Module 3: Invariant Estimation

$\underline{X} = (X_1, \dots, X_n)$: a random sample with $\underline{X}$ having an unknown d.f. $F \in \mathcal{P}$. (or $F_\theta \in \mathcal{P}$, $\theta \in \Theta$)

Goal: For a given function g defined on $\mathcal{P}$ or Θ, to estimate $g(F)$ (or $g(\underline{\theta})$) based on data $\underline{X}$. here g may take values in a subset of $\mathbb{R}^k$.

$\mathcal{X}$: sample space of $\underline{X}$;

$g(F)$ or $g(\theta)$: Estimand

$\mathcal{A}$: action space

$L(F, a)$ or $L(\underline{\theta}, a)$: Loss function

Assumption: No two different actions have the same value of loss function $L(F, a)$ (or $L(\underline{\theta}, a)$) for every $F \in \mathcal{P}$ ($\theta \in \Theta$).

Principle of Rational Invariance: The actions/decisions taken in an estimation/decision problem should correspond (i.e. they do not depend on unit of measurements of $\underline{X}$).

Invariance Principle: If two decision problems have the same formal structure (in terms of $\mathcal{X}$, $\mathcal{P}$, $\mathcal{A}$, L) then the same decision rule should be used in each problem.

Example $X \sim N(\theta, 1)$ $\theta \in \Theta = \mathbb{R}$. For any $c \in \mathbb{R}$ estimation of θ and $\theta + c$ are identical. If δ is estimate for θ then $\delta + c$ should be estimate for $\theta + c$.

Consider For any $c \in \mathbb{R}$, $X + c \sim N(\theta + c, 1)$. Estimate of $\theta + c$ based on $X + c$ should correspond to estimate of θ based on X. $\Rightarrow \delta(X + c) = \delta(X) + c$, $\forall c \in \mathbb{R}$

G: a class of 1-1 and onto transformations of $\mathcal{X}$ into itself.

Definition: $\mathcal{G}$ is called a group if

(i) the identity transformation $e \in \mathcal{G}$ (here $e(x) = x$, $\forall x \in \mathcal{X}$).

(ii) $g_1, g_2 \in \mathcal{G} \Rightarrow g_1 \circ g_2 \in \mathcal{G}$

(iii) $g \in \mathcal{G} \Rightarrow \exists\, g^{-1} \in \mathcal{G}$ such that $g^{-1}(g(x)) = x$ $\forall x \in \mathcal{X}$, here $g^{-1}(\cdot)$ is called inverse transformation of $g(\cdot)$

Example (multiplicative group or the group of scale transformations)

$\mathcal{X} = \mathbb{R}$ or $\mathcal{X} = (0, \infty)$.

$\mathcal{G} = \{ g_c : c > 0 \}$, where $g_c(x) = cx$, $c \in (0, \infty)$, $x \in \mathcal{X}$.

Clearly $\mathcal{G}$ is a group with

$$e(x) = g_1(x) = x, \quad x \in \mathcal{X}$$

$$g_c^{-1}(x) = \frac{x}{c} = g_{\frac{1}{c}}(x), \quad x \in \mathcal{X}, \quad c \in (0, \infty)$$

Example (Additive group or location group on $\mathbb{R}$)

$\mathcal{X} = \mathbb{R}$

$\mathcal{G} = \{ g_c : c \in \mathbb{R} \}$, where $g_c(x) = x + c$, $x \in \mathcal{X}$, $c \in \mathbb{R}$.

Clearly $\mathcal{G}$ is a group with

$$e(x) = g_0(x) \quad x \in \mathcal{X}$$

and $g_c^{-1}(x) = x - c = g_{-c}(x)$, $x \in \mathcal{X}$, $c \in \mathbb{R}$.

Example (Affine Group)

$\mathcal{X} = \mathbb{R}$

$\mathcal{G} = \{ g_{b,c} : -\infty < b < \infty, c > 0 \}$, where $g_{b,c}(x) = cx + b$, $x \in \mathcal{X}$, $c > 0$, $b \in \mathbb{R}$.

Clearly $\mathcal{G}$ is a group with

$$e(x) = g_{0,1}(x), \quad x \in \mathbb{R}$$

$$g^{-1}_{b,c}(x) = \frac{x-b}{c} = g_{-\frac{b}{c},\frac{1}{c}}(x), \quad c>0, \ b \in \mathbb{R}.$$

Example (Permutation Group):

$$\mathcal{X} = \mathcal{X}_0 \times \cdots \times \mathcal{X}_0 = \mathcal{X}_0^n$$

$$\mathcal{G} = \{ g_{\underline{\alpha}} : \underline{\alpha} = (\alpha_1, \ldots, \alpha_n) \text{ is a permutation of } (1,\ldots,n) \},$$

where

$$g_{\underline{\alpha}}(\underline{x}) = (x_{\alpha_1}, \ldots, x_{\alpha_n}), \quad \underline{x} \in \mathcal{X}, \quad \underline{\alpha} \in S_n = \text{set of all permutations of } (1,\ldots,n)$$

It can be verified that $\mathcal{G}$ is a group with

$$e(\underline{x}) = g_{\underline{1}}(\underline{x}), \quad \underline{x} \in \mathcal{X}, \quad \text{where } \underline{1} = (1, 2, \ldots, n)$$

$$g^{-1}_{\underline{\alpha}}(\underline{x}) = g_{\underline{\alpha}^{-1}}(\underline{x}), \quad \underline{x} \in \mathcal{X},$$

where $\underline{\alpha}^{-1} = (\alpha_1^{-1}, \ldots, \alpha_n^{-1})$ is the inverse permutation of $\underline{\alpha}$ (if $y_i = x_{\alpha_i}$, $i=1,\ldots,n$, then $y_{\alpha_i^{-1}} = x_i$, $i=1,\ldots,n$)

Definition: $\mathcal{P}$ is said to be invariant under $\mathcal{G}$ if for every $g \in \mathcal{G}$ and $F \in \mathcal{P}$ if $\underline{X} \sim F$ then there is a unique $F^* \in \mathcal{P}$ such that $Y = g(\underline{X}) \sim F^*$. We denote F^* by $\bar{g}(F)$.

Remark: (i) If $\mathcal{P}$ is invariant under $\mathcal{G}$ then

$$P_F(g(\underline{X}) \in A) = P_{\bar{g}(F)}(\underline{X} \in A), \quad F \in \mathcal{P}, \ A \subseteq \mathcal{X}, \ g \in \mathcal{G}$$

and for any function $h : \mathcal{X} \to \mathbb{R}^d$

$$E_F(h(g(\underline{X}))) = E_{\bar{g}(F)}(h(\underline{X})), \quad F \in \mathcal{P}, \ g \in \mathcal{G}.$$

Home Assignment:

Show that $\bar{G} = \{\bar{g} : g \in G\}$ is a group of transformations of $\mathcal{P}$ into itself.

Definition (a) Let $\mathcal{P}$ be invariant under G. The loss function $L(F, a)$ is said to be invariant under G if for every $g \in G$ and $a \in \mathcal{A}$ there exists an $a^* \in \mathcal{A}$ such that

$$L(F, a) = L(\bar{g}(F), a^*), \quad \forall F \in \mathcal{P}.$$

We denote a^* by $\tilde{g}(a)$.

(b) A decision problem is said to be invariant under the group G if $\mathcal{P}$ is invariant under G and $L(F, a)$ is invariant under G.

Home Assignment:

Show that $\tilde{G} = \{\tilde{g} : g \in G\}$ is a group of transformations of $\mathcal{A}$ into itself.

Example: Let $\mathcal{P} = \{F_\theta : \theta > 0\}$ where

$$F_\theta(x) = F\left(\frac{x}{\theta}\right), \quad x \in \mathbb{R}, \quad \theta > 0$$

for some d.f. F.

Goal: To estimate θ under the squared error loss function

$$L(\theta, a) = \left(\frac{a}{\theta} - 1\right)^2, \quad a > 0, \quad \theta > 0.$$

Consider the scale group

$$G = \{g_c : g_c(x) = cx, \ x \in \mathcal{X}, \ c > 0\}$$

If $X \sim F_\theta$ then $Y = g_c(X) = cX$ has d.f.

$$F\left(\frac{x}{c\theta}\right) = F_{c\theta}(x)$$

$$\Rightarrow \quad \bar{g}_c(\theta) = c\theta, \quad c > 0, \quad \theta > 0$$

Thus $\mathcal{P}$ is invariant under $\mathcal{G}$

Let $\tilde{g}_c(a) = ca$, $c>0$, $a \in \mathcal{A} = (0, \infty)$. Then

$$L(\theta, a) = \left(\frac{a}{\theta}-1\right)^2 = \left(\frac{ca}{c\theta}-1\right)^2 = L(\bar{g}_c(\theta), \tilde{g}_c(a))$$

Thus the decision / estimation problem is invariant under $\mathcal{G}$.

