

ADMISSIBILITY OF LINEAR ESTIMATORS IN THE ONE PARAMETER EXPONENTIAL FAMILY

BY MALAY GHOSH AND GLEN MEEDEN

Iowa State University

For estimating the mean in the one parameter exponential family with quadratic loss, Karlin (1958) gave sufficient conditions for admissibility of estimators of the form aX . Later, Ping (1964) and Gupta (1966) gave sufficient conditions for admissibility of estimators of the form $aX + b$ for the same problem. Zidek (1970) gave sufficient conditions for the admissibility of X for estimating an arbitrary piecewise continuous function of the parameter, say $\gamma(\theta)$, not necessarily the mean. In this paper it is shown that Karlin's argument yields sufficient conditions for the admissibility of estimators of the form $aX + b$ for estimating $\gamma(\theta)$. The results are then extended to the case when the parameter space is truncated.

1. Introduction. Let X be a real valued random variable with probability density function $p_\theta(x) = \beta(\theta) \exp(\theta x)$ with respect to some σ -finite measure μ . Also, $\theta \in \Theta = \{\theta: \int \exp(\theta x) d\mu(x) < \infty\}$. From the convexity of the exponential function, Θ is an interval. The upper and lower end points of Θ are denoted respectively by $\bar{\theta}$ and $\underline{\theta}$, which may or may not belong to Θ . The problem is the estimation of $\gamma(\theta)$, some specified piecewise continuous function of θ with squared error loss. Conditions are stated under which linear estimators of the form $aX + b$ are admissible for estimating $\gamma(\theta)$.

When $\gamma(\theta) = E_\theta(X) = -\beta'(\theta)/\beta(\theta)$, sufficient conditions for admissibility of estimators of the form aX were given by Karlin (1958). Implicit in Karlin's argument is the fact that the estimator is essentially generalized Bayes with respect to some (possibly improper) prior distribution; sufficient conditions on the tail behavior of the prior density guarantee the admissibility of estimators. A different proof using the Cramér-Rao inequality was given by Ping (1964) under the same sufficient conditions.

Zidek (1970) provides sufficient conditions for admissibility of X in estimating $\gamma(\theta)$ using Stein's (1965) technique of approximating improper priors by proper ones. Unlike Karlin, Zidek needs an explicit condition to guarantee a formal Bayes representation of the estimator, that is, its representation as the ratio of two integrals. This extra condition is not necessary. Examples 1 and 2 in Section 3 illustrate that an estimator may be admissible, even though Zidek's conditions are not met.

The main theorem of the paper is proved in Section 2 showing that Karlin's

Received June 1975; revised November 1976.

AMS 1970 subject classifications. 62C15, 62F10.

Key words and phrases. Admissibility, one parameter exponential family, linear estimators, squared error loss, generalized Bayes estimators, Cramér-Rao inequality, truncated parameter space.

technique can be used to prove the admissibility of linear estimators for estimating $\gamma(\theta)$, not necessarily the mean. Several applications of the theorem are considered in Section 3. Finally, in Section 4, the argument is modified to include cases where the parameter space is truncated, thereby getting a generalized version of a theorem of Katz (1961).

2. Admissibility of linear estimators. First, a heuristic argument is given leading to a (possible improper) distribution with respect to which $aX + b$ is generalized Bayes for estimating $\gamma(\theta)$ (if certain integrals exist), under squared error loss. The prior distribution is assumed to be absolutely continuous with respect to Lebesgue measure with a Radon-Nikodym derivative, say $\Pi(\theta)$. Then with the notations $a = (\lambda + 1)^{-1}$, $b = \alpha(\lambda + 1)^{-1}$, ($\lambda \neq -1$), $h_1(\theta) = \gamma(\theta)\beta(\theta)\Pi(\theta)$, $h_2(\theta) = \beta(\theta)\Pi(\theta)$, one has the formal representation

$$(2.1) \quad (x + \alpha)(\lambda + 1)^{-1} = \int \exp(\theta x)h_1(\theta) d\theta / \int \exp(\theta x)h_2(\theta) d\theta .$$

Rearrangement of terms and integration by parts now lead to

$$(2.2) \quad (\lambda + 1) \int \exp(\theta x)h_1(\theta) d\theta = - \int \exp(\theta x)h_2'(\theta) d\theta + \alpha \int \exp(\theta x)h_2(\theta) d\theta .$$

The uniqueness property of Laplace transforms now gives $(\lambda + 1)h_1(\theta) = -h_2'(\theta) + \alpha h_2(\theta)$, or $h_2'(\theta)/h_2(\theta) = \alpha - (\lambda + 1)\gamma(\theta)$. Integrating, it follows that

$$(2.3) \quad \Pi(\theta) = \beta^{-1}(\theta) \exp[\alpha\theta - (\lambda + 1) \int_a^\theta \gamma(t) dt] ,$$

where d is an interior point of Θ .

