SOME CONTRIBUTIONS TO BAYES, MINIMAX AND ADMISSIBLE DECISION RULES IN ONE PARAMETER AND MULTIPARAMETER FAMILIES

ANIRBAN DAS GUPTA

Thesis submitted to the Indian Statistical Institute
in partial fulfilment of the requirements
for the award of the degree of
Doctor of Philosophy
CALCUTTA
1982

ACKNOWLEDGMENT

My greatest sincere thanks of gratitude are due to Professor J.K. Ghosh. This work was done under his supervision. It was a pleasure and a privilege to work under him. During the past four years in which I worked on this thesis, he took parental care of my research, my well-being and even my health, making invaluable suggestions with critical insight and stimulating interest at moments of crisis. I am particularly grateful to him for the benefit of his precious time and attention I could draw upon so liberally.

Sincere thanks are due to Dr. Bimal K. Sinha who introduced me to admissibility problems and took personal interest in my work althrough. I must also thank him for permitting me to include our joint works in this thesis.

I had the good fortune of discussing some of my problems with Professor Malay Ghosh. At various stages of this work, he made many extremely useful comments. My heartfelt thanks are due to him.

Friends and colleagues have been extremely kind and inspiring over the years. Decades will go before I can forget the pleasure of my association with R.L. Karandikar, Srinivas Bhogle, N.T. Sreekumar, Preshanta Pathak, C.N. Rao, and many others. My teachers K.K. Roy, Somesh Bagchi and B.V. Rao have graciously listened to my problems on several occasions. I am grateful to all of them.

Mr. Dilip Bardhan, who did a particularly elegant job of typing this manuscript, has my sincere appreciation. Thanks also extend to Mr. Mukta Lal Bag for his efficient duplicating of this thesis.

Anirban Das Gupta

C O N T E N T S

Page

*****

INTRODUCTION AND SUMMARY

The study of admissible, minimax, and Bayes procedures has been of primary importance ever since the pioneering work of Wald (50). Since early seventies, new directions have been opening up, and not merely new techniques, completely new interpretations and interrelations have come to be known.

To prove admissibility of estimates the most commonly used technique is to show that it is extended Bayes and approximate its risk by the risk of the corresponding Bayes estimates. This technique is due to Blyth (51). A sort of converse result, which essentially shows that this technique must work for all admissible estimates is due to Farrell and Stein who derive a necessary and sufficient condition for admissibility.

Karlin (58) was among the foremost statisticians to have evolved a general technique of proving admissibility in one dimension. In the one parameter exponential family with a density $p(x,w) = e^{xw}\beta(w)d\mu(x)$, $\underline{w} < w < \bar{w}$, Karlin showed that a linear estimate of the form $\frac{X}{\lambda+1}$, $\lambda \geq 0$, is admissible for $E_w(X)$, if

$$\int_{\underline{w}}^{a} \beta^{-\lambda}(w)dw = \infty = \int_{b}^{\bar{w}} \beta^{-\lambda}(w)dw, \text{ for } \underline{w} < a, b < \bar{w}.$$

Karlin's conjecture that the converse of this result is also true is open till this day. In this thesis, we shall use Karlin's technique quite extensively in Chapters 2 and 3, for deriving sufficient conditions of admissibility of probably non-linear estimates in the one parameter regular exponential and non-regular families. A unified proof of the admissibility of some standard generalized Bayes estimates of the mean in the exponential family has recently been given by Brown and Hwang (81) in the lines of Blyth (51).

From late sixties, interest in admissibility work shifted from particular problems to very general problems. Some leading stones in this direction were Brown (66), Brown (71) and a series of articles due to Berger (76). Brown (66) showed that the best invariant estimate of a location parameter is, under very general conditions, admissible in dimension 1 and inadmissible for dimensions 3 and more. Subsequently, in the context of simultaneous estimation of independent normal means, Brown (71) discovered a novel relation between admissibility and the recurrence of a related diffusion process. Brown (71) also practically characterised all the admissible estimates of the multivariate normal mean in the generalized Bayes class. Some important work in the spirit of Brown (71) was subsequently done by Srinivasan (81) and in the context of control problems by Srinivasan (82). While Brown (66) dwelt, in broad generality, on the problem of estimating a full location vector, Berger (76a,76b) considered the question of admissibility or otherwise of generalized Bayes estimators of coordinates of a location vector.

The most popular and elegant tool now in use of proving inadmissibility or improving upon inadmissible estimators is the method of solving differential inequalities. The tool should be primarily attributed to Stein who achieved a major breakthrough by proving what is popularly known as " Stein's identity ". It was later generalized by Hudson (7), Berger (80), Hwang ( ). Some very important works in multiparameter inadmissibility and differential inequalities were also done by Brown (79), Brown (81), and Ghosh and Parsian (80). The tool of differential inequalities will be used in this thesis in Chapters 4 and 5, in the context of multiparameter estimation.

Although the thrust in decision-theoretic works is now on multiparameter problems, even now there are important and interesting problems in the one parameter families as well. We shall consider in this thesis different aspects of Bayes, minimax, and admissible estimation in single parameter as well as multiparameter problems.

In Chapter 1, we have shown that if $S \subseteq \mathbb{R}^p$ is a bounded convex set containing $\underset{\sim}{0}$, $\underset{\sim}{a} \in \mathbb{R}^p$, $\underset{\sim}{x}$ is an observation from an arbitrary multiparameter density $p(\underset{\sim}{x} \mid \underset{\sim}{\theta})\ (d\mu)$, and $\underset{\sim}{\theta} \in S_b = \{bs + a : s \in S\}$, then for $b > 0$ small, the Bayes estimate with respect to the least favourable prior on the boundary of S_b is minimax for quite general $\mu(\theta)$ under typically quadratic losses provided some regularity conditions, which are true in the multiparameter exponential family, hold (Theorem 1.2.1). This generalizes the work of Casella and Strawderman (81).

In Chapter 2, following Karlin's technique, we have derived sufficient conditions for admissibility of polynomial estimators

$\delta_n(X) = a_n X^n + \dots + a_1 X + a_0$ (Theorem 2.3.1); important parametric functions for which no linear admissible estimates are known are considered in the examples and a class of admissible estimates have been proposed. The result of Ghosh and Meeden (77) comes out as a special case.

In Chapter 3, we consider the problem of finding admissible estimates of general non-negative and monotone functions $h(w)$ of the basic parameter w in a one-parameter non-regular family with density $p(x,w) = r(x)q(w)$, $\underline{w} < w < \bar{w}$, $\underline{w} < x \leq w$. The mean and all quantiles of the distribution fall in this class if $\underline{w} = 0$. It has been shown (Theorem 3.2.1) that in general $\frac{1+\varepsilon}{\varepsilon} h(X)$ is admissible for $h(w)$ provided h satisfies a Karlin-type condition

$$\int_{\underline{w}}^{a} h'(w)\, h^{\varepsilon-2}(w)\, q(w)dw = \infty = \int_{b}^{\bar{w}} h'(w)\, h^{\varepsilon-2}(w)\, q(w)dw .$$

More generally, we have derived sufficient conditions for admissibility of generalized Bayes estimates of $h(w)$ with respect to more general priors $\pi(w) = |h'(w)|\, f(h(w))/q(w)$, where $f : [0,\infty) \to [0,\infty)$ (Theorem 3.3.1). Theorem 3.2.1 is a special case of Theorem 3.3.1 when $f(u) = u^{-2-\varepsilon}$, $\varepsilon > 0$. Also, it has been shown (Theorem 3.4.1) that $\log q(X) - 1$, the Pitman estimate of $\log q(w)$, is admissible if $q(\bar{w}) = 0$. As an application, we have shown that $X + 1$ is an admissible estimator of w in the truncated exponential distribution with density $e^{-(x-w)}$, $x \geq w$ (0 otherwise).

In Chapter 4, we have switched back to multiparameter problems. Hwang (81) - type bounds on admissible estimators of the scale parameters in independent simple exponential populations have been obtained. It is shown (Theorem 4.2.4) that any estimate $\delta(x) = (\delta_1(x), \ldots, \delta_p(x))$ of the vector of scale-parameters is inadmissible, provided for some $0 < c < 2(p-1)$, and some $M > 0$,

$$\sum_{i=1}^{p} x_i^{-3} \delta_i(x) \leq \sum_{i=1}^{p} x_i^{-3} \delta_{c,i}^{B}(x) \quad \text{for every } x \in (0,M]^p,$$

where $\delta_{c,i}^{B}(x) = \frac{x_i}{2}\left[1 + cx_i^{-4}/2(\sum_{j=1}^{p} x_j^{-2})^2\right]$ = Berger's estimate (80).

A similar necessary condition for admissibility of estimators of independent gamma shape parameters is also given. Finally, we have considered the problem of estimating independent gamma scale-parameters under the invariant loss $L(\theta,a) = \sum_{i=1}^{p} a_i\theta_i - \sum_{i=1}^{p} \log a_i\theta_i - p$, in contrast to the weighted quadratic losses of Berger (80), and have shown that the 'standard' estimate is inadmissible for $p \geq 3$. Some observations are then made relating this result to Brown (80), Brown (66), Brown and Fox (74), and Berger (80).

In Chapter 5, we continue with differential inequalities and multiparameter admissibility problems in terms of an approximating risk within a class of estimates which arise naturally in the context of second order efficiency. Formally, the concept of second order admissibility (Ghosh and Sinha (81)) is extended to general multiparameter families. It is shown that estimates unbiased upto

$o(n^{-1})$ are second order admissible in dimension 2, if the components are both second order admissible in dimension 1 (Theorem 5.2.1), and always inadmissible if $p \geq 3$ (Theorem 5.3.1). Examples are given to show that in general, the method of generating admissible estimates by constructing Bayes solutions may fail even in one dimension, but Bayes solutions do indeed exist if the class of estimates is suitably reduced (Proposition 5.4.1).

In Chapter 6, we consider the problem of estimating $|\Sigma|$ under the loss $(a - |\Sigma|)^2 \, |\Sigma|^{-2}$ on the basis of $S \sim W_p(k,\Sigma)$. Using Brown (66), it is shown that $|S|.(k-p+2)!/(k+2)!$, the best fully equivariant estimate of $|\Sigma|$, is admissible in the class of estimates depending on $|S|$ only (Theorem 6.3.2).

CHAPTER 1

BAYES MINIMAX ESTIMATION IN MULTIPARAMETER FAMILIES WHEN THE PARAMETER SPACE IS RESTRICTED TO A BOUNDED CONVEX SET

1.1 Introduction

Let $X \sim N(\theta,1)$, $-\infty < \theta < \infty$. It is well known that X is generalized Bayes for θ with respect to the uniform prior under squared error loss, and is an admissible minimax estimator. However, if it is known that the mean θ lies in a compact interval $[a,b]$, then X becomes inadmissible and also ceases to remain minimax; Ghosh (64) proved the existence of a unique minimax estimate $u_o(x)$ for θ and provided a sequence of estimates $\{u_n(x)\}$ in the space of estimators with uniformly bounded risk whose maximum risk converges to the minimax value of the problem. His argument, however, did not spell out the exact form of the minimax estimate $u_o(x)$. Casella and Strawderman (81) recently came out with a relatively simple form of $u_o(x)$ and the least favourable prior when $b-a$ is sufficiently small. There is no loss in assuming the interval $[a,b]$ to be symmetric around 0, say, $[-a,a]$. It was observed by Casella and Strawderman (81) that if $a \leq 1.05$ approximately, then the Bayes estimate $\delta_a(x)$ of θ against the prior putting probability $\frac{1}{2}$ each at $\pm a$ is minimax for θ. However, as a increases, the two-point prior ceases to be least favourable. Recently, Bickel (81) exhibited estimates which are asymptotically minimax upto $o(a^{-2})$ and obtained approximations to the least favourable prior density upto the same order.

It is clear from the pictures of risk functions of $\delta_a(x)$ in Casella and Strawderman (81), and can be shown mathematically that if a is a bit smaller ($a \leq 0.643$ approximately), then the risk function of $\delta_a(x)$ is convex in the interval $[-a,a]$, and hence attains its maximum at $\pm a$. $\delta_a(x)$ is thus minimax for Θ. In this chapter, it has been shown in Section 1.2 that variations of this convexity argument, motivated by the normal example, carries over to very general multiparameter families, and the Bayes estimate against the least favourable prior on the boundary of the parameter space continues to be minimax. Some remarks are then made illustrating various aspects of the hypotheses and application of this result. In Section 1.3, some explicit examples are given in one parameter set-ups, where the regions of convexity are actually worked out. Throughout, we work with squared error loss; generalization in this direction is indicated (Remark 4).

1.2 Main Result

Let $S \subseteq \mathbb{R}^p$ $(p \geq 1)$ be any compact convex set containing $\underset{\sim}{0}$ in its interior. Let $\underset{\sim}{a} \in \mathbb{R}^p$. We let

$$S_b = \left\{ b\underset{\sim}{s} + \underset{\sim}{a} : \underset{\sim}{s} \in S \right\}, \; b \geq 0$$

$$\partial S_b = \text{Boundary of } S_b .$$

Note S_b is convex. We take ∂S_b to be compact. Let $\underset{\sim}{x}$ be an observation from a multiparameter (family of) prob. measure

dP_θ which has a density $p(x|\theta)$ with respect to some σ-finite measure $d\mu$. Let $d\pi_b(\theta)$ be the leastfavourable prior distribution if θ lies in the compact set ∂S_b. Let $\delta_{a,b}(x)$ be the Bayes estimate with respect to the prior $d\pi_b(\theta)$. In the following analysis it is implicitly assumed that the marginal distribution of X dominates dP_θ for every θ. $\delta_{a,b}(x)$ is minimax on ∂S_b. Let $\mu(\theta) = (\mu_1(\theta),\ldots,\mu_p(\theta))$ be such that, with

$$D(\theta) = ((\frac{\partial}{\partial\theta_i}\mu_j(\theta))),\ \operatorname{tr} D^T(\theta)D(\theta)>0\,\forall\theta.$$

Further assume that $\frac{\partial^2}{\partial\theta_i^2}\mu_j(\theta)$ exists and is continuous for $1 \leq i,\ j \leq p$.

Theorem 1.2.1. There exists $b > 0$ such that for $\theta \varepsilon S_b$, $\delta_{a,b}(X)$ is minimax for $\mu(\theta)$, provided $R(\theta,\delta_{a,b})$, the risk function of $\delta_{a,b}(X)$, is twice differentiable in θ_i $(1 \leq i \leq p)$ and the p-fold sum of these second-order derivatives $\sum_{i=1}^{p}\frac{\partial^2}{\partial\theta_i^2}R(\theta,\delta_{a,b})$ is jointly continuous in θ and b at the points $(\theta,0)$.

Proof : Define for $b \geq 0$,

$$g(\theta,a,b) = R(\theta,\delta_{a,b})$$

$$f(\theta,a,b) = \sum_{i=1}^{p}\frac{\partial^2}{\partial\theta_i^2}g(\theta,a,b) \qquad \ldots(2.1)$$

Clearly, $g(\theta,a,0) = \sum_{i=1}^{p}[\mu_i(\theta)-\mu_i(a)]^2$.

By direct computation, $\lim_{\theta\to a} f(\theta,a,0)$ exists and is given by

$$\lim_{\theta \to a} f(\theta,a,0) = 2\mathrm{tr}\, D^T(a)D(a) > 0. \qquad \ldots(2.2)$$

Therefore, by the joint continuity of $f(\theta,a,b)$ in θ,b at the points $(\theta,0)$, there exists $b_o, r > 0$ such that,

$$f(\theta,a,b) \geq 0 \quad \forall\, \theta \in S(a,r),\ b \leq b_o .$$

Choosing b_o small, we get $f(\theta,a,b) \geq 0$ $\forall\, \theta \in S_b$, if $b \leq b_o$. Consequently, $g(\theta,a,b)$ is sub-harmonic in θ if $b \leq b_o$, and hence, by the Maximum Principle, attains its maximum on ∂S_b. Since $\delta_{a,b}(X)$ is minimax for variations of θ in ∂S_b, the proof is complete.

Remarks on Theorem 1.2.1

1. If $\mu_i(\theta)$ is a strictly monotone function of θ_i alone for $1 \leq i \leq p$, then the condition $\mathrm{tr}\, D^2(\theta) > 0$ is satisfied for every θ. In particular, the mean in a one-parameter MIR family is handled.

2. By the Mean-value Theorem of multivariate calculus, it can be shown that $\delta_{a,b}(x) \to \mu(a)$ as $b \to 0$, for almost all x. This fact can be used to show quite easily that the condition of joint continuity of $f(\theta,a,b)$ in θ and b holds in the multi-parameter exponential family.

3. For a symmetric distribution, like the Multinormal, the leastfavourable prior is uniform on the boundary of the parameter space. There is no loss in assuming a to be 0 and S a symmetric convex set.

4. Theorem 1.2.1 can be generalized to losses of the form $W(\sum_{i=1}^{p} (\delta_i - \mu_i)^2)$ where W is symmetric, convex, and $W'(0) > 0$.

5. In the special one-parameter families, the argument of sub-harmonicity reduces to plain convexity. It may be mentioned that although the phenomenon of Casella and Strawderman (81) holds under very general conditions, their proof will not carry over to such general situations.

6. In the following section, we actually work out the regions of convexity in some specific distributions. The minimax estimates are equalizer rules on the support of the leastfavourable priors concentrated on the boundary of the parameter space. The Bayes risk of the estimates are the same as their maximum risks and they are thus minimax.

1.3 Examples

In this section, we actually work out the regions of convexity for some important distributions. In distributions, like the Normal or the Binomial, where some kind of symmetry is present, the least-favourable prior puts equal probability on the boundary of the parameter space. In other distributions, finding out the relevant least-favourable priors and subsequently the zones of convexity may become extremely difficult without the aid of a computer. We start with the important Normal case.

Example 1. Let $X \sim N(\theta, 1)$, where $\theta \varepsilon [-n, n]$. The Bayes estimate $\delta_n(x)$ against the prior assigning mass $\frac{1}{2}$ each at $\pm n$ is given by

$$\delta_m(x) = m\,\frac{e^{mx} - e^{-mx}}{e^{mx} + e^{-mx}} = m\tanh mx \qquad \text{(a.e. Lebesgue)} \qquad \ldots(3.1)$$

Therefore, $\delta_m'(x) = m^2 - \delta_m^2(x)$ and $\delta_m''(x) = -2\delta_m(x)\delta_m'(x)$.

Now, $$R(\theta,\delta_m) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left\{\delta_m(x) - \theta\right\}^2 e^{-(x-\theta)^2/2}\,dx \qquad \ldots(3.2)$$

Differentiating under the integral and using Stein's identity, one has,

$$\frac{d}{d\theta}R(\theta,\delta_m) = \frac{1}{\sqrt{2\pi}}\,2\int_{-\infty}^{\infty}\left[x - \delta_m(x) - \delta_m'(x)\left\{x + \delta_m(x)\right\}\right]e^{-(x-\theta)^2/2}\,dx. \qquad \ldots(3.3)$$

Differentiating once again under the integral in (3.3) and then integrating by parts,

$$\frac{d^2}{d\theta^2}R(\theta,\delta_m) = \frac{1}{\sqrt{2\pi}}\,2\int_{-\infty}^{\infty} g'(x)\,e^{-(x-\theta)^2/2}\,dx,$$

where $$g(x) = x - \delta_m(x) - \delta_m'(x)\left\{x + \delta_m(x)\right\} \qquad \ldots(3.4)$$

It is easy to see that $g(-x) = -g(x)$ for every x.

Hence, $$\frac{d^2}{d\theta^2}R(\theta,\delta_m) = \frac{1}{\sqrt{2\pi}}\left[2\int_0^{\infty} g'(x)e^{-(x-\theta)^2/2}\,dx + 2\int_0^{\infty} g'(x)e^{-(x+\theta)^2/2}\,dx\right] \qquad \ldots(3.5)$$

Therefore, it will follow that $R(\theta,\delta_m)$ is convex in θ if $g'(x) \geq 0$ for $x \geq 0$.

Now, from (3.4), and the fact $\delta_m'(x) = m^2 - \delta_m^2(x)$, $\delta_m''(x) = -2\delta_m(x)\delta_m'(x)$, we have,

$$
\begin{aligned}
g'(x) &= 1 - \delta_m'(x) - \delta_m''(x)\left\{x + \delta_m(x)\right\} - \delta_m'(x)\left\{1 + \delta_m'(x)\right\} \\
&= -3\delta_m^4(x) \quad -2x\,\delta_m^3(x) + 2(m^2+1)\delta_m^2(x) + 2m^2x\delta_m(x) \\
&\qquad + 2m^2\,\delta_m^2(x) \quad + 1 - 2m^2 - m^4 \\
&\geq -3m^2\delta_m^2(x) + 2m^2\delta_m^2(x) - 2m^2x\delta_m(x) + 2(m^2+1)\delta_m^2(x) \\
&\qquad + 2m^2x\delta_m(x) + 1 - 2m^2 - m^4
\end{aligned}
$$

(since $0 \leq \delta_m(x) \leq m$ if $x \geq 0$)

$$
\begin{aligned}
&\geq (2+m^2)\delta_m^2(x) + 1 - 2m^2 - m^4 \\
&\geq 0 \quad \text{if} \quad 1 - 2m^2 - m^4 \geq 0 .
\end{aligned}
$$

It follows that $1 - 2m^2 - m^4 \geq 0$ if $m \leq \sqrt{\sqrt{2} - 1} = 0.643$ (approximately), and note that $m^2 > -2$ always holds.

Hence, for $m^2 \leq \sqrt{2} - 1$, $R(\theta, \delta_m)$ is convex in $\theta \in [-m, m]$, and therefore, $\delta_m(x) = m\tanh mx$ is minimax for θ.

If $X \sim N(\theta, \sigma^2)$, where σ is known, then the interval of convexity is approximately 1.29σ wide. Although our calculations allow only $m \leq 0.643$, the zone of convexity is most probably wider.

Example 2. Let $X \sim \text{Bin}(1, \theta)$. Assume $\theta \in [\alpha, 1-\alpha]$ for some $0 < \alpha < \frac{1}{2}$. Consider the two-point prior putting probability $\frac{1}{2}$ each at α and $1-\alpha$. We estimate the mean θ. The Bayes estimate $\delta_\alpha(x)$ of θ for this prior is given by

$$
\begin{aligned}
\delta_\alpha(0) &= 2\alpha(1-\alpha) \\
\delta_\alpha(1) &= \alpha^2 + (1-\alpha)^2 = 1 - \delta_\alpha(0) \qquad \ldots(3.6)
\end{aligned}
$$

By definition, $R(\theta,\delta_\alpha) = [\theta - \delta_\alpha(0)]^2 (1-\theta) + [\theta - \delta_\alpha(1)]^2 \theta$ ...(3.7)

Straightforward calculation shows,

$$\frac{d}{d\theta} R(\theta,\delta_\alpha) = E_\theta \left[\{\theta - \delta_\alpha(X)\} \{2 - \nabla\delta_\alpha(0)\}\right], \qquad ...(3.8)$$

where $\nabla\delta_\alpha(0) = \dfrac{\delta_\alpha(1) - \delta_\alpha(0)}{\delta_\alpha(0)}$.

Adding $\pm x$ and differentiating once more in (3.8), one has,

$$\frac{d^2}{d\theta^2} R(\theta,\delta_\alpha) = 2\delta_\alpha(0) [2 - \nabla\delta_\alpha(0)] > 0$$

$$\text{iff} \quad \nabla\delta_\alpha(0) = \frac{\delta_\alpha(1) - \delta_\alpha(0)}{\delta_\alpha(0)} < 2$$

$$\text{iff} \quad 4\delta_\alpha(0) > 1$$

$$\text{iff} \quad 8\alpha(1-\alpha) > 1$$

$$\text{iff} \quad \frac{1}{2} > \alpha > \frac{1}{2} - \frac{1}{2\sqrt{2}}$$

Therefore, if $\theta \in [\alpha, 1-\alpha]$ for $\frac{1}{2} - \frac{1}{2\sqrt{2}} < \alpha < \frac{1}{2}$, then the Bayes estimate $\delta_\alpha(X)$ given by (3.6) is minimax for θ. Roughly speaking, $\delta_\alpha(X)$ is minimax for θ if θ varies in any sub-interval of $[0.147, 0.853]$ symmetric about $\frac{1}{2}$. This example shows that the convexity argument allows a fairly large subset of the natural parameter space $[0,1]$.

