HONAM MATHEMATICAL JOURNAL
Volume 13, Number 1. July 1991.

Best Invariant Estimators in the Scale Parameter Problem

Kuey−Chung Choi

Department of Computer Science and Statistics, College of Natural Science, Chosen University.
Kwangju, 501—759, Korea

Abstract

In this paper we first present the elements of the theory of families of distributions and corresponding estimators having structual properties which are preserved under certain groups of transformations, called "Invariance Principle".

The invariance principle is an intuitively appealing decision principle which is frequently used, even in classical statistics. It is interesting not only in its own right, but also because of its strong relationship with several other proposal approaches to statistics, including the fiducial inference of Fisher [3, 4], the structural inference of Fraser [5], and the use of noninformative priors of Jeffreys [6]. Unfortunately, a space precludes the discussion of fiducial inference and structural inference. Many of the key ideas in these approaches will, however, be brought out in the discussion of invarience and its relationship to the use of noninformatives priors.

This principle is also applied to the problem of finding the best scale invariant estimator in the scale parameter problem. Finally, several examples are subsequently given.

1. Introduction

The invariance principle involves groups of transformations over the following three spaces : the parameter space $\boldsymbol{\Theta}$, the decision space $\mathbf{A}$ and the sample space $\mathbf{X}$. The most basic is the group of transformations of $\mathbf{X}$ onto $\mathbf{X}$, that is, if for every $x_1 \in \mathbf{X}$ there exists an $x_2 \in \mathbf{X}$ such that $g(x_2) = x_1$. A transformation g from $\mathbf{X}$ into itself is to be one−to−one if $g(x_1)=g(x_2)$ implies $x_1=x_2$.

Let G denote a group of measurable transformations from $\mathbf{X}$ into itself. The basic group operation is composition : if g_1 and g_2 are transformations from $\mathbf{X}_t$ into itself, $g_2 \circ g_1$ is defined as the transformation $x \to g_2(g_1(x))$, $x \in \mathbf{X}$. A set of transformations on a space is a group if it is closed under composition operator · which is associative and inverse, and has the identity e. The inverse of a transformation g exist if and only if g one−to−one and onto. Hence all transformations in G are automatically one−to−one and onto. The assumption $g \in G$ be measurable is made into ensure that whenever X is a random variable in $\mathfrak{X}$, then $g(x)$ is also a random variable. Comprehensive treatment can be found in the books of Ferguson [2], Berger [1], and Lehmann [8].

The purpose of this paper is to study some basic results for invariant estimation in the scale parameter problem.

In section 2 we treat some definitions and preliminary results containing a group of measurable transformations.

In section 3 we obtain best invariant estimators for the scale parameter in the one dimensional case as well as the n−dimensional case, and treat the best scale invariant estimator (Pitman's estimator) as a special case. Finally, section 4 contains sevral examples for the results of Section 3.

2. Preliminaries

Definition 2.1. Let $\mathfrak{X}$ be any non−empty set.

G is a group of (measurable) transformation on $\mathfrak{X}$ if

i) $g \in G \Rightarrow g$ is a one to one and onto function from $\mathfrak{X}$ to $\mathfrak{X}$;

ii) the identity function $e(x)=x$ for all x belongs to G ;

iii) $g \in G \Rightarrow g^{-1} \in G$;

iv) $g \in G,\ h \in G \Rightarrow g \circ h \in G$ where

$(g \circ h)(x) \equiv gh(x) \equiv g \circ h(x) = g(h(x))$

Let $\{P_\theta^X : \theta \in \Theta\}$ be the family of distributions of X taking on values in $\mathfrak{X}$ with the corresponding Borel field $\mathfrak{B}$. Assume that $\theta_1 \neq \theta_2$ implies $P_{\theta_1}^X \neq P_{\theta_2}^X$ (Identifiability). Suppose G is a group of transformations on $\mathfrak{X}$ such that

i) if $g \in G$, then g is a measurable function from $(\mathfrak{X}, \mathfrak{B})$ to $(\mathfrak{X}, \mathfrak{B})$. (Note that since $g^{-1} \in G$ and is measuarable we see that g is bimeasurable.)

