UNIT-VI

Order Statistics

(a) order statistics

(b) weak law of large numbers

(c) strong law of large numbers

(d) central limit theorem

(a) Order Statistics

Independent and Identically Distributed Random Variables

Let $X_1, X_2, \ldots X_n$ are i.i.d.r.v

if

$$F_{X_1, X_2, \ldots X_n}(x_1, x_2 \ldots x_n) = F_{X_1}(x_1) F_{X_2}(x_2) \ldots F_{X_n}(x_n)$$ (Independen)

and $F_{X_i}(x) = F(x) \quad \forall i = 1, 2, \ldots n$ (Identical)

where $F_{X_i}(x)$ is c.d.f of X_i $\quad i = 1, 2, \ldots n$

$F_{X_1, X_2 \ldots X_n}(x_1 x_2 \ldots x_n)$ is J.c.d.f of $X_1, X_2 \ldots X_n$

Order Statistics

Let $x_1, x_2, \ldots x_n$ be a random sample from a population with c.d.f $F(x)$. Then

$x_{(1)}$ = Smallest of $x_1, x_2 \ldots x_n$

$x_{(2)}$ = Second Smallest of " "

$x_{(3)}$ = Third Smallest of " "

$\vdots$ $\qquad\vdots$ $\qquad\vdots$

$x_{(n)}$ = largest of $x_1, x_2 \ldots x_n$

The order values $x_{(1)}, x_{(2)}, \ldots x_{(n)}$ are known as the order statistics (O.S) of the r.v's $x_1, x_2 \ldots x_n$

Note:

① O.S are r.v's itself

② $x_{(1)} \le x_{(2)} \le x_{(3)} \le \cdots \le x_{(n)}$

③ $x_1, x_2, \dots x_n$ are i.i.d. r.v's

but $x_{(1)}, x_{(2)} \dots x_{(n)}$ not i.i.d. r.v's

Distributions of O.S in continuous

let $x_1, x_2, \dots x_n$ be a random sample from a continuous with c.d.f $F(x)$ and p.d.f $f(x)$

Marginal distributions

For $X_{(n)}$

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

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

For $X_{(1)}$

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

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

For $X_{(j)}$

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

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

Joint Distributions

for $1 \le i < j \le n$

J.p.d.f of $X_{(i)}$ and $X_{(j)}$

$$f_{X_{(i)} X_{(j)}}(u, v)$$

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

J.p.d.f of $X_{(1)}\, X_{(2)} \cdots X_{(n)}$ is

$$f_{X_{(1)} \cdots X_{(n)}}(x_1, x_2 \cdots x_n) = n!\, f(x_1) f(x_2) \cdots f(x_n)$$

Problems on order statistics

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 $X_{(1)}, X_{(3)}, X_{(4)}$

Sol: Given $X \sim U[0, \theta]$

its p.d.f $\boxed{f(x) = \frac{1}{\theta}}$

c.d.f $F(x) = \int_0^x f(x)\,dx$

$$= \int_0^x \frac{1}{\theta}\,dx = \frac{1}{\theta}\int_0^x 1\,dx$$

$$\boxed{F(x) = \frac{x}{\theta}}$$

$\boxed{n=4}$

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

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

Here $n=4$

$$f(x) = \frac{1}{\theta}$$

$$F(x) = \frac{x}{\theta}$$

$$= 4\left[1-\frac{x}{\theta}\right]^{4-1} \frac{1}{\theta}$$

$$\boxed{f_{X_{(1)}}(x) = \frac{4}{\theta}\left(1-\frac{x}{\theta}\right)^{3}}$$

c.d.f of $X_{(1)}$

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

Here $n=4$

$$F(x) = \frac{x}{\theta}$$

$$\therefore \boxed{F_{X_{(1)}}(x) = 1 - \left(1 - \frac{x}{\theta}\right)^4}$$

P.d.f of $X_{(4)}$

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

Here $n = 4$

$$f(x) = \frac{1}{\theta}$$

$$F(x) = \frac{x}{\theta}$$

$$f_{X_{(4)}}(x) = 4\left(\frac{x}{\theta}\right)^{4-1} \frac{1}{\theta}$$

$$\boxed{f_{X_{(4)}}(x) = \frac{4}{\theta}\left(\frac{x}{\theta}\right)^3}$$

c.d.f of $X_{(4)}$

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

Here $n = 4$

$F(x) = \frac{x}{\theta}$

$$\boxed{F_{X_{(4)}}(x) = \left(\frac{x}{\theta}\right)^4}$$

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

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

Here $n = 4$, $j = 3$

$F(x) = \frac{x}{\theta}$, $f(x) = \frac{1}{\theta}$

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

$$\boxed{f_{X_{(3)}}(x) = 12\,\frac{x^2}{\theta^3}\left(1-\frac{x}{\theta}\right)}$$