Example 3. Let $X \sim R[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$ and suppose $\theta \in [-a, a]$. We consider the case when $a < \frac{1}{2}$, and we estimate the mean θ.

The Bayes estimate $\delta_a(x)$ against the prior with mass $\frac{1}{2}$ each at $\pm a$ is given by

$$\begin{aligned}\delta_a(x) &= a \quad \text{if} \quad -a+\tfrac{1}{2} < x \leq a+\tfrac{1}{2}\\ &= 0 \quad \text{if} \quad a-\tfrac{1}{2} \leq x \leq -a+\tfrac{1}{2}\\ &= -a \quad \text{if} \quad -a-\tfrac{1}{2} \leq x < a-\tfrac{1}{2} \qquad \ldots(3.9)\end{aligned}$$

$R(\theta,\delta_a)$ is convex in θ if $0 < a \leq \frac{1}{4}$ and $\delta_a(X)$ has equal risk at $\pm a$. Hence, $\delta_a(X)$ in (3.9) is minimax for θ if $\theta \in [-a,a]$ where $0 < a \leq \frac{1}{4}$.

Curiously, for $\frac{1}{4} < a < \frac{1}{2}$, the maximum risk is no longer attained at $\pm a$, but at 0. This indicates that if a so large that the risk function of $\delta_a(X)$ cases to be convex on $[-a,a]$, then the two-point prior no longer serves our purpose.

<u>Example 4</u>. In the previous examples, some kind of symmetry in the underlying distributions was available and the uniform prior on the boundary worked. In this example, we actually find out the leastfavourable prior. Let $X \sim R(0,\theta)$ where $\theta \in [a,b]$. We estimate θ.

The Bayes estimate $\delta_{a,b,}(x)$ of θ against a prior that puts mass π and $1-\pi$ at a and b respectively, is given by

$$\begin{aligned}\delta_{a,b,}(x) &= \frac{ab}{b\pi + a(1-\pi)} \quad \text{if} \quad 0 < x \leq a\\ &= b \qquad\qquad\qquad \text{if} \quad a < x < b \qquad \ldots(3.10)\end{aligned}$$

Our first step is to find out the value of π that forces the risk of $\delta_{a,b,}$ (X) at a and b to be equal.

By direct calculation, it turns out that π satisfies the relation

$$\frac{\pi}{1-\pi} = \sqrt{\frac{a}{b}} \,. \qquad \ldots(3.11)$$

Therefore, the estimate in (3.10) reduces to

$$\begin{aligned} \delta_{a,b}(x) &= \sqrt{ab} \quad \text{if} \quad 0 < x \leq a \\ &= b \quad \text{if} \quad a < x \leq b \end{aligned} \qquad \ldots(3.12)$$

Once again, straightforward calculations give

$$R(\theta,\delta_{a,b}) = \theta^2 - 2b\theta + b^2 + 2ab - 2a\sqrt{ab} + \frac{a^2b - ab^2}{\theta}, \qquad \ldots(3.13)$$

and therefore, $\frac{\partial^2}{\partial\theta^2} R(\theta,\delta_{a,b}) = 2 + \frac{2ab(a-b)}{\theta^3}$ $\qquad \ldots(3.14)$

It follows from (3.14) that $\frac{\partial^2}{\partial\theta^2} R(\theta,\delta_{a,b}) \geq 0$ if $a^2 + ab - b^2 \geq 0$, or equivalently, $a \geq \frac{(\sqrt{5}-1)b}{2}$.

Therefore, for a given b, the region of convexity is approximately [0.62b,b], and for a given a, it is approximately, [a, 1.62a].

CHAPTER 2

ADMISSIBILITY OF POLYNOMIAL ESTIMATORS IN ONE-PARAMETER EXPONENTIAL FAMILY

2.1 Introduction

In Chapter 1, we considered the problem of finding minimax estimates for certain parametric functions when the 'natural' parameter space is restricted. In particular, results in Chapter 1 could be used to obtain minimax estimates of $E_\theta(X)$ when X is an observation from a member of the one-parameter exponential family and θ varies in a small compact sub-interval of the natural parameter space. In this chapter, we consider the problem of finding admissible estimates of parametric functions more general than the mean, in a one-parameter exponential family, when θ varies in the entire natural parameter space.

Let the distribution of X admit a density function $p(x,w) = \beta(w)e^{wx}$ with respect to some σ-finite measure μ on the real line. w represents a typical point in the natural parameter space $\Omega = \left\{ w : \int e^{wx} \, d\mu(x) < \infty \right\}$. It is well known that Ω is an interval $(\underline{w}, \bar{w})$, which may be finite or infinite. Also $\mu(w) = E_w(X) = -\beta'(w)/\beta(w)$ in the interior of Ω and $E_w(X)$ is an increasing function of w. In fact, $\mu'(w) = \text{Var}_w(X) > 0$.

Karlin (58) considered the problem of finding linear admissible estimates of the form $\frac{X}{\lambda+1}$ (for $\lambda \geq 0$) for the mean of X, under squared-error loss. It was shown by Karlin (58) that $\frac{X}{\lambda+1}$ is admissible for $\mu(w)$ if

$$\int_{a}^{\bar{w}} \beta^{-\lambda}(w)dw = \infty = \int_{\underline{w}}^{b} \beta^{-\lambda}(w)dw \qquad \ldots(1.1)$$

for $\underline{w} < a, b < \bar{w}$.

Many interesting admissibility results for the " contraction " linear estimates followed from Karlin (58). Karlin conjectured that his conditions (1.1) are also necessary for the admissibility of $\frac{X}{\lambda+1}$. Although this question has been answered to some extent subsequently [Joshi (69), Morton and Raghavachari (66), Johnstone (81)], in its full generality the question raised by Karlin still remains open. Since then, many people have obtained generalizations of Karlin's result in several directions. By using the Rao - Cramer inequality, Ping (64) obtained sufficient conditions for admissibility of $aX + b$ for estimating the mean. Zidek (70) addressed the problem of finding sufficient conditions for the admissibility of X when the parametric function is any arbitrary piece-wise continuous function $\gamma(w)$. Ghosh and Meeden (77) later used Karlin's argument to find admissible estimates of the form $aX + b$ for estimating the same parametric functions as Zidek's.

An essential step in Karlin's argument is the derivation of a prior $\pi(w)$ through a differential equation. The works following Karlin (58) concentrate on finding only linear admissible estimates of various parametric functions. Sometimes it goes against intuition to use linear estimates for parametric functions which are highly non-linear. Not unexpectedly, there are interesting parametric functions in important distributions for which no linear estimate satisfies the known sufficient conditions of admissibility, and perhaps for which no linear admissible estimates exist. In this chapter, we have considered the problem of finding polynomial admissible estimates using Karlin's argument. However, it becomes extremely difficult to derive the form of the prior if the parametric functions to be estimated are as general as Zidek's. In what follows, we have practically restricted attention to parametric functions of the form $\gamma_m(w) = c_m E_w(X^m) + \ldots + c_1 E_w(X)$. However, in many important members of the one-parameter exponential family, this class includes interesting parametric functions, including $Var_w(X)$. Example 4 in Section 2.4 considers a different type of parametric function.

In Section 2.2, we derive the form of the prior. Next, in Section 2.3, explicit sufficient conditions are obtained for the admissibility of estimates of the form $\delta_m(X) = a_0 + a_1 X + \ldots + a_m X^m$. The technique used is Karlin's. The main result is stated for $m = 2$, although exactly similar sufficient conditions can be written down, by using the same technique, for any m (i.e., $m = 3, 4, \ldots$ etc). Finally, in Section 2.3, the result is applied to concrete examples.

2.2 Derivation of the prior

Let

$$h(w) = E_w(X) = \frac{-\beta'(w)}{\beta(w)} \qquad \ldots(2.1)$$

Using this notation,

$$E_w(X^2) = h^2(w) + h'(w)$$

and $$E_w(X^3) = h^3(w) + 3h(w)h'(w) + h''(w). \qquad \ldots(2.2)$$

Let $$\gamma(w) = c_o h'(w) + c_1 h(w) + c_2 h^2(w) \qquad \ldots(2.3)$$

be the parametric function to be estimated and

$$\delta(X) = a_o + a_1 X + a_2 X^2 \qquad \ldots(2.4)$$

its estimate.

Note that $\gamma(w)$ is slightly more general than $\gamma_2(w)$ introduced in Section 2.1

Let dG be a prior on Ω which is absolutely continuous with respect to the Lebesgue measure on Ω and let $\pi(w)$ denote the Radon-Nikodym derivative.

If $\delta(X)$ were to be generalized Bayes with respect to dG, for estimating $\gamma(w)$, then one has,

$$\delta(x) = \frac{\int \gamma(w) e^{xw} \beta(w) \pi(w) dw}{\int e^{xw} \beta(w) \pi(w) dw} \quad \text{(a.e. } d\mu) \qquad \ldots(2.5)$$

where the integrals are over $\Omega = (\underline{w}, \bar{w})$.

Let $\rho_o(w) = \beta(w)\pi(w)$, for $w \in \Omega$.

Hence, from (2.5),

$$\delta(x) \int e^{xw} \rho_o(w) dw = \int \gamma(w) e^{xw} \rho_o(w) dw$$

$$\Rightarrow (a_2 x^2 + a_1 x + a_o) \int e^{xw} \rho_o(w) dw = \int \gamma(w) e^{xw} \rho_o(w) dw \quad \text{(a.e. } d\mu)$$

$\ldots(2.6)$

Now, for $\underline{w} \leq a < b \leq \bar{w}$, integrating by parts,

$$\int_a^b e^{xw} \rho_o'(w)dw = \left[e^{xb} \rho_o(b) - e^{xa} \rho_o(a) \right] - x \int_a^b e^{xw} \rho_o(w)dw \qquad \ldots(2.7)$$

Similarly,

$$\int_a^b e^{xw} \rho_o''(w)dw = \left[e^{xb} \rho_o'(b) - e^{xa} \rho_o'(a) \right] - x \left[e^{xb} \rho_o(b) - e^{xa} \rho_o(a) \right] + x^2 \int_a^b e^{xw} \rho_o(w)dw \qquad \ldots(2.8)$$

Hence, if the quantities $\left[e^{x\bar{w}} \rho_o(\bar{w}) - e^{x\underline{w}} \rho_o(\underline{w}) \right]$ and $\left[e^{x\bar{w}} \rho_o'(\bar{w}) - e^{x\underline{w}} \rho_o'(\underline{w}) \right]$ are zero, then, from (2.6) one gets,

$$a_2 \int e^{xw} \rho_o''(w)dw - a_1 \int e^{xw} \rho_o'(w)dw + a_o \int e^{xw} \rho_o(w)dw = \int \gamma(w) e^{xw} \rho_o(w)dw \quad (\text{a.e. } d\mu) \qquad \ldots(2.9)$$

Using now the uniqueness property of the Laplace transform, one has,

$$a_2 \rho_o''(w) - a_1 \rho_o'(w) = (\gamma(w) - a_o) \rho_o(w) \quad \text{for} \quad w \in (\underline{w}, \bar{w}) \qquad \ldots(2.10)$$

It is clear from (2.10) that a solution to it for an arbitrary $\gamma(w)$ is very difficult to obtain if $a_2 \neq 0$, i.e., estimates are non-linear. In view of the specific form of $\gamma(w)$ under consideration, we suggest the "trial solution" given by

$$\rho_o(w) = e^{d_1 w + d_2 \int h(w)dw} \qquad \ldots(2.11)$$

where $h(w)$ can be taken as any differentiable function, not necessarily the mean, $\int h(w)dw$ is to be interpreted as a primitive, and d_1, d_2 are two constants to be suitably chosen later.

One therefore has,

$$\rho_0'(w) = (d_1 + d_2 h(w))\ \rho_0(w)$$

$$\text{and}\quad \rho_0''(w) = \left[(d_1 + d_2 h(w))^2 + d_2 h'(w) \right] \rho_0(w) \qquad \ldots(2.12)$$

Hence, from (2.10) and (2.12), one gets

$$a_2(d_1^2 + 2d_1 d_2 h + d_2^2 h^2) + a_2 d_2 h' - a_1(d_1 + d_2 h) + a_0 = c_2 h^2 + c_1 h + c_0 h' \qquad \ldots(2.13)$$

where the variable 'w' has been suppressed.

Clearly,

$$a_2 d_2^2 - c_2 = 0,\ 2a_2 d_1 d_2 - a_1 d_2 - c_1 = 0,\ a_2 d_1^2 - a_1 d_1 + a_0 = 0,$$

$$a_2 d_2 - c_0 = 0, \qquad \ldots(2.14)$$

are sufficient for $\rho_0(w)$ given by (2.11) to be a solution to (2.10). Henceforth, we shall refer to the system of equations (2.14) as the 'consistency conditions'.

2.3 Finding admissible estimates : the main result

Theorem 2.3.1 Let X have a density given by $p(x,w) = e^{xw}\beta(w)$, with respect to some σ-finite measure μ on the real-line. Let $\Omega = \left\{ w : \int e^{xw}\, d\mu(x) < \infty \right\}$ denote the natural parameter space. Consider any differentiable function $h(w)$ on Ω. Let

$\gamma(w) = c_2h^2(w) + c_1h_1(w) + c_oh'(w)$, where c_o, c_1, c_2 are any fixed constants. Let

$$\rho_o(w) = e^{d_1 w + d_2 \int h(w)\, dw},$$

where $d_1, d_2 \neq 0$ are some constants. Let $\pi(w) = \beta^{-1}(w)\, \rho_o(w)$, $w \in \Omega$. Let $\delta(X) = a_2X^2 + a_1X + a_o$, where a_o, a_1, a_2 are real numbers. Suppose a_o, a_1, a_2, d_1, d_2 subject to the consistency conditions (2.14) are such that

$$\text{(i)} \quad \int_b^{\bar{w}} \pi^{-1}(w)\, f^{-1}(w)dw = \infty \quad \text{for } \underline{w} < b < \bar{w}$$

and

$$\text{(ii)} \quad \int_{\underline{w}}^{c} \pi^{-1}(w)\, f^{-1}(w)dw = \infty \quad \text{for } \underline{w} < c < \bar{w}, \qquad \ldots(3.1)$$

where $f(w)$ is as in (3.10). Then $\delta(X)$ is an admissible estimator of $\gamma(w)$, under squared - error loss.

Proof : Suppose $\delta(X)$ is not admissible. Then there exists another estimator $\delta'(X)$ such that

$$E_w(\delta'(X) - \gamma(w))^2 \leq E_w(\delta(X) - \gamma(w))^2 \quad \forall\, w \in \Omega, \qquad \ldots(3.2)$$

with strict inequality for at least one 'w'. (3.2) can be rewritten as

$$\int (\delta'(x) - \gamma(w))^2\, e^{xw}\, \beta(w)d\mu(x) \leq \int (\delta(x) - \gamma(w))^2\, e^{xw}\beta(w)d\mu(x) \quad \forall\, w \in \Omega$$

$$\Longleftrightarrow \int (\delta'(x) - \delta(x))^2\, e^{xw}\beta(w)d\mu(x)$$

$$\leq 2\int (\delta(x) - \delta'(x))(\delta(x) - \gamma(w))\, e^{xw}\beta(w)d\mu(x) \quad \forall\, w \in \Omega, \quad \ldots(3.3$$

Hence, for $(a,b) \subseteq (\underline{w}, \bar{w})$,

$$\int_a^b \left\{ \int \left[\delta'(x) - \delta(x)\right]^2 e^{xw} \beta(w)d\mu(x)\right\} \pi(w)dw$$

$$\leq 2 \int_a^b \left\{ \int (\delta(x) - \delta'(x))(\delta(x) - \gamma(w))\ e^{xw} \beta(w)d\mu(x)\right\} \pi(w)dw$$

$$= 2 \int \left[\delta(x) - \delta'(x)\right] \left\{ \int_a^b \left[\delta(x) - \gamma(w)\right] e^{xw} \rho_o(w)dw \right\} d\mu(x), \quad \ldots(3.4)$$

on interchanging the order of integration.

The inner integral on the RHS of (3.4) is

$$\int_a^b [\delta(x) - \gamma(w)]\ e^{xw} \rho_o(w)dw$$

$$= (a_2x^2 + a_1x + a_o) \int_a^b e^{xw} \rho_o(w)dw - \int_a^b [c_2h^2(w) + c_1h(w) + c_oh'(w)] \times e^{xw} \rho_o(w)dw \quad \ldots(3.5)$$

$$= a_2 \int_a^b e^{xw} \rho_o''(w)dw - a_1 \int_a^b e^{xw} \rho_o'(w)dw + a_o \int_a^b e^{xw} \rho_o(w)dw$$

$$- a_2 [e^{xb} \rho_o'(b) - e^{xa} \rho_o'(a)] + a_2x [e^{xb} \rho_o(b) - e^{xa} \rho_o(a)]$$

$$+ a_1 [e^{xb} \rho_o(b) - e^{xa} \rho_o(a)]$$

$$- \int_a^b [c_2h^2(w) + c_1h(w) + c_oh'(w)] \rho_o(w)dw \quad \ldots(3.6)$$

(using (2.7) and (2.8))

$$= a_2 \int_a^b \left[d_1^2 + 2d_1 d_2 h(w) + d_2^2 h^2(w) + d_2 h'(w)\right] e^{xw} \rho_o(w)dw$$

$$- a_1 \int_a^b \left[d_1 + d_2 h(w)\right] e^{xw} \rho_o(w)dw + a_o \int_a^b e^{xw} \rho_o(w)dw$$

$$- \int_a^b \left[c_2 h^2(w) + c_1 h(w) + c_o h'(w)\right] e^{xw} \rho_o(w)dw$$

$$+ e^{xb} \rho_o(b) \left[a_2 x + a_1 - a_2 d_1 - a_2 d_2 h(b)\right]$$

$$- e^{xa} \rho_o(a) \left[a_2 x + a_1 - a_2 d_1 - a_2 d_2 h(a)\right] \qquad \ldots(3.7)$$

(using (2.12))

$$= \int_a^b \left[(a_2 d_1^2 - a_1 d_1 + a_o) + (2a_2 d_1 d_2 - a_1 d_2 - c_1)h(w) + (a_2 d_2^2 - c_2)h^2(w) + (a_2 d_2 - c_o)h'(w)\right] e^{xw} \rho_o(w)dw$$

$$+ e^{xb} \rho_o(b) \left[\alpha_1 x + \alpha_2 + \alpha_3 h(b)\right] - e^{xa} \rho_o(a) \left[\alpha_1 x + \alpha_2 + \alpha_3 h(a)\right], \quad \ldots(3.8)$$

(where $\alpha_1 = a_2$, $\alpha_2 = a_1 - a_2 d_1$, $\alpha_3 = - a_2 d_2$)

$$= \left[\alpha_1 x + \alpha_2 + \alpha_3 h(b)\right] e^{xb} \rho_o(b) - \left[\alpha_1 x + \alpha_2 + \alpha_3 h(a)\right] e^{xa} \rho_o(a), \qquad \ldots(3.9)$$

using the consistency conditions (2.14).

Setting $T(w) = \int \left[\delta(x) - \delta'(x)\right]^2 e^{xw} \beta(w) d\mu(x)$,

one has, from (3.4),

$$\int_a^b T(w)\pi(w)dw \leq K \left[\sqrt{\pi(b)f(b)} \cdot \sqrt{T(b)\pi(b)} + \sqrt{\pi(a)f(a)} \cdot \sqrt{T(a)\pi(a)}\right]$$

by Cauchy-Schwartz's inequality,

where K is a constant, and

$$f(w) = E_w(a_1X + a_2 + a_3h(w))^2 . \qquad \ldots(3.10)$$

The proof now follows in the lines of Karlin (58).

Remarks on Theorem 2.3.1

1. The proof of Theorem 2.3.1 depends heavily on the 'matching procedure' and the consistency conditions obtained thereof in (2.14). To show that the technique of proof works for any general m, we observe that the same matching procedure as in (2.14) can be carried out for every $m \geq 1$. For an arbitrary $m \geq 1$, the fundamental differential equation analogous to (2.10) would involve in the LHS functions $\rho_o^{(j)}/\rho_o$ for $j = 1,2,\ldots,m$. A careful analysis shows that $\rho_o^{(j)}/\rho_o$ can be expressed as the jth moment of a certain distribution whose first j cumulants are $d_1 + d_2k_1, d_2k_2, \ldots, d_jk_j$ respectively, where k_i's are cumulants of X. Since $\gamma(w)$ is a linear combination of the raw moments, the matching procedure can be carried out using the well-known relations between moments and cumulants, and theoretically, the consistency conditions analogous to (2.14) can be written down, although algebraically it becomes difficult with increasing m. For $m=3$, $\gamma_3(w) = c_oh'(w) + c_1h(w) + c_2h^2(w) + c_3h^3(w) + c_4h(w)h'(w) + c_5h''(w)$, and $\delta_3(X) = a_3X^3 + a_2X^2 + a_1X + a_o$, the consistency conditions are

$$a_3d_2^3 + c_3 = 0$$

$$-3a_3d_1d_2^2 + a_2d_2^2 - c_2 = 0$$

$$- 3a_3d_1^2d_2 + 2a_2d_1d_2 - a_1d_2 - c_1 = 0$$

$$- a_3d_1^3 + a_2d_1^2 - a_1d_1 + a_o = 0$$

$$- 3a_3d_1d_2 + a_2d_2 - c_o = 0$$

$$3a_3d_2^2 + c_4 = 0$$

$$a_3d_2 + c_5 = 0 \qquad \ldots(3.11)$$

An application is provided in Section 2.4.

2. In applications, if $h(w) = E_w(X)$, then h, h', and h^2 often are linearly related, and the consistency conditions (2.14) can be rewritten so that effectively the number of independent equations reduces and a larger class of admissible estimates can be obtained.

3. If $h(w)$ is taken as w in Theorem 2.3.1, $\gamma(w)$ becomes a polynomial in w for a general m. However, in all the standard one-parameter exponential set - ups, including the important normal case, $\pi(w)$ in this situation becomes proper and our theorem yields only proper Bayes admissible estimators.

4. For a completely arbitrary piece-wise continuous $\gamma(w)$, Theorem 2.3.1 does not provide any polynomial admissible estimates. The result of Ghosh and Meeden (77), however, follows from Theorem 2.3.1 on taking $a_2 = 0$, $c_2 = c_o = 0$, $c_1 = 1$. Differentiability of $h(w)$ is not necessary now as the coefficient of $h'(w)$ is taken as 0. Also, as $f(w)$ turns out to be a constant, (3.1) reduces exactly to the condition of Ghosh and Meeden (77).

5. Theorem 2.3.1 can also be used to generate a class of parametric functions for which a given polynomial in X is admissible. Each such $\gamma(w)$ is obtained by choosing the c_i's in a way that they satisfy the consistency conditions (2.14) and the divergence requirements (3.1).

2.4 Examples In this section, we use Theorem 2.3.1 to find polynomial admissible estimates of various parametric functions in some important one-parameter exponential set-ups.

Example 1 Let $X \sim$ Poisson (e^w), $-\infty < w < \infty$.

(a) Consider $\gamma(w) = b_2 e^{2w} + b_1 e^w$, where b_1, b_2 are real numbers.

Therefore, $\gamma(w) = c_2 h^2(w) + c_1 h(w) + c_o h'(w)$, with $h(w) = e^w$, $c_2 = b_2$, and $c_o + c_1 = b_1$.

Therefore, the consistency conditions (2.14) reduce to

$$a_2 d_2^2 - b_2 = 0$$

$$2a_2 d_1 d_2 - a_1 d_2 + a_2 d_2 - b_1 = 0$$

$$a_2 d_1^2 - a_1 d_1 + a_o = 0 . \qquad \text{...(4.1)}$$

In this case, $\rho_o(w) = e^{d_1 w + d_2 e^w} \Longrightarrow \pi(w) = e^{d_1 w + (1+d_2)e^w}, -\infty < w < \infty$.

Also, $f(w)$ defined in (3.10) is the mean of a quadratic in X, and therefore,

$$f(w) = ae^{2w} + be^w + c \quad \text{for some constants } a, b, c .$$

Now, $\pi^{-1}(w)f^{-1}(w) = \dfrac{e^{-d_1 w - (1+d_2)e^w}}{ae^{2w} + be^w + c} \longrightarrow \dfrac{1}{c}$ as $w \longrightarrow -\infty$

$$\longrightarrow \infty \text{ as } w \longrightarrow \infty, \quad \ldots(4.2)$$

if $d_1 = 0$, and $1 + d_2 < 0$.