ii) if $g \in G$, then $P_\theta^{gx} \in \{P_{\theta'}^x : \theta' \in \Theta\}$ where $P_\theta^{gx}(B) = P_\theta(gX \in B)$, $B \in \mathfrak{B}$, (Such a θ' is unique by identifiability)

i.e., for every $g \in G$ and every $\theta \in \Theta$ $\exists$ unique $\theta' \in \Theta \cdot \ni \cdot$ the distribution of $g(X)$ is given by $P_{\theta'}^x$ whenever the distribution of X is given by P_θ^x. (For example, we start with a normal problem, then it remains a normal problem under any transformation).

When this happens, we say that G leaves the probability structure "invariant", i.e., a family of distributions P_θ^X, $\theta \in \Theta$, on $\mathfrak{X}$ is "invariant" under G.

For each $g \in G$ let $\bar{g}$ be a function from Θ into Θ defined by $P_\theta^{gX} = P_{\bar{g}\theta}^X = P_{\theta'}X$ (Such a $\theta' = \bar{g}\theta$ is uniquely defined by g and θ).

Theorem 2.1. If the probability structure is invariant under G, then

$$E_\theta \phi(gX) = E_{\bar{g}\theta} \phi(X)$$

for all measurable function ϕ (integrable real function).

Outline of Proof)

Let $\phi = I_B$, $B \in \mathfrak{B}$. Next ϕ simple, nonnegative and take the limit. Finally, arbitrary measurable function.

Theorem 2.2. If a family of distributions P_θ^X, $\theta \in \Theta$, is invariant under G, then $\bar{G}$ is a group of transformations on Θ where

$$\bar{G} = \{\bar{g} : g \in G\}.$$

Furthermore, $\overline{gh} = \bar{g} \circ \bar{h}$, i.e., $g \to \bar{g}$ is a homomorphism, and hence $\bar{e} = e$ and $\overline{g^{-1}} = \bar{g}^{-1}$.

Proof. To show $\overline{gh} = \bar{g} \circ \bar{h}$:

$$\begin{aligned} B \in \mathfrak{B},\ P_{\overline{gh}\theta}^X(B) &= P_\theta^{ghX}(B) \\ &= P_\theta(gh(X) \in B) \\ &= P_\theta(h(X) \in g^{-1}(B)) \\ &= P_\theta^{hX}(g^{-1}(B)) \\ &= P_{\bar{h}\theta}^X(g^{-1}(B)) \\ &= P_{\bar{h}\theta}(X \in g^{-1}(B)) \end{aligned}$$

Since this is true for all $\theta \in \Theta$ and $B \in \mathfrak{B}$, we have $\overline{gh} = \bar{g} \circ \bar{h}$

To show $\bar{g}$ is onto Θ :

$$e = g \circ g^{-1} \Rightarrow \bar{e} = \overline{gg^{-1}} = \bar{g} \circ \overline{g^{-1}}$$

$$\theta = \bar{e}(\theta) = \bar{g}(\overline{g^{-1}}(\theta)) = \bar{g}(\theta^*) \text{ where } \theta^* = \overline{g^{-1}}(\theta)$$

To show $\bar{g}$ is one to one :

$$\begin{aligned} \bar{g}(\theta_1) = \bar{g}(\theta_2) &\Rightarrow \overline{g^{-1}}\,\bar{g}(\theta_1) = \overline{g^{-1}}\,\bar{g}(\theta_2) \\ &\Rightarrow \overline{g^{-1}g}(\theta_1) = \overline{g^{-1}g}(\theta_2) \end{aligned}$$

$\Rightarrow \bar{e}(\theta_1)=\bar{e}(\theta_2)$

$\Rightarrow \theta_1=\theta_2$

These facts also show that $\bar{G}$ is closed under inverses and $\bar{e}$ is the identity in $\bar{G}$. Now, we want to estimate $r(\theta)$ with $\{r(\theta) : \theta\in \boldsymbol{\Theta}\}\subset D$ where D is decision space.