The following theorem shows that the tail behavior of $\Pi(\theta)$ essentially guarantees admissibility of $(X + \alpha)/(\lambda + 1)$.

THEOREM. Let $p_\theta(x) = \beta(\theta) \exp(\theta x)$ be the density of the exponential family wrt the σ -finite measure μ ; $\theta \in \Theta = \{\theta: \beta^{-1}(\theta) = \int \exp(\theta x) d\mu(x) < \infty\}$. Let $\gamma(\theta)$ be some specified piecewise continuous function to be estimated with squared error loss. It is also assumed that $\int_a^b \gamma(\theta) d\theta$ exists for all $[a, b]$ in Θ .

If $\Pi(\theta)$ defined in (2.3) satisfies

$$(2.4) \quad \int_u^{\bar{\theta}} \Pi^{-1}(\theta) d\theta \rightarrow \infty \text{ as } u \rightarrow \bar{\theta} , \quad \int_u^c \Pi^{-1}(\theta) d\theta \rightarrow \infty \text{ as } u \rightarrow \underline{\theta} ,$$

where $\bar{\theta}$ and $\underline{\theta}$ are the upper and lower end points of Θ , and c is an interior point of Θ , then $(X + \alpha)/(\lambda + 1)$ ($\lambda \neq -1$) is admissible for estimating $\gamma(\theta)$.

PROOF. Suppose $(X + \alpha)/(\lambda + 1)$ is not admissible. Then there exists some $\delta(X)$ such that

$$(2.5) \quad \begin{aligned} & \int_{-\infty}^{\infty} [\delta(x) - \gamma(\theta)]^2 \beta(\theta) \exp(\theta x) d\mu(x) \\ & \leq \int_{-\infty}^{\infty} [(x + \alpha)/(\lambda + 1) - \gamma(\theta)]^2 \beta(\theta) \exp(\theta x) d\mu(x) , \\ \text{or} \\ & \int_{-\infty}^{\infty} [\delta(x) - (x + \alpha)/(\lambda + 1)]^2 \beta(\theta) \exp(\theta x) d\mu(x) \\ & \leq 2 \int_{-\infty}^{\infty} [(x + \alpha)/(\lambda + 1) - \delta(x)] \\ & \quad \times [(x + \alpha)/(\lambda + 1) - \gamma(\theta)] \beta(\theta) \exp(\theta x) d\mu(x) , \end{aligned}$$

for all $\theta \in \Theta$, the inequality being strict for some $\theta \in \Theta$. Next, it is shown that $\delta(x) = (x + \alpha)/(\lambda + 1)$ a.e. (μ), and the theorem will follow by contradiction.

To this end, first multiply both sides of (2.5) by $\Pi(\theta)$ (defined in (2.3)), and then integrate wrt θ over (a, b) , (a, b) being some proper subset of Θ . Then, on reversing the order of integration on the RHS, one gets

$$(2.6) \quad \int_a^b [\int_{-\infty}^{\infty} [\delta(x) - (x + \alpha)/(\lambda + 1)]^2 \beta(\theta) \exp(\theta x) d\mu(x)] \Pi(\theta) d\theta \\ \leq 2 \int_{-\infty}^{\infty} [(x + \alpha)/(\lambda + 1) - \delta(x)] \\ \times \left\{ \int_a^b ((x + \alpha)/(\lambda + 1) - \gamma(\theta)) \exp[(x + \alpha)\theta - (\lambda + 1) \\ \times \int_a^b \gamma(t) dt] d\theta \right\} d\mu(x).$$

