M. Ragab

Sufficiency & Completeness

$\underset{\sim}{X} \sim P_\theta$, where $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$

Let $T : \Omega \longrightarrow \mathbb{R}^s$, $T(\underset{\sim}{x})$ is a reduction of the data. That is; a statistic $T(\underset{\sim}{x})$ is fully as informative as the original sample $\underset{\sim}{X}$.

<u>Def.:</u> T is sufficient for θ iff $P_\theta(\underset{\sim}{X} = \underset{\sim}{x} \mid T = t)$ does not depend on θ.

<u>Note:</u> If T is sufficient for $P \in \mathcal{P}$, then T is also sufficient for $P \in \mathcal{P}_0 \subset \mathcal{P}$ but not necessarily sufficient for $P \in \mathcal{P}_1 \supset \mathcal{P}$

<u>Example:</u> Let $\underset{\sim}{X} = (X_1, \ldots, X_n) \overset{r.s.}{\sim}$ Bernoulli(θ) with pdf $f_\theta(x) = \theta^x (1-\theta)^{1-x}$, $x = 0, 1$. Show that $T = \sum^n X_i$ is sufficient stat. for θ.

$$P(\underset{\sim}{X} = \underset{\sim}{x} \mid T = t) = \frac{P(\underset{\sim}{X} = \underset{\sim}{x}, T = t)}{P(T = t)}$$

$$= \frac{\theta^{\sum x_i} (1-\theta)^{n - \sum x_i}}{\binom{n}{t} \theta^t (1-\theta)^{n-t}} = \frac{1}{\binom{n}{t}} \text{ - free of } \theta$$

<u>Factorization</u> <u>Theorem:</u> Suppose $\underset{\sim}{X}$ is a sample from $P \in \mathcal{P}$. Then $T(\underset{\sim}{X})$ is sufficient for $P \in \mathcal{P}$ iff there are functions h (does not depend on P) and g_P (which depends on P) on the range of T such that

$$p_\theta(x) = \frac{dP(x)}{dx} = g_P(T(\underset{\sim}{x}))\, h(\underset{\sim}{x}).$$

<u>Example</u> (<u>Truncation families</u>): Let $\phi(x)$ be a positive measurable function on $(\mathbb{R}, \mathcal{B})$, $\mathcal{F} = \mathcal{B}$ such that $\int_a^b \phi(x)\,dx < \infty$ for any a and b, $-\infty < a < b < \infty$. Let $\underset{\sim}{\theta} = (a, b) \in \Theta = \{(a,b) \in \mathbb{R}^2 : a < b\}$ and

$$f_\theta(x) = c(\theta)\phi(x) I_{(a,b)}(x),$$

where $c(\theta) = \left[\int_a^b \phi(x)\,dx\right]^{-1}$. Then $f_\theta(x)$, called a truncation family, is a parametric family dominated by the Lebesgue measure on $\mathbb{R}$. Let $X_1, \dots, X_n$ be iid r.v.'s from $f_\theta(x)$. Then the joint pdf of $\underset{\sim}{X}$ is

$$f_\theta(\underset{\sim}{x}) = \prod_{i=1}^n f_\theta(x_i) = [c(\theta)]^n \prod_{i=1}^n \phi(x_i) I_{\{a < X_{(1)} < X_{(n)} < b\}}(\underset{\sim}{x})$$

where $X_{(i)}$ is the ith smallest value of $X_1, \dots, X_n$. Let $T(\underset{\sim}{X}) = (X_{(1)}, X_{(n)})$

$$g_\theta(t_1, t_2) = [c(\theta)]^n I_{\{a < X_{(1)} < X_{(n)} < b\}}(\underset{\sim}{x})$$

$$h(\underset{\sim}{x}) = \prod_{i=1}^n \phi(x_i)$$

By Factorization Thm, $T(\underset{\sim}{X})$ is sufficient for $\theta \in \Theta$.