Definition 2.2. We say that G is invariant for the estimation of $r(\theta)$

if $r(\theta_1)=r(\theta_2)$ implies $r(\bar{g}\,\theta_1)=r(\bar{g}\,\theta_2)\ \forall \bar{g}\in\bar{G}$.

We assume that there exists a group of transformations $\bar{\bar{G}}$ acting on D such that $\bar{\bar{g}}(r(\theta))=r(\bar{g}(\theta))\ \forall\theta\in\boldsymbol{\Theta}$ and $\forall g\in G$. Also, assume that $g\to\bar{\bar{g}}$ is a map such that $\overline{\overline{gh}}=\bar{\bar{g}}\circ\bar{\bar{h}}$ for all $g, h\in G$. Note that if $D=\{r(\theta) : \theta\in\boldsymbol{\Theta}\}$, then such a $\bar{\bar{G}}$ exists which satisfies $\overline{\overline{gh}}=\bar{\bar{g}}\circ\bar{\bar{h}}$.

Definition 2.3. An estimator δ is said to be invariant under G if

$\bar{\bar{g}}\delta(X)=\delta(gX)$

or $\delta(X)=\bar{\bar{g}}^{-1}\delta(gX)\ \forall g\in G,\ \forall X$

Finally, we assume that the loss function is invariant under G, i.e., $L(e, d)=L(\bar{g}\theta, \bar{\bar{g}}d)\ \forall g\in G$, $\theta\in\boldsymbol{\Theta}$, $d\in D$ ($d'=\bar{\bar{g}}d$ is unique).

Definition 2.4. Two points θ_1 and θ_2 are equivalent under G if there exists a $\bar{g}\in\bar{G}$ such that $\bar{g}(\theta_1)=\theta_2$.

(This breaks $\boldsymbol{\Theta}$ up into equivalent classes, call it "orbits")

Theorem 2.3. Let δ be an invariant estimator, then the risk function is constant along orbits of $\bar{G}$, i.e.,

$$R(\theta, \delta)=R(\bar{g}\theta, \delta)\ \forall\bar{g}\in\bar{G},\ \forall\theta\in\boldsymbol{\Theta}$$

Proof. $R(\theta, \delta)=E_\theta L(\theta, \delta(X))$

$=E_\theta L(\bar{g}\theta, \bar{\bar{g}}\delta(X))$ ($\because$ loss invariant)

$=E_\theta L(\bar{g}\theta, \delta(g(X))$ ($\because$ δ invariant)

$=\int L(\bar{g}\theta, \delta(g(X)))dP_\theta^X$

$=E_{\bar{g}\theta}L(\bar{g}\theta, \delta(X))$ ($\because$ $E_\theta\phi(\bar{g}X)=E_{\bar{g}\theta}\phi(X)$)

$=R(\bar{g}\theta, \delta)\ \forall\bar{g}\in\bar{G},\ \forall\theta\in\boldsymbol{\Theta}$

3. Best Invariant Estimation in the Scale Parameter Problem

Let X be a real valued random variable with density $f_\theta(X)$, $\theta \in \Theta = (0, \infty)$, with respect to some σ-finite measure μ.

Definition 3.1. θ is called a scale parameter if

$$f_\theta(x) = \frac{1}{\theta} f_{\theta=1}\left(\frac{x}{\theta}\right) = \frac{1}{\theta} f\left(\frac{x}{\theta}\right)$$

for some known function f which vanishes unless $x > 0$.