Home Assignment: (i) Let $X \sim N(\theta, 1)$, where $\theta \in \Theta = \mathbb{R}$ is unknown. Let $\mathcal{A} = \Theta$ and consider estimating θ under the loss function

$$L(\theta, a) = W(|a-\theta|) \quad a \in \mathcal{A}, \ \theta \in \Theta.$$

Show that the estimation/decision problem is invariant under the additive group $\mathcal{G} = \{g_c : c \in \mathbb{R}\}$, where $g_c(x) = x + c$, $x \in \mathcal{X}$, $c \in \mathbb{R}$.

(ii) Let $X \sim Bin(n, \theta)$, $\theta \in \Theta = (0,1)$. Let $\mathcal{A} = \Theta = (0,1)$ and consider problem of estimating θ under squared error loss function

$$L(\theta, a) = (a-\theta)^2, \quad a \in \mathcal{A}, \ \theta \in \Theta.$$

Find a suitable group of transformations under which the problem of estimating θ under the squared error loss function $L(\theta, a)$ is invariant.

Invariant Estimators: Consider an estimation problem $(\mathcal{P}, \mathcal{A}, L)$ which is invariant under a group $\mathcal{G}$ of transformations on $\mathcal{X}$. Then formal structures of problem involving X and $Y = g(X)$ $(\forall g \in \mathcal{G})$ are the same (in terms of $\mathcal{X}$, $\mathcal{P}$ and L)

(a) invariance principle $\Rightarrow$ estimators δ and δ^* used in X and Y problems should be the same

(b) Principle of Rational Invariance

$\Rightarrow$ Action taken in two problems should correspond

(a) + (b) $\Rightarrow$ a decision rule (estimator) δ should satisfy $\Rightarrow$

$$\delta(A|x) = \delta(\tilde{g}(A)|g(x)), \quad \forall g \in \mathcal{G},\ A \subseteq \mathcal{A}. \quad \ldots\ldots (I),$$

where $\tilde{g}(A) = \{\tilde{g}(a) : a \in \mathcal{A}\}, g \in \mathcal{G}$.

For given $X = x$, let Y_x be the random variable (element) corresponding to probability distribution $\delta(\cdot|x)$. Then

δ is invariant (i.e. satisfies (I))

$$\Leftrightarrow \delta(A|x) = Pr(Y_x \in A)$$
$$= \delta(\tilde{g}(A)|g(x))$$
$$= Pr(Y_{g(x)} \in \tilde{g}(A))$$
$$= Pr(\tilde{g}^{-1}(Y_{g(x)}) \in A), \quad \forall A \subseteq \mathcal{A},\ g \in \mathcal{G},\ x \in \mathcal{X}$$

$$\Leftrightarrow Y_x \overset{d}{=} \tilde{g}^{-1}(Y_{g(x)}), \quad \forall x \in \mathcal{X}, g \in \mathcal{G}.$$

$$\Leftrightarrow \tilde{g}(Y_x) \overset{d}{=} Y_{g(x)}, \quad \forall x \in \mathcal{X}, g \in \mathcal{G}.$$

Definition: If a decision problem is invariant under a group $\mathcal{G}$ of transformations then a decision rule (estimator) δ is said to be invariant under $\mathcal{G}$ iff, for all $x \in \mathcal{X}$, $g \in \mathcal{G}$ and $A \subseteq \mathcal{A}$

$$\delta(A|x) = \delta(\tilde{g}(A)|g(x)).$$

Remark: A non-randomized estimator δ is invariant under $\mathcal{G}$ iff

$$\delta(g(x)) = \tilde{g}(\delta(x)), \quad \forall x \in \mathcal{X}, g \in \mathcal{G}.$$

Home Assignment: (i) Let $\mathcal{P} = \{F_\theta : \theta > 0\}$ where $F_\theta(x) = F(\frac{x}{\theta})$, $x \in \mathcal{X} = (0, \infty)$, $\theta > 0$, for some d.f. F. Consider problem of estimating θ under scaled squared error loss function

$$L(\theta, a) = \left(\frac{a}{\theta} - 1\right)^2, \quad a \in \mathcal{A} = (0, \infty), \quad \theta \in \Theta = (0, \infty).$$

Show that a non-randomized estimator δ is invariant under scale group $\mathcal{G} = \{g_c : c > 0\}$ where $g_c(x) = cx$, $c > 0$, $x \in \mathcal{X}$, iff $\delta(x) = dx$, $x \in (0, \infty)$, for some $d > 0$.

(ii) Let $\mathcal{P} = \{F_\theta : \theta \in \mathbb{R}\}$, where $F_\theta(x) = F(x - \theta)$, $x \in \mathcal{X} = \mathbb{R}$, $\theta \in \Theta = \mathbb{R}$, for some d.f. F. Consider problem of estimating θ under squared error loss function

$$L(\theta, a) = (a - \theta)^2, \quad a \in \mathcal{A} = \mathbb{R}, \quad \theta \in \Theta = \mathbb{R}.$$

(or in general $L(\theta, a) = W(|a - \theta|)$, $a \in \mathcal{A} = \mathbb{R}$, $\theta \in \Theta = \mathbb{R}$, for some non-negative function $W(\cdot)$ defined on $[0, \infty)$.) Show that the estimation problem is invariant under additive group $\mathcal{G} = \{g_c : c \in \mathbb{R}\}$ where $g_c(x) = x + c$, $x \in \mathcal{X}$, $c \in \mathbb{R}$. Also show that a non-randomized estimator δ is invariant under additive group $\mathcal{G}$ iff $\delta(x) = x + d$, $x \in \mathbb{R}$, for some $d \in \mathbb{R}$.

(c) **Theorem:** Consider a decision (or estimation) problem $(\mathcal{P}, \mathcal{A}, L)$ invariant under a group of transformations $\mathcal{G}$. Let δ be an invariant decision rule under $\mathcal{G}$. Then

$$R_\delta(F) = R_\delta(\bar{g}F), \quad \forall F \in \mathcal{P}, \ \forall g \in \mathcal{G}.$$

Proof. For given $X = x$ ($x \in \mathcal{X}$) let Z_x be the random element corresponding to probability distribution $\delta(\cdot|x)$. Then

$$\delta \text{ is invariant} \Rightarrow Z_x \stackrel{d}{=} \tilde{g}^{-1} Z_{g(x)}$$

and, for $g \in G$

$$L^*(\delta, F, x) = E(L(F, Z_x))$$
$$= E(L(F, \bar{g}^{-1} Z_{g(x)}))$$
$$= E(L(\bar{g}F, Z_{g(x)}))$$

$$\Rightarrow \quad R_\delta(F) = E_F(L^*(\delta, F, X))$$
$$= E_F(E(L(\bar{g}F, Z_{g(X)})))$$
$$= E_{\bar{g}F}(E(L(\bar{g}F, Z_X)))$$
$$= R_\delta(\bar{g}F)$$

<u>Remark</u>: If for every $F, F^* \in \mathcal{P}$ $\exists$ $\bar{g}(\cdot) \in \bar{G}$ s.t. $\bar{g}(F^*) = F$ (if such a condition holds we say that $\bar{G}$ acts transitively on $\mathcal{P}$) then

$$R_\delta(F) = \text{constant}, \quad \forall \ F \in \mathcal{P}.$$

Let

$\mathcal{D}_I$ = class of all invariant decision rules under a group of transformations.

We will see that in many situations a best decision rule with in the class $\mathcal{D}_I$ exists.

Location Invariant Estimation

$X = (X_1, \ldots, X_n)$ has a joint d.f. $F_\mu(x_1, \ldots, x_n) = F(x_1 - \mu, \ldots, x_n - \mu)$, $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $\mu \in \mathbb{R}$, for some d.f. F defined on $\mathbb{R}^n$. Let $f(y_1, \ldots, y_n)$ be the joint p.d.f. corresponding to $F(y_1, \ldots, y_n)$.

Then $X = (X_1, \ldots, X_n)$ has the p.d.f.

$$f_\mu(x_1, \ldots, x_n) = f(x_1 - \mu, \ldots, x_n - \mu), \quad x = (x_1, \ldots, x_n) \in \mathbb{R}^n, \ \mu \in \mathbb{R}$$

here

$\mathcal{P} = \{F_\mu : \mu \in \mathbb{R}\}$, where $F_\mu(x_1, \ldots, x_n) =$

$F(x_1-\mu, \dots, x_n-\mu)$, $\mu \in \mathbb{R}$,

<u>Goal:</u> To estimate μ under the loss function

$$L(\mu, a) \quad \mu \in \mathbb{R}, \quad a \in \mathcal{A} = \mathbb{R}.$$

Assume that, $\forall \mu \in \mathbb{R}$, $L(\mu, a)$ is convex in $a \in \mathbb{R}$.

Then one may consider only non-randomized estimators (or decision rules).

