INDEX

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

* facts =>

- $a < b \Rightarrow \frac{1}{b} < \frac{1}{a}$, when a & b are of same sign.
- for $0 < x < \frac{\pi}{2}$,

$$0 < \cos x < \frac{\sin x}{x} < \frac{1}{\cos x}$$

- $\big||a| - |b|\big| \leq |a-b|$
- $\int_0^\infty x^m e^{-ax^2} dx = \frac{\Gamma\left(\frac{m+1}{2}\right)}{2(a)^{\frac{m+1}{2}}}$, $a>0$, $m>-1$
- $E(X^n) = \int_0^\infty n x^{n-1} P(X>x) dx$, for non-ve RV X (cont)
- If $X_1, \ldots, X_n$ r.s., [pop^n var $= \sigma^2$]

$$Var(\hat{S}^2) = \frac{1}{n}\left(\mu_4 - \frac{n-3}{n-1}\sigma^4\right), \; n>1$$

- $f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\phi(t)\, dt$ [$\phi \to$ cf].
- Fermat's Thm: If a is an integer & p be a prime, then $a^{p-1} \equiv 1 \pmod p$.
- Euler:- $a^{\phi(n)} \equiv 1 \pmod n$, if $\gcd(a,n) = 1$

* $\int_0^\infty e^{-ax^2} dx = \frac{1}{2}\sqrt{\frac{\pi}{a}}$.

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

(Σu_n) Series – Tests → Convergent, if →

- Ratio: $Lt \frac{u_{n+1}}{u_n} < 1$ [for +ve series]
- Cauchy (Root test): $Lt\ u_n^{1/n} < 1$. | Also, Comparison Test, Telescopic Sum.
- Raabe's: $Lt\ n\left(\frac{u_n}{u_{n+1}} - 1\right) > 1$
- Log: $Lt\left(n \log \frac{u_n}{u_{n+1}}\right) > 1$
- Gauss: $\frac{u_n}{u_{n+1}} = 1 + \frac{\mu}{n} + O\left(\frac{1}{n^2}\right) \rightarrow \mu > 1$
- De Morgan & Bertrand's,

$$Lt\left[\left\{n\left(\frac{u_n}{u_{n+1}} - 1\right) - 1\right\} \log n\right] > 1.$$

- Integral: $\int^{\infty} u_n\, dx$ is finite [non increasing terms]
- Pringshiem's thm: ($u_n \downarrow$) + Conv $\Rightarrow$ $Lt\ n u_n = 0$
- ✓ Cauchy Condensaⁿ: $\Sigma f(n)$ & $\Sigma a^n f(a^n)$ same nature [$a > 1$; $f(\cdot) \downarrow$]
- Leibnitz: $u_n \downarrow$ +ve seq & $\lim u_n = 0$.

↓ then $u_1 - u_2 + u_3 - u_4 \dots$ is conv.

Arbitrary Series:-

- Abel: $\{b_n\}$ monotone bdd & Σa_n is conv

Then, $\Sigma a_n b_n$ is conv.

- Dirichlet: $\{b_n\}$ monotone conv to 0, Σa_n bdd

(a_n / b_n +/+/-) Then, $\Sigma a_n b_n$ is conv.

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

Leibnitz' Rule for differentiating an integral:-

$$I(t) = \int_{g(t)}^{h(t)} f(x;t)\,dx$$

Then

$$\frac{dI}{dt} = \int_{g(t)}^{h(t)} \frac{\partial f}{\partial t}\,dx + f(h,t)\frac{dh}{dt} - f(g,t)\frac{dg}{dt}.$$

Part. Case:- $f(x;t)$ indep. of t:

$$\frac{d}{dt}\int_{g(t)}^{h(t)} f(x)\,dx = f(h)\frac{dh}{dt} - f(g)\frac{dg}{dt}.$$

Scanned by CamScanner
ABOUT.ME/SOURAV.NANDI

*** ⇒

(A) $(1-x)^{-n} = \sum_{j=0}^{\infty} \binom{-n}{j} (-x)^j$

$[-1<x<1]$

$= \sum_{j=0}^{\infty} \binom{n+j-1}{j} x^j$

(B) 2-Var Taylor Series :-

$f(x,y) = f(a,b) + f_x(a,b)(x-a)$

$+ f_y(a,b)(y-b) + \frac{1}{\lfloor 2}\Big[f_{xx}(a,b)(x-a)^2$

$+ 2 f_{xy}(a,b)(x-a)(y-b) + f_{yy}(a,b)(y-b)^2\Big]$

$+ \cdots + \frac{1}{\lfloor 3}\Big[\cdots\Big] + \cdots$

$\cdots + \frac{1}{\lfloor n}\left((x-a)\frac{\partial}{\partial x} + (y-b)\frac{\partial}{\partial y}\right)^n f\big(a+\theta(x-a),\ b+\theta(y-b)\big)$

(C) H be HM, $\frac{1}{H} = E\left(\frac{1}{x}\right)$

G be GM, $\ln G = E(\ln x)$

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

Distⁿ of fⁿ of R.V. ⇒

X & Y are indep. cont. R.V.s, then →

(i) pdf of $\overset{u}{U} = \overset{v}{X} + Y$ is given by,

$$h(u) = \int_{-\infty}^{\infty} f_X(v)\, f_Y(u-v)\, dv$$

(ii) pdf of $\overset{u}{U} = \overset{v}{X} - Y$ is,

$$h(u) = \int_{-\infty}^{\infty} f_X(v)\, f_Y(v-u)\, dv .$$

(iii) pdf of $\overset{u}{U} = \overset{v}{X}Y$ is,

$$h(u) = \int_{-\infty}^{\infty} f_X(v)\, f_Y\left(\frac{u}{v}\right) \cdot \frac{dv}{|v|}$$

(iv) pdf of $\overset{u}{U} = \frac{X}{Y} \overset{u}{}$, $Y \leftarrow v$

$$h(u) = \int_{-\infty}^{\infty} f_X(uv)\, f_Y(v)\, |v|\, dv .$$

Scanned by CamScanner
ABOUT.ME/SOURAV.NANDI

Disc/Cont Distr's

$X \sim U[a,b]$.