<u>Example</u> (<u>Order Stat's</u>): Let $\underset{\sim}{X} = (X_1, \dots, X_n) \overset{iid}{\sim} P_\theta \in \mathcal{P}$, where $\mathcal{P}$ is the family of dist's. on $\mathbb{R}$ having Lebesgue pdf's. Let $X_{(1)}, \dots, X_{(n)}$ be order statistics (OSs) from the same sample. Note that

$$f(x_1)\cdots f(x_n) = f(x_{(1)})\cdots f(x_{(n)})$$

$\Rightarrow T(\underset{\sim}{X}) = (X_{(1)}, \dots, X_{(n)})$ is Suff. Stat. for $P \in \mathcal{P}$.

Remark: If $T(\underset{\sim}{x})$ is a sufficient stat. and $T = \psi(S)$, where ψ is measurable and S is another statistic, then S is sufficient. For example (Bernoulli data), $\left(\sum_{i=1}^{m} x_i, \sum_{m+1}^{n} x_i\right)$ is suff. for θ.

Remark: If $T(\underset{\sim}{x})$ is a suff. stat. & $T = \psi(S)$ and ψ is not 1-1 function, then $\sigma(T) \subset \sigma(S)$ and T is more useful than S, since it gives more reduction.

Def. (Minimal sufficiency): Let T be a suff. stat. for $P \in \mathcal{P}$,

T is a minimal suff. for $P \in \mathcal{P}$ iff, for other statistic S sufficient for $P \in \mathcal{P}$, there is a measurable function $\psi \ni T = \psi(S)$ a.s. $\mathcal{P}$ [holds except for a set of measure 0]

Example: Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} U(\theta, \theta+1)$, $\theta \in \mathbb{R}$.

$$f_\theta(\underset{\sim}{x}) = I_{\{\theta < x_{(1)} < x_{(n)} < \theta+1\}}(\underset{\sim}{x}) \implies T = (X_{(1)}, X_{(n)}) \text{ suff. for } \theta$$

Note that
$$x_{(1)} = \sup\{\theta : f_\theta(x) > 0\}$$
$$x_{(n)} = 1 + \inf\{\theta : f_\theta(x) > 0\}.$$

If $S(\underset{\sim}{x})$ is any other suff. stat. for θ, then for x with $h(x) > 0$
$$f_\theta(\underset{\sim}{x}) = g_\theta(S(\underset{\sim}{x}))\, h(\underset{\sim}{x}).$$

$$\implies x_{(1)} = \sup\{\theta : g_\theta(S(\underset{\sim}{x})) > 0\}$$
$$x_{(n)} = 1 + \inf\{\theta : g_\theta(S(\underset{\sim}{x})) > 0\}$$

$\implies \exists$ a measurable function $\psi \ni T(\underset{\sim}{x}) = \psi(S(\underset{\sim}{x}))$,

$\implies$ T is minimal suff. stat. for θ.

Theorem: If $P_\theta(\underset{\sim}{x}) \underset{\theta}{\propto} P_\theta(\underset{\sim}{y}) \Longrightarrow T(\underset{\sim}{x}) = T(\underset{\sim}{y})$, then T is minimal.

[$P_\theta(\underset{\sim}{x}) \underset{\theta}{\propto} P_\theta(\underset{\sim}{y})$ means $P_\theta(\underset{\sim}{x}) = \text{Const.} \cdot P_\theta(y)$ viewed as fn. of θ]

Example (Exp. family): $X_1, \ldots, X_n \overset{i.i.d.}{\sim} \text{Exp}(\theta)$ with pdf

$f_\theta(x) = \theta e^{-\theta x}$, $x > 0$. Show that $T(\underset{\sim}{x}) = \sum^n X_i$ is ~~minimal~~ minimal suff. for θ.

$$\frac{f_\theta(\underset{\sim}{x})}{f_\theta(\underset{\sim}{y})} = \frac{\theta^n e^{-\theta \sum^n x_i}}{\theta^n e^{-\theta \sum y_i}} = e^{-\theta\left[\sum^n x_i - \sum^n y_i\right]}$$

does not depend on θ iff $\sum x_i = \sum y_i$. That is, $\sum^n X_i$ is minimal suff. for θ.