It is interesting to note that $f(w)$ does not play any role in the divergence of $\int \pi^{-1}(w)f^{-1}(w)dw$. The same phenomenon occurs so long as $f(w)$ is the mean of any polynomial in X.

Now, also note that if $d_1 = 0$, $\pi(w) \longrightarrow 1$ as $w \longrightarrow -\infty$, implying π is improper. Also (4.2)guarantees the divergence conditions of Theorem 2.3.1 are satisfied.

Since $d_1 = 0$, (4.1) reduces to

$$a_2 d_2^2 - b_2 = 0$$

$$a_2 d_2 - a_1 d_2 - b_1 = 0$$

$$a_o = 0 \quad \ldots(4.3)$$

$$\Longrightarrow a_2 = \frac{b_2}{d_2^2}; \text{ also, } a_1 = a_2 - \frac{b_1}{d_2} = \frac{b_2}{d_2^2} - \frac{b_1}{d_2}. \quad \ldots(4.4)$$

Hence, $\delta(X) = \dfrac{b_2}{d_2^2} X^2 + \left(\dfrac{b_2}{d_2^2} - \dfrac{b_1}{d_2}\right) X$ is an admissible estimator of $\gamma(w) = b_2 e^{2w} + b_1 e^w$ for every $d_2 < -1$.

In particular, when $b_2 = 1$, $b_1 = 0$, $\gamma(w) = e^{2w} = E_w\{X(X-1)\}$.

It follows that $a(X^2 + X)$ is admissible for e^{2w} for every $0 < a < 1$. It is known from Ghosh and Meeden (77) that the linear estimate X is admissible for e^{2w}.

(b) We now give an example in the Poisson distribution when $m = 3$, and use the consistency conditions (3.11) in Remark 1 to obtain non-linear admissible estimates of $e^{3w} = E_w\{X(X-1)(X-2)\}$. As before, writing $h(w) = e^w$, and identifying $\gamma(w)$ as

$$\gamma_3(w) = c_o h'(w) + c_1 h(w) + c_2 h^2(w) + c_3 h^3(w) + c_4 h(w) h'(w) + c_5 h''(w),$$

where $c_2 + c_4 = 0$, $c_1 + c_5 + c_o = 0$, $c_3 = 1$,

(3.11) reduces to

$$\begin{aligned} a_o &= 0 \\ a_3 d_2^3 &= -1 \\ 3a_3 d_2^2 - a_2 d_2^2 &= 0 \\ a_1 d_2 + a_3 d_2 - a_2 d_2 &= 0 . \end{aligned} \qquad \ldots(4.5)$$

As before, if $d_1 = 0$ and $1 + d_2 < 0$, $\pi(w)$ is improper and the conditions of Theorem 2.3.1 are satisfied, since $f(w)$ remains the mean of a polynomial in X. While writing (4.5) using (3.11), d_1 has been set equal to 0.

Now, from (4.5),

$$a_2 = 3a_3 , \; a_1 = 2a_3 , \; a_3 = -\frac{1}{d_2^3} < 1 \;\; (\text{as } d_2 < -1), \;\; a_o = 0 \qquad \ldots(4.6)$$

Hence, $\delta(X) = a(X^3 + 3X^2 + 2X)$ is an admissible estimator of $E_w\{X(X-1)(X-2)\}$ for every $0 < a < 1$.

Example 2

Let $X \sim \text{Bin}(n, \frac{e^w}{1+e^w})$, $-\infty < w < \infty$.

Let $h(w) = E_w(X) = \frac{ne^w}{1+e^w}$ ...(4.7)

Then $h'(w) = \text{Var}_w(X) = h(w) - \frac{h^2(w)}{n}$ ...(4.8)

Let $\gamma(w) = \text{Var}_w(X) = h'(w)$.

From (4.8) it follows that $\gamma(w)$ can be written in the form

$$\gamma(w) = \frac{-h^2(w)}{n+\delta} + \frac{nh(w)}{n+\delta} + \frac{\delta h'(w)}{n+\delta}, \text{ for } 0 < \delta < 1. \qquad ...(4.9)$$

In this case, $\pi(w) = e^{d_1 w} \cdot (1+e^w)^{n(1+d_2)}$ ...(4.10)

$f(w)$, being the mean of a quadratic in X, is given by

$$f(w) = a\left(\frac{e^w}{1+e^w}\right)^2 + b\,\frac{e^w}{1+e^w} + c, \qquad ...(4.11)$$

for some constants a,b,c

Hence, from (4.10) and (4.11)

$$\pi^{-1}(w)f^{-1}(w) = \frac{e^{-d_1 w}\,(1+e^w)^{-n(1+d_2)}}{a\left(\frac{e^w}{1+e^w}\right)^2 + b\,\frac{e^w}{1+e^w} + c} \begin{array}{l} \longrightarrow \frac{1}{c} \text{ as } w \longrightarrow -\infty \\ \longrightarrow \infty \text{ as } w \longrightarrow \infty \end{array}$$

...(4.12)

if we let $d_1 = 0$, $1 + d_2 < 0$.

Also, from (4.10) it is clear that $\pi(w) \longrightarrow 1$ as $w \longrightarrow -\infty$ so that π is improper. (4.12) on the other hand shows the conditions of Theorem 2.3.1 hold.

Using (4.9), the consistency conditions boil down to

$$a_2 d_2^2 = -\frac{1}{n+\delta}, \quad a_2 d_2 = \frac{\delta}{n+\delta}, \; a_1 d_2 = -\frac{n}{n+\delta}, \; a_o = 0 \qquad \ldots(4.13)$$

Hence, from (4.13), $d_2 = -\frac{1}{\delta} < -1$ (as $0 < \delta < 1$), as is required in (4.12).

(4.13) now gives

$$a_2 = -\frac{\delta^2}{n+\delta}, \quad a_1 = \frac{n\delta}{n+\delta} \qquad \ldots(4.14)$$

Therefore, $\delta(X) = -\frac{\delta^2}{n+\delta} X^2 + \frac{n\delta}{n+\delta} X$

$$= \frac{\delta}{n+\delta}(nX - \delta X^2)$$

is an admissible estimator of $\mathrm{Var}_w(X)$ for every $0 < \delta < 1$.

It may be mentioned in this regard that no linear estimate satisfies the sufficient conditions of Ghosh and Meeden (77) if $\mathrm{Var}_w(X)$ is to be estimated.

Example 3

Let X have the Negative Binomial distribution given by

$$P_w(X = x) = \binom{r+x-1}{x} p^r q^x, \quad x = 0,1,2,\ldots, r > 1 \text{ known}, \; 0 < p < 1.$$

To express it in the standard one-parameter exponential set up, we equivalently write $X \sim NB(e^w)$, $-\infty < w < 0$.

Let $h(w) = E_w(X) = \frac{r\,e^w}{1-e^w}$ ...(4.15)

Then $h'(w) = \mathrm{Var}_w(X) = \frac{r\,e^w}{(1-e^w)^2} = h(w) + \frac{h^2(w)}{r}$ ...(4.16)

We estimate $\gamma(w) = \mathrm{Var}_w(X) = h'(w)$.

Hence, using (4.16),

$$\gamma(w) = c_2 h^2(w) + c_1 h(w) + c_o h'(w)$$

$$= (c_2 - \frac{c_1}{r})h^2(w) + (c_1 + c_o)h'(w) \qquad \text{...(4.17)}$$

where $c_2 - \frac{c_1}{r} = 0$, $c_1 + c_o = 1$.

Here, $\pi(w) = e^{d_1 w}\,.\,(1 - e^w)^{-r(1+d_2)}$ ...(4.18)

As in the previous examples, $f(w)$ is the mean of a quadratic in X, say,

$$f(w) = a.\frac{re^w}{1-e^w} + b.\,\frac{re^w}{(1-e^w)^2} + c.\,\frac{re^{2w}}{(1-e^w)^2} + d \qquad \text{...(4.19)}$$

From (4.18) and (4.19),

$$\pi^{-1}(w)f^{-1}(w) = \frac{(1-e^w)^{r(1+d_2)}\;e^{-d_1 w}}{a\,\frac{re^w}{1-e^w} + b\,\frac{re^w}{(1-e^w)^2} + c\,\frac{re^{2w}}{(1-e^w)^2} + d}$$

$$\longrightarrow \frac{1}{d} \text{ as } w \longrightarrow -\infty,$$

$\int_{\omega_0}^{b} \pi^{-1}(\omega)\,f^{-1}(\omega)\,d\omega \longrightarrow \infty$ as $w_0 \longrightarrow 0$, ...(4.20)

if $d_1 = 0$, and $r(1 + d_2) + 3 < 0$.

From (4.17), the consistency conditions determining the coefficients a_o, a_1, a_2 are

$$a_2 d_2^2 + \frac{a_1 d_2}{r} = 0,\ a_2 d_2 - a_1 d_2 = 1,\ a_o = 0 \qquad \ldots(4.21)$$

From (4.21), one has,

$$a_2 = -\frac{a_1}{r d_2}$$

$$\Longrightarrow -a_1\left(\frac{1}{r d_2} + 1\right) = \frac{1}{d_2}$$

$$\Longrightarrow a_1 = -\frac{r}{1 + r d_2} > 0 .$$

Since $d_2 < -1 - \frac{3}{r}$, if follows that $a_1 < \frac{r}{r+2}$. $\qquad \ldots(4.22)$

Also, $a_2 = a_1 + \frac{1}{d_2} = \frac{1}{d_2(1 + r d_2)} < \frac{r}{(r+3)(r+2)}$ $\qquad \ldots(4.23)$

From (4.22) and (4.23) we have,

$\delta(X) = a_2 X^2 - a_1 X$ is admissible for $\mathrm{Var}_w(X)$ for every $0 < a_2 < \frac{r}{(r+1)(r+2)}$, $0 < a_1 < \frac{r}{r+1}$.

Example 4

In all the preceding examples $h(w)$ was taken as $E_w(X)$ and d_1 as 0, although in the proof of the theorem, $h(w)$ can be taken as any differentiable function. In this example, these conditions are removed.

Let $X \sim \mathrm{Bin}(n, \frac{e^w}{1+e^w})$, $-\infty < w < \infty$.

Let $\gamma(w) = e^{w}$ (i.e., we are now estimating the odds $\frac{p}{q}$ of heads against tails).

Let $h(w) = e^{w/2}$. $\gamma(w)$ can then be written as

$$\gamma(w) = c_2 h^2(w) + c_1 h(w) + c_o h'(w),$$

where $c_2 = 1$, $c_1 + c_o/2 = 0$ ...(4.24)

Here, $$\pi(w) = \frac{(1+e^{w})^{n} \cdot e^{2d_2 e^{w/2}}}{e^{w/2}}$$ ...(4.25)

if we let $d_1 = -\frac{1}{2}$.

Note that $\pi(w) \longrightarrow \infty$ as $w \longrightarrow -\infty$, so that $\pi(w)$ is an improper prior.

Also, by definition,

$f(w) = E_w(a_2 X - a_2 d_2\, e^{w/2})^2$ (if we choose a_1, a_2 so that $a_1 - a_2 d_1 = a_1 + a_2/2 = 0$)

$$= a_2^2 \left[\frac{ne^{w}}{(1+e^{w})^2} + \frac{n^2 e^{2w}}{(1+e^{w})^2} - \frac{2nd_2 e^{3w/2}}{1+e^{w}} + d_2^2\, e^{w} \right]$$

$$= \frac{a_2^2}{(1+e^{w})^2} \Big[(n + d_2^2)\, e^{w} + (n^2 + 2d_2^2)\, e^{2w} + d_2^2\, e^{3w} - 2nd_2\, e^{3w/2} - 2nd_2\, e^{5w/2} \Big]$$

$$= \frac{a_2^2\, e^{w}}{(1+e^{w})^2} \Big[(n + d_2^2) - 2nd_2 e^{w/2} + (n^2 + 2d_2^2) e^{w} - 2nd_2 e^{3w/2} + d_2^2 e^{2w} \Big]$$

...(4.26)

Hence, from (4.25) and (4.26),

$$\pi^{-1}(w)f^{-1}(w) = \frac{(1+e^w)^2 \, e^{-2d_2 e^{w/2}} \quad e^{w/2}}{a_2^2 \, e^w (1+e^w)^n \left[(n+d_2^2) - 2nd_2 e^{w/2} + (n^2+2d_2^2)e^w - 2nd_2 e^{3w/2} + d_2^2 \, e^{2w} \right]}$$

$$\longrightarrow \infty \quad \text{as} \quad w \longrightarrow \pm \infty \qquad \ldots(4.27)$$

if $d_2 < 0$.

(4.27) ensures that the divergence conditions of Theorem 2.3.1 are satisfied.

From (4.24), the consistency conditions reduce to

$$a_2 d_2^2 = 1 ,$$

$$- a_2 d_2 + a_2 d_2/2 + a_2 d_2/2 = 0 ,$$

$$a_2/4 - a_2/4 + a_o = 0 \qquad \ldots(4.28)$$

(since $a_1 + a_2/2$ was taken as 0 in (4.26)).

In (4.28) the last two consistency conditions are automatically satisfied if $a_o = 0$.

Hence, $\delta(X) = a(2X^2 - X)$ is admissible for $\gamma(w) = e^w$, for every $a > 0$.

CHAPTER 3

ADMISSIBILITY OF GENERALIZED BAYES AND PITMAN ESTIMATES IN THE NON-REGULAR FAMILY

3.1 Introduction

In Chapter 2, it was shown how Karlin's technique can be exploited to obtain admissible estimates of many parametric functions in the one-parameter regular exponential family. In this chapter we shall consider in detail the question of finding admissible estimates for a fairly general class of parametric functions in the so called "non-regular" type of densities. Let X have a density (with respect to Lebesgue measure) of the form $p(x,w) = r(x)q(w)$, where $\underline{w} < x \leq w$, and $w \in (\underline{w}, \bar{w})$, which may be an infinite interval. It is clear that the support of $p(x,w)$ depends on w, and in fact $P_w(X \leq w) = 1$, for all $w \in (\underline{w}, \bar{w})$. It is well known that $q(w)$ is a monotone decreasing function of w and $q(\underline{w}) = \infty$. Also, $q(x)$ has a derivative, and $r(x) = -q'(x)/q^2(x)$. If $X_1,\ldots,X_n$ are iid with a uniform distribution on $[0,w]$, then $X_{(n)}$, the maximum of $X_1,\ldots,X_n$, is sufficient for w, and its sampling distribution is of the form stated in the preceding paragraph. Primarily with the objective of estimating 'w' on the basis of a linear estimate $\gamma.X_{(n)}$, Karlin (58) considered the problem of finding a general admissible estimate of $q^{-\alpha}(w)$ for $\alpha > 0$, and showed that $\frac{2\alpha+1}{\alpha+1} q^{-\alpha}(X)$, the best estimator of $q^{-\alpha}(w)$ of the form $\gamma.q^{-\alpha}(X)$, is admissible for $q^{-\alpha}(w)$, for every $\alpha > 0$.

Karlin, however, did not address the problem of finding an admissible estimate for more general parametric functions, like the mean. It is quite easy to see that the mean is a monotone function of the parameter 'w', just like Karlin's parametric functions are. In fact, quantiles and all moments of a non-regular density are monotone functions of w. Question arises if it is not possible to find admissible estimates for all these parametric functions. It has been shown in this chapter that Karlin's technique of finding admissible estimates is quite general, and for monotone increasing (decreasing) functions of w, which are also non-negative, Karlin's argument gives us sufficient conditions for admissibility of generalized Bayes estimates.

It was pointed out in Chapter 2 that deriving the form of a crucial prior is an essential step of Karlin's analysis. Towards this end, if $h(w)$ is any non-negative increasing function of w, we have considered (possibly improper) priors of the form $\pi_f(w) = h'(w)f(h(w))/q(w)$, where f is a non-negative function defined on the range of h. Since h is monotone, it is almost everywhere differentiable. For the sake of simplicity, we have assumed that $h'(w)$ exists everywhere. The prior under consideration has been assumed absolutely continuous, with the density described above. Throughout we work on squared-error loss.

In the following sections, $\delta_f(x)$ will stand for the generalized Bayes estimate of $h(w)$ with respect to the prior $\pi_f(w)$. It is easy to see that $\delta_f(x)$ takes the form

$$\delta_f(x) = \int_{h(x)}^{h(\bar{w})} uf(u)du \Big/ \int_{h(x)}^{h(\bar{w})} f(u)du \qquad \ldots(1.1)$$

From (1.1) it is clear that certain integrability conditions would have to be imposed on f for $\delta_f(x)$ being well defined. But for this, π_f is quite general. Sufficient conditions on f are derived for $\delta_f(X)$ to be admissible.

If, in particular, $f(u) = u^{-2-\varepsilon}$, $\varepsilon > 0$, is considered, then (1.1) reduces to $\delta_f(x) = \frac{1+\varepsilon}{\varepsilon} h(x)$. For obvious reasons, we have called these " substitution estimates ". To get an insight into the more general problem when f is any non-negative function, in Section 3.2 we have derived sufficient conditions for the admissibility of " substitution estimates " for a non-negative increasing function. In particular, the result of Karlin (58) follows from this. Theorem 3.2.1 is a special case of Theorem 3.3.1 in the next section.

In Section 3.3, we have dealt with the problem of finding sufficient conditions on f for δ_f to be admissible, when $h \geq 0$ and increasing. Sufficient conditions on f can be obtained in the same spirit when $h \geq 0$ and decreasing.

Finally, in Section 3.4, we have considered the problem of estimating $\log q(w)$. This is also a differentiable monotone function of w, but not necessarily non-negative. It has been proved that $\log q(X) - 1$, the Pitman estimate of $\log q(w)$, is an admissible estimate. An application has been made to the estimation of the mean of a truncated exponential.

3.2 Admissibility of substitution estimates

Theorem 3.2.1 Let X have a density of the form

$$p(x,w) = r(x)q(w) \;, \quad \underline{w} < x \leq w \;,$$
$$= 0 \qquad \text{otherwise} \;, \qquad \ldots(2.1)$$

where $w \in (\underline{w}, \bar{w})$.

Let $h(w)$ be any non-negative increasing function of w, everywhere differentiable. Suppose $h(\bar{w}) = \lim_{w \to \bar{w}} h(w) = \infty$.

If there exists an $\varepsilon > 0$ such that

$$\text{i)} \quad \int_{\underline{w}}^{a} h'(w)h^{\varepsilon-2}(w)q(w)dw = \infty$$

$$\text{ii)} \quad \int_{b}^{\bar{w}} h'(w)h^{\varepsilon-2}(w)q(w)dw = \infty \qquad \ldots(2.2)$$

for $\underline{w} < a$, $b < \bar{w}$,

then $\delta(x) = \frac{1+\varepsilon}{\varepsilon} h(x)$ is an admissible estimate of $h(w)$.

Proof : The proof closely resembles that of Karlin (58).

If $\delta(X)$ is not admissible, there exists another estimator $\delta'(X)$ such that

$$E_w[\delta'(X) - h(w)]^2 \leq E_w[\delta(X) - h(w)]^2 \;\; \forall \, w \in (\underline{w}, \bar{w})$$

(strict inequality for at least one w)

$\ldots(2.3)$

(2.3) can be rewritten as

$$\int_{\underline{w}}^{w} [\delta'(x) - h(w)]^2 r(x)q(w)dx \le \int_{\underline{w}}^{w} [\delta(x) - h(w)]^2 r(x)q(w)dx \quad \forall\, w \in (\underline{w}, \bar{w})$$

$$\Longleftrightarrow \int_{\underline{w}}^{w} [\delta'(x) - \delta(x)]^2 r(x)q(w)dx \le 2\int_{\underline{w}}^{w} [\delta(x) - \delta'(x)]\,[\delta(x) - h(w)] \times r(x)q(w)dx \quad \forall\, w \in (\underline{w}, \bar{w}) \qquad \ldots(2.4)$$

Therefore, if $\pi(w)$ is any prior on $(\underline{w}, \bar{w})$, and $\underline{w} < a < b < \bar{w}$, then (2.4) implies

$$\int_{a}^{b}\int_{\underline{w}}^{w} [\delta'(x) - \delta(x)]^2 r(x)q(w)\pi(w)dxdw$$

$$\le 2\int_{a}^{b}\int_{\underline{w}}^{w} [\delta(x) - \delta'(x)]\,[\delta(x) - h(w)]\, r(x)q(w)\pi(w)dxdw \qquad \ldots(2.5)$$

Interchanging the order of integration, we have,

RHS of (2.5)

$$= 2\int_{a}^{b} [\delta(x) - \delta'(x)]\, r(x)\left\{\int_{x}^{b} [\delta(x) - h(w)]\, q(w)\pi(w)dw\right\} dx$$

$$+ \int_{\underline{w}}^{a} [\delta(x) - \delta'(x)]\, r(x)\left\{\int_{a}^{b} [\delta(x) - h(w)]\, q(w)\pi(w)dw\right\} dx \qquad \ldots(2.6)$$

because if $\underline{w} < x \le a$, then $a = \max(x,a) \le w \le b$, and if $a < x \le b$, then $x = \max(x,a) \le w \le b$.

In (2.6), write

$$\int_a^b [\delta(x) - \delta'(x)]\, r(x) \left\{ \int_x^b [\delta(x) - h(w)]\, q(w)\pi(w)dw \right\} dx$$

$$= \int_{\underline{w}}^b [\delta(x) - \delta'(x)]\, r(x) \left\{ \int_x^b [\delta(x) - h(w)]\, q(w)\pi(w)dw \right\} dx$$

$$- \int_{\underline{w}}^a [\delta(x) - \delta'(x)]\, r(x) \left\{ \int_x^b [\delta(x) - h(w)]\, q(w)\pi(w)dw \right\} dx \qquad \ldots(2.7)$$

Therefore, combining (2.5), (2.6), and (2.7), one has,

$$\int_a^b \int_{\underline{w}}^w [\delta'(x) - \delta(x)]^2 r(x)q(w)\pi(w)dxdw$$

$$\leq 2 \left[\int_{\underline{w}}^b [\delta(x) - \delta'(x)]\, r(x) \left\{ \int_x^b [\delta(x) - h(w)]\, q(w)\pi(w)dw \right\} dx \right.$$

$$\left. - \int_{\underline{w}}^a [\delta(x) - \delta'(x)]\, r(x) \left\{ \int_x^a [\delta(x) - h(w)]\, q(w)\pi(w)dw \right\} dx \right] \qquad \ldots(2.8)$$

Note that till (2.8), $\delta(x)$, $\delta'(x)$, $h(w)$, $\pi(w)$ have been completely arbitrary, and the hypothesis of this theorem has not been used.

