Equivariance Estimation

We restrict attention to estimators that satisfy restrictions. The location family of dist's is given by

$$P_\theta = \{f(x-\theta),\ \theta \in \Theta\}.$$

In general, let $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ be a family of dist's. Let $\mathcal{G} = \{g\}$ be a class of transformations of the sample space, i.e $g : \Omega \longrightarrow \Omega$.

Def.: When $X \sim P_\theta$ and $X' = gX \sim P_{\theta'} \in \mathcal{P}$, then $\mathcal{P}$ is called invariant under $\mathcal{G}$.

Example: Let F be a fixed dist. on $\mathbb{R}^n$, $\mathcal{P} = \{F(\underset{\sim}{x} - \theta) : \theta \in \mathbb{R}\}$ with $g_a(\underset{\sim}{x}) = \underset{\sim}{x} + a$. Then $\mathcal{P}$ is invariant under $\mathcal{G}$. Let us take F be the cdf of $N(\theta, 1)$ with pdf

$$p_\theta = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\theta)^2}, \quad \theta \in \Theta = \mathbb{R}$$

The induced maps on Θ are given by

$$g_a : X \sim P_\theta \longrightarrow X + a \sim P_{\theta+a} = P_{\theta'}, \quad \theta' = \theta + a$$

That is $\bar{g}\theta = \theta' = \theta + a$.

The groups $\mathcal{G}, \bar{\mathcal{G}}$ are related by

$$P_\theta(gX \in A) = P_{\bar{g}\theta}(X \in A)$$

Since $X \sim P_\theta \Rightarrow gX \sim P_{\bar{g}\theta}$, or equivalently, $P_\theta(\bar{g}^{-1}(A)) = P_{\bar{g}\theta}(A)$

Def.: An estimator δ for the location θ in a location family is called <u>equivariant</u> (location equivariant) if
$$\delta(\underset{\sim}{x} + a\underset{\sim}{1}) = \delta(\underset{\sim}{x}) + a \quad \forall\, a \in \mathbb{R},\ \underset{\sim}{x} \in \mathbb{R}^n.$$

Def.: An estimator δ is a location invariant if
$$\delta(\underset{\sim}{x} + a\underset{\sim}{1}) = \delta(\underset{\sim}{x}) \quad \forall\, a \in \mathbb{R}\ \&\ \underset{\sim}{x} \in \mathbb{R}^n$$

<u>Remark:</u> Let δ_1 & δ_2 be 2 location equivariant, then $\delta = \delta_1 - \delta_2$ is location invariant. This is due to
$$\begin{aligned}\delta(\underset{\sim}{x} + a\underset{\sim}{1}) &= \delta_1(\underset{\sim}{x} + a\underset{\sim}{1}) - \delta_2(\underset{\sim}{x} + a\underset{\sim}{1})\\ &= (\delta_1(\underset{\sim}{x}) + a) - (\delta_2(\underset{\sim}{x}) + a)\\ &= \delta_1(\underset{\sim}{x}) - \delta_2(\underset{\sim}{x})\\ &= \delta(\underset{\sim}{x})\end{aligned}$$

<u>Def:</u> Let $\mathcal{P}$ be a location family and $L(\theta, d)$ a loss function. The loss function is called location invariant if
$$L(\theta + a, d + a) = L(\theta, d), \text{ for } \theta \in \Theta,\ d, a \in \mathbb{R}.$$
If L is location invariant, the estimation problem is called location invariant.

<u>Theorem:</u> Assume δ is an equivariant estimator in a location invariant problem (loss is inv.). Then the bias, variance and risk of δ are all constant (do not depend on θ)

<u>Proof:</u> $\text{Bias}(\delta) = E_\theta\, \delta(\underset{\sim}{x}) - \theta = E_0(\delta(\underset{\sim}{x} + \theta)) - \theta$
$= E_0\, \delta(\underset{\sim}{x})$ does not depend on θ.

The risk is $E_\theta\left[L(\theta, \delta(\underset{\sim}{X}))\right] \overset{\text{Inv.}}{\underset{\text{model}}{=}} E_0\left[L(\theta, \delta(\underset{\sim}{X} + \theta\underset{\sim}{1}))\right]$

$\overset{\text{Equiv.}}{=} E_0\left[L(\theta, \delta(\underset{\sim}{X}) + \theta)\right]$

$\overset{\text{Inv.}}{\underset{\text{loss}}{=}} E_0\left[L(0, \delta(\underset{\sim}{X}))\right]$

Let us provide a characterization of equivariant estimators.