Example (more general) Suppose P is an s-parameter exp. family with $P_\theta(\underset{\sim}{x}) = e^{\eta(\theta) T(\underset{\sim}{x}) - A(\theta)} h(x)$, $\theta \in \Theta$

By factorization Thm, $T(\underset{\sim}{x})$ is suff. Let

$$P_\theta(\underset{\sim}{x}) \underset{\theta}{\propto} P_\theta(\underset{\sim}{y}) \Longleftrightarrow e^{\eta(\theta) T(\underset{\sim}{x})} \underset{\theta}{\propto} e^{\eta(\theta) T(\underset{\sim}{y})}$$

$$\Longleftrightarrow \eta(\theta) T(\underset{\sim}{x}) = \eta(\theta) T(\underset{\sim}{y}) + c$$

depends on X's

If θ_0 & θ_1 are any 2 pts in Θ,

$$[\eta(\theta_0) - \eta(\theta_1)] T(\underset{\sim}{x}) = [\eta(\theta_0) - \eta(\theta_1)] T(\underset{\sim}{y})$$

$$[\eta(\theta_0) - \eta(\theta_1)][T(\underset{\sim}{x}) - T(\underset{\sim}{y})] = 0$$

$$\Rightarrow T(\underset{\sim}{x}) = T(\underset{\sim}{y}) \Longrightarrow T(\underset{\sim}{x}) \text{ is minimal suff.}$$

<u>Example</u>: $X_1, X_2, \ldots, X_n \overset{iid}{\sim} f_\theta(x) = \frac{1}{2} e^{-|x-\theta|}$.

The joint pdf is $f_\theta(\underset{\sim}{x}) = \frac{1}{2^n} e^{-\sum_{}^{n} |x_i-\theta|}$

$T(\underset{\sim}{X}) = (X_{(1)}, \ldots, X_{(n)})$ is suff. stat.

$f_\theta(\underset{\sim}{x}) \underset{\theta}{\propto} f_\theta(\underset{\sim}{y}) \iff \sum |X_i-\theta| = \sum |Y_i-\theta| + c$

$\iff \sum |X_i-\theta| - \sum |Y_i-\theta| = c$

The difference can only be const. in θ iff $\underset{\sim}{X}$ & $\underset{\sim}{Y}$ have the same OSs. $\Rightarrow T(\underset{\sim}{X}) = (X_{(1)}, \ldots, X_{(n)})$ is minimal suff.

Completeness

<u>Def.</u>: A stat. T is complete for $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ if

$E_\theta f(T) = c$, for all $\theta \Rightarrow f(T) = 0$ (a.s. $\mathcal{P}$)

<u>Example</u>: $X_1, \ldots, X_n \overset{iid}{\sim} U(0, \theta)$.

By factorization theorem, $T = X_{(n)}$ is suff.

$G(t) = P(X_{(n)} \le t) = P(X_1 \le t, \ldots, X_n \le t) = \prod^n P(X_i \le t) = \left(\frac{t}{\theta}\right)^n$

The pdf of T is

$g(t) = \frac{n}{\theta^n} t^{n-1}, \quad 0 < t < \theta$

$$E f(T) = 0 \iff \int_0^\theta f(t) \cdot \frac{n t^{n-1}}{\theta^n} dt = 0$$

$$\iff \int_0^\theta f(t) t^{n-1} dt = 0$$

$$\overset{\text{Diff.}}{\iff} f(\theta)\theta^{n-1} = 0$$

$$\iff f(\theta) = 0, \ \forall \theta \in \Theta$$

$\Rightarrow T = X_{(n)}$ is complete suff. stat.