① **Uniform** :- $f_X(x) = \frac{1}{b-a}, \quad a \le x \le b$

$\rightarrow E(X) = \frac{b+a}{2}, \quad V(X) = \frac{(b-a)^2}{12}$

$$M_X(t) = \begin{cases} \frac{e^{bt} - e^{at}}{t(b-a)}, & t \neq 0 \\ 1, & t = 0. \end{cases}$$

$$F(x) = \begin{cases} \frac{x-a}{b-a}, & a \le x \le b \\ 0, & x < a \\ 1, & x > b. \end{cases}$$

$\mu_n'(a) = \frac{(b-a)^n}{n+1} \Rightarrow \mu_3 = 0,$

$\mu_4 = \frac{(b-a)^4}{80}$

$\therefore \gamma_1 = 0, \quad \gamma_2 = -1.2.$

So, symm & platykurtic.

$\Rightarrow \gamma_1 = \frac{\mu_3}{\mu_2^{3/2}}$; $\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3$

$\uparrow$ Skewness ; $\uparrow$ kurtosis

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

$\to \gamma(n, \theta)$

② Gamma Distⁿ : [n=1 → exponential distⁿ]

$f_X(x) = \frac{1}{\Gamma n} e^{-\frac{x}{\theta}} \left(\frac{x}{\theta}\right)^{n-1} \cdot \frac{1}{\theta}$, $x > 0$

$X \sim Gamma(\theta, n)$ $[\theta > 0, n > 0.]$

→ $E(X) = n\theta$, $V(X) = n\theta^2$

$\mu_r' = \theta^r \frac{\Gamma r+n}{\Gamma n}$, $r > -n$. [all non -ve moments exist]

VVR

$HM = [\mu_{-1}']^{-1} = \frac{1}{E(\frac{1}{X})} = \theta \cdot (n-1)$ $(n>1)$

→ $M_X(t) = (1-\theta t)^{-n}$, $(\theta t < 1)$

→ $K_r = n\theta^r \lfloor r-1$

Thus, $\gamma_1 = \frac{2}{\sqrt{n}}$, $\gamma_2 = \frac{6}{n}$.

Thus, +vely skew & leptokurtic

Scanned by CamScanner ABOUT.ME/SOURAV.NANDI

③ Beta Dist^n :→ MGF exists, but no compact form

$$f_X(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad \underline{0<x<1}$$

$\alpha, \beta > 0$

$$\mu_r' = \frac{B(\alpha+r, \beta)}{B(\alpha,\beta)}, \quad \longrightarrow \quad r > -\alpha$$

$\alpha > 1$

$$E(X) = \frac{\alpha}{\alpha+\beta}, \quad HM = \frac{\alpha-1}{\alpha+\beta-1}$$

So, $E(X) > V(X)$

$$V(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

GM be G, $\quad \ln G = \frac{\partial}{\partial\alpha}\left[\ln B(\alpha,\beta)\right]$

④. 2nd kind Beta :-

$$f_X(x) = \frac{1}{B(\alpha,\beta)} \cdot \frac{x^{\alpha-1}}{(1+x)^{\alpha+\beta}}, \quad x > 0$$

$\alpha, \beta > 0$

$$\mu_r' = \frac{B(\alpha+r, \beta-r)}{B(\alpha,\beta)}, \quad E(X) = \frac{\alpha}{\beta-1}$$

$-\alpha < r < \beta$ →

$$V(X) = \frac{\alpha(\alpha+\beta-1)}{(\beta-1)^2(\beta-2)}, \quad HM = \frac{\alpha-1}{\beta}.$$

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

If $X \sim$ 2nd kind $B(\alpha, \beta)$

Then, $Y = \frac{X}{1+X} \sim$ 1st kind $B(\alpha, \beta)$.

$B(m,n) = \int_0^{\pi/2} 2 \sin^{2m-1}\theta \cos^{2n-1}\theta \, d\theta$.

(5) Normal Distn :- $X \sim N(\mu, \sigma^2)$

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

$\mu_{2r+1} = 0$; $\mu_{2r} = \sigma^{2r}(2r-1)(2r-3)\cdots 3\cdot 1$

$\therefore \gamma_1 = 0, \ \gamma_2 = 0$

$MGF = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$

$Var(s^2) = \frac{2\sigma^4}{n-1}$

$MD_\mu = \sigma\sqrt{\frac{2}{\pi}}$; $V(|X-\mu|) = \sigma^2\left(1 - \frac{2}{\pi}\right)$

$$1 - \frac{1}{x^2} \le \frac{x\{1 - \Phi(x)\}}{\phi(x)} \le 1.$$

$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}, \ a > 0$

Distn of $|X|$ $(= Y)$:- $f_Y(y) = \begin{cases} 0, & y \le 0 \\ 2 f_X(y), & y > 0 \end{cases}$

$X \sim N(0, \sigma^2)$ ↳ folded normal distn.

Scanned by CamScanner
ABOUT.ME/SOURAV.NANDI

⑥ Log-Normal distⁿ :- ⊙

$X \sim \Lambda(\mu, \sigma^2)$ iff $\log_e x \sim N(\mu, \sigma^2)$

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2}, \; x > 0$$

$\mu_r' = e^{\mu r + \frac{1}{2}\sigma^2 r^2}$; $HM = e^{\mu - \frac{1}{2}\sigma^2}$

$\therefore E(X) = e^{\mu + \frac{1}{2}\sigma^2}$, $GM = e^{\mu}$.

