Some Moment Inequalities :-

Markov :-

- $P[h(X) \geq a] \leq \frac{E[h(X)]}{a}$

When $h(X) \geq 0$, $E[h(X)]$ exists and $a>0$.

- Chebychev :- $P[X \geq \mu + t\sigma] \leq \frac{1}{1+t^2}$, $(t>0)$

$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$ $(\forall k>0)$.

- Markov :- $P(|X| \geq a) \leq \frac{E(|X|)}{a}$ $(a>0)$

- X be a RV with $E(X)=0$, $V(X)=\sigma^2$.

Then, $P(X \geq x) \leq \frac{\sigma^2}{x^2+\sigma^2}$, if $x>0$

and $P(X \geq x) \geq \frac{x^2}{\sigma^2+x^2}$, if $x<0$.

- Jensen :-

equality :- If $P[g(X) = a+bx] = 1$

$E(g(X)) \geq g(E(X))$ if g = convex.

(finite) $(g'' \geq 0)$

$g, h \to$ fn of x. (real valued)

• Cauchy Schwarz :- $E(g^2)E(h^2) \geq E^2(gh)$

equality :- $P[g(x) = c\,h(x)] = 1$.

• Let $E|X|^4 < \infty$, $E(X) = 0$, $V(X) = \sigma^2$

Then, for $k > 1$, $[\mu_4 = E(X^4)]$

$$P(|X| \geq k\sigma) \leq \frac{\mu_4 - \sigma^4}{\mu_4 + \sigma^4 k^4 - 2k^2\sigma^4}$$

• Liapounov :- [Provided exists].

$$\gamma_a^{b-c} \cdot \gamma_c^{a-b} \geq \gamma_b^{a-c}$$

$\forall\ a > b > c \geq 0$, $\gamma_t = E|X|^t$.

Also, $\gamma_k^{1/k}$ is a non $\downarrow$ fn of k.

i.e., $\gamma_{k-1}^{\frac{1}{k-1}} \leq \gamma_k^{\frac{1}{k}}$.

• $|E(X)| \leq E(|X|)$

Bool :- (a) $P\left(\bigcup A_i\right) \leq \sum P(A_i)$

(b) $P\left(\bigcap_1^n A_i\right) \geq \sum_1^n P(A_i) - (n-1)$.

• The prob that exactly m of the n events A_i $(i=1,2,\dots,n)$ will occur is:

$$P_{[m]} = \sum_{k=0}^{n-m} (-1)^k \binom{m+k}{m} S_{m+k}.$$

$\hookrightarrow S_2 = \sum_{1\le i_1<i_2\le n} P[A_{i_1} \cap A_{i_2}], \dots$

• The prob that atleast m will occur (out of n events A_i) $(i=1(1)n)$:

$$P_m = \sum_{k=0}^{n-m} (-1)^k \binom{m+k-1}{m-1} S_{m+k}.$$

• Poincare Formula (atleast one event),

$$P_1 = \sum_{k=0}^{n-1} (-1)^k S_{k+1} = \sum_{k\ge 1}^{n} (-1)^{k-1} S_k = S_1 - S_2 + \dots$$

Here, $S_j = \sum_{1\le i_1<i_2<\dots<i_j\le n} P[A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_j}]$

(*) For a RV with $E(X)=\mu$; $V(X)=\sigma^2$ $(<\infty)$

we have, $F(x) \le \dfrac{\sigma^2}{\sigma^2+(x-\mu)^2}$, if $x\le\mu$

$\ge \dfrac{(x-\mu)^2}{\sigma^2+(x-\mu)^2}$, if $(x\ge\mu)$

Results from Outline ⇒

- For $a, b > 0$, $(a+b)^r \le C_r(a^r + b^r)$

  where $C_r = 1$ if $r \le 1$

  $= 2^{r-1}$ if $r \geq 1$.

↳ C_r Ineql.: $E|X+Y|^r \le C_r\{E|X|^r + E|Y|^r\}$

- Holder's Ineql.: for $r > 1$, $\frac{1}{r} + \frac{1}{s} = 1$,

  $E|XY| \le E^{1/r}|X|^r \, E^{1/s}|X|^s$

- Minkowski's ineql.: for $r \geq 1$,

  $E^{1/r}|X+Y|^r \le E^{1/r}|X|^r + E^{1/r}|Y|^r$.

MGK 6.25

If f & g are monotone in same direc^n, $E(f(x)g(x)) \geq E(f(x)) E(g(x))$

If opposite direc^n, → $E(fg) \le E(f)\, E(g)$

(i) Stein's Lemma: If $X \sim N(\mu, \sigma^2)$,

Then $E[g(x)(x-\theta)] = \sigma^2 E[g'(x)]$

derivative ↗

→ $f^{(n-1)}$ cont on $[a, a+h]$; $f^{(n)}$ exist on $(a, a+h)$

Taylor's Thm:- $\exists\, \theta$, $0<\theta<1$, [See PB Khata]

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{\lfloor 2} f''(a) + \dots + \underbrace{\frac{h^n}{\lfloor n} f^{n}(a+\theta h)}_{R_n}$$

[$R_n \to 0$ as $n \to \infty$]

VVI distn's

Degrees of Freedom

(A) Chi Squared (p)