<u>Theorem:</u> If T is complete & suff., then T is minimal sufficient.

<u>Proof:</u> Need to show if $\tilde{T}$ is a minimal suff. stat. then $T = g(\tilde{T})$ for some g.

Define $g(\tilde{T}) = E(T|\tilde{T})$ indep. of θ since $\tilde{T}$ is suff.

$$\Rightarrow E\, g(\tilde{T}) = EE(T|\tilde{T}) = ET$$

$$\Rightarrow E\left(g(\tilde{T}) - T\right) = 0 \xRightarrow{\text{completeness}} T = g(\tilde{T}) \Rightarrow T \text{ is minimal}$$

<u>Def.:</u> An exponential density $p_\theta(x) = e^{\eta(\theta)T(x) - A(\theta)} h(x)$, $\theta \in \Theta$ is said to be full rank if $T_1, T_2, \ldots, T_s$ do not satisfy a linear constraint of the form $\underset{\sim}{u}'\underset{\sim}{T} = c$.

<u>Theorem:</u> In an exponential family of full rank, T is complete.

<u>Def.</u> A stat. V is called ancillary if its distribution does not depend on θ

<u>Theorem (Basu Thm):</u> If T is complete & suff. for $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$, and V is ancillary, then T and V are indep. under P_θ for any $\theta \in \Theta$.

Proof: Define $q_A(T) = P_\theta(V \in A | T)$, $p_A = P_\theta(V \in A)$ $\Big\rangle$ not depend on θ

$$p_A = P_\theta(V \in A) = E_\theta P_\theta(V \in A | T) = E_\theta q_A(T)$$

$\xRightarrow{\text{completeness}}$ $q_A(T) = p_A \quad (\text{a.s. } \mathcal{P})$

$$P(T \in B, V \in A) = E[I_B(T) I_A(V)]$$

$$= EE[I_B(T) \cdot I_A(V) | T]$$

$$= E[I_B(T) E(I_A(V) | T)]$$

$$= E I_B(T) \cdot q_A(T)$$

$$= E I_B(T) \cdot p_A$$

$$= P(T \in B) \cdot P(V \in A)$$, for A, B arbitrary Borel sets

$\Rightarrow$ T & V are indep.

Example: $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ and $\mathcal{P} = \{P_\sigma : P_\sigma$ is family of normal dist's with fixed $\sigma\}$. The joint density is

$$p_\mu = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{\frac{n\mu}{\sigma^2}\bar{X} - \frac{n\mu^2}{2\sigma^2} - \frac{1}{2\sigma^2}\sum X_i^2}$$

These densities for P_σ form a full rank exp. family

$\Rightarrow$ $\bar{X}$ is complete suff. for $\mathcal{P}_\sigma$. define

$$S^2 = \frac{1}{n-1}\sum^n (X_i - \bar{X})^2$$ — sample Var.

To see, S^2 is ancillary, let $Y_i = X_i - \mu$, $i=1,2,\cdots,n \sim N(0,\sigma^2)$

$\Rightarrow Y_1, Y_2, \cdots, Y_n \overset{iid}{\sim} N(0,\sigma^2)$

$\Rightarrow Y_i - \bar{Y} \overset{\mathcal{D}}{=} X_i - \bar{X}$

$\Rightarrow S^2 = \frac{1}{n-1}\sum^{n}(Y_i - \bar{Y})^2$ is ancillary for $\mathcal{P}_\sigma$.

By Basu's Thm, $\bar{X}$ & S^2 are indep.

Rao-Blackwell Theorem

Let T be a sufficient stat. for $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$, let δ be an estimator of $g(\theta)$, and define $T^* = E(\delta(X)|T)$.
If $R(\theta,\delta) < \infty$ and $L(\theta,\cdot)$ is convex, then

$$R(\theta, T^*) \le R(\theta,\delta)$$

Furthermore, if $L(\theta,\cdot)$ is strictly convex, the inequality will be strict unless $\delta(\underset{\sim}{X}) = T^*$ (a. e. $\mathcal{P}_\theta$).