$V(X) = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)$

$M_e = e^{\mu}$; M_0 (mode) $= e^{\mu - \sigma^2}$.

pth quantile, $\xi_p = e^{\mu + \sigma Z_p}$ (✓)

[Z_p : pth quantile of $N(0,1)$ distⁿ].

+vely skew.

[$\because QD > 0$]

$AM > GM > HM$

$e^{\mu + \frac{1}{2}\sigma^2} > e^{\mu} > e^{\mu - \frac{1}{2}\sigma^2}$

$\mu > M_e > M_0$

$e^{\mu - \sigma^2}$

$\boxed{\text{Discrete}} \Rightarrow$

$$\mu_{[r]} = \frac{d^r}{dt^r} P_x(t) \Big|_{t=1}.$$

$$\mu_r' = \frac{d^r}{dt^r} M_x(t) \Big|_{t=0}.$$

$\hookrightarrow$ coeff of $\frac{t^r}{\lfloor r}$ in $M_x(t) = E(e^{tx})$

$p(x) =$ Coeff of t^x in $P_x(t) = E(t^x)$

$K_r =$ Coeff of $\frac{t^r}{\lfloor r}$ in $K_x(t) = \ln M_x(t)$

$\mu_1' = K_1$; $\mu_2 = K_2$; $\mu_3 = K_3$; $\mu_4 = K_4 + 3K_2^2$

(7) Degenerate dist^n :

$p(x) = \begin{cases} 1, & x = c \\ 0, & \text{o.w.} \end{cases}$	$\mu_r' = c^r$
$K_x(t) = ct$	$E(X) = c$
$K_r = 0 \ \forall r > 1.$	$V(X) = 0$
$\Delta \ K_1 = ct.$	mgf $= e^{ct}.$

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

⑧ Uniform distⁿ:

$$p(x) = \begin{cases} \frac{1}{n}, & x = a + ih, \; i = 1(1)n \\ 0, & o.w. \end{cases}$$

$$E(X) = a + \frac{h(n+1)}{2} \; ; \; V(X) = \frac{h^2(n^2-1)}{12}$$

$$\mu_r = \frac{h^r}{n} \sum_{1}^{n} \left(i - \frac{n+1}{2}\right)^r$$

$$\mu_3 = 0$$

$$\mu_4 = \frac{h^4 (n^2-1)(3n^2-7)}{240}.$$

Thus, $\gamma_1 = 0$, $\gamma_2 = \beta_2 - 3 < 0$

where $\beta_2 = 1.8 \times \frac{n^2 - \frac{7}{3}}{n^2-1}$

$$MD_\mu = \begin{cases} |h| \times \frac{n^2-1}{4n}, & n = odd \\ |h| \times \frac{n}{4}, & n = even. \end{cases}$$

⑨ Bernoulli → Bin(1, p).

→ $\sum_{i=1}^{n} i^4 = \left(\sum i^2\right) \times \left(\frac{3n^2+3n-1}{5}\right)$

Scanned by CamScanner

(9) Binomial Distⁿ :-

$$p(x) = \begin{cases} \binom{n}{x} p^x q^{n-x}, & x = 0(1)n \\ 0, \text{ o.w.} & \begin{pmatrix} 0<p<1 \\ q=1-p \end{pmatrix} \end{cases}$$

$\mu'_{[r]} = (n)_r \, p^r$

$\mu'_r \Rightarrow \mu'_{[r]}$ [Coeff. : 1 ($\mu'_1 = \mu'_{[1]}$)
11, 131, 1671]

$\mu'_4 = \mu'_{[4]} + 6\mu'_{[3]} + 7\mu'_{[2]} + \mu'_{[1]}$

$\mu_{r+1} = pq\left[nr\mu_{r-1} + \frac{d\mu_r}{dp}\right], r = 1, 2 \ldots$

$\therefore \mu'_1 = np, \quad \mu_2 = npq \quad [\mu_0 = 1, \mu_1 = 0]$

$\mu_3 = npq(q-p)$, &

$\mu_4 = 3n^2p^2q^2 + npq(1-6pq)$

$\therefore \gamma_1 = \frac{q-p}{\sqrt{npq}}, \quad \gamma_2 = \frac{1-6pq}{npq}$

$\therefore$ +vely skew / symm / −vely skew as $p \lesseqgtr \frac{1}{2}$

l/m/p as $pq \lesseqgtr \frac{1}{6}$; As $n \to \infty$,
$(\gamma_1, \gamma_2) \to (0, 0)$.

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

For $p \neq \frac{1}{2}$, $\beta_2 = c\beta_1 + 3$,

$$c = 1 - \frac{2pq}{(q-p)^2} < 1.$$

$$\text{Mode} = \begin{cases} k \text{ \& } k-1 & \text{if } k = \text{integer} \\ [k] & \text{, if } k = \text{frac}^n \end{cases}$$

$k = (n+1)p$

$M_x(t) = (q + pe^t)^n$; $P_x(t) = (q + pt)^n$

$MD_\mu = 2npq \times \binom{n-1}{k} p^k q^{n-1-k}$

$\rightarrow \sqrt{\frac{2npq}{\pi}}$, as $n \to \infty$ $\quad (k = [np])$

$\rightarrow$ $P(X \leq k) = P(Y \leq q)$, $[q = 1-p]$

where $X \sim Bin(n,p)$; $Y \sim \beta(n-k, k+1)$

$\hookrightarrow$ use $\binom{n}{x} = \binom{k}{x} \times \frac{B(n-k, k-x+1)}{B(n-k, k+1)}$

$X \sim Bin(n,p) \Rightarrow n - X \sim Bin(n, q)$

$K_{r+1} = pq \frac{d}{dp}(K_r)$ with $K_1 = np$.