Consider location group defined on $\mathcal{X} = \mathbb{R}^n$

$$\mathcal{G} = \{g_c : c \in \mathbb{R}\},$$

where

$$g_c(\underline{x}) = g_c(x_1, \dots, x_n) = (x_1+c, \dots, x_n+c), \quad \underline{x} \in \mathcal{X} = \mathbb{R}^n, \quad c \in \mathbb{R}.$$

Then, for $c \in \mathbb{R}$, $\underline{Y} = g_c(\underline{X}) = (X_1+c, \dots, X_n+c)$ has p.d.f.

$$f(y_1-c-\mu, \dots, y_n-c-\mu) = f(y_1-(\mu+c), \dots, y_n-(\mu+c))$$

Since $\mu+c \in \mathbb{R}$, $\mathcal{P}$ is invariant under $\mathcal{G}$ with

$$\bar{g}_c(\mu) = \mu+c, \quad \mu \in \mathbb{R}, \ c \in \mathbb{R}.$$

Let $\tilde{g}_c(a) = a+c$, $a, c \in \mathbb{R}$. Then,

L is location invariant $\Leftrightarrow$ $L(\mu, a) = L(\bar{g}\mu, \tilde{g}a)$, $\forall g, \mu, a$

$\Leftrightarrow$ $L(\mu, a) = L(\mu+c, a+c)$, $\forall c \in \mathbb{R}, \mu, a \in \mathbb{R}$

$\Leftrightarrow$ $L(\mu, a) = L(0, a-\mu)$, $\forall c \in \mathbb{R}, \mu, a \in \mathbb{R}$

$\Leftrightarrow$ $L(\mu, a) = W(a-\mu)$, $a, \mu \in \mathbb{R}$,

for some convex function $W(\cdot)$ on $\mathbb{R}$.

We take $L(\mu, a) = W(a-\mu)$, $a, \mu \in \mathbb{R}$, for some convex and non-negative fn $W(\cdot)$.

A non-randomized estimator δ is invariant (location invariant) iff

$$\delta(g_c(\underline{X})) = \tilde{g}_c(\delta(\underline{X})), \quad \forall c \in \mathbb{R}$$

$\Leftrightarrow$ $\delta(X_1+c, \dots, X_n+c) = \delta(X_1, \dots, X_n)+c$, $\forall c \in \mathbb{R}$

$\Leftrightarrow$ $\delta(x_1+c, \dots, x_n+c) = \delta(x_1, \dots, x_n)+c$, $\forall \underline{x} \in \mathbb{R}^n, c \in \mathbb{R}$

For $\underline{x}=(x_1,\dots,x_n)\in\mathbb{R}^n$, let $d_i=x_i-x_n$, $i=1,\dots,n-1$, and let $D_i=X_i-X_n$, $i=1,\dots,n-1$, $\underline{d}=(d_1,\dots,d_{n-1})$ and $\underline{D}=(D_1,\dots,D_{n-1})$.

Note: (i) Sample mean $\bar{X}$ and the sample median M are location invariant.

(ii) A necessary and sufficient condition for an estimator δ to be location invariant is that $\delta(x_1,\dots,x_n)=\delta(x_1-x_n,\dots,x_{n-1}-x_n,0)+x_n=x_n+\delta(d_1,\dots,d_{n-1})$.

Lemma: Let δ_0 be a given location invariant estimator of μ. Then a necessary and sufficient condition for a non-randomized estimator δ to be location invariant is that there exists a function $u(\cdot)$ on $\mathbb{R}^{n-1}$ ($u\equiv a$ constant, for $n=1$) such that

$$\delta(\underline{x})=\delta_0(\underline{x})-u(\underline{d}),\qquad\forall\ \underline{x}\in\mathbb{R}^n.\qquad(*)$$

Proof. First suppose that δ is of form $(*)$. Then, $\forall\ c\in\mathbb{R}$,

$$\begin{aligned}\delta(x_1+c,\dots,x_n+c)&=\delta_0(x_1+c,\dots,x_n+c)-u(x_1+c-(x_n+c),\dots,x_{n-1}+c-(x_n+c))\\&=(\delta_0(x_1,\dots,x_n)+c)-u(x_1-x_n,\dots,x_{n-1}-x_n)\\&=\delta_0(x_1,\dots,x_n)-u(\underbrace{x_1-x_n}_{D_1},\dots,\underbrace{x_{n-1}-x_n}_{D_{n-1}})+c\\&=\delta(\underline{x})+c\end{aligned}$$

$\Rightarrow$ δ is location invariant.

Conversely suppose that δ is location invariant. Define

$$\tilde{u}(\underline{x})=\delta_0(\underline{x})-\delta(\underline{x}),\qquad\underline{x}\in\mathbb{R}^n$$

Then, $\forall\ \underline{x}=(x_1,\dots,x_n)\in\mathbb{R}^n$ and $c\in\mathbb{R}$

$$\tilde{u}(x_1+c,\dots,x_n+c)=\delta_0(x_1+c,\dots,x_n+c)-\delta(x_1+c,\dots,x_n+c)$$

$= (\delta_0(x_1, \dots, x_n) + c) - (\delta(x_1, \dots, x_n) + c)$

(δ_0 and δ_1 are location invariant)

$= \delta_0(x_1 \dots x_n) - \delta(x_1 \dots x_n)$

$= \tilde{u}(x_1 \dots x_n)$

Putting $c = -x_n$

$\tilde{u}(x_1 \dots x_n) = \tilde{u}(x_1 - x_n, \dots, x_{n-1} - x_n, 0)$

$= \tilde{u}(d_1 \dots d_{n-1}, 0)$

$= u(d_1, \dots, d_{n-1}), \quad \forall \underline{x} \in \mathbb{R}^n.$

$\Rightarrow u(d_1 \dots d_{n-1}) = \delta_0(\underline{x}) - \delta(\underline{x}), \quad \forall \underline{x} \in \mathbb{R}^n$

$\Rightarrow \delta(\underline{x}) = \delta_0(\underline{x}) - u(d_1 \dots d_{n-1}) \quad \forall \underline{x} \in \mathbb{R}^n$

We take our loss function to be

$L(\mu, a) = W(a - \mu) \quad a, \mu \in \mathbb{R},$

for some convex function $W(\cdot)$ defined on $[0, a]$.

Then the problem is invariant under G with

$\bar{g}_c(\mu) = \mu + c$ and $\tilde{g}_c(a) = a + c$

Clearly $\bar{G} = \{\bar{g}_c : c \in \mathbb{R}\}$ acts transitively on Ω.

Thus for any location invariant estimator δ

$R_\delta(\mu) = \underline{\text{Constant}} \quad \forall \mu \in \mathbb{R}.$

Let δ_0 be a given location invariant estimator/decision rule.

Define

$\mathcal{D}_I = \{\delta_u : \delta_u(\underline{x}) = \delta_0(\underline{x}) - u(d_1 \dots d_{n-1}), \underline{x} \in \mathbb{R}^n,$ for some function $u : \mathbb{R}^{n-1} \to \mathbb{R}\}$.

$\to$ class of all location invariant estimators / decision rules

Remark: The problem of finding the best location invariant estimator reduces to comparing constants instead of risk functions.

Definition Consider an invariant estimation problem in which all invariant estimators have constant risks (not depending on $F \in \mathcal{P} / \theta \in \Theta$). An invariant estimator δ is called the Minimum Risk Invariant Estimator (MRIE) if δ has the smallest constant risk among all invariant estimators.

Lemma Let $\psi: D \to \mathbb{R}$ be a convex and non-monotone function, where D is an interval in $\mathbb{R}$. Then ψ achieves its minimum on D and this minimum is unique if ψ is strictly convex. Moreover $\lim_{|x| \to \infty} \psi(x) = \infty$.

Theorem Let δ_0 be a location invariant estimator of μ with finite risk.

(a) Suppose that, for each $\underline{d} \in \mathbb{R}^{n-1}$, there exists $u_*(\underline{d})$ that minimizes

$$E_{\mu=0}\left(W(\delta_0(\underline{x}) - u_*(\underline{d})) \mid \underline{D} = \underline{d}\right)$$

over all functions $u(\cdot)$. Then the MRIE of μ exists and is given by

$$\delta_*(\underline{x}) = \delta_0(\underline{x}) - u_*(D).$$

(b) If W is convex and non-monotone then the function $u_*(\cdot)$ in (a) is ~~unique~~ exists and is unique if W is strictly convex.

(c) If δ_0 and $\underline{D}$ are independent then u_* is a constant that minimizes

$$E_{\mu=0}\left(W(\delta_0(\underline{x}) - u)\right)$$

If in addition the distribution of δ_0 is symmetric about μ and W is convex and even then $u_* = 0$.

Proof. (a) Let δ be an arbitrary location invariant estimator. Then

$$\delta(\underline{x}) = \delta_0(\underline{x}) - u(\underline{d}), \quad \underline{x} \in \mathcal{X} = \mathbb{R}^n, \quad \underline{d} = (x_1 - x_n, \ldots, x_{n-1} - x_n) = (d_1, \ldots, d_{n-1}),$$

for some function $u: \mathcal{X} \to \mathbb{R}$.

