UNIT-I

Definition: Binomial Distribution:

A r.v X is said to follow B.D if it assumes only non-negative values and its p.m.f is given by:

$$P(X=x)=p(x)=\begin{cases}\binom{n}{x}p^{x}q^{n-x}, & x=0,1,2\dots n;\ q=1-p\\ 0, & \text{otherwise}\end{cases}$$

The two independent constants n and p in the distribution are known as the parameters of the distribution.

Moments of Binomial Distribution:

$$\mu_1' = E(X) = \sum_{x=0}^{n} x\binom{n}{x}p^{x}q^{n-x} = \sum_{x=1}^{n}\binom{n-1}{x-1}p^{x-1}q^{n-x} = np(q+p)^{n-1} = np$$

$$\mu_2' = E(X^2) = \sum_{x=0}^{n} x^2\binom{n}{x}p^{x}q^{n-x} = \sum_{x=0}^{n}\{x(x-1)+x\}\frac{n(n-1)}{x(x-1)}p^{x}q^{n-x}$$

$$= n(n-1)p^2\left\{\sum_{x=2}\binom{n-2}{x-2}p^{x-2}q^{n-x}\right\}+np = n(n-1)p^2(q+p)^{n-2}+np$$

$$\mu_2' = n(n-1)p^2+np$$

$$V(x) = E(X^2)-[E(X)]^2 = n(n-1)p^2+np-(np)^2 = np(1-p) = npq$$

$$\mu_3' = E(X^3) = \sum_{x=0}^{n} x^3 p(x) = \sum_{x=0}^{n}\{x(x-1)(x-2)+3x(x-1)+x\}\binom{n}{x}p^{x}q^{n-x}$$

$$= n(n-1)(n-2)p^3\sum_{x=3}^{n}\binom{n-3}{x-3}p^{x-3}q^{n-x}+3n(n-1)p^2\sum_{x=2}^{n}\binom{n-2}{x-2}p^{x-2}q^{n-x}+np$$

$$= n(n-1)(n-2)p^3+3n(n-1)p^2+np$$

lly

$$\therefore \mu_4' = E(X^4) = \sum_{x=0}^{n}\{x(x-1)(x-2)(x-3)+6x(x-1)(x-2)+7x(x-1)+x\}\binom{n}{x}p^{x}q^{n-x}$$

$$= n(n-1)(n-2)(n-3)p^4+6n(n-1)(n-2)p^3+7n(n-1)p^2+np$$

Central Moments of B.D:

$$\mu_2 = \mu_2'-\mu_1'^2 = npq\ ;\ \mu_3 = \mu_3'-3\mu_2'\mu_1'+2\mu_1'^3 = npq(q-p)$$

$$\mu_4 = \mu_4'-4\mu_3'\mu_1'+6\mu_2'\mu_1'^2-3\mu_1'^4 = npq\{1+3(n-2)pq\}$$

Hence $\beta_1 = \dfrac{\mu_3^2}{\mu_2^3} = \dfrac{(1-2p)^2}{npq}$, $\beta_2 = \dfrac{\mu_4}{\mu_2^2} = 3+\dfrac{1-6pq}{npq}$

$$\gamma_1 = \sqrt{\beta_1} = \frac{1-2p}{\sqrt{npq}},\quad \gamma_2 = \beta_2-3 = \frac{1-6pq}{npq}$$

Moment Generating Function:

$$M_x(t) = \sum_{x=0}^{n} E(e^{tx}) = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x q^{n-x} = (q+pe^t)^n$$

Characteristic Function of B.D:

$$\phi_x(t) = E[e^{itx}] = \sum_{x=0}^{n} \binom{n}{x} (pe^{it})^x q^{n-x} = (q+pe^{it})^n$$

Cumulants of the Binomial Distribution:

$$K_x(t) = \log M_x(t) = \log(q+pe^t)^n = n\log(q+pe^t) = n\log[q+p(1+t+t^2/2!+t^3/3!+\cdots)]$$

$$= n\log[1+p(t+t^2/2!+t^3/3!+\cdots)]$$

$$= n[p(t+t^2/2!+t^3/3!+\cdots) - p^2/2(t+t^2/2!+\cdots)^2 + \frac{p^3}{3}(t+t^2/2!+\cdots)^3 - \frac{p^4}{4}(t+t^2/2!+t^3/3!+\cdots)^4 + \cdots]$$

Mean k_1 = coefficient of t in $k_x(t) = np$,

$\mu_2 = k_2 = n(p-p^2) = npq$, $\mu_3 = npq(q-p) = k_3$

$k_4 = npq(1-6pq)$, $\mu_4 = k_4 + 3k_2^2 = npq[1+3pq(n-2)]$

Mode of B.D:

We have,

$$\frac{p(x)}{p(x-1)} = \frac{\binom{n}{x} p^x q^{n-x}}{\binom{n}{x-1} p^{x-1} q^{n-x+1}} = 1 + \frac{(n+1)p-x}{xq} \quad \text{—(1)}$$

The Recurrence Relation for between Probability of B.D

If $X \sim B(n,p)$

Then the p.m.f of the r.v X is

$p(x) = \binom{n}{x} p^x q^{n-x}$ —(1), $x = x+1$ in (1) we get

$$p(x+1) = \binom{n}{x+1} p^{x+1} q^{n-(x+1)} \quad \text{—(2)}$$

$$\frac{p(x+1)}{p(x)} = \frac{\binom{n}{x+1} p^{x+1} q^{(n-x-1)}}{\binom{n}{x} p^x q^{n-x}} = \frac{p}{q} \frac{x!(n-x)(n-x-1)!}{(x+1)x!(n-x-1)!}$$

$$= p/q \, \frac{n-x}{x+1} \, p(x)$$

Recurrence Relation for the Moments of B.D:

$$\mu_r = E\{x - E(x)\}^r = E(x-np)^r = \sum_{x=0}^{n} (x-np)^r \binom{n}{x} p^x (1-p)^{n-x} \quad \text{— (1)}$$

Differentiating with r. to 'p' we get

$$\frac{d\mu_r}{dp} = \sum_{x=0}^{n} \binom{n}{x} \left[-nr(x-np)^{r-1} p^x q^{n-x} + (x-np)^r \{x p^{x-1} q^{n-x} - (n-x) p^x q^{n-x-1}\} \right]$$

$$= -nr \sum_{x=0}^{n} \binom{n}{x} (x-np)^{r-1} p^x q^{n-x} + \sum_{x=0}^{n} \binom{n}{x} (x-np)^r p^x q^{n-x} \left(\frac{x}{p} - \frac{n-x}{q} \right)$$

$$= -nr \sum_{x=0}^{n} (x-np)^{r-1} p(x) + \frac{1}{pq} \sum_{x=0}^{n} (x-np)^{r+1} p(x)$$

$$= -nr\mu_{r-1} + \frac{1}{pq}\mu_{r+1}$$

$$\mu_{r+1} = pq\left[nr\mu_{r-1} + \frac{d\mu_r}{dp}\right] \quad \text{— (1)} \qquad \because \mu_0 = 1 \ \& \ \mu_1 = 0$$

Putting $r = 1, 2$ and 3 we get.

$r=1$

$$\mu_2 = pq\left(n(1)\mu_0 + \frac{d\mu_1}{dp}\right) = npq$$

$$\mu_3 = pq\left[2n\mu_1 + \frac{d\mu_2}{dp}\right] = pq\left[0 + \frac{d(npq)}{dp}\right] = npq(q-p)$$

$$\mu_4 = pq\left[3n\mu_2 + \frac{d\mu_3}{dp}\right] = pq\left[3n(npq) + \frac{d}{dp}\{npq(q-p)\}\right]$$

$$= npq\left[1 + 3pq(n-2)\right]$$

Additive property of Binomial Distribution:

Let $X \sim B(n_1, p_1)$ and $y \sim B(n_2, p_2)$ be independent r.v's then.

$$M_x(t) = (q_1 + p_1 e^t)^{n_1}, \quad M_y(t) = (q_2 + p_2 e^t)^{n_2} \quad \text{— (1)}$$

we have.

$$M_{x+y}(t) = M_x(t) \cdot M_y(t) \Rightarrow (q_1 + p_1 e^t)^{n_1} \cdot (q_2 + p_2 e^t)^{n_2} \quad \text{— (2)}$$

Since (2) can't be expressed in the form $(q+pe^t)^n$ from Uniqueness theorem of MGF's it follows that $x+y$ is not B. Variable. Hence in general the sum of two independent binomial variate.

However, if we take $p_1 = p_2 = p$ (say) then from (2)

$$M_{x+y}(t) = (q+pe^t)^{n_1+n_2}$$

which is the MGF of a binomial variate with parameters $(n_1 + n_2, p)$.

Poisson Distribution: (Definition)

A r.v X is said to follow a poisson distribution if it assumes only non-negative values and its p.m.f is given by,

$$P(x,\lambda) = P(X=x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!} & , x=0,1,2\ldots \lambda>0 \\ 0 & , \text{otherwise} \end{cases}$$

Here λ is known as the parameter of the distribution.