Now consider the first term on the RHS of (2.8). The inner integral in this term is

$$\int_x^b [\delta(x) - h(w)]\, q(w)\pi(w)dw = \phi_b(x) \quad \text{(say)} \qquad \ldots(2.9)$$

Now use the form of the prior $\pi(w) = h'(w)f(h(w))/q(w)$ introduced in Section 3.1. Since $f(u)$ is taken as $f(u) = u^{-2-\varepsilon}$, $u > 0$, $\varepsilon > 0$, one has,

$$\pi(w) = \frac{h'(w)h^{-2-\varepsilon}(w)}{q(w)} , \; w \in (\underline{w}, \overline{w}) \qquad \ldots(2.10)$$

Now $$\psi_b(x) = \frac{1+\varepsilon}{\varepsilon} h(x) \int_x^b h'(w)h^{-2-\varepsilon}(w)dw - \int_x^b h'(w)h^{-1-\varepsilon}(w)dw$$

$$= \frac{1}{\varepsilon} h(x) [h^{-1-\varepsilon}(x) - h^{-1-\varepsilon}(b)] + \frac{1}{\varepsilon} [h^{-\varepsilon}(b) - h^{-\varepsilon}(x)]$$

$$= \frac{h^{-\varepsilon}(b)}{\varepsilon} - h(x) \frac{h^{-1-\varepsilon}(b)}{\varepsilon}$$

$$= \frac{h^{-\varepsilon}(b)}{\varepsilon} \left[1 - \frac{h(x)}{h(b)}\right] \qquad \ldots(2.11)$$

Since h is non-negative and increasing, from (2.11) it follows that

$$0 \leq \psi_b(x) \leq \frac{h^{-\varepsilon}(b)}{\varepsilon} \qquad \ldots(2.12)$$

Similarly the inner integral in the second term of the RHS of (2.8) has the bound

$$0 \leq \psi_a(x) \leq \frac{h^{-\varepsilon}(a)}{\varepsilon} \qquad \ldots(2.13)$$

Hence, from (2.8),

$$\int_a^b \int_{\underline{w}}^{w} [\delta'(x) - \delta(x)]^2 r(x)q(w)\pi(w)dxdw$$

$$\leq \frac{2}{\varepsilon} \left[\int_{\underline{w}}^{b} |\delta(x) - \delta'(x)| r(x) \, h^{-\varepsilon}(b)dx + \int_{\underline{w}}^{a} |\delta(x) - \delta'(x)| r(x)h^{-\varepsilon}(a)dx \right]$$

$$\leq \frac{2}{\varepsilon}\left[\left\{\int_{\underline{w}}^{b}|\delta(x)-\delta'(x)|^2 r(x)dx\right\}^{1/2}\left\{\int_{\underline{w}}^{b} r(x)dx\right\}^{1/2} h^{-\varepsilon}(b)\right.$$

$$\left.+\left\{\int_{\underline{w}}^{a}|\delta(x)-\delta'(x)|^2 r(x)dx\right\}^{1/2}\left\{\int_{\underline{w}}^{a} r(x)dx\right\}^{1/2} h^{-\varepsilon}(a)\right] \quad \ldots(2.14)$$

Defining $T(w) = \int_{\underline{w}}^{w} |\delta(x) - \delta'(x)|^2 r(x)q(w)dx$, for $w \in (\underline{w}, \bar{w})$,

(2.14) gives

$$\int_a^b T(w)\pi(w)dw \leq K\left[\sqrt{T(b)\pi(b)} \Big/ \sqrt{h'(b)h^{\varepsilon-2}(b)q(b)}\right.$$

$$\left. + \sqrt{T(a)\pi(a)} \Big/ \sqrt{h'(a)h^{\varepsilon-2}(a)q(a)}\right] \quad \ldots(2.15)$$

(in (2.14), note that $\left\{\int_{\underline{w}}^{b} r(x)dx\right\}^{1/2} = \frac{1}{\sqrt{q(b)}}$, and then multiply and divide by $\sqrt{\pi(b)}$ and use (2.10)).

Proceeding now as in Karlin (58), the proof is complete.

<u>Remarks on Theorem 3.2.1</u>

1. A major consequence of Theorem 3.2.1 is the result of Karlin (58). Taking $h(w) = q^{-\alpha}(w)$, and $\varepsilon = (\alpha+1)/\alpha$, one has

$$\int_{\underline{w}}^{a} h'(w)h^{\varepsilon-2}(w)q(w)dw = -\alpha \int_{\underline{w}}^{a} \frac{q'(w)}{q(w)} dw = \infty \quad \text{as} \quad q(\underline{w}) = \infty .$$

Similarly, $\int_b^{\bar{w}} h'(w)h^{\varepsilon-2}(w)q(w)dw = \infty$ if $q(\bar{w}) = 0$.

Hence, $\delta(x) = \frac{1+\varepsilon}{\varepsilon} h(x) = \frac{2\alpha+1}{\alpha+1} q^{-\alpha}(X)$ is an admissible estimator of $q^{-\alpha}(w)$.

2. In the context of Theorem 3.2.1, we recall a similar result due to Sinha, Ghosh, and Banerjee (78) that if X is a discrete variable with a p.m.f. $\psi(x,\theta) = q(\theta)r(x) > 0$ for $x = d, d+1, \ldots, \theta$, where $\theta \in \text{(H)} = \{d, d+1, \ldots\}$ for some $-\infty < d < \infty$ then $g(X)$ is admissible for $g(\theta)$, where g is any measurable-function.

3. The " cut-off " point w is itself an interesting parametric function. Theorem 3.2.1 shows that if $\int w^{\varepsilon-2} q(w)dw$ is ∞ on both the tails, then $\frac{1+\varepsilon}{\varepsilon} \cdot X$ is an admissible estimator of w.

4. In a general non-regular density, the mean of X is given as

$$m(w) = E_w(X) = w - q(w) \int_0^w \frac{1}{q(x)}\, dx ,$$

if $\underline{w}$ is taken as 0. Because of this complicated form, it is difficult to obtain a general admissible estimate of the mean from Theorem 3.2.1.

However, denoting $\int_0^w \frac{1}{q(x)}\, dx$ by $I(w)$, it follows that if either $\liminf_{w \to \underline{w}} q^2(w)I(w)$ and $\liminf_{w \to \bar{w}} q^2(w)I(w)$ are positive, and $q(\bar{w}) = 0$, or if $\limsup_{w \to \underline{w}} q^2(w)I(w)$ and $\limsup_{w \to \bar{w}} q^2(w)I(w)$ are finite, and $\int_{\underline{w}}^{a} q(w)dw = \infty = \int_b^{\bar{w}} q(w)dw$, then $\frac{3}{2}m(X)$ is in general admissible for $E_w(X)$. Note that for the $R(0,w)$ distribution, both these conditions are satisfied. Also, the condition that $q^2(w)I(w)$ tends to a non-zero finite limit can be rewritten as $w - m(w) = O(\frac{1}{q(w)})$ as $w \to \underline{w}, \bar{w}$.

5. For the special case $h(w) = q^{-a}(w)$, if one takes an observation X from the distribution of $\max(X_1, \ldots, X_n)$ where $X_1, \ldots, X_n$ form a random sample from the original density $p(x,w) = r(x)q(w)(x \leq w)$, then a new multiple of $h(X)$ continues to be admissible for $h(w)$, i.e., if $\exists\, \varepsilon > 0$ is such that $\int h'(w)h^{\varepsilon-2}(w)q(w)dw = \infty$ at the two tails, one can find another $\varepsilon_1 > 0$ such that $\int h'(w)h^{\varepsilon_1-2}(w)q^n(w)dw = \infty$ at the two tails, and hence $\frac{1+\varepsilon_1}{\varepsilon_1} h(X)$ is admissible for $h(w)$.

<u>Example</u>

We now give an application of Theorem 3.2.1 in the estimation of quantiles in a specific distribution. In the important rectangular and truncated exponential distributions, the quantiles are either multiples or shifts of w.

Let X be as in Theorem 3.2.1 with $q(w) = \dfrac{1}{\log(1+w)}$, $0 < w < \infty$.

We estimate the parametric functions $h(w) = (1+w)^p$, $0 < p < 1$, which, but for a constant, are the quantiles of the distribution of X. Note that,

$$\begin{aligned}
\int_0^a h'(w)h^{\varepsilon-2}(w)q(w)dw &= (\text{constant}) \int_0^a (1+w)^{p-1}(1+w)^{p(\varepsilon-2)} \frac{1}{\log(1+w)}dw \\
&= (\text{constant}) \int_0^a (1+w)^{p(\varepsilon-1)-1} \frac{1}{\log(1+w)}dw \\
&\geq (\text{constant}) \int_0^a \frac{1}{w}dw \\
&= \infty .
\end{aligned}$$

Also, $\int_b^\infty h'(w)h^{\varepsilon-2}(w)q(w)dw = (\text{constant}) \int_b^\infty (1+w)^{p(\varepsilon-1)-1} \frac{1}{\log(1+w)}\, dw$

$$= \infty,$$

if $p(\varepsilon-1) \geq 1$.

Hence, $\frac{1+\varepsilon}{\varepsilon}(1+X)^p$ is an admissible estimator of $(1+w)^p$

if $p(\varepsilon-1) \geq 1$, i.e., $\varepsilon \geq \frac{p+1}{p}$.

3.3 Admissibility of generalized Bayes estimates.

We now proceed to obtain sufficient conditions for admissibility of generalized Bayes estimates with respect to general priors $\pi_f(w)$ introduced in Section 3.1. Although (2.8) holds in general, inequalities corresponding to (2.12) and (2.13) require more difficult argument. Lemma 3.3.2 serves this purpose. Theorem 3.2.1 follows from the following theorem by taking $f(u) = u^{-2-\varepsilon}$, $\varepsilon > 0$, $u > 0$.

Theorem 3.3.1. Let X be as in Theorem 3.2.1. Let $h(w)$ be any non-negative increasing function of w, everywhere differentiable. Let $f \geq 0$ be defined on $[0,\infty)$ such that

$\int_a^{h(\bar{w})} f(u)du$ and $\int_a^{h(\bar{w})} u\, f(u)du$ are finite for every $a > 0$.

Then the following is a sufficient condition for the admissibility of the generalized Bayes estimate $\delta_f(x)$ for estimating $h(w)$:

$$\int_b^{\bar{w}} \frac{h'(w)f(h(w))q(w)}{\left(\int_{h(w)}^{h(\bar{w})} uf(u)du\right)^2}\, dw = \infty = \int_{\underline{w}}^{a} \frac{h'(w)f(h(w))q(w)}{\left(\int_{h(w)}^{h(\bar{w})} uf(u)du\right)^2}\, dw \qquad \ldots(2.16)$$

for $\underline{w} < a$, $b < \bar{w}$.

We state below a well known fact in the form of a lemma to be subsequently used in the proof of Theorem 3.3.1.

Lemma 3.3.2. Let $f \geq 0$ defined on $[0,\infty)$ be such that $\int_0^\delta f(u)du$, $\int_0^\delta uf(u)du$ are finite for every $0 < \delta < \infty$. Let $0 < a < \infty$ be fixed. Let $\phi(c) = \int_c^a uf(u)du \Big/ \int_c^a f(u)du$. Then $\phi(c)$ increases with c for $c < a$.

Proof of Theorem 3.3.1. As pointed out before, (2.8) holds without any change, δ denoting δ_f. Define

$$\phi_b(x) = \int_x^b [\delta(x) - h(w)]\, q(w)\pi(w)dw \qquad \ldots(2.17)$$

Then, $\phi_b(x) = \left(\int_{h(x)}^{h(\bar{w})} uf(u)du \Big/ \int_{h(x)}^{h(\bar{w})} f(u)du\right) \cdot \int_{h(x)}^{h(b)} f(u)du - \int_{h(x)}^{h(b)} uf(u)du$

(using (1.1) and the form of $\pi(w) = \pi_f(w)$)

$$= \int_{h(b)}^{h(\bar{w})} uf(u)du - \left(\int_{h(x)}^{h(\bar{w})} uf(u)du \Big/ \int_{h(x)}^{h(\bar{w})} f(u)du\right) \cdot \int_{h(b)}^{h(\bar{w})} f(u)du \qquad \ldots(2.18)$$

By Lemma 3.3.2, the facts that $h(b) \geq h(x)$, and h, f are non-negative, it follows that

$$0 \leq \phi_b(x) \leq \int_{h(b)}^{h(\bar{w})} uf(u)du \qquad \ldots(2.19)$$

Similarly, one has,

$$0 \leq \psi_a(x) \leq \int_{h(a)}^{h(\bar{w})} uf(u)du \qquad \ldots(2.20)$$

Hence, from (2.8),

$$\int_a^b \int_{\underline{w}}^{w} [\delta'(x) - \delta(x)]^2 r(x)q(w)\pi(w)dxdw$$

$$\leq 2\left[\int_{\underline{w}}^{b} |\delta(x) - \delta'(x)| r(x) \left(\int_{h(b)}^{h(\bar{w})} uf(u)du\right)dx + \int_{\underline{w}}^{a} |\delta(x) - \delta'(x)| r(x) \times \left(\int_{h(a)}^{h(\bar{w})} uf(u)du\right)dx\right] \qquad \ldots(2.21)$$

Using Cauchy-Schwartz's inequality exactly as in (2.14), and defining $T(w)$ once again as $\int_{\underline{w}}^{w} |\delta(x) - \delta'(x)|^2 r(x)q(w)dx$, for $w \in (\underline{w}, \bar{w})$, the proof is complete in the lines of Karlin (58) and Theorem 3.2.1.

Remark : Broadly in the same spirit, the case when h is non-negative and decreasing can also be handled. However, we do not go into it.

3.4 Admissibility of the Pitman estimate.

In this section, we consider the question of estimating $\log q(w)$ under squared-error loss. It is easy to see that $\log q(w)$ is a location parameter for the distribution of $\log q(X)$. There exists a best estimate of $\log q(w)$ of the form $\log q(X) + C$, and it is given by $\delta(X) = \log q(X) - 1$. Once again, following Karlin's technique, we have shown that $\delta(X)$ is an admissible estimator of $\log q(w)$. An application is subsequently made.

Theorem 3.4.1. $\log q(X) - 1$ is an admissible estimator of $\log q(w)$, under squared-error loss, provided $q(\bar{w}) = 0$.

Proof : As in (2.8), if $\delta(X) = \log q(X) - 1$ is not admissible, then there exists another estimator such that

$$\int_a^b \int_{\underline{w}}^{w} [\delta(x) - \delta'(x)]^2 r(x) q(w) \pi(w) dx dw$$

$$\leq 2 \Big[\int_{\underline{w}}^{b} [\delta(x) - \delta'(x)] r(x) \Big\{ \int_x^b [\log q(x) - 1 - \log q(w)] q(w) \pi(w) dw \Big\} dx$$

$$- \int_{\underline{w}}^{a} [\delta(x) - \delta'(x)] r(x) \Big\{ \int_x^a [\log q(x) - 1 - \log q(w)] q(w) \pi(w) dw \Big\} dx \Big] \quad \ldots(4.1)$$

Taking $\pi(w) = -\frac{q'(w)}{q(w)}$, we now proceed to simplify the RHS of (4.1).

Note that $\int_x^b [\log q(x) - 1 - \log q(w)] q(w) \pi(w) dw$

$$= - \int_x^b [\log q(x) - 1 - \log q(w)] q'(w) dw$$

$$= [\log q(x) - 1] [q(x) - q(b)] + \int_x^b \log q(w) \cdot q'(w) dw$$

$$= [\log q(x) - 1][q(x) - q(b)] + q(w)\log q(w)\Big|_x^b + [q(x) - q(b)]$$

$$= q(b)[\log q(b) - \log q(x)] \qquad \ldots(4.2)$$

Similarly, $\int_x^a [\log q(x) - 1 - \log q(w)]\, q(w)\pi(w)dw$

$$= q(a)[\log q(a) - \log q(x)] \qquad \ldots(4.3)$$

By letting $T(w) = \int_{\underline{w}}^{w} [\delta(x) - \delta'(x)]^2 r(x)q(w)dx$, and applying Schwartz's inequality at (4.1), we now have,

$$\int_a^b T(w)\pi(w)dw \le 2\Big[\sqrt{T(b)\pi(b)}\sqrt{\frac{q(b)}{\pi(b)}}\Big(\int_{\underline{w}}^{b}\Big|\log\frac{q(b)}{q(x)}\Big|^2 r(x)dx\Big)^{1/2}$$

$$+ \sqrt{T(a)\pi(a)}\cdot\sqrt{\frac{q(a)}{\pi(a)}}\Big(\int_{\underline{w}}^{a}\Big|\log\frac{q(a)}{q(x)}\Big|^2 r(x)dx\Big)^{1/2}\Big]$$

$$= 2\Big[\sqrt{T(b)\pi(b)}\sqrt{-\frac{q^2(b)}{q'(b)}}\Big(\int_{\underline{w}}^{b}\Big|\log\frac{q(b)}{q(x)}\Big|^2 r(x)dx\Big)^{1/2}$$

$$+ \sqrt{T(a)\pi(a)}\sqrt{-\frac{q^2(a)}{q'(a)}}\Big(\int_{\underline{w}}^{a}\Big|\log\frac{q(a)}{q(x)}\Big|^2 r(x)dx\Big)^{1/2}\Big] \qquad \ldots(4.4)$$

Now, if $x \le b$, then $0 \le q(b) \le q(x) \Longrightarrow \frac{q(b)}{q(x)} \le 1$.

Hence, for every $0 < \varepsilon < 1$, $\Big|\log\frac{q(b)}{q(x)}\Big|^2\Big(\frac{q(b)}{q(x)}\Big)^{\varepsilon}$ is uniformly bounded by some constant $M^2(\varepsilon)$. Fix such an $\varepsilon > 0$ and call this bound M^2. A similar bound applies for $\left|\log \frac{q(a)}{q(x)}\right|^2 \left(\frac{q(a)}{q(x)}\right)^{\varepsilon}$.

Hence, from (4.4),

$$\int_a^b T(w)\pi(w)dw \leq 2M \left[\sqrt{T(b)\pi(b)} \left\{ \frac{q^{2-\varepsilon}(b)}{q'(b)} \int_{\underline{w}}^{b} \frac{q'(x)}{q^{2-\varepsilon}(x)} dx \right\}^{1/2} \right.$$

$$\left. + \sqrt{T(a)\pi(a)} \left\{ \frac{q^{2-\varepsilon}(a)}{q'(a)} \int_{\underline{w}}^{a} \frac{q'(x)}{q^{2-\varepsilon}(x)} dx \right\}^{1/2} \right]$$

$$= \frac{2M}{\sqrt{1-\varepsilon}} \left[\sqrt{T(b)\pi(b)} \sqrt{-\frac{q(b)}{q'(b)}} + \sqrt{T(a)\pi(a)} \sqrt{-\frac{q(a)}{q'(a)}} \right] \quad \ldots(4.5)$$

(since $q(\underline{w}) = \infty$).

Since $\int_{\underline{w}}^{a} -\frac{q'(w)}{q(w)} dw = \infty = \int_{b}^{\bar{w}} -\frac{q'(w)}{q(w)} dw$ for $\underline{w} < a$, $b < \bar{w}$, the proof now follows in the lines of Karlin (58).

Example

Let X have the truncated exponential distribution with density $p(x,w) = e^{-(x-w)}$, $x \geq w$. Then $\log q(w) = w$. Therefore, by the preceding theorem $X+1$ is an admissible estimator of w.

Remarks. 1. The admissibility of the Pitman estimate $\log q(X) - 1$ in Theorem 3.4.1 under squared error loss probably also follows from Brown (66).

2. Typically all the theorems in this chapter go through with some obvious modifications for the density $p(x,w) = r(x)q(w)$, $\bar{w} > x \geq w$, $\underline{w} < w < \bar{w}$. In fact, the example in Section 3.4 is from one such distribution.

CHAPTER 4

ADMISSIBILITY AND INADMISSIBILITY IN THE MULTIPARAMETER EXPONENTIAL FAMILY WITH APPLICATIONS IN THE GAMMA DISTRIBUTION

4.1 Introduction

In Chapters 2 and 3, we dealt with admissibility problems in one-parameter (exponential and non-regular) families. In this chapter, we consider some aspects of admissibility in the multiparameter exponential family.

It follows from the result of Karlin (58) that if $X \sim N(\theta, 1)$, $-\infty < \theta < \infty$, then X is admissible for θ under squared-error loss. However, the natural estimate X ceases to be admissible for θ under sum of squared-error losses if $X \sim N_p(\theta, I)$ and $p \geq 3$. This was shown by Stein (55) and formally by James and Stein (60). Since then the 'Stein-effect' has been found to be present in many non-normal populations, notably the Poisson and the Gamma distributions. A major breakthrough on multiparameter admissibility was achieved by Stein when he proved what is now popularly known as 'Stein's identity'. It was then generalized by Hudson (78), Berger (80), and recently by Hwang (81). The Stein-type identities reduce the difference in risk $R(\theta, \delta_1) - R(\theta, \delta_0)$ of two estimates to the form $E_\theta[\Delta(x)]$, where $\Delta(x)$ is a differential expression involving δ_0, δ_1, and their partial derivatives. If δ_1 could be so chosen that

$\Delta(x) < 0$ for all x, it will then follow that δ_o is inadmissible. Solving differential inequalities thus play a very important role in multiparameter admissibility problems. Some very important works on solving relevant differential inequalities were done by Berger (80), Ghosh and Parsian (80), Brown (81). In this chapter we use some of these tools to obtain inadmissibility results in the multiparameter exponential family.

Brown (71) showed that the admissibility of an estimate of the multinormal mean can often be settled by 'comparing' it with the James - Stein estimate

$$\delta_c(X) = \left(1 - \frac{c}{\sum_{i=1}^{p} X_i^2}\right) X, \qquad \ldots(1.1)$$

with $c = p-2$.

Formally, Brown (71) showed that

(a) An estimate $\delta(\cdot)$ of θ is inadmissible if for some $c < p-2$, and $M > 0$,

$$\sum_{i=1}^{p} x_i\delta_i(x) \geq \sum_{i=1}^{p} x_i\delta_{ci}(x) \quad \text{for } \|x\| \geq M \qquad \ldots(1.2)$$

(b) If $\delta(x)$ is a generalized Bayes estimate of θ with bounded risk, then $\delta(x)$ is admissible if for $c = p-2$, and some $M > 0$,

$$\sum_{i=1}^{p} x_i\delta_i(x) \leq \sum_{i=1}^{p} x_i\delta_{ci}(x) \quad \text{for } \|x\| \geq M. \qquad \ldots(1.3)$$

Note that (1.2) and (1.3) virtually imply that the James - Stein estimate with $c = p-2$ stands as a dividing line between admissible and inadmissible estimates of the multinormal mean. Hwang (81) extended Brown's result (a) for finding necessary conditions of admissibility of estimates of the vector of natural parameters in the general continuous exponential family.

In Section 4.2, we consider the problem of obtaining Hwang - type bounds on admissible estimates of the vector of reciprocals $(\theta_1^{-1}, \theta_2^{-1}, \ldots, \theta_p^{-1})$ of the natural parameters $\theta_1, \ldots, \theta_p$ in the continuous exponential family; as an application, it has been proved that if $X_1, \ldots, X_p$ are independent simple exponential random variables with scale parameters $\theta_1, \ldots, \theta_p$, then any estimate $\delta(x) = (\delta_1(x), \ldots, \delta_p(x))$ of the mean vector $(\theta_1^{-1}, \ldots, \theta_p^{-1})$ is inadmissible if for some $0 < c < 2(p-1)$ and some $M > 0$,

$$\sum_{i=1}^{p} x_i^{-3}\, \delta_i(x) \leq \sum_{i=1}^{p} x_i^{-3}\, \delta_{c,i}^{B}(x), \text{ for every } x \in (0,M]^p$$

where $\delta_{c,i}^{B}(x) = \frac{x_i}{2}\left[1 + c\,x_i^{-4}/2\,\left(\sum_{i=1}^{p} x_i^{-2}\right)^2\right]$ is the improved estimate of θ_i^{-1} suggested by Berger (80).

Hwang's technique, essentially an integration by parts, will not carry over to the problem of estimating the mean vector in the general continuous exponential family. It is implicit, however, in Hwang's work, that his bounds are true in the special Hudson (78) sub-family. In Section 4.3, these bounds are used to give necessary conditions of admissibility of estimates of the vector of gamma shape parameters when the scale parameters are known.

Finally, in Section 4.4, we address the problem of simultaneous estimation of gamma scale parameters. In contrast with the weighted quadratic losses considered by Berger (80) and Ghosh and Parsian (80), we have considered another invariant loss function $L(\theta, a) = \sum_{i=1}^{p} a_i\theta_i - \sum_{i=1}^{p} \log a_i\theta_i - p$, and have shown that the "natural" estimate for this loss is inadmissible for $p \geq 3$. We then make some remarks relating this inadmissibility result to Berger (80), Brown (80b), and Brown (66).

4.2 Estimation of reciprocals of the natural parameters

Hwang (81) generalized Berger's (80) unbiased estimate of $E_\theta[\theta\, h(X)]$ for step functions. In this section, we will need similar unbiased estimates of $E_\theta[\theta^{-1}h(X)]$ for step functions, in continuous exponential family. We start with notations and preliminaries.