We wish to estimate $r(\theta) = \theta$ under the loss

$$L(\theta, d) = L\left(\frac{d}{\theta}\right),\ d \in D = (0, \infty)$$

$G = \{g_b : b > 0\}$, $g_b(x) = bx$, is a group.

$$X \sim f_\theta(x) = \frac{1}{\theta} f_{\theta=1} f\left(\frac{x}{\theta}\right) = \frac{1}{\theta} f\left(\frac{x}{\theta}\right),\ \theta \in \Theta$$

$$g_b(X) = bX = b\theta \cdot \frac{X}{\theta} \sim \frac{1}{b\theta} f\left(\frac{x}{b\theta}\right),\ b\theta \in \Theta$$

$\bar{g}_b\theta = b\theta \Rightarrow$ there is only one orbit in Θ

$$r(\theta_1) = r(\theta_2) \Rightarrow \theta_1 = \theta_2 \Rightarrow b\theta_1 = b\theta_2.\ \forall b > 0$$

$$\Rightarrow r(\bar{g}_b\theta_1) = r(\bar{g}_b\theta_2)$$

$$\bar{\bar{g}}_b(\theta) = \bar{\bar{g}}_b r(\theta) = \bar{r}(\bar{g}_b\theta) = r(b\theta) = b\theta \Rightarrow \bar{\bar{g}}_b(\theta) = \bar{\bar{g}}_b(\theta_1) = b\theta$$

$$\therefore\ \bar{\bar{g}}(d) = bd$$

Here $G = \bar{G} = \bar{\bar{G}}$

$$\bar{\bar{g}}_c(\delta(x)) = \delta(g_c(x))\ \ \forall x > 0,\ \forall c > 0$$

$$\Rightarrow c\,\delta(x) = \delta(g_c(x))\ \ \forall x > 0,\ \forall c > 0$$

Take $\quad c = 1/x \Rightarrow \frac{1}{X}\,\delta(x) = \delta(1)$

$$\Rightarrow \delta(x) = x\delta(1) = bx \text{ where } b = \delta(1)$$

$$\therefore\ \delta_b(x) = bx\ \ \forall b > 0,\ \forall x > 0$$

Hence, every nonrandomized invariant estimator in of the form

$$\delta_b(X) = bX\ \ \forall b > 0$$

$$L(\bar{g}\theta, \bar{\bar{g}}d) = L(\theta, d)$$

$$R(\theta, \delta_b)=E_\theta L(\frac{bX}{\theta})=E_1 L(bX)=R(1, \delta_b) \ \forall\theta, \ b>0$$

(δ_b has constant risk $\forall b>0$).

Suppose $E_1L(bX)$ exists and is finite for some $b>0$, and assume $\exists$ b · ∈ · $E_1L(b_0X)=\inf_{b>0} E_1L(bX)$. Then $\delta(X)=b_0X$ is "best invariant". Here, it should be remarked that if every nonrandomized invariant estimator has constant risk, then the nonrandomized invariant estimators from an essentially complete class among the class of all randomized invariant estimators.(Hence, in trying to find a best invariant estimator, attention may be restricted to the nonrandomized estimators)

Now let $X=(X_1, X_2, \cdots, X_n)$ be a random vector in R^n with the density $f_\theta(X)$ with respect to some σ-finite measure μ, $\theta\in\Theta=(0, \infty)$ satisfying

$$f_\theta(X)=(\frac{1}{\theta})^n f_{\theta=1}(\frac{X_1}{\theta}, \cdots, \frac{X_n}{\theta})=(\frac{1}{\theta})^n f(\frac{X_1}{\theta}, \cdots, \frac{X_n}{\theta})$$

where f is some known function which vanishes unless all coordinates are positive. (the X_i's need not be independent)

It is desired to estimate $r(\theta)=\theta$ under the loss

$$L(\theta, d)=L(d/\theta), \ d\in D=(0, \infty).$$

We solve this n-dimensinal problem by reducing to the 1-dimensinal problem as follows :

Define $Y_i=\frac{X_i}{X_n}$, $i=1, 2, \cdots, n-1$.