[Gamma, $n = \frac{p}{2}$; $\theta = 2$]

$$f(x) = \frac{e^{-\frac{x}{2}}\, x^{\frac{p}{2}-1}}{\Gamma\frac{p}{2} \times 2^{p/2}}, \quad 0 \le x < \infty, \quad (p = 1, 2, \dots)$$

Then, $E(X) = p$, Mode $= p - 2$ $(p>2)$

$$Var(X) = 2p$$

Also, mgf,

$$M_X(t) = (1-2t)^{-\frac{p}{2}}, \quad t < \frac{1}{2}$$

$X_i \sim N(0,1)$ indep

$\Rightarrow X_1^2 + \dots + X_\nu^2 \sim \chi^2(\nu)$

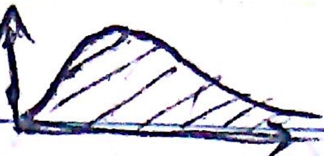
If $X \sim \chi^2(p)$,

$$\frac{X-\mu}{\sigma} = \frac{X-p}{\sqrt{2p}} \xrightarrow{n\to\infty} N(0,1)$$

$U_i \sim \chi^2(\nu_i)$ indep

$\Rightarrow \Sigma U_i \sim \chi^2(\Sigma \nu_i)$

(B) F distⁿ :-



$$f(x) = \frac{\Gamma\left(\frac{v_1+v_2}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)} \times \left(\frac{v_1}{v_2}\right)^{v_1/2} \times \frac{x^{\frac{v_1-2}{2}}}{\left(1+\frac{v_1}{v_2}x\right)^{\frac{v_1+v_2}{2}}}$$

$0 \le x < \infty$; $v_1, v_2 \ge 1$.

$$\hookrightarrow = \frac{1}{B\left(\frac{v_1}{2}, \frac{v_2}{2}\right)} \times \frac{\left(\frac{v_1}{v_2}x\right)^{\frac{v_1}{2}}}{x\left(1+\frac{v_1}{v_2}x\right)^{\frac{v_1+v_2}{2}}}$$

$$E(X) = \frac{v_2}{v_2-2} \quad [v_2 > 2]$$

$$Var(X) = 2\left(\frac{v_2}{v_2-2}\right)^2 \times \frac{(v_1+v_2-2)}{v_1(v_2-4)}, \quad \hookrightarrow [v_2 > 4]$$

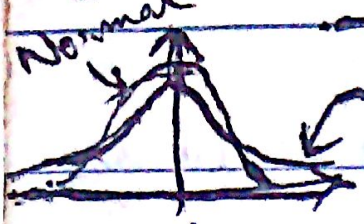
mgf → doesn't exist.

$$F_{1,v} = t_v^2 \; ; \; F_{v_1,v_2} = \left(\frac{\chi^2_{v_1}}{v_1}\right) \Big/ \left(\frac{\chi^2_{v_2}}{v_2}\right)$$

$$E(X^n) = \frac{\Gamma\left(\frac{v_1+2n}{2}\right) \times \Gamma\left(\frac{v_2-2n}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right) \times \Gamma\left(\frac{v_2}{2}\right)} \times \left(\frac{v_2}{v_1}\right)^n, \quad \left(n < \frac{v_2}{2}\right)$$

$X \sim F_{p,q} \Rightarrow \frac{1}{X} \sim F_{q,p}$ | $Y \sim t_v \Rightarrow Y^2 \sim F_{1,v}$

* $X \sim F(m,n) \Rightarrow W = \frac{mX}{n+mX} \sim \beta\left(\frac{m}{2}, \frac{n}{2}\right)$



→ ht at 0 ↑, as DF ↑

$\boxed{\text{t dist}^n}$ → Symm (0)

→ $-\infty < x < \infty$

$\nu = 1, \dots$

$$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \times \frac{1}{\sqrt{\nu\pi}} \times \frac{1}{\left(1+\frac{x^2}{\nu}\right)^{\frac{\nu+1}{2}}}$$

$\nu = 1$ → Cauchy

$E(X) = 0 \quad (\nu > 1)$,

$V(X) = \frac{\nu}{\nu-2} \quad (\nu > 2)$.

mgf ↓ doesn't exist. (as all moments are not there)

For $n < \nu$,

$$E(x^n) = \begin{cases} 0, & \text{if } n \text{ is odd} \\ \dfrac{\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{\nu-n}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \nu^{\frac{n}{2}}, & \text{if } n \text{ is odd} \end{cases}$$

$$T = \frac{Y}{\sqrt{Z/\nu}}$$

$Y \to N(0,1)$, $Z \to \chi^2(\nu)$

$T \downarrow t(\nu)$

→ $\left(T = \frac{Z}{\sqrt{\chi^2/\nu}}\right)$

↳ $$f(x) = \frac{1}{B\left(\frac{\nu}{2}, \frac{1}{2}\right)} \times \frac{1}{\sqrt{\nu} \times \left(1+\frac{x^2}{\nu}\right)^{\frac{\nu+1}{2}}}$$

Relat^ns of Percentile values :-

- $F_{1-p} = \dfrac{1}{F_p}$ (→ p th percentile)

- $F_{1-p;1,v} = t^2_{1-\frac{p}{2};v}$

- $F_{p;v,\infty} = \dfrac{\chi^2_{p;v}}{v}$

Results on χ^2 distn :-

- If $X \sim \chi^2_{n_1}$ & $Y \sim \chi^2_{n_2}$ (indep)

then → $\dfrac{X}{Y} \sim \beta_2\left(\dfrac{n_1}{2}, \dfrac{n_2}{2}\right)$.