Moments of the Poisson Distribution:

$$\mu_1' = E(x) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} = \lambda e^{-\lambda}\left\{\sum_{x=0}^{\infty}\frac{\lambda^{(x-1)}}{(x-1)!}\right\} =$$

$$= \lambda e^{-\lambda}(1+\lambda+\lambda^2/2!+\lambda^3/3!+\cdots) = \lambda e^{-\lambda}e^{\lambda} = \lambda$$

$$\mu_2' = E(x^2) = \sum_{x=0}^{\infty}\{x(x-1)+x\}\frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=0}^{\infty}\left[x(x-1)\frac{e^{-\lambda}\lambda^x}{x!} + x\frac{e^{-\lambda}\lambda^x}{x!}\right]$$

$$= \lambda^2 e^{-\lambda}e^{\lambda} + \lambda = \lambda^2+\lambda$$

$$\mu_3' = E(x^3) = \sum_{x=0}^{\infty} x(x-1)(x-2)\frac{e^{-\lambda}\lambda^x}{x!} + \sum_{x=0} 3x(x-1)\frac{e^{-\lambda}\lambda^x}{x!} + \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{x=0}^{\infty} x(x-1)(x-2)\frac{e^{-\lambda}\lambda^x}{x!} + 3\sum_{x=0}^{\infty}x(x-1)\frac{e^{-\lambda}\lambda^x}{x!} + \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda}\lambda^3\left\{\sum_{x=3}^{\infty}\frac{\lambda^{x-3}}{(x-3)!}\right\} + 3e^{-\lambda}\lambda^2\left\{\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{(x-2)!}\right\} + \lambda$$

$$= e^{-\lambda}\lambda^3 e^{\lambda} + 3e^{-\lambda}\lambda^2 e^{\lambda} + \lambda =$$

$$= \lambda^3 + 3\lambda^2 + \lambda$$

$$\underline{V(x)} = E(x) - \{E(x)\}^2 = \lambda^2+\lambda-\lambda^2 = \lambda$$

$$\mu_4' = E(x^4) = \sum_{x=0}^{\infty}\left\{x(x-1)(x-2)(x-3) + 6x(x-1)(x-3) + 7x(x-1) + x\right\}\frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda}\lambda^4\left\{\sum_{x=4}^{\infty}\frac{\lambda^{x-4}}{(x-4)!}\right\} + 6e^{-\lambda}\lambda^3\left\{\sum_{x=3}^{\infty}\frac{\lambda^{x-3}}{(x-3)!}\right\} + 7e^{-\lambda}\lambda^2\left\{\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{(x-2)!}\right\} + \lambda.$$

$$= \lambda^4(e^{-\lambda}e^{\lambda}) + 6\lambda^3(e^{-\lambda}e^{\lambda}) + 7\lambda^2(e^{-\lambda}e^{\lambda}) + \lambda$$

$$= \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda$$

The four central moments are now obtained as follows. ⑤

$\mu_2 = \mu_2' - \mu_1'^2 = (\lambda^2+\lambda) - \lambda^2 = \lambda$

$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3 = (\lambda^3+3\lambda^2+\lambda) - 3\lambda(\lambda^2+\lambda) + 2\lambda^3 = \lambda$

$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4$

$= (\lambda^4+6\lambda^3+7\lambda^2+\lambda) - 4\lambda(\lambda^3+3\lambda^2+\lambda) + 6\lambda^2(\lambda^2+\lambda) - 3\lambda^4 = 3\lambda^2+\lambda$

$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\lambda^2}{\lambda^3} = 1/\lambda$ and $\gamma_1 = \sqrt{\beta_1} = 1/\sqrt{\lambda}$

$\beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + 1/\lambda$ and $\gamma_2 = \beta_2 - 3 = 1/\lambda$

Hence the poisson distribution is always a skewed distribution. Proceeding to the limit as $\lambda \to \infty$ $\beta_1 = 0$ and $\beta_2 = 3$.

Mode of the Poisson Distribution:

$$\frac{P(x)}{P(x-1)} = \frac{\frac{e^{-\lambda}\lambda^x}{x!}}{\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}} = \lambda/x$$

MGF of the P.D:

$$M_X(t) = \sum_{x=0}^{\infty} e^{tx}\frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\cancel{\left[\right.} = \cancel{\sum_{x=0}^{\infty} e^{-\lambda}(e^{t}\lambda}$$

$$= \sum_{x=0}^{\infty} e^{-\lambda}\frac{(\lambda e^t)^x}{x!} = e^{-\lambda}\left\{1+\lambda e^t + \frac{(\lambda e^t)^2}{2!}+\cdots\right\} = e^{-\lambda}e^{\lambda e^t}$$

$$= e^{\lambda(e^t-1)}$$

Characteristic Function of the P.D:

$$\phi_X(t) = \sum_{x=0}^{\infty} e^{itx}\frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{(\lambda e^{it})^x}{x!} = e^{-\lambda}e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}$$

Cumulants of the P.D:

$$K_X(t) = \log\left[e^{\lambda(e^t-1)}\right] = \lambda\left[(e^t-1)\right] = \lambda\left[\left(1+t+t^2/2!+t^3/3!+\cdots\right)-1\right]$$

$$= \lambda\left[(t+t^2/2!+\cdots+t^r/r!+\cdots\right]$$

K_r = r^{th} cumulant = Coefficient of $t^r/r!$ in $K_X(t) = \lambda$

$\Rightarrow \quad K_r = \lambda \;, r = 1, 2, 3 \ldots$

Mean $= k_1 = \mu_1 = \lambda$, $\mu_2 = \lambda$, $\mu_3 = \lambda$ and $\mu_4 = k_4 + 3k_2^2 = \lambda + 3\lambda^2$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \lambda^2/\lambda^3 = 1/\lambda \text{ and } \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\lambda + 3\lambda^2}{\lambda^2} = 1/\lambda + 3$$

Recurrence relation for Moments of the P.D

$$\mu_r = E[X - E(X)]^r = \sum_{x=0}^{\infty} (x-\lambda)^r \frac{e^{-\lambda}\lambda^x}{x!}$$

diff with respect to λ, we get

$$\frac{d\mu_r}{d\lambda} = \sum_{x=0}^{\infty} r(x-\lambda)^{r-1}(-1)\frac{e^{-\lambda}\lambda^x}{x!} + \sum_{x=0}^{\infty} \frac{(x-\lambda)^r}{x!}\left\{x\lambda^{x-1}e^{-\lambda} - \lambda^x e^{-\lambda}\right\}$$

$$= -r\sum_{x=0}^{\infty}(x-\lambda)^{r-1}\frac{e^{-\lambda}\lambda^x}{x!} + 1/\lambda \sum_{x=0}^{\infty}(x-\lambda)^{r+1}\frac{e^{-\lambda}\lambda^x}{x!}$$

$$= -r\mu_{r-1} + 1/\lambda\, \mu_{r+1}$$

$$\mu_{r+1} = r\lambda\mu_{r-1} + \lambda\frac{d\mu_r}{d\lambda}$$

Putting $r = 1, 2$ and 3 successively we get.

$r=1$

$$\mu_2 = r\lambda\mu_0 + \lambda\frac{d\mu_1}{d\lambda} = \lambda$$

$$\mu_3 = 2\lambda\mu_1 + \lambda\frac{d\mu_2}{d\lambda} = \lambda, \quad \mu_4 = 3\lambda\mu_2 + \frac{\lambda\, d\mu_3}{d\lambda} = 3\lambda^2 + \lambda$$

Reccurrence relation ~~of~~ for probability of P.D

$$P(x) = \frac{e^{-\lambda}\lambda^x}{x!} \text{ — (1)}, \quad P(x+1) = \frac{e^{-\lambda}\lambda^{x+1}}{(x+1)} \text{ — (2)}$$

$$\frac{P(x+1)}{P(x)} = \frac{e^{-\lambda}\lambda^{x+1}/(x+1)!}{e^{-\lambda}\lambda^x/x!} = \lambda x!/(x+1)x! \Longrightarrow P(x+1) = \frac{\lambda}{x+1}P(x),$$

Additive property of Poisson Distribution

Sum of independent Poisson variate more elaborately, if $x_i\,(i=1,2,\ldots,n)$ are independent poisson variates with parameters $\lambda_i; i=1,2,\ldots n$ respectively then $\sum_{i=1}^{n} x_i$ is also a poisson variate with parameter $\sum_{i=1}^{n}\lambda_i$

Proof: $M_{x_i}(t) = e^{\lambda_i(e^t-1)}$

$$M_{x_1+x_2+\ldots+x_n}(t) = M_{x_1}(t)\,M_{x_2}(t)\ldots M_{x_n}(t) \Rightarrow e^{\lambda_1(e^t-1)}e^{\lambda_2(e^t-1)}\ldots e^{\lambda_n(e^t-1)}$$

$$= e^{(\lambda_1+\lambda_2+\cdots+\lambda_n)(e^t-1)}$$

which is the m.g.f of a Poisson Variate with parameters $\lambda_1+\lambda_2+\ldots+\lambda_n$. Hence by uniqueness theorem of m.g.f's $\sum_{i=1}^{n} x_i$ is also a poisson Variable with parameter $\sum_{i=1}^{n}\lambda_i$

Geometric Distribution:

A r.v X is said to have a geometric distribution if it assumes only non-negative Value and its p.m.f is given by:

$$P(X=x)=\begin{cases} q^x p & , x=0,1,2\ldots \text{ at } 0<p\le 1; \; q=1-p \\ 0 & , \text{otherwise.} \end{cases}$$

Moments of Geometric function:

$$\text{Mean } \mu_1' = \sum_{x=0}^{\infty} x\, p q^x = pq \sum_{x=1}^{\infty} x\, q^{x-1} = pq(1-q)^{-2} = q/p$$

$$V(x) = E(x^2) - [E(x)]^2 = E[x(x-1)] + E(x) - [E(x)]^2 \quad \text{--- (1)}$$

$$E[X(X-1)] = \sum_{x=2}^{\infty} x(x-1)pq^x = 2pq^2 \sum_{x=2}^{\infty}\left[\frac{x(x-1)}{2\times1} q^{x-2}\right] = 2pq^2(1-q)^{-3} = \frac{2q^2}{p^2}$$

$$V(x) = \mu_2 = \frac{2q^2}{p^2} + q/p - q^2/p^2 = q/p^2$$

MGF:

$$M_x(t) = E(e^{tx}) = \sum_{x=0}^{\infty} e^{tx} q^x p = p(1-qe^t)^{-1} = \frac{p}{(1-qe^t)}$$

$$\mu_1' = \left[\frac{d}{dt} M(t)\right]_{t=0} = \left[\frac{d}{dt}[p(1-qe^t)^{-1}]\right]_{t=0} \Rightarrow pq(1-q)^{-2} = q/p$$

$$\mu_2' = \left[\frac{d^2}{dt^2} M(t)\right]_{t=0} = q/p + 2q^2/p^2$$

$$\mu_2 = \mu_2' - \mu_1'^2 = q/p + 2q^2/p^2 - q^2/p^2 = \frac{q^2+pq}{p^2} = q/p^2$$