Let $X_1,\ldots,X_p$ be p independent random variables, X_i having density (Lebesgue) $f_{\theta_i}(x_i) = e^{-\theta_i r_i(x_i)}\, p(\theta_i)\, t_i(x_i)$, $1 \leq i \leq p$; $a < x_i < b$. For $a \leq M \leq b$, and functions g such that $\lim_{x \to b} g(x) = 0$, we define

$$\int_a^b g(x)\left\{\frac{d}{dx} I_{(M,b)}(x)\right\} dx = g(M) \qquad \ldots(2.1)$$

More generally, if $h(x)$, $g(x)$ are two differentiable functions of x, then, we define

$$\frac{d}{dx}\left[h(x) + g(x)\, I_{(M,b]}(x)\right] = h'(x) + g'(x)\, I_{(M,b]}(x) + g(x)\,\frac{d}{dx} I_{(M,b]}(x) \qquad \ldots(2.2)$$

In (2.1) and (2.2), $(M,b]$ is to be interpreted as (M,∞) if $b = \infty$. From (2.2) it follows that

$$\frac{d}{dx} I_{[a,M]}(x) = -\frac{d}{dx} I_{(M,b]}(x) \qquad \ldots(2.3)$$

From (2.1), therefore,

$$E_{\theta_i}\left[g(X_i)\,\frac{d}{dX_i} I_{(M,b]}(X_i)\right]$$

$$= \int_a^b g(x)\left\{\frac{d}{dx} I_{(M,b]}(x)\right\} f_{\theta_i}(x)\, dx$$

$$= g(M)\, f_{\theta_i}(M) \qquad \ldots(2.4)$$

Analogously, $E_{\theta_i}\left[g(X_i)\,\frac{d}{dX_i} I_{[a,M]}(X_i)\right]$

$$= -\, g(M)\, f_{\theta_i}(M) \qquad \ldots(2.5)$$

It is understood that in (2.4) (and similarly in (2.5)), g satisfies a limiting condition $\lim_{x \to b} g(x) f_{\theta_i}(x) = 0$.

In the multivariate situation, A will be taken as a product

$\prod_{i=1}^{p} (M_i\,, b]$ or $\prod_{i=1}^{p} [a, M_i]$ as is the situation. The rule for finding partial derivatives of step-functions in the above sense will be

$$\frac{\partial}{\partial x_j} \prod_{i=1}^{p} g_i(x_i) \quad \left[\frac{d}{dx_j} g_j(x_j) \right] \prod_{i \neq j} g_i(x_i)\,, \quad 1 \leq j \leq p\,, \quad ..(2.6)$$

where $\frac{d}{dx_j} g_j(x_j)$ is already defined in (2.1), (2.2), (2.3).

Now we quote below a slightly modified version of a result of Hwang (81), to be used in the subsequent analysis.

<u>Theorem 4.2.1</u>. (Hwang (81)) Let $X = (X_1,\ldots,X_p)$ have an arbitrary multivariate distribution depending on a parameter $\theta = (\theta_1\,, \theta_2,\ldots,\theta_p)$. Let $\underset{\sim}{\gamma}(\theta) = (\gamma_1(\theta),\ldots,\gamma_p(\theta))$ be any parametric function and let $\underset{\sim}{\delta}_1(X)$ and $\underset{\sim}{\delta}_2(X)$ be two estimators such that $R(\theta,\underset{\sim}{\delta}_2(X)) \leq R(\theta,\underset{\sim}{\delta}_1(X))\ \forall\ \theta$ (with strict inequality for some θ), where $R(\theta,\underset{\sim}{\delta}(X)) = E_\theta \sum_{i=1}^{p} (\delta_i(X) - \gamma_i(\theta))^2$ for any estimator $\underset{\sim}{\delta}(X)$.

Let $\underset{\sim}{d}(X) = \underset{\sim}{\delta}_2(X) - \underset{\sim}{\delta}_1(X)$.

Let $\delta(X)$ be any estimator of $\underset{\sim}{\gamma}(\theta)$ such that

$$\underset{\sim}{d}(X).\underset{\sim}{\delta}(X) \leq \underset{\sim}{d}(X).\underset{\sim}{\delta}_1(X) \quad \text{for all } x\,.$$

Then $\underset{\sim}{\delta}(X)$ is inadmissible.

<u>Proof</u> : We show $\underset{\sim}{\delta}'(X) = \underset{\sim}{\delta}(X) + \underset{\sim}{d}(X)$ is better than $\underset{\sim}{\delta}(X)$.

For, $R(\theta,\underset{\sim}{\delta}'(X)) - R(\theta,\underset{\sim}{\delta}(X))$

$$= E \sum_{i=1}^{p} (\delta_i(X) + d_i(X) - \gamma_i(\theta))^2 - E \sum_{i=1}^{p} (\delta_i(X) - \gamma_i(\theta))^2$$

$$= E \sum_{i=1}^{p} d_i^2(X) + 2E \sum_{i=1}^{p} d_i(X)(\delta_i(X) - \gamma_i(\theta))$$

$$\leq E \sum_{i=1}^{p} d_i^2(X) + 2E \sum_{i=1}^{p} d_i(X)(\delta_{1i}(X) - \gamma_i(\theta)) \quad \text{(by hypothesis)}$$

$$= R(\theta, \underset{\sim}{\delta}_2(X)) - R(\theta, \underset{\sim}{\delta}_1(X)).$$

This completes the proof.

Theorem 4.2.1 essentially asserts that Hwang's basic lemma is applicable for any parametric function so long as the loss is squared error. With these preliminaries, we state below two lemmas, which are later used in Theorem 4.2.4.

Lemma 4.2.2 Let X have a density $f(x|\theta) = e^{-\theta r(x)} p(\theta) t(x)$, $\theta > 0$ $(\theta < 0)$, with respect to Lebesgue measure on $[a,b]$. Let $g(x)$ be an absolutely continuous function such that $\lim_{x \to b \, (x \to a)} g(x) e^{-\theta r(x)} = 0$ for every $\theta > 0$ $(\theta < 0)$. Then

$$E_\theta \left[\theta^{-1} \frac{\frac{\partial}{\partial X}\{g(X) I_A(X)\}}{t(X)} \right] = E_\theta \left[\frac{g(X) r'(X)}{t(X)} I_A(X) \right], \qquad \ldots(2.7)$$

where $A = [a, M]$ or $(M, b]$.

This lemma is followed by its usual multivariate analog below.

Lemma 4.2.3 Let X_i have a density $f(x_i|\theta_i) = e^{-\theta_i r_i(x_i)} p(\theta_i)$ $t_i(x_i)$, $\theta_i > 0$ $(\theta_i < 0)$, with respect to Lebesgue - measure on $[a,b]$. Let $g(X) = (g_1(X), \ldots, g_p(X))$ be an absolutely continuous function such that $\lim_{\substack{x_i \to b \\ (x_i \to a)}} g_i(x)\, e^{-\theta_i r_i(x_i)} = 0$ for every $\theta_i > 0$ ($\theta_i < 0$). Then

$$E_\theta\left[\theta_i^{-1} \frac{\frac{\partial}{\partial x_i}\left\{g_i(X)\, I_A(X)\right\}}{t_i(X_i)}\right] = E_\theta\left[\frac{g_i(X)\, r_i'(X_i)}{t_i(X_i)}\, I_A(X)\right], \quad ..(2.8)$$

where A is a p-dimensional rectangle $[a,M]^p$ or $(M,b]^p$.

Lemma 4.2.2 is straightforward integration by parts and Lemma 4.2.3 is proved by the usual technique of taking conditional expectations given $X_j = x_j$, $j \neq i$ (Hudson (78), Stein (81)).

Remark:

Suppose now $h_i(x)$ is a given function of x, for $i = 1, 2, \ldots, p$. If $g_i(x)$ is defined as the indefinite integral of $h_i(x)\, t_i(x_i)$ with respect to x_i, then (2.8) can be rewritten as

$$E_\theta\left[\theta_i^{-1} h_i(X) I_A(X)\right] = E_\theta\left[\frac{g_i(X)\, r_i'(X_i)}{t_i(X_i)}\, I_A(X)\right]$$

$$- E_\theta\left[\theta_i^{-1} \frac{g_i(X)}{t_i(X_i)} \frac{\partial}{\partial X_i} I_A(X)\right] \quad ...(2.9)$$

(2.9) will frequently be useful in expressing the difference in risk of two estimators as a differential operator plus a negative quantity.

We now go into actually obtaining Hwang type bounds on admissible estimators of $(\theta_1^{-1}, \ldots, \theta_p^{-1})$ under squared error loss in the general continuous exponential family.

Let $\lambda^*(X)$ and $\lambda(X)$ be two estimators of $(\theta_1^{-1},\ldots,\theta_p^{-1})$ defined by

$$\lambda_i^*(x) = \frac{x_i}{\alpha_i+1}\left[1+\phi_i^*(x)\, I_A(x)\right],\ 1 \le i \le p \qquad \ldots(2.10)$$

$$\lambda_i(x) = \frac{x_i}{\alpha_i+1}\left[1+\phi_i(x)\, I_A(x)\right],\ 1 \le i \le p \qquad \ldots(2.11)$$

where $\alpha_1,\ldots\alpha_p$ are any constants, and ϕ,ϕ^* are functions satisfying certain conditions indicated later in this section, and A is a p-dimensional rectangle. The calculations closely resemble that of Hwang (81). The idea is to express $R(\theta,\lambda^*)-R(\theta,\lambda)$ as $E(\Delta\phi^*(x)-\Delta\phi(x))$ plus a negative quantity, where $\Delta(\cdot)$ is a differential operator, and choose ϕ,ϕ^* suitably such that $\Delta\phi^*(x) < \Delta\phi(x)$.

Define $\delta^0(X)$ by

$$\delta_i^0(x) = \frac{x_i}{\alpha_i+1},\quad 1 \le i \le p \qquad \ldots(2.12)$$

Now, $R(\theta,\lambda^*) - R(\theta,\delta^0)$

$$= E\sum_{i=1}^{p}(\lambda_i^*(x)-\theta_i^{-1})^2 - E\sum_{i=1}^{p}(\delta_i^0(x)-\theta_i^{-1})^2$$

$$= E\sum_{i=1}^{p}\left(\lambda_i^*(x)-\frac{x_i}{\alpha_i+1}\right)^2 + 2E\sum_{i=1}^{p}\left(\frac{x_i}{\alpha_i+1}-\theta_i^{-1}\right)\left(\lambda_i^*(x)-\frac{x_i}{\alpha_i+1}\right)$$

$$= E\left[\sum_{i=1}^{p} \frac{x_i^2}{(\alpha_i+1)^2} \phi_i^{*2}(x)\, I_A(x) + 2 \sum_{i=1}^{p} \frac{x_i^2}{(\alpha_i+1)^2} \phi_i^{*}(x)\, I_A(x) \right.$$

$$\left. - 2 \sum_{i=1}^{p} \theta_i^{-1} \frac{x_i}{\alpha_i+1} \phi_i^{*}(x)\, I_A(x) \right]$$

$$= E\left[\sum_{i=1}^{p} \frac{x_i^2}{(\alpha_i+1)^2} \phi_i^{*2}(x)\, I_A(x) + 2 \sum_{i=1}^{p} \frac{x_i^2}{(\alpha_i+1)^2} \phi_i^{*}(x)\, I_A(x) \right.$$

$$\left. - 2 \sum_{i=1}^{p} \frac{g_i^{*}(x) r_i'(x_i)}{t_i(x_i)} I_A(x) + 2 \sum_{i=1}^{p} \theta_i^{-1} g_i^{*}(x) \frac{\partial}{\partial x_i} I_A(x) \Big/ t_i(x_i) \right] \qquad \ldots(2.13)$$

The last equality in (2.13) is a consequence of (2.9), where $g_i^*(x)$ is to be taken as an indefinite integral (with respect to x_i) of $\frac{x_i}{\alpha_i+1} \phi_i^*(x)\, t_i(x_i)$. At this stage it is worth noting that $\phi_i^*(x)$ (and similarly $\phi_i(x)$, for (2.15) to be valid) should be such that

$$\lim_{\substack{x_i \to b \\ (x_i \to a)}} g_i^*(x)\, e^{-\theta_i r_i(x_i)} = 0$$

$$\lim_{\substack{x_i \to b \\ (x_i \to a)}} g_i(x)\, e^{-\theta_i r_i(x_i)} = 0\,, \qquad \ldots(2.14)$$

for every θ_i ; $g_i(x)$ is formally defined below.

Similarly,

$$R(\theta,\lambda) - R(\theta,\delta^{0})$$

$$= E\Big[\sum_{i=1}^{p} \frac{x_i^2}{(\alpha_i+1)^2} \phi_i^2(x)\, I_A(x) + 2\sum_{i=1}^{p} \frac{x_i^2}{(\alpha_i+1)^2} \phi_i(x)\, I_A(x)$$

$$- 2\sum_{i=1}^{p} \frac{g_i(x) r_i'(x_i) I_A(x)}{t_i(x_i)} + 2\sum_{i=1}^{p} \theta_i^{-1}\, g_i(x) \frac{\partial}{\partial x_i} I_A(x) \Big/ t_i(x_i)\Big] \qquad \ldots(2.15)$$

where $g_i(x)$ is an indefinite integral (with respect to x_i) of $\frac{x_i}{\alpha_i+1}\, \phi_i(x)\, t_i(x_i)$. (2.13) and (2.15) now enable us to write

$$R(\theta,\lambda^{*}) - R(\theta,\lambda)$$

$$= \left\{ R(\theta,\lambda^{*}) - R(\theta,\delta^{0}) \right\} - \left\{ R(\theta,\lambda) - R(\theta,\delta^{0}) \right\}$$

$$= E\Big[(\Delta\phi^{*}(x) - \Delta\phi(x)) I_A(x) + 2\sum_{i=1}^{p} B_i(x,\theta) \Big] \qquad \ldots(2.16)$$

where $$\Delta\phi^{*}(x) = \sum_{i=1}^{p} \frac{x_i^2}{(\alpha_i+1)^2} \phi_i^{*2}(x) + 2\sum_{i=1}^{p} \frac{x_i^2}{(\alpha_i+1)^2} \phi_i^{*}(x)$$

$$- 2\sum_{i=1}^{p} \frac{g_i^{*}(x) r_i'(x_i)}{t_i(x_i)} \qquad \ldots(2.17)$$

$$\Delta\phi(x) = \sum_{i=1}^{p} \frac{x_i^2}{(\alpha_i+1)^2} \phi_i^2(x) + 2\sum_{i=1}^{p} \frac{x_i^2}{(\alpha_i+1)^2} \phi_i(x)$$

$$- 2\sum_{i=1}^{p} \frac{g_i(x) r_i'(x_i)}{t_i(x_i)} \qquad \ldots(2.18)$$

and $$B_i(x,\theta) = \theta_i^{-1}(g_i^{*}(x) - g_i(x)) \frac{\partial}{\partial x_i} I_A(x) \Big/ t_i(x_i) \qquad \ldots(2.19)$$

Therefore, if there exists a rectangle A such that $\Delta\phi^*(x) < \Delta\phi(x)$ for every $x \in A$, and $E_\theta(B_i(x, \theta)) \le 0 \;\; \forall\, \theta$, then $R(\theta, \lambda^*) < R(\theta, \lambda) \;\; \forall\, \theta$, and Hwang's lemma (Theorem 4.2.1) is applicable.

Hence, any estimator $\delta(X) = (\delta_1(X), \ldots, \delta_p(X))$ of $(\theta_1^{-1}, \ldots, \theta_p^{-1})$ will be inadmissible if

$$\sum_{i=1}^{p} \delta_i(x) \frac{x_i}{\alpha_i+1}(\phi_i^*(x) - \phi_i(x)) \le \sum_{i=1}^{p} \lambda_i(x) \frac{x_i}{\alpha_i+1}(\phi_i^*(x) - \phi_i(x)) \qquad \ldots(2.20)$$

for almost all x in A.

We can now state the following theorem.

<u>Theorem 4.2.4</u> Let $\lambda^*(X)$ and $\lambda(X)$ be two estimators as defined in (2.10) and (2.11), satisfying the conditions (2.14). Let $\Delta\phi^*(x)$, $\Delta\phi(x)$, $B_i(x,\theta)$ be as in (2.17), (2.18) and (2.19) respectively. If there exists an A such that

$\Delta\phi^*(x) < \Delta\phi(x)$ for (almost all) $x \in A$

and $E_\theta(B_i(x,\theta)) \le 0$ for every θ,

then any estimate $\delta(x)$ of $(\theta_1^{-1}, \ldots, \theta_p^{-1})$, satisfying (2.20) for (almost all) $x \in A$, is inadmissible.

<u>Remark</u>. Theorem 4.2.4 could also be stated by starting with an arbitrary estimate $\delta^0(X)$ rather than $\delta_i^0(X) = \frac{X_i}{\alpha_i+1}$. But since the motivation behind Theorem 4.2.4 is in estimating scale -

parameters in independent gamma distributions, in which case $\frac{X_i}{\alpha_i+1}$ is the standard estimate of θ_i^{-1}, we have conveniently chosen $\delta_i^o(X)$ as $\frac{X_i}{\alpha_i+1}$, for some constants $\alpha_1, \alpha_2, \ldots, \alpha_p$.

<u>Example</u> Let $X_1, \ldots, X_p$ be independent with $f_{\theta_i}(x_i) = \theta_i e^{-\theta_i x_i}$, $\theta_i > 0$, $x_i > 0$, $i = 1,2,\ldots,p$. In this case $E_{\theta_i}(X_i) = \theta_i^{-1}$, $i = 1,2,\ldots,p$. The natural estimate of $(\theta_1^{-1}, \ldots, \theta_p^{-1})$ is $(\frac{X_1}{2}, \ldots, \frac{X_p}{2})$.

Let $$\phi_i(x) = \frac{c\, x_i^{-4}}{2(\sum_{j=1}^{p} x_j^{-2})^2}, \quad 1 \leq i \leq p, \quad 0 < c < 2(p-1).$$

By definition, $g_i(x)$ is the indefinite integral of

$$\frac{c\, x_i^{-3}}{4(\sum_{j=1}^{p} x_j^{-2})^2}$$ with respect to x_i.

Hence, $$g_i(x) = \frac{c}{8(\sum_{j=1}^{p} x_j^{-2})}, \quad 1 \leq i \leq p.$$

Indeed, $$\frac{\partial}{\partial x_i} g_i(x) = \frac{2\, c x_i^{-3}}{8(\sum_{j=1}^{p} x_j^{-2})^2} = \frac{c x_i^{-3}}{4(\sum_{j=1}^{p} x_j^{-2})^2}, \quad 1 \leq i \leq p.$$

Now, using (2.18),

$$\mathcal{R}(\phi(x)) = \sum_{i=1}^{p} \frac{x_i^2}{4} \frac{c^2 x_i^{-8}}{4(\sum_{j=1}^{p} x_j^{-2})^4} + 2 \sum_{i=1}^{p} \frac{x_i^2}{4} \frac{c\, x_i^{-4}}{2(\sum_{j=1}^{p} x_j^{-2})^2}$$

$$- 2 \sum_{i=1}^{p} \frac{c}{8 \sum_{j=1}^{p} x_j^{-2}}$$

$$= \frac{1}{16}\left[- 4\, p\, c\, D^{-1} + 4\, c\, D^{-1} + c^2 \frac{\sum_{i=1}^{p} x_i^{-6}}{(\sum_{j=1}^{p} x_j^{-2})^4} \right]$$

$$= \frac{1}{16}\left[(c^2 + 4c - 4pc) D^{-1} + c^2 \left\{ \frac{\sum_{i=1}^{p} x_i^{-6}}{(\sum_{i=1}^{p} x_i^{-2})^4} - \frac{1}{\sum_{i=1}^{p} x_i^{-2}} \right\} \right] ,$$

...(2.21)

where $D = \sum_{i=1}^{p} x_i^{-2}$.

It is easy to see that $(c^2 + 4c - 4pc)D^{-1}$ is minimized at $c = 2(p-1)$.

Also, as $(\sum_{i=1}^{p} x_i^{-6}) (\sum_{i=1}^{p} x_i^{-2}) < (\sum_{i=1}^{p} x_i^{-2})^4$ for every x, it follows that

$$c^2 \left\{ \frac{\sum_{i=1}^{p} x_i^{-6}}{(\sum_{i=1}^{p} x_i^{-2})^4} - \frac{1}{\sum_{i=1}^{p} x_i^{-2}} \right\} > 4(p-1)^2 \left\{ \frac{\sum_{i=1}^{p} x_i^{-6}}{(\sum_{i=1}^{p} x_i^{-2})^4} - \frac{1}{\sum_{i=1}^{p} x_i^{-2}} \right\}$$

...(2.22)

for every $0 < c < 2(p-1)$.

Therefore, if $\phi_i^*(x)$ is taken as $\phi_i(x)$ with $c = 2(p-1)$, then

(2.22) gives $\triangle(\phi^*(x)) < \triangle(\phi(x))$ for every x.

Also, with this choice of $\phi_i^*(x)$,

$$g_i^*(x) - g_i(x) = \frac{2(p-1)-c}{8\sum_{i=1}^{p} x_i^{-2}} . \qquad \ldots(2.23)$$

If now, $A = (0,M]^p$, then by using (2.5), (2.6), (2.23) and the fact that $c < 2(p-1)$, $E_\theta\left[\theta_i^{-1}\left\{g_i^*(x) - g_i(x)\right\}\frac{\partial}{\partial x_i} I_A(x)\right] \leq 0 \quad \forall\, \theta > 0$.

Therefore, we have,

<u>Corollary</u> Let $X_1,\ldots,X_p$ be independent simple exponentials, with $E_{\theta_i}(X_i) = \theta_i^{-1}$. Then any estimate $\delta(X) = (\delta_1(X),\ldots,\delta_p(x))$ of $(\theta_1^{-1},\ldots,\theta_p^{-1})$ is inadmissible, provided for some $0 < c < 2(p-1)$, some $M > 0$,

$$\sum_{i=1}^{p} x_i^{-3}\,\delta_i(x) \leq \sum_{i=1}^{p} x_i^{-3}\,\delta^B_{c,i}(x) \quad \text{for every } x \in (0,M]^p \qquad \ldots(2.24)$$

where $\delta^B_{c,i}(x) = \frac{x_i}{2}\left[1 + cx_i^{-4}/2(\sum_{i=1}^{p} x_i^{-2})^2\right]$ = Berger's estimate (80).

<u>Remark</u> : The corollary above seems to give evidence that Berger's estimate with $c = 2(p-1)$ stands as the dividing line between admissible and inadmissible estimates of the mean-vector in independent simple exponential distributions. In particular, this corollary also shows Berger's observation that the standard estimate $X/2$ is inadmissible if $p \geq 2$. For the general gamma case, however, the calculations corresponding to (2.21) and (2.22) get complicated and it is not clear if a similar result holds there too. Hwang (81) obtained a similar bound for admissible estimates of natural parameters in independent gamma distributions.

4.3 Bounds on admissible estimates in Hudson's family : An example

Hwang (81) obtained bounds on admissible estimates of the vector of natural parameters in the continuous exponential family. Since the mean in Hudson's family is essentially the natural parameter of a continuous exponential, it is implicit in Hwang (81) that such bounds hold in the special Hudson family. Specializing to the Hudson family seems to be necessary for the crucial integration by parts to go through. We will use in this section the bounds of Hwang to give another example in the gamma distribution.

The Hudson (78) family is characterized by the density function

$$f_\theta(x) = e^{\mu(\theta)\int a^{-1}(x)dx - \psi(\theta)} \quad a^{-1}(x)\ e^{-\int x a^{-1}(x)dx} \qquad \underline{w} < x < \bar{w} \quad ..(3.1)$$

where $a(x) > 0$ for every x, the integrals are interpreted as primitives, and $\mu(\theta)$ denotes $E_\theta(X)$.

Defining $B(x) = \int a^{-1}(x)dx$, it is seen that B has a density of the form

$$f^*_\mu(b) = e^{\mu b - \phi(\mu)} \quad k(b) \qquad ...(3.2)$$

We now quote below the essential steps for the sake of completeness, with the notations of the previous section.

<u>Lemma 4.3.1</u> Let $g(X)$ be an absolutely continuous function such that $\lim_{x \to \bar{w}} a(x)g(x)\, f_\theta(x) = \lim_{x \to \underline{w}} a(x)g(x)f_\theta(x) = 0 \quad \forall\, \theta$, and $E_\theta\left[\,|g'(X)a(X)|\,\right] < \infty \quad \forall\, \theta$. Then,

$$E_\theta\left[(X-\mu)g(X)\ I_A(X)\right] = E_\theta\left[\frac{d}{dX}\left\{g(X)I_A(X)\right\} a(X)\right],$$

where $A = (M_1, M_2)$, $\underline{w} \le M_1 < M_2 \le \bar{w}$.