<u>Lemma</u>: Let δ_0 be a fixed equivariant estimator. The set of (location) equivariant estimators is given by

$$\Delta = \left\{ \delta = \delta_0 + u : u(\underset{\sim}{x}) = u(\underset{\sim}{x} + a),\ \forall \underset{\sim}{x} \in \mathbb{R}^n, a \in \mathbb{R} \right\}$$

<u>Proof</u>: Define $\delta(\underset{\sim}{x}) = \delta_0(\underset{\sim}{x}) + u(\underset{\sim}{x})$

$$\begin{aligned} \delta(\underset{\sim}{x} + a) &= \delta_0(\underset{\sim}{x} + a) + u(\underset{\sim}{x} + a) \\ &= \delta_0(\underset{\sim}{x}) + a + u(\underset{\sim}{x}) \\ &= \delta(\underset{\sim}{x}) + a \end{aligned}$$

i.e. δ is equivariant.

<u>Lemma</u>: The set of invariant loss function is

$$\{ L(\theta, d) = \rho(d - \theta) : \rho : \mathbb{R} \longrightarrow \mathbb{R}^+, \rho(0) = 0 \}$$

<u>Proof</u>: $\Rightarrow$ Assume L is invariant, i.e. $L(\theta + a, d + a) = L(\theta, d)$, for all $a \in \mathbb{R}$. Putting $a = -\theta$ to get

$$L(0, d - \theta) = \rho(d - \theta)$$

$\Leftarrow$ Assume $L(\theta, d) = \rho(d - \theta)$. Then

$$\begin{aligned} L(\theta + a, d + a) &= \rho(d + a - \theta - a) \\ &= \rho(d - \theta) \\ &= L(\theta, d) \end{aligned}$$

$\Rightarrow$ L is invariant.

The following is a useful strategy for deriving minimum risk equivariant (MRE) estimator.

Theorem: If the decision problem is location invariant, δ_0 is location equivariant with finite risk, and if $v^*(\underset{\sim}{y})$ minimizes $E_0\left[\rho(\delta_0 - v(\underset{\sim}{y})) \mid \underset{\sim}{Y} = \underset{\sim}{y}\right]$ for each y, then an MRE estimator is $\delta_0(\underset{\sim}{X}) - v^*(\underset{\sim}{Y})$, where
$\underset{\sim}{Y} = (X_1 - X_n, \ldots, X_{n-1} - X_n)$.

Proof: Consider $\delta^*(\underset{\sim}{X}) = \delta_0(\underset{\sim}{X}) - v^*(\underset{\sim}{Y})$. Clearly δ^* is equivariant. Let $\delta(\underset{\sim}{X}) = \delta_0(\underset{\sim}{X}) - v(\underset{\sim}{Y})$ be arbitrary equivariant estimator. Using the fact that the risk functions of equiv. estimators are constant, we need to minimize

$$R(\theta, \delta) = R_0(\delta) = E_0\left[\rho(\delta_0(\underset{\sim}{X}) - v(\underset{\sim}{Y}) - \theta)\right]$$

Now,
$$\begin{aligned} R(\theta,\delta) &= E_0\left[\rho(\delta_0(\underset{\sim}{X}) - v(\underset{\sim}{Y}))\right] \\ &= E_0 E_0\left[\rho(\delta_0(\underset{\sim}{X}) - v(\underset{\sim}{Y})) \mid \underset{\sim}{Y} = \underset{\sim}{y}\right] \\ &\geq E_0 E_0\left[\rho(\delta_0(\underset{\sim}{X}) - v^*(\underset{\sim}{Y})) \mid \underset{\sim}{Y} = \underset{\sim}{y}\right] \\ &= E_0\left[\rho(\delta_0(\underset{\sim}{X}) - v^*(\underset{\sim}{Y}))\right] \\ &= R(\theta, \delta^*) \end{aligned}$$

$\Rightarrow \delta^*$ is MRE estimator.