c.d.f of $X_{(3)}$

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

Here $n=4$, $F(x)=\frac{x}{\theta}$, $j=3$

$$= \sum_{i=3}^{4} {}^{4}c_i \left(\frac{x}{\theta}\right)^{i}\left(1-\frac{x}{\theta}\right)^{4-i}$$

$$= {}^{4}c_3\left(\frac{x}{\theta}\right)^{3}\left(1-\frac{x}{\theta}\right)^{1} +$$

$$ {}^{4}c_4\left(\frac{x}{\theta}\right)^{4}\left(1-\frac{x}{\theta}\right)^{0}$$

$$= 4\left(\frac{x}{\theta}\right)^{3}\left(1-\frac{x}{\theta}\right) + \left(\frac{x}{\theta}\right)^{4}$$

$$\boxed{F_{X_{(3)}}(x) = \left(\frac{x}{\theta}\right)^3 \left[4 - \frac{3x}{\theta}\right]}$$

Note:

Ily we can do c.d.f and p.d.f of $X_{(2)}$

② For exponential Distribution $f(x) = e^{-x}$; $(x \geq 0)$ Then find p.d.f and c.d.f of $X_{(1)}$ and $X_{(n)}$

③ If $f(x) = \lambda e^{-\lambda x}$; $x > 0, \lambda > 0$ Then find c.d.f and p.d.f of $X_{(1)}$ and $X_{(n)}$

Convergence of seq of r.v

$\{X_n\}$ is a seq of r.v and

X is a r.v

1. Almost sure convergence

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

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

In this case we write

$$X_n \xrightarrow{a.s} X$$

(or)

$X_n \longrightarrow X$ with probability 1

2. Convergence in probability

The seq of r.v's $\{X_n\}$ is said to convergence in probability to a r.v X if

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

for every $\epsilon > 0$

In this case we write

$$X_n \xrightarrow{P} X$$

3. Convergence in rth mean

$\{X_n\}$ convergence in rth mean to a r.v X if

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

In this case we write

$$X_n \xrightarrow{r} X$$

If $r=2$ we call it as Convergence in Quadratic Mean it is denoted by

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

4. Convergence in Distribution

let $\{F_n\}$ be a Seq of c.d.f's if there exists a c.d.f F such that as $n \to \infty$

$$F_n(x) \longrightarrow F(x) \quad \forall x$$

Then we say that F_n converges weakly to F and it is denoted by $F_n \xrightarrow{w} F$

Note: If $F_n \xrightarrow{w} F$ Then

$$X_n \xrightarrow{d} X \quad \text{(Distr.)}$$

(or)

$$X_n \xrightarrow{L} X \quad \text{(law)}$$

Problem:

let $X_1, X_2, \dots, X_n$ be i.i.d.r.v's

with common p.d.f

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

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

Then show that $X_{(n)} \xrightarrow{L} X$

solution: given $f(x) = \frac{1}{\theta}$

$$F(x) = \int_0^x f(x)\, dx$$

$$= \int_0^x \frac{1}{\theta}\, dx$$

$$= \frac{x}{\theta}$$

Hence

$$\boxed{F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{\theta} & 0 \le x < \theta \\ 1 & x \ge \theta \end{cases}}$$

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

$$F_{(n)}(x) = [F(x)]^n = \begin{cases} 0 & x < 0 \\ \left(\frac{x}{\theta}\right)^n & 0 \le x < \theta \\ 1 & x \ge \theta \end{cases}$$