Lemma 4.3.2 (Multivariate analog of Lemma 4.3.1) Let $X_1, \ldots, X_p$ be independent with X_i having density $f_{\theta_i}(x_i)$. Let $\mu_i = E_{\theta_i}(X_i)$. Let $g = (g_1, \ldots, g_p)$ be any absolutely continuous function such that

$$\lim_{\substack{x_i \to \bar{w} \\ (x_i \to \underline{w})}} a(x_i)\, g_i(x) f_{\theta_i}(x_i) = 0 \quad \forall\, \theta_i, \quad \forall\, i = 1, \ldots, p$$

and $E_\theta \left[\left| \frac{\partial}{\partial X_i} g_i(X) \right| a(X_i) \right] < \infty \quad \forall\, \theta$. Then,

$$E_\theta \left[(X_i - \mu_i) g_i(X) I_A(X) \right] = E_\theta \left[\frac{\partial}{\partial X_i} \left\{ g_i(X) I_A(X) \right\} a(X_i) \right],$$

where $A = (M_1, M_2)^p$, $\underline{w} \le M_1 < M_2 \le \bar{w}$.

Lemma 4.3.1 and 4.3.2 are essentially proved in Hwang (81), and are straightforward integration by parts. Let now $\lambda^o(X)$ and $\lambda^*(X)$ defined by

$$\lambda_i^o(X) = X_i + g_i^o(X) I_A(X) \qquad \ldots(3.3)$$

$$\lambda_i^*(X) = X_i + g_i^*(X) I_A(X) \qquad \ldots(3.4)$$

be two estimators of μ, where it is understood g_i and g_i^* are such that they satisfy the conditions of Lemma 4.3.2. Then, proceeding as in Hwang (81),

$$R(\theta, \lambda^*) - R(\theta, \lambda^o)$$

$$= E_\theta \left[\left\{ \Delta g^*(X) - \Delta g^o(x) \right\} I_A(X) \right] + 2 E_\theta \left[\sum_{i=1}^{p} a(X_i) \left\{ g_i^*(X) - g_i^o(X) \right\} \frac{\partial}{\partial X_i} I_A(X) \right]$$

$$\ldots(3.5)$$

where $\Delta g^o(x) = \sum_{i=1}^{p} g_i^{o2}(x) + 2 \sum_{i=1}^{p} a(x_i) \frac{\partial}{\partial x_i} g_i^o(x)$ ...(3.6)

and $\Delta g^*(x) = \sum_{i=1}^{p} g_i^{*2}(x) + 2 \sum_{i=1}^{p} a(x_i) \frac{\partial}{\partial x_i} g_i^*(x)$ ...(3.7)

Let also, $B_i(x) = a(x_i) \left\{ g_i^*(x) - g_i^o(x) \right\} \frac{\partial}{\partial x_i} I_A(x)$ ...(3.8)

Therefore, if there exists a set A of positive Lebesgue measure such that $\Delta g^*(x) < \Delta g^o(x)$ a.e. on A ...(3.9)

and $E_\theta [B_i(x)] \leq 0 \quad \forall \theta$ ...(3.10)

then $R(\theta, \lambda^*) < R(\theta, \lambda^o)$, and Theorem 4.2.1 is applicable.

<u>Theorem 4.3.3</u> Let $g^o(X)$ and $g^*(X)$ be two estimators of μ satisfying the hypotheses of Lemma 4.3.2. If there is a set $A = (M_1, M_2)^p$ for $\underline{w} \leq M_1 < M_2 \leq \bar{w}$, such that (3.9) and (3.10) hold, then any estimator $\delta(X)$ of μ is inadmissible if

$$\sum_{i=1}^{p} \delta_i(x) \left[g_i^*(x) - g_i^o(x) \right] \leq \sum_{i=1}^{p} (x_i + g_i^o(x)) \left[g_i^*(x) - g_i^o(x) \right],$$

a.e. on A ...(3.11)

<u>Remark</u> $\lambda^o(X)$ and $\lambda^*(X)$ could be defined by starting with any estimate $\delta^o(X)$ rather than X. Once again, since we are interested in the application of Theorem 4.3.3 to estimation of the gamma shape-parameter, we have conveniently started with X.

<u>Example</u> Let $X_1, X_2, \ldots, X_p$ be independent, X_i having p.d.f.

$$f_{\mu_i}(x_i) = e^{-x_i} x_i^{\mu_i - 1} / \Gamma(\mu_i), \quad 1 \leq i \leq p, \quad x_i > 0, \ \mu_i > 0.$$

Note that $f_\mu(x)$ is a member of the Hudson family, with $a(x) = x$.

Take $g_i^o(x) = -c\log x_i / \sum_{i=1}^{p} (\log x_i)^2$, motivated by Hudson (78).

Then $$\Delta g^o(X) = \frac{c^2}{\sum_{i=1}^{p}(\log x_i)^2} - \frac{2cp}{\sum_{i=1}^{p}(\log x_i)^2} + \frac{4c}{\sum_{i=1}^{p}(\log x_i)^2}$$

$$= \frac{c^2 - 2c(p-2)}{S}, \quad \text{where } S = \sum_{i=1}^{p}(\log x_i)^2 \qquad \ldots(3.12)$$

Now, $c^2 - 2c(p-2)$ is minimized at $c = p-2$. Therefore, if we let $g_i^*(x)$ as $g_i^o(x)$ with $c = p-2$, then from (3.12) it follows that $\Delta g^*(x) < \Delta g^o(x)$ for all x. Also,

$$E_\theta[B_i(X)] = E_\theta\left[(c-(p-2))\, X_i \log X_i \frac{\partial}{\partial X_i} I_A(X) \Big/ \sum_{j=1}^{p}(\log X_j)^2\right] \qquad \ldots(3.13)$$

If $c < p-2$, then from (3.13) it is clear that $E_\theta(B_i(X)) \le 0 \;\; \forall\, \theta$ if $A = (0,M]^p$ with $M < 1$, or if $A = (M,\infty)^p$ with $M > 1$. Thus (3.9) and (3.10) both hold, and we have the following corollary.

<u>Corollary</u>. Let $X_1,\ldots,X_p$ be independently distributed, X_i having p.d.f.

$$f_{\mu_i}(x_i) = e^{-x_i} x_i^{\mu_i - 1} \Big/ \Gamma(\mu_i), \; 1 \le i \le p, \; x_i > 0, \; \mu_1,\ldots,\mu_p$$

positive and unknown. Let $\delta(X)$ be any estimator of $(\mu_1,\ldots,\mu_p)$. Suppose there exists $0 < c < p-2$ such that

$$\sum_{i=1}^{p} \delta_i(x)\log x_i \ge \sum_{i=1}^{p}\left[x_i - \frac{c\log x_i}{\sum_{j=1}^{p}(\log x_j)^2}\right]\log x_i \qquad \ldots(3.14)$$

for all $x \in (0,M]^p$ for some $M < 1$, or for all $x \in (M,\infty)^p$ for some $M > 1$. Then $\delta(X)$ is inadmissible.

4.4 Estimation of the gamma scale parameter

Let $X_1, X_2, \ldots, X_p$ be independent, with X_i having density $f_{\theta_i}(x_i) = e^{-\theta_i x_i}\,\theta_i^{\alpha_i}\,x_i^{\alpha_i-1} / \Gamma(\alpha_i)$, $x_i > 0$, $1 \le i \le p$, where $\alpha_i > 0$ are known, and θ_i's (> 0) are considered unknown. Berger (80) considered weighted quadratic losses $\sum_{i=1}^{p} \theta_i^{-m} (\delta_i\theta_i - 1)^2$ for $m = 0, 2, 1, -1$, and showed that the standard estimate of $(\theta_1^{-1}, \theta_2^{-1}, \ldots, \theta_p^{-1})$, namely, $(\frac{X_1}{\alpha_1+1}, \ldots, \frac{X_p}{\alpha_p+1})$ is inadmissible for $p \ge 2$ except when $m = 0$, in which case it is inadmissible for $p \ge 3$. Ghosh and Parsian (80) also discussed this problem for the same weighted quadratic losses. In this section, we consider a typically different loss $\sum_{i=1}^{p} \delta_i\theta_i - \sum_{i=1}^{p} \log\delta_i\theta_i - p$; the vector of unbiased estimates $(\frac{X_1}{\alpha_1}, \ldots, \frac{X_p}{\alpha_p})$ is a natural estimate of the mean-vector for this loss. We show that this estimate is inadmissible for $p \ge 3$ and relate this inadmissibility result to some observations of Berger and Brown.

The usual technique of integration by parts (Berger's (80) identity) and a theorem of Ghosh and Parsian (80) are stated below for future reference. Also, a technical lemma, to be used subsequently, is also proved.

Lemma 4.4.1 (Berger (80)) Let $h(x) = (h_1(x), \ldots, h_p(x))$ be a function such that $\lim_{x_i \to 0} h_i(x)\, x_i^{\alpha_i - 1} e^{-\theta_i x_i} = 0$

and $\lim_{x_i \to \infty} h_i(x)\, x_i^{\alpha_i - 1} e^{-\theta_i x_i} = 0$

for every $\theta_i > 0$. Assume $h_i(x)$ has all partial derivatives of first order. Then

$$E_\theta[\theta_i h_i(X)] = E_\theta\left[h_i^{i(1)}(X) + \frac{(\alpha_i - 1)h_i(X)}{X_i}\right],$$

where $h_i^{i(1)}(X) = \frac{\partial}{\partial X_i} h_i(X)$, $1 \leq i \leq p$.

It is implicitly assumed in the above identity that

$E_\theta\left[|h_i^{i(1)}(X) + (\alpha_i - 1) h_i(X)/X_i|\right] < \infty$, for every θ.

Lemma 4.4.2 (Ghosh and Parsian (80)) For given functions $\upsilon_i(x_i)$, $\psi(x) > 0$, $w_i(x)$, define $\xi_i'(x_i) = \upsilon_i^{-1}(x_i)$ and $S = \sum_{j=1}^{p} d_j |\xi_j(x_j)|^\beta$ where d_j and β are positive constants to be chosen later. If

$$\sum_{i=1}^{p} w_i(x)\xi_i^2(x_i)/\psi(x) \leq KS \quad \text{for some } K, d_j \text{ and } \beta \quad \ldots(4.1)$$

(all positive) and for all $x \in \mathbb{R}^p$,

then

$$\phi_i(x) = \frac{-c\,\xi_i(x_i)}{S+b} \quad \ldots(4.2)$$

provides a solution to $\Delta(x) < 0$ for all $b > 0$ and $0 < c < K^{-1}(p-\beta)$

where $\Delta(x) = \psi(x) \sum_{i=1}^{p} \upsilon_i(x_i)\, \phi_i^{i(1)}(x) + \sum_{i=1}^{p} w_i(x)\, \phi_i^2(x) \quad \ldots(4.3)$

Remark Such solutions to $\Delta(x) < 0$ were first obtained by Berger (80). The constant c can be generalized to a non-decreasing function $\tau(S)$, with $0 < \tau(S) < K^{-1}(p-\beta)$.

<u>Lemma 4.4.3</u> For $|x| < \frac{1}{2}$, $\log(1+x) \geq x - \frac{3x^2}{2}$.

<u>Proof</u> : Define $f(x) = \log(1+x) - x + \frac{3x^2}{2}$.

Then $f'(x) = \frac{x(3x+2)}{1+x}$

$$\geq 0 \quad \text{for} \quad 0 \leq x < \tfrac{1}{2}$$

$$\leq 0 \quad \text{for} \quad -\tfrac{1}{2} < x \leq 0 .$$

Consequently, $f(x) \geq f(0) = 0$, for $|x| < \frac{1}{2}$.

<u>Theorem 4.4.4</u> Let $X_1,\ldots,X_p$ be independent gamma variables with $E(X_i) = \frac{\alpha_i}{\theta_i}$, α_i known. Consider the loss $L(\theta^{-1},\delta)$ $= \sum_{i=1}^{p} \delta_i\theta_i - \sum_{i=1}^{p} \log \delta_i\theta_i - p$. Then $(\frac{X_1}{\alpha_1},\ldots,\frac{X_p}{\alpha_p})$ is an inadmissible estimator of $(\theta_1^{-1},\ldots,\theta_p^{-1})$ for $p \geq 3$.

<u>Proof</u> : Let $\delta(X)$ be a competitor to the natural estimate

$$\delta_o(X) = (\frac{X_1}{\alpha_1}, \ldots, \frac{X_p}{\alpha_p}) .$$

Write $\delta_i(x) = \frac{x_i}{\alpha_i} + h_i(x)$, $1 \leq i \leq p$. We assume $h_i(x)$ are such that Lemma 4.4.1 holds.

Then, $\alpha(\theta) = R(\theta,\delta) - R(\theta,\delta_o)$

$$= \sum_{i=1}^{p} E\left\{ \delta_i(X)\theta_i - \log(\delta_i(X)\theta_i) - \frac{X_i}{\alpha_i}\theta_i + \log(\frac{X_i}{\alpha_i}\theta_i)\right\}$$

$$= \sum_{i=1}^{p} E\left\{ \theta_i h_i(X) - \log \frac{\alpha_i\delta_i(X)}{X_i}\right\}$$

$$= \sum_{i=1}^{p} E\left\{ \theta_i h_i(X) - \log(1 + \frac{\alpha_i h_i(X)}{X_i})\right\} \qquad \ldots(4.4)$$

If the competitor $\delta(X)$ is such that $\left|\frac{\alpha_i h_i(x)}{x_i}\right| < \frac{1}{2}$ uniformly in x, for every $1 \leq i \leq p$, then by Lemma 4.4.3,

$$\alpha(\theta) \leq \sum_{i=1}^{p} E\left\{ \theta_i h_i(X) - \frac{\alpha_i h_i(X)}{X_i} + \frac{3}{2} \frac{\alpha_i^2 h_i^2(X)}{X_i^2} \right\}$$

$$= \sum_{i=1}^{p} E\left\{ h_i^{i(1)}(X) + \frac{(\alpha_i - 1) h_i(X)}{X_i} - \frac{\alpha_i h_i(X)}{X_i} + \frac{3}{2} \frac{\alpha_i^2\, h_i^2(X)}{X_i^2} \right\}$$

(by Lemma 4.4.1)

$$= \sum_{i=1}^{p} E\left\{ h_i^{i(1)}(X) - \frac{h_i(X)}{X_i} + \frac{3}{2} \frac{\alpha_i^2\, h_i^2(X)}{X_i^2} \right\} \qquad \ldots(4.5)$$

Now make the transformation

$$h_i(x) = x_i \phi_i(x), \quad 1 \leq i \leq p.$$

Then $$h_i^{i(1)}(x) = \phi_i(x) + x_i \phi_i^{i(1)}(x)$$

$$= \frac{h_i(x)}{x_i} + x_i \phi_i^{i(1)}(x).$$

Hence, (4.5) gives,

$$\alpha(\theta) \leq E\left[\sum_{i=1}^{p} x_i \phi_i^{i(1)}(x) + \frac{3}{2} \sum_{i=1}^{p} \alpha_i^2\, \phi_i^2(x) \right] \qquad \ldots(4.6)$$

Note now the differential expression within braces in (4.6) is of the form (4.3), and (4.1) of Lemma 4.4.2 is satisfied with $K = \frac{3}{2}$, $\beta = 2$, $d_j = \alpha_j^2$.

Therefore, for $0 < c < \frac{2}{3}(p-2)$, $b > 0$,

$$\phi_i(x) = \frac{-c \log x_i}{\sum_{j=1}^{p} \alpha_j^2 (\log x_j)^2 + b} \qquad \ldots(4.7)$$

is a solution to $\sum_{i=1}^{p} x_i \phi_i^{i(1)}(x) + \frac{3}{2} \sum_{i=1}^{p} \alpha_i^2 \phi_i^2(x) < 0$.

Also observe that $|\alpha_i \phi_i(x)|^2 = \dfrac{c^2 \alpha_i^2 (\log x_i)^2}{\left\{\sum_{j=1}^{p} \alpha_j^2 (\log x_j)^2 + b\right\}^2} < \frac{1}{4}$ if $b > 4c^2$.

Hence, if $\delta_i(x) = \dfrac{x_i}{\alpha_i} - \dfrac{c\, x_i \log x_i}{\sum_{j=1}^{p} \alpha_j^2 (\log x_j)^2 + b}$, $1 \le i \le p$, $\qquad \ldots(4.8)$

where $0 < c < \frac{2}{3}(p-2)$, $b > 4c^2$, then, $\alpha(\theta) < 0$ for all θ.

This proves the theorem.

Some remarks on Theorem 4.4.4

1. It is easy to check that the tail condition required on $h_i(x)$ for Lemma 4.4.1 to hold is satisfied by the solutions eventually obtained in Theorem 4.4.4.

2. It was shown by Berger (80) that the critical dimension of inadmissibility of the natural estimate of the gamma scale-parameters is frequently 2, rather than 3. Berger contended that the critical dimension of inadmissibility is typically 2, and 3 dimension is required only in special situations. Brown (80b) discussed Berger's phenomenon and some of its peripheral aspects in the context of simultaneous estimation of independent normal means, and

gave examples to assert that the critical dimension of inadmissibility depends on the loss, rather than the underlying coordinate distributions. Theorem 4.4.4 gives another example of a natural and invariant loss for which the critical dimension of inadmissibility could be 3 in the gamma distribution itself, although the peripheral aspects relating to the point of shrinkage are not illustrated by this example. Interestingly, Berger (80) also required 3 dimension for inadmissibility only for the invariant quadratic loss. This is probably expected from Brown (66) and Brown and Fox (74). It follows from Brown (66) that under the loss described in Theorem 4.4.4, the standard estimate is admissible if $p=1$, although at this moment we do not know if admissibility for $p=2$ follows readily from Brown and Fox (74). We conjecture the standard estimate is admissible in two dimension.

3. The improved estimate in Theorem 4.4.4 bears similarity to the James-Stein estimate of the multinormal mean. This is expected since on making a log transform, the problem reduces to the estimation of a location vector. One should also observe that our improved estimate is practically the same as Berger's (80) for the other invariant loss $\sum_{i=1}^{p} (\delta_i\theta_i - 1)^2$.

CHAPTER 5

SECOND ORDER ADMISSIBILITY IN MULTIPARAMETER FAMILIES

5.1 Introduction

In Chapter 4, we discussed some aspects of multiparameter admissibility under exact risks through differential inequalities. In this chapter, we discuss another aspect via an approximating risk, and its interrelations with differential inequalities.

Let $X_1, X_2, \ldots, X_k$ be independent $\text{Bin}(n_i, \pi_i)$ random-variables, $i = 1, 2, \ldots, k$,

where $$\pi_i = \left\{1 + e^{-(\theta + \beta d_i)}\right\}^{-1} , \qquad \ldots(1.1)$$

θ is the only unknown parameter. In an asymptotic formulation of Berkson's problem of estimating θ on the basis of X_i's, Ghosh and Sinha (81) recently proved that the mle $(\hat{\theta})$ is always inadmissible, although in general neither of the mle $(\hat{\theta})$ and the Rao-Blackwellized version of Berkson's estimate $(\tilde{\theta})$ dominates the other. Kariya, Sinha, and Subramanyan (81) recently showed that $\tilde{\theta}$ cannot dominate $\hat{\theta}$ for large values of θ, even when one replaces the approximate risks by exact risks.

In this chapter, we extend the concept of second-order admissibility to multiparameter families satisfying the usual Cramer-Rao regularity conditions. We restrict attention to estimates of the form

$$(\hat{\theta}_1 + \frac{d_1(\hat{\theta})}{n}, \ldots, \hat{\theta}_p + \frac{d_p(\hat{\theta})}{n}) ,$$

where n is the (current) sample size, and $d_1, \ldots, d_p$ have one continuous partial derivatives. Restriction to this class can be justified by invoking the second-order efficiency of the mle. Formally, we declare an estimate $(\hat{\theta}_1 + \frac{c_1(\hat{\theta})}{n}, \ldots, \hat{\theta}_p + \frac{c_p(\hat{\theta})}{n})$ to be inadmissible if there exists another estimate

$(\hat{\theta}_1 + \frac{d_1(\hat{\theta})}{n}, \ldots, \hat{\theta}_p + \frac{d_p(\hat{\theta})}{n})$ such that

$$\sum_{i=1}^{p} E(\hat{\theta}_i + \frac{d_i(\hat{\theta})}{n} - \theta_i)^2 \leq \sum_{i=1}^{p} E(\hat{\theta}_i + \frac{c_i(\hat{\theta})}{n} - \theta_i)^2 \text{, upto } o(n^{-2}), \quad \ldots(1.2)$$

$\forall\ \theta$ (with strict inequality for some θ).

By a usual Taylor expansion, it follows that (1.2) is equivalent to

$$\sum_{i=1}^{p} g_i^2(\theta) + 2\sum_{i=1}^{p} g_i(\theta) b_i(\theta) + 2\sum_{i=1}^{p}\sum_{j=1}^{p} g_i^{j(1)}(\theta)\, I^{ij}(\theta) \leq 0 \ \forall\ \theta, \quad \ldots(1.3)$$

[strict inequality for some θ is understood],

where $g_i(\theta) = d_i(\theta) - c_i(\theta)$

$$\frac{b_i(\theta)}{n} = \frac{b_{oi}(\theta) + c_i(\theta)}{n} = \text{bias of } \hat{\theta}_i + \frac{c_i(\hat{\theta})}{n} \text{ upto } o(n^{-1}),$$

where $\frac{b_{oi}(\theta)}{n}$ = bias of the mle upto $o(n^{-1})$,

$$g_i^{j(1)}(\theta) = \frac{\partial}{\partial \theta_j} g_i(\theta),$$

$$((I^{ij}(\theta))) = I^{-1}(\theta),$$

where $I(\theta)$ is the Fisher information matrix which we assume is p.d.

We now specialize to the case of a p-fold product of independent one-parameter families, and $c_i(\theta) = c_i(\theta_i)$, in which case (1.3) reduces to

$$\sum_{i=1}^{p} g_i^2(\theta) + 2\sum_{i=1}^{p} g_i(\theta) b_i(\theta_i) + 2\sum_{i=1}^{p} g_i^{i(1)}(\theta)/I_{ii}(\theta_i) \leq 0 \quad \forall\, \theta . \qquad \ldots(1.4)$$

We assume $I_{ii}(\theta_i)$ is continuous in θ_i, $1 \leq i \leq p$. Henceforth, unbiased estimates will mean estimates unbiased upto $o(n^{-1})$ and admissibility will mean second-order admissibility.

In Section 5.2, by a slight variation of Stein's technique, we show that any unbiased estimate is admissible in dimension 2, if the components are separately admissible in dimension 1. In Section 5.3, sufficient conditions on the bias of an estimate are obtained for it to be inadmissible when $p \geq 3$. It follows that unbiased estimates are always inadmissible if $p \geq 3$. Finally, in Section 5.4, we give an example to illustrate the difficulty of generating admissible estimates by constructing Bayes solutions, even in one-dimension. These difficulties suggest that we should restrict ourselves to a suitable subclass of estimators, although the right form of restriction is not yet clear. In Section 5.4, we suggest some restrictions.

5.2 Admissibility of unbiased estimates in 2-dimension

Theorem 5.2.1 Let $\hat{\theta}_c = (\hat{\theta}_1 + \frac{c_1(\hat{\theta}_1)}{n} , \hat{\theta}_2 + \frac{c_2(\hat{\theta}_2)}{n})$ be an unbiased estimate of θ, such that $\hat{\theta}_i + \frac{c_i(\hat{\theta}_i)}{n}$ is admissible for θ_i, $i = 1,2$. Then $\hat{\theta}_c$ is admissible for θ.

<u>Proof</u> : Since $\hat{\theta}_c$ is unbiased, it follows from (1.4) that $\hat{\theta}_c$ is inadmissible if and only if there is a g such that

$$R(\theta,g) = \sum_{i=1}^{2} g_i^2(\theta) + 2 \sum_{i=1}^{2} g_i^{i(1)}(\theta)/I_{ii}(\theta_i) \leq 0 \quad \forall\, \theta \in \mathbb{R}^2 \qquad \ldots(2.1)$$

Defining $\eta_i = \eta_i(\theta_i) = \int_{\theta_o}^{\theta_i} I_{ii(u)}du$, $i = 1,2$, it follows that (2.1) holds if and only if

$$R(\eta,f) = \sum_{i=1}^{2} f_i^2(\eta) + 2 \sum_{i=1}^{2} f_i^{i(1)}(\eta) \leq 0 \quad \forall\, \eta \in \mathbb{R}^2, \qquad \ldots(2.2)$$

where $f_i(\eta) = g_i(\theta)$, $i = 1,2$.