Lack of Memory:

The geometric distribution is said to lack memory in a certain sense, suppose an event E can occur at one of the time $t=0,1,2\ldots$ and the occurrence distribution time X has a g.d with parameter p. Thus, $P(X=t) = q^t p;\ t=0,1,\ldots$

Suppose we know that the event E has not occurred before k. (ie) $X \geq k$. Let $Y = X-k$. Thus Y is the amount

$$\frac{1}{f(x)} f'(x) = -\frac{1}{\sigma^2}(x-\mu) \Rightarrow f'(x) = -1/\sigma^2 (x-\mu) f(x)$$

and $$f''(x) = -\frac{1}{\sigma^2}[1 \cdot f(x) + (x-\mu) f'(x)] = -\frac{f(x)}{\sigma^2}\left[1 - \frac{(x-\mu)^2}{\sigma^2}\right] \quad (1)$$

$f'(x) = 0 \Rightarrow x-\mu = 0 \Rightarrow x = \mu$. At the point $x=\mu$, we have from (1)

$$f''(x) = -1/\sigma^2 [f(x)]_{x=\mu} \Rightarrow -1/\sigma^2 \frac{1}{\sigma\sqrt{2\pi}} < 0.$$

Median of N.D:

If M is the median of the N.D, we have.

$$\int_{-\infty}^{M} f(x)\,dx = \frac{1}{2} \Rightarrow \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{M} \exp\{-(x-\mu)^2/2\sigma^2\}\,dx = 1/2$$

$$\Rightarrow \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\mu} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}dx + \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu}^{M} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}dx = \frac{1}{2}$$

But $$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\mu} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} \exp(z^2/2)\,dz = 1/2 \quad (1)$$

$\therefore$ Form (1) we have $\cancel{1/2} + \cancel{\frac{1}{2\sqrt{2\pi}}}\ \frac{1}{2} + \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu}^{M} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}dx = 1/2$

$$\Rightarrow \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu}^{M} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}dx = 0 \quad \text{ie, } \boxed{\mu = M}$$

M.G.F:

$$M_x(t) = \int_{-\infty}^{\infty} e^{tx} \exp\{-(x-\mu)^2/2\sigma^2\}\,dx \qquad \because z = \frac{x-\mu}{\sigma}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\{t(\mu+\sigma z)\} \exp(-z^2/2)\,dz$$

$$= e^{\mu t} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\{-1/2(z^2 - 2t\sigma z)\}\,dz$$

$$= e^{\mu t} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-1/2\{(z-\sigma t)^2 - \sigma^2 t^2\}\right]dz$$

$$= e^{\mu t + t^2\sigma^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\{-1/2(z-\sigma t)^2\}\,dz$$

$$= e^{\mu t + t^2\sigma/2} \times \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-u^2/2)\,du \Rightarrow e^{\mu t + t^2\sigma^2/2}$$

$$= e^{\cancel{\mu t + t^2\sigma/2}}$$

Cumulant Generating function:

$$K_x(t) = \log M_x(t) = \log\left(e^{\mu t + t^2\sigma^2/2}\right) = \mu t + \frac{t^2\sigma^2}{2}$$

$\therefore$ Mean $= k_1 =$ Coefficient of t in $K_x(t) = \mu$

Variance $= k_2 =$ " $t^2/2!$ " $= \sigma^2$

and $k_r =$ Coefficient of $t^r/r!$ in $K_x(t) = 0$, $r = 3, 4 \cdots$

Thus $\mu_3 = k_3 = 0$ and $\mu_4 = k_4 + 3k_2^2 = 3\sigma^4$

Hence $\beta_1 = \mu_3^2/\mu_2^3 = 0$ and $\beta_2 = \frac{\mu_4}{\mu_2^2} = 3$

Moments of N.D: odd order moments about mean are given by:

$$\mu_{2n+1} = \int_{-\infty}^{\infty}(x-\mu)^{2n+1} f(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(x-\mu)^{2n+1}\exp\{-(x-\mu)^2/2\sigma^2\}\,dx \quad \because z = \frac{x-\mu}{\sigma}$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\sigma z)^{2n+1}\exp(-z^2/2)\,dz$$

$$\therefore \mu_{2n+1} = \frac{\sigma^{2n+1}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z^{2n+1}\exp(-z^2/2)\,dz = 0 \qquad \text{—(1)}$$

Since the integrand $z^{2n+1}e^{-z^2/2}$ is an odd function of z.

Even order moments about mean are given by:

$$\mu_{2n} = \int_{-\infty}^{\infty}(x-\mu)^{2n} f(x)\,dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\sigma z)^{2n}\exp(-z^2/2)\,dz$$

$$= \frac{\sigma^{2n}}{\sqrt{2\pi}}\, 2\int_0^{\infty} z^{2n}\exp(-z^2/2)\,dz \qquad [t = z^2/2]$$

$$= \frac{2\sigma^{2n}}{\sqrt{2\pi}}\int_0^{\infty}(2t)^n e^{-t}\,\frac{dt}{\sqrt{2t}}$$

$$\therefore \mu_{2n} = \frac{2^n\sigma^{2n}}{\sqrt{\pi}}\int_0^{\infty} e^{-t}\, t^{(n+1/2)-1}\,dt \Rightarrow = \frac{2^n\sigma^{2n}}{\sqrt{\pi}}\,\Gamma(n+1/2)$$

Changing n to $(n-1)$, we get.

$$\mu_{2n-2} = \frac{2^{n-1}\sigma^{2n-2}}{\sqrt{\pi}}\,\Gamma(n-1/2)$$

$$\therefore \frac{\mu_{2n}}{\mu_{2n-2}} = 2\sigma^2\,\frac{\Gamma(n+1/2)}{\Gamma(n-1/2)} = 2\sigma^2(n-1/2) = \sigma^2(2n-1)\mu_{2n-2} \qquad \text{—(2)}$$

which gives the recurrence relation for the moment of N.D

$$\mu_{2n} = [(2n-1)\sigma^2][(2n-3)\sigma^2]\mu_{2n-4} = 1.3.5\cdots(2n-1)\sigma^{2n} \qquad \text{—(3)}$$

from (1) & (2)

A Linear Combination of independent normal Variates is also a normal Variate:

Let X_i, $(i=1,2,\dots,n)$ be n independent normal Variates with mean μ_i and variance σ_i^2 respectively. Then.

$$M_{X_i}(t) = \exp\{\mu_i t + (t^2\sigma_i^2/2)\} \quad \text{—— (1)}$$

The m.g.f of their linear combination $\sum_{i=1}^{n} a_i X_i$, where $a_1,\dots,a_n$ are constants, is given by.

$$M_{\sum a_i x_i}(t) = \prod_{i=1}^{n} M_{a_i x_i}(t) = M_{X_1}(a_1 t).M_{X_2}(a_2 t)\dots M_{X_n}(a_n t) \quad \text{(2)}$$

From (1), we have $M_{X_1}(a_i t) = e^{\mu_i a_i t + t^2 a_i^2 \sigma_i^2/2}$

$$\therefore M_{\sum_i a_i x_i}(t) = [e^{\mu_1 a_1 t + t^2 a_1^2\sigma_1^2/2} \times \dots \times e^{\mu_n a_n t + t^2 \sigma_n^2 \sigma_n^2/2}]$$

$$= \exp[(\sum_{i=1}^{n} a_i\mu_i)t + t^2(\sum_{i=1}^{n} a_i^2\sigma_i^2)/2]$$

Which is the m.g.f of a normal Variate with mean $\sum_{i=1}^{n} a_i\mu_i$ and variance $\sum_{i=1}^{n} \sigma_i^2\sigma_i^2$.