as $n \to \infty$

$$F(x) = \begin{matrix} 0 & x < \theta \\ 1 & x \ge \theta \end{matrix}$$

Hence we see that

$$\boxed{\lim_{n \to \infty} F_{(n)}(x) = F(x)}$$

Hence $F_n \xrightarrow{w} F$

$\therefore \; X_n \xrightarrow{L} X$

② let $\{X_n\}$ be a seq of r.v's

defined by

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

$n = 1, 2, 3, \ldots$

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

where $P(X = 0) = 1$

Sol. given $P(X_n = 0) = 1 - \frac{1}{n}$

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

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

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

$$E(X_n) = \frac{1}{n}$$

$$E(x_n)^r = 0^r\left(1-\frac{1}{n}\right) + 1^r\left(\frac{1}{n}\right)$$

$$= \frac{1}{n}$$

Now $E(|x_n - 0|^r) = E(|x_n|^r)$

Here $E(|x_n|^r) = \frac{1}{n}$

$E(|x_n|^r) = \frac{1}{n} \to 0$ as $n \to \infty$

Hence $x_n \xrightarrow{a.m} x$

Weak law of large numbers

let $\{X_n\}$ be the seq of r.v's

let $\bar{X}_n = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$

$$= \frac{1}{n} \sum_{i=1}^{n} X_i$$

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

mean of first n r.v's

<u>Note</u>: Some time

$$S_n = X_1 + X_2 + \cdots + X_n$$

$$= \sum_{i=1}^{n} X_i$$

Hence $\dfrac{S_n}{n} = \bar{X}_n$

* Def of WLLN

A seq $\{x_n\}$ of r.v's is said to satisfy the WLLN if

$$\lim_{n \to \infty} P\left[\left| \frac{S_n}{s} - E\left(\frac{S_n}{n}\right) \right| < \epsilon \right] = 1 \quad \forall \epsilon > 0$$

where $S_n = \sum_{i=1}^{n} x_i$

ie $\frac{S_n}{n} \xrightarrow{P} E\left(\frac{S_n}{n}\right)$

Theorem :

Let $\{X_n\}$ be a seq of r.v's and

let $S_n = X_1 + X_2 + \cdots + X_n$ with

$B_n = V(S_n) < \infty$,

If $\frac{B_n}{n^2} \to 0$ as $n \to \infty$, Then for any $\epsilon > 0$

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

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

Proof: We know that chebychev's inequality

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$

Here $X = \frac{S_n}{n}$

$$\mu = E\left(\frac{S_n}{n}\right)$$

$$\sigma^2 = V(X_n) = V\left(\frac{S_n}{n}\right)$$

$$\therefore P\left(\left|\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right| \geq \epsilon\right) \leq \frac{V\left(\frac{S_n}{n}\right)}{\epsilon^2}$$

$$\boxed{\begin{array}{l} V(ax) = a^2 V(x) \\ V\left(\frac{S_n}{n}\right) = \frac{1}{n^2} V(S_n) \end{array}}$$

$$\therefore P\left(\left|\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right| \geq \epsilon\right) \leq \frac{V(S_n)}{n^2 \epsilon^2}$$

$$= \frac{B_n}{n^2 \epsilon^2}$$

as $n \to \infty$ the above $\longrightarrow 0$

$$\therefore \lim_{n \to \infty} P\left(\left|\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right| \geq \epsilon\right) = 0$$

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

Hence $\{X_n\}$ satisfies WLLN

Problem: let $\{X_n\}$ be i.i.d.r.v's with mean μ and variance σ^2

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

as $n \to \infty$ for some constant c

Then find 'c'.

Sol: Here $E(X_i) = \mu$, $V(X_i) = \sigma^2$ $\forall i$

$$S_n = X_1^2 + X_2^2 + \cdots + X_n^2$$

$$E(S_n) = E(X_1^2) + E(X_2^2) + \cdots + E(X_n^2)$$

$$= n\,E(X_1^2)$$ (Since X_i's are i.i.d.r.v)

$$E(S_n) = n\,E(X_1^2)$$

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

$$= n\left[\sigma^2 + \mu^2\right]$$

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

By Khinchine's WLLN

we have $\frac{S_n}{n} \to E\left(\frac{S_n}{n}\right)$

$$\therefore \quad \frac{S_n}{n} \to \sigma^2 + \mu^2$$

Hence $\boxed{C = \sigma^2 + \mu^2}$

Strong Law of Large Numbers

A seq of r.v's $\{X_n\}$ is said to satisfy the strong law of large numbers (SLLN) if

$$\frac{S_n - E(S_n)}{n} \xrightarrow{a.s} 0 \quad \text{as } n \to \infty$$

Kolmogorov's SLLN

Let $\{X_n\}$ be a seq of i.r.v's with $E(X_i) = \mu$ and $V(X_i) = \sigma_i^2$ for $i = 1, 2, \ldots$

If $\sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} < \infty$, Then SLLN Holds

Problem: let $\{X_n\}$ be a seq of i.r.v 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 SLLN Holds for $\{X_n\}$