Proof: By Jensen's inequality

$$L(\theta, T^*) = L(\theta, E(\delta(\underset{\sim}{X})|T)) \le E[L(\theta,\delta)|T]$$

$$\overset{\text{Expectation}}{\Longrightarrow} R(\theta, T^*) \le R(\theta,\delta)$$

Unbiasedness

$\hat{\theta} = \delta(\underset{\sim}{X})$ is unbiased estimator of $g(\theta)$ if $E(\hat{\theta}) = g(\theta)$

- How large is the class of all unbiased estimators of θ?
- What functions $g(\theta)$ have unbiased estimators?
- Is it possible to find a better unbiased estimator than $\hat{\theta}$?

Example: $X \sim U(0,\theta)$. δ is unbiased of $g(\theta)$ if

$$\int_0^\theta \delta(x)\frac{1}{\theta}dx = g(\theta),\ \forall\ \theta \quad \text{or} \quad \int_0^\theta \delta(x)\,dx = \theta\, g(\theta)$$

$$\overset{\text{Diff.}}{\Longrightarrow} \quad \delta(\theta) = \theta g'(\theta) + g(\theta) \quad \text{or} \quad \delta(x) = x g'(x) + g(x)$$

For example, if $g(\theta) = \theta \Longrightarrow \delta(X) = 2X$.

Example: $X \sim B(n,\theta)$, $g(\theta) = \sin\theta$

$\delta(x)$ is unbiased if $\underbrace{\sum_{k=0}^{n} \delta(x)\binom{n}{k}\theta^k(1-\theta)^{n-k}}_{\text{Polynomial of } \theta} = \sin\theta \quad \forall\theta$

$\Rightarrow g(\theta) = \sin\theta$ is not U-estimable.

Remark: By considering SE loss function, $L(\theta,\delta) = (d - \delta(x))^2$, the risk of an unbiased estimator δ is

$$R(\theta,\delta) = E_\theta\big(\delta(x) - g(\theta)\big)^2 = \text{Var. of } \delta(x) = Var(\delta(x))$$

Our goal is to minimize $R(\theta,\delta)$.

Def.: An unbiased estimator δ is uniformly minimum variance unbiased (UMVU) est. if

$$Var(\delta) \le Var(\delta^*),\ \forall\ \theta \in \Theta,$$

for any competing unbiased estimator δ^*.

<u>Def.</u>: for all $U, V \in L_2(P_\theta)$, define $\langle U, V\rangle_\theta = Cov_\theta(U,V)$
inner product

$$\|U\|_\theta = \sqrt{\langle U, U\rangle_\theta} = \sqrt{Var_\theta(U)}$$

<u>Notes</u>: (1) $(L_2(P_\theta), \|\ \|)$ is complete in the sense

$$\|U\| = 0 \Rightarrow U \text{ is const. a.s.}$$

(2) Such a "complete" normed linear space is called Banach space.

Define $$\mathcal{U} = \{U \in L_2(P) \mid E_\theta U = 0, \forall \theta \in \Theta\}$$

$$\mathcal{U}^\perp = \{V \in L_2(P) \mid \langle U, V\rangle_\theta = 0, \forall U \in \mathcal{U}\}$$

<u>Theorem</u>: Let $\delta \in L_2(P)$. Then δ is UMVUE iff $\delta \in \mathcal{U}^\perp$

<u>Theorem</u>: Suppose g is U-estimable and S is complete. Then there is unique unbiased estimator based on S (UMVU).