Since $\bar{G}$ acts transitively on Θ ($\forall\, \theta_1, \theta_2 \in \Theta$, $\theta_1 = \theta_2 + (\theta_1 - \theta_2) = g_{\theta_1 - \theta_2}(\theta_2)$), for $\mu \in \Theta$,

$$\begin{aligned}
R_\delta(\mu) &= R_\delta(0) \\
&= E_{\mu=0}\left(L(\delta(\underline{X}), 0)\right) \\
&= E_{\mu=0}\left(W(\delta(\underline{X}))\right) \\
&= E_{\mu=0}\left(W(\delta_0(\underline{X}) - u(\underline{D}))\right) \\
&= E_{\mu=0}\left(E_{\mu=0}\left(W(\delta_0(\underline{X}) - u(\underline{D})) \mid \underline{D}\right)\right) \\
&\geq E_{\mu=0}\left(E_{\mu=0}\left(W(\delta_0(\underline{X}) - u_*(\underline{D})) \mid \underline{D}\right)\right) \\
&= E_{\mu=0}\left(W(\delta_0(\underline{X}) - u_*(\underline{D}))\right) \\
&= E_{\mu=0}\left(W(\delta_*(\underline{X}) - 0)\right) \\
&= R_{\delta_*}(\mu)
\end{aligned}$$

(b) Since W is (strictly) convex and non-monotone. Then, for any $\underline{d} \in \mathcal{X}$,

$$\psi_0(a) := E_{\mu=0}\left(L(\delta_0(\underline{X}) - a) \mid \underline{D} = \underline{d}\right)$$

is (strictly) convex and non-monotone. Hence ψ_0 achieves its minimum (minimum is unique)

(c) If δ_0 and $\underset{\sim}{D}$ are independent then, for any $\delta(\underset{\sim}{X}) = \delta_0(\underset{\sim}{X}) - u(\underset{\sim}{D}) \in \theta_I$,

$$E_{\mu=0}\left(W\left(\delta_0(\underset{\sim}{X}) - u(\underset{\sim}{d})\right) \mid \underset{\sim}{D} = \underset{\sim}{d} \right)$$

$$= E_{\mu=0}\left(L\left(\delta_0(\underset{\sim}{X}) - u(\underset{\sim}{d})\right)\right)$$

and the minimizing $u(\cdot)$ will not depend on $\underset{\sim}{d}$.

Now suppose that $\delta_0(\underset{\sim}{X}) - \mu \overset{d}{=} \mu - \delta_0(\underset{\sim}{X})$, $\forall \mu$. Consider

$$\psi_1(u) = E_{\mu=0}\left(W\left(\delta_0(\underset{\sim}{X}) - u\right)\right)$$

$$= E_{\mu=0}\left(W\left(u - \delta_0(\underset{\sim}{X})\right)\right) \quad (W(t) = W(-t))$$

$$= E_{\mu=0}\left(W\left(u + \delta_0(\underset{\sim}{X})\right)\right) \quad \left(\text{under } \mu = 0 \;\; \delta_0(\underset{\sim}{X}) \overset{d}{=} -\delta_0(\underset{\sim}{X})\right)$$

$$= \psi_1(-u).$$

Thus, $\forall u \in \mathbb{R}$

$$\psi_1(0) = \psi_1\left(\frac{u + (-u)}{2}\right) \le \frac{\psi_1(u) + \psi_1(-u)}{2} = \psi_1(u)$$

Note: $\underset{\sim}{D}$ is ancillary. Thus if $\delta_0(\underset{\sim}{X})$ is a function of complete and sufficient statistic, then δ_0 and $\underset{\sim}{D}$ are independent.

Theorem: Under the assumptions of last theorem, suppose that $W(t) = t^2$, $t \in \mathbb{R}$, so that $L(\mu, a) = (a - \mu)^2$, $a \in \mathcal{A}$, $\mu \in \Theta$. Then

(a) The MRIE of μ is

$$\delta_P(\underset{\sim}{X}) = \frac{\int_{-\infty}^{\infty} t\, f(x_1 - t, \ldots, x_n - t)\, dt}{\int_{-\infty}^{\infty} f(x_1 - t, \ldots, x_n - t)\, dt},$$

which is also known as the Pitman estimator of μ.

(b) The MRIE δ_* is unbiased

Proof (a) Let $\delta_0(\underline{X}) = X_n$. Clearly $\delta_0 \in \mathcal{D}_I$. The MRIE is $\delta_*(\underline{X}) = X_n - u_*(\underline{D})$, where $u_*(\underline{d})$ minimizes

$$\psi_0(u(\underline{d})) = E_{\mu=0}\left(\left(\delta_0(\underline{X}) - u(\underline{d})\right)^2 \mid \underline{D} = \underline{d}\right),$$

for every $\underline{d} \in \mathcal{A}$.

Clearly minimizing $u(\underline{d})$ $(\underline{d} \in \mathcal{A})$ is

$$u_*(\underline{d}) = E_{\mu=0}(X_n \mid \underline{D} = \underline{d})$$

Define $D_n = X_n$. Then, under $\mu=0$, joint p.d.f. of $\underline{D}^* = (D_1, \dots, D_{n-1}, D_n)$ is

$$f_{\underline{D}^*}(\underline{d}) = f(d_1 + d_n, \dots, d_{n-1} + d_n, d_n), \quad \underline{d} \in \mathbb{R}^n$$

and the marginal p.d.f. of $\underline{D} = (D_1, \dots, D_{n-1})$ is

$$f_{\underline{D}}(\underline{d}) = \int_{-\infty}^{\infty} f(d_1 + t, \dots, d_{n-1} + t, t)\, dt, \quad \underline{d} \in \mathbb{R}^{n-1}$$

Under $\mu = 0$, the conditional p.d.f. of $X_n \,(= D_n)$ given $\underline{D} = \underline{d}$ is

$$f_{X_n \mid \underline{D}}(t \mid \underline{d}) = \frac{f(d_1 + t, \dots, d_{n-1} + t, t)}{\int_{-\infty}^{\infty} f(d_1 + u, \dots, d_{n-1} + u, u)\, du}, \quad t \in \mathbb{R}$$

and

$$u_*(\underline{d}) = \frac{\int_{-\infty}^{\infty} t\, f(d_1 + t, \dots, d_{n-1} + t, t)\, dt}{\int_{-\infty}^{\infty} f(d_1 + t, \dots, d_{n-1} + t, t)\, dt}$$

Thus the MRIE is

$$\delta_*(\underline{X}) = X_n - u_*(X_1 - X_n, \dots, X_{n-1} - X_n)$$

$$= X_n - \frac{\int_{-\infty}^{\infty} t\, f(x_1 - x_n + t, \dots, x_{n-1} - x_n + t, t)\, dt}{\int_{-\infty}^{\infty} f(x_1 - x_n + t, \dots, x_{n-1} - x_n + t, t)\, dt}$$

$$= X_n - \frac{\int_{-\infty}^{\infty} (x_n - u)\, f(x_1 - u, \dots, x_{n-1} - u, x_n - u)\, du}{\int_{-\infty}^{\infty} f(x_1 - u, \dots, x_{n-1} - u, x_n - u)\, du}$$

$$= \frac{\int_{-\infty}^{\infty} u\, f(x_1 - u, \dots, x_{n-1} - u, x_n - u)\, du}{\int_{-\infty}^{\infty} f(x_1 - u, \dots, x_{n-1} - u, x_n - u)\, du}.$$

(b) We have

$$\delta_*(\underline{X}) = X_n - E_{\mu=0}(X_n \mid \underline{D})$$

Note that

$$(X_1, \dots, X_n)_{\mu=0} \overset{d}{=} (X_1 - \mu, \dots, X_n - \mu)_\mu$$

Thus

$$E_\mu(\delta_*(\underline{X})) = E_\mu(X_n) - E_\mu(E_{\mu=0}(X_n \mid \underline{D}))$$

$$= E_\mu(X_n) - E_\mu(E_\mu(X_n - \mu \mid \underline{D}))$$

$$= E_\mu(X_n) - \cancel{E_\mu(E_\mu(\dots))}\; E_\mu(X_n - \mu)$$

$$= \mu.$$

Home Assignment:

Example 1 Let $X_1, \dots, X_n$ be i.i.d. $N(\mu, \sigma_0^2)$, where $\mu \in \mathbb{R} = \Theta$ is unknown and $\sigma_0 > 0$ is known. Consider estimation of μ under the loss function

$$L(\mu, a) = W(a - \mu), \quad a \in \mathcal{A} = \mathbb{R},\ \mu \in \Theta,$$

where $W : \mathbb{R} \to \mathbb{R}$ is a convex function with $W(t) = W(-t)\ \forall t \in \mathbb{R}$.

Show that $\delta_0(\underset{\sim}{X}) = \bar{X}$ is the MRIE of μ

Example 2 Let $X_1, \dots, X_n$ be i.i.d. $E(\mu, \theta_0)$ with p.d.f.

$$f_\mu(x) = \frac{1}{\theta_0} e^{-\frac{(x-\mu)}{\theta_0}}, \quad x > \mu,$$

where $\mu \in \Theta = \mathbb{R}$ is unknown and $\theta_0 > 0$ is known.
Consider estimation of μ under the loss function

$$L(\mu, a) = |a-\mu|, \quad a \in \mathcal{A} = \mathbb{R}, \ \mu \in \Theta.$$

Show that the MRIE of μ is

$$\delta_*(\underset{\sim}{X}) = X_{(1)} - \theta_0 \ln \frac{2}{n},$$

where $X_{(1)} = \min\{X_1, \dots, X_n\}$.