Note that (η_1, η_2) span the whole of $\mathbb{R}^2$ by Ghosh and Sinha (81) since the components are admissible in dimension 1, and therefore $\eta_i(\theta_i) \longrightarrow \infty$ as $\theta_i \longrightarrow \infty$, and $\eta_i(\theta_i) \longrightarrow -\infty$ as $\theta_i \longrightarrow -\infty$. Also $\eta_i(\theta_i)$ is a strictly increasing function of θ_i, with $\eta_i'(\theta_i) = I_{ii}(\theta_i)$.

Therefore, $\hat{\theta}_c$ will be inadmissible if and only if (2.2) holds for some f (with one continuous partial derivatives). We now return to our original notation $R(\theta,g)$ in place of $R(\eta,f)$.

As in Stein (55), we now show that (2.2) holds for some g, if and only if it holds for a spherically symmetric h, where $h(\cdot)$ will be called spherically symmetric if $Lh(L'\theta) = h(\theta)$ for all orthogonal L, and all θ in $\mathbb{R}^2$.

The 'if' part of the claim is immediate. In order to prove the 'only if' part, we adopt a slight variation of Stein's (55) technique since our analysis is in terms of an approximation to the risk upto $o(n^{-2})$. If $g(\theta)$ is as in (2.2), define

$h(\theta) = \int_{O_2} Lg(L'\theta)\, d\mu(L)$, where $d\mu$ stands for the invariant probability-measure on the orthogonal group O_2. h is spherically symmetric since μ is invariant. ...(2.3)

Define $\frac{\partial g}{\partial \theta} = ((\frac{\partial g_i}{\partial \theta_j}))_{2\times 2}$,

$$\frac{\partial h}{\partial \theta} = ((\frac{\partial h_i}{\partial \theta_j}))_{2\times 2}. \quad \text{...(2.4)}$$

Then
$$\begin{aligned} R(\theta,h) &= \|h(\theta)\|^2 + 2\,\mathrm{tr}\frac{\partial h}{\partial \theta}(\theta) \\ &= \|\int_{O_2} Lg(L'\theta)\, d\mu(L)\|^2 + 2\,\mathrm{tr}\int_{O_2} L\frac{\partial g}{\partial \theta}(L'\theta)L'\, d\mu(L) \\ &\leq \int_{O_2} \|g(L'\theta)\|^2 d\mu(L) + 2\int_{O_2}\mathrm{tr}\frac{\partial g}{\partial \theta}(L'\theta)\, d\mu(L) \\ &= \int_{O_2} R(L'\theta,g)\, d\mu(L) \\ &\leq 0 \;\forall\, \theta \;\;(< 0 \text{ for atleast one } \theta). \end{aligned} \quad \text{...(2.5)}$$

In (2.5) above, the second equality follows because differentiation can be carried inside the integral by a straightforward application of Dominated Convergence Theorem. This proves the claim.

The proof of Theorem 2.1 now follows in the lines of Stein (55)

Remark : Theorem 5.2.1 indeed also follows from a result of Peng (75), who shows that the only solution to (2.2) is $f \equiv 0$.

5.3 Inadmissibility for $p \geq 3$.

In this section, we proceed to show that any unbiased estimate is inadmissible for $p \geq 3$. As pointed out in Section 5.1, in order to establish inadmissibility, we need to explicitly solve the differential inequality (1.4) in g. Towards this end, we will use Lemma 4.4.2 of Chapter 4.

Theorem 5.3.1 Let $\hat{\theta}_c = (\hat{\theta}_1 + \frac{c_1(\hat{\theta}_1)}{n}, \ldots, \hat{\theta}_p + \frac{c_p(\hat{\theta}_p)}{n})$ be such that

$$\inf_{\theta_i} \int_{\theta_o}^{\theta_i} b_i(u)\, I_{ii}(u) = K_i > -\infty, \text{ for all } i = 1,2,\ldots,p \qquad \ldots(3.1)$$

Then $\hat{\theta}_c$ is inadmissible for $p \geq 3$.

Proof : Let $g_i(\theta) = h_i(\theta)\, e^{-\int_{\theta_o}^{\theta_i} b_i(u)\, I_{ii}(u)du}$, $i = 1,2,\ldots,p$.

Then,
$$g_i^{i(1)}(\theta) = -b_i(\theta_i)I_{ii}(\theta_i)g_i(\theta) + h_i^{i(1)}(\theta)\, e^{-\int_{\theta_o}^{\theta_i} b_i(u)I_{ii}(u)du} \qquad \ldots(3.2)$$

Therefore, (1.4) reduces to

$$\sum_{i=1}^{p} h_i^2(\theta)\, e^{-2\int_{\theta_o}^{\theta_i} b_i(u)I_{ii}(u)du} + 2\sum_{i=1}^{p} h_i^{i(1)}(\theta)\, \frac{e^{-\int_{\theta_o}^{\theta_i} b_i(u)I_{ii}(u)du}}{I_{ii}(\theta_i)} \leq 0$$

$$\forall\, \theta \in \mathbb{R}^p \qquad \ldots(3.3)$$

We now identify h_i as ϕ_i, $e^{-2\int_{\theta_o}^{\theta_i} b_i(u)I_{ii}(u)du}$ as w_i,

$e^{-\int_{\theta_o}^{\theta_i} b_i(u)I_{ii}(u)du} \Big/ I_{ii}(\theta_i)$ as υ_i and ψ as 2 in Lemma 4.4.2.

It follows that (4.1) of Lemma (4.4.2) in Chapter 4 holds, if (3.1) is satisfied, for one can then take

$2K = e^{-2m}\min(K_1, K_2, \ldots, K_p)$, $\beta = 2$, $d_j = 1 \;\; \forall\, j = 1,2,\ldots,p$.

Therefore, $h_i(\theta) = \dfrac{-c\;\xi_i(\theta_i)}{\sum_{j=1}^{p} \xi_j^2(\theta_j) + b}$

is a solution to (3.3) for $0 < c < K^{-1}(p-2)$, and $b > 0$, by an application of Lemma 4.4.2 in Chapter 4,

where, $\xi_i(\theta_i) = \int_{\theta_o}^{\theta_i} I_{ii}(u)\, e^{\int_{\theta_o}^{u} b_i(z)I_{ii}(z)dz}\, du$, $i = 1,2,\ldots,p$.

Equivalently,

$$g_i(\theta) = e^{-\int_{\theta_o}^{\theta_i} b_i(u)I_{ii}(u)du} \times -\frac{c\;\xi_i(\theta_i)}{\sum_{j=1}^{p} \xi_j^2(\theta_j) + b}, \quad i = 1,2,\ldots,p,$$

is a solution to (1.4) for $p \geq 3$, whenever $\inf_{\theta_i} \int_{\theta_o}^{\theta_i} b_i(u)I_{ii}(u)du > -\infty$, $1 \leq i \leq p$.

This proves the theorem.

<u>Remarks</u> 1. The transformation in (3.2), motivated by Ping (64), was made to reduce (1.4) to the form suggested in Lemma 4.4.2 of Chapter 4.

2. In the special case when we consider an unbiased estimate, $\int_{\theta_0}^{\theta_i} b_i(u)\, I_{ii}(u)du \equiv 0$. Therefore, unbiased estimates are always inadmissible for $p \geq 3$.

3. It was shown in Ghosh and Sinha (81), that a necessary and sufficient condition for admissibility of an estimate with bias $b(\theta)$ is that

$$\int_{\theta_0}^{\infty} I(\theta)\, e^{-\int_{\theta_0}^{\theta} b(u)I(u)du}\, d\theta = \infty$$

$$\text{and } \int_{-\infty}^{\theta_0} I(\theta)\, e^{\int_{\theta}^{\theta_0} b(u)I(u)du}\, d\theta = \infty,$$

$-\infty < \theta_0 < \infty$. It is suggested by these conditions that for an estimate to be admissible, $\int_{\theta_0}^{\theta} b(u)\, I(u)du$ should be large negative as $\theta \longrightarrow \pm\infty$. The condition (3.1) of Theorem 5.3.1 seems to suggest exactly the opposite of that. This is primarily because Lemma 4.4.2 of Chapter 4 is essentially a tool for proving inadmissibility of the unbiased estimates.

5.4 Existence of Bayes estimates

In dimension 1, taking the risk of an estimate $\hat{\theta} + \frac{c(\hat{\theta})}{n}$ as $c^2(\theta) + 2c(\theta)b_0(\theta) + \frac{2c'(\theta)}{I(\theta)}$ (see Ghosh and Sinha (81)), a natural question to be asked is does there exist a function $c(\theta)$ in $C^1(\mathbb{R})$ which minimizes the integrated Bayes risk

$$\int \left\{ c^2(\theta) + 2c(\theta)b_o(\theta) + \frac{2c'(\theta)}{I(\theta)} \right\} dG(\theta)$$

where dG is a sufficiently smooth prior? For simplicity, let us take $I(\theta) \equiv 1$. It is easy to see that then this minimization problem is equivalent to minimizing $\int \left\{ b^2(\theta) + 2b'(\theta) \right\} q(\theta)d\theta$ over $C^1(\mathbb{R})$, where $\frac{dG(\theta)}{d\theta} = q(\theta)$, and $b(\theta) = b_o(\theta) + c(\theta)$; it is implicitly assumed $\int \left\{ b_o^2(\theta) + 2b_o'(\theta) \right\} q(\theta)d\theta$ exists. If a minimizing $b(\theta)$ exists in $C^1(\mathbb{R})$, then it must necessarily be equal to $\frac{q'(\theta)}{q(\theta)}$ at all points θ in whose neighbourhood q'/q is continuous. This is shown by essentially verifying that the Euler equation for this minimization problem is actually a necessary condition for the minima. This is done in the following way. We assume $dG(\theta)$ is so smooth that $\frac{q'(\theta)}{q(\theta)}$ is a.e. continuous. If $b^*(\theta)$ is a minima, then for every $\varepsilon(\theta)$ in $C^1(\mathbb{R})$,

$$\int \left\{ \varepsilon^2(\theta) + 2b^*(\theta)\varepsilon(\theta) + 2\varepsilon'(\theta) \right\} q(\theta)d\theta \geq 0 . \qquad \ldots(4.1)$$

Hence, writing $\varepsilon(\theta) = \varepsilon\, \eta(\theta)$, for any $\eta(\theta) \in C_c(\mathbb{R})$ such that $\int \left\{ 2b^*(\theta)\,\varepsilon(\theta) + 2\varepsilon'(\theta) \right\} q(\theta)d\theta \neq 0$, (4.1) can be violated by choosing ε suitably (as positive or negative).

Therefore, $\int \left\{ b^*(\theta)\,\varepsilon(\theta) + \varepsilon'(\theta) \right\} q(\theta)d\theta = 0 , \qquad \ldots(4.2)$

for every $\varepsilon(\theta) \in C_c(\mathbb{R})$.

If now $\varepsilon(\theta)$ is any function in $C^1(\mathbb{R})$ with a compact support, then by integrating by parts, (4.2) gives

$$\int \varepsilon(\theta) \left\{ b^*(\theta) - \frac{q'(\theta)}{q(\theta)} \right\} q(\theta) d\theta = 0 \quad \ldots(4.3)$$

$$\Rightarrow b^*(\theta) = \frac{q'(\theta)}{q(\theta)} \text{ a.e.} \quad \ldots(4.4)$$

(at all points θ in whose neighbourhood $q'(\theta)/q(\theta)$ is continuous)

Since $b^*(\theta) \in C^1(\mathbb{R})$, it follows that (4.4) must actually be true for every θ. Note that we have tacitly assumed that the integral in (4.2) exists for all $\varepsilon(\theta) \in C_c(\mathbb{R})$, which will be true if for example $\frac{q'(\theta)}{q(\theta)}$ is continuous.

If $q(\theta)$ has a compact support, say, $[0,1]$, but does not have a continuous terminal contact at the end-points 0,1, then it was shown by Ghosh, Sinha, and Joshi (81) (See their example 2, Section 6) that the above infimum could be $-\infty$. If $q(\theta)$ has a compact support, say, $[0,1]$, and also a smooth terminal contact at 0,1, then $\frac{q'(\theta)}{q(\theta)}$ cannot be continuous in the closed interval $[0,1]$. However, by restricting to closed sub-intervals $[\varepsilon, 1-\varepsilon]$, $\varepsilon > 0$, the argument leading to (4.4) once again shows that $b^*(\theta)$ must be $\frac{q'(\theta)}{q(\theta)}$ for every θ, which is a contradiction. However, it follows from Theorem 5.1 of Ghosh, Sinha, and Joshi (81) that the integrated Bayes risk in the sense considered in this section has a finite lower bound, and hence a finite minimum. The preceding argument then says that this infimum cannot be attained. In this section, we give an example of a $q(\theta)$ not having a compact support, namely the standard normal, such that $\int \{b^2(\theta) + 2b'(\theta)\} q(\theta)d\theta$ need not even exist in the sense of Lebesgue, thus illustrating the difficulties in generating admissible estimates via Bayesian methods. We conjecture that the integral in the following example does not exist in the Riemann sense as well. Finally, in Proposition§4.1, we give some sufficient conditions for the existence of a Bayes estimate.

Example Let $b(\theta) = e^{\theta^2/2} \text{Sin}(e^{\theta^2/2}), -\infty < \theta < \infty$.

Then $b'(\theta) = \theta e^{\theta^2/2} \text{Sin}(e^{\theta^2/2}) + \theta e^{\theta^2} \cos(e^{\theta^2/2})$ ...(4.5)

Let $dG(\theta) = e^{-\theta^2/2} d\theta$.

Then, $\int_{\mathbb{R}} \{b^2(\theta) + 2b'(\theta)\} dG(\theta)$

$$= \int_{\mathbb{R}} e^{\theta^2/2} \left\{ \text{Sin}^2(e^{\theta^2/2}) + 2\theta\cos(e^{\theta^2/2}) + 2\theta e^{-\theta^2/2}\text{Sin}(e^{\theta^2/2}) \right\} d\theta \quad \text{...(4.6)}$$

In order to show that the integral in (4.6) does not exist, we consider the integral over the measurable subset $[0,\infty)$. Making the transformation $e^{\theta^2/2} = y$, one has,

$$\int_{[0,\infty)} \{b^2(\theta) + 2b'(\theta)\} e^{-\theta^2/2} d\theta$$

$$= \int_{[1,\infty)} \left\{ \frac{\text{Sin}^2 y}{\sqrt{2\log y}} + 2\cos y + \frac{2\,\text{Sin}\, y}{y} \right\} dy \quad \text{...(4.7)}$$

Considering successively sequences of intervals

$I_n = \left[(4n+1)\frac{\pi}{2} - \delta,\ (4n+1)\frac{\pi}{2} + \delta\right]$ of fixed length 2δ such that $\text{Sin}\, y > 1 - \varepsilon$ ($\varepsilon > 0$ fixed) for $y \in I_n$, $n \geq n_o$ (say), it is easy to show that (4.7) = + ∞. On the other hand, considering similar intervals of fixed length 2α with centre at $(2n+1)\pi$, $n > n_0$, such that $\cos y \leq -1 + \varepsilon$, it can be shown that (4.7) = −∞. Thus the integral in (4.6) cannot exist in the sense of Lebesgue. The example above suggests that if $b(\theta)$ varies in the entire $C^1(\mathbb{R})$, then the technique of generating admissible estimates by obtaining Bayes solutions may not succeed even in one dimension. It is therefore necessary to take a suitable sub-class of estimators previously considered. That the infinum exists if we allow only well-behaved b's is demonstrated by the following proposition. $q(\theta)$ is taken as $e^{-\theta^2/2}$.

<u>Proposition 5.4.1</u> If $b(\theta)$ varies in a sub-class of $C^1(\mathbb{R})$ such that $\int \left\{ b^2(\theta) + 2b'(\theta) \right\} e^{-\theta^2/2}\, d\theta$ exists, and

either i) 0 is a limit point of $b(\theta)\, e^{-\frac{1}{2}\theta^2}$ as $|\theta| \to \infty$

or ii) $\lim\limits_{|\theta| \to \infty} \dfrac{b'(\theta)}{b^2(\theta)}$ exists (with possibly different limits),

then $\inf \int \left\{ b^2(\theta) + 2b'(\theta) \right\} e^{-\theta^2/2}\, d\theta$ is finite, and the infimum is attained by $b_0(\theta) = -\theta = \dfrac{q'(\theta)}{q(\theta)}$.

<u>Proof</u>: Let i) hold. Then for any $b \in C^1(\mathbb{R})$ such that $\int \left\{ b^2(\theta) + 2b'(\theta) \right\} e^{-\theta^2/2}\, d\theta$ exists,

$$\int \left\{ b^2(\theta) + 2b'(\theta) \right\} e^{-\theta^2/2}\, d\theta$$

$$= \int \left\{ b^2(\theta) + 2\theta\, b(\theta) \right\} e^{-\theta^2/2}\, d\theta$$

(integrating by parts and taking limits along sequences such that $b(\theta)\, e^{-\frac{1}{2}\theta^2} \longrightarrow 0$)

$$= \int \left\{ (\theta + b(\theta))^2 - \theta^2 \right\} e^{-\theta^2/2}\, d\theta$$

$$\geq \int -\theta^2\, e^{-\theta^2/2}\, d\theta$$

$$= \int (\theta^2 - 2)\, e^{-\theta^2/2}\, d\theta$$

$$= \int \left\{ b_o^2(\theta) + 2b_o'(\theta) \right\} e^{-\theta^2/2}\, d\theta . \qquad \ldots(4.8)$$

Suppose now **ii)** holds but not **i)**.

Define $K(\theta) = b(\theta)\, e^{-\theta^2/2}$

Since **i)** does not hold, $K(\theta)$ is bounded away from 0 as $|\theta| \longrightarrow \infty$. Note now

$$\frac{b'(\theta)}{b^2(\theta)} = \frac{K'(\theta)\, e^{-\frac{1}{2}\theta^2}}{K^2(\theta)} + \frac{\theta\, e^{-\frac{1}{2}\theta^2}}{K(\theta)} . \qquad \ldots(4.9)$$

Integrating by parts,

$$\int_{|\theta| > a} \frac{K'(\theta)}{K^2(\theta)}\, e^{-\theta^2/2}\, d\theta = c - \int_{|\theta| > a} \frac{\theta\, e^{-\theta^2/2}}{K(\theta)}\, d\theta , \qquad \ldots(4.10)$$

since $K(\theta)$ is bounded away from 0 as $|\theta| \longrightarrow \infty$. Also,

$\int_{|\theta| > a} \frac{\theta\, e^{-\theta^2/2}}{K(\theta)}\, d\theta < \infty$, implying, by (4.9) and (4.10),

$\int_{|\theta| > a} \frac{b'(\theta)}{b^2(\theta)}\, d\theta$ exists and is finite. But, by hypothesis,

$\lim\limits_{|\theta| \to \infty} \dfrac{b'(\theta)}{b^2(\theta)}$ exist, and, now these limits must necessarily be 0.

Now observe,

$$\int\limits_{|\theta|>a} \left\{ b^2(\theta) + 2b'(\theta) \right\} e^{-\frac{\theta^2}{2}} d\theta$$

$$= \int\limits_{|\theta|>a} b^2(\theta) \left\{ 1 + \frac{2b'(\theta)}{b^2(\theta)} \right\} e^{-\frac{\theta^2}{2}} d\theta$$

$$\geq \frac{1}{2} \int\limits_{|\theta|>a} b^2(\theta)\, e^{-\theta^2/2} d\theta \quad \text{(by choosing a large enough)}$$

$$= \frac{1}{2} \int\limits_{|\theta|>a} K^2(\theta)\, e^{\theta^2/2} d\theta$$

$$= \infty .$$

This completes the proof of Proposition 5.4.1.

Remark Hypothesis i) of Proposition 5.4.1 is to be compared to a similar condition imposed in Brown (81), namely, $b(\theta) \in L^2(e^{-\theta^2/2\sigma^2} d\theta)$ for some $\sigma > 0$, in which Brown has discussed problems relating to a general differential inequality arising out of estimation problems, although from a motivation different from ours.

CHAPTER 6

ESTIMATION OF THE GENERALIZED VARIANCE THROUGH THE WISHART MATRIX

6.1 Introduction

In Chapters 4 and 5, we discussed some aspects of multiparameter admissibility through the study of differential inequalities. The problems discussed in Chapters 4 and 5 typically involved estimation of a multidimensional parametric function in a multiparameter set up. Although we now have a considerable literature on this aspect of multiparameter admissibility, much is yet to be done regarding Stein effect on one dimensional parametric functions in a multiparameter family. In this chapter, we consider one such problem.

Let $X_1, \ldots, X_N$ be iid p variate normal with mean μ and dispersion matrix Σ (p.d.). Consider the problem of estimating $|\Sigma|$ on the basis of $X_1, \ldots, X_N$ under the loss $(a - |\Sigma|)^2 |\Sigma|^{-2}$. The sample mean $\bar{X}$ and the Wishart matrix S are jointly sufficient for (μ, Σ) and the problem of estimating $|\Sigma|$ remains invariant under the transformations.

$$\bar{X} \longrightarrow A\bar{X} + b, \quad S \longrightarrow ASA',$$

where A is full rank and the equivariant estimates are of the form $\phi(\bar{X}, S) = c|S|$, $c > 0$. There exists a best multiple of $|S|$, given by

$$c = (K - p + 2)! \;/\; (K+2)!, \text{ where } K = N - 1.$$

For $p = 1$, Stein (64) showed that the standard estimate of $|\Sigma|$ $(= \sigma^2)$

is inadmissible, dominating it by a 'testimator'. The idea of Stein was exploited by Shorrock and Zidek (76) who proved the same result for all $p \geq 1$. Recently, Weerahandi, Tsui and Zidek (80) proved that the standard estimate of the generalized residual variance is inadmissible. It is not known, however, if in the class of estimates that depend on the Wishart matrix S alone, or equivalently when the mean μ is known, this estimate of $|\Sigma|$ is admissible or not. This problem, apart from the reason mentioned in the first paragraph, is interesting also because of its difficulty. In this chapter we have proved using Brown (66) that this natural estimate of $|\Sigma|$ is admissible in a further sub-class of estimates which are functions of $|S|$ alone. The original question raised before is, however, still open.

6.2 Preliminaries

It is well known that $\frac{|S|}{|\Sigma|}$, under Σ, is distributed as the product of p independent chi-square variables $X_1, \ldots, X_p$, with $E(X_i) = K - p + i$, $1 \leq i \leq p$. Therefore, $|\Sigma|$ is a scale parameter for the distribution of $|S|$. Hence, $\log |\Sigma|$ is a location parameter for the distribution of $\log |S|$. Also, note that the loss $L(a, |\Sigma|) = (a - |\Sigma|)^2 |\Sigma|^{-2}$ is a function $a/|\Sigma|$ alone. It is well known that estimating $|\Sigma|$ on the basis of $|S|$ under the loss $L(a, |\Sigma|) = W(a/|\Sigma|)$ is equivalent to estimating $\theta = \log|\Sigma|$ on the basis of $\log|S|$ under the loss $L'(\theta, a) = W'(\theta - a)$, where $W'(x) = W(e^x)$. In our problem, $W(x) = (x - 1)^2$.
Hence, $W'(x) = (e^x - 1)^2$.

We will let $p(x)$ denote the density of $X = |S|$ when $|\Sigma| = 1$. Consequently, the density of $Y = \log X$ when $\theta = 0$ is given by

$$q(y) = e^y \, p(e^y), \quad -\infty < y < \infty; \qquad \ldots(2.1)$$

A best invariant estimate of θ on the basis of the observation y exists under the loss $L'(\theta - a) = (e^{\theta - a} - 1)^2$, and as pointed out in Brown (66), there is no loss in assuming that the variable Y has been so transformed that this best invariant estimate is Y itself, or in other words,

$$E_{\theta=0}\,(e^{2Y}) = E_{\theta=0}\,(e^{Y}) \qquad \ldots(2.2)$$

In terms of $X = |S|$, the meaning of this is that a scale change has been done to X so that $E(X) = E(X^2)$ under the density $p(x)$. The original density $q(y)$ in (2.1) changes to $c_1 e^y p(c_2 e^y)$ because of this change in X; but the constants c_1, c_2 are not going to affect the subsequent analysis and we will use the same notation $p(x)$ and $q(y)$ for the densities of X and Y (under $|\Sigma| = 1$ and $\theta = 0$ respectively). We will further assume without loss that scale changes have been done to $X_1, \ldots, X_p$ so that for $1 \leq i \leq p$, $E(X_i) = E(X_i^2)$, i.e., the density of X_i is of the form

$$f(x_i) = K_i \, e^{-\alpha_i x_i} \, x_i^{\beta_i - 1}, \quad \text{where } \beta_i + 1 = \alpha_i \qquad \ldots(2.3)$$

As already pointed out, $\frac{(K-p+2)!}{(K+2)!} |S|$ is admissible for $|\Sigma|$ if and only if Y is admissible for θ in the transformed problem. We are now in a position to appeal to Brown (66). We quote below a theorem of Brown (66) for future reference.