<u>Proof</u>: $\Longrightarrow$

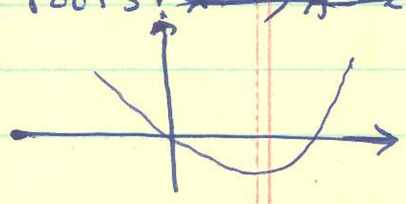
Fix $U \in \mathcal{U}$, $\theta \in \Theta$ and for any $\lambda \in \mathbb{R}$, let $\delta' = \delta + \lambda U$. Then $E_\theta(\delta') = g(\theta)$ and $Var(\delta') \geq Var(\delta)$

$$Var(\delta') = Var(\delta) + 2\lambda Cov(\delta, U) + \lambda^2 Var(U) \geq Var(\delta)$$

$$\Rightarrow h(\lambda) = \lambda^2 Var(U) + 2\lambda Cov(\delta, U) \geq 0$$

$\Rightarrow$ The quadratic eq. in λ has 2 real roots: $\cancel{\lambda = 0, \lambda = -2Cov}$

$\lambda = 0$, $\lambda = -2\,Cov(\delta, U)/Var(U)$.



$h(\lambda)$ takes on negative values unless $Cov(\delta, U) = 0 \Rightarrow \delta \in \mathcal{U}^\perp$

$\Leftarrow$ Suppose $Cov(\delta, U) = 0 \Rightarrow E(\delta U) = 0, \ \forall \theta \in \Theta$.
To show that δ is UMVUE, let δ' be any unbiased of $E(\delta) = g(\theta)$.

$\Longrightarrow \delta - \delta' \in \mathcal{U} \Rightarrow \cancel{E\delta^2 = E(\delta}$

$\Longrightarrow E\,\delta(\delta - \delta') = 0$ by assumpt.

$\Longrightarrow E\delta^2 = E(\delta\delta')$

$\Longrightarrow E\delta^2 - (g(\theta))^2 = E(\delta\delta') - (g(\theta))^2$

$\Longrightarrow Var(\delta) = Cov(\delta, \delta')$

From $\dfrac{Cov(\delta, \delta')}{\sqrt{Var(\delta) \cdot Var(\delta')}} \leq 1$, we have $Var(\delta) \leq Var(\delta')$

$\Longrightarrow \delta$ is UMVUE.

Example: Let $X_1, X_2, \cdots, X_n \overset{iid}{\sim} U(0, \theta)$. $T = \max(X_1, \cdots, X_n)$ is complete suff. stat. Suppose $\delta(T)$ is unbiased for $g(\theta)$. Then

$$\int_0^\theta \delta(t) \frac{n t^{n-1}}{\theta^n} dt = g(\theta), \ \theta > 0$$

$$\Rightarrow \quad n \int_0^\theta \delta(t) t^{n-1} dt = \theta^n g(\theta)$$

$$\overset{\text{Diff.}}{\Longrightarrow} \quad n\,\delta(\theta)\,\theta^{n-1} = \theta^n g'(\theta) + n\theta^{n-1} g(\theta)$$

$$\delta(\theta) = \frac{\theta}{n} g'(\theta) + g(\theta)$$

OR $\delta(t) = g(t) + \dfrac{t\,g'(t)}{n}$

If $g(\theta) = \theta$, $\delta(t) = t + \dfrac{t}{n} = \dfrac{n+1}{n} t$ UMVUE of θ.

Comparisons

$$\delta(t) = \frac{n+1}{n} t, \quad \delta^*(\cdot) = 2\bar{X}$$

- $E\,\delta(t) = \theta$, $E\,\delta^*(\cdot) = \theta$
- $Var(\delta^*) = Var(2\bar{X}) = 4\frac{\sigma^2}{n} = 4\frac{\theta^2}{12n} = \frac{\theta^2}{3n}$
- $\frac{X_{(n)}}{\theta} \sim Beta(n, 1)$

$$Var(X_{(n)}) = \theta^2 \frac{n}{(n+1)^2(n+2)} \Rightarrow Var(\delta(t)) = \frac{\theta^2}{n(n+2)}$$

- $Ratio(\delta^*, \delta) = \frac{Var(\delta)}{Var(\delta^*)} = \frac{3}{n+2} \longrightarrow 0$, as $n \to \infty$

$\Longrightarrow$ $\delta(t)$ is more effective than δ^*.