Example 3 Let $X_1, \dots, X_n$ be i.i.d. $U(\mu - \frac{1}{2}, \mu + \frac{1}{2})$, where $\mu \in \mathbb{R} = \Theta$ is unknown. Consider estimation of μ under the loss function

$$L(\mu, a) = (a-\mu)^2, \quad a \in \mathcal{A} = \mathbb{R}, \ \mu \in \Theta.$$

Find the MRIE of μ.

Scale Invariant Estimation

$$\underset{\sim}{X} = (X_1, \dots, X_n) \sim F_\theta \in \mathcal{P} = \{F_\theta : \theta > 0\},$$

where

$$F_\theta(x_1, \dots, x_n) = F\left(\frac{x_1}{\theta}, \dots, \frac{x_n}{\theta}\right), \quad \underset{\sim}{x} \in \mathbb{R}^n = \mathcal{X},$$

for some known d.f. $F(\cdot)$. Let $f(\cdot)$ be the pdf corresponding to F. So that pdf corresponding to $F_\theta(\cdot)$ is $f_\theta(x_1, \dots, x_n) = \frac{1}{\theta^n} f(\frac{x_1}{\theta}, \dots, \frac{x_n}{\theta})$, $\underset{\sim}{x} \in \mathcal{X}$.

Consider scale group of transformations

$$\mathcal{G} = \{g_c : c > 0\}.$$

where

$$g_c(x_1, \dots, x_n) = (cx_1, \dots, cx_n), \quad \underline{x} \in \mathcal{X} = \mathbb{R}^n, \quad c>0.$$

Consider estimation of θ^b, for some given $b \neq 0$.

Clearly $\mathcal{P}$ is invariant under G with $\bar{G} = \{\bar{g}_c : c>0\}$

$\bar{g}_c(\theta) = c\theta$, $\theta \in \Theta = (0, \infty)$, $c>0$.

Consider $\tilde{G} = \{\tilde{g}_c : c>0\}$, where $\tilde{g}_c(a) = c^b a$, $a \in \mathcal{A} = (0, \infty)$, $c>0$.

$$L \text{ is invariant} \iff L(\underline{\theta}, a) = L(\bar{g}_c(\underline{\theta}), \tilde{g}_c(a)) \ \forall c>0, \theta>0, a>0$$

$$\iff L(\theta, a) = L(c\theta, c^b a), \ \forall c>0, \theta>0, a>0$$

$$\iff L(\theta, a) = L\left(1, \frac{a}{\theta^b}\right), \ \forall \theta>0, a>0$$

$$\iff L(\theta, a) = W\left(\frac{a}{\theta^b}\right), \quad \theta>0, a>0,$$

for some function $W : (0, \infty) \to \mathbb{R}$

Assume that $W(t)$ is convex so that it suffices to consider only non randomized estimators.

$$\delta \text{ is invariant} \iff \delta(g_c(\underline{x})) = \tilde{g}_c(\delta(\underline{x})), \ \forall c>0, \underline{x} \in \mathcal{X}$$

$$\iff \delta(cx_1, \dots, cx_n) = c^b \delta(x_1, \dots, x_n), \ \forall c>0, \underline{x} \in \mathcal{X}$$

$$\iff \delta(x_1, \dots, x_n) = |x_n|^b \delta\left(\frac{x_1}{|x_n|}, \dots, \frac{x_{n-1}}{|x_n|}, \frac{x_n}{|x_n|}\right)$$

$$= |x_n|^b \, v(z_1, \dots, z_{n-1}, z_n) = \frac{|x_n|^b}{u(z_1, \dots, z_{n-1}, z_n)}$$

where $z_i = \frac{x_i}{x_n}$, $i = 1, \dots, n-1$ and $z_n = \frac{x_n}{|x_n|}$

(since $\left(\frac{x_1}{|x_n|}, \dots, \frac{x_{n-1}}{|x_n|}, \frac{x_n}{|x_n|}\right)$ and $\left(\frac{x_1}{x_n}, \dots, \frac{x_{n-1}}{x_n}, \frac{x_n}{|x_n|}\right)$ are 1-1 fns).

Thus

$\mathcal{D}_I =$ class of non-randomized scale invariant estimators

$= \{ \delta_u : \text{ for some } u : \mathbb{R}^n \to \mathbb{R}_+ \}$,

where

$$\delta_u(\underline{x}) = \frac{|x_n|^b}{u(z_1, \dots, z_{n-1}, z_n)}, \quad \underline{x} \in \mathbb{R}^n,$$

$z_i = \dfrac{x_i}{x_n}$, $i = 1, \dots, n-1$, and $z_n = \dfrac{x_n}{|x_n|}$.

Now let $\delta_0 \in \mathcal{D}_I$ be given and let $\delta \in \mathcal{D}_I$.

Then

$$\delta_0(\underline{x}) = \frac{|x_n|^b}{u_0(z_1, \dots, z_{n-1}, z_n)}, \quad \underline{x} \in \mathbb{R}^n$$

for some $u_0 : \mathbb{R}^n \to \mathbb{R}_+$

and

$$\delta(\underline{x}) = \frac{|x_n|^b}{u_1(z_1, \dots, z_{n-1}, z_n)}, \quad \underline{x} \in \mathbb{R}^n$$

for some $u_1 : \mathbb{R}^n \to \mathbb{R}_+$.

Then

$$\delta(\underline{x}) = \frac{u_0(\underline{z})}{u_1(\underline{z})} \delta_0(\underline{x})$$

$$= \frac{\delta_0(\underline{x})}{u(\underline{z})}, \quad \underline{x} \in \mathbb{R}^n \quad \dots\dots (*)$$

Conversely any estimator δ of form $(*)$ is in $\mathcal{D}_I$.

Thus, for a given $\delta_0 \in \mathcal{D}_I$,

$\mathcal{D}_I = \{ \delta_u : \text{ for some } u : \mathbb{R}^n \to \mathbb{R}_+ \}$,

where $\delta_u(\underline{x}) = \dfrac{\delta_0(\underline{x})}{u(\underline{z})}$, $\underline{x} \in \mathbb{R}^n$, $z_i = \dfrac{x_i}{x_n}$, $i = 1, \dots, n$

$z_n = \dfrac{x_n}{|x_n|}$.

Let $Z_i = \frac{X_i}{X_n}$, $i = 1, \dots, n-1$, $Z_n = \frac{X_n}{|X_n|}$ and

$\underline{Z} = (Z_1, \dots, Z_{n-1}, Z_n)$.

Note: $\underline{Z}$ is ancillary [Define $T_i = \frac{X_i}{\theta}$, $i = 1, \dots, n$ $\underline{T} = (T_1, \dots, T_n)$. The jt pdf of $\underline{T}$ is $f_{\underline{T}}(\underline{t}) = f(t_1, \dots, t_n)$, $\underline{t} \in \mathbb{R}^n$

Clearly $\bar{G}$ is acts transitively on.β. Thus for any $\delta \in \mathcal{D}_I$

$$R_\delta(\mu) = R_\delta(1), \quad \forall\ \mu \in \Theta.$$

Let $\delta_0 \in \mathcal{D}_I$ be fixed. Then, for any $\delta \in \mathcal{D}_I$

$$\delta(\underline{X}) = \frac{\delta_0(\underline{X})}{U(\underline{Z})},$$

for some $u: \mathbb{R}^{n-1} \to \mathbb{R}$, where $\underline{Z} = (Z_1, \dots, Z_{n-1}, Z_n)$, $Z_i = \frac{X_i}{|X_n|}$, $i = 1, \dots, n-1$ and $Z_n = \frac{X_n}{|X_n|}$.

$$\begin{aligned}
R_\delta(\theta) &= E_\theta\left(L(\theta, \delta(\underline{X}))\right) \\
&= E_\theta\left(W\left(\frac{\delta(\underline{X})}{\theta}\right)\right) \\
&= R_\delta(1) \\
&= E_{\theta=1}\left(W(\delta(\underline{X}))\right) \\
&= E_{\theta=1}\left(W\left(\frac{\delta_0(\underline{X})}{U(\underline{Z})}\right)\right) \\
&= E_{\theta=1}\left[E_{\theta=1}\left[W\left(\frac{\delta_0(\underline{X})}{U(\underline{Z})}\right) \Big| \underline{Z}\right]\right]
\end{aligned}$$

Thus the MRIE is

$$\delta_*(\underline{X}) = \frac{\delta_0(\underline{X})}{U_*(\underline{Z})},$$

where $U_*(\underline{z})$ minimizes (for given $\underline{Z} = \underline{z}$)

and $W(t) = (t-1)^2,\ t \in \mathbb{R}$

With $\delta_0(\underline{X}) = X_n^b$ the MRIE is given by

$$\delta_*(\underline{X}) = \frac{X_n^b}{u_*(\underline{Z})},$$

where $u_* \equiv u_*(\underline{z})$ minimizes

$$\psi(u) = E_{\theta=1}\left(\left(\frac{X_n^b}{u} - 1\right)^2 \Big| \underline{Z} = \underline{z}\right)$$
$$= \frac{1}{u^2} E_{\theta=1}(X_n^{2b} | \underline{Z} = \underline{z}) - \frac{2}{u} E_{\theta=1}(X_n^b | \underline{Z} = \underline{z}) + 1$$