Now,

$$(2.7) \quad \int_a^b [(x + \alpha)/(\lambda + 1) - \gamma(\theta)] \exp[(x + \alpha)\theta - (\lambda + 1) \int_a^b \gamma(t) dt] d\theta \\ = (\lambda + 1)^{-1} [\exp\{(x + \alpha)b - (\lambda + 1) \int_a^b \gamma(t) dt\} \\ - \exp\{(x + \alpha)a - (\lambda + 1) \int_a^b \gamma(t) dt\}].$$

Using (2.7), the notation $T(\theta) = \int_{-\infty}^{\infty} [\delta(x) - (x + \alpha)/(\lambda + 1)]^2 \beta(\theta) \exp(\theta x) d\mu(x)$, and Schwarz's inequality, one finds that

$$(2.8) \quad \int_a^b T(\theta) \Pi(\theta) d\theta \\ \leq 2|\lambda + 1|^{-1} [\Pi(b) \int_{-\infty}^{\infty} |\delta(x) - (x + \alpha)/(\lambda + 1)| \exp(bx) \beta(b) d\mu(x) \\ + \Pi(a) \int_{-\infty}^{\infty} |\delta(x) - (x + \alpha)/(\lambda + 1)| \exp(ax) \beta(a) d\mu(x)] \\ \leq 2|\lambda + 1|^{-1} [\Pi(b) T^{\frac{1}{2}}(b) + \Pi(a) T^{\frac{1}{2}}(a)].$$

It suffices to show that $T(\theta) = 0$, for all θ , since then $\delta(x) = (x + \alpha)/(\lambda + 1)$ a.e. (μ). This is accomplished by considering the following two cases:

I. $\liminf_{b \rightarrow \bar{\theta}} \Pi(b) T^{\frac{1}{2}}(b) = \Delta (> 0)$.

Fix a , and let $H(b) = \int_a^b \Pi(\theta) T(\theta) d\theta$. Then (2.8) leads to $H(b) \leq K [H'(b) \Pi(b)]^{\frac{1}{2}}$, where, in the above and in what follows, K is a generic constant. Then, choosing $b_1 < b_2$ and $H(b_1) > 0$, one gets

$$\int_{b_1}^{b_2} [H'(b)/H^2(b)] db \geq K \int_{b_1}^{b_2} \Pi^{-1}(b) db$$

or,

$$H^{-1}(b_1) - H^{-1}(b_2) \geq K \int_{b_1}^{b_2} \Pi^{-1}(b) db \rightarrow \infty \quad \text{as } b_2 \rightarrow \bar{\theta} \text{ from (2.4).}$$

But $H(b_1) > 0$, so that the left-hand side remains bounded, a contradiction.

II. $\liminf_{b \rightarrow \bar{\theta}} \Pi(b) T^{\frac{1}{2}}(b) = 0$.

Let $G(a) = \int_a^{\bar{\theta}} \Pi(\theta) T(\theta) d\theta$. Then, $G(a) \leq K \Pi^{\frac{1}{2}}(a) (-G'(a))^{\frac{1}{2}}$. Hence, if a_0 is such that $G(a_0) > 0$, then, taking $a_1 < a_0$,

$$\int_{a_1}^{a_0} (-G'(a)/G^2(a)) da \geq K \int_{a_1}^{a_0} \Pi^{-1}(a) da \rightarrow \infty \quad \text{as } a_1 \rightarrow \underline{\theta}.$$

But LHS = $G^{-1}(a_0) - G^{-1}(a_1)$ which remains bounded. Hence $G(a) = 0$ for all a which implies $T(\theta) = 0$ for all $\theta \in \Theta$. Hence, the theorem.

The above theorem includes Ping's (1964) theorem as a special case when $\gamma(\theta) = E_{\delta}(X) = -\beta'(\theta)/\beta(\theta)$. Note that in this case, choosing d such that $\beta(d) = 1$, $\Pi(\theta) = \exp(\alpha\theta - (\lambda + 1) \int_a^b (-\beta'(t)/\beta(t) dt) \beta^{-1}(\theta) = \beta^{\lambda}(\theta) \exp(\alpha\theta)$, and condition (2.4) reduces to Ping's condition.