Hence by uniqueness theorem of m.g.f

$$\sum_{i=1}^{n} a_i x_i \sim N\left[\sum_{i=1}^{n} a_i\mu_i, \sum_{i=1}^{n} a_i\sigma_i^2\right]$$

Gamma Distribution:

A r.V X is said to have a gamma distribution with parameter $\lambda>0$, if its p.d.f is given by:

$$f(x) = \begin{cases} \dfrac{e^{-x} x^{\lambda-1}}{\Gamma(\lambda)}; & \lambda>0,\ 0<x<\infty \\ 0, & \text{otherwise} \end{cases}$$

M.G.F of Gamma Distribution:

$$M_X(t) = E(e^{tx}) = \int_0^\infty e^{tx} e^{-x} x^{\lambda-1} dx$$

$$= \frac{1}{\Gamma(\lambda)} \int_0^\infty e^{-(1-t)x} x^{\lambda-1} dx = \frac{1}{\Gamma(\lambda)} \frac{\Gamma(\lambda)}{(1-t)^\lambda}, \ |t|<1$$

$$M_X(t) = (1-t)^{-\lambda}$$

Cumulant Generating Function:

$$K_x(t) = \log(1-t)^{-\lambda} = -\lambda \log(1-t) \; ; \; |t|<1$$

$$= \lambda\left(t + t^2/2 + t^3/3 + t^4/4 + \cdots\right)$$

Mean $= k_1 =$ Coefficient of t in $K_x(t) = \lambda$

Variance $= \mu_2 = k_2$ " " $t^2/2!$ $K_x(t) = \lambda$

$$\mu_3 = k_3 = 2\lambda, \quad k_4 = 6\lambda,$$

$$\mu_4 = k_4 + 3k_2^2 = 6\lambda + 3\lambda^2$$

Hence $\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{4\lambda^2}{\lambda^3} = 4/\lambda$ and $\beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + 6/\lambda$

Additive property of Gamma Distribution:

The sum of independent Gamma Variates is also a Gamma Variate. More precisely, if $x_1, \cdots, x_k$ are i.G.V with parameters $\lambda_1, \lambda_2, \ldots, \lambda_k$ respectively then $x_1 + x_2 + \cdots + x_k$ is also a G.V with parameter $\lambda_1 + \lambda_2 + \cdots + \lambda_k$

Proof:

Since x_i is a $\gamma(\lambda_i)$ Variate, $M_{x_i}(t) = (1-t)^{-\lambda_i}$

The m.g.f of the sum $x_1 + x_2 + \cdots + x_k$ is given by.

$$M_{x_1+x_2+\cdots+x_k}(t) = M_{x_1}(t)\, M_{x_2}(t) \cdots M_{x_k}(t)$$

$$= (1-t)^{-\lambda_1} \cdots (1-t)^{-\lambda_k}$$

$$= (1-t)^{-(\lambda_1 + \cdots + \lambda_k)}$$

which is m.g.f of a G.V with parameter $\lambda_1 + \lambda_2 + \cdots + \lambda_k$.

Hence the result follows by the uniqueness theorem of m.g.f.'s.

Exponential Distribution:

A r.v. X is said to have an exponential distribution with parameter $\theta > 0$, if its p.d.f is given by:

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

MGF:

$$M_x(t) = E[e^{tx}) = \theta \int_0^{\infty} e^{tx} e^{-\theta x} dx = \theta \int_0^{\infty} \exp\{-(\theta - t)x\} dx$$

$$= \frac{\theta}{(\theta - t)} = (1 - t/\theta)^{-1} = \sum_{r=0}^{\infty} (t/\theta)^r, \theta > t.$$

$\mu_r' = E(X^r) =$ Coefficient of $t^r/r!$ in $M_x(t) = \frac{r!}{\theta^r}$; $r = 1, 2, \ldots$

Mean $= \mu_1' = 1/\theta$ and Variance $= \mu_2 = \mu_2' - \mu_1'^2 = \frac{2}{\theta^2} - 1/\theta^2$
$= 1/\theta^2$

Hence, if $x \sim \exp(\theta)$, then Mean $= 1/\theta$ & Variance $= 1/\theta^2$

STANDARD LAPLACE (double Exponential) Distribution:

A Continuous r.V X is said to have standard Laplace (double exponential) distribution if its p.d.f is given by: $f(x) = \frac{1}{2} \exp.(-|x|)$; $-\infty < x < \infty$

Characteristic Function of standard Laplace Distribution:

$$\cancel{\phi_x(t) = \int_{-\infty}^{\infty} e^{itx} e^{-|x|} dx}$$

$$\phi_x(t) = 1/2 \int_{-\infty}^{\infty} e^{itx} e^{-|x|} dx$$

$$= 1/2 \left[\int_{-\infty}^{\infty} \cos tx . e^{-|x|} dx + i \int_{-\infty}^{\infty} \sin tx \, e^{-|x|} dx \right]$$

$$= \frac{1}{2} \, 2 \int_0^{\infty} \cos tx \, e^{-|x|} dx$$

Since the integrands in the first and second integrals are even and odd function of x respectively.

$$\therefore \phi_x(t) = \int_0^\infty e^{-x}\cos tx \, dx = 1 - t^2 \int_0^\infty e^{-x} \cos tx \, dx = 1 - t^2 \phi_x(t)$$

$$\Rightarrow \phi_x(t) = \frac{1}{1+t^2}$$

$$\therefore \phi_x(t) = 1 - t^2 + t^4 + \cdots = 1 + 2\frac{(it)^2}{2!} + 4! \frac{(it)^4}{4!} + \cdots$$

Hence $k_1 = k_3 = 0$; $k_2 = 2$; $k_4 = 4! = 24$

$\therefore$ Mean $= k_1 = 0$; Variance $= \mu_2 = k_2 = 2$

$\mu_3 = k_3 = 0$; $\mu_4 = k_4 + 3k_2^2 = 24 + 12 = 36$

$$\Rightarrow \beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0 \; ; \; \beta = \frac{\mu_4}{\mu_2^2} = 9$$

Weibul Distribution:

A r.v X ~~have~~ has a Weibul distribution with three parameters $c\,(>0)$ $\alpha\,(>0)$ and μ if the r.v.

$$Y = \left(\frac{X - \mu}{\sigma}\right)^c$$

has the exponential distribution with p.d.f

$$p_Y(y) = e^{-y}, \; y > 0$$

Definition:

A Continuous r.v. X has a Weibul distribution with parameters $c\,(>0)$, $a\,(>0)$, and μ if its p.d.f is:

$$f(x; c, \alpha, \mu) = \frac{c}{\alpha}\left(\frac{x-\mu}{\alpha}\right)^{c-1} \exp\left\{\left(\frac{x-\mu}{\alpha}\right)^c\right\}; \quad x > \mu, \; c > 0$$

The standard weibul distribution is obtained on taking $\alpha = 1$ & $\mu = 0$, so that the p.d.f of S.W.D which depends only on a single parameter C is given by:

$$p_X(x) = C x^{c-1} \cdot \exp(-x^c) \; ; \; x > 0, \; c > 0.$$

Moments of Standard Weibul Distribution:

$$\mu_r' = E(X^r) = E(Y^{1/c})^r = E(Y^{r/c}) = \int_0^\infty e^{-y} y^{r/c}\, dy$$
$$= \Gamma(r/c + 1)$$

$\therefore$ Mean $= E(X) = \Gamma(1/c + 1)$

and $Var(X) = E(X^2) - [E(X)]^2 = \Gamma(2/c + 1) - [\Gamma(1/c + 1)]^2$

Similarly, we can obtain expressions for higher order moments and hence for β_1 and β_2 for Large c, The mean is approximated by:

$$E(X) \simeq 1 - \gamma/c + \frac{1}{2c^2}\left(\frac{\pi^2}{6} + \gamma^2\right) \cancel{+ 0.57722c^{-1} + 0}$$
$$= 1 - 0.57722c^{-1} + 0.989055c^{-2}$$

Where $\gamma = 0.57722$ is Euler's Constant.

Cauchy Distribution:

A r.v X is said to have a Standard Cauchy distribution if its p.d.f is given by:

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

and X is termed as S.C.V

More generally, C.D with parameters λ and μ has the p.d.f

$$g_Y(y) = \frac{\lambda}{\pi[\lambda^2 + (y-\mu)^2]}, \quad -\infty < y < \infty;\ \lambda > 0$$

and we write $X \sim C(\lambda, \mu)$

Characteristic Function of (Standard) C.D:

$$\phi_X(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{e^{itx}}{1+x^2}\, dx \quad \text{—— (1)}$$

To evaluate (1) consider Standard Laplace distribution

$$f_1(z) = \frac{1}{2} e^{-|z|}, \quad -\infty < z < \infty.$$

Then $\phi_1(t) = \phi_Z(t) = E(e^{itZ}) = \frac{1}{1+t^2}$

Since $\phi_1(t)$ is absolutely integrable in $(-\infty, \infty)$, we have by Inversion theorem.

$$\frac{1}{2}e^{-|z|} = f_1(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{itz}\phi_1(t)\,dt = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-itz}}{1+t^2}\,dt$$

$$\Rightarrow e^{-|z|} = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{e^{itz}}{1+t^2}\,dt = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{e^{itz}}{1+t^2}\,dt$$

[Changing t to $-t$]

On interchanging t and z, we have.

$$e^{-|t|} = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{e^{itz}}{1+z^2}\,dz \quad \text{--- (2)}$$

from (1) and (2) we get.

$$\phi_X(t) = e^{-|t|}$$

Moments of Cauchy Distribution:

$$E(y) = \int_{-\infty}^{\infty} y f(y)\,dy = \frac{\lambda}{\pi}\int_{-\infty}^{\infty}\frac{y}{\lambda^2+(y-\mu)^2}\,dy$$

$$= \frac{\lambda}{\pi}\int_{-\infty}^{\infty}\frac{(y-\mu)+\mu}{\lambda^2+(y-\mu)^2}\,dy = \mu\frac{\lambda}{\pi}\int_{-\infty}^{\infty}\frac{dy}{\lambda^2+(y-\mu)^2}+\frac{\lambda}{\pi}\int_{-\infty}^{\infty}\frac{(y-\mu)}{\lambda^2+(y-\mu)^2}\,dy$$

$$= \mu\cdot 1 + \frac{1}{\pi}\int_{-\infty}^{\infty} z\,\frac{z}{\lambda^2+z^2}\,dz$$

Although the integral $\int_{-\infty}^{\infty}\frac{z}{\lambda^2+z^2}\,dz$, is not completely convergent, i.e., $\lim_{\substack{n\to\infty \\ n'\to\infty}}\int_{-n}^{n'}\frac{z}{\lambda^2+z^2}\,dz$, does not exist, its principal value, $\lim_{n\to\infty}\int_{-n}^{n}\frac{z}{\lambda^2+z^2}\,dz$, exists and is equal to zero.