$\Rightarrow$

$$\Rightarrow \frac{1}{u_*(\underline{z})} = \frac{E_{\theta=1}(X_n^b | \underline{Z} = \underline{z})}{E_{\theta=1}(X_n^{2b} | \underline{Z} = \underline{z})}$$

Under $\theta = 1$, the joint p.d.f. of $\underline{V} = (Z_1, \ldots, Z_{n-1}, X_n)$ is $h(z_1, \ldots, z_{n-1}, u) = |u|^{n-1} f(z_1 u, \ldots, z_{n-1} u, u)$, $(z_1, \ldots, z_{n-1}, u) \in \mathbb{R}^n$ and the conditional p.d.f. of X_n given $Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1}$ is

$$g_{\underline{z}}(u) = \frac{|u|^{n-1} f(z_1 u, \ldots, z_{n-1} u, u)}{\int_{-\infty}^{\infty} |t|^{n-1} f(z_1 t, \ldots, z_{n-1} t, t)\, dt},$$

For $z_n = -1$ (i.e. $x_n < 0$)

$$\frac{1}{u_*(\underline{z})} = \frac{\int_{-\infty}^{0} t^{n-1+b} f(z_1 t, \ldots, z_{n-1} t, t)\, dt}{\int_{-\infty}^{0} t^{n-1+2b} f(z_1 t, \ldots, z_{n-1} t, t)\, dt}$$

$$= \frac{1}{x_n^b} \frac{\int_0^{\infty} t^{n-1+b} f(x_1 t, \ldots, x_{n-1} t, x_n t)\, dt}{\int_0^{\infty} t^{n-1+2b} f(x_1 t, \ldots, x_{n-1} t, x_n t)\, dt}$$

Similarly, for $z_n = 1$, (i.e. $x_n > 0$),

$$\frac{1}{u_*(\underline{z})} = \frac{1}{x_n^b} \frac{\int_0^{\infty} t^{n-1+b} f(x_1 t, \ldots, x_{n-1} t, x_n t)\, dt}{\int_0^{\infty} t^{n-1+2b} f(x_1 t, \ldots, x_{n-1} t, x_n t)\, dt}$$

$\boxed{22/3}$

Thus we have the following result.

Theorem Suppose that the loss function is given by

$$L(\theta, a) = \left(\frac{a}{\theta^b} - 1\right)^2, \quad a \in \mathcal{A} = (0, \infty) = \Theta, \quad \theta \in \Theta,$$

i.e. $W(t) = (t-1)^2$, $t \in \mathbb{R}$. Let δ_0 be a given scale invariant estimator of θ^b.

Then the unique MRIE of θ^b is

$$\delta_*(\underline{X}) = \delta_0(\underline{X}) \frac{E_{\theta=1}(\delta_0(\underline{X}) \mid \underline{Z})}{E_{\theta=1}(\delta_0^2(\underline{X}) \mid \underline{Z})}$$

$$= \frac{\int_0^\infty t^{n+b-1} f(tx_1, \dots, tx_n)\, dt}{\int_0^\infty t^{n+2b-1} f(tx_1, \dots, tx_n)\, dt},$$

which is also known as the Pitman estimator of θ^b.

Home Assignment:

(1) Let $X_1, \dots, X_n$ be i.i.d. $N(0, \sigma^2)$, where $\sigma > 0$. Find the MRIE of σ^2 under the loss function $L(\sigma, a) = \left(\frac{a}{\sigma^2} - 1\right)^2$ (Consider scale group of transformations).

(2) Let $X_1, \dots, X_n$ be i.i.d. $U(0, \theta)$ where $\theta > 0$. Find the MRIE of θ under the loss function $L(\theta, a) = \left(\frac{a}{\theta} - 1\right)^2$, $a \in \mathcal{A} = (0, \infty)$, $\theta \in \Theta = (0, \infty)$.

(3) Let $X_1, \dots, X_n$ be i.i.d. $DE(0, \theta)$ where $\theta > 0$. Find the MRIE of θ under the loss function $L(\theta, a) = \left|\frac{a}{\theta} - 1\right|$, $a \in \mathcal{A} = (0, \infty)$, $\theta \in \Theta = (0, \infty)$

(4) Let $X_1, \dots, X_m$ be iid $N(\theta_1, 1)$ and let $Y_1, \dots, Y_n$ be iid $N(\theta_2, 1)$ where $\theta_1, \theta_2 \in \mathbb{R}$. Let $G = \{g_{c_1, c_2} : c_1, c_2 \in \mathbb{R}\}$ where $g_{c_1 c_2}(x_1, \dots, x_m; y_1, \dots, y_n) = (x_1 + c_1, \dots, x_m + c_1, y_1 + c_2, \dots, y_n + c_2)$, $c_1, c_2 \in \mathbb{R}$. Find the MRIE of $\theta = \theta_1 - \theta_2$ under the loss function $L(\theta, a) = (a - \theta)^2$, $a \in \mathcal{A} = \mathbb{R}$, $\theta \in \Theta = \mathbb{R}$.

Ram

Find the MRIE of θ under the loss function

$$L(\theta, a) = W(a-\theta), \quad a \in \mathcal{A}, \ \theta \in \Theta,$$

$$W(-t) = W(t) \quad t \in \mathbb{R}.$$

Where W is strictly convex and $W(-t) = W(t)$, $t \in \mathbb{R}$.

Also find MRIE of θ_1 under the loss function $L(\theta, a) = W(a-\theta)$, $a \in \mathcal{A}$, $\theta \in \Theta$, where W is even and convex.

(5) Let $X_1, \ldots, X_n$ be iid $E(0, \theta_1)$ and let $Y_1, \ldots, Y_n$ be iid $E(0, \theta_2)$, $\theta_1, \theta_2 \in (0, \infty) = \Theta$ and let $G = \{g_{c_1, c_2} : c_1, c_2 \in \mathbb{R}\}$ where $g_{c_1 c_2}(x_1, \ldots, x_n, y_1, \ldots, y_n) = (c_1 x_1, \ldots, c_1 x_n, c_2 y_1, \ldots, c_2 y_n)$, $c_1, c_2 \in \mathbb{R}$. Find the MRIE of $\theta = \frac{\theta_1}{\theta_2}$ under the loss function

$$L(\theta, a) = \left(\frac{a}{\theta} - 1\right)^2, \quad a \in \mathcal{A}, \ \theta \in \Theta.$$

Also derive the MRIE of θ_1 under the loss function

$$L(\theta_1, a) = \left(\frac{a}{\theta_1} - 1\right)^2, \quad a \in \mathcal{A}, \ \theta_1 \in (0, \infty).$$

Location - Scale Invariant Estimation

$\underline{X} = (X_1, \ldots, X_n) \sim F_{\underline{\theta}} \in \mathcal{P} = \{F_{\underline{\theta}} : \underline{\theta} = (\mu, \sigma) \in \mathbb{R} \times (0, \infty)\}$, and

$F_{\underline{\theta}}(x_1, \ldots, x_n) = F\left(\frac{x_1}{\theta}, \ldots, \frac{x_n}{\theta}\right)$, $\underline{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$

where F is known with Lebesgue p.d.f. $f(\cdot)$. Then the Lebesgue pdf of $F_{\underline{\theta}}$ is

$$f_{\theta}(\underline{x}) = \frac{1}{\sigma^n} f\left(\frac{x_1 - \mu}{\sigma}, \ldots, \frac{x_n - \mu}{\sigma}\right), \quad \underline{x} \in \mathbb{R}^n = \mathcal{X},$$

$$\underline{\theta} = (\mu, \sigma) \in \mathbb{R} \times (0, \infty) = \Theta$$

Consider

$$G = \{g_{b,c} : b > 0, \ c \in \mathbb{R}\},$$

Where

$$g_{b,c}(x_1, \ldots, x_n) = (bx_1 + c, \ldots, bx_n + c), \quad \underline{x} \in \mathbb{R}^n, \ b > 0, \ c \in \mathbb{R}.$$

$$\bar{G} = \{\bar{g}_{b,c} : b > 0, \ c \in \mathbb{R}\},$$

Where

$\bar{g}_{b,c}(\underline{\theta}) = (b\mu+c,\ b\sigma), \quad b>0 \quad c\in\mathbb{R}.$] $\bar{G}$ acts transitively on Θ

Goal I: For given $r\neq 0$ to estimate σ^r.

Consider
$\tilde{g}_{b,c}(\underline{\theta}) = b^r\sigma^r, \quad b>0 \quad c\in\mathbb{R}.$

For $L(\underline{\theta}, a)$ to be invariant

$$L((\mu,\sigma), a) = L(\bar{g}_{b,c}(\mu,\sigma), \tilde{g}_{b,c}(a)), \quad \forall\ \underline{\theta}\in\mathbb{R}\times(0,\infty)$$

$$\Leftrightarrow \quad L((\mu,\sigma), a) = L(b\mu+c,\ b^r a), \quad \forall\ \underline{\theta}\in\mathbb{R}\times(0,\infty)$$

$$\Leftrightarrow \quad L((\mu,\sigma), a) = L(0, \tfrac{a}{\sigma^r}), \quad \underline{\theta}\in\mathbb{R}\times(0,\infty)$$

(taking $b=\frac{1}{\sigma}$ and $c=-\frac{\mu}{\sigma}$)

$\underline{\theta}=(\mu,\sigma)\in\mathbb{R}\times(0,\infty)$.

$$\Leftrightarrow \quad L((\mu,\sigma), a) = W(\tfrac{a}{\sigma^r}), \text{ say,}$$

for some $W: [0,\infty) \to [0,\infty)$.

