The Annals of Statistics
1979, Vol. 7, No. 5, 960–994

A HEURISTIC METHOD FOR DETERMINING ADMISSIBILITY OF ESTIMATORS—WITH APPLICATIONS[1]

BY LAWRENCE D. BROWN

Rutgers University

Questions of admissibility of statistical estimators are reduced to considerations involving differential inequalities. The coefficients of these inequalities involve moments of the underlying distributions; and so are, in principle, not difficult to derive.

The methods are "heuristic" because it is necessary to verify on an ad-hoc basis that error terms are small. Some conditions on the structure of the problem are given which we believe will guarantee that these error terms are small.

Several different statistical estimation problems are discussed. Each problem is transformed (if necessary) so as to meet the above mentioned structure conditions. Then the heuristic method is applied in order to generate conjectures concerning the admissibility of certain generalized Bayes procedures in these problems.

Introduction. Section 1 of this study describes a very general heuristic method of discovering admissible (and reasonable) estimators in a variety of statistical problems. It is our belief that this method—or minor modifications of it—can also be used to check the admissibility of various commonly proposed procedures.

The method is applied in Section 2 to generate conjectures concerning the admissibility of certain generalized Bayes procedures in some common problems. The applications discussed include: estimation of several location parameters, estimation of several Poisson means, estimation of the largest of several ordered translation parameters, and estimation of a normal variance when the mean is unknown. Few of the results described in the applications are rigorously proven. Nonetheless we feel that they are all plausible for the reasons given in this study.

The question of admissibility of estimators of location parameters has already received considerable attention in the literature, and many results have been obtained. Various references are mentioned in the course of this paper—see especially Section 2.2. Some progress has also been made in a few other situations. Most of these are also discussed in Section 2.

Two interesting general features stand out in results so far obtained. One is that the admissibility of a generalized Bayes estimator seems to depend on the general

Received June 1974; revised October 1977.

[1]This research was supported in part by NSF research grant number MCS75-23343-A01. This paper is a revised version of the author's invited address to the Institute of Mathematical Statistics meeting in May 1973. The material concerning "Method II" for proving inadmissibility was added in revisions prepared in August 1977.

AMS 1970 *subject classifications.* Primary 62C15, 62F10; secondary 65M99, 65P05, 62H99, 62C10.

Key words and phrases. Estimation, admissibility, (nonlinear) differential inequalities, generalized Bayes estimators, estimating location parameters, estimating Poisson means, estimating the largest mean, estimating a normal variance (mean unknown).

structure of the problem (the general type of problem-location, scale, etc.—and the dimension of the parameter space) and on the generalized Bayes prior, but not on other more specific features of the problem such as the loss function used or the exact shape of the densities (normal, exponential, etc.). A similar comment holds for best invariant estimators in problems where they exist. A more striking feature is that certain estimators which have in the past seemed intuitively reasonable have turned out to be inadmissible.

This study gives an explanation of both these features of admissibility problems. A few preliminary comments may be useful.

The admissibility problem is related to a differential inequality. The coefficients of this differential inequality are very much dependent on the general type of problem involved, and on the generalized prior used. On the other hand, in many problems the important qualitative relations among the coefficients are not affected by the other more specific features of the problem. For this reason only the general features rather than the specific features usually determine the admissibility of generalized Bayes procedures for given generalized priors. Certain inadmissibility results, which appear very surprising when the statistical problem is viewed intuitively, appear natural when one instead looks at the form of the related differential inequality. A striking example of this occurs in the problem of several Poisson means, which is treated in Section 2.3.

1. Setting and methods.

1.1. *Setting*. Observe random variables having a probability density $p_{\xi,\eta}(x,y)$ with respect to Lebesgue measure (dx, dy) on $R^k \times R^l$. The parameters, ξ, η, are unknown with $\xi, \eta \in \Omega \subset R^k \times R^l$, $k \geq 1$, $l \geq 0$. The case $l = 0$ is to be interpreted that only X is observed and no variable Y or parameter η is present in the problem. The assumption of the existence of a density for the variables X, Y is for convenience and notational uniformity only.

It is desired to estimate ξ only, with η to be considered as a nuisance parameter. Hence we consider (nonrandomized) estimation rules of the form $\delta : R^k \times R^l \to R^k$. That is to say, $\delta(x,y) = d \in R^k$.

A loss function, $L = L((\xi, \eta), d) \geq 0$, is specified. The possible dependence of L on η will be considerably restricted by the regularity conditions to be described below.

As usual, a rule δ is *inadmissible* if there is a rule δ' such that $\Delta_{\delta,\delta'}(\xi, \eta) = E_{\xi,\eta}(L((\xi, \eta), \delta(x,y))) - E_{\xi,\eta}(L((\xi, \eta), \delta'(x,y))) \geq 0$ but not $\equiv 0$, and otherwise δ is *admissible*.

For simplicity the following derivations are given only for the case where $k = l = 1$ and $L((\xi, \eta), d) = W(d - \xi)$. Results for other situations can be similarly derived. Some special cases are presented in the examples in Section 2. For convenience, we will always assume that W has at least three continuous derivatives.

Special case. An important specialization is provided by the case of *squared error*, where $w(d - \xi) = (d - \xi)^2$. (When $k > 1$ *squared error* is defined by $L((\xi, \eta), d) = W(d - \xi) = \|d - \xi\|^2$.)

Regularity assumptions. The above setting is extremely general. In fact the methods to follow do not apply in such generality. We are not able to construct a precise, broad formulation of regularity conditions. The following is a brief, and rough, description of the principal conditions which we believe are sufficient to justify the analyses in this paper.

It appears to be necessary to assume

$$(1)\qquad L((\xi, \eta), d) \leqslant V(\|\xi - d\|) \quad \text{and} \quad \left|\frac{\partial}{\partial d} L((\xi, \eta), d)\right| \leqslant V(\|\xi - d\|)$$

where V is a function satisfying a certain moment condition, as set forth in (2). We will also assume that $(\partial^k/\partial d^k)L((\xi, \eta), d)$ exists for certain moderate values of k, and that these partial derivatives are also bounded in absolute value by $V(\|\xi - d\|)$ for all (ξ, η), d. Note that condition (1) considerably restricts the dependence of L on η.

The essential condition on V is that for some moderately large integer K there is a $B < \infty$ such that for all ξ, η

$$(2)\qquad E_{\xi, \eta}((\|(x, y) - (\xi, \eta)\|^K + 1)(1 + V^K(\|x - \xi\|))) < B < \infty.$$

The constant K can (hopefully) be determined from an examination of the structure of the error terms which appear in the following analysis.

While the above conditions may be relaxed, they cannot be eliminated entirely. Many problems which do not satisfy (1)–(2) in their usual formulations may—via changes of variables, etc.,—be reformulated to satisfy (1)–(2). A description of some of the possible reformulations is one of the goals of Section 2.

1.2. *Proving inadmissibility*: *The approach.* Begin by investigating the difference in risks, Δ, between two given estimators δ and δ' when $|\delta - \delta'|$ is small. Under this nearness assumption and some other similar ones we produce an approximate differential expression for $\Delta_{\delta, \delta'}(\xi, \eta)$. It is intended that from this differential expression one can discover how, given δ, to find an estimator δ' which is better than the given δ.