Sol: Given

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

Now $E(x_n) = \sum x_n \, P(x_n)$

$$= (-2^n)\left(\frac{1}{2^{2n+1}}\right) + 0\left(1 - \frac{1}{2^{2n}}\right) + 2^n\left(\frac{1}{2^{2n+1}}\right)$$

$$= 0$$

$$\sigma^2 = V(x_n) = E(x_n^2) - (E(x_n))^2$$

$$V(x_n) = E(x_n^2)$$

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

$$= 2^{2n} \frac{1}{2^{2n+1}} + 2^{2n} \frac{1}{2^{2n+1}}$$

$$= 1$$

$$\sigma^2 = V(x_n) = 1$$

$\boxed{\sigma^2 = 1}$

Hence $\sum_{n=1}^{\infty} \frac{\sigma^2}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2}$

$\boxed{\sum \frac{1}{n^p} \text{ converges if } p > 1}$

$\Rightarrow \sum_{n=1}^{\infty} \frac{1}{n^2}$ converges

$\therefore$ SLLN holds for $\{X_n\}$

Example problem

① For what values of α does the SLLN hold for the sequence

$$P(X_K = \pm K^{\alpha}) = \frac{1}{2}$$

② SLLN Hold or not for

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

Central Limit Theorem

let $\{X_n\}$ be a seq of i.r.v's

let $S_n = \sum_{i=1}^{n} x_i$

By WLLN $\frac{S_n}{n} \xrightarrow{P} E\left(\frac{S_n}{n}\right)$

SLLN $\frac{S_n - E(S_n)}{n} \xrightarrow{a.s} 0$

By CLT $\frac{S_n}{n} \longrightarrow Z$

where Z is normal variate

Def :-

A seq of i.r.v's $\{X_i\}$ with mean $E(X_i) = \mu_i$ and $V(X_i) = \sigma_i^2$ $\forall i$ is said to follow C.L.T under certain conditions. if the r.v $S_n = X_1 + X_2 + \cdots + X_n$ is Asymptotically Normal (AN) with mean μ, variance σ^2

where $\mu = \sum_{i=1}^{n} \mu_i$, $\sigma^2 = \sum_{i=1}^{n} \sigma_i^2$

Notation: $S_n \sim AN(\mu, \sigma^2)$

Note: Standard Normal Distri.

$Z_n \sim AN(0, 1)$

where $Z_n = \dfrac{S_n - \mu}{\sigma}$

Problem:- let $x_1\ x_2 \cdots$ be i.i.d Poisson variate with parameter λ; use CLT to estimate $P(120 \le S_n \le 160)$ where

$S_n = x_1 + x_2 + \cdots + x_n$

$\lambda = 2, \quad n = 75$

Sol. Since x_i's are i.i.d.r.v

$E(x_i) = V(x_i) = \lambda \quad \forall i$

$S_n = x_1 + x_2 + \cdots + x_n$

$E(S_n) = E(x_1) + \cdots + E(x_n)$

$= n\,E(x_i)$

$\boxed{E(S_n) = n\lambda} = \mu$

$V(S_n) = V\left[\sum_{i=1}^{n} x_i\right]$

$$= \sum_{i=1}^{n} V(x_i)$$

$$= \sum_{i=1}^{n} \lambda$$

$$\boxed{V(S_n) = n\lambda} = \sigma^2$$

Here $S_n \sim AN(\mu; \sigma^2)$

$$\Rightarrow S_n \sim AN(n\lambda, n\lambda)$$

$n = 2, \lambda = 75$

$$S_n \sim AN(150, 150)$$

Now we Evaluate

$P(120 \leq S_n \leq 160)$:-

$$P(120 \le S_n \le 160) = P(z_1 \le z \le z_2)$$

$$z_1 = \frac{x_1 - \mu}{\sigma} \qquad z_2 = \frac{x_2 - \mu}{\sigma}$$

$$= \frac{120 - 150}{\sqrt{150}} \qquad = \frac{160 - 150}{\sqrt{150}}$$

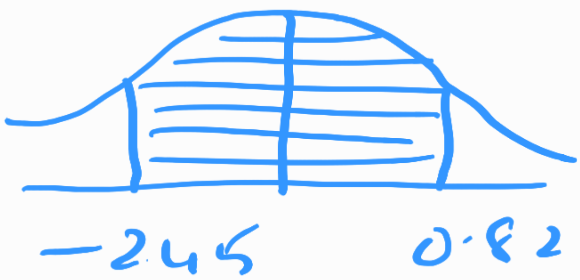
$$= -2.45 \qquad = 0.82$$

$$P(120 \le S_n \le 160) = P(-2.45 \le z \le 0.82)$$

$$= A(2.45) + A(0.82)$$

$$= 0.4929 + 0.2938$$

$$= 0.7868$$



$$\boxed{P(120 \le S_n \le 160) = 0.7868}$$