3. Examples. In all the examples considered, the loss is squared error.

EXAMPLE 1. Suppose $X \sim \text{Binomial}(n, \exp(\theta)/(1 + \exp(\theta)))$, $\theta \in (-\infty, \infty)$. (Note that the distribution is suitably reparametrized to write the pdf in the form given in this paper.) The parameter of interest is $\gamma(\theta) = \exp(2\theta)/(1 + \exp(\theta))^2$, and the estimator is X/n . Then, from (2.3),

$$(3.1) \quad \begin{aligned} \Pi^{-1}(\theta) &= (1 + \exp(\theta))^{-n} \exp\{n \int_a^\theta [\exp(2t)/(1 + \exp(t))^2] dt\} \\ &= C(d) \exp(-n \exp(\theta)/(1 + \exp(\theta))), \end{aligned}$$

where $C(d)$ is a positive constant depending on d . Since $\Pi^{-1}(\theta) \rightarrow C(d)$ or $C(d)\exp(-n)$ as $\theta \rightarrow -\infty$ or ∞ , (2.4) is satisfied, so that X/n is an admissible estimator of $\gamma(\theta)$.

Note however that in this case, Zidek's (1970) condition regarding the representation of the estimator as the ratio of two integrals is not met. To guarantee admissibility, Zidek needs, in addition to (2.4),

$$(3.2) \quad \exp(x\theta - n \int_a^\theta \gamma(t) dt) \rightarrow 0,$$

as $\theta \rightarrow \pm\infty$ for all $x = 0, 1, \dots, n$. But the expression involved in (3.2) is $\exp(\theta x)(1 + \exp(\theta))^{-n}\Pi(\theta)$ which at $x = 0$ tends to $C^{-1}(d)$ as $\theta \rightarrow -\infty$, and at $x = n$ tends to $C^{-1}(d) \exp(n)$ as $\theta \rightarrow \infty$ so that (3.2) is not satisfied. This point is illustrated with another example.

EXAMPLE 2. Suppose $X \sim \text{Poisson}(\exp(\theta))$, where $\theta \in (-\infty, \infty)$. The problem is estimation of $\gamma(\theta) = \exp(g\theta)$ for $g > 1$, while the estimator is X . In this case, from (2.3),

$$(3.3) \quad \Pi^{-1}(\theta) = C(d) \exp[-\exp(\theta) + g^{-1} \exp(g\theta)],$$

where $C(d)$ is once again a positive constant depending on d . In this case, $\Pi^{-1}(\theta) \rightarrow \infty$ or $C(d)$ according as $\theta \rightarrow \infty$ or $-\infty$. Thus, (2.4) is satisfied, and so X is admissible for estimating $\gamma(\theta)$. However, the expression involved in (3.2) is $\exp(\theta x - \exp(\theta))\Pi(\theta)$ which at $x = 0$ tends to $C^{-1}(d)$ as $\theta \rightarrow -\infty$, so that (3.2) is not satisfied.

Three more examples are presented to illustrate the theorem.

EXAMPLE 3. Suppose X has a general power series distribution (see Roy and Mitra (1957)) with pdf of the form

$$(3.4) \quad f_\theta(x) = a_x \exp(\theta x)\beta(\theta), \quad x = 0, 1, \dots, \quad \theta \in (-\infty, \infty),$$

where $\beta^{-1}(\theta) = \sum_{x=0}^\infty a_x \exp(\theta x)$ (assumed to be finite). Roy and Mitra (1957) have considered the problem of minimum variance unbiased estimation of $\exp(g\theta)$ for all positive integers g . For $g > 1$ typically such estimators are non-linear estimators. In considering admissibility of estimators of the form $(X + \alpha)/(\lambda + 1)$, as in the case of Example 2, one can readily compute from (2.3),

$$\Pi^{-1}(\theta) = C(d)\beta(\theta) \exp[-\alpha\theta + (\lambda + 1)g^{-1} \exp(g\theta)],$$

and a sufficient condition to guarantee admissibility is (2.4) with $\bar{\theta} = \infty$, $\underline{\theta} = -\infty$. On returning to the important special case of a Poisson distribution one finds that for $g > 1$, any estimator $(X + \alpha)/(\lambda + 1)$ with $\lambda > -1$, $\alpha \geq 0$ is admissible for estimating $\exp(g\theta)$, while any estimator $(X + \alpha)/(\lambda + 1)$ with $\lambda > 0$, $\alpha > 0$ or $\lambda = \alpha = 0$ is admissible for estimating $\exp(\theta)$. The latter conforms with Gupta's (1966) result concerning the estimation of $E_\theta(X) = \exp(\theta)$.