$MD_\mu = 2p(n-k) f(k) \rightarrow$ pmf of $Bin(n,p)$.

Fitting :- $\frac{p(x)}{p(x-1)} = \frac{n-x+1}{x} \times \frac{p}{q}$, $p(0) = q^n$

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

(10) Poisson Distⁿ :

$$p(x) = \begin{cases} e^{-\lambda}\dfrac{\lambda^x}{\lfloor x}, & \lambda > 0,\ x = 0,1,2\ldots \\ 0, & o.w. \end{cases}$$

We can take $\lambda = \nu t$ ($\nu \to$ mean rate ~~of occurrence~~)

$\lambda \to$ mean no of occ.

$\mu'_{[r]} = \lambda^r$, $\boxed{K_r = \lambda}$; $\gamma_1 = \frac{1}{\sqrt{\lambda}}$, $\gamma_2 = \frac{1}{\lambda}$

$\mu'_1 = \lambda$, $\mu_2 = \lambda$, $\mu_3 = \lambda$, $\mu_4 = 3\lambda^2 + \lambda$

$M_X(t) = e^{-\lambda} \cdot e^{\lambda e^t}$; $P_X(t) = e^{-\lambda} \cdot e^{\lambda t}$

$MD_\mu = 2\lambda e^{-\lambda}\frac{\lambda^k}{\lfloor k}$, $k = [\lambda]$.

$= 2\lambda f(k)$

$$P(X \leq k) = P(Y > \lambda) \left[= 1 - \overbrace{I\left(\frac{\lambda}{\sqrt{k+1}}, k\right)}^{P(Y \leq \lambda)} \right]$$

$X \sim Poisson(\lambda)$; $Y \sim Gamma(k+1)$.

where $I(u,k) = \dfrac{\int_0^\lambda e^{-t} t^k dt}{\Gamma k+1} = P(Y \leq \lambda)$

$$u = \frac{\lambda}{\sqrt{k+1}}$$

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

If $X_1 \sim \text{Poisson}(\lambda_1)$
$X_2 \sim \text{Poisson}(\lambda_2)$ } indep.

Then, (i) $X_1 + X_2 \sim \text{Poisson}(\lambda_1 + \lambda_2)$

(ii) $X_1 | (X_1 + X_2) \sim \text{Bin}\left(x, \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)$

$$\mu'_{r+1} = \lambda \sum_{j=0}^{r} \binom{r}{j} \mu'_j .$$

$$\text{Mode} = \begin{cases} \lambda, \lambda - 1 & \text{if } \lambda \in I^+ \\ [\lambda], & \text{o.w.} \end{cases}$$

$$\mu_{r+1} = \lambda \left[r \mu_{r-1} + \frac{d\mu_r}{d\lambda} \right]$$

$p(x-1) \lesseqgtr p(x)$, x .....

Fitting: $\dfrac{p(x)}{p(x-1)} = \dfrac{\lambda}{x}$, $p(0) = e^{-\lambda}$

Scanned by CamScanner ABOUT.ME/SOURAV.NANDI

$\rightarrow H(N, Np, n)$

(11) Hypergeometric Dist^r :

$$pmf = \frac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}}, \; x = 0(1)n$$

$Hyp(N, n, p)$

$E(X) = np$; $\rightarrow V(X) = npq \cdot \frac{N-n}{N-1}$

$$\mu'_{[r]} = \frac{(n)_r \cdot (Np)_r}{(N)_r}$$

Mode :- $\boxed{k = \frac{(n+1)(Np+1)}{N+2}}$

$\hookrightarrow \begin{cases} k, k-1 & \text{when } k \in \mathbb{I}^+ \\ [k] & \text{, otherwise} \end{cases}$

$$MD_\mu = \frac{2(Np-k)(n-k)}{N} f(k)$$

$= 2(p - \frac{k}{N})(n-k) f(k)$ Hyp with $x = k$

~~If X_1 & Y are~~

If $X_1 \sim Bin(n_1, p)$; $X_2 \sim Bin(n_2, p)$ (indep.), Then $(X_1 | X_1 + X_2 = x)$

$x_1 = 0(1)x$ $\Big|$ $\sim Hyp(n_1 + n_2, x, \frac{n_1}{n_1+n_2})$

Scanned by CamScanner
ABOUT.ME/SOURAV.NANDI

(12) Geometric :- ($r=1$ in -ve Bin) $\rightarrow$ Geometric(p)

$f(x) = pq^x, \; x = 0, 1, 2, \ldots$

$F(x) = 1 - q^{x+1}$, $\sum_{x=0}^{\infty} q^x = (1-q)^{-1}$

$E(x) = \frac{q}{p}$ $\therefore \sum_{1}^{\infty} x q^{x-1} = (1-q)^{-2}$

$V(x) = \frac{q}{p^2}$ $\therefore \sum_{2}^{\infty} x(x-1) q^{x-2} = 2(1-q)^{-3}$.

Mode, $\mu_0 = 0$

Memory Loss Prop :-

$$P(X \leq t+k \mid X \geq t) = P(X \leq k)$$

$$MGF = \frac{p}{1-qe^t} \; ; \; PGF = \frac{p}{1-qt}.$$

① X_1, X_2 are iid geometric Γ pmf, $f(y) = \frac{1}{1+x}$, $y = 0(1)x$.