$\dfrac{X}{X+Y} \sim \beta_1\left(\dfrac{n_1}{2}, \dfrac{n_2}{2}\right)$.

- In a random large sample,

$\sum_{i=1}^{k}\left[\dfrac{(n_i - np_i)^2}{np_i}\right] \sim \chi^2(k-1)$.

Here, n_i = observed freq for ith class

np_i = expected freq $\left[\sum_{1}^{k} n_i = n\right]$.

Sampling Th. :- Relationship $\Rightarrow$

Reg $\Rightarrow$ 1. To test sample reg. coeff β is equal to b, we use that the statistic

$$t = \frac{\beta - b}{s_e / s_x}\sqrt{n-2}$$ has t distn with DF $(n-2)$.

2. If $y = y_p$ for popn & y_0 from sample,

Then $$t = \frac{(y_0 - y_p)\sqrt{n-2}}{s_e\sqrt{n+1+\left[\frac{n(x_0-\bar{x})^2}{s_x^2}\right]}}$$

is $t(n-2)$. [$n \to$ sample size]

3. for predicted mean value,

$$t = \frac{(y_0 - \bar{y}_p)\sqrt{n-2}}{s_e\sqrt{1+\left[(x_0-\bar{x})^2/s_x^2\right]}} \sim t(n-2)$$

i.e., $\bar{y}_p \to E[Y|X = x_0]$

Corr ⇒ ① To test $\rho = 0$, we use

that $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim t$ distⁿ $(\underset{DF}{n-2})$.

VVI

② $\rho \neq 0 \rightarrow Z = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right) \sim N(\mu_Z, \sigma_Z^2)$

Fisher's Z transformation

where, $\mu_Z = \frac{1}{2}\ln\left(\frac{1+\rho}{1-\rho}\right)$; $\sigma_Z = \frac{1}{\sqrt{n-3}}$

③ To test whether r_1 & r_2, drawn from samples of size n_1 & n_2, differ significantly, we use that,

$$Z = \frac{Z_1 - Z_2 - \mu_{Z_1 - Z_2}}{\sigma_{Z_1 - Z_2}} \sim N(0,1)$$

[see ② for expression]

Standard error for large samples, [∵ Normal popⁿ]

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$; $\sigma_p = \sqrt{\frac{pq}{n}}$ (proportion) ; $\sigma_{med} = \sigma\sqrt{\frac{\pi}{2n}}$ (Median)

σ_s (SD) $= \begin{cases} \frac{\sigma}{\sqrt{2n}} & \text{of normal pop}^n \\ \sqrt{\frac{\mu_4 - \sigma^4}{4n\sigma^2}}, & \text{o.w.} \end{cases}$ | σ_{S^2} (var) $= \begin{cases} \sigma^2\sqrt{\frac{2}{n}}, & \text{of normal pop}^n \\ \sqrt{\frac{\mu_4 - \sigma^4}{n}}, & \text{o.w.} \end{cases}$

Schaum's Outline (Sampling) ⇒

for ∞ pop^n / WR sampling,

$$V(\bar{x}) = \frac{\sigma^2}{n} \leftarrow \text{sample size} \quad // \quad \text{pop}^n \text{ size}$$

for WOR sampling, $V(\bar{x}) = \frac{\sigma^2}{n} \times \frac{N-n}{N-1}$

⓪ Sampling distn of proportions :- $(\mu_p = p)$

For ∞ bin (p) pop^n, $\sigma_p = \sqrt{\frac{pq}{n}}$.

sample size

for finite pop^n,

$$\sigma_p^2 = \frac{pq}{n} \times \left(\frac{N-n}{N-1}\right)$$

⓪ Sampling distn of differences of -

① mean → $Z = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$ ~ Normal [any pop^n]

② proportions → $\dfrac{(P_1 - P_2) - (p_1 - p_2)}{\sqrt{\dfrac{p_1 q_1}{n_1} + \dfrac{p_2 q_2}{n_2}}}$

Normal ⇒

So, (1) $\hat{S}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$

(2) $E(\hat{S}^2) = \begin{cases} \sigma^2, & \text{for WR} \\ \frac{N\sigma^2}{N-1}, & \text{for WOR} \end{cases}$

(3)* $\frac{(n-1)\hat{S}^2}{\sigma^2} \sim \chi^2(n-1)$.

(4) $\frac{\bar{x}-\mu}{\hat{S}/\sqrt{n}} \sim t(n-1)$ ← ✓

(5) $\frac{\hat{S}_1^2/\sigma_1^2}{\hat{S}_2^2/\sigma_2^2} \sim F(m-1, n-1)$.