EXAMPLE 4. Suppose $X \sim N(\theta, 1)$ where θ is real. Then $\beta(\theta) = \exp(-\frac{1}{2}\theta^2)$, $\gamma(\theta) = \theta^3$. Taking $d = 0$ (without loss of generality), $\Pi(\theta) = \exp(\alpha\theta + \frac{1}{2}\theta^2 - \frac{1}{4}(\lambda + 1)\theta^4)$. So for any $\lambda > -1$, $\int_u^\infty \Pi^{-1}(\theta) d\theta \rightarrow \infty$ as $u \rightarrow +\infty$ and $\int_u^c \Pi^{-1}(\theta) d\theta \rightarrow \infty$ as $u \rightarrow -\infty$. Thus any estimator $(X + \alpha)/(\lambda + 1)$, α real, $\lambda > -1$ is admissible for estimating θ^3 .

The conditions are not, however, met if $\lambda < -1$. It is easy to see that in this case the estimator $(X + \alpha)/(\lambda + 1)$ goes in the opposite direction as X , and we can expect the former to be inadmissible for estimating θ^3 . In fact, it is dominated by $-(X + \alpha)/(\lambda + 1)$.

It can be easily seen in the normal case that the above conclusion holds when $\gamma(\theta)$ is any odd-degree polynomial in θ , the coefficient of the highest power of θ being positive. The conclusion is not, however, true for an even degree polynomial. To see a simple example, let $\gamma(\theta) = \theta^2$. Then, $\Pi(\theta) = \exp(\alpha\theta + \frac{1}{2}\theta^2 - \frac{1}{3}(\lambda + 1)\theta^3)$. Then for $\lambda > -1$, the condition $\int_u^\infty \Pi^{-1}(\theta) d\theta \rightarrow \infty$ fails as $u \rightarrow -\infty$, while for $\lambda < -1$, the condition $\int_u^c \Pi^{-1}(\theta) d\theta \rightarrow \infty$ fails as $u \rightarrow \infty$.

EXAMPLE 5. Again suppose $X \sim N(\theta, 1)$, $\beta(\theta) = \exp(-\frac{1}{2}\theta^2)$. We are interested in finding sufficient conditions on γ guaranteeing X as an admissible proper Bayes estimator. Thus, we require conditions under which Π is a proper distribution or equivalently $g(\theta) = \exp[-\int_0^\theta (\gamma(t) - t) dt]$ is integrable. It is easily seen that if for some constant $M > 0$, and for some $\varepsilon > 0$, $\gamma(t) \geq t + (1 + \varepsilon)t^{-1}$ for $t > M$, $\gamma(t) \leq t + (1 + \varepsilon)t^{-1}$ for $t < -M$ and γ is bounded on $[-M, M]$, then $g(\theta)$ is a proper prior. On the other hand a sufficient condition for X to be an admissible estimator of γ is that for some constant $M > 0$, γ is bounded on $[-M, M]$, $\gamma(t) \geq t - t^{-1}$ for $t > M$, $\gamma(t) \leq t - t^{-1}$ for $t < -M$. As a special case, one gets the admissibility of X in estimating $\gamma(\theta) = \theta(1 + K/(1 + \theta^2))$, a fact mentioned in Blyth (1974).

From the above theorem and examples, one might end up with the misleading conclusion that for the one parameter exponential family, any generalized Bayes estimator with respect to some improper prior satisfying (2.4) type conditions is admissible. This is however, not true. To see why this is so, note that a major step in the proof of Theorem 1 is the reduction (2.7). In general if $\delta^*(X)$ is a generalized Bayes estimator with respect to a prior $\Pi(\theta)$ which satisfies (2.4), and $p_\theta(x)$ is a density of X (exponential or not) with respect to some σ -finite measure μ , then defining

$$(3.5) \quad M(x; a, b) = \int_a^b [\delta^*(x) - \gamma(\theta)] p_\theta(x) \Pi(\theta) d\theta / [p_b(x) \Pi(b) + p_a(x) \Pi(a)],$$

the condition $E_b M^2(X; a, b) \leq K$, $E_a M^2(X; a, b) \leq K$ uniformly in a and b guarantees the admissibility of $\delta^*(X)$ for estimating $\gamma(\theta)$ when the loss is squared error. The proof repeats the arguments of Theorem 1. This fact was also noticed by Katz (1961), who, however, set the stronger condition $|M(x; a, b)| \leq K$, uniformly in a and b .