$\Rightarrow (X_1 \mid X_1 + X_2 = x)$ is uniform

ABOUT.ME/SOURAV.NANDI

(13) Negative Binomial distⁿ ⇒

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

$$= \binom{-r}{x} p^r (-q)^x \rightarrow -ve\ Bin(r, p)$$

from $(Q-P)^{-r} \rightarrow$ $$= \binom{r+x-1}{x} P^x Q^{-(r+x)} \quad \left[p = \frac{1}{Q} \quad q = \frac{P}{Q} \right]$$

$$\text{Mode} = \begin{cases} k \ \& \ k-1 & \text{if } k \in I^+ \\ [k] & \text{o.w.} \end{cases}, \text{ where}$$

$$k = (r-1)\frac{q}{p}.$$

$$MD_\mu = 2\frac{(k+1)}{p} f(k+1) = 2(r+k)\frac{q}{p} f(k)$$

$$\mu'_{[k]} = (r+k-1)_{k} \left(\frac{q}{p}\right)^k$$

R

$$\mu = \frac{rq}{p} \quad , \quad Var = \frac{rq}{p^2}$$

$$\mu_{k+1} = q\left[\frac{rk}{p^2}\mu_{k-1} - \frac{d\mu_k}{dp}\right]$$

$\rightarrow$ $$P(X \le k) = P(Y \le p) = I_p(r, k+1)$$

$$M_X(t) = (Q - Pe^t)^{-r} = p^r(1 - qe^t)^{-r}$$

$$P_X(t) = p^r(1 - qt)^{-r}$$

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

$$\mu_3 = \frac{rq(1+q)}{p^3},$$

$$\mu_4 = \frac{rq}{p^4}\left(p^2 + 3rq + 6q\right)$$

Fitting :-

$$\frac{p(x)}{p(x-1)} = \frac{r+x-1}{x}\,q \; ; \; p(0) = p^r$$

∴ +vely skew & leptokurtic

Scanned by CamScanner
ABOUT.ME/SOURAV.NANDI

$k, k-1$ if $k+1$ ... $\}$; $[k]$, o.w. — if $k = [\mu]$

	Mode	MD μ
Bin ⇒	$k = (n+1)p$,	$2p(n-k)f(k)$
Poisson ⇒	$k = \lambda$,	$2\lambda f(k)$
Hyp ⇒	$k = (n+1)\left(\frac{Np+1}{N+2}\right)$,	$2\left(p - \frac{k}{N}\right)(n-k) f(k)$
-ve Bin ⇒	$k = (r-1)\frac{q}{p}$,	$2\frac{(k+1)}{p} f(k+1)$

⊛ $$E(X^n) = \int_0^\infty n \cdot x^{n-1} P[X > x]\, dx$$

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

(14) Power Series Dist^n:-

$$p(x) = a_x \frac{\theta^x}{g(\theta)}, \quad a_x, \theta, g(\theta) > 0, \quad \boxed{x = 0, 1, 2, \ldots}$$

Cases:- (i) Bin:- $\theta = \frac{p}{q}$, $g(\theta) = (1+\theta)^n$

(ii) Poisson:- $\theta = \lambda$, $g(\theta) = e^{+\theta}$.

(iii) -ve Bin:- $\theta = q$, $g(\theta) = (1-\theta)^{-r}$

Here, $E(X) = \theta \frac{d}{d\theta}[\ln g(\theta)]$,

$$V(X) = E(X) + \theta^2 \frac{d^2}{d\theta^2}[\ln g(\theta)]$$

$MGF = \frac{g(\theta e^t)}{g(\theta)}$; $PGF = \frac{g(\theta t)}{g(\theta)}$

Recursion Rela^n:- $\mu_{k+1} = \theta \frac{d\mu_k}{d\theta} + k\mu_2 \mu_{k-1}$.

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

Cont. Dist^n s ⇒

(15) Cauchy Dist^n: $X \sim C(\mu, \sigma)$

$$f_X(x) = \frac{\sigma}{\pi[\sigma^2 + (x-\mu)^2]}, \quad x \in \mathbb{R}$$

$(\mu \in \mathbb{R}, \sigma > 0)$

$$F_X(x) = \frac{1}{\pi}\left[\tan^{-1}\left(\frac{x-\mu}{\sigma}\right) + \frac{\pi}{2}\right]$$

Median $= \mu$ $=$ mode | Heavy tailed distn.

⊙ Only fractional order moments exist: $E(X^r) = \frac{1}{\pi} B\left(\frac{1-r}{2}, \frac{r+1}{2}\right)$, $-1 < r < 1$

↳ If $X \sim C(0,1) \Rightarrow$

Standard Cauchy dist^n, $\rightarrow F_X(x) = \frac{1}{\pi}\left[\tan^{-1}x + \frac{\pi}{2}\right]$

(platykurtic dist^n) $f_X(x) = \frac{1}{\pi(1+x^2)}$, $x \in \mathbb{R}$.

MGF does not exist $[CF = e^{-|t|}]$

Mean & Var doesn't exist

① (16) Shifted Exponential distⁿ ⇒

$$f_X(x) = \frac{1}{\theta} e^{-\frac{x-a}{\theta}}, \; x \geq a \; (\theta > 0).$$

$$F_X(x) = \begin{cases} 1 - e^{-\frac{x-a}{\theta}}, & x > a \\ 0, & x \leq a \end{cases} \quad \Big\| \quad \boxed{\mu = a + \theta}$$

$\xi_p = a - \ln(1-p)\theta$; $MD_\mu = \theta \ln 2$

$QD = \frac{\theta}{2} \ln 3$; $\mu_r'(a) = \theta^r \lfloor r$ $(r > -1)$

$\mu = a + \theta$; $Var = \theta^2$; $Me = a + \theta \ln 2$,

$\mu_3 = 2\theta^3$, $\mu_4 = 9\theta^4$

$\gamma_1 = 2$, $\gamma_2 = 6$

~~$M_X(t) = f(t)\theta$~~ → $X \sim exp(\theta)$

Exp Distⁿ :- $\frac{1}{\theta} e^{-\frac{x}{\theta}}$, $x > 0$ also

$M_X(t) = (1-\theta t)^{-1}$, $\left(t < \frac{1}{\theta}\right)$ ⇒ $K_r = \theta^r \lfloor r-1$.