$\Uparrow$

* (MGB-P250) → $\boxed{\text{If } X_1, \ldots, X_n \text{ r.s. from } N(\mu, \sigma^2)}$ ✓

$$\frac{\sqrt{n(n-1)}\,(\bar{X}-\mu)}{\sqrt{\Sigma(X_i-\bar{X})^2}} \sim t(n-1).$$

*(3) Note that, $\frac{\sum_{1}^{n}(x_i-\mu)^2}{\sigma^2} \sim \chi^2(n)$ [Mean Known]

$\hat{S}^2 = \frac{1}{2n(n-1)}\sum_{i,j=1}^{n}\sum(x_i-x_j)^2$;

Scanned by CamScanner

Estimation (MGB)

(i) $T_n = T(x_1, \ldots, x_n)$ is unbiased (estimate) for $\gamma(\theta)$, if $E(T_n) = \gamma(\theta)$, $\forall\ \theta \in \Theta$.

(ii) Efficiency :- If T_1 most efficient (var min), then $E(T_2) = \frac{V_1}{V_2}$ $[E \leq 1]$

(iii) (A) Consistency :- [See ix + ②]

If $T_n \xrightarrow{p} \gamma(\theta)$, as $n \to \infty$, then T_n is consistent for $\gamma(\theta)$.

$\Rightarrow$ If $E_\theta(T_n) \to \gamma(\theta)$ & $V_\theta(T_n) \to 0$, as $n \to \infty$, then also T_n is consistent for $\gamma(\theta)$.

(B) Mean Square Consistency :-

(iv) $\lim_{n \to \infty} E_\theta[T_n - \gamma(\theta)]^2 = 0\ \forall\ \theta \in \Theta$

Note :- $B \Rightarrow A$, but not conversely.

(iii) $\mapsto S_1, \ldots, S_n$ jointly suff iff $L = g_\theta(S_1, \ldots, S_n)\, h(x)$

⊙ Sufficiency :- $T = t(x)$ is sufficient for θ, iff $L = g_\theta[t(x)] \times h(x)$

$L \downarrow$ Joint Density f^n.

(v) ⊙ Cramer-Rao Inequality :-

t be unbiased for $\gamma(\theta)$,

$$Var(t) \geq \frac{\{\gamma'(\theta)\}^2}{E\left(\frac{\partial}{\partial\theta}\log L\right)^2} \longrightarrow I(\theta)$$

$= nE\left(\frac{\partial}{\partial\theta}\ln f\right)^2$; pdf

$/ -E\left(\frac{\partial^2}{\partial\theta^2}\ln L\right)$

(information on θ supplied by sample)

Equality :- $L \rightarrow$ Likelihood f^n.

Iff, $\frac{\partial}{\partial\theta}\log L = \frac{t - \gamma(\theta)}{\lambda(\theta)}$

Then,
- t is unbiased for $\gamma(\theta)$
- t is MVB estimator for $\gamma(\theta)$
- $V(t) = |\gamma'(\theta) \times \lambda(\theta)|$.

[Certain Regularity Condi^{n}s are Assumed]

(vi) **Completeness** :- $T = t(x)$ is complete for θ, if

$$E_\theta[h(T)] = 0 \ \forall \theta \Rightarrow P_\theta[h(T)=0] = 1$$

$\hookrightarrow \int h(t) f(t,\theta) dt$; $\rightarrow$ pdf of T

(1) **Maximum Likelihood Estimators**

For random sample $X_1, \ldots, X_n$,

likelihood f^n, $L(\theta) = \prod_i f(x_i;\theta)$

$L(\theta_1, \ldots, \theta_k) = \prod_i f(x_i; \theta_1, \ldots, \theta_k)$

MLE is the sol^n of $\frac{\partial L}{\partial \theta_i} = 0 \ \forall i \rightarrow$ (1)

(maybe $L \rightarrow \log L$) [simultaneous system of eq^n]

$\hookrightarrow$ check if not a minima

Prop :-

(i) Cramer-Rao : MLEs are consistent

(ii) Any consistent sol^n of (1) provides MLE with prob $\rightarrow 1$ as $n \rightarrow \infty$ (sample size) [Hazoor Bazar]

(iii) MLE is most efficient

(iv) Sufficient estimator is a f^n of MLE.

(v) T MLE of $\theta \Rightarrow \psi(T)$ MLE of $\psi(\theta)$ [~~ψ~~, one-one ψ] [Invariance Prop]

(2) (i) Minimum Variance U.E. :)

Minimise $V(t)$ subj. to $E(t) = \gamma(\theta)$

i.e., Min $\int_{-\infty}^{\infty} [t - \gamma(\theta)]^2 L \, dx$, $\quad \hookrightarrow \int t \cdot L \, dx = \gamma(\theta)$

subj to $\int_{-\infty}^{\infty} \{t - \gamma(\theta)\} L \, dx = 0$.

(3) (i) Method of Moments,

$\mu_r' = m_r'$

(4) (i) Method of Least Squares

[~~if fitted curve linear → MVUE~~]

(5) (i) Method of Minimum χ^2 :- [MGB, P-286]

$$\text{Min } \chi^2 = \sum_{j=1}^{k} \frac{[n_j - n p_j(\theta)]^2}{n p_j(\theta)},$$

where $S_1, \ldots, S_k \rightarrow$ partition of range of X

$X_1, \ldots, X_n \rightarrow$ r.s. from density $f_X(x, \theta)$.

$p_j(\theta)$ = prob that a obs falls in S_j.