Thus in the general sense the mean of Cauchy dist. does not exist. But, if we conventionally agree to assume that the mean of C.D exists, then it is located at $x=\mu$. Also obviously, the p.d.f. curve is symmetrical about the point $x=\mu$. Hence for this dist the mean, median, mode, coincide at the point $x=\mu$

$$\mu_2 = E(y-\mu)^2 = \int_{-\infty}^{\infty}(y-\mu)^2 f(y)\,dy = \frac{\lambda}{\pi}\int_{-\infty}^{\infty}\frac{(y-\mu)^2}{\lambda^2+(y-\mu)^2}\,dy$$

which does not exist since the integral is not convergent. Thus, in general, for the Cauchy's distribution the moments μ_r $(r \geq 2)$ do not exist.

Additive Property of Cauchy Distribution

If X_1 and X_2 are independent Cauchy variates with parameters (λ_1, μ_1) and (λ_2, μ_2) respectively, then X_1+X_2 is a Cauchy with parameters $(\lambda_1+\lambda_2, \mu_1+\mu_2)$.

Proof:

$$\phi_{X_j}(t) = \exp\{i\mu_j t - \lambda_j |t|\}, \quad (j=1,2)$$

$$\therefore \phi_{X_1+X_2}(t) = \phi_{X_1}(t)\cdot\phi_{X_2}(t) = \exp[it(\mu_1+\mu_2)-(\lambda_1+\lambda_2)|t|]$$

and the result follows by uniqueness theorem of characteristic functions.

Erlang Distribution:

A continuous r.v x is said to follow an erlang dist. with parameters $\lambda > 0$ and $k > 0$, if p.d.f is given by.

$$f(x) = \begin{cases} \dfrac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma k} & \text{for } x \geq 0,\ k \geq 0 \\ 0 & \text{, otherwise} \end{cases}$$

Mean and Variance of erlang distribution:

The raw moment μ_r' about origin of erlang distribution are given by

$$\mu_r' = E(x^r) = \int_0^\infty \frac{\lambda^k}{\Gamma k} x^r x^{k-1} e^{-\lambda x}\,dx = \frac{\lambda^k}{\Gamma k}\int_0^\infty x^{k+r-1} e^{-\lambda x}\,dx$$

put, $\lambda x = t$, $\lambda dx = dt$, $dx = dt/\lambda$

when $x=0 \to t=0$, $x=\infty \to t=\infty$

$$\mu_r' = \frac{\lambda^k}{\Gamma k}\int_0^\infty \frac{t^{k+r-1}}{\lambda^{k+r-1}} e^{-t}\,dt/\lambda \Rightarrow \frac{1}{\lambda^r}\frac{\Gamma k+r}{\Gamma k} \text{ //}$$

$$\text{mean} = E(x) = \mu_1' = 1/\lambda^1 \frac{\Gamma(k+1)}{\Gamma k} = \frac{1}{\lambda}\frac{k\Gamma k}{\Gamma k} = k/\lambda \text{ //}$$

$$E(x^2) = \mu_2' = 1/\lambda^2 \frac{(k+1)\Gamma k+1}{\Gamma k} \text{ //}$$

$$Var(x) = \frac{k(k+1)}{\lambda^2} - \frac{k^2}{\lambda^2} = k/\lambda^2 \text{ //}$$

Additive property of Erlang distribution:

The sum of a finite number of independent erlang variables is also an erlang variable. i.e, if $x_1, x_2, \ldots x_n$ are independent erlang variables with parameters $(\lambda_1, k_1) \ldots (\lambda_n k_n)$ then $x_1+x_2+\cdots+x_n$ is also an erlang variable with parameter $(\lambda_1, k_1+k_2+\cdots+k_n)$

Proof:-

Let us first find with parameter λ and k and use it.

$$M_x(t) = E(e^{tx}) = \int_0^\infty \frac{\lambda^k}{\Gamma k} x^{k-1} e^{-\lambda x} e^{tx} dx$$

$$= \frac{\lambda^k}{\Gamma k} \int_0^\infty x^{k-1} e^{-(\lambda-t)x} dx$$

Put $(\lambda-t)x = y \Rightarrow x = y/_{\lambda-t} \Rightarrow dx = dy/_{\lambda-t}$

when $x=0, y=0$ / $x=\infty, y=\infty$

$$M_x(t) = \frac{\lambda^k}{\Gamma k (\lambda-t)^k} \int_0^\infty y^{k-1} e^{-y} dy = \frac{1}{\Gamma k} \Gamma k \frac{\lambda^k}{(\lambda-t)^k} = \left(\frac{\lambda}{\lambda-t}\right)^k$$

$$= \left(\frac{\lambda-t}{\lambda}\right)^{-k} = (1-t/\lambda)^{-k}$$

Now, $M_{x_1+\cdots+x_n}(t) = M_{x_1}(t) \ldots M_{x_n}(t)$

$$= (1-t/\lambda)^{-k_1} \ldots (1-t/\lambda)^{-k_n}$$

$$= (1-t/\lambda)^{-(k_1+k_2+\cdots+k_n)}$$

which is the m.g.f of an erlang variable with parameter $(\lambda_1, k_1+k_2+\cdots+k_n)$

UNIT-II

Bivariate Normal Distribution: (BND)

The B.N.D is a generalization of a N.D for a single variate, let x and y be two normally correlated variables with correlation coefficient ρ and $E(x) = \mu_1$, $V(x) = \sigma_1^2$; $E(y) = \mu_2$, $V(y) = \sigma_2^2$. The following three assumptions.

(i) The regression of y and x is linear, since of each arrays is on the line of regression y

The mean value of y is $\rho(\sigma_2/\sigma_1)x$, for different value of x.

ii) The arrays are homoscedastic, i.e variance in each array is same. The common variance of estimate of y in each array is then given by $\sigma_2^2(1-\rho^2)$, ρ being the correlation coefficient between variables x and y is independent of x.

iii) The distribution of y in different arrays is normal.

Suppose that one of the variates x, is distributed normally with mean 0 and S.D σ_1 so that the probability that a random random value of x will fall in the small interval dx is

$$g(x)\,dx = \frac{1}{\sigma_1\sqrt{(2\pi)}}\exp\left(-x^2/2\sigma_1^2\right)dx$$

The probability that a value of y, taken at random in an assigned vertical array will fall in the interval dy is:

$$h(y/x)\,dy = \frac{1}{\sigma_2\sqrt{2\pi(1-\rho^2)}}\exp\left\{-\frac{1}{2\sigma_2^2(1-\rho^2)}\left(y-\rho\frac{\sigma_2}{\sigma_1}x\right)^2\right\}$$

The J.p.d of x and y is given by.

$$dp(x,y) = g(x)\,h(y/x)\,dx\,dy.$$

$$= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}}\exp\left(-\frac{1}{2\sigma_1^2}x^2\right)\exp\left[-\frac{1}{2\sigma_2^2(1-\rho^2)}\left(y-\rho\frac{\sigma_2}{\sigma_1}x\right)^2\right]dx\,dy$$

$$= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2}-\frac{2\rho xy}{\sigma_1\sigma_2}+\frac{y^2}{\sigma_2^2}\right)\right\}dx\,dy.$$

Shifting the origin to (μ_1,μ_2), we get.

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}}\exp\left[-\frac{1}{2(1-\rho^2)}\left\{\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right\}\right],$$

Where $\mu_1, \mu_2, \sigma_1(>0), \sigma_2(>0)$, and $\rho\ (-1<\rho<1)$ are five parameters of the distribution.

MGF of BND:

Let $(X,Y) \sim BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, By def.

$$M_{XY}(t_1,t_2) = E\left[e^{t_1 x + t_2 y}\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp(t_1 x + t_2 y) f(x,y)\, dx\, dy.$$

Put $\frac{x-\mu_1}{\sigma_1} = u$, $\frac{y-\mu_2}{\sigma_2} = v$, $-\infty < (u,v) < \infty$

ie, $x = \sigma_1 u + \mu_1$, $y = \sigma_2 v_2 + \mu_2 \Rightarrow |J| = \sigma_1\sigma_2$

$$\therefore M_{x,y}(t_1,t_2) = \frac{\exp(t_1\mu_1 + t_2\mu_2)}{2\pi\sqrt{1-\rho^2}} \times \int_u\int_v \exp\left[t_1\sigma_1 u + t_2\sigma_2 v - \frac{1}{2(1-\rho^2)}\{u^2 - 2\rho uv + v^2\}\right] du\, dv$$

$$= \frac{\exp(t_1\mu_1 + t_2\mu_2)}{2\pi\sqrt{1-\rho^2}} \int_u\int_v \exp\left[\frac{1}{2(1-\rho^2)} \times \{(u^2 - 2\rho uv + v^2) - 2(1-\rho^2)(t_1\sigma_1 u + t_2\sigma_2 v)\}\right] du\, dv$$

We have, $(u^2 - 2\rho uv + v^2) - 2(1-\rho^2)(t_1\sigma_1 u + t_2\sigma_2 v)$

$$= \{(u - \rho v) - (1-\rho^2)t_1\sigma_1\}^2 + (1-\rho^2)\{(v - \rho t_1\sigma_1 - t_2\sigma_2)^2 - t_1^2\sigma_1^2 - t_2^2\sigma_2^2 - 2\rho t_1 t_2 \sigma_1\sigma_2\} \quad \text{--- (1)}$$

By taking

$$\left.\begin{aligned} u - \rho v - (1-\rho^2)t_1\sigma_1 &= \omega(1-\rho^2)^{1/2} \\ \text{and} \quad v - \rho t_1\sigma_1 - t_2\sigma_2 &= z \end{aligned}\right\} \Rightarrow du\, dv = \sqrt{1-\rho^2}\, d\omega\, dz$$

and using (1), we get.