We take

$$L((\mu,\sigma), a) = W(\tfrac{a}{\sigma^r}), \quad \underline{\theta}=(\mu,\sigma)\in\mathbb{R}\times(0,\infty)$$

Where $W: [0,\infty)\to[0,\infty)$ is convex. Then the given decision problem is invariant with $\bar{g}_{b,c}(\mu,\sigma)=(b\mu+c, b\sigma)$

It suffices to consider non-randomized invariant decision rules.

Note that $\bar{G}$ acts transitively on Θ. Let

$\mathcal{D}_I$ = class of all (location-scale) invariant estimators (decision rules).

Then

$$\delta\in\mathcal{D}_I \Leftrightarrow \delta(g_{b,c}(\underline{x})) = \tilde{g}_{b,c}(\delta(\underline{x})), \quad \forall\ b>0,\ c\in\mathbb{R},\ \underline{x}\in\mathbb{R}^n$$

$$\Leftrightarrow \delta((bx_1+c,\dots,bx_n+c)) = b^r\,\delta(x_1,\dots,x_n), \quad \forall\ b>0,\ c\in\mathbb{R},\ \underline{x}\in\mathbb{R}^n$$

$\Rightarrow \delta(x_1+c, \ldots, x_n+c) = \delta(x_1, \ldots, x_n), \quad \forall\, c \in \mathbb{R},\ \underline{x} \in \mathbb{R}^n$

$\Rightarrow \delta(\underline{x})$ is a function of $\underline{d} = (d_1, \ldots, d_{n-1})$, where $d_i = x_i - x_n$, $i = 1, \ldots, n-1$.

$\Rightarrow \delta(\underline{x}) = V(d_1, \ldots, d_{n-1})$, for some function $V: \mathbb{R}^{n-1} \to \mathbb{R}$

Also

$$\delta(bx_1+c, \ldots, bx_n+c) = b^r \delta(x_1, \ldots, x_n), \quad \forall\, b>0,\ c \in \mathbb{R},\ \underline{x} \in \mathbb{R}^n$$

$$\Rightarrow V(bd_1, \ldots, bd_{n-1}) = b^r V(d_1, \ldots, d_{n-1})$$

$$\Rightarrow \delta(x_1, \ldots, x_n) = V(d_1, \ldots, d_{n-1})$$
$$= \frac{V(bd_1, \ldots, bd_{n-1})}{b^r}, \quad \forall\, b>0,\ \underline{x} \in \mathbb{R}^n, \quad (*)$$

for some function $V: \mathbb{R}^{n-1} \to \mathbb{R}$.

Conversely if δ is of the form $(*)$ then $\delta \in \mathcal{D}_I$

The p.d.f. of $\underline{D} = (D_1, \ldots, D_{n-1})$ $(D_i = X_i - X_n,\ i = 1, \ldots, n-1)$ is

$$f_{\underline{D}}(\underline{d}) = \frac{1}{\sigma^{n-1}} \int_{-\infty}^{\infty} f\left(\frac{d_1}{\sigma}+t, \ldots, \frac{d_{n-1}}{\sigma}+t\right) dt$$

$$= \frac{1}{\sigma^{n-1}}\, g\left(\frac{d_1}{\sigma}, \ldots, \frac{d_{n-1}}{\sigma}\right), \quad \underline{d} \in \mathbb{R}^{n-1}$$

It follows from the results for scale families that if $\delta_0(\cdot)$ is any finite risk scale invariant estimator of σ^r based on $\underline{D}$, then the MRIE of σ^r is

$$\delta_*(\underline{D}) = \frac{\delta_0(\underline{D})}{u_*(\underline{W})},$$

where $\underline{W} = (W_1, \ldots, W_{n-1})$, $W_i = \frac{D_i}{D_{n-1}}$, $i=1, \ldots, n-1$,

$W_{n-1} = \frac{D_{n-1}}{|D_{n-1}|}$, $i=1,\ldots,n-1$, $W_{n-1} = \frac{D_{n-1}}{|D_{n-1}|}$ and $u_*(\underline{w})$

minimises

$$E_{\sigma=1}\left(W\left(\frac{\delta_0(\underline{D})}{u_*(\underline{w})} \right) \Big| \underline{W} = \underline{w} \right).$$

Under $W(\lambda) = (\lambda-1)^2$, $\lambda \in \mathbb{R}$ (i.e., $L(\underline{\theta}, a) = \left(\frac{a}{\sigma}-1\right)^2$, $a, \sigma \in (0,\infty)$)

$$u_*(\underline{w}) = \frac{E_{\mu=0, \sigma=1}\left(\delta_0^2(\underline{D}) \,|\, \underline{W} = \underline{w} \right)}{E_{\mu=0, \sigma=1}\left(\delta_0(\underline{D}) \,|\, \underline{W} = \underline{w} \right)}.$$

Note: $\underline{W}$ is ancillary. Thus if $\delta_0(\underline{D})$ is a function of complete sufficient statistic, then $\delta_0(\underline{D})$ and $\underline{W}$ are independent.

Assignment Problem: (1) Let $X_1, \ldots, X_n$ be i.i.d. $N(\mu, \sigma^2)$ where $\underline{\theta} = (\mu, \sigma) \in \mathbb{R} \times (0,\infty) = \Theta$ is unknown. For estimating σ^2, under the loss function $L(\underline{\theta}, a) = \left(\frac{a}{\sigma^2}-1\right)^2$, $a, \sigma \in (0,\infty)$, find best affine invariant estimator.

(2) Let $X_1, \ldots, X_n$ be i.i.d. $E(\mu, \sigma)$ where $(\mu, \sigma) \in \mathbb{R}\times(0,\infty)$ is unknown. For estimating σ, under the loss function $L(\underline{\theta}, a) = \left(\frac{a}{\sigma}-1\right)^2$, $a, \sigma \in (0,\infty)$, find the best affine invariant estimator.

(3) Let $X_1, \ldots, X_n$ be i.i.d $U(\mu-\sigma, \mu+\sigma)$, where $(\mu, \sigma) \in \mathbb{R}\times(0,\infty)$ is unknown. For estimating σ, under the loss function $L(\underline{\theta}, a) = \left(\frac{a}{\sigma}-1\right)^2$, $a, \sigma \in (0,\infty)$, find the best affine invariant estimator.

To estimate μ

Goal II.

$\bar{g}_{b,c}(\underline{\theta}) = (b\mu+c, b\sigma)$

$\tilde{g}_{b,c}(a) = ba+c.$

For $L(\underline{\theta}, a)$ to be invariant

$$L(\underline{\theta}, a) = L(\bar{g}_{b,c}(\underline{\theta}), \tilde{g}_{b,c}(a)) \quad \forall\, b>0,\ c\in\mathbb{R},\ a\in\mathbb{R},\ \underline{\theta}\in\Theta$$

$$\Rightarrow L((\mu,\sigma), a) = L((b\mu+c, b\sigma), (ba+c)) \quad \forall\, b>0,\ c\in\mathbb{R},\ a\in\mathbb{R},\ \underline{\theta}\in\Theta$$

$$\Leftrightarrow L((\mu,\sigma), a) = L\left((0,1), \frac{a-\mu}{\sigma}\right) \qquad \left(c=-\frac{\mu}{\sigma},\ b=\frac{1}{\sigma}\right)$$

$$\Leftrightarrow L((\underline{\theta}, a)) = W\left(\frac{a-\mu}{\sigma}\right), \quad \underline{\theta}\in\Theta,\ a\in\mathbb{R}.$$

We assume that $W(\cdot)$ is convex so that we can restrict to the class $\mathcal{D}_I$ of non randomized decision rules.