N_j = # X_i (in sample) in S_j

(n_j)

(vi) Minimum Distance :-

e.g., $d(F, G) = \sup_x |F(x) - G(x)|$

find $\hat{\theta}$ s.t. $d(F(x, \theta), F_n(x))$ is minimised

Pop^n CDF ↳ sample CDF $(x_1, \ldots, x_n)$

(vii) Closeness :- T' more concentrated for $\tau(\theta)$ than T, iff,

$$P_\theta[\tau(\theta) - \lambda < T' \leq \tau(\theta) + \lambda] \geq P_\theta[\tau(\theta) - \lambda < T \leq \tau(\theta) + \lambda]$$

$\forall \lambda > 0$ & $\forall \theta \in \Theta$

Most Conc :- More conc than any other estimator

→ Pitman Closer :- T' pitman closer estimator of $\tau(\theta)$ than T iff,

$$P_\theta[|T' - \tau(\theta)| < |T - \tau(\theta)|] \geq \frac{1}{2} \quad \forall \theta \in \Theta$$

(2) → Pitman Closest :- P. Closer than others.

(viii) Mean Squared Error :-

$$MSE_t(\theta) = E_\theta[T - \tau(\theta)]^2$$

$$= \int \cdots \int [t(x_1, \dots, x_n) - \tau(\theta)]^2 \prod_{i=1}^{n} f(x_i, \theta)\, dx_i$$

(ix) Best Asymptotically Normal / Consistent Asymp. Normal Efficient estimator (BAN/CANE) ⇒

A seq of estimators $T_1^*, \dots, T_n^*, \dots$ of $\tau(\theta)$ is BAN iff —

① $\sqrt{n}\,[T_n^* - \tau(\theta)] \to N(0, \sigma^{*2}(\theta))$, as $n \to \infty$.

② $\forall\, \epsilon > 0$, $\forall\, \theta \in \Theta$,

$$\lim_{n\to\infty} P_\theta[|T_n^* - \tau(\theta)| > \epsilon] = 0$$

③ T_n be another seq of simple consistent estimators with ① $\to N(0, \sigma^2(\theta))$. [②]

Then, $\sigma^{*2}(\theta) \le \sigma^2(\theta)$

(x) t be estimate of $\tau(\theta)$. Then a real valued f^n satisfying → (2)

(1) $l(t,\theta) \geq 0$ ($\forall t, \forall \theta$)

& (2) $l(t,\theta) = 0$ for $t = \tau(\theta)$,

is called <u>Loss f^n</u> + $l(T,\theta)$.

⇒ <u>Risk f^n</u>, $R_t(\theta) = \underset{\text{(over } x)}{E_\theta}[l(T,\theta)]$

⇒ t_1 is a <u>better</u> estimator than t_2, iff $R_{t_1}(\theta) \leq R_{t_2}(\theta)\ \forall \theta$ (+) strict for $\geq$ one θ.

⇒ <u>Admissible</u>, iff no better est is there

(xi) <u>Minimax</u>:- An estimator t^* is minimax iff $\sup_\theta R_{t^*}(\theta) \leq \sup_\theta R_t(\theta)$, for every t.

(xii) <u>Minimal Sufficient</u>:- iff a set of jt. suff stat is a f^n of every other set of sufficient stat.

(*) K-parameter exponential family :-

$$f(x_i;\theta_1,\dots,\theta_k) = a(\theta_1,\dots,\theta_k)\, b(x)\, e^{\sum c_j(\theta_1 \dots \theta_k) d_j(x)}$$

Note :- for 1 parameter $a(\theta) b(x) e^{c(\theta) d(x)}$,

$\sum d(X_i)$ is a complete Minimal Sufficient Statistics.

(xviii) Rao-Blackwell :- $S_1, \dots, S_k$ be a set of jointly sufficient statistics, & T be unbiased for $\tau(\theta)$. ~~$T' = E[T|S_1 \dots S_k]$~~

Define $T' = E[T | S_1, \dots, S_k]$

Then → ① $E[T'] = \tau(\theta)$ ∴ T' also unbiased

② $V(T') \leq V(T)$ $\forall \theta$,

and $V(T') < V(T)$ for some θ [unless $T = T'$, with $p = 1$]

(*) UMVUE :- unbiased & variance not larger than any other unbiased estimators.

Scanned by CamScanner

(xiv) Lehmann-Scheffé :- $X_1, \ldots, X_n$ r.s. from $f(\cdot, \theta)$. If $S = s(X_1, \ldots, X_n)$ be a complete sufficient stat & if $T^* = t^*(S)$ is an u.e. of $\tau(\theta)$, then T^* is an UMVUE of $\tau(\theta)$

[u.e. based on complete suff stat is UMVUE]

(xv) Locaⁿ invariance :- $\forall x_1, \ldots, x_n$ & c,

$$t(x_1+c, \ldots, x_n+c) = t(x_1, \ldots, x_n) + c$$

⇒ Pitman Estimator for locaⁿ,

$$t(x_1, \ldots, x_n) = \frac{\int_{-\infty}^{\infty} \theta \prod_1^n f(x_i, \theta)\, d\theta}{\int_{-\infty}^{\infty} \prod_1^n f(x_i, \theta)\, d\theta}$$

It has uniformly smallest MSE [in (xv) estimator].