Loss of memory ~~distⁿ~~ : $P[X > a+b \mid X > a] = P[X > b]$

$\mu_r' = \theta^r \lfloor r$ $(r > -1)$. [↳ Gamma$(\theta, 1)$]

Scanned by CamScanner
ABOUT.ME/SOURAV.NANDI

(17) Laplace / Double-Exp distⁿ (μ, λ)

$$f_X(x) = \frac{1}{2\lambda} e^{-\left|\frac{x-\mu}{\lambda}\right|}, \ x \in \mathbb{R}$$

$$(\mu \in \mathbb{R}, \lambda > 0)$$

lepto-kurtic

$$= \begin{cases} \frac{1}{2\lambda} e^{-\left(\frac{x-\mu}{\lambda}\right)}, & x \geq \mu \\ \frac{1}{2\lambda} e^{\left(\frac{x-\mu}{\lambda}\right)}, & x < \mu \end{cases}$$

$$\mu_r'(\mu) = \begin{cases} 0, & r = \text{odd} \\ \lambda^r \, \underline{|r}, & r = \text{even} \end{cases}$$

$[\mu_r]$

Mode = median = mean = μ, $Var = 2\lambda^2$

$$F_X(x) = \begin{cases} \frac{1}{2} e^{\frac{x-\mu}{\lambda}}, & x < \mu \\ 1 - \frac{1}{2} e^{-\frac{x-\mu}{\lambda}}, & x \geq \mu \end{cases}$$

$$M_X(t) = \frac{e^{\mu t}}{1-\lambda^2 t^2}$$

$|t| < \frac{1}{\lambda}$

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

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

$\mu = 0, \lambda = 1 \Rightarrow$ Std. DE / Laplace distⁿ

$X \sim$ std Laplace $\Rightarrow |X| \sim$ std exp.

$\hookrightarrow M_X(t) = (1-t^2)^{-1}$

Scanned by CamScanner
ABOUT.ME/SOURAV.NANDI

18 Pareto Distn :- Parameters $P(x_0, \alpha)$

$$f_X(x) = \frac{\alpha}{x_0} \times \left(\frac{x_0}{x}\right)^{\alpha+1}, \quad \forall\ x \geq x_0.$$

$$P(x_0, \alpha) = \frac{\alpha\, x_0^{\alpha}}{x^{\alpha+1}} \quad \begin{pmatrix} x_0 > 0 \\ \alpha > 0. \end{pmatrix}$$

$$F_X(x) = \begin{cases} 0, & x < x_0 \\ 1 - \left(\frac{x_0}{x}\right)^{\alpha}, & x \geq x_0. \end{cases}$$

$\alpha > 1$

$$E(X) = \frac{\alpha x_0}{\alpha - 1}; \quad V(X) = \frac{\alpha x_0^2}{(\alpha-1)^2(\alpha-2)}, (\alpha > 2)$$

$$HM = x_0\left(1 + \frac{1}{\alpha}\right); \quad GM = x_0 \times e^{1/\alpha}.$$

$$M_e = x_0 \times 2^{1/\alpha}, \ \alpha > 0. \quad \Big| \quad M_0 = x_0$$

MGF does not exist | +vely skew

$$E|X|^r = \frac{\alpha x_0^r}{\alpha - r}, \quad r < \alpha$$

μ_r'

$$\cancel{\text{AM } x_0 \times \frac{\alpha}{\alpha-1} > \overset{GM}{x_0 \times e^{1/\alpha}} > \overset{HM}{x_0 \times \left(\frac{\alpha+1}{\alpha}\right)}}$$

$$(\mu) \text{ ,, } > \underset{(M_e)}{x_0 \times 2^{1/\alpha}} > x_0 \ (M_0)$$

Scanned by CamScanner
ABOUT.ME/SOURAV.NANDI

(19) <u>Logistic distⁿ</u> ; $L(\alpha, \beta)$

$$F_X(x) = \left[1 + e^{-\frac{x-\alpha}{\beta}}\right]^{-1}, \quad x \in \mathbb{R}, \quad (\alpha \in \mathbb{R}, \beta > 0)$$

(*) <u>Std. Logistic</u> :- Symm distⁿ.

$\alpha = 0, \beta = 1 \rightarrow F_Y(y) = \dfrac{1}{1+e^{-y}}$ ✓

$$\therefore f_Y(y) = \frac{e^{-y}}{(1+e^{-y})^2}, \quad y \in \mathbb{R}.$$

$MD_\mu = 2 \ln 2$. $[\mu = 0]$ ✓

MGF :- $M_Y(t) = \dfrac{\pi t}{\sin \pi t}$.

μ_2'

$Var = \dfrac{\pi^2}{3}$; $\mu_{2r+1} = 0 \ \forall r$.

Leptokurtic $\left[\because \mu_4 = \mu_4' = \dfrac{7\pi^4}{15} > 3\right]$.

Symmetrical

— ☺ —

Scanned by CamScanner
ABOUT.ME/SOURAV.NANDI

Scanned by CamScanner
ABOUT.ME/SOURAV.NANDI

Descriptive Measures

- Median, $\tilde{x} = l + \dfrac{\frac{n}{2} - F_l}{f_{\tilde{x}}} \times c$

l → lower boundary of the class interval containing median; F_l → cumulative freq; $f_{\tilde{x}}$ → freq; c → width

- Mode, $\check{x} = l + \dfrac{f_m - f_{m-1}}{2f_m - f_{m+1} - f_{m-1}} \times c$.