Since $Y_i=\frac{X_i/\theta}{X_n/\theta}$ and the distribution of $\frac{X_i}{\theta}$ and $\frac{X_n}{\theta}$ does not depend on θ, the joint distribution of $Y=(Y_1, Y_2, \cdots, Y_{n-1})$ does not depend on θ.

So we may pretend that $(Y_1, Y_2, \cdots, Y_{n-1})$ are observed first, and then X_n is chosen from the conditional distribution of X_n given Y, a distribution with θ as a scale parameter. This conditional problem is intuitively equivalent to the original problem. By 1-dimensinal, case, the best (nonrandomized) invariant estimator of the conditional problem is $\delta_0(X)=b_0(Y)X_n$ (in the sence that for each $Y=y$, $b_0(y)$ is the number (provided it exists) for which

$$E_1(L(b_0(Y)X_n \mid Y=y))=\inf_b E_1(L(b(Y)X_n) \mid Y=y).$$

In fact the estimator δ_b is best invariant for the original problem. This can be easily shown as follows :

$$G=\{g_c : g_c(n)=cX \text{ for } c>0\} \text{ is a group.}$$

$g_c(X)=(cX_1, \cdots, cX_n)\sim(c\theta)^{-n}f(\frac{X_n}{c\theta}, \cdots, \frac{X_n}{c\theta})$

$\bar{g}_c(\theta)=c\theta$ ($\bar{G}$ has only one orbit in $\boldsymbol{\Theta}$)

$\bar{\bar{g}}_c(d)=cd$

Then, the original problem is invariant under G with $\bar{g}_c$ and $\bar{\bar{g}}_c$

$(G=\bar{G}=\bar{\bar{G}})$

δ is invariant$\Rightarrow\bar{\bar{g}}_c\delta(x)=\delta(g_c x)$

$\Rightarrow c\delta(x)=\delta(cx_1, \cdots, cx_n)\ \forall x_i>0,\ \forall c>0$

$\Rightarrow\frac{1}{x_n}\delta(x)=\delta(\frac{x_1}{x_n}, \cdots, \frac{x_{n-1}}{x_n}, 1)\ (\because c=\frac{1}{x_n})$

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

$\Rightarrow\delta(x_1, \cdots, x_n)=x_n b(y_1, \cdots, y_{n-1})$

where $b(y_1, \cdots, y_{n-1})=\delta(\frac{x_1}{x_n}, \cdots, \frac{x_{n-1}}{x_n}, 1)$

Every nonrandomized invariant estimator is of the form

$\delta(X)=X_n b(Y)$

$R(\theta, \delta)=E_\theta L(X_n b(Y)/\theta)$

$=E_\theta\{E_\theta[L(X_n b(Y)/\theta \mid Y]\}$

$=E_\theta\{E_1[L(X_n b(Y)/\theta \mid Y]\}$

$=E_1\{E_\theta[L(X_n b(Y)/\theta \mid Y]\}$ (By Fubini Theorem)

$=E_1(L(X_n b(Y)))$

$=R(1, \delta)\ \forall\theta\in\boldsymbol{\Theta}$, and $\forall\delta$

$R(\theta, \delta)=R(1, \delta)=E_1 L(X_n b(Y))$

$=E_1\{E_\theta[L(X_n b(Y)) \mid Y]\}$

$=E_\theta\{E_1[L(X_n b(Y)) \mid Y]\}$ (By Fubini Theorem)

$\geq E_\theta\{E_1[L(X_n b_0(Y)) \mid Y]\}$

$=E_1\{E_\theta[L(X_n b_0(Y)) \mid Y]\}$ (By Fubini Theorem)

$=E_1 L(X_n b_0(Y))$

$=R(1, \delta_0)$

$=R(\theta, \delta_0)$ (provided only that $X_n b_0(Y)$ is measurable considered as a function of Y and $R(1, \delta_0)<\infty$).