$$M_{x,y}(t_1,t_2) = \exp\left[t_1\mu_1 + t_2\mu_2 + \tfrac{1}{2}(t_1^2\sigma_1^2 + t_2^2\sigma_2^2 + 2\rho t_1 t_2 \sigma_1\sigma_2)\right] \times \left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\omega^2/2}\, d\omega\right] \times \left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-z^2/2}\, dz\right]$$

$$= \exp\left[t_1\mu_1 + t_2\mu_2 + \tfrac{1}{2}(t_1^2\sigma_1^2 + t_2^2\sigma_2^2 + 2\rho t_1 t_2 \sigma_1\sigma_2)\right]$$

In particular if $(X,Y) \sim BVN(0,0,1,1,\rho)$, then

$$M_{x,y}(t_1,t_2) = \exp\left[\tfrac{1}{2}(t_1^2 + t_2^2 + 2\rho t_1 t_2)\right]$$

Marginal Distribution of Bivariate Normal Distribution:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy$$

Put $\frac{y-\mu_2}{\sigma_2} = u$, then $dy = \sigma_2\,du$

$$\therefore f_X(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}}\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho u\left(\frac{x-\mu_1}{\sigma_1}\right) + u^2\right\}\right]\sigma_2\,du$$

$$= \frac{1}{2\pi\sigma_1\sqrt{(1-\rho^2)}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right]\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2(1-\rho^2)}\left\{u-\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\right\}^2\right]du$$

Put $\frac{1}{\sqrt{(1-\rho^2)}}\left[u-\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\right] = t$, then $du = \sqrt{(1-\rho^2)}\,dt$

$$\therefore f_X(x) = \frac{1}{2\pi\sigma_1}\exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right]\int_{-\infty}^{\infty}\exp(-t^2/2)\,dt$$

$$= \frac{1}{2\pi\sigma_1}\exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right]\sqrt{2\pi}$$

$$= \frac{1}{\sigma_1\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right]$$

Similarly, we get,

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx = \frac{1}{\sigma_2\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]$$

Hence $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$

Conditional Distributions:

The conditional distribution of X for fixed Y is given by.

$$f_{X/Y}(x/y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

$$= \frac{1}{\sigma_1\sqrt{2\pi}\sqrt{(1-\rho^2)}}\exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\left(1-(1-\rho^2)\right)\right\}\right]$$

$$= \frac{1}{\sigma_1\sqrt{2\pi}\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2(1-\rho^2)\sigma_1^2}\left\{(x-\mu_1) - \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)\right\}^2\right]$$

$$= \frac{1}{\sigma_1 \sqrt{2\pi} \sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)\sigma_1^2}\left\{x - \left(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)\right\}^2\right]\right.$$

which is the p.f of a univariate normal distribution with mean and variance given by.

$$E(X/Y=y) = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2) \text{ and } V(X/Y=y) = \sigma_1^2(1-\rho^2)$$

Hence the conditional distribution of X for fixed Y is also normal, given by.

$$(X/Y=y) \sim N\left[\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2), \sigma_1^2(1-\rho^2)\right]$$

Similarly the conditional distribution of r.v's Y for a fixed X is:

$$f_{Y/X}(y/x) = f_{XY}(x,y) / f_X(x)$$

$$= \frac{1}{\sigma_2\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)\sigma_2^2}\left\{(y-\mu_2) - \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right\}^2\right]$$

Thus the conditional distribution of Y for fixed X is also normal, given by.

$$(Y/X=x) \sim N\left[\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1), \sigma_2^2(1-\rho^2)\right]$$

UNIT-IV

Students t distribution

Let x_i $(i=1,2,\ldots n)$ be a random sample of size n from a normal population with mean μ and variance σ^2. Then Student's t is defined by the statistics.

$$t = \frac{\bar{x}-\mu}{S/\sqrt{n}}$$

where, $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, the sample mean. $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$ is an unbiased estimate of the population variance σ^2, and it follows Student's t population with $\nu = (n-1)$ d.f with probability density function

$$f(t) = \frac{1}{\sqrt{\nu}\, B\left(\frac{1}{2},\frac{\nu}{2}\right)} \cdot \frac{1}{\left[1+\frac{t^2}{\nu}\right]^{(\nu+1)/2}} \quad ; -\infty < t < \infty$$

Derivation of Students t-distribution

The expression can be rewritten as,

$$t^2 = \frac{n(\bar{x}-\mu)^2}{S^2} = \frac{n(\bar{x}-\mu)^2}{nS^2/(n-1)} \qquad [\because nS^2 = (n-1)S^2]$$

$$\Rightarrow \frac{t^2}{(n-1)} = \frac{(\bar{x}-\mu)^2}{\sigma^2/n}\,\frac{1}{nS^2/\sigma^2} = \frac{(\bar{x}-\mu)^2/(\sigma^2/n)}{nS^2/\sigma^2}$$

Since x_i, $(i=1,2,\ldots n)$ is a random sample from the normal population with mean μ and variance σ^2,

$$\bar{x} \sim N(\mu, \sigma^2/n) \Rightarrow \frac{(\bar{x}-\mu)}{\sigma/\sqrt{n}} \sim N(0,1)$$

Hence $\frac{(\bar{x}-\mu)^2}{\sigma^2/n}$, being the square of a standard normal variate is a chi square variate with 1 d.f. Also ns^2/σ^2 is a χ^2 variate with $(n-1)$ d.f.

Further since $\bar{x}$ and s^2 are independently distributed. $\frac{t^2}{n-1}$ being the ratio of two independent χ^2 variates with 1 and $(n-1)$ d.f. respectively, is a $\beta_2(\frac{1}{2}, \frac{n-1}{2})$ variate and its distribution is given by

$$dF(t) = \frac{1}{B(\frac{1}{2}, \frac{v}{2})} \cdot \frac{(t^2/v)^{\frac{1}{2}-1}}{[1 + t^2/v]^{(v+1)/2}} d(t^2/v), \quad 0 \leq t^2 < \infty$$

$[\text{where } v = (n-1)]$

$$= \frac{1}{\sqrt{v}\, B(\frac{1}{2}, \frac{v}{2})} \frac{1}{[1 + t^2/v]^{(v+1)/2}} dt; \quad -\infty < t < \infty$$

the factor 2 disappearing since the integral from $-\infty$ to ∞ must be unity. This is the required probability function as given in Student's 't' distribution with $v = (n-1)$ d.f.

Application of t-distribution

The t-distribution has a wide number of application in Statistics, some of which are enumerated below,

i) The t-distribution. To test the sample mean $(\bar{x})$ differs significantly from hypothesis value μ of the population mean.

ii) To test the significance of the difference between the two sample means.

iii) To test the significance of an observed sample correlation co-efficient and sample regression coefficient.

iv) To test the significance of observed partial correlation coefficient.

Chi-square distribution

The square of a standard normal variate is known as a chisquare variate (pronounced ki sky without s) with 1 d.f. $X \sim N(\mu, \sigma^2)$, then,

$$Z = \frac{\bar{x}-\mu}{\sigma} \sim N(0,1)$$

$Z^2 = \left(\frac{x-\mu}{\sigma}\right)^2$ is a chisquare variate with 1 d.f. $\chi^2 = \sum_{i=1}^{n}\left(\frac{X_i-\mu_i}{\sigma_i}\right)^2$ is a chisquare variate with n d.f.

Derivation of chisquare distribution

If X_i, $(i=1,2,\ldots n)$ are independent $N(\mu_i, \sigma_i^2)$ the distribution of,

$$\chi^2 = \sum_{i=1}^{n}\left[\frac{X_i-\mu_i}{\sigma_i}\right]^2 = \sum_{i=1}^{n} u_i^2 \quad \text{where, } u_i = \frac{X_i-\mu_i}{\sigma_i} \sim N(0,1)$$

Since x_i's are independent u_i's are also independent.

$$M_{u_i^2}(t) = M_{\sum u_i^2}(t) = \prod_{i=1}^{n} M_{u_i^2}(t)$$

$$= \left[M_{u_i^2}(t)\right]^n$$

$$M_{u_i^2}(t) = E\left[\exp\{t u_i^2\}\right]$$

$$= \int_{-\infty}^{\infty} \exp(t u_i^2) f(x_i)\, dx_i$$

$$= \int_{-\infty}^{\infty} \exp(tu_i^2) \frac{1}{\sigma\sqrt{2\pi}} \exp\{-(x_i-\mu)^2/2\sigma^2\} dx_i$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(tu_i^2) \exp(-u_i^2/2) du_i \qquad \left[u_i = \frac{x_i-\mu}{\sigma}\right]$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\left(\frac{1-2t}{2}\right) u_i^2\right\} du_i$$

$$= \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\pi}}{\left(\frac{1-2t}{2}\right)^{1/2}} = (1-2t)^{-1/2}$$

$$\therefore M_{\chi^2}(t) = (1-2t)^{-n/2}$$

Which is the m.g.f of a gamma variate with parameters $\frac{1}{2}$ and $\frac{1}{2}n$.

Hence by uniqueness theorem of m.g.f's.

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{x_i-\mu_i}{\sigma_i}\right)^2$$ is a gamma variate with parameters $\frac{1}{2}$

and $\frac{1}{2}n$.

$$\therefore dP(\chi^2) = \frac{(1/2)^{n/2}}{\Gamma(n/2)} \cdot \left[\exp(-\tfrac{1}{2}\chi^2)\right] (\chi^2)^{(n/2)-1} d\chi^2$$

$$= \frac{1}{2^{n/2}\Gamma(n/2)} \left[\exp(-\chi^2/2)\right] (\chi^2)^{(n/2)-1} d\chi^2, \quad 0 \le \chi^2 < \infty.$$

Which is the required probability density function of chi-square distribution with n degrees of freedom.