Example: Let $X_1, \dots, X_n \overset{iid}{\sim} f_\theta(x) = e^{-(x-\theta)}$, $x > \theta \in \mathbb{R}$. Obtain the MRE estimator of θ.

$$L(\theta) = f(\underset{\sim}{x};\theta) = e^{-\sum^n x_i} e^{n\theta} I_{\{X_{(1)} > \theta\}}(\underset{\sim}{x}) \Rightarrow X_{(1)} \text{ is minimal suff.}$$

Consider $\underset{\sim}{Y} = \begin{pmatrix} X_{(2)} - X_{(1)} \\ X_{(3)} - X_{(2)} \\ \vdots \\ X_{(n)} - X_{(n-1)} \end{pmatrix}$ spacings of OS_s are indep.

It is easy to show $X_{(j+1)} - X_{(j)} \sim Exp(n-j)$ [Try to show that]

- $\delta_0(\underset{\sim}{X}) = X_{(1)}$ is equivariant
- $R(\theta, \delta) = E_{\theta=0}\left[\rho(X_{(1)} - v) \mid \underset{\sim}{Y}\right]$

$$= E_{\theta=0}\left[\rho(X_{(1)} - v)\right]$$

$$= \int_0^\infty \rho(x-v) \cdot n e^{-nx} \, dx$$

- when $\rho(u) = u^2 \Rightarrow R(\theta, \delta) = \int_0^\infty (x-v)^2 \, dF(x)$

$\Rightarrow v^* = \int x \, dF(x) = EX = \frac{1}{n}$ (our case)

$\Rightarrow \boxed{\delta^*(\underset{\sim}{X}) = X_{(1)} - \frac{1}{n}}$

- when $\rho(u) = |u| \Rightarrow R(\theta, \delta) = \int_0^\infty |x-u| \, dF(x)$

$\Rightarrow v^* = med(X) = \frac{\ln 2}{2}$ in our case

$\Rightarrow \boxed{\delta^*(\underset{\sim}{X}) = X_{(1)} - \frac{\ln 2}{2}}$

COR.: Under the assumption of previous theorem, we have

(i) If $\rho(u) = u^2$ (quadratic loss function), then
$$\delta^*(\underset{\sim}{X}) = \delta_0(\underset{\sim}{X}) - E_0(\delta_0(\underset{\sim}{X}) \mid \underset{\sim}{Y})$$

(ii) If $\rho(u) = |u|$ (absolute loss func.), then

$$\delta^*(\underset{\sim}{X}) = \delta_0(\underset{\sim}{X}) - \text{med}(\delta_0(\underset{\sim}{X}) \mid \underset{\sim}{Y})$$

where $\text{med}(\delta_0(\underset{\sim}{X}) \mid Y)$ is the conditional median of $\delta_0(\underset{\sim}{X})$ given $\underset{\sim}{Y} = \underset{\sim}{y}$.

COR.: Under the previous assumption, suppose that ρ is convex and not monotone. Then an MRE estimator of θ exists; it is unique if ρ is strictly convex.

Example: Let us consider $n = 1$ (one observation). Then since an arbitrary equivariant estimator can be written as
$$\delta(X) = \delta_0(X) - v(X - X) = \delta_0(X) + c$$
$$= X + c \text{, by taking } \delta_0(X) = X$$

$$v^* = \min E_0 \rho(X - v)$$

(i) If $\rho(u) = u^2$, then $v^* = E_0 X \Rightarrow$ MRE est. $= X - E_0(X)$

(ii) If $\rho(u) = |u|$, " $v^* = \text{med}_0(X) \Rightarrow$ " $= X - \text{Med}_0(X)$

(iii) If $\rho(u) = I_{\{|u| > k\}}$, then minimizing

$$E_0 \rho(X - v) = P_0(|X - v| > k) \text{ or maximizing } P_0(|X - v| \leq k).$$

* For example, if f is symm. around 0 and unimodal, the MRE est. is $\delta^*(X) = X - 0$ since $v^*(\underset{\sim}{y}) = 0$
* If f is symm. around 0 and U-shaped with support $[-c, c]$. Then $v_1^* = c - k$ & $v_2^* = k - c$ are both minimizers and $\delta_1^* = X - c + k$ & $\delta_2^* = X + c - k$ are both MRE est.'s.