6.3 Main Result

Theorem 6.3.1 (Brown (66)) Let θ be a real location parameter for the distribution of Y in dimension 1. Let R_0 = risk of the best invariant estimate (Y) of θ under a loss $L'(\theta,a) = W'(\theta-a)$. Assume $R_0 < \infty$ and that

(i) $R(Y+c_i, \theta) \longrightarrow R(Y,\theta)$ as $i \longrightarrow \infty$

$\Rightarrow c_i \longrightarrow 0$ as $i \longrightarrow \infty$

($Y + c_i$ and Y are, incidentally, invariant estimates, and have constant risks)

(ii) $$\int_0^\infty \left\{ \sup_\gamma \int_{-\lambda}^{\lambda} \left[W'(y) - W'(y+\gamma) \right] q(y)dy \right\} d\lambda < \infty$$

(iii) $$\int |y|\, W'(y)\, q(y)\, dy < \infty .$$

Then the best invariant estimate (Y) is admissible for estimating θ.

Theorem 6.3.2 Let $S \sim W_p(K,\Sigma)$, $K \geq p$, Σ p.d. Then $\frac{(K-p+2)!}{(K+2)!}|S|$ is an admissible estimator of $|\Sigma|$ under the loss $(a - |\Sigma|)^2|\Sigma|^{-2}$ in the class of estimates that depend on $|S|$ alone.

Proof : We verify conditions (i), (ii) and (iii) of Theorem 6.3.1 in the transformed problem. (i) and (iii) are relatively easy to verify, and we verify them first.

Note $R(Y+c_i, \theta) \longrightarrow R(Y,\theta)$ as $i \longrightarrow \infty$

$$\Leftrightarrow E_{\theta=0}\,(e^{c_i} e^{Y} - 1)^2 \longrightarrow E_{\theta=0}\,(e^{Y}-1)^2 \text{ as } i \longrightarrow \infty$$

$$\Leftrightarrow E_{\theta=0}\left[(e^{2c_i} - 1)\, e^{2Y} - 2(e^{c_i} - 1)\, e^{Y}\right] \longrightarrow 0 \text{ as } i \longrightarrow \infty$$

$$\Leftrightarrow \quad (e^{c_i}-1)\, E_{\theta=0}\left[(e^{c_i}+1)\, e^{2Y}-2e^{Y}\right] \longrightarrow 0 \text{ as } i \longrightarrow \infty \quad \ldots(3.1)$$

If $c_i \not\to 0$ as $i \longrightarrow \infty$, let a subsequence converge to $c \neq 0$. We denote this subsequence by $\{c_i\}$ itself. Note $e^{c_i}-1 \to e^{c}-1 \neq 0$.

Hence from (3.1), $e^{c_i}+1 \longrightarrow \dfrac{2E_{\theta=0}(e^{Y})}{E_{\theta=0}(e^{2Y})}$

$$\Leftrightarrow \quad e^{c}+1 = 2 \qquad \text{(by (2.2))}$$

$$\Leftrightarrow \quad e^{c} = 1,$$

a contradiction. This verifies (i).

To verify (iii), we observe that

$$\int |y|\; W'(y)\; q(y)\, dy$$

$$= \int |y| (e^{y}-1)^2\, e^{y} p(e^{y})\, dy$$

$$= \int_{y>0} y(e^{y}-1)^2\, e^{y} p(e^{y})\, dy - \int_{y\leq 0} y(e^{y}-1)^2\, e^{y} p(e^{y})\, dy$$

$$= \int_{\log x>0} \log x\, (x-1)^2\, p(x) dx - \int_{\log x \leq 0} \log x\, (x-1)^2 p(x)\, dx$$

$$= \int |\log x|\, (x-1)^2 p(x)\, dx \quad \ldots(3.2)$$

Note that for $x > 1$, $|\log x| = \log x \leq A_1\, x^{\varepsilon}$ and for $x \leq 1$, $|\log x| = \log \frac{1}{x} \leq A_2\, x^{-\varepsilon}$, for any $\varepsilon > 0$, where A_1, A_2 are two constants depending on ε. Therefore, (3.2) is finite if

$$\int x^{2+\varepsilon}\, p(x)\, dx < \infty \quad \ldots(3.3)$$

$$\int x^{-\varepsilon}\, p(x)\, dx < \infty \quad \ldots(3.4)$$

for some $\varepsilon > 0$.

Now recall that X is distributed as $\prod_{i=1}^{p} X_i$, where $X_1,\ldots,X_p$ are independent with densities as given by (2.3). It is clear $\varepsilon > 0$ can be chosen so that $E(X_i^{-\varepsilon}) < \infty$, for $1 \leq i \leq p$. Thus (3.4) is true. (3.3) is immediate.

Finally, we now go on to verifying (ii).

For a fixed γ, let $e^{\gamma} = c$. Fix $\lambda > 0$.

Then $\sup_{\gamma} \int_{-\lambda}^{\lambda} [W'(y) - W'(y+\gamma)]\, q(y)dy$

$$= \sup_{c>0} \int_{-\lambda}^{\lambda} [(e^y - 1)^2 - (c\,e^y - 1)^2] e^y\, p(e^y)dy$$

$$= \sup_{c>0} \int_{e^{-\lambda}}^{e^{\lambda}} [(x-1)^2 - (cx-1)^2]\, p(x)dx \qquad \ldots(3.5)$$

It is easy to verify that for a fixed $\lambda > 0$,

$\int_{e^{-\lambda}}^{e^{\lambda}} [(x-1)^2 - (cx-1)^2] p(x)dx$ is maximized when

$$c = \int_{e^{-\lambda}}^{e^{\lambda}} xp(x)dx \Big/ \int_{e^{-\lambda}}^{e^{\lambda}} x^2 p(x)dx \,. \qquad \ldots(3.6)$$

Therefore, defining, $g(\lambda) = \int_{e^{-\lambda}}^{e^{\lambda}} xp(x)dx$

$$h(\lambda) = \int_{e^{-\lambda}}^{e^{\lambda}} x^2\, p(x)\, dx \,, \qquad \ldots(3.7)$$

from (3.5), (3.6), it follows that

$$\sup_{y} \int_{-\lambda}^{\lambda} \left[w'(x) - w'(x+y) \right] q(x)dx = \frac{[g(\lambda) - h(\lambda)]^2}{h(\lambda)} \quad \ldots(3.8)$$

(ii) will be verified if we can show $\int_0^\infty \frac{[g(\lambda) - h(\lambda)]^2}{h(\lambda)} d\lambda < \infty$

$$\ldots(3.9)$$

For λ in an interval $0 < \lambda < a$, $h(\lambda)$, $g(\lambda)$, and $\frac{g^2(\lambda)}{h(\lambda)}$ are bounded. Therefore, it is enough to show that the integral in (3.9) is finite at the tail near ∞. Since $h(\lambda)$ is increasing, and by (2.2), $g(\lambda) - h(\lambda) \longrightarrow 0$ as $\lambda \longrightarrow \infty$, it is enough to show that $\int_0^\infty |g(\lambda) - h(\lambda)| d\lambda < \infty$. We will show this now. We prove this when $p = 2$. For a general p, the proof is similar, and will be indicated.

Note $g(\lambda) = E(X_1 X_2 \, I_{e^{-\lambda} < X_1 X_2 < e^{\lambda}})$,

$$h(\lambda) = E(X_1^2 X_2^2 \, I_{e^{-\lambda} < X_1 X_2 < e^{\lambda}}),$$

where X_1, X_2 are as before.

Therefore, $h(\lambda) - g(\lambda)$

$$= E\left\{ X_2^2 I_{\frac{e^{-\lambda}}{X_1} < X_2 < \frac{e^{\lambda}}{X_1}} \right\} X_1^2 - E\left\{ X_1 I_{\frac{e^{-\lambda}}{X_2} < X_1 < \frac{e^{\lambda}}{X_2}} \right\} X_2$$

$$= E\left\{ (X_2^2 - X_2) I_{\frac{e^{-\lambda}}{X_1} < X_2 < \frac{e^{\lambda}}{X_1}} \right\} X_1^2 + E\left\{ (X_1^2 - X_1) I_{\frac{e^{-\lambda}}{X_2} < X_1 < \frac{e^{\lambda}}{X_2}} \right\} X_2 \quad \ldots(3.10)$$

Now, $K_2 \int_{\frac{e^{-\lambda}}{x_1}}^{\frac{e^{\lambda}}{x_1}} (x_2^2 - x_2)\, e^{-\alpha_2 x_2}\, x_2^{\beta_2 - 1}\, dx_2$

$$= -\frac{K_2}{\alpha_2} e^{-\alpha_2 x_2} x_2^{\beta_2 + 1} \Bigg|_{\frac{e^{-\lambda}}{x_1}}^{\frac{e^{\lambda}}{x_1}} \qquad \ldots(3.11)$$

integrating by parts and using $\beta_2 + 1 = \alpha_2$.

One similarly has,

$$E\left\{ (X_1^2 - X_1)\, I_{\frac{e^{-\lambda}}{x_2} < X_1 < \frac{e^{\lambda}}{x_2}} \right\}$$

$$= -\frac{K_1}{\alpha_1} e^{-\alpha_1 x_1} x_1^{\beta_1 + 1} \Bigg|_{\frac{e^{-\lambda}}{x_2}}^{\frac{e^{\lambda}}{x_2}} \qquad \ldots(3.12)$$

Now, from (3.10), taking conditional expectations given X_1, X_2 respectively in the first and the second term, and using independence of X_1, X_2, one has,

$$\int_0^\infty |h(\lambda) - g(\lambda)|\, d\lambda$$

$$\leq \int_0^\infty \left\{ \frac{K_2}{\alpha_2} E\left[e^{-\alpha_2 \frac{e^{-\lambda}}{X_1}} \left(\frac{e^{-\lambda}}{X_1}\right)^{\beta_2 + 1} + e^{-\alpha_2 \frac{e^{\lambda}}{X_1}} \left(\frac{e^{\lambda}}{X_1}\right)^{\beta_2 + 1} \right] X_1^2 \right.$$

$$\left. + \frac{K_1}{\alpha_1} E\left[e^{-\alpha_1 \frac{e^{-\lambda}}{X_2}} \left(\frac{e^{-\lambda}}{X_2}\right)^{\beta_1 + 1} + e^{-\alpha_1 \frac{e^{\lambda}}{X_2}} \left(\frac{e^{\lambda}}{X_2}\right)^{\beta_1 + 1} \right] X_2 \right\} d\lambda \qquad \ldots(3.13)$$

We handle the first term in (3.13); the second term is similarly handled.

The first term in (3.13)

$$= \frac{K_1K_2}{\alpha_2} \int_0^\infty \Big[\int_0^\infty e^{-\alpha_2 \frac{e^{-\lambda}}{x_1}} \left(\frac{e^{-\lambda}}{x_1}\right)^{\beta_2+1} x_1^2 \, e^{-\alpha_1 x_1} \, x_1^{\beta_1-1} \, dx_1$$

$$+ \int_0^\infty e^{-\alpha_2 \frac{e^{\lambda}}{x_1}} \left(\frac{e^{\lambda}}{x_1}\right)^{\beta_2+1} x_1^2 \, e^{-\alpha_1 x_1} \, x_1^{\beta_1-1} \, dx_1 \Big] d\lambda \quad ..(3.14)$$

In (3.14), we integrate first with respect to λ, and then with respect to x_1, by interchanging the order of integration.

Now, $\int_0^\infty e^{-\alpha_2 \frac{e^{-\lambda}}{x_1}} \left(\frac{e^{-\lambda}}{x_1}\right)^{\beta_2+1} d\lambda$

$$= \int_0^{\frac{1}{x_1}} e^{-\alpha_2 u} \, u^{\beta_2} \, du \,, \quad \text{by substituting } \frac{e^{-\lambda}}{x_1} = u. \quad ...(3.15)$$

Similarly, $\int_0^\infty e^{-\alpha_2 \frac{e^{\lambda}}{x_1}} \left(\frac{e^{\lambda}}{x_1}\right)^{\beta_2+1} d\lambda$

$$= \int_{\frac{1}{x_1}}^\infty e^{-\alpha_2 u} \, u^{\beta_2} \, du \,, \quad \text{by substituting } \frac{e^{\lambda}}{x_1} = u. \quad ..(3.16)$$

Therefore, the first term in (3.13)

$$= \frac{K_1 K_2}{\alpha_2} \left[\int_0^\infty \left\{ \int_0^{\frac{1}{x_1}} e^{-\alpha_2 u} u^{\beta_2} du + \int_{\frac{1}{x_1}}^\infty e^{-\alpha_2 u} u^{\beta_2} du \right\} x_1^2 e^{-\alpha_1 x_1} x_1^{\beta_1 - 1} dx_1 \right]$$

$$= \frac{K_1 K_2}{\alpha_2} \frac{\Gamma(\beta_2+1)}{\alpha_2^{\beta_2+1}} \int_0^\infty x_1^{\beta_1+1} e^{-\alpha_1 x_1} dx_1$$

$$= \frac{K_1 K_2}{\alpha_2} \frac{\Gamma(\alpha_2)}{\alpha_2^{\alpha_2}} \frac{\Gamma(\beta_1+2)}{\alpha_1^{\beta_1+2}}$$

$$= \frac{K_1 K_2 \, \Gamma(\alpha_2) \, \Gamma(\alpha_1+1)}{\alpha_2^{\alpha_2+1} \, \alpha_1^{\alpha_1+1}} < \infty \qquad \ldots(3.17)$$

The second term in (3.13) is similarly handled.

This proves $\int_0^\infty |h(\lambda) - g(\lambda)| d\lambda < \infty$, and the proof of Theorem 6.3.2 is complete, when $p = 2$.

For a general p, we write $h(\lambda) - g(\lambda)$ as

$$E\left\{ (X_1^2 - X_1) I_{\frac{e^{-\lambda}}{X_2 \ldots X_p} < X_1 < \frac{e^{\lambda}}{X_2 \ldots X_p}} \right\} X_2^2 \ldots X_p^2$$

$$+ E\left\{ (X_2^2 - X_2) I_{\frac{e^{-\lambda}}{X_1 X_3 \ldots X_p} < X_2 < \frac{e^{\lambda}}{X_1 X_3 \ldots X_p}} \right\} X_1 X_3^2 \ldots X_p^2$$

$$+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$+ E\left\{ (X_p^2 - X_p) I_{e^{-\lambda}/X_1 X_2 \ldots X_{p-1} < X_p < e^{\lambda}/X_1 X_2 \ldots X_{p-1}} \right\} X_1 X_2 \ldots X_{p-1} \cdot$$

and the same technique goes through, and infact as in (3.17), we shall be able to exactly calculate the individual terms.

A remark on Theorem 6.3.2 The problem of estimating $|\Sigma|$ on the basis of the loss $(a-|\Sigma|)^2/|\Sigma|^2$ remains invariant under the group of transformations $S \longrightarrow ASA'$, where $|A| = \pm 1$. $|S|$ is a maximal invariant on the sample space with respect to this group of transformations. This is a partial justification for estimating $|\Sigma|$ on the basis of $|S|$, although this group is not compact, and therefore this does not settle the main problem.

REFERENCES

AMEMIYA, T. :

The n^{-2}-order mean squared errors of the maximum likelihood and the minimum chi-square estimator. Ann. Statist. 8, 488-505 (1980).

BERGER, JAMES O.

Inadmissibility results for generalized Bayes estimators of coordinates of a location vector. Ann. Statist. 4, 302-333 (1976a).

Admissibility results for generalized Bayes estimators of coordinates of a location vector. Ann. Statist. 4, 334-356 (1976b).

Inadmissibility result for the best invariant estimator of two coordinates of a location vector. Ann. Statist. 4, 1065-1076 (1976c).

Improving on inadmissible estimators in continuous exponential families with applications to simultaneous estimation of Gamma scale parameters. Ann. Statist. 8, 545-571 (1980).

Statistical Decision Theory : Foundations, concepts and methods. Springer-Verlag (1980).

BERKSON, J. :

Maximum likelihood and minimum chi-square estimates of the logistic function Jour. Amer. Statist. Assoc. 50, 130-162 (1955).

BICKEL, P.J. :

Minimax estimation of the mean of a normal distribution when the parameter space is restricted. Ann. Statist. 9, 1301-1309 (1981).

BICKEL, P.J. and COLLINS, J.R.

Minimizing Fisher information over mixtures of distributions (1982) (To appear)

BLYTH, COLIN R. .

On minimax statistical decision procedures and their admissibility. Ann. Math. Statist. 22, 22-42 (1951)

BROWN, L.D. :

On the admissibility of invariant estimators of one or more location parameters. Ann. Math. Statist. 37, 1087-1136 (1966).

Admissible estimators, recurrent diffusions and insoluble boundary value problems. Ann. Math. Statist. 42, 855-903 (1971).

A heuristic method for determining admissibility of estimators with applications. Ann. Statist. 7, 960-994 (1979).

A necessary condition for admissibility. Ann. Statist. 8, 540-544 (1980a).

Example of Berger's phenomenon in the estimation of independent normal means. Ann. Statist. 8, 572-585 (1980b).

The differential inequality of a statistical estimation problem. Pre-print (1981).

BROWN, L.D. and FOX, M. :

Admissibility of procedures in two-dimensional location parameter problems. Ann. Statist. 2, 248-266 (1974).

BROWN, L.D. and HWANG, J.T. :

Unified admissibility proof, Pre-print (1981).

CASELLA, GEORGE and STRAWDERMAN, WILLIAM E.

Estimating a bounded normal mean. Ann. Statist. 9, 870-878 (1981).

CLEVENSON, M.L. and ZIDEK, J.V. :

Simultaneous estimation of the means of independent poisson laws. Jour. Amer. Statist. Assoc. 70, 698-705 (1975).

DAS GUPTA, ANIRBAN :

Bayes Minimax estimation in multiparameter families when the parameter space is restricted to a bounded convex set (1982a) (Submitted for publication)

Inadmissibility results in the Gamma distribution : Two examples (1982b) (Submitted for publication).

Is the best equivariant estimator of the generalized variance admissible ? Technical Report, Stat-Math. Division, Indian Statistical Institute (1982c).

DAS GUPTA, ANIRBAN and GHOSH, J.K. :

Some remarks on second-order admissibility in the multiparameter case (1982) (To appear).

DAS GUPTA, ANIRBAN and SINHA, BIMAL KUMAR :

On the admissibility of polynomial estimators in the one-parameter exponential family. Sankhyā : Ser. B, 42, 129-142 (1980).

Admissibility for estimation in non-regular densities. Technical Report, Stat-Math. Division, Indian Statistical Institute (1979).

FARRELL, R.H. :

Estimators of a location parameter in the absolutely continuous case. Ann. Math. Statist. 35, 949-998 (1964).

A necessary and sufficient condition for admissibility when strictly convex loss is used. Ann. Math. Statist. 39, 24-29 (1968).

FERGUSON, T.S. :

Methmatical Statistics : A Decision-theoretic approach. Academic Press (1967).

GHOSH, J.K. and SINHA, BIMAL K. :

A necessary and sufficient condition for second order admissibility with applications to Berkson's problem. Ann. Statist. 11, 1334-1338 (1981).

GHOSH, J.K., SINHA, B.K., and JOSHI, S.N. :

Expansion of posterior probability and integrated Bayes risks. (1981) (To appear).

GHOSH, J.K. and SUBRAMANYAM, K.

Second order efficiency of maximum likelihood estimators. Sankhyā : Ser. A, 36, 325-358 (1974).

GHOSH, M. and MEEDEN, G. :

Admissibility of linear estimators in the one-parameter exponential family. Ann. Statist. 5, 772-778 (1977).

GHOSH, MALAY and PARSIAN, AHMAD .

Admissible and minimax multiparameter estimation in exponential families.Jour. Mult. Analysis. 10, 551-564 (1980).

GHOSH, M.N. :

Uniform approximation of minimax point estimates. Ann. Math. Statist. 35, 1031-1047 (1964).

HAFF, L.R. .

An identity for the Wishart distribution with applications. Jour. Mult. Analysis. 9, 531-544 (1979).

HUDSON, H.M. :

A natural identity for exponential families with applications in multiparameter estimation. Ann. Statist. 6, 473-484 (1978).

HWANG, J.T. :

Certain bounds on the class of admissible estimators in continuous exponential families (1981) (To appear).

JAMES, W. and STEIN, C. :

Estimation with quadratic loss. Proc. Fourth Berkeley Symp. Math. Statist. Prob. 1, 361-379 (1960).

JOHNSTONE, I. :

A converse to Karlin's theorem on linear estimates (1981) (To appear).

JOSHI, V.M. :

On a theorem of Karlin regarding admissible estimates for exponential populations. Ann. Math. Statist. 40, 216-223 (1969)

KARIYA, TAKEKAI, SINHA, BIMAL K. and SUBRAMANYAM, K. :

Berkson's bioassay problem-revisited. (1981) (To appear).

KARLIN, S. :

Admissibility for estimation with quadratic loss. Ann. Math. Statist. 29, 406-436 (1958).

KIEFER, J. :

Invariance, minimax sequential estimation, and continuous time processes, Ann. Math. Statist. 28, 573-601 (1957).

LEVIT, B.Ya. :

On the theory of the asymptotic minimax property of second order. Theory Prob. Appl. 24, 435-437 (1979).

On asymptotic minimax estimates of the second order. Theory Prob. Appl. 25, 552-568 (1980).

MORTON, and RAGHAVACHARI, :

On a theorem of Karlin regarding admissibility of linear estimates in exponential populations. Ann. Math. Statist. 37, 1809-1813 (1966).

PENG, J.C. :

Simultaneous estimation of the parameters of independent Poisson distributions.- Technical Report, Department of Statistics, Stanford University (1975).

PING, C. :

Minimax estimates of parameters of distributions belonging to the exponential family. Chinese Math. Acta. 5, 277-299 (1964).

SELLIAH, J.B.

Estimation and testing problems in a Wishart distribution.- Technical Report, Department of Statistics, Stanford University (1964).

SHORROCK, R.W. and ZIDEK, J.V. :

An improved estimator of the generalized variance. Ann. Statist. 4, 627-638 (1976).

SINHA, BIMAL K., GHOSH, M. and BANERJEE, P.K. :

An admissibility result and its applications. Pre-print (1978).

SRINIVASAN, C. :

Admissible generalized Bayes estimators and exterior boundary value problems. Sankhyā, Ser. A 43, 1 - 25 (1981).

Admissibility in the control problem. Pre-print (1982).

STEIN, C.

A necessary and sufficient condition for admissibility. Ann. Math. Statist. 26, 518-522 (1955).

STEIN, C. :

Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. Proc. Third Berkeley Symp. Math. Statist. Prob. 1, 197-206 (1955).

The admissibility of Pitman's estimators of a single location parameter. Ann. Math. Statist. 30, 970-979 (1959).

Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. Ann. Inst. Statist. Math. 16, 155-160 (1964).

Estimation of the mean of a multivariate normal distribution. Proc. Prague Symp. Asymptotic Statist. 345-381 (1973).

Estimation of the mean of a multivariate normal distribution. Ann. Statist. 9, 1135-1151 (1981).

STRAWDERMAN, W.E. :

Proper Bayes minimax estimators for the mean of a multivariate normal population. Ann. Math. Statist. 42, 385-388 (1971).

On the existence of proper Bayes minimax estimators of the mean of a multivariate normal distribution. Proc. Sixth Berkeley Symp. 1, 51-55 (1972).

WALD, A. :

Statistical Decision Functions. John Wiley (1950).

ZIDEK, J.V. :

Sufficient conditions for the admissibility under squared-error loss of formal Bayes estimators. Ann. Math. Statist. 41, 446-456 (1970).