Hence, $\delta^0(X)=X_n b_0(Y)$ is best invariant for the original problem. As a special case, consider the problem of estimating the scale parameter θ under the loss $L(d/\theta)=\{(d/\theta)-1\}^2=(\theta-d)^2/\theta^2$. The resulting best invariant estimator is called Pitman's estimator.

$$\theta=1:\ L(d)=(1-d)^2,\ X\sim f(x_1, \cdots, x_n)$$

$$Y_1=X_1/X_n, \qquad X_1=Y_1Y_n$$

$$\cdots\cdots\cdots\cdots$$

$$Y_{n-1}=X_{n-1}/X_n, \qquad X_{n-1}=Y_{n-1}Y_n$$

$$Y_n=X_n, \qquad X_n=Y_n$$

$$|J|=y_n^{n-1}=x_n^{n-1}$$

$$b_0(y)=\inf_b E_1[(1-X_n b(y))^2 \mid Y=y]$$

$$=\frac{E_1(X_n \mid Y=y)}{E_1(X_n^2 \mid Y=y)}$$

$$=\frac{\int_0^\infty x_n g(y_1, \cdots, y_{n-1}, x_n)dx_n}{\int_0^\infty x_n^2 g(y_1, \cdots, y_{n-1}, x_n)dx_n}$$

where $g(y_1, \cdots, y_{n-1}, x_n)=f(y_1x_n, \cdots, y_{n-1}x_n, x_n)x_n^{n-1}$ is the joint desity of $(Y_1, \cdots, Y_{n-1}, X_n)$ when $\theta=1$

$$=\frac{\int_0^\infty x_n f(y_1x_n, \cdots, y_{n-1}x_n, x_n)x_n^{n-1}dx_n}{\int_0^\infty x_n^2 f(y_1x_n, \cdots, y_{n-1}x_n, x_n)x_n^{n-1}dx_n}$$

$$=\frac{\int_0^\infty \omega f(y_1\omega, \cdots, y_{n-1}\omega, \omega)\omega^{n-1}d\omega}{\int_0^\infty \omega^2 f(y_1\omega, \cdots, y_{n-1}\omega, \omega)\omega^{n-1}d\omega} \quad (\omega=x_n/\theta)$$

$$=\frac{\int_0^\infty (x_n/\theta)f(x_1/\theta, \cdots, x_{n-1}/\theta, x_n/\theta)(x_n/\theta)^{n-1}\cdot x_n/\theta^2\cdot d\theta}{\int_0^\infty (x_n/\theta)^2 f(x_1/\theta, \cdots, x_{n-1}/\theta, x_n/\theta)(x_n/\theta)^{n-1}\cdot x_n/\theta^2\cdot d\theta}$$

$$=\frac{1}{x_n}\ \frac{\int_0^\infty \theta^{-(n+2)}f(x_1/\theta, \cdots, x_n/\theta)d\theta}{\int_0^\infty \theta^{-(n+3)}f(x_1/\theta, \cdots, x_n/\theta)d\theta}$$

$$\therefore\ \delta^0(x)=X_n b_0(y)$$

$$=\frac{\int_0^\infty \theta^{-(n+2)}f(x_1/\theta, \cdots, x_n/\theta)d\theta}{\int_0^\infty \theta^{-(n+3)}f(x_1/\theta, \cdots, x_n/\theta)d\theta}.$$

This is the (generalized) Bayes estimator when the prior is

$g(\theta)=\theta^{-3}$, $\theta>0$ under squared error loss $L(\theta, d)=(\theta-d)^2$.

4. Examples

Example 4.1. $X=(X_1, \cdots, X_n)$

X_i's iid uniform $(\theta, 2\theta)$, $\theta \in \Theta=(0, \infty)$.