Additive property of χ^2- variates

The sum of independent chisquare variate is also χ^2 variate more precisely, if x_i $(i=1,2\ldots k)$ are independent χ^2 variate with n_i d.f. respectively then the sum $\sum_{i=1}^{k} x_i$ is also a chisquare variate with $\sum_{i=1}^{k} n_i$ d.f.

Proof

We have,

$$M_{X_i}(t) = (1-2t)^{-n_i/2} \qquad i = 1, 2, \dots k$$

The m.g.f of the sum $\sum_{i=1}^{k} X_i$ is given by,

$$M_{\Sigma X_i}(t) = M_{X_1}(t)\, M_{X_2}(t) \dots M_{X_n}(t) \qquad [\because X_i\text{'s are independent}]$$

$$= (1-2t)^{-n_1/2}\,(1-2t)^{-n_2/2} \dots (1-2t)^{-n_k/2}$$

$$= (1-2t)^{-(n_1+n_2+\dots+n_k)/2}$$

Which is the m.g.f of a χ^2 variate with $(n_1+n_2+\dots+n_k)$ d.f. Hence by uniqueness theorem of m.g.f's $\sum_{i=1}^{k} X_i$ is a χ^2 variate with $\sum_{i=1}^{k} n_i$ d.f

F - distribution

If X and Y are two independent chisquare variates with v_1 and v_2 d.f respectively then F statistics is defined by,

$$F = \frac{X/v_1}{Y/v_2}$$

In other words, F is defined as the ratio of two independent chisquare variates divided by their respective degrees of freedom and it follows Snedecor's F distribution with (v_1, v_2) d.f.

$$f(F) = \frac{(v_1/v_2)^{v_1/2}}{B\left(\frac{v_1}{2}, \frac{v_2}{2}\right)} \frac{F^{v_1/2 - 1}}{\left[1 + F\frac{v_1}{v_2}\right]^{(v_1+v_2)/2}} \qquad 0 \le F < \infty$$

Derivation of Snedecor's F distribution

Since x and y are independent chi square variates with v_1 and v_2 d.f respectively, their joint probability differential is given by,

$$dF(x,y) = \left\{\frac{1}{2^{v_1/2}\Gamma(v_1/2)} \exp(-x/2)\, x^{(v_1/2)-1}\, dx\right\}$$

$$\times$$

$$\left\{\frac{1}{2^{v_2/2}\Gamma(v_2/2)} \exp(-y/2)\, y^{(v_2/2)-1}\, dy\right\}$$

$$= \frac{1}{2^{(v_1+v_2)/2}\,\Gamma(v_1/2)\,\Gamma(v_2/2)} \exp\{-(x+y)/2\} \times x^{(v_1/2)-1}\, y^{(v_2/2)-1}\, dx\, dy$$

$$0 \le (x,y) < \infty$$

Let us make the following transformation of variables.

$$F = \frac{x/v_1}{y/v_2} \text{ and } u = y, \text{ so that } 0 \le F < \infty,\ 0 < u < \infty.$$

$$\therefore x = v_1/v_2\, Fu = \frac{v_1}{v_2} Fu \text{ and } y = u$$

Jacobian of transformation J is given by,

$$J = \frac{\partial(x,y)}{\partial(F,u)} = \begin{vmatrix} v_1/v_2\, u & 0 \\ v_1/v_2\, F & 1 \end{vmatrix} = \frac{v_1 u}{v_2}$$

Thus the distribution of the transformed variable is,

$$dG(F,u) = \frac{1}{2^{(v_1+v_2)/2}\,\Gamma(v_1/2)\,\Gamma(v_2/2)} \exp\left\{-u/2\left(1+v_1/v_2\, F\right)\right\} \times \left(v_1/v_2\, Fu\right)^{(v_1/2)-1}$$

$$u^{(v_2/2)-1}\, |J|\, du\, dF.$$

$$= \frac{(v_1/v_2)^{v_1/2}}{2^{(v_1+v_2)/2}\,\Gamma(v_1/2)\,\Gamma(v_2/2)} \exp\left\{-u/2\left(1+v_1/v_2 F\right)\right\} \times u^{\{(v_1+v_2)/2\}-1} F^{(v_1/2)-1}$$

$du\, dF \quad 0<u<\infty,\ 0\leq F<\infty$

Integrating out u over the range 0 to ∞, the distribution of F becomes

$$g_1(F)dF = \frac{(v_1/v_2)^{(v_1/2)} F^{(v_1/2)-1}}{2^{(v_1+v_2)/2}\,\Gamma(v_1/2)\,\Gamma(v_2/2)} \times \frac{\Gamma[(v_1+v_2)/2]}{\left[\frac{1}{2}\left(1+v_1/v_2 F\right)\right]^{(v_1+v_2)/2}} dF$$

$$\therefore g_1(F) = \frac{(v_1/v_2)^{v_1/2}}{B\left(\frac{v_1}{2},\frac{v_2}{2}\right)} \cdot \frac{F^{(v_1/2)-1}}{\left[1+v_1/v_2 F\right]^{(v_1+v_2)/2}}, \quad 0\leq F<\infty$$

Which is the required probability function of F distribution with (v_1, v_2) d.f

Relation between F and χ^2 distribution

In $F(n_1, n_2)$ distribution if we, Let $n_2\to\infty$ then $\chi^2 = n_1 F$ follows χ^2 distribution with n_1 d.f.

Proof

$$f(F) = \frac{(n_1/n_2)^{n_1/2} F^{(n_1/2)-1}}{\Gamma n_1/2\ \Gamma n_2/2} \cdot \frac{\Gamma(n_1+n_2)/2}{(1+n_1/n_2 F)^{(n_1+n_2)/2}} \quad 0<F<\infty$$

$$\frac{\Gamma(n_1+n_2)/2}{n_2^{n_1/2}\,\Gamma n_2/2} \to \frac{(n_2/2)^{n_1/2}}{n_2^{n_1/2}} = \frac{1}{2^{n_1/2}} \longrightarrow ①$$

$$\lim_{n_2\to\infty} (1+n_1/n_2 F)^{n_1+n_2/2} = \lim_{n_2\to\infty} \left[(1+n_1/n_2 F)^{n_2}\right]^{1/2} \times \lim_{n_2\to\infty} (1+n_1/n_2 F)$$

$$\Rightarrow \exp(n_1 F/2) = \exp(\chi^2/2) \longrightarrow ②$$

Relation between t and F distribution

In F distribution with (ν_1, ν_2) d.f, take $\nu_1 = 1$, $\nu_2 = \nu$ and $t^2 = F$

i.e., $dF = 2t\,dt$. Thus, the probability differential of F transforms to

$$dG(t) = \frac{(1/\nu)^{1/2}}{B(1/2, \nu/2)} \cdot \frac{(t^2)^{1/2 - 1}}{[1 + t^2/\nu]^{(\nu+1)/2}} \, 2t\,dt \qquad 0 \le t^2 < \infty$$

$$= \frac{1}{\sqrt{\nu}\, B(1/2, \nu/2)} \cdot \frac{1}{[1 + t^2/\nu]^{(\nu+1)/2}} \, dt, \qquad -\infty < t < \infty$$

the factor 2 disappearing since the total probability in the range $(-\infty, \infty)$ is unity. This is the probability function of Student's t distribution with ν d.f. Hence we have following relation between t and F distributions. If a statistic t follows student's t distribution with n d.f then t^2 follows Snedecor's F distribution with (1,n) d.f

Symbolically,

if $t \sim t(n)$

then $t^2 \sim F(1,n)$

Theorem

If X_1 and X_2 are independent χ^2 variates with n_1 and n_2 d.f respectively $u = \frac{X_1}{X_1 + X_2}$ and $v = X_1 + X_2$ are independent distribution u as a $\beta_1(n_1/2, n_2/2)$ and v as χ^2 variate with $(n_1 + n_2)$ d.f

Proof

$$dp(x_1, x_2) = \frac{1}{2^{(n_1+n_2)/2}\,\Gamma n_1/2\,\Gamma n_2/2}\ \exp\left\{-\frac{(x_1+x_2)}{2}\right\} \times (x_1)^{(n_1/2)-1}(x_2)^{n_2/2-1}$$

$$dx_1\, dx_2 \quad 0 \leq (x_1, x_2) < \infty$$

Unit - V

Order statistics

Let $x_1, x_2, \ldots x_n$ be n independent and identically distributed variance each with Cumulative distribution $F_X(x)$ in this variable arrange ascending order of magnitude and written as, $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$. We call $X_{(r)}$ as the rth order statistic $r = 1, 2, \ldots n$. The $X_{(r)}$'s because of the inequality relations among then are necessarily dependent.

Remark

If we write these ordered values as,

$Y_1 \leq Y_2 \leq \ldots \leq Y_n$ then,

$Y_r = X_r = r$th smallest of $X_1, X_2, \ldots X_n$

$Y_1 = X_1 =$ The smallest of $X_1, X_2, \ldots X_n$

$Y_n = X_{(n)} =$ The largest of $X_1, X_2, \ldots X_n$

Cumulative distribution Function of a single order Statistic

Let $F_r(x)$, $r = 1, 2 \ldots n$ denote the c.d.f of the rth order Statistic $X_{(r)}$. Then the c.d.f of the largest order statistic $X_{(n)}$ is given by,

$$F_n(x) = P(X_{(n)} \leq x) = P(X_i \leq x;\ i = 1, 2, \ldots n)$$

$$= P(X_1 \leq x \cap X_2 \leq x \cap \ldots \cap X_n \leq x)$$

$$= P(X_1 \leq x) \cdot P(X_2 \leq x) \ldots P(X_n \leq x) \quad (\because X_i\text{'s are independent})$$

$$= [F(x)]^n$$

Since $X_1, X_2 \dots X_n$ are identically distributed.