[fⁿ of suff. stat)

(xvi) Scale invariance :- $t(cx_1, \ldots, cx_n) = c\, t(x_1, \ldots, x_n)$ $\forall x_1, \ldots, x_n$ & $c > 0$.

Pitman Estimator for scale has uniformly smallest risk for loss fⁿ $\ell(t, \theta) = \frac{(t-\theta)^2}{\theta^2}$.

Given by, $t = \dfrac{\int_0^\infty \frac{1}{\theta^2} \prod_{i=1}^{n} f(x_i, \theta)\, d\theta}{\int_0^\infty \frac{1}{\theta^3} \prod_{i=1}^{n} f(x_i, \theta)\, d\theta}$

(Pitman Estimator for Scale)

⊛ Ellipsoid of Concentration :-

$(T_1, \ldots, T_r)$ U.E. of $(\tau_1(\theta), \ldots, \tau_r(\theta))$. [vector]

$\sigma^{ij}(\theta) \rightarrow$ ijth element of inverse of the covariance matrix of $(T_1, \ldots, T_r)$

Then, ellipsoid of conc. is the interior & boundary of $\sum_{i=1}^{r} \sum_{j=1}^{r} \sigma^{ij}(\theta) [t_i - \tau_i(\theta)] [t_j - \tau_j(\theta)]$

$= r + 2$ [ellipsoid]

⊛ Wilks generalised variance :-

Determinant of covariance matrix of $(T_1, \ldots, T_r)$ [ijth element is, $\sigma_{ij}(\theta) = Cov_\theta[T_i, T_j]$]

Bayes Estimator:-

Prior distⁿ be $\pi(\theta)$, sampling dist $f(x|\theta)$

then, the Posterior distⁿ, the conditional distⁿ of θ given the sample, x, is,

$$\pi(\theta|x) = \frac{f(x|\theta)\,\pi(\theta)}{m(x)} \quad \{\, \rightarrow = f(x,\theta)\}$$

where, $m(x) = \int f(x|\theta)\,\pi(\theta)\,d\theta$.

(marginal of x).

Optimum properties of MLE :-

$\hat{\theta}_n$ be mle of θ for rs of size n.

from $f(x;\theta)$ [satisfying certain regularity conditions]

Then → (i) $\{\hat{\theta}_n\}$ is BAN

(ii) $\hat{\theta}_n$ is asymptotically normal

with mean θ & variance $= \dfrac{1}{n\,E\left[\frac{\partial}{\partial\theta}\log f(\cdot)\right]^2}$

Casella:-

•) **Minimal Sufficient statistics:**

$T(X)$ min. suff. for θ, iff for every two sample pts x & y, (vectors)
the ratio $\dfrac{f(x|\theta)}{f(y|\theta)}$ is const as a f^n of θ iff $T(x) = T(y)$.
[i.e., indep of θ].

*) **Ancillary Statistics**:- A statistic whose distn does not depend on the parameter θ.

•) **Basu's Theorem**:- If $T(X)$ is complete (minimal) sufficient, then it is indep of every ancillary statistic.

•) If minimal suff stat. exists, Then any complete stat. is also a minimal sufficient statistic.

Order Statistics [$Y_\alpha \to \alpha$ th order stat]

For order stat from CDF $F(\cdot)$, $\alpha = 1(1)n$.

•) $F_{Y_\alpha}(y) = \sum_{j \ge \alpha}^{n} \binom{n}{j} [F(y)]^j [1-F(y)]^{n-j}$

$\hookrightarrow = \sum_{j=\alpha}^{n} P$[exactly j of n order stats $\le y$]

•) X_i be iid with density $f(\cdot)$. $\xi_p \to p$th quantile (unique)

p_n is s.t. np_n is integer & $n|p_n - p|$ is bounded.

$Y^{(n)}_{np_n} \to (np_n)^{th}$ order stat from rs of size n.

Then, this is asymptotically $N\left(\xi_p, \frac{p(1-p)}{n[f(\xi_p)]^2}\right)$

•) for changing sample size n,

k^{th} largest order stat $\to Y^{(n)}_{n-k+1}$. ($Y_1^{(n)}, Y_n^{(n)} \to$ extreme value stat)

•) Let $F_n(x)$ denote the sample cdf of a random sample of size n from $F(\cdot)$,

then $\to$ $P\left[F_n(x) = \frac{k}{n}\right] = \binom{n}{k} [F(x)]^k [1-F(x)]^{n-k}$, $k = 0(1)n$

Scanned by CamScanner

(Contd) :-

Sample / Empirical DF :- (no of X_k's $\le x$, $1 \le k \le n$)

$$F_n(x) = \frac{1}{n}\sum_{j=1}^{n} \varepsilon(x - X_j) \leftarrow \text{(Def}^n\text{)}$$

•) Glivenko-Cantelli Thm :-

$F_n(x)$ converges uniformly to $F(x)$.

i.e, for $\epsilon > 0$, $\lim_{n\to\infty} P\left\{\sup_x |F_n(x) - F(x)| > \epsilon\right\} = 0$

Limit Thms :-

•) WLLN :- X_i iid with $V(X_i) < \infty$ ($= \sigma^2$), ($E X_i = \mu$), $i = 1(1)n$. $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Then $\forall \epsilon > 0$, $\overline{X}_n \xrightarrow{P} \mu$, i.e, $\lim_{n\to\infty} P(|\overline{X}_n - \mu| < \epsilon) = 1$

•) SLLN :- Under same^{d} condi^{n}s, $\overline{X}_n \xrightarrow{a.s} \mu$

i.e, $\forall \epsilon > 0$, $P\left(\lim_{n\to\infty} |\overline{X}_n - \mu| < \epsilon\right) = 1$

•) Conv in dist^{n}/law/weakly :- (d / L / W)

$\{X_n\}$ converges in dist^{n} to X, if at every continuity pt of $F(x)$, $\lim_{n\to\infty} F_n(x) = F(x)$.

•) CLT :- Under the same condi^{n} (of WLLN),

$$\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow[(d)]{L} N(0,1).$$

•) Slutsky's Thm :- If $X_n \xrightarrow{L} X$ & $Y_n \xrightarrow{P} a$ (const)

Then $X_n Y_n \xrightarrow{L} aX$ & $X_n + Y_n \xrightarrow{L} X + a$

Delta Mtd :- (assumptions :- derivatives exist)

•) Taylor :- $g(x) \simeq \sum_{i=0}^{n} \frac{(x-a)^i}{\underline{|i}} g^{(i)}(a) = T_n(x)$

If $g^n(a)$ exists, Then → $\lim_{x \to a} \frac{g(x) - T_n(x)}{(x-a)^n} = 0$.

•) So, if $\underset{\sim}{T} = (T_1, \ldots, T_k)$ & $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_k)$,

then $g(\underset{\sim}{t}) \approx g(\underset{\sim}{\theta}) + \sum_{i=1}^{k} g_i'(\theta)(t_i - \theta_i)$.

$\hookrightarrow = \frac{\partial}{\partial t_i} g(t) \Big|_{t = \theta}$, $i = 1(1)k$

So, $E g(\underset{\sim}{T}) \approx g(\underset{\sim}{\theta})$.

$\operatorname{Var} g(\underset{\sim}{T}) \simeq E\left(\left[\sum_i^k g_i'(\theta)(T_i - \theta_i)\right]^2\right)$

$= \sum (g_i'(\theta))^2 \operatorname{Var} T_i + 2 \sum_{i > j} g_i'(\theta) g_j'(\theta) \operatorname{Cov}(T_i, T_j)$

•) ~~Delta mtd~~ :- So in 1st order,

$g(X) = g(\mu) + g'(\mu)(X - \mu) \Rightarrow E g(X) \approx g(\mu)$

& $\operatorname{Var} g(X) \approx [g'(\mu)]^2 \operatorname{Var} X$ ←

Scanned by CamScanner

ABOUT.ME/SOURAV.NANDI

◎ Delta Mtd :- If $\sqrt{n}(Y_n - \theta) \xrightarrow{d} N(0, \sigma^2)$,

Then $\sqrt{n}[g(Y_n) - g(\theta)] \xrightarrow{d} N(0, \sigma^2 [g'(\theta)]^2)$

provided $g'(\theta)$ exists and $\neq 0$.

◦) Second Order :- If $\sqrt{n}(Y_n - \theta) \xrightarrow{d} N(0, \sigma^2)$,

$g'(\theta) = 0$ & $g''(\theta)$ exists & non-zero for some θ.

Then, $n[g(Y_n) - g(\theta)] \xrightarrow{d} \sigma^2 \dfrac{g''(\theta)}{2} \chi_1^2$

◦) Multivariate :- $\underset{\sim}{X_1}, \ldots, \underset{\sim}{X_n}$ r.s. with $EX_{ij} = \mu_i$

& $Cov(X_{ik}, X_{jk}) = \sigma_{ij}$. For a given g

with cont first partial derivatives & $\underset{\sim}{\mu} = (\mu_1, \ldots, \mu_p)$,

for which $\tau^2 = \sum\sum \sigma_{ij} \dfrac{\partial g(\underset{\sim}{\mu})}{\partial \mu_i} \times \dfrac{\partial g(\underset{\sim}{\mu})}{\partial \mu_j} > 0$,

$\sqrt{n}[g(\bar{X}_1, \ldots, \bar{X}_p) - g(\mu_1, \ldots, \mu_p)] \xrightarrow{d} N(0, \tau^2)$

•) Sample Survey:-

(A) $Var(\bar{x}) = \begin{cases} \frac{\sigma^2}{n} & \text{in SRSWR} \\ \frac{N-n}{N-1} \times \frac{\sigma^2}{n}, & \text{SRSWOR} \end{cases}$

(B) u.e of $\sigma^2 = \begin{cases} s^2 = \frac{1}{n-1} \sum_{1}^{n} (x_i - \bar{x})^2, \rightarrow \text{SRS WR} \\ \frac{N-1}{N} s^2, \rightarrow \text{SRSWOR} \end{cases}$

(C) So, u.e of $Var(\bar{x})$ =

(i) $\frac{s^2}{n}$ in SRSWR

(ii) $\frac{N-n}{N} \times \frac{s^2}{n}$ in SRSWOR.

Hypothesis Testing

$\rightarrow \theta = \theta_0$

$H_0 \rightarrow$ Null hypothesis , $H_1 \rightarrow$ Alternative "

$\hookrightarrow \theta = \theta_1$

$W \rightarrow$ Critical Region .

•) $\alpha = P(\text{type I error}) = P(x \in W \mid H_0)$

= Prob of rejecting H_0, when H_0 is true .

•) $\beta = P(\text{type II error}) = P(x \in W^c \mid H_1)$

= prob of accepting H_0 when H_0 is false .

$\Rightarrow \alpha$: Level of Significance $= P(x \in W \mid H_0)$

Size of Critical Region $\rightarrow \sup \beta(\theta)$.

$\Rightarrow 1-\beta$: Power of the test $= P(x \in W \mid H_1)$

$\rightarrow \alpha$: Producer's Risk ; β : Consumer's Risk

•) Uniformly Most Powerful (UMP) test :-

$\Rightarrow$ W is MP (of size α) for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$,

$\hookrightarrow$ simple

$\exists \rightarrow P(x \in W \mid H_0) = \int_W L_0 \, dx = \alpha$

& $P(x \in W \mid H_1) \geq P(x \in W_1 \mid H_1)$

for every other W_1 of size α .

$\Rightarrow$ W is UMP of size α, for testing $H_0: \theta = \theta_0$ against $H_1: \theta \neq \theta_0$ (Composite)
(& corresponding test is UMP of level α), if

(i) $P(x \in W \mid H_0) = \int_W L_0 \, dx = \alpha$ [ie, size α]

(ii) and $P(x \in W \mid H_1) \geq P(x \in W_1 \mid H_1)$ $\forall \theta \neq \theta_0$,

whatever W_1 (satisfying (i)) may be.

9) Neyman - Pearson Lemma :-

Let $k > 0$ be a const. & W critical region of size α st :-

$$W = \left\{ x \in S : \frac{f(x, \theta_1)}{f(x, \theta_0)} > k \right\} \Rightarrow W = \left\{ x \in S : \frac{L_1}{L_0} > k \right\} \to (i)$$

and $\overline{W} = \left\{ x \in S : \frac{L_1}{L_0} \leq k \right\} \to (ii)$

(L_0 & L_1 are likelihood f^ns of $x = (x_1, \ldots, x_n)$ under H_0 & H_1 resp). Then, W is the MP critical region of test hypothesis

$H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$.

*) A test (or corr. critical region) is said to be **unbiased**, if "Power $\geq$ Size of CR"

i.e, $1-\beta \geq \alpha \Leftrightarrow P_\theta(W) \geq P_{\theta_0}(W) \ \forall \ \theta \in \Theta \ (\neq \theta_0)$

*) Every MP/UMP Critical Region (CR) is necessarily unbiased.

ⓐ Likelihood Ratio Test (LRT) :-

Criterion for LRT ; (likelihood ratio) → $\lambda = \dfrac{\sup_{\theta \in \Theta_0} L(x,\theta)}{\sup_{\theta \in \Theta} L(x,\theta)}$

where $\Theta_0 \subset \Theta$

$\ni 0 \leq \lambda \leq 1$. $\left[P[\lambda < \lambda_0 | H_0] = \alpha \right]$

Critical Region is, $\boxed{0 < \lambda < \lambda_0}$

*) If λ is LR for testing H_0 (simple Hypothesis) and $U = \phi(\lambda)$ monotonic ↑(↓) f^n of λ, Then the test based on U is equv to LRT. The CR for test based on U is,

$\phi(0) < U < \phi(\lambda_0)$ $\left(\phi(0) > U > \phi(\lambda_0) \right)$

•) $x_1, \ldots, x_n$ be r.s. from a pop^n with pdf $f(x;\theta_1,\ldots,\theta_k)$. $H_0: \theta_1=\theta_1', \ldots, \theta_r=\theta_r'$, $r<k$

→ known

When H_0 is true, $-2\ln\lambda \sim \chi^2_r$, for large n

Thus at level α Reject H_0 if $-2\ln\lambda > \chi^2_r(\alpha)$

where $P[\chi^2 > \chi^2_r(\alpha)] = \alpha$ (upper α pt)

→ r = # known parameters in H_0

Prop of LRT: → generally UMP, if at all exists

→ Under certain condi^{n}s, $-2\ln\lambda \sim \chi^2$ dist^n.

→ " LR Test is consistent

•) p-value: The prob of observing under H_0, a sample outcome at least as extreme as the one observed is known as p-value (in the direction of H_1)

(*) $\beta(\theta)$ be the power f^n,

size of the test $= \sup_{\theta \in \Omega_0} \beta(\theta)$.

$= \sup_{\theta \in \Omega_0} P_\theta[\underset{\sim}{X} \in \omega] \rightarrow CR$ $\quad$ $\Omega_0 \rightarrow$ Null, $\Omega_1 \rightarrow$ Alt.

rs $\underset{\sim}{X} = (x_1 \ldots x_n)'$

ie, Power f^n, $\beta(\theta) = \begin{cases} P_\theta[\underset{\sim}{X} \in \omega] = P[\text{type I error}], & \text{if } \theta \in \Omega_0 \\ 1 - P_\theta[\underset{\sim}{X} \in \overline{\omega}] = 1 - P[\text{type II error}] & \text{if } \theta \in \Omega_1 \end{cases}$

→ If size $\leq \alpha$, then "level α test".

Scanned by CamScanner $\quad$ ABOUT.ME/SOURAV.NANDI