To estimate $r(\theta)=\theta$ under the loss

$$L(\theta, d)=L(\frac{d}{\theta})=(1-\frac{d}{\theta})^2=\frac{1}{\theta^2}(\theta-d)^2,\ d \in (0, \infty).$$

$$f_\theta(x_1, \cdots, x_n)=\frac{1}{\theta_n} \quad I\{\theta \leq \min x_i \leq \max x_i \leq 2\theta\}$$

$$f(x_1, \cdots, x_n)= \begin{cases} 1 & \text{when } 1 \leq \min x_i \leq \max x_i \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

$$f_\theta(x)=\frac{1}{\theta^n} f(\frac{x_1}{\theta}, \cdots, \frac{x_n}{\theta})$$

$$\delta^0(x)=\frac{\int_A \theta^{-(n+2)} d\theta}{\int_A \theta^{-(n+3)} d\theta}$$

Where $A=\{\theta \in \Theta \mid 1 \leq \min(\frac{x_i}{\theta}) \leq \max(\frac{x_i}{\theta}) \leq 2\}$

$$=\frac{\int_{½\max x_i}^{\min x_i} \theta^{-(n+2)}\, d\theta}{\int_{½\max x_i}^{\min x_i} \theta^{-(n+3)}\, d\theta}$$

$$=\frac{n+2}{n+1} \frac{(½ \max x_i)^{-(n+1)}-(\min x_i)^{-(n+1)}}{(½ \max x_i)^{-(n+2)}-(\min x_i)^{-(n+2)}}$$

Remark 4.1. Admissibility of $\delta^0(X)$ in estimating θ was shown in Kim [7] using Karlin's method. In fact he showed that

$$\delta^*(X)=\frac{n+2k}{n+k}\left[\frac{\{\max X_i/(s+1)\}^{-(n+k)}-\{\min X_i/s\}^{-(n+k)}}{\{\max X_i/(s+1)\}^{-(n+2k)}-\{\min X_i/s\}^{-(n+2k)}}\right]$$

is the (generalized) Bayes estimator of θ^k, $k>0$, with respect to the prior $g(\theta)=\theta^{-2k-1}$, $\theta>0$, and also is admissible under squared error loss $L(\theta, d)=(\theta-d)^2$ when $X_1, \cdots, X_n$ are iid as

$$f_\theta(x)= \begin{cases} 1/\theta, & s\theta \leq x \leq (s+1)\theta \\ 0, & \text{otherwise} \end{cases}$$

where s is known positive constant, $\theta>0$.

Remark 4.2. It should be also remarked that it is not necessarily true that a best invariant estimator of a scale parameter is admissible. For example, see [1], [2].

Example 4.2. $X=(X_1, \cdots, X_n)$

X_i's iid $N(0, \sigma^2)$ with

$$f_\sigma(x_i)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}x_i^2}$$

$$f_\theta(x)=\frac{1}{\sigma^n}f_{\sigma=1}(\frac{x}{\sigma})=\frac{1}{\sigma^n}f(\frac{x}{\sigma}),\quad f(y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\sum y_i^2}$$

$$\theta=\sigma:\ f_\theta(x)=\frac{1}{\theta^n}f(\frac{x_1}{\theta}, \cdots, \frac{x_n}{\theta})$$

$$\delta^0(x)=\frac{\int_0^\infty \sigma^{-(n+2)}f(x_1/\sigma, \cdots, x_n/\sigma)d\sigma}{\int_0^\infty \sigma^{-(n+3)}f(x_1/\sigma, \cdots, x_n/\sigma)d\sigma}$$

$$=\frac{\int_0^\infty (\sigma^2)^{-1/2\,(n+2)}e^{-1/2\,\sigma^2\Sigma x_i^2}\,d\sigma}{\int_0^\infty (\sigma^2)^{-1/2\,(n+3)}e^{-1/2\,\sigma^2\Sigma x_i^2}\,d\sigma}$$

$$=\frac{\int_0^\infty (2\omega)^{\frac{1}{2}(n-1)}e^{\beta\omega}\,d\omega}{\int_0^\infty (2\omega)^{\frac{1}{2}n}e^{\beta\omega}\,d\omega}\quad(\omega=1/2\delta^2,\ \beta=\Sigma x_i^2)$$