The c.d.f of the smallest order statistic $X_{(1)}$ is given by,

$$F_1(x) = P(X_{(1)} \leq x)$$

$$= 1 - P(X_{(1)} > x)$$

$$= 1 - P[X_i > x ; i = 1, 2 \dots n]$$

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

$$= 1 - [1 - F(x)]^n$$

Since $X_1, X_2 \dots X_n$ are i.i.d. rv's. In general, the c.d.f of the rth order statistic $X_{(r)}$ is given by,

$$F_r(x) = P(X_{(r)} \leq x)$$

$$= P[\text{At least } r \text{ of the } X_i\text{'s are } \leq x]$$

$$= \sum_{j=r}^{n} P[\text{exactly } j \text{ of the } n, X_i\text{'s are } \leq x]$$

$$= \sum_{j=r}^{n} \binom{n}{j} F^j(x) [1 - F(x)]^{n-j}$$

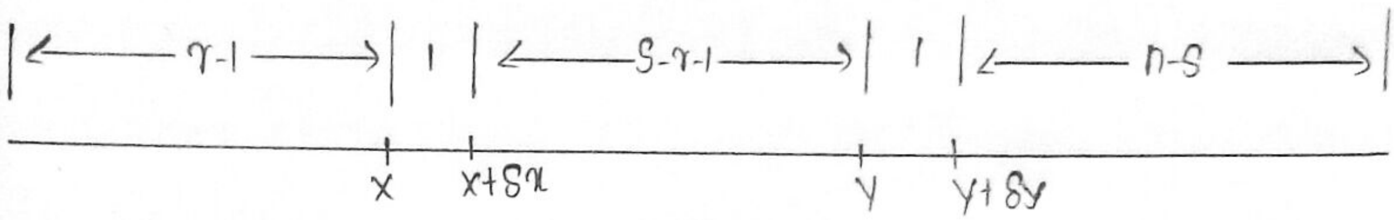
By using Binomial probability model.

Joint p.d.f of two order Statistics

Let us denote the joint p.d.f of $X_{(r)}$ and $X_{(s)}$, where $1 \leq r < s \leq n$ by $f_{rs}(x, y)$ then,

$$f_{rs}(x,y) = \lim_{\substack{\delta x \to 0 \\ \delta y \to 0}} \frac{P\left[x \le X_{(r)} \le x+\delta x \cap y \le X_{(s)} \le y+\delta y\right]}{\delta x\, \delta y} \longrightarrow (1)$$

The event $E = \{x \le X_{(r)} \le x + \delta x \cap y \le X_{(s)} \le y = \delta y\}$ can materialise as follows.



$X_i \le x$ for $r-1$ of the X_i's

$x < X_i \le x + \delta x$ for one X_i

$x + \delta x < X_i \le y$ for $(s-r-1)$ of X_i's

$y < X_i \le y + \delta y$ for one X_i, and $X_i > y + \delta y$ for $(n-s)$ of the X_i's.

Hence using multinomial probability law, we get

$$P(E) = \frac{P\left[x \le X_{(r)} \le x + \delta x \cap y \le X_{(s)} < y + \delta y\right]}{(r-1)!\,1!\,(s-r-1)!\,1!\,(n-s)!} P_1^{r-1} \cdot P_2 P_3^{s-r-1} P_4 P_5^{n-s} \longrightarrow (2)$$

where,

$P_1 = P(X_i \le x) = F(x)$

$P_2 = P(x < X_i \le x + \delta x) = F(x + \delta x) - F(x)$

$P_3 = P(x + \delta x < X_i \le y) = F(y) - F(x + \delta x)$

$P_4 = P(y < X_i \le y + \delta y) = F(y + \delta y) - F(y)$

$P_5 = P(X_i > y + \delta y) = 1 - P(X_i \leq y + \delta y) = 1 - F(y + \delta y)$

Substituting in ② and using ① we get,

$$f_{rs}(x, y) = \lim_{\substack{\delta x \to 0 \\ \delta y \to 0}} \frac{P(E)}{\delta x \, \delta y}$$

$$= \frac{n!}{(r-1)!\,(s-r-1)!\,(n-s)!} \times F^{r-1}(x) \times \lim_{\delta x \to 0} \frac{[F(x+\delta x) - F(x)]}{\delta x}$$

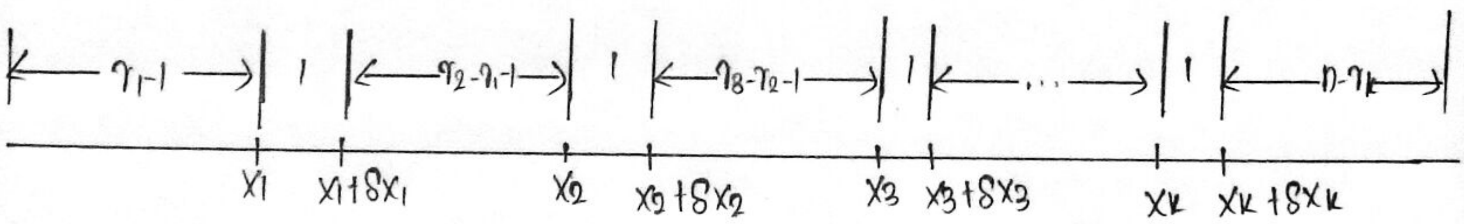
$$\times \lim_{\delta y \to 0} \left[\frac{F(y+\delta y) - F(y)}{\delta y}\right] \times \lim_{\delta y \to 0} [1 - F(y+\delta y)]^{n-s}$$

$$\times \lim_{\delta x \to 0} [F(y) - F(x+\delta x)]^{s-r-1}$$

$$= \frac{n!}{(r-1)!\,(s-r-1)!\,(n-s)!} F^{r-1}(x) . f(x) . [F(y) - F(x)]^{s-r-1} f(y) . [1 - F(y)]^{n-s}$$

Joint p.d.f of k order statistics

The joint p.d.f of k order statistics $X_{(r_1)}, X_{(r_2)} \dots X_{(r_k)}$ 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,



$$f_{r_1, r_2, \ldots r_k}(x_1, x_2, \ldots x_k) = \frac{n!}{(r_1-1)!\,(r_2-r_1-1)!\ldots(r_k-r_{k-1}-1)!\,(n-r_k)!}$$

$$\times F^{r_1-1}(x_1) \times f(x_1) \times [F(x_2)-F(x_1)]^{r_2-r_1-1} \times f(x_2) \times [F(x_3)-F(x_2)]^{r_3-r_2-1}$$

$$\times f(x_3) \times \ldots \times f(x_k)\,[1-F(x_k)]^{n-r_k} \longrightarrow ①$$

Joint p.d.f of all n - order Statistics

In particular the joint p.d.f of all the n order statistics is obtained on taking $k=n$ in ①. This implies that $r_i = l$ for $l = 1, 2 \ldots n$. Hence joint p.d.f of $X_{(1)}, X_{(2)} \cdots X_{(n)}$ is given by,

$$f_{1,2,\ldots n}(x_1, x_2 \ldots x_n) = n!\, f(x_1)\, f(x_2) \ldots f(x_n)$$

Distribution of range and other systematic statistics

Let us obtain the p.d.f the statistic $W_{rs} = X_{(s)} - X_{(r)}; r<s$

We start with the joint p.d.f of $X_{(r)}$ and $X_{(s)}$ given in and transform $[X_{(r)}, X_{(s)}]$ to the new variables W_{rs} and $X_{(r)}$ s.t.

$$W_{rs} = y - x; \quad x = x^{i} \quad \text{s.t}; \quad y = x + W_{rs} \text{ and } \bar{x} = x$$

$$\therefore J = \frac{\partial(x, y)}{\partial(x, W_{rs})} = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1 \Rightarrow |J| = 1$$

The joint p.d.f $f_{rs}(x,y)$ in transforms to the joint p.d.f of $X_{(r)}$ and W_{rs} as given below,

$$g(x, w_{rs}) = C_{rs} \cdot F^{r-1}(x) \cdot f(x) \left[F(x+w_{rs}) - F(x)\right]^{s-r-1} \times f(x+w_{rs}) \times \left[1 - F(x+w_{rs})\right]^{n-s} \longrightarrow (i)$$

where $$C_{rs} = \frac{n!}{(r-1)!\,(s-r-1)!\,(n-s)!} \longrightarrow (ii)$$

Integrating (i) w.r. to x from $-\infty$ to ∞ we obtain the p.d.f of W_{rs} as;

$$g(w_{rs}) = C_{rs} \int_{-\infty}^{\infty} \left\{F^{r-1}(x) f(x) \left[F(x+w_{rs}) - F(x)\right]^{s-r-1} \cdot f(x+w_{rs}) \left[1 - F(x+w_{rs})\right]^{n-s}\right\} dx.$$

Distribution of range $W = X_{(n)} - X_{(1)}$

Taking $r=1$ and $s=n$ in (ii) we obtain the p.d.f of the range $W = X_{(n)} - X_{(1)}$ as,

$$g(w) = n(n-1) \int_{-\infty}^{\infty} f(x) \left[F(x+w) - F(x)\right]^{n-2} f(x+w)\,dx \ ; \quad w \geq 0$$

The c.d.f of W is rather simple as given below,

$$G(w) = P(W \leq w) = \int_0^w g(u)\,du$$

$$= \int_0^w \left\{ n(n-1) \int_{-\infty}^{\infty} f(x) \left[F(x+u) - F(x) \right]^{n-2} f(x+u)\, dx \right\} du$$

$$= n \int_{-\infty}^{\infty} \left[f(x) \left\{ \int_0^w (n-1) f(x+u) (F(x+u) - F(x))^{n-2} du \right\} dx \right.$$

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