$MD_c = \frac{1}{n} \sum_i |x_i - c| f_i$

- $MD_{\bar{x}} = \frac{2}{n} \sum_{x_i > \bar{x}} (x_i - \bar{x}) = \frac{2}{n} \sum_{x_i < \bar{x}} (\bar{x} - x_i)$

- $s^2 = \frac{1}{n} \sum x_i^2 f_i - \left(\frac{1}{n} \sum x_i f_i\right)^2$.

→ For r grs of values, with ith gr having n_i values with mean $\bar{x}_i$, sd s_i,

$$s^2 = \frac{\sum n_i s_i^2}{\sum n_i} + \frac{\sum n_i (\bar{x}_i - \bar{x})^2}{\sum n_i} \quad \left[\bar{x} = \frac{\sum \bar{x}_i n_i}{\sum n_i} \right]$$

- $RMSD_c \geq S_x \geq MD_{\bar{x}} \geq |\bar{x} - me|$

$R \geq S \; ; \; \frac{R^2}{4} \geq s^2 \geq \frac{R^2}{2n}$.

→ $s^2 = \frac{1}{2n^2} \sum_i \sum_j (x_i - x_j)^2$

var of $x_1, x_2, \ldots, x_n$

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

•) Moments, [A=0 → raw, A= $\bar{x}$ → central]

$$m_r'(A) = \frac{1}{n}\sum(x_i - A)^r f_i \qquad (n = \sum_{i=1}^{k} f_i)$$

→ $m_1'(0) = \bar{x}$; $m_2 = s^2$; $\underline{m_r \to A = \bar{x}}$

•) $m_r = m_r'(a) - \binom{r}{1} m_{r-1}'(a)\, m_1'(a) + \cdots + (-1)^r m_1'^r(a)$

↳ $m_0 = 1$; $m_1 = 0$; $m_2 = m_2' - m_1'^2$,

$m_3 = m_3' - 3m_2' m_1' + 2m_1'^3$, ...

•) $m_r'(a) = m_r + \binom{r}{1} m_{r-1} d + \binom{r}{2} m_{r-2} d^2 + \cdots + d^r$

↳ $m_0' = 1$; $m_1' = d = \bar{x} - A$; $m_2' = m_2 + d^2$,

$m_3' = m_3 + 3m_2 d + d^3$, - -

•) Sheppard's Correction :- Givn P-181

•) factorial moment

$$m_{[r]}'(0) = \frac{1}{n}\sum_{1}^{n} x_i(x_i - 1) \cdots (x_i - r + 1)$$

•) For +vely skew distrⁿ, $\bar{x} > \tilde{x} > \hat{x}$

•) Skewness :- $Sk_2 = \frac{\bar{x} - \hat{x}}{s} \simeq \frac{3(\bar{x} - \tilde{x})}{s} = Sk_3$

$Sk_1 = \frac{m_3}{s^3} = g_1 = \sqrt{b_1}$, $Sk_4 = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}$ (Bowley)

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

*) Kurtosis :- $g_2 = b_2 - 3 = \frac{m_4}{m_2^2} - 3$.

$\rightarrow$ $b_2 \geq b_1 + 1$. [equality :- 2 diff values, same freq]

	B	β
A	a	b
α	c	d

Association

$$Q_{AB} = \frac{f_{AB} f_{\alpha\beta} - f_{A\beta} f_{\alpha B}}{f_{AB} f_{\alpha\beta} + f_{A\beta} f_{\alpha B}} = \frac{ad - bc}{ad + bc}$$

yule

$$Y_{AB} = \frac{\sqrt{ad} - \sqrt{bc}}{\sqrt{ad} + \sqrt{bc}} \quad ; \quad \delta_{AB} = f_{AB} - \frac{f_A \cdot f_B}{n}.$$

$$V_{AB} = \frac{ad - bc}{\sqrt{f_A \cdot f_\alpha \cdot f_B \cdot f_\beta}}$$ ← by pearson

Dispersion

$$\Delta_1 = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} |x_i - x_j| f_i f_j$$

↳ Gini's mean difference (for grouped data)

$$\Delta_2 = \frac{1}{n^2} \sum_i \sum_j (x_i - x_j)^2 f_i f_j$$

$$= 2s^2$$

Gini's Coeff of conc, $c = \frac{\Delta}{2\bar{x}}$.

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

→ p^{th} quantile,

$$Z_p = x_l + \frac{Np - F_l}{f_p} \times c$$

[See median formula ∵ $p = \frac{1}{2}$]

Diagrammatic Representation

(A) Freq type data :-

(i) Attribute :→ Bar diagram (Horizontal, divided)

• Pie Chart.

(ii) Variable :-

→ Grouped (Cont) :• Histogram, • Ogive,

• Freq. Polygon

→ Ungrouped (Discrete) :→ Column diagram,

• Step diagram,

(B) Non (freq type data) - (Variable) ⇒

• Bar diagram (Vertical, Horizontal, Multiple, Component)

• Line diagram

• Ratio chart.