$$=\frac{2^{\frac{1}{2}(n-1)}}{2^{\frac{1}{2}n}}\ \frac{\int_0^\infty \omega^{\frac{1}{2}(n-1)}e^{-\beta\omega}\,d\omega}{\int_0^\infty \omega^{\frac{1}{2}n}e^{-\beta\omega}\,d\omega}$$

$$=\frac{1}{\sqrt{2}}\ \frac{\dfrac{\Gamma(\frac{1}{2}n+\frac{1}{2})}{\beta^{\frac{1}{2}n+\frac{1}{2}}}}{\dfrac{\Gamma(\frac{1}{2}n+1)}{\beta^{\frac{1}{2}n+1}}}$$

$$=\frac{1}{\sqrt{2}}\ \frac{\Gamma(\frac{1}{2}n+\frac{1}{2})}{\Gamma(\frac{1}{2}n+1)}\ \sqrt{\Sigma x_i^2}$$

Example 4.3. $X=(X_1, \cdots, X_n)$, X_i's iid exponential (θ) with

$$f_\theta(x_i)=1/\theta\cdot\exp\{-x_i/\theta\},\ x_i>0$$

$$f(x)=\frac{1}{\theta^n}f_{\theta=1}(\frac{x}{\theta})=\frac{1}{\theta^n}f(\frac{x}{\theta}),\quad f(y)=e^{-\Sigma y_i},\ y>0$$

$$\delta^0(x)=\frac{\int_0^\infty \theta^{-(n+2)}f(x_1/\theta, \cdots, x_n/\theta)d\theta}{\int_0^\infty \theta^{-(n+3)}f(x_1/\theta, \cdots, x_n/\theta)d\theta}$$

$$=\frac{\int_0^\infty \theta^{-(n+2)}e^{-(\theta^{-1})\cdot\Sigma x_i}d\theta}{\int_0^\infty \theta^{-(n+3)}e^{-(\theta^{-1})\cdot\Sigma x_i}d\theta}$$

$$=\frac{\int_0^\infty \omega^{n+2}e^{-\omega\Sigma x_i}\cdot\omega^{-2}d\omega}{\int_0^\infty \omega^{n+3}e^{-\omega\Sigma x_i}\cdot\omega^{-2}d\omega}\qquad(\omega=1/\theta)$$

$$=\frac{\int_0^\infty \omega^{n}e^{-\omega\Sigma x_i}d\omega}{\int_0^\infty \omega^{n+1}e^{-\omega\Sigma x_i}d\omega}$$

$$=\frac{\gamma(n+1)/(\Sigma x_i)^{n+1}}{\gamma(n+2)/(\Sigma x_i)^{n+2}}=\frac{1}{n+1}\Sigma x_i$$

References

[1] J.O. Berger, Statistical Decision Theory and Bayesian Analysis, 2nd ed., springer-verlag, New York, 1985.

[2] T.S. Ferguson, Mathemetical Statistics, A Decision Theoretic Approach, Academic Press, New York, 1967.

[3] R.A. Fisher, On The Mathemetical Foundation of Theoretical Statisfics, Philos. Trans. Roy. Soc., A, 222, 309~368, 1922.

[4] R.A. Fisher, Invariance Probability, Proc. Camb. Rhil. Soc., 26, 528, 1930.

[5] D.A.S., Fraser, The Structure of Inference, John Wiley, New York, 1968.

[6] H., Jeffreys, Theorey of Probability, 3rd ed., Clarendon Press, Oxford, 1961.

[7] B.H. Kim, Admissibility of Generalized Bayes Estimators in An One Parameter Nonregular Family, Submitted to Metrika, 1990.

[8] E.L. Lehmann, Theory of Point Estimation, John Willy and Sons, Inc., New York, 1983.

[9] E.J.G. Pitman, The Estimation of Location and Scale Parameters of A Continuous Population of any given from, Biometrica, 30, 391~421, New York, 1939.