It is also possible to prove Theorem 2.1 using a Cramér-Rao type inequality (see Blyth (1974), (8), page 469). This was essentially used by Ping (1964) in proving admissibility of estimators of the form $aX + b$ in estimating $E_\theta(X) = -\beta'(\theta)/\beta(\theta)$. It is also clear from there that the Cramér-Rao technique presupposes the knowledge of $\Pi(\theta)$. However, once $\Pi(\theta)$ is known, Karlin's (1958) method seems to be much more intuitive than the Cramér-Rao method in as much as the former imitates the admissibility proof of Bayes estimators.

4. Admissibility estimators when the parameter space is truncated. Assume as before that X has an exponential density $p_\theta(x)$ with respect to some σ -finite measure μ , but now $\theta \in \Theta_0 = \{\theta | \theta \geq a\} \subset \Theta = (-\infty, \infty)$, a being an interior point of Θ . For simplicity, take $a = 0$. For estimating $\gamma(\theta)$, using the same prior $\Pi(\theta)$ as before, one is led to the generalized Bayes estimator $\delta(X)$, where $\delta(x) = (x + \alpha)/(\lambda + 1) + g(x)$, $g(x) = (\lambda + 1)^{-1} \exp(-(\lambda + 1) \int_a^0 \gamma(t) dt) / \int_0^\infty \exp\{(x\theta + \alpha\theta) - (\lambda + 1) \int_a^\theta \gamma(t) dt\} d\theta$ (assuming the appropriate integrals to exist) being the "correction factor" due to the truncation of the parameter space. Once again, an explicit condition similar as (3.2) is not needed. An argument similar to the proof of Theorem 1 establishes the admissibility of $\delta(X)$ for estimating $\gamma(\theta)$. Also this extends a result of Katz (1961) who proved the admissibility of δ in estimating $\gamma(\theta) = E_\theta X$ using both Karlin's (1958) and Blyth's (1951) techniques. Note that in the present case, the condition (2.4) can be simplified to $\int_a^b \Pi^{-1}(\theta) d\theta \rightarrow \infty$ as $b \rightarrow \infty$, since the parameter space is truncated.

Acknowledgment. Thanks are due to the Associate Editor and the referees for their constructive criticisms leading to substantial improvement on an earlier version of the paper.

REFERENCES

- [1] BLYTH, C. R. (1951). On minimax statistical decision procedures and their admissibility. *Ann. Math. Statist.* **22** 22-42.
- [2] BLYTH, C. R. (1974). Necessary and sufficient conditions for inequalities of Cramér-Rao type. *Ann. Statist.* **2** 464-473.
- [3] GUPTA, M. K. (1966). On the admissibility of linear estimates for estimating the mean of distributions of the one parameter exponential family. *Calcutta Statist. Assoc. Bull.* **20** 14-19.
- [4] KARLIN, S. (1958). Admissibility for estimation with quadratic loss. *Ann. Math. Statist.* **29** 406-436.
- [5] KATZ, M. W. (1961). Admissible and minimax estimates of parameters in truncated spaces. *Ann. Math. Statist.* **32** 136-142.
- [6] PING, C. (1964). Minimax estimates of parameters of distributions belonging to the exponential family. *Chinese Math.-Acta* **5** 277-299.

- [7] ROY, J. and MITRA, S. K. (1957). Unbiased minimum variance estimation in a class of discrete distributions. *Sankhyā* **18** 371-378.
- [8] STEIN, C. (1965). Approximation of improper prior measures by prior probability measures. *Bernoulli, Bayes, Laplace Anniversary Volume*. Springer-Verlag, New York.
- [9] ZIDEK, J. V. (1970). Sufficient conditions for the admissibility under squared error loss of formal Bayes estimators. *Ann. Math. Statist.* **41** 446-456.

DEPARTMENT OF STATISTICS
SNEDECOR HALL
IOWA STATE UNIVERSITY
AMES, IOWA 50010