Notation. Define

$$\gamma(x, y) = \delta(x, y) - x,$$

and

$$\lambda(x, y) = \delta'(x, y) - \delta(x, y).$$

Thus $\delta(x, y) = x + \gamma(x, y)$ and $\delta'(x, y) = x + \gamma(x, y) + \lambda(x, y)$. For the following computations assume that γ is bounded, and that λ is small and is a "smooth" function. The definition of "smooth" will be more apparent later.

As noted, the following derivations are given for the one-dimensional case where $L = W(d - \xi)$ and W has three continuous derivatives. The results for the general case can be derived by analogy, as described at the end of this section.

First Taylor expansion. Write

$$\Delta_{\delta,\delta'}(\xi,\eta) = E_{\xi,\eta}(W(x+\gamma-\xi) - W(x+\gamma+\lambda-\xi)).$$

(For simplicity of display many of the variables are suppressed in the above expressions and in similar ones below.) Then expand in a Taylor series about $x + \gamma - \xi$. Keeping terms to the second order in λ and indicating the remainder by $+ \cdots$ yields

$$\Delta = E_{\xi,\eta}\left(-\lambda W'(x+\gamma-\xi) - \frac{\lambda^2}{2} W''(x+\gamma-\xi) + \cdots\right). \tag{1}$$

Expanding about $x - \xi$ (with an exact Taylor expansion) then yields

$$\Delta = E_{\xi,\eta}\left(-\lambda W'(x-\xi) - \lambda\gamma W''((x-\xi)_\gamma) - \frac{\lambda^2}{2} W''(x-\xi) + \cdots\right) \tag{2}$$

where $(x-\xi)_\gamma$ is between $x - \xi$ and $x - \xi + \gamma$.

For convenience rewrite (2) as

$$\Delta = -\mathbf{D}(\lambda) - \mathbf{E}_\gamma(\gamma\lambda) - \mathbf{E}(\lambda^2/2) + \cdots \tag{2'}$$

where $\mathbf{D}$, $\mathbf{E}$, and $\mathbf{E}_\gamma$ are the linear operators defined by

$$\mathbf{D}(\lambda)(\xi,\eta) = E_{\xi,\eta}(\lambda(x,y)W'(x-\xi))$$
$$\mathbf{E}_\gamma(\lambda)(\xi,\eta) = E_{\xi,\eta}(\lambda(x,y)W''((x-\xi)_\gamma))$$
$$\mathbf{E}(\lambda)(\xi,\eta) = E_{\xi,\eta}(\lambda(x,y)W''(x-\xi)).$$

The case $\gamma \equiv 0$. Here $\mathbf{E}_\gamma(\gamma\lambda) \equiv 0$. For this reason the expression (2′) is generally easier to compute and analyze in this case.

Now, suppose that for each y, $\delta(x,y)$ is a one-one function of x. This situation is typical in the examples of interest. Then the transformation $(x,y) \to (\delta(x,y),y)$ is one-one. The statistical problem with coordinates $(x',y) = (\delta(x,y),y)$ is thus isomorphic to the original problem. If the original problem satisfies the moment conditions of Section 1.1, so will the transformed problem. In this transformed problem one has that $\gamma \equiv 0$ for the estimator corresponding to the original one. Therefore, the condition $\gamma \equiv 0$ can usually be obtained without essential loss of generality, but at the cost of considerable algebraic and analytic manipulations. (See Section 2.5 for an example.)

Special case—squared error loss. In this case $W'(x-\xi) = 2(x-\xi)$ and $W'' \equiv 2$. Furthermore $W''' \equiv 0$ so that there are no error terms in (1) and $\mathbf{E} = \mathbf{E}_\gamma$. Hence in this case (2′) is an exact expression with $\mathbf{E}(h) = \mathbf{E}_\gamma(h) = 2E(h)$ and $\mathbf{D}(h) = E_{\xi,\eta}(h(x,y)(x-\xi))$.

Second Taylor expansion. Expanding λ in a Taylor series and denoting error terms as before yields

$$(3)\qquad \mathbf{D}(\lambda) = E(\lambda(x,y)W'(x-\xi)) = m_{10}\lambda + m_{20}\lambda_1 + m_{11}\lambda_2 + \cdots + \left(\tfrac{1}{2}\right)m_{12}\lambda_{22} + \cdots$$

where

$$\lambda_i = \lambda_i(\xi,\eta) = \frac{\partial}{\partial x_i}\lambda(x_1, x_2)\Big|_{x_1=\xi,\, x_2=\eta}$$

$$\lambda_{ij} = \lambda_{ij}(\xi,\eta) = \frac{\partial^2}{\partial x_i\, \partial x_j}\lambda(x_1, x_2)\Big|_{x_1=\xi,\, x_2=\eta}$$

and

$$m_{ij} = m_{ij}(\xi,\eta) = E_{\xi,\eta}\left((x-\xi)^{i-1}(y-\eta)^j W'(x-\xi)\right).$$

The subscripts on m are chosen as above because *for squared error loss* $m_{ij} = 2E_{\xi,\eta}((x-\xi)^i(y-\eta)^j)$.

Similarly,

$$\mathbf{E}(\lambda^2/2) = E_{\xi,\eta}((\lambda^2(x,y)/2)W''(x-\xi)) = n\lambda^2/2 + \cdots,$$

and

$$(4)\qquad \mathbf{E}_\gamma(\lambda\gamma) = E_{\xi,\eta}(\lambda(x,y)\gamma(x,y)W''((x-\xi)_\gamma)) = \lambda\psi + \cdots$$

where

$$n = n(\xi,\eta) = E_{\xi,\eta}(W''(x-\xi)), \qquad \text{and}$$

$$\psi_\gamma(\xi,\eta) = \psi(\xi,\eta) = E_{\xi,\eta}(\gamma(x,y)W''((x-\xi)_\gamma)).$$

The fundamental expression. Substituting (3) and (4) into (2′) yields

$$(5)\qquad \Delta = \left[(-m_{10}-\psi)\lambda - m_{20}\lambda_1 - m_{11}\lambda_2 - \left(\tfrac{1}{2}\right)m_{12}\lambda_{22}\right] - n\lambda^2/2 + \cdots = \mathbf{Q}\lambda - n\lambda^2/2 + \cdots$$

Note that $\mathbf{Q}$ is a linear differential operator—in fact in many examples it is a parabolic partial differential operator. In special cases it may even be the negative of the heat operator (in which x is thought of as the time variable and y as the space variable).

Method I. To show that δ is inadmissible try to find a solution to the inequality

$$(6)\qquad \Delta_{\text{est}} = \mathbf{Q}\lambda - n\lambda^2/2 \geqslant 0.$$

If a solution to the inequality can be found then it may be that another solution can be found such that the error terms (indicated by $+\cdots$) in (5) are negligible for all ξ, η. Past experience and some of the general arguments in the following paragraphs concerning error terms indicate why this may often be the case. If such a solution can be found then of course δ is inadmissible, since $\delta' = \delta + \lambda$ is better.

(Usually λ will be very small. Hence δ' will not be much of an improvement over δ. It is virtually certain that δ' will itself be inadmissible. Therefore, even if the above approach proves that δ is not admissible, it does not seem likely to provide an *admissible* alternative which is better than δ. Nevertheless, information about an admissible alternative can be provided by combining (5) above and the results of Section 1.4.)

There is another approach which can be used when a solution to (6) can be found. This method is described in the following paragraphs.

Method II. The following result holds under regularity conditions weaker than those assumed previously.

THEOREM. *Let $S \subset \Omega$ be compact. Let $H : S \to (0, \infty)$ and assume that*

$$\inf_{(\xi, \eta) \notin S} H(\xi, \eta)\Delta(\xi, \eta) > 0. \tag{7}$$

Then either δ_0 is inadmissible or δ_0 is generalized Bayes for some (generalized) prior G satisfying

$$\int_\Omega H^{-1}(\xi, \eta) G(d(\xi, \eta)) < \infty. \tag{8}$$

A precise statement and proof of this result may be found in Brown (1980).

In many applications Ω is closed and it is possible to find a solution, λ, to (6) for which the error terms are asymptotically negligible in the sense that $\liminf_{\|(\xi, \eta)\| \to \infty} \Delta(\xi, \eta)/\Delta_{\text{est}}(\xi, \eta) \geqslant \zeta$. Then H may be chosen as any convenient positive continuous function such that $\liminf_{\|(\xi, \eta)\| \to \infty} H\Delta_{\text{est}} > 0$. In order to prove that δ is inadmissible it then suffices to show that it cannot be generalized Bayes for any prior satisfying (8). Sometimes, as in the Poisson example in Section 2.3, this can be shown directly. Otherwise a further study of properties of generalized Bayes estimators is required. This study is initiated in Section 1.3. The inadmissibility results in Sections 2.2 and 2.3 illustrate some possible applications of this method. (The above procedure is nonconstructive in the sense that it does not by itself produce a formula for an estimator better than the given δ.)

The unimportance of n. Usually $1/\varepsilon > n > \varepsilon$ for some $\varepsilon > 0$. If this is the case then it is not necessary to know the exact value of n to investigate the existence of a solution to (6).

Suppose $1/\varepsilon > n > \varepsilon$. If λ' is a solution to (6) then $\lambda = \varepsilon\lambda'/2$ is a solution to

$$Q\lambda - \lambda^2 \geqslant 0. \tag{9}$$

Conversely, if λ is a solution to (9) then $\lambda' = 2\varepsilon\lambda$ is a solution to (6). Hence under the above condition the existence of a solution to (9) is equivalent to the existence of a solution to (6).

The error terms when $\gamma = 0$. In most of the common applications $m_{20} > \varepsilon > 0$. Also assume that the coefficients which appear in (6) and in the error terms are smooth, bounded functions of ξ, η. (The boundedness actually follows from $|m_{ij}| < B$, $0 \leqslant i + j \leqslant K - 1$.)

Suppose a solution, λ, to (6) can be found with the following properties:

$$(10)\qquad \begin{aligned} &\text{(i)} \quad \Delta_{\text{est}} = \mathbf{Q}\lambda - n\lambda^2/2 \geqslant 0 \\ &\text{(ii)} \quad |\lambda_{ij}| < c\Delta_{\text{est}}, \ |\lambda_{jkl}| < c\Delta_{\text{est}} \\ &\text{(iii)} \quad \Delta_{\text{est}}(x_1, y_1)/\Delta_{\text{est}}(x_2, y_2) \leqslant 2 + \|(x_1, y_1) - (x_2, y_2)\|^8. \end{aligned}$$

Here, c will be some appropriate small constant; for example $c = 1/100B$. The coefficient "8" which appears in (iii) can certainly be reduced to a smaller value, but this, again, is a refinement we will not pursue here. (The coefficient K in the assumption (2) of Section 1.1 is related in a monotone way to the coefficient "8" which appears here, so a reduction of this coefficient will lead to a reduction of the value K needed in (2) of Section 1.1.)

When the coefficients which appear in the definition of Q are constants (or very nearly so) the technique of randomization of the origin can often be used to derive λ satisfying (6) and (ii), (iii) from any given solution of (6). See Brown (1975) and Berger (1975) for some examples.

With this background consider, for example, the following error term which appears in (3):

$$a = \left(\tfrac{1}{2}\right)E_{\xi,\eta}\left((x-\xi)^2(y-\eta)\lambda_{112}(\hat{x},\hat{y})W'(x-\xi)\right)$$

where $(\hat{x}, \hat{y})$ is on the line joining (ξ, η) to (x, y). By (ii), (iii) and (2) of Section 1.1.

$$\begin{aligned} |a| &\leqslant (c/2)E_{\xi,\eta}\left((x-\xi)^2|y-\eta|\Delta(\hat{x},\hat{y})|W'(x-\xi)|\right) \\ &\leqslant c\Delta(\xi,\eta)E_{\xi,\eta}\left((x-\xi)^2|y-\eta|(1 + \|(x,y)-(\xi,\eta)\|^8)|W'(x-\xi)|\right) \\ &\leqslant c\Delta(\xi,\eta)E_{\xi,\eta}\left((1 + \|(x,y)-(\xi,\eta)\|^{11})|W'(x-\xi)|\right) \\ &\leqslant cB\Delta(\xi,\eta) = \Delta(\xi,\eta)/100 \qquad (\text{for } c = 1/100B). \end{aligned}$$

There are twelve such error terms in (1)–(4). Each of them can be treated in a similar fashion. It follows that the magnitude of the total error is certainly bounded by $\Delta/2$ and that $\delta'(x,y) = x + \gamma(x,y) + \lambda(x,y)$ will, therefore, satisfy $\Delta_{\delta,\delta'}(\xi,\eta) > 0$ for all ξ, η. Hence δ is inadmissible.

If λ satisfies conditions (10) asymptotically as $\|\xi,\eta\| \to \infty$ rather than for all ξ, η then the above considerations indicate that the error terms should be negligible as $\|(\xi,\eta)\| \to \infty$, and Method II may then be applicable to prove inadmissibility of δ.

An approximation to ψ. Consider the approximation

$$(11)\qquad \psi = n\gamma + \cdots .$$

If this is substituted for ψ in (5) the basic inequality (6) becomes

$$(12)\qquad \hat{\mathbf{Q}}\lambda - n\lambda^2/2 \geqslant 0$$

where

$$(13)\qquad \hat{\mathbf{Q}}\lambda = (-m_{10} - n\gamma)\lambda - m_{20}\lambda_1 - m_{11}\lambda_2 - \left(\tfrac{1}{2}\right)m_{12}\lambda_{22}.$$

Computation of (12) is usually somewhat simpler than computation of (6).

In many examples the error in the approximation (11) is asymptotically negligible as $\|(x, y)\| \to \infty$. If this is so (and Ω is closed) then (6) and (12) will be asymptotically equivalent; equation (12) can be used in place of equation (6) when using Method II. It is also true that general considerations indicate that when the error is asymptotically negligible then (12) is solvable if and only if (6) is solvable. Therefore it is often more convenient to study the solubility of (6) by first examining the somewhat simpler inequality (12). The nature of the approximation (11) and the uses of (12) are clarified by some of the applications in Section 2.

Summary. The methods we have indicated above for proving inadmissibility of a given estimator, $\delta = x + \gamma$ (with γ bounded), are as follows:

(a) write down the fundamental inequality, (6) (or (9)) or (12). The coefficients of this inequality involve only γ and various (mixed) moments of $(x - \xi)$, $(y - \eta)$, and $W'(x - \xi)$. Then either:

(b) find a solution to this inequality,

(c) use the solution from (b) to find a "smooth" solution, λ, to (6)—one satisfying (10). This can be done by guesswork, by randomizing the origin, or by any other appropriate method.

(d) Write down the exact error terms from the expansions (1)–(4), substitute the λ from (c) and thus check directly that λ can be used to define an estimator better than δ. Or

(b′) find an "asymptotic" solution to (6) or (12)—that is, a solution which is valid on the complement of a compact subset, S, of Ω.

(c′) Check that the asymptotic solution mentioned in (b′) is asymptotically smooth (which it usually will be)—that is, that it satisfies conditions (10) on the complement of S. If the asymptotic solution in (b′) is not asymptotically smooth, try to modify it as in (c) to find one which is.

(d′) Write down the error terms from the expansion (1)–(4) to check directly that $\inf_{(\xi, \eta) \notin S} \Delta / \Delta_{\text{est}} > 0$.

(e′) Let $H > 0$ be any continuous function satisfying $\inf_{(\xi, \eta) \notin S} H \Delta_{\text{est}} > 0$ and then show that δ cannot be generalized Bayes for any generalized prior satisfying (8) (see Section 1.3).

In most problems Ω is closed in R^{kl}. If so then there is no need to explicitly identify S; instead one only needs to show that

$$\liminf_{\|(\xi, \eta)\| \to \infty} \Delta / \Delta_{\text{est}} > 0$$

and choose H so that

$$\liminf_{\|(\xi, \eta)\| \to \infty} H \Delta_{\text{est}} > 0.$$

In the applications of Method II which follow we mainly proceed in this latter fashion.

The program (a), (b), (c), (d) has been carried out explicitly by several authors for particular classes of examples (see Section 2.1 for references). It is our hope that reasonable regularity conditions can be found so that the steps (d) or (d′) can be carried out without the explicit computations mentioned above. This hope is supported by some general work which has been done on differential inequalities like (6). See, for example, Fujita (1966) and Portnoy (1975b).

Generalizations. (A) When $L((\xi, \eta), d)$ is not of the form $W(d - \xi)$, the above method can still be applied. The derivation of (6) is analogous and leads to the same fundamental equations except that $(\partial/\partial d)L((\xi, \eta), d)$ replaces $W'(d - \xi)$ throughout. The error terms are analogous to those discussed above for the case $L = W(d - \xi)$. Applications involving this situation can be found in Sections 2.6 and 2.3.

(B) When $k \geqslant 1$ or $l > 1$ the same general method of deriving a fundamental differential inequality is valid. Because one is now dealing with multidimensional Taylor series in place of ordinary ones (as in (1), (2)) or bivariate ones (as in (3)) the resulting expressions are notationally more complex. The basic principles are the same. In particular, in the analog of (3) all terms higher than the second order and all terms involving more than one partial derivative with respect to the "ξ" coordinates should be considered to be error terms. The coefficient "8" in the condition (10)(iii) may have to be increased but otherwise a discussion of the error terms in this multivariate case can proceed like the discussion above after making the obvious changes of notation. For examples of the basic equations which result in such cases see 2.2(1), 2.3(1), and 2.5(1).

1.3. *Generalized Bayes estimators.* This section contains an approximate formula for the functional value of a generalized (or, formal) Bayes estimator. The techniques used derive from those in the preceding section; the resulting formula is intimately related to the basic expressions derived there.

The result here serves three purposes. First, in order to prove inadmissibility of a generalized Bayes estimator by the method of the preceding section one needs to know the value of $\gamma(x, y) = \delta(x, y) - x$ for this estimator. The formula here provides an approximation to γ which is sufficiently accurate to prove inadmissibility in many applications. Since this formula only involves the same moments as formula 1.2(6) it is generally much easier to compute the approximate γ from it than it is to compute γ exactly from the definition. Second, formula (2) can sometimes be used in Method II to verify that the prior does not satisfy 1.2(8). Third, the formulae of this section form a central part of the method to be described in the next section for proving admissibility of certain generalized Bayes estimators.

Expansion for γ_g. Let G be a given nonnegative measure on Ω (generalized prior). If it is absolutely continuous let $g \geqslant 0$ denote its density relative to Lebesgue measure. Assuming that W is differentiable and strictly convex (and often without this assumption) the generalized Bayes estimator for g—to be denoted by $\delta_g(x, y) = x + \gamma_g(x, y)$—is the unique root of

(1) $$0 = \int W'(x + \gamma_g - \xi)p_{\xi, \eta}(x, y)g(\xi, \eta)\, d\xi\, d\eta.$$

An exact Taylor expansion yields

$$\text{(2)} \qquad \gamma_g(x,y) = -\frac{\int W'(x-\xi)p_{\xi,\eta}(x,y)g(\xi,\eta)\,d\xi\,d\eta}{\int W''((x-\xi)_\gamma)p_{\xi,\eta}(x,y)g(\xi,\eta)\,d\xi\,d\eta} = -\frac{\mathbf{D}^*(g)}{\mathbf{E}_\gamma^*(g)}.$$

The above expression defines $\mathbf{D}^*$ and $\mathbf{E}_\gamma^*$.

The notation is chosen to indicate the essential fact that $\mathbf{D}^*$ and $\mathbf{E}_\gamma^*$ are the Hilbert space duals of $\mathbf{D}$ and $\mathbf{E}_\gamma$—that is

$$\text{(3)} \quad \langle s, \mathbf{D}^*g\rangle = \int s(x,y)\mathbf{D}^*(g)(x,y)\,dx\,dy = \int \mathbf{D}(s)(\xi,\eta)g(\xi,\eta)\,d\xi\,d\eta = \langle \mathbf{D}s, g\rangle$$

for continuous functions, s, with compact support, and similarly for $\mathbf{E}_\gamma$ and $\mathbf{E}_\gamma^*$.

Equation 1.2(3) states that $\mathbf{D}$ is approximately equal to a certain linear partial differential operator. It is therefore plausible that $\mathbf{D}^*$ should be approximately equal to the dual of that operator. Thus

$$\text{(4)} \qquad \mathbf{D}^* \approx \left(m_{10} + m_{20}\frac{\partial}{\partial\xi} + m_{11}\frac{\partial}{\partial\eta} + \left(\tfrac{1}{2}\right)m_{12}\frac{\partial^2}{\partial\eta^2}\right)^* = \left(m_{10} - \frac{\partial}{\partial\xi}m_{20} - \frac{\partial}{\partial\eta}m_{11} + \left(\tfrac{1}{2}\right)\frac{\partial^2}{\partial\eta^2}m_{12}\right).$$

Similarly

$$\text{(4)}' \qquad \mathbf{E}_\gamma^* \approx (n_\gamma)^* = (n_\gamma),$$

where $n_\gamma = E_{\xi,\eta}(W''(x-\xi)_\gamma)$. Note that if γ is small $n_\gamma \approx E(W''(x-\xi)) = n$.

Collecting the above yields

$$\text{(5)} \qquad \gamma_g \approx -\frac{m_{10}g - \dfrac{\partial}{\partial\xi}(m_{20}g) - \dfrac{\partial}{\partial\eta}(m_{11}g) + \left(\tfrac{1}{2}\right)\dfrac{\partial^2}{\partial\eta^2}(m_{12}g)}{ng} = \tilde{\gamma}_g \text{ (def.)}.$$

The expressions (4)–(5) are not exact equalities. Without further discussion it is possible only to interpret the symbol "$\approx$" as, "is approximately equal to". With this crude interpretation (5) has little practical value. However, if g is well behaved near ∞ then (5) should also be an asymptotic ratio-equality. (The symbol, "$\approx$", can then be interpreted as, "$\sim$".) A suitable requirement seems to be $\|\nabla g\|/g \to 0$ as $\|(\xi,\eta)\| \to \infty$. The approximation in (5) then appears to be sufficiently accurate for many applications of the methods described in the preceding section. Some of these applications are described in Sections 2.2 and 2.3 and also in Berger (1976c).

The heuristic asymptotic value for γ derived above is more accurate than those proved in Meeden and Isaacson (1977) and Umbach (1976), but involves more restrictive regularity conditions.

Proving inadmissibility by method II. Suppose γ is a given smooth function. In order to prove that $\delta = x + \gamma$ is inadmissible by Method II it must be shown that if δ is generalized Bayes for prior G then $\int H^{-1}(\xi, \eta)G(d(\xi, \eta)) = \infty$. If G has density g it must, therefore, be shown that

$$\int H^{-1}(\xi, \eta)g(\xi, \eta)d(\xi, \eta) = \infty. \tag{6}$$

Suppose g is sufficiently smooth. Then it must very nearly satisfy the differential equation

$$\begin{aligned} \mathbf{Q}^*g = -m_{10}g + \frac{\partial}{\partial\xi}(m_{20}g) + \frac{\partial}{\partial\eta}(m_{11}g) \\ -\tfrac{1}{2}\frac{\partial^2}{\partial\eta^2}(m_{12}g) - \psi g = 0 \end{aligned} \tag{7}$$

for large values of $\|\xi, \eta\|$. The operator $\mathbf{Q}^*$ defined by (7) is the dual of the operator $\mathbf{Q}$ introduced in 1.2(5). If $\gamma \to 0$ as $\|\xi, \eta\| \to \infty$ then ψ may be replaced by $n\gamma$ in the preceding assertion. (The operator which results should be denoted by $\hat{\mathbf{Q}}^*$, but we will simplify the notation by using the symbol $\mathbf{Q}^*$ for both cases.) Suppose any solution to (7) for large values of $\|(\xi, \eta)\|$ satisfies (6) (in applications it is often easy to ascertain whether this is the case.) Then this is strong evidence that δ is inadmissible. Of course it is not a proof of this.

The following line of argument provides further evidence. Possibly some regularity conditions may be supplied so that this line can be converted to a proof. Let s be any probability density for which the derivatives of $\ln s$ are small. Then if δ is generalized Bayes for G, having density g,

$$\begin{aligned} 0 &= \langle s, -\mathbf{D}^*g - \gamma\mathbf{E}_\gamma^*g\rangle \\ &= \langle(-\mathbf{D} - \mathbf{E}_\gamma)s, g\rangle \\ &\leqslant \langle\mathbf{Q}s + \rho s, g\rangle \\ &= \langle s, (\mathbf{Q}^* + \rho)g\rangle \end{aligned} \tag{8}$$

where ρ denotes an error bound which is small as $\|(\xi, \eta)\| \to \infty$, as described in Section 1.2. Formula (8) also has a valid interpretation if g is the generalized function $g = dG$. Depending on the specific context of the problem it may be possible to verify that ρ is of a sufficiently small order as $\|(\xi, \eta)\| \to \infty$ so that any solution to

$$\mathbf{Q}^*g + \rho g \geqslant 0 \tag{9}$$

satisfies $\int H^{-1}(\xi, \eta)g(\xi, \eta)d(\xi, \eta) = \infty$. (Again, this last assertion involves only a standard partial differential equation, and so should be relatively easy to check.) The inequality (8) may be interpreted as saying that (9) holds in the mean, and ought to also imply that $\int H^{-1}(\xi, \eta)g(\xi, \eta)d(\xi, \eta) = \infty$. (We think!)

The location parameter estimation problem is treated in Section 2.2. That section contains an outline for a proof of inadmissibility by Method II. That outline involves statements resembling (8). The precise reasoning discussed following (9) is not used there; instead a judicious choice of the density s in (8) makes possible a different approach.

Expansion for $\gamma_{hg} - \gamma_g$. For proving admissibility as discussed in the next section a closely related approximation is needed. Let g be a generalized prior density, as before; let hg be another one, where h is uniformly a smoothly differentiable function. Let h_i', $h_{i,j}''$, etc., denote the partial derivatives of h. Then, from (5),

$$(10) \quad \gamma_{hg} - \gamma_g \approx \tilde{\gamma}_{hg} - \tilde{\gamma}_g = \left(m_{20}h_1' + m_{11}h_2' - ((m_{12}g)_2/g)h_2' - \left(\tfrac{1}{2}\right)m_{12}h_{22}''\right)/nh.$$

(Here the symbol, "$\approx$", may again often be interpreted as "$\sim$".)

Actually, all that is required for most applications of the method in the next section is the weaker inequality

$$(11) \quad |\gamma_{hg} - \gamma_g| = 0((m_{20}h_1' + m_{11}h_2' - ((m_{12}g)_2/g)h_2' - m_{12}h_{22}''/2)/nh).$$

The regularity conditions for (11) to hold appear to be significantly weaker than those required for (5) or (10). In particular (11) appears to us to be valid if the moments m_{ij} and the function h are appropriately smooth functions as $\|(\xi, \eta)\| \to \infty$; with only very minimal conditions required on g.

Generalizations. As in the previous section the above results generalize to cases where $L((\xi, \eta), d)$ is not of the form $W(d - \xi)$ or where the dimensions k, l do not satisfy $k = l = 1$.

1.4. *Proving admissibility.* Make use of Stein's necessary and sufficient condition for admissibility, see Stein (1955), and also Farrell (1968, Section 3). In the current context a necessary and sufficient condition for admissibility of an estimator, δ, is the existence of a sequence $\{g_i\}$ of generalized prior densities each having finite mass ($\int g_i(\xi, \eta)\, d\xi\, d\eta < \infty$) and satisfying $\int_{\|\xi, \eta\| < 1} g_i(\xi, \eta)\, d\xi\, d\eta \geq \varepsilon > 0$ for all i and

$$(1) \quad \Lambda_i = \int \left[W(x + \gamma(x, y) - \xi) - W\left(x + \gamma_{g_i}(x, y) - \xi\right)\right] p_{\xi, \eta}(x, y) \times g_i(\xi, \eta)\, dx\, dy\, d\xi\, d\eta \to 0.$$

Expanding the first part of the integrand about $x + \gamma_{g_i} - \xi$, interchanging orders of integration, and using 1.3(1) yields:

$$(2) \quad \Lambda_i = \tfrac{1}{2}\int \left(\gamma(x, y) - \gamma_{g_i}(x, y)\right)^2 \left(\int W''((x - \xi)_*) p_{\xi, \eta}(x, y) g_i(\xi, \eta)\, d\xi\, d\eta\right) dx\, dy$$

where $(x - \xi)_*$ lies between $x + \gamma_{g_i}(x, y) - \xi$ and $x + \gamma(x, y) - \xi$.

Now, suppose γ is a generalized Bayes procedure, say $\gamma = \gamma_g$. Let $g_i = h_i g$ and suppose h_i is a uniformly smoothly differentiable function so that the bound 1.3(11) applies. Similar to the situation in Section 1.3 one typically has

$$\int W''((x - \xi)_*) p_{\xi, \eta}(x, y) h_i(\xi, \eta) g(\xi, \eta)\, d\xi\, d\eta \approx n h_i(x, y) g(x, y)$$

for all $i = 1, 2, \cdots$. We actually need only the weaker and easier to verify the fact that

$$(3)\qquad \int W''((x-\xi)_*)p_{\xi,\eta}(x,y)h_i(\xi,\eta)g(\xi,\eta)\,d\xi\,d\eta \leqslant B_1 h_i(x,y)g(x,y)$$

for some $B_1 < \infty$. The above yields

$$(4)\qquad \begin{aligned}\Lambda_i &\leqslant B\int(\tilde{\gamma}_g(x,y)-\tilde{\gamma}_{h_ig}(x,y))^2 h_i(x,y)g(x,y)\,dx\,dy\\ &= B\tilde{\Lambda}_i \text{ (definition)}.\end{aligned}$$

Equation (4), together with 1.3(10) which defines $\tilde{\gamma}_g - \tilde{\gamma}_{h_ig}$, is the basic equation of this development. If, for a given g, it is possible to find a suitable sequence of function h_i such that $\tilde{\Lambda}_i \to 0$ then γ_g should be admissible.

As with the method outlined in Section 1.2 the above development can actually be used to *prove* that γ_g is admissible if one writes down suitable error terms in the above development and shows that they are indeed negligible. If so, then Λ_i (as defined in (1)) tends to zero, and this proves admissibility of $\delta_g = x + \gamma_g$.

Generalizations. Again, the above method is adaptable to the case of general loss functions, L, and to arbitrary dimensions k, l.

The method is suitable for certain types of situations where g is not smooth—even for situations where the generalized prior is described by a measure G which does not have a density with respect to Lebesgue measure. In those cases h_i still must be sufficiently smooth. The factor $h_i(x,y)g(x,y)$ in (4) should be replaced by

$$\int h_i(\xi,\eta)p_{\xi,\eta}(x,y)G(d\xi\,d\eta).$$

(Similarly g in 1.3(5) should be replaced by $\int p_\xi(x,y)G(d\xi\,d\eta)$.) The admissibility proof in Brown (1971) provides a detailed argument of this type for the estimation of several normal means with squared error loss, in which certain error terms disappear by virtue of the special nature of the problem.

2. Applications and conjectures.

2.1. *Summary of published examples.* The methods outlined in Section 1 have been successfully used in the past to prove admissibility or inadmissibility in a number of special cases. The interested reader may refer to these examples to reinforce the plausibility of our contentions concerning insignificance of error terms, and to discover various techniques of computation, approximation, and smoothing which can be used to prove admissibility or inadmissibility in specific situations via the methods we have described (except for our Method II, which is new).

C. Stein's (1956) original proof of the inadmissibility of the best invariant estimator of three normal means follows the pattern outlined in Section 1.2 for the case $k \geqslant 3$, $l = 0$. The pattern can perhaps be more clearly seen from reading the account in Brown (1975). Stein's inadmissibility proof was generalized in Brown (1966, Part 2) using the same method. Berger (1976a) uses an interesting adaptation of our method to find reasonable alternatives to the best invariant estimator in this situation. See also Berger and Srinivasan (1977).

The admissibility proof for the best invariant estimator of a single location parameter given in Stein (1959) uses essentially the method outlined in Section 1.4. Of course the special features of the problem he treats simplify some of the expressions, especially the approximate formula 1.3(5) for γ_g. R. Farrell (1964) generalized some of Stein's results. In particular he was able to prove some results using this method for certain generalized Bayes estimators. The admissibility results of Brown (1971) also follow essentially the method outlined above, with certain modifications and simplifications made possible by the special nature of the problem treated there (normal means, squared error loss).

Berger (1976b, c, d) uses our methods to prove both admissibility and inadmissibility results for the problem of estimating one or two out of several unknown location parameters.

Stein (1965) gave a heuristic discussion of the admissibility problem for dimension $k = 1, l = 0$ which he hoped would lead to a proof of admissibility for certain generalized Bayes estimators under suitable regularity conditions. This hope was realized in Zidek (1970). Stein's method and Żidek's proof are closely related to the method we have outlined. For the problem he considers, which involves a squared error loss function and no nuisance parameters, Zidek's equation (3.6) is analogous to our 1.3(10), and involves less restrictive regularity conditions. Under our smoothness and boundedness assumptions and the condition $m_{20} > \varepsilon > 0$ the two equations become effectively equivalent. Their derivations differ—Zidek's derivation exploits the special properties of squared error loss and uses the Cauchy-Schwarz inequality. No Taylor expansions or error terms are involved at this stage of his argument. Zidek's result was slightly generalized and was applied to certain ANOVA problems by Portnoy (1971). See also Zidek (1973, Section 3).

In the following we apply our heuristic methods to several statistical problems. While we have not proved the assertions concerning admissibility which we make below, we have examined the error terms in these examples enough to convince ourselves of the plausibility that these assertions are valid.

2.2. *Estimating location parameters* (*without nuisance parameters*): *Setup.* Suppose one observes $x \in R^k$ with density $p(x - \xi)$, $\xi \in R^k$. It is desired to estimate ξ by $d \in R^k$. The loss function is $W(d - \xi)$. Assume $E(\|x - \xi\|^K) < \infty$ for some sufficiently large value of K in order to satisfy 1.1(1). Suppose $\delta = x + \gamma$ is a given estimator and $\delta' = x + \gamma + \lambda$ ($\lambda \in R^k$).

In the following we make the conjecture that for generalized priors of the form $g(\xi) \sim (\xi' D \xi)^{b/2}$ as $\|\xi\| \to \infty$ the generalized Bayes estimator is inadmissible if $b + k - 2 > 0$ and admissible if $b + k - 2 \leqslant 0$. This generalized Bayes estimator is conjectured to have the asymptotic form (4).

A change of location of the original problem yields $E_0((\partial/\partial x_i)W(x)) = 0$, $1 \leqslant i \leqslant k$. Make this assumption without loss of generality. Let $w''_{ij} = (\partial^2/\partial x_i\, \partial x_j)W$.

The appropriate version of 1.2(6) is

(1) $$\Delta_{\text{est}} = -\psi^{\tau}\lambda - \nabla \cdot (M\lambda) - (\tfrac{1}{2})\lambda^{\tau}N\lambda = \mathbf{Q}\lambda - (\tfrac{1}{2})\lambda^{\tau}N\lambda$$

where λ is a column vector, τ denotes transpose, "$\nabla \cdot$ " denotes divergence;

$$M = (m_{ij}),\ m_{ij} = E_0\left(x_i \frac{\partial}{\partial x_j} W(x)\right);$$

$$N = (n_{ij}),\ n_{ij} = E_0(w_{ij}''(x));$$

$$\psi = (\psi_i),\ \psi_i(\xi) = E_\xi\left(\Sigma_{j=1}^{k} w_{ij}''((x-\xi)_{\gamma,i})\gamma_j(x)\right)$$

where $(x-\xi)_{\gamma,i}$ is on the line joining $x-\xi$ to $x-\xi+\gamma(x)$, and depends also on the coordinate, i. (The terms involving M form the ànalog to that involving m_{20} in 1.2(6). There are no terms analogous to those in 1.2(6) which involve m_{10} because of the assumption $E_0((\partial/\partial x_i)W(x)) = 0$. No second derivative terms appear in (1) because there are no nuisance parameters or observations. The appropriate versions of the operators defined in 1.2(2′) are

$$\mathbf{D}\lambda = \int \Sigma_{i=1}^{k} \lambda_i(x) \frac{\partial}{\partial x_i} W(x-\xi)p(x-\xi)\,dx$$

$$\mathbf{E}_\gamma(\lambda, \gamma) = \int \lambda^{\tau}(x)w''(x,\xi)\gamma(x)p(x-\xi)\,dx$$

where $(w''(x,\xi))_{ij} = W_{ij}''(x-\xi)_{\gamma,i}$; and $\mathbf{E}(\lambda, \lambda)$ is similarly defined.)

Assume N is positive definite and M is nonsingular. Note that if W is squared error loss $M = 2\Sigma$ and $N = 2I$, but in general M need not be symmetric. (The assumption that M be nonsingular first appears in Brown (1966).)

Generalized Bayes procedures. The analog of 1.3(5) is

(2) $$\gamma_g \approx \frac{N^{-1}M^{\tau}\nabla g}{g} = \tilde{\gamma}_g.$$

Assume a prior generalized density of the form $g(\xi) \sim (\xi^{\tau}D\xi)^{b/2}$ as $\|\xi\| \to \infty$, and whose first two derivatives are asymptotic to the corresponding derivatives of $(\xi^{\tau}D\xi)^{b/2}$. Then

(3) $$\tilde{\gamma}_g(x) = \frac{bN^{-1}M^{\tau}Dx}{x^{\tau}Dx}.$$

The location family structure of the problem makes it relatively straightforward to directly write down error terms for the approximation indicated by (2) and (3). Having done this for some special cases we feel confident in asserting that $\|\tilde{\gamma}_g - \gamma_g\| = 0(\|x\|^{-2})$ for the above situation and that $|(\partial/\partial x_i)(\tilde{\gamma}_g - \gamma_g)_j| = 0(\|x\|^{-3})$. Note that this assertion is consistent with the discussion following 1.3(5).)

Inadmissibility results. Suppose

(4) $$\gamma(x) = \frac{bN^{-1}M^{\tau}Dx}{x^{\tau}Dx} + 0(\|x\|^{-2}) \qquad \text{as} \quad \|x\| \to \infty$$

where D is positive definite. Suppose also that $(\partial/\partial x_j)\gamma_i = 0(\|x\|^{-2})$. This form of γ is motivated by (3).

Suppose $b + k - 2 > 0$. Let $\nu = \max(x^\tau M^{\tau^{-1}} N M^{-1} x / x' D x) < \infty$, and $0 < c < 2(b + k - 2)/\nu$. The assumptions on W and γ yield $\psi = N\gamma + 0(\|\xi\|^{-2})$. Let

$$\lambda = \lambda(x) = -\frac{cM^{-1}x}{1 + x^\tau D x}. \tag{5}$$

Then

$$\begin{aligned} \Delta_{\text{est}} &= -\psi^\tau \lambda + c\frac{k + (k-2)\xi^\tau D\xi - (c/2)\xi^\tau (M^\tau)^{-1} N M^{-1}\xi}{(1 + \xi^\tau D \xi)^2} \\ &\geqslant c\left(\frac{k + (b + k - 2 - c\nu/2)\xi^\tau D\xi}{(1 + \xi^\tau D\xi)^2}\right) \\ &\quad + 0\left((\xi^\tau D\xi)^{-\frac{3}{2}}\right) \sim c(b + k - 2 - c\nu/2)(\xi^\tau D \xi)^{-1}. \end{aligned} \tag{6}$$

Choose $H(\xi) = 1 + \xi^\tau D\xi$. A careful investigation of error terms should show that

$$\lim_{\xi^\tau\xi\to\infty} H(\xi)\Delta(\xi) = c(b + k - 2 - c\nu/2) > 0, \tag{7}$$

as can be seen from similar computations in Brown (1966).

The appropriate version of 1.3(7) is

$$\mathbf{Q}^* g = + M^\tau \nabla g - g\psi = 0.$$

For large values of $\|\xi\|$ this equation becomes approximately

$$+ M^\tau \nabla g - g\frac{bM^\tau D\xi}{\xi^\tau D\xi} = 0 \tag{8}$$

which has the solution $g(\xi) = (\xi^\tau D\xi)^{b/2}$ in accordance with the derivation of (3). Now

$$\int H^{-1}(\xi) g(\xi)\, d\xi = \int (\xi^\tau D\xi)^{(b-2)/2}\, d\xi = \infty \tag{9}$$

since $b + k - 2 > 0$ by assumption. This motivates the conjecture that δ is inadmissible according to the discussion following 1.3(7).

(We note parenthetically that it is possible to find functions λ so that $\Delta_{\text{est}}(\xi) > 0$ for all $\xi \in R^k$. For example the function $\lambda = -(\alpha(x)M^{-1}x)/(1 + x\tau Dx)$ with $\alpha(x) = c\exp(-a\cot^{-1}(\|x\|/d))$ yields such a result whenever a and d are sufficiently large. (How large these constants must be depends in part on $\sup\|\psi(x)\|$.) However such choices of λ probably do not have Δ_{est} sufficiently near Δ because for small and moderate $\|x\|$ some second derivatives of λ are not small compared to Δ_{est}.)

For some special given estimators it is possible to find a λ satisfying $\Delta > 0$ by a direct approach. The simplest example is the case of the best invariant estimator where $\gamma \equiv 0$ and $k \geqslant 3$. Then one may choose $\lambda = -(cM^{-1}x)(a + x^\tau x)$. It follows from (6) that $\Delta_{\text{est}} > 0$ for all $\xi \in R^k$. It can be proven that the error terms are insignificant when a is sufficiently large so that also $\Delta > 0$ for all $\xi \in R^k$. A complete treatment of all the error terms is contained in Brown (1966).

A proof of inadmissibility. The following argument involves steps similar to 1.3(8). Even though the argument is only for the one dimensional case we feel that it demonstrates the general plausibility and feasibility of Method II for proving inadmissibility as described in Section 1.2 and 1.3. The final part of the following argument differs from that indicated in 1.3(9), and so illustrates an alternative to that part of Section 1.3.

Make the following special assumption in addition to the general assumptions of this section: (i) $k = 1$. (ii) The convex hull of the support of p is a bounded set. Call it $[\underline{K}, \overline{K}]$. (iii) Let $m = m_{11} > 0$. Define

$$v(t) = m^{-1}\textstyle\int_t^{\overline{K}} W'(x)p(x)\,dx. \tag{10}$$

Assume $v(t) > 0$ for $t \in (\underline{K}, \overline{K})$. (This assumption will be satisfied if W is strictly bowl shaped, and often otherwise so long as $W'(\underline{K}) < 0 < W'(\overline{K})$.) (iv) $-\overline{K} < \inf \gamma(x)$ or $\sup \gamma(x) < -\underline{K}$. (This assumption is needed in order to guarantee that δ is not a Bayes procedure for some prior having compact support.)

Assume, without loss of generality, that $N = 1$ and $D = 1$ in the definition, (4), of γ. In this case $\gamma \sim mb/\xi$ and the basic linear operator is $\mathbf{Q}\lambda = -m\lambda' - \psi\lambda$ where $\psi \sim n\gamma \sim mb/\xi$ with $b > 1$. The asymptotic form of the dual is $\mathbf{Q}^*g \sim mg' - mbg/\xi$. This has the solution $g(\xi) = \xi^b$ which satisfies $\int H^{-1}(\xi)g(\xi)\,d\xi = \int(1 + \xi^2)^{-1}\xi^b\,d\xi = \infty$.

We wish to show that

$$-\mathbf{D}^*g - \gamma\mathbf{E}_\gamma^*g = 0 \tag{11}$$

implies $\int(1 + \xi^2)^{-1}g(\xi)\,d\xi = \infty$. (In order to maintain complete generality it is necessary to allow the interpretation here that $g = dG$ is a generalized function.) To this end, assume (11) holds, and also suppose that

$$\int(1 + \xi^2)^{-1}g(\xi)\,d\xi < \infty. \tag{12}$$

Let

$$\begin{aligned} s_{\alpha,\sigma}^{(B)}(x) = s_{\alpha,\sigma}(x) &= \frac{e^{-\alpha x}}{1 + |x|/\sigma}\operatorname{sgn} x \quad |x| \geq B \\ &= 0 \qquad\qquad |x| < B, \quad 0 \leq \alpha \leq 1 \leq \sigma. \end{aligned}$$

Then there is a $B' < \infty$ such that for any $B > B'$

$$-\mathbf{D}s_{0,\sigma}(\xi) - \mathbf{E}_\gamma\gamma s_{0,\sigma}(\xi) < 0 \qquad \text{for} \quad \xi > B - \underline{K} \text{ and } \xi < -B - \overline{K}, \tag{13}$$

for all σ. This assertion could be proven using arguments like those displayed in the discussion of error terms in Section 1.2. Here is an even more efficient argument:

note that $v(t) = 0$ for $t \notin [\underline{K}, \bar{K}]$ and that $\int v(t)\,dt = 1$. Then for ξ as in (13)

$$(14)\qquad \begin{aligned} -\mathbf{D}s &= -\int s(x) W'(x-\xi) p(x-\xi)\,dx \\ &= -m \int s'(x) v(x-\xi)\,dx \\ &\sim m\sigma^{-1}(1+|\xi|/\sigma)^{-2} \text{ uniformly in } \sigma. \end{aligned}$$

Similarly

$$(15)\qquad \begin{aligned} -\mathbf{E}_\gamma \gamma s &= -\int \gamma(x) s(x) W''((x-\xi)_\gamma) p(x-\xi)\,dx \\ &\sim -mb|\xi|^{-1}(1+|\xi|/\sigma)^{-1} \text{ uniformly in } \sigma, \end{aligned}$$

since $n = \int W''(x-\xi)p(x-\xi)\,dx = 1$ by assumption. This verifies (13).

Calculations similar to (14), (15) verify that for fixed B, σ

$$(16)\qquad \sup_\alpha |-\mathbf{D}s_{\alpha,\sigma} - \mathbf{E}_\gamma \gamma s_{\alpha,\sigma}| = 0\big((1+\xi^2)^{-1}\big).$$

Thus,

$$(17)\qquad \begin{aligned} 0 &= \lim_{\alpha\to 0} \int_{-\infty}^{\infty} (-\mathbf{D}s_{\alpha,\sigma} - \mathbf{E}_\gamma \gamma s_{\alpha,\sigma}) G(d\xi) \\ &= \int_{-\infty}^{\infty} (-\mathbf{D}s_{0,\sigma} - \mathbf{E}_\gamma \gamma s_{0,\sigma}) G(d\xi) \end{aligned}$$

by (12), (16) and the dominated convergence theorem.

It follows from (13) that

$$(18)\qquad \Big\{\int_{B-\underline{K}}^{\infty} + \int_{-\infty}^{B-\bar{K}}\Big\}(-\mathbf{D}s_{0,\sigma} - \mathbf{E}_\gamma \gamma s_{0,\sigma}) G(d\xi) < 0.$$

Now, note that

$$(19)\qquad \begin{aligned} \int_0^{B-\underline{K}} -\mathbf{D}s_{0,\sigma}(\xi) G(d\xi) &= m \int_{B-\bar{K}}^{B-\underline{K}} \Big[-(1+B/\sigma)^{-1} v(B-\xi) \\ &\qquad + \int_B^{\bar{K}+\xi} \sigma^{-1}(1+x/\sigma)^{-2} v(x-\xi)\,dx\Big] G(d\xi) \\ &= -(1+B/\sigma)^{-1} m \int v(B-\xi) G(d\xi) \\ &\qquad + \sigma^{-1} 0\big(G([B-\bar{K}, B\,\underline{K}])\big) \end{aligned}$$

by an integration by parts similar to that involved in (14). Also

$$(20)\qquad \begin{aligned} \int_0^{B-\underline{K}} -\mathbf{E}_\gamma \gamma s_{0,\sigma}(\xi) G(d\xi) &= 0\big(B^{-1} G([B-\bar{K}, B-\underline{K}])\big) \\ &< \frac{CmG([B-\bar{K}, B-\underline{K}])}{B} \end{aligned}$$

for an appropriate $C < \infty$.

Fix $\varepsilon > 0$. Note that $v(B-\xi) > c_\varepsilon > 0$ for $\xi \in (B - \bar{K} + \varepsilon, B - \underline{K} - \varepsilon)$ by assumption (iii). Hence

$$(21a)\qquad \int v(B-\xi) G(d\xi) > c_\varepsilon G\big((B - \bar{K} + \varepsilon, B - \underline{K} - \varepsilon)\big).$$

Note that condition (iii) on $v(\cdot)$ also implies that $W'(\underline{K}) < 0 < W'(\bar{K})$. It can be shown that, as a consequence, there are $\varepsilon_1, \varepsilon_2 > 0$ such that our assumption that

$|\gamma(x)| \to 0$ as $x \to \infty$ implies

(21b) $$\frac{\max\{G([B-\bar{K}, B-\bar{K}+\varepsilon_2]), G([B-\underline{K}-\varepsilon_2, B-\underline{K}])\}}{G([B-\bar{K}, B-\underline{K}])} < 1-\varepsilon_1$$

for all B sufficiently large. (Reasonable assumptions on the form of W other than (iii) will also yield this statement.)

Let $\varepsilon = \varepsilon_2/2$ and choose $B > C/c^2\varepsilon_1^2$ with $0 < c < \min(c_\varepsilon, \frac{1}{2})$. Let B' satisfy $\varepsilon < B' - B < \bar{K} - \underline{K} - 2\varepsilon$ and define $a = G([B-\bar{K}, B-\underline{K}])$, $a_1 = G([B-\bar{K}, B-\bar{K}+\varepsilon])$, $a_2 = G((B-\bar{K}+\varepsilon, B-\underline{K}-\varepsilon))$, $a_3 = G([B-\underline{K}-\varepsilon, B-\underline{K}]) = a - a_1 - a_2$ and $a', \cdots, a_3'$ similarly with B' replacing B. Note that $a_1' \leqslant a_2$ and $a_2' \geqslant a_3$ by set inclusion. Suppose

(21c) $$\