Example: Let $X_1, \dots, X_n \overset{r.s.}{\sim} N(\theta, \sigma^2)$, σ is known. Now, $\delta_0 = \bar{X}$ is equivariant. It follows from Basu's theorem, δ_0 is indep. of $\underset{\sim}{Y} = (X_1 - X_n, \dots, X_{n-1} - X_n)$ and hence $v^*(\underset{\sim}{y}) = v^*$

is constant determined by minimizing $E_0\, \rho(\bar{X} - v)$

By H.W.1: For ~~convex~~ Convex, even ρ, then the value minimizes $E_0\, \rho(\bar{X} - v)$ is $v = 0 \Longrightarrow$

$$\delta^*(\underset{\sim}{x}) = \bar{X} - 0 = \bar{X}.$$

Pitman Estimator of ξ

Let $L(\xi, d) = (d - \xi)^2$, $\rho(w) = w^2$. The MRE est.

$$\delta^*(\underset{\sim}{x}) = \delta_0(\underset{\sim}{x}) - E\left(\delta_0(\underset{\sim}{x}) \mid \underset{\sim}{y}\right) \text{ is}$$

$$\delta^*(\underset{\sim}{x}) = \frac{\int u\, f(x_1 - u, \dots, x_n - u)\, du}{\int f(x_1 - u, \dots, x_n - u)\, du}$$

It is called the Pitman est. of ξ.

<u>Example</u>: $X_1, \ldots, X_n \overset{iid}{\sim} U(\xi - \frac{b}{2}, \xi + \frac{b}{2})$, b is known.

$$f(x_1 - \xi, \ldots, x_n - \xi) = \frac{1}{b^n} I_{\{\xi - \frac{b}{2} < X_{(1)} < X_{(n)} < \xi + \frac{b}{2}\}}$$

$$\delta^*(\underset{\sim}{X}) = \frac{\int_{x_{(n)} - \frac{b}{2}}^{x_{(1)} + \frac{b}{2}} u \, du}{\int_{x_{(n)} - \frac{b}{2}}^{x_{(1)} + b/2} du} =$$

$$= \frac{\frac{1}{2}\left[\left(x_{(1)} + \frac{b}{2}\right)^2 - \left(x_{(n)} - \frac{b}{2}\right)^2\right]}{x_{(1)} - x_{(n)} + b}$$

$$= \frac{\frac{1}{2}\left[X_{(1)}^2 + X_{(n)}^2 + b\left(X_{(1)} + X_{(n)}\right)\right]}{X_{(1)} + X_{(n)} + b}$$

$$= \frac{X_{(1)} + X_{(n)}}{2}\left[\frac{X_{(1)} + X_{(n)} + b}{X_{(1)} + X_{(n)} + b}\right]$$

$$= \frac{X_{(1)} + X_{(n)}}{2}$$

<u>Remarks</u>: For SE loss

(1) If $\delta(\underset{\sim}{X})$ is any equivariant with constant bias b, then $\delta(\underset{\sim}{X}) - b$ is equivariant, unbiased and has smaller risk than $\delta(\underset{\sim}{X})$

(2) The unique MRE estimator is unbiased.

(3) If UMVUE exists and is location equivariant, then it is also MRE.

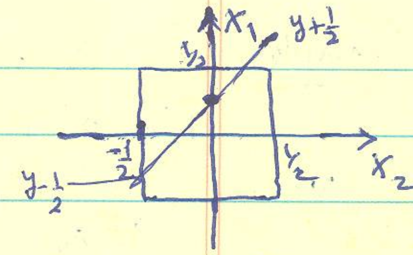
(4) Unlike UMVUEs which frequently inadmissible, the Pitman est. is admissible under mild conditions.

<u>Example:</u> Let X_1 and X_2 be 2 indep. r.v.'s with joint pdf
$$f(x_1,x_2) = I_\xi(x_1)\cdot I_\xi(x_2)$$
where
$$I_\xi(x) = \begin{cases} 1 & \text{if } |x-\xi| < \frac{1}{2} \text{ or } \xi-\frac{1}{2} < x < \xi+\frac{1}{2} \\ 0 & \text{e.w.} \end{cases}$$
That is $X_i \sim U(\xi-\frac{1}{2}, \xi+\frac{1}{2})$. Consider the following loss function
$$L(\xi,\delta) = \rho(\delta-\xi) = \begin{cases} 1 & \text{if } |\delta-\xi| > k \\ 0 & \text{e.w.} \end{cases}$$
Since $\delta_0(\underset{\sim}{X}) = X_1$ equivariant with finite risk, the MRE est. is
$$\delta^*(\underset{\sim}{X}) = X_1 - v^*(Y),\ Y = X_1 - X_2,$$
where $v^*(y)$ is the value minimizing $E_0\left[\rho(X_1 - v(Y))\,|\,Y\right]$.
$$E_0\left[\rho(X_1-v(Y))\,|\,Y\right] = P_0(|X_1-v(Y)| > k\,|\,Y)$$
$$= 1 - P_0(v-k < X_1 < v+k\,|\,Y)$$
Need the dist. of $X_1|Y=y$.