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

Rank Correlation $[d_i = u_i - v_i]$

•) Spearman's

$$r_R = 1 - \frac{6 \sum d_i^2}{n(n^2-1)} \quad \text{[no tie case]}$$

$$= \frac{\frac{n^2-1}{12} - \frac{T_u + T_v}{2} - \frac{1}{2n} \sum d_i^2}{\sqrt{\frac{n^2-1}{12} - T_u} \times \sqrt{\frac{n^2-1}{12} - T_v}} \quad \text{(for tie)}$$

where $T_u = \sum_{j=1}^{t} \frac{k_j(k_j^2 - 1)}{12n}$; $T_v = \ldots$

$\hookrightarrow$ for t ties of length $k_1, \ldots, k_t$

•) Kendall's τ

$$= \frac{\sum \sum_{i<j} a_{ij} b_{ij}}{\binom{n}{2}}$$

$$\left[a_{ij} = \begin{cases} 1, & \text{if } u_i < u_j \\ -1, & \text{o.w.} \end{cases} \quad b_{ij} = \begin{cases} 1, & v_i < v_j \\ -1, & \text{o.w.} \end{cases} \right.$$

for two set of ranks $\{u_1, u_2, u_3, \ldots, u_n\}$ & $\{v_1, v_2, \ldots, v_n\}$

For tie

$$\tau = \frac{\sum \sum_{i<j} a_{ij} b_{ij}}{\sqrt{\binom{n}{2} - T_u'} \cdot \sqrt{\binom{n}{2} - T_v'}}$$

$$\left[a_{ij} = \begin{cases} 1, & u_i < u_j \\ 0, & u_i = u_j \\ -1, & \text{o.w.} \end{cases} \right.$$

$b_{ij} = \ldots$

$T_u' = \sum_{j=1}^{t} \binom{k_j}{2} \rightarrow$ for t ties of length $k_1, \ldots, k_t$

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

-) Intraclass Correlation, | Data :- for both x & y, x_{ij} with freq $k-1$, $i=1(1)p$; $j=1(1)k$

$$r_I = \frac{1}{k-1}\left[k\frac{V_B}{V_T}-1\right],$$

$$V_T = \frac{1}{pk}\sum_i\sum_j (x_{ij}-x_{00})^2 ; \quad V_B = \frac{1}{pk}\sum_i k(\bar{x}_{i0}-\bar{x}_{00})^2$$

$$-\frac{1}{k-1} \le r_I \le 1.$$

Multivariate : See G-K

Regression (sth deg LS pol fitting)

$$y_i = {}_sY_i + {}_se_i \quad \left[{}_sY_i = \hat{a}_0 + \hat{a}_1 x_i + \dots + \hat{a}_s x_i^s\right]$$

Normal eq^{n}s :- $\sum_{i=1}^{n} x_i^j \, {}_se_i = 0$, $j = 0(1)s$.

sth deg Corr Index, $r_s^2 = \frac{V({}_sY)}{V(y)}$

$Cov(y, {}_sY) = V({}_sY)$;

$V({}_se) = s_y^2(1-r_s^2)$; $r(y, {}_sY) = r_s$

$Cov({}_sY, {}_se) = 0$, $Cov(x, {}_se) = 0$.

$\overline{{}_sY} = \bar{y}$; $0 \le r_1^2 \le r_2^2 \le \dots \le r_s^2$.

Array Data > Corr Ratio, $e_{yx}^2 = \frac{V(\bar{y}_{i0})}{V(y)}$ } See Khata

Scanned by CamScanner
ABOUT.ME/SOURAV.NANDI

Measures Of Association

(A) For Ordinal Data :-

(1) Goodman Kruskal's $G = \frac{C-D}{C+D} = \frac{\pi_c - \pi_D}{1-\pi_t}$

$$= \frac{\sum\sum a_{ij}' b_{ij}'}{\left[n^2 - \sum f_{p0}^2 - \sum f_{0q}^2 + \sum\sum f_{pq}^2\right]}$$

(2) Kendall's $\tau_a = \frac{C-D}{\binom{n}{2}} = \frac{\sum\sum_{i<j} a_{ij}' b_{ij}'}{\binom{n}{2}}$
(no tie)

(3) Kendall $\tau_b = \frac{C-D}{\sqrt{\binom{n}{2}-T_A}\sqrt{\binom{n}{2}-T_B}}$

$T_A = \sum_{p=1}^{k}\binom{f_{p0}}{2}$; $T_B = \sum_{q=1}^{l}\binom{f_{0q}}{2}$

(4) Somer's $d = \frac{C-D}{\binom{n}{2}-T_A}$ → Tie case

C → Concordant pairs   D → discordant

(5) $\alpha = \frac{C}{D}$, $\delta = \frac{C-D}{n}$ [In 2×2, $Q = \frac{\alpha-1}{\alpha+1}$]

(6) Odds Ratio, $\theta_{AB} = \frac{\Omega_A}{\Omega_\alpha} = \frac{\pi_{AB}\,\pi_{\alpha\beta}}{\pi_{A\beta}\,\pi_{\alpha B}}$.

Ω_α → Odds for success of B when A is at α level

Scanned by CamScanner
ABOUT.ME/SOURAV.NANDI

(B) For Nominal data :- [Also for Ordinal]

- Mean Square Contingency, $\phi^2 = \frac{\chi^2}{n} = \sum\sum \frac{f_{ij}^2}{f_{io} f_{oj}} - 1$,
- Sqr. Cont. $\rightarrow \chi^2 = \sum\sum \frac{\delta_{ij}^2}{f_{io} f_{oj}/n}$, $\delta_{ij} = f_{ij} - \frac{f_{io} f_{oj}}{n}$
- Pearson's Coeff of Cont., $C = \sqrt{\frac{\phi^2}{1+\phi^2}} = \sqrt{\frac{\chi^2}{n+\chi^2}}$

  $[0 \leq C < 1]$
- Tchuprow's measure, $T = \left[\frac{\phi^2}{\sqrt{(k-1)(l-1)}}\right]^{1/2}$
- Cramer's $V^2 = \frac{\phi^2}{\min(k-1, l-1)}$

for 2×2 case, $\chi^2 = n V^2$

$\therefore \phi^2 = V^2$ (V → pearson's coeff of assn)

Categorical Data Analysis

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI