Hölder $|\mathbb{E}XY| \le \mathbb{E}|XY| \le (\mathbb{E}|X|^p)^{1/p} \cdot (\mathbb{E}|Y|^q)^{1/q}$ $(\frac{1}{p}a^p + \frac{1}{q}b^q \ge ab,\ \frac{1}{p}+\frac{1}{q}=1)$

Cauchy Schwarz & Covariance Ineq $(p,q=2)$ $\mathbb{E}(g(X)h(X)) \le \mathbb{E}g(X)\,h(X)$ $\nearrow\searrow$ $\ge$ $\nearrow\nearrow$ or $\searrow\searrow$

Minkowski $(\mathbb{E}|X+Y|^p)^{1/p} \le (\mathbb{E}|X|^p)^{1/p} + (\mathbb{E}|Y|^p)^{1/p}$ $p \ge 1$

Jensen $\mathbb{E}g(X) \ge g(\mathbb{E}X)$ g convex $M_H \le M_G \le M_A \le M_Q \le \text{Norm}\ (p>2)$

Chebyshev $r P(\varphi(X) \ge r) \le \mathbb{E}\varphi(X)$ $P(|X-\mathbb{E}X| \ge \varepsilon) \le \frac{\operatorname{Var}X}{\varepsilon^2}$ $\mathbb{E}e^{tX}$

By integrating the tails of any dist / By MGF: $P(|X| \ge a) \le e^{-at}M_X(t)$

Bernoulli (p): $\{p,\ p(1-p),\ (1-p)+pe^t\}$ $\Rightarrow$ Binomial (n,p): $\{n\cdots,\ [\cdot\cdot]^n\}$

Poisson (λ): $\{\lambda,\ \lambda,\ \exp[\lambda(e^t-1)]\}$. $P(X=x+1) = \frac{\lambda}{x+1}P(X=x)$

Geometric (p): $\{\frac{1}{p},\ \frac{1-p}{p^2},\ pe^t/(1-(1-p)e^t)\}$, $P(X>s \mid X>t) = P(X>s-t)$

$H(N,M,K)$: $\{\frac{KM}{N},\ \frac{KM}{N}\frac{(N-M)(N-K)}{N(N-1)},\ ?\}$

$NB(r,p)$: $\{\frac{r(1-p)}{p},\ \frac{r(1-p)}{p^2},\ [p/(1-(1-p)e^t]^n\ \ t<-\log(1-p)\}$

Beta (α,β): $\{\frac{\alpha}{\alpha+\beta},\ \alpha\beta/(\alpha+\beta)^2(\alpha+\beta+1),\ \cdots\}$ $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt$

$\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$, $\Gamma(1)=1$, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, $\Gamma(n) = (n-1)!$

Cauchy (θ,σ): $\frac{1}{\pi\sigma}\frac{1}{1+(\frac{x-\theta}{\sigma})^2}$ $\{\infty, \infty, NA\}$ fat-tail ~~exp family~~ $\frac{N}{N}$

χ^2_p: $\{p,\ 2p,\ (\frac{1}{1-2t})^{p/2}\ \ t<\frac{1}{2}\}$ $\sum \chi^2_{p_i} = \chi^2_{\sum p_i}$ $\sum N^2$ $Exp(2) = \chi^2_2$

DExp (μ,σ): $\{\mu,\ 2\sigma^2,\ e^{\mu t}/(1-(\sigma t)^2)\ \ |t|<\frac{1}{\sigma}\}$

Exp (β): $\{\beta,\ \beta^2,\ \frac{1}{1-\beta t}\ \ t<\frac{1}{\beta}\}$ memory loss $\approx$ Geometric.

$F_{m,n}$: $\{\frac{n}{n-2},\ 2(\frac{n}{n-2})^2(m+n-2)/m(n-4),\ NA\}$ $\frac{1}{m}\chi^2_m / \frac{1}{n}\chi^2_n$

Gamma (α,β): $\{\alpha\beta,\ \alpha\beta^2,\ (1/(1-\beta t))^\alpha\ \ t<\frac{1}{\beta}\}$ α shape β scale

Logistic (μ,β): $\{\mu,\ \pi^2\beta^2/3,\ \cdots\}$ cdf $F(x\mid\mu,\beta) = 1/(1+\exp[-(x-\mu)/\beta])$

$N(\mu,\sigma^2)$: $\{\mu,\ \sigma^2,\ \exp\{\mu t + \frac{1}{2}\sigma^2t^2\}\}$

t_n: $\{0,\ \frac{n}{n-2},\ NA\}$ $t_n^2 = F_{1,n}$

Uniform (a,b): $\{(b+a)/2,\ (b-a)^2/12,\ (e^{bt}-e^{at})/(b-a)t\}$

$X \sim N(\mu,\sigma^2) \Rightarrow (n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$ | Any X, $\mathbb{E}S^2 = \sigma^2$

$Z = X+Y$ convolution $\int_{-\infty}^{+\infty} f_X(w)f_Y(z-w)dw$

$[X_1,\cdots,X_n]^T$, $U = AX \in \mathbb{R}^\ell$, $V = BX \in \mathbb{R}^m$ $\operatorname{Cov}(U,V) = [\sum_{i=1}^n a_{ji}b_{ki}\sigma_i^2]_{\ell\times m}$

$\{X_i\} \sim N(\mu_i, \sigma_i^2)$ then $\operatorname{Cov}(U,V) = \vec{0} \Leftrightarrow U \perp\!\!\!\perp V$

$f(x|\theta) = h(x)c(\theta)\exp\{\sum_{i=1}^k w_i(\theta)t_i(x)\}1_S(x)$ support set. $\boxed{\text{Exp Family}}$

$\Rightarrow (T_1,\cdots,T_k) \sim f_T(u_1,\cdots,u_k|\theta) = H(u_1,\cdots,u_k)[c(\theta)]^n\exp\{\sum_{i=1}^k w_i(\theta)u_i\}$ RECHECK.

$f(x|\eta) = h(x)c^*(\eta)\exp\{\sum_{i=1}^d \eta_i t_i(x)\}1_S(x)$ natural paramization

$\mathcal{H} = \{\eta = (\eta_1,\cdots,\eta_d) : \int_{-\infty}^{+\infty} h(x)\exp\{\sum_{i=1}^d \eta_i t_i(x)\}1_S(x) < \infty\}$

MGF (Not for generate moments but characterizing a dist.)

Thm. Restrictions for ensuring the uniqueness of moments (Chf Uniqueness)

(i) $F_X = F_Y$ iff X, Y bdd support & $\mathbb{E}X^r = \mathbb{E}Y^r$ for $r=0,1,\cdots$ (iv) $F_X = F_Y \Leftarrow \phi_X(t) = \phi_Y(t)$ for $\forall t \in \mathbb{R}^k$

(ii) $\cdots$ $\Leftarrow M_X(t) = M_Y(t)$ for $\forall t \in U(0,h)\ \exists h>0$.

(iii) Convergence of MGF. $\lim_{n\to\infty} F_{X_n} = F_X \Leftarrow \lim_{n\to\infty} M_{X_n} = M_X$ for $|t|<h$

Thm. Linear Op $M_{aX+b}(t) = \exp(bt)M_X(at)$

Eg. Poisson Approximation

$X \sim \text{Binomial}(n,p)$ $Y \sim \text{Poisson}(\lambda)$ $np \to \lambda$ $n \to \infty$

$M_X(t) = [pe^t + (1-p)]^n$, $M_Y(t) = \exp\{\lambda(e^t-1)\}$

$= [1 + \frac{1}{n}(e^t-1)np]^n \to \exp\{\lambda(e^t-1)\}$

$\{\lim_{n\to\infty}(e^t-1)np = (e^t-1)\lim_{n\to\infty} np = \lambda(e^t-1)\}$

(Lemma if $\lim_{n\to\infty} a_n = a$, then $\lim_{n\to\infty}(1+\frac{a_n}{n})^n = e^a$ help.

Interchangability between a $\frac{d}{dx}$ and $\int dx$

Thm Lebesgue's Dominated Convergence Theorem

A. $|h(x,y)| \le g(x)$ for $\forall x,y$ & $\int_{-\infty}^\infty g(x)dx < \infty$

$\Rightarrow \lim_{y\to y_0}\int_{-\infty}^\infty h(x,y)dx = \int_{-\infty}^\infty \lim_{y\to y_0} h(x,y)dx$

B. $|(f(x,\theta_0+\delta) - f(x,\theta_0))/\delta| \le g(x,\theta_0)$ for $\forall x, |\delta| < \delta_0 > 0$ Lipschitz condition

& $\int_{-\infty}^\infty g(x,\theta_0)dx < \infty$

$\Rightarrow \frac{d}{d\theta}\int_{-\infty}^\infty f(x,\theta)dx\big|_{\theta=\theta_0} = \int_{-\infty}^\infty \frac{\partial}{\partial\theta}f(x,\theta)\big|_{\theta=\theta_0}dx$

C. $|\frac{\partial}{\partial\theta}f(x,\theta)|_{\theta=\theta'}| \le g(x,\theta)$ for $\forall\theta' \in \{\theta' : |\theta'-\theta| \le \delta_0\}$

& $\int_{-\infty}^\infty g(x,\theta)dx < \infty$

$\Rightarrow \frac{d}{d\theta}\int_{-\infty}^\infty f(x,\theta)dx = \int_{-\infty}^\infty \frac{\partial}{\partial\theta}f(x,\theta)dx$

$\frac{\partial}{\partial\theta}f(x,\theta)\big\vert_{\theta=\theta'}$ (target)	$\xrightarrow{\text{scale}}$	construct a g, s.t. target $\le g(x,\theta)$ (i)
	$\xrightarrow{\text{integrate}}$	justify the dominated g with finite Integration (ii)

Order Statistic

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

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

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

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

Disc. $P(X_{(k)} \le x_i) = \sum_{j=k}^n \binom{n}{k}P_i^j(1-P_i)^{n-j}$ $\begin{cases} P_0 = 0 \\ P_1 = p_1 \\ \vdots \\ P_i = \sum_{j=1}^i p_j \end{cases}$ $P(X=x_i) = p_i$ $x_1<x_2<\cdots<x_n$

$P(X_{(k)} = x_i) = \sum_{j=k}^n \binom{n}{k}[P_i^j(1-P_i)^{n-j} - P_{i-1}^j(1-P_{i-1})^{n-j}]$

Convergence Concepts.

1. $X_n \xrightarrow{P} X$: $\lim_{n\to\infty} P(|X_n - X| \ge \varepsilon) = 0$ or $\lim_{n\to\infty}P(|X_n-X|<\varepsilon) = 1$
   a. WLLN : $\{X_i\}$ iid $\mathbb{E}X_1 = \sigma^2 < \infty$, then $\bar{X}_n \xrightarrow{P} \mathbb{E}X_1$ (Chebyshev's)
   b. $X_n \xrightarrow{P} X \Rightarrow h(X_n) \xrightarrow{P} h(X)$ for $h \in C$
2. $X_n \xrightarrow{a.s.} X$: $P(\lim_{n\to\infty}|X_n - X| < \varepsilon) = 1$ pointwise convergence first
   a. SLLN : $\{X_i\}$ iid $\mathbb{E}X_1 = \sigma^2 < \infty$, then $\bar{X}_n \xrightarrow{a.s.} \mathbb{E}X_1$
3. $X_n \xrightarrow{d} X$: $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$ for $\forall x \in \{$cont. points of $F_X\}$
   $\Leftrightarrow \lim_{n\to\infty} P(X_n < x) = \tau(x)$

Thm. $X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{P} X$, $X_n \xrightarrow{L^r} X \Rightarrow X_n \xrightarrow{P} X$, $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$,

$X_n \xrightarrow{P} \mu \Leftrightarrow X_n \xrightarrow{d} \mu$, μ is a contant.

Thm : CLT : Any $\{X_i\}$ iid, $\operatorname{Var}X_1 = \sigma^2 < \infty$, MGF exist

then $\sqrt{n}(\bar{X}_n - \mu)/\sigma \xrightarrow{d} X \sim N(0,1)$ Proof by MGF

CLT (stronger) : $\cdots$ No need of MGF existance.

the $\sqrt{n}(\bar{X}_n-\mu)/\sigma \xrightarrow{d} X \sim N(0,1)$ Proof by Chf

comparing with order statistic

Thm. Slutsky's If $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{P} a$, then

$X_nY_n \xrightarrow{d} aX$ and $X_n + Y_n \xrightarrow{d} X+a$

(NOTE : if a is r.v., Slutsky's does not hold)

Delta Method for approximating the function of r.v.

Thm Tayler Series $g(x) = T_r(x) + R_r(x)$

$= \sum_{i=0}^r \frac{g^{(i)}(a)}{i!}(x-a)^i + R_r(x)$, $\lim_{x\to a}\frac{R_r(x)}{(x-a)^r} = 0$

$R_r(x) = \int_a^x \frac{g^{(r+1)}(t)}{r!}(x-t)^r dt$

For multi-variables $g(\vec{t}) \approx g(\vec{\theta}) + \sum_{i=1}^k g_i'(\vec{\theta})(t_i - \theta_i)$ $g_i' = \frac{\partial}{\partial t_i}g(\vec{t})\big|_{t_1=\theta_1,\cdots,t_k=\theta_k}$

then $\mathbb{E}g(\vec{T}) \approx \mathbb{E}g(\vec{\theta}) + \sum_{i=1}^k g_i'(\vec{\theta})\mathbb{E}(T_i - \theta_i)$

$= g(\vec{\theta})$

then $\operatorname{Var}g(\vec{T}) \approx \mathbb{E}(g(\vec{T}) - g(\vec{\theta}))^2$

$\approx \mathbb{E}(\sum_{i=1}^k g_i'(\vec{\theta})(T_i - \theta_i)^2)$

$= \sum_{i=1}^k [g_i'(\vec{\theta})]^2\operatorname{Var}T_i + 2\sum_{i<j}g_i'(\vec{\theta})g_j'(\vec{\theta})\operatorname{Cov}(T_i,T_j)$ Local Linear.

Thm 1st order $\sqrt{n}(Y_n - \theta) \xrightarrow{d} N(0,\sigma^2) \Rightarrow \sqrt{n}(g(Y_n) - g(\theta)) \xrightarrow{d} N(0, g'(\theta)^2\sigma^2)$

2nd order $\sqrt{n}(Y_n - \theta) \xrightarrow{d} N(0,\sigma^2) \Rightarrow n(g(Y_n) - g(\theta)) \xrightarrow{d} \frac{g''(\theta)}{2}\sigma^2\chi_1^2$

ratio of Normal dist. = Cauchy dist. (No $\mathbb{E}X$ No $\operatorname{Var}X$).

Data Reduction

- Sufficient : $P(X=y \mid T(X) = T(x))$ is constant wrt θ

Thm : $T(X)$ is sufficient statistic

1. $P_\theta(X=x \mid T(X)=T(x)) = \frac{P(X=x|\theta)}{q(T(X)=T(x)|\theta)} = C_\theta$
2. NFFC (Factorization) : $p(x_{1:n};\theta) = g_\theta(T(x_{1:n}))h(x_{1:n})$

Minimal: $\{T_i(x)\}$ & $\{\eta_i\}$ are both not linear correlation. : Exp dist. minimal.

- Minimal sufficient. T is a func. of T' $\forall T' \Leftrightarrow T(x) = T(y)$ if $T'(x) = T'(y)$

Thm : $T(X)$ is minimal sufficient. [For Exp. Fam. full rank, $\vec{T}(X)$ m.s.] ☆ Thm

$\boxed{\text{Lehmann Scheffe}}$ T s.t. $p(x;\theta) = C_{x,y}P(y;\theta)$ iff $T(x) = T(y)$ for $\forall\theta$ & $C_{x,y} \in \mathbb{R}$

$\frac{P(x;\theta)}{P(y;\theta)} = \frac{\cdots}{\cdots}$ to be constant wrt θ hold, iff $T = \cdots$, then the found T is minimal statistic

(More to find rather than verify) exp-fam. s-dimension 's $(T_i(x))_{i=1}^s$ is minimal statistic

- Complete : if $\mathbb{E}_\theta g(T(X)|\theta) = 0$ implies $P_\theta(g(T|\theta)=0) = 1$ holds for $\forall$ such g & θ.

then T is complete.

- Anillary: $P(S(X)=s) = f_S(s) \perp \theta$ Loc Fam: $R = X_{(i)} - X_{(j)}$

computed from $f_X(x)$ Sca Fam: $Z = X_{(i)}/X_{(j)}$.

1st order Anicillary $\mathbb{E}_\theta(A(X)) \perp \theta$

Thm 1. Complete sufficient $\Rightarrow$ minimal sufficient [Bahadur's theorem]

2. complete sufficient T, $\phi(T) = \mathbb{E}(W(X)|T(X))$ UMVUE

Thm Lehmann-Scheffé

T complete & Sufficient, $\mathbb{E}_\theta h(T(X)) = g(\theta)$ (unbiased for $g(\theta)$)

then $h(T(X))$ is

(a) only function of $T(X)$ unbiased

(b) an UMRUE under any convex func.

(c) the unique UMRUE under any strictly convex func.

Remark find a educated guess of statistic

1. Rao-Blackwellisation
2. solve for δ s.t. $\mathbb{E}_\theta\delta(T(X)) = g(\theta)$ for all $\theta \in \Omega$ directly.
3. Guess

Thm, Basu's V ancillary & T bdd complete & sufficient then $V \perp\!\!\!\perp T$

Thm. complete sufficient $\Rightarrow$ minimal sufficient.

Thm. $\mathcal{P}$ finite family with densities $\{p_i\}_{i=0}^k$ and same support.

(TPE 6.12) $T(X) = (\frac{p_1(X)}{p_0(X)},\cdots,\frac{p_k(X)}{p_0(X)})$ is minimal sufficient

Lemma. $\mathcal{P}_0 \subset \mathcal{P}$, if T is minimal sufficient for $\mathcal{P}_0$ and sufficient for $\mathcal{P}$, then T m.s. for $\mathcal{P}$

Thm One param. family (Exp) : $\{Y_i\}_{i=1}^n \overset{iid}{\sim} E(\eta,1)$.

If $X_i = e^{-Y_i}$ and $\theta = e^{-\eta}$, then $\{X_i\}_{i=1}^n \overset{iid}{\sim} U(0,\theta)$ and $X_{(n)}$ or $Y_{(1)}$ s.c.

Multi-dimensional FI

$\theta \in \mathbb{R}^k$

$I(\theta)$

$= \mathbb{E}_\theta[\nabla_\theta\log f(X|\theta)][\nabla_\theta\log f(X|\theta)]^T$

$= \operatorname{Cov}(\nabla_\theta\log f(X|\theta))$

$= -\mathbb{E}_\theta\nabla_\theta^2\log f(X|\theta)$

$\operatorname{Var}_\theta W(X)$

$\ge [\nabla_\theta\tau(\theta)]^T I^{-1}(\theta)[\nabla_\theta\tau(\theta)]$.

Identity:

1. Starling's formula $n! = \sqrt{2\pi n}(\frac{n}{e})^n$

($\ln n! = \sum_{i=1} \ln i \approx \int_1^n \ln x\,dx$, $n! = \int_0^\infty e^{-x}x^n dx = \Gamma(n+1)$)

2. $\ln(\frac{1+x}{1-x}) = 2\sum_{i=1}^\infty x^{2i-1}/(2i-1)$ (if $x = 1/(2n+1)$, then $\cdots$)

* $X_n = O_p(Y_n)$ iff $\forall\varepsilon>0, \exists C_\varepsilon > 0$ s.t. $\sup P(|X_n| \ge C_\varepsilon|Y_n|) < \varepsilon$

$X_n = o_p(Y_n)$ iff $X_n/Y_n \xrightarrow{P} 0$. X_n/Y_n remains bdd (tightness if $O(1)$)

Exp. Fam. (Extension)

* $\mathbb{E}(\sum_{i=1}^k \frac{\partial w_i(\theta)}{\partial\theta_j}t_i(X)) = -\frac{\partial}{\partial\theta_j}\log c(\theta)$

* $\operatorname{Var}(\sum_{i=1}^k \frac{\partial w_i(\theta)}{\partial\theta_j}t_i(X)) = -\frac{\partial^2}{\partial\theta_j^2}\log c(\theta) - \mathbb{E}(\sum_{i=1}^k \frac{\partial^2 w_i(\theta)}{\partial\theta_j^2}t_i(X))$

Lemma : Stein's $X \sim N(\theta,\sigma^2)$ & $\mathbb{E}|g'(X)| < \infty \Rightarrow \mathbb{E}g(X)(X-\theta) = \sigma^2\mathbb{E}g'(X)$

Thm : Hwang $\mathbb{E}g(X)$, $g(-1)$ bdd (i) $X \sim \text{Poisson}(\lambda)$, $\mathbb{E}\lambda g(X) = \mathbb{E}(Xg(X-1))$

(ii) $X \sim NB(r,p)$, $\mathbb{E}((1-p)g(X)) = \mathbb{E}(\frac{X}{r+X-1}g(X-1))$

Thm : Dominated Convergence

$\{f_n\}$ s.t. $|f_n| < g$ a.e. x, If $\int g\,dx < \infty$ and $\lim_{n\to\infty}f_n(x) = f(x)$ for a.e $x \Rightarrow \int f_n dx = \int f dx$

Thm : CLT (median $\tilde{X}$)

$\sqrt{n}(\tilde{X}_n - \mu) \sim N(0, 1/4[F'(\mu)]^2)$

Thm : CLT ($\lfloor rn \rfloor$-th order statistic $\tilde{\theta}_n$

$\sqrt{n}(\tilde{\theta}_n - \mu) \sim N(0, r(1-r)/[F'(\theta)]^2$

Thm. Supporting Hyperplane theorem.

Find a estimator

1. Sample Moments = Population Moments. $\frac{1}{n}\sum_{i=1}^n X_i^k = \mathbb{E}X^k$ $(\tau(\theta))$

Thm Satterthwaite Approximation

1. avoid minus 2. use obs. to est. expectation of a r.v.
3. keep nominator & denominator same moment.

2. MLE's Invariance : $\hat{\theta} = MLE(\theta) \Rightarrow \mu(\hat{\theta}) = MLE(\mu(\theta))$
3. Bayes Est. $\pi(\theta|x) = f(x|\theta)\pi(\theta)/m(x) \leftarrow$ marginal dist. of x.

Beta-Binomial $f(y,p) = B(n,p)(y) \times Beta(\alpha,\beta)(p) \leftarrow$ model ratio

$= \binom{n}{y}p^y(1-p)^{n-y} \times \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1}$

$f(p|y) \sim Beta(y+\alpha, n-y+\beta)$ $\hat{p}_B = \mathbb{E}(p|Y) = \frac{n}{\alpha+\beta+n}\cdot\frac{y}{n} + \frac{\alpha+\beta}{\alpha+\beta+n}\cdot\frac{\alpha}{\alpha+\beta}$

Normal-Normal(θ) $f(x,\theta) = N(\theta,\sigma^2)(x) \times N(\mu,\tau^2)(\theta)$

$\hat{\theta} = \mathbb{E}(\theta|X=x) = \frac{\tau^2}{\tau^2+\sigma^2}x + \frac{\sigma^2}{\tau^2+\sigma^2}\mu$. $\operatorname{Var}(\theta|X=x) = \frac{\sigma^2\tau^2}{\sigma^2+\tau^2}$

Eval. a estimator

1. MSE
2. Best Unbiased Estimators (UMVUE)

Lemma (*) if $\frac{d}{d\theta}\mathbb{E}_\theta(\frac{\partial}{\partial\theta}\log f(X|\theta)) = \int\frac{\partial}{\partial\theta}\{[\frac{\partial}{\partial\theta}\log f(x|\theta)]f(x|\theta)\}dx$

then $\mathbb{E}_\theta(\frac{\partial}{\partial\theta}\log f(X|\theta))^2 = -\mathbb{E}_\theta(\frac{\partial^2}{\partial\theta^2}\log f(X|\theta))$

Thm Cramér-Rao Ineq. (gives a lower bound of variance of a estimator).

$\{X_i\} \sim$ pdf $f(x|\theta)$ estimator $W(X)$ s.t.

$\begin{cases} \frac{d}{d\theta}\mathbb{E}_\theta W(X) = \int\frac{\partial}{\partial\theta}[W(x)f(x|\theta)]dx & \text{(interchange)} \\ \operatorname{Var}_\theta W(X) < \infty \end{cases}$

then $\operatorname{Var}_\theta W(X) \ge \frac{[\frac{d}{d\theta}\mathbb{E}_\theta W(X)]^2}{\mathbb{E}_\theta(\frac{\partial}{\partial\theta}\log f(X|\theta))^2}$ $\mathbb{E}_\theta W(X) = \tau(\theta)$. $\leftarrow$ Fisher Information $I(\theta)$

if iid $\operatorname{Var}_\theta W(X) \ge \frac{\cdots}{n\mathbb{E}_\theta(\frac{\partial}{\partial\theta}\log f(X|\theta))^2}$ (smaller lower bound)

if Lemma (*) $\cdots \ge \frac{\cdots}{-\mathbb{E}_\theta(\frac{\partial^2}{\partial\theta^2}\log f(X|\theta))}$ (May unattainable).

Cramér-Rao / Information Bound.

3. Sufficiency & Unbiasedness.

Thm. Rao-Blackwell $\{\mathbb{E}X = \mathbb{E}[\mathbb{E}(X|Y)],\ \operatorname{Var}X = \operatorname{Var}[\mathbb{E}(X|Y)] + \mathbb{E}[\operatorname{Var}(X|Y)]\}$

$\mathbb{E}W(X) = \tau(\theta)$, $T(X)$ complete & sufficient, let $\phi(T) = \mathbb{E}(W|T)$ function of only sample.

then $\phi(T)$ is UMVUE of $\tau(\theta)$

Steps 1. See estimation target $\tau(\theta)$

2. (Guess) establish estimator $W(X)$ s.t. $\mathbb{E}_\theta W(X) = \tau(\theta)$
3. Choose a Complete & sufficient statistic $T(X)$
4. $\phi(T) = \mathbb{E}_\theta(W(X)|T(X))$
5. UMVUE$(\tau(\theta)) = \phi(T) = \phi(T(X))$

Thm : $\mathbb{E}W(X) = \tau(\theta)$, $W(X)$ is UMVUE iff $\operatorname{Cov}_\theta(W,U) = 0$ for $\forall U$ s.t. $\mathbb{E}_\theta U = 0$

Loss Function Optimality

$R(\theta,\delta) = \mathbb{E}_\theta L(\theta,\delta(X))$ eg. $L = MSE$ then $R(\theta,\delta) = \mathbb{E}_\theta(\delta(X) - \theta)^2$

$= \operatorname{Var}_\theta\delta(X) + \operatorname{Bias}_\theta^2\delta(X)$

Stein's Loss for Variance est. $L(\sigma^2,a) = \frac{a}{\sigma^2} - 1 - \log\frac{a}{\sigma^2}$

general v. $= K_1\frac{a}{\sigma^2} + K_2 + K_3\log\frac{a}{\sigma^2}$

Bayes Risk $= \int_\Theta R(\theta,\delta(X))\pi(\theta)d\theta = \mathbb{E}R(\theta,\delta(X)) = \mathbb{E}\mathbb{E}_\theta L(\theta,\delta(X))$

$= \int_\Theta \mathbb{E}L(\theta,\delta(X))\pi(\theta)d\theta$

$= \int_\Theta[\int_X L(\theta,\delta(x))f(x|\theta)dx]\pi(\theta)d\theta$

$= \underbrace{\int_X[\int_\Theta L(\theta,\delta(x))\pi(\theta|x)d\theta]}_{\text{A func. of } x \text{ : Bayes Optimal Statistic}}m(x)dx$