- $P_0(X_1=x_1, X_1-X_2=y) = P(X_1=x_1, X_2=x_1-y)$
$$= P(X_1=x_1)\,P(X_2=x_1-y)$$
$$= I(x_1)\cdot I(x_1-y)$$
$$= \begin{cases} 1 & \text{if } |x_1| < \frac{1}{2},\ |x_1-y| < \frac{1}{2} \\ 0 & \text{e.w.} \end{cases}$$
$$= \begin{cases} 1 & \text{if } -\frac{1}{2} < x_1 < \frac{1}{2},\ y-\frac{1}{2} < x_1 < y+\frac{1}{2} \\ 0 & \text{e.w.} \end{cases}$$
- The dist. of $Y=y$, $P(X_1-X_2=y) = P(X_1 = X_2+y)$

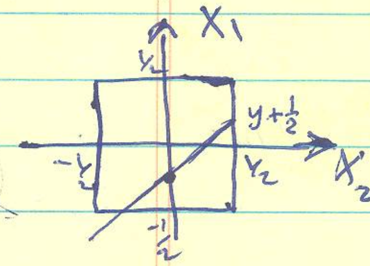
For $y>0$, $P(Y=y) = \int_{y-\frac{1}{2}}^{1/2} 1 \, dx_1 = 1-y$

For $y<0$, $P(Y=y) = \int_{-1/2}^{y+\frac{1}{2}} 1 \, dx_1 = 1+y$



$\Rightarrow$

$$P(Y=y) = \begin{cases} 1-y & \text{if } y>0 \\ 1+y & \text{if } y<0 \end{cases}$$



For $y>0$, $P(X_1=x_1 \mid Y=y) = \dfrac{1}{1-y}$, $y-\frac{1}{2} < x_1 < \frac{1}{2}$

$y<0$, $\qquad = \dfrac{1}{1+y}$, $-\frac{1}{2} < x_1 < y+\frac{1}{2}$

$\Rightarrow$ For $y>0$, $X_1 \mid Y=y \sim U\left(y-\frac{1}{2}, \frac{1}{2}\right)$

$y<0$, $X_1 \mid Y=y \sim U\left(-\frac{1}{2}, y+\frac{1}{2}\right)$

It is clearly to note that $v^*(y) = \text{med}(X_1 \mid Y=y) = \dfrac{Y}{2}$

$\Rightarrow$

$$\delta^*(\underset{\sim}{X}) = X_1 - \frac{X_1 - X_2}{2} = \frac{X_1+X_2}{2}.$$

Notes: (1) Let $X_i \sim U\left(\xi-\frac{1}{2}, \xi+\frac{1}{2}\right)$, $i=1,2$ $\Rightarrow$ $\delta^* = \dfrac{X_1+X_2}{2}$ is unbiased equiv. est. with $E\delta^* = \xi$ and $Var(\delta^*) = \frac{1}{4}\left(\frac{1}{12}+\frac{1}{12}\right) = \frac{1}{24}$.

(2) One can think in other estimators (not equiv., unbiased, small risk)

Let $\delta_a = \begin{cases} a\delta^* & \text{if } Z=1 \\ 0 & \text{if } Z=0 \end{cases}$ with $P(Z=1) = \frac{1}{a} = 1-P(Z=0)$, $a>1$

$$E_\xi(\delta_a) = EE(\delta_a \mid Z) = E(\delta_a \mid Z=1)\frac{1}{a} + E(\delta_a=0 \mid Z=0)\left(1-\frac{1}{a}\right)$$
$$= a\xi \frac{1}{a} + 0 = \xi \Rightarrow \delta_a \text{ is unbiased}$$

However,

$$Var_0(\delta^*) = \frac{1}{24} \overset{?}{<} Var_0(\delta_a) = E\, Var_0(\delta_a \mid Z) + Var\, \cancel{E_0}(\delta_a \mid Z)$$
$$= \frac{a^2}{24}\frac{1}{a} = \frac{a}{24} \quad \text{True for } a>1,$$

But for $a<1$, $Var_0(\delta^*) > Var_0(\delta_a)$.