Note that $\bar{G}$ acts transitively on Θ. Thus the risk of any affine invariant decision rule (estimator) is constant.

$$\delta\in\mathcal{D}_I \Leftrightarrow \delta((bx_1+c,\ldots,bx_n+c)) = b\,\delta(x_1,\ldots,x_n)+c, \quad \forall\, b>0,\ c\in\mathbb{R},\ \underline{x}\in\mathbb{R}^n$$

$$\Rightarrow \delta(x_1+c,\ldots,x_n+c) = \delta(x_1,\ldots,x_n)+c, \quad \forall\, c\in\mathbb{R},\ \underline{x}\in\mathbb{R}^n$$

$$\Rightarrow \delta(x_1,\ldots,x_n) = x_n + \delta(x_1-x_n,\ldots,x_{n-1}-x_n)$$

$$= x_n + \psi(D_1,\ldots,D_{n-1}) \quad \text{----- } (*)$$

Also $\delta(bx_1,\ldots,bx_n) = b\,\delta(x_1,\ldots,x_n)$

$$\Rightarrow bx_n + \psi(bD_1,\ldots,bD_{n-1}) = bx_n + b\,\psi(D_1,\ldots,D_{n-1})$$

$$\Rightarrow \psi(bD_1,\ldots,bD_{n-1}) = b\,\psi(D_1,\ldots,D_{n-1}), \quad \forall\, b>0$$

$$\Rightarrow \cancel{\psi(D_1,\ldots,D_{n-1}) = |D_{n-1}|\,\psi\left(\frac{D_1}{|D_{n-1}|},\ldots,\frac{D_{n-1}}{|D_n|}\right)}$$

$\Rightarrow \psi(\cdot)$ is some affine equivariant estimator of σ

- Let $\delta_1(\cdot)$ be a given affine ~~eq~~ invariant estimator of σ. Then

$$\psi(D_1, \ldots, D_{n-1}) = U_1(\underline{W})\, \delta_1(\underline{D}),$$

for some function $U_1 : \mathbb{R}^{n-1} \to \mathbb{R}$.

Thus, for a given affine invariant estimator δ_1 of σ,

(class of all non-randomized affine invariant estimators of μ)

$\delta \in \Theta_I \Rightarrow \delta(x_1, \ldots, x_n) = x_n + U_1(\underline{W})\, \delta_1(\underline{D})$. ---- (*)

Conversely if δ is any estimator of form (*) then $\delta \in \Theta_I$ (class of all non-randomized affine estimators of μ)

Let δ_0 be a given affine invariant estimator of μ and let δ_1 be a given affine equivariant estimator of σ.

If $\delta \in \Theta_I$, then

$$\delta(\underline{X}) = X_n + U_1(\underline{W})\, \delta_1(\underline{D})$$

$$\delta_0(\underline{X}) = X_n + U_0(\underline{W})\, \delta_1(\underline{D}) \qquad (\delta_0 \in \Theta_I)$$

for some functions $U_1(\cdot)$ and $U_0(\cdot)$.

$$\Rightarrow \quad \delta(\underline{X}) = \delta_0(\underline{X}) - (U_0(\underline{W}) - U_1(\underline{W}))\, \delta_1(\underline{D})$$

$$\Rightarrow \quad \delta(\underline{X}) = \delta_0(\underline{X}) - U(\underline{W})\, \delta_1(\underline{D}),$$

for some function $U(\cdot)$.

Lemma: Let δ_0 be any affine invariant estimator of μ and let δ_1 be any affine invariant estimator of σ (i.e. $\delta_0(bx_1+c, \ldots, bx_n+c) = b\,\delta_0(x_1, \ldots, x_n) + c$, and $\delta_1(bx_1+c, \ldots, bx_n+c) = b\,\delta_1(x_1, \ldots, x_n) + c$ $\forall\, b>0$, $c \in \mathbb{R}$, $\underline{x} \in \mathbb{R}^n$).

Then an estimator δ of μ is affine invariant iff there is a function $U(\cdot)$ such that

$$\delta(\underline{X}) = \delta_0(\underline{X}) - U(\underline{W})\, \delta_1(\underline{X})$$

where $\underline{W} = (W_1, \ldots, W_{n-1})$, $W_i = \dfrac{D_i}{D_{n-1}}$, $i = 1, \ldots, n-2$, $W_{n-1} = \dfrac{D_{n-1}}{|D_{n-1}|}$ and $D_i = X_i - X_n$, $i = 1, \ldots, n-1$.

Since $\mathcal{G}$ acts transitively on $\underset{\sim}{\theta}$ the risk function of any $\delta \in \mathcal{D}_I$ is constant. Thus we have the following theorem.

Theorem Let δ_0 be a given affine invariant estimator of μ and let δ_1 be a given affine invariant estimator of σ. Then an MRIE of μ is

$$\delta_*(\underset{\sim}{X}) = \delta_0(\underset{\sim}{X}) - u_*(\underset{\sim}{W})\,\delta_1(\underset{\sim}{X}),$$

where, for every fixed $\underset{\sim}{w}$, $u_* \equiv u_*(\underset{\sim}{w})$ minimizes

$$\psi(u) = E_{\mu=0,\sigma=1}\left(W\left(\delta_0(\underset{\sim}{X}) - u\,\delta_1(\underset{\sim}{X})\right) \big| \underset{\sim}{W} = \underset{\sim}{w}\right).$$

Since $\underset{\sim}{W}$ is ancillary it may be helpful to choose δ_0 and δ_1 in a manner that they are functions of complete and sufficient statistic so that (δ_0, δ_1) and $\underset{\sim}{W}$ are independent.

Corollary Under the assumptions of above theorem let $W(t) = t^2$, $t \in \mathbb{R}$, so that

$$L(\underset{\sim}{\theta}, a) = \frac{(a-\mu)^2}{\sigma^2}, \quad a, \mu \in \mathbb{R}, \ \sigma > 0.$$

Then the MRIE is

$$\delta_*(\underset{\sim}{X}) = \delta_0(\underset{\sim}{X}) - u_*(\underset{\sim}{W})\,\delta_1(\underset{\sim}{X}),$$

where, for fixed $\underset{\sim}{w}$

$$u_* \equiv u_*(\underset{\sim}{w}) = \frac{E_{\mu=0,\sigma=1}\left(\delta_0(\underset{\sim}{X})\,\delta_1(\underset{\sim}{X}) \big| \underset{\sim}{W} = \underset{\sim}{w}\right)}{E_{\mu=0,\sigma=1}\left(\delta_1^2(\underset{\sim}{X}) \big| \underset{\sim}{W} = \underset{\sim}{w}\right)}$$

Assignment Problems

(1) Let $X_1, \dots, X_n$ be iid from $N(\mu, \sigma^2)$ where $\mu \in \mathbb{R}$ and $\sigma > 0$ are unknown. Consider estimation of μ under the loss function

$$L(\underset{\sim}{\theta}, a) = W\left(\frac{a-\mu}{\sigma}\right), \quad a, \mu \in \mathbb{R}, \ \sigma > 0,$$

where $W(\cdot)$ is convex and even. Find the MRIE of μ.

(2) Let $X_1, \dots, X_n$ be i.i.d. from $U(\mu - \frac{\sigma}{2}, \mu + \frac{\sigma}{2})$, where $\mu \in \mathbb{R}$ and $\sigma > 0$ are unknown. Consider estimation of $\mu \in \mathbb{R}$ under the loss function

$$L(\underset{\sim}{\theta}, a) = \left(\frac{a-\mu}{\sigma}\right)^2, \quad a, \mu \in \mathbb{R}; \ \sigma > 0$$

Find the MRIE of μ. Can your findings be generalized to more general loss function.

(3) Same as (1) with $N(\mu, \sigma^2)$ replaced by $E(\mu, \sigma)$ and ~~W replaced B~~ $W(t) = t^2$, $t \in \mathbb{R}$.

(4) Same as Problem 1 except that you estimate pth ($p \in (0,1)$) quantile of $N(\mu, \sigma^2)$ and $W(t) = t^2$.

<u>Example</u> (Two Sample Location-Scale Problem) The joint p.d.f. of $\underset{\sim}{X} = (X_1, \dots, X_m)$, $\underset{\sim}{Y} = (Y_1, \dots, Y_n)$. The joint p.d.f. of $\underset{\sim}{Z} = (\underset{\sim}{X}, \underset{\sim}{Y})$ is

$$\frac{1}{\sigma_1^m \sigma_2^n} f\left(\frac{x_1-\mu_1}{\sigma_1}, \dots, \frac{x_m-\mu_1}{\sigma_1}, \frac{y_1-\mu_2}{\sigma_2}, \dots, \frac{y_n-\mu_2}{\sigma_2}\right),$$

where $\underset{\sim}{\theta} = (\mu_1, \mu_2, \sigma_1, \sigma_2) \in \Theta = \mathbb{R}^2 \times \mathbb{R}_+^2$ is unknown.

Consider

$$G = \{g_{a,b,c,d} : a > 0, \ c > 0, \ b \in \mathbb{R}, \ d \in \mathbb{R}\},$$

where

$$g_{a,b,c,d}(\underset{\sim}{x}, \underset{\sim}{y}) = (ax_1+b, \dots, ax_m+b, cy_1+d, \dots, cy_n+d)$$

Then β is invariant under $\bar{\bar{G}}$ with $\bar{G} = \{\bar{g}_{\underline{a},\underline{b},\underline{c},\underline{d}} : a>0, c>0, b\in\mathbb{R}, d\in\mathbb{R}\}$ and

$$\bar{g}_{\underline{a},\underline{b},\underline{c},\underline{d}}(\underline{\theta}) = (a\mu_1+b,\ c\mu_2+d,\ a\sigma_1,\ c\sigma_2)$$

Following possibilities may arise: (i) $\sigma_1=\sigma_2$, $a=c$ (ii) $\mu_1=\mu_2$, $a=c$, $d=b$

__Case I:__ Estimation of $\Delta = \mu_1-\mu_2$ when $\sigma_1=\sigma_2=\sigma$ (say)

__Case II:__ Estimation of $\psi = \sigma_2/\sigma_1$.