-\frac{3}{2}} \text{ as } n \rightarrow \infty$$

Since Liapounoff's condition is not satisfied, we can't say CLT holds or not.

P4:

Let $X_1, \dots, X_n$ be i.i.d. standardized variates with $E(X_i^n) < \infty$. Find the limiting distribution of

$$Z_n = \frac{\sqrt{n}(X_1X_2 + X_3X_4 + \dots + X_{2n-1}X_{2n})}{(X_1^2 + X_2^2 + \dots + X_{2n}^2)}$$

P4:

Let $X_1, \dots, X_n$ be i.i.d. standardized variates with $E(X_i^n) < \infty$. Find the limiting distribution of

$$Z_n = \frac{\sqrt{n}(X_1X_2 + X_3X_4 + \dots + X_{2n-1}X_{2n})}{(X_1^2 + X_2^2 + \dots + X_{2n}^2)}$$

Solution:

Since X_i s are i.i.d standardized variates, we have

$$E(X_i) = 0, V(X_i) = E(X_i^2) = 1, i = 1, 2, \dots, n$$

Let $Y_i = X_{2i-1}X_{2i}, i = 1, 2, \dots, n$

$$\Rightarrow E(Y_i) = E(X_{2i-1})E(X_{2i}) = 0 \quad (\because X_i \text{s are independent})$$

$$\text{and } V(Y_i) = E(Y_i^2) = E[X_{2i-1}^2 X_{2i}^2] = E[X_{2i-1}^2] E[X_{2i}^2] = 1 \cdot 1 = 1$$

Hence $Y_i, i = 1, 2, \dots, n$ are also i.i.d.r.vs. Hence CLT holds for $S_n = \sum_{i=1}^n Y_i$. Further,

$$E(S_n) = \sum_{i=1}^n E(Y_i) = 0 \text{ and}$$

$$V(S_n) = \sum_{i=1}^n V(Y_i) = n$$

Then by CLT

$$U_n = \frac{S_n - E(S_n)}{\sqrt{V(S_n)}} = \frac{X_1X_2 + \dots + X_{2n-1}X_{2n}}{\sqrt{n}} \xrightarrow{L} Z \text{ where } Z \sim N(0,1) \text{ as } n \rightarrow \infty.$$

Also $E(X_i^2) = 1$ (finite), $i = 1, 2, \dots, n$.

Hence, by Khinchine's theorem, WLLN applies to the sequence $\{X_i^2\}$, $i = 1, 2, \dots, 2n$ so that

$$V_n = \frac{X_1^2 + X_2^2 + \dots + X_{2n}^2}{2n} \xrightarrow{P} E(X_i^2) = 1 \text{ as } n \rightarrow \infty.$$

Hence, by Slutsky's theorem, we have

$$\lim_{n \rightarrow \infty} (U_n \cdot V_n) = \frac{2\sqrt{n}(X_1X_2 + \cdots + X_{2n-1}X_{2n})}{X_1^2 + \cdots + X_n^2} \xrightarrow{L} \frac{Z}{1} = Z \sim N(0,1)$$

$$\text{Thus, } \frac{\sqrt{n}(X_1X_2 + \cdots + X_{2n-1}X_{2n})}{X_1^2 + \cdots + X_n^2} \xrightarrow{L} \frac{Z}{2} \sim N\left(0, \frac{1}{4}\right) \text{ (If } Z \sim N(0,1), \text{ then } cZ \sim N(0, c^2)).$$