Example: Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim}$ Bernoulli(θ) with

$$P(X_i = 1) = \theta = 1 - P(X_i = 0), \; i = 1, 2, \ldots, n$$

$$f_\theta(x) = \theta^x (1-\theta)^{1-x}, \; x = 0, 1$$

$$f_\theta(\underset{\sim}{x}) = \theta^{\sum x_i} (1-\theta)^{n - \sum x_i} \quad \text{Exp. family}$$

$$\Longrightarrow T(\underset{\sim}{X}) = \sum^n X_i \sim B(n, \theta) \text{ complete suff. stat.}$$

Let us consider unbiased estimator of $g(\theta) = \theta^2$. One unbiased estimator is $\delta = X_1 X_2$. The UMVUE must be

$$\delta(t) = E(X_1 X_2 \mid T)$$
$$= P(X_1 = X_2 = 1 \mid T)$$

$$P(X_1 = X_2 = 1, T = t) = P\left(X_1 = X_2 = 1, \sum_{i=3}^{n} X_i = t-2\right)$$
$$= \theta^2 \binom{n-2}{t-2} \theta^{t-2} (1-\theta)^{n-t}$$

$$P(X_1 = X_2 = 1 \mid T = t) = \frac{\theta^2 \binom{n-2}{t-2} \theta^{t-2} (1-\theta)^{n-t}}{\binom{n}{t} \theta^{t} (1-\theta)^{n-t}} = \frac{t(t-1)}{n(n-1)}$$

$\Rightarrow \delta(t) = \dfrac{T(T-1)}{n(n-1)}$ is UMVUE of θ^2

Example (Biased estimator may compete UMVUE)

$X_1, X_2, \ldots, X_n \overset{iid}{\sim} U(0, \theta)$, $T = X_{(n)}$ is complete suff. stat.

$\delta(t) = \frac{n+1}{n} T$ is UMVUE of θ

Consider $\delta_a = aT$, $ET = \dfrac{\theta n}{n+1}$

$$R(\theta, \delta_a) = E(aT - \theta)^2 = Var(aT - \theta) + (aET - \theta)^2$$
$$= a^2 Var(T) + \underbrace{(aET - \theta)^2}_{(\text{Bias})^2}$$

$ET = \theta E\left(\dfrac{X_{(n)}}{\theta}\right) = \dfrac{\theta n}{n+1}$ since $\dfrac{X_{(n)}}{\theta} \sim Beta(n, 1)$

$$Var(T) = \theta^2 \frac{n}{(n+1)^2 (n+2)}$$

$$\Rightarrow R(\theta, \delta_a) = \frac{a^2 \theta^2 n}{(n+1)^2 (n+2)} + \theta^2 \left(\frac{na}{n+1} - 1\right)^2$$

The risk is minimized when $\dfrac{2an}{(n+1)^2(n+2)} + 2\left(\dfrac{na}{n+1} - 1\right)\dfrac{n}{n+1} = 0$

$$\Leftrightarrow \frac{a}{(n+1)(n+2)} + \frac{na}{n+1} - 1 = 0$$

$$\Leftrightarrow a + n(n+2)a - (n+1)(n+2) = 0$$

$$\Leftrightarrow (n+1)^2 a = (n+1)(n+2) \Rightarrow \boxed{a = \frac{n+2}{n+1}}$$

$$\Rightarrow R(\theta, \delta_a) = \frac{\theta^2}{(n+1)^2} < \frac{\theta^2}{n(n+2)} = R(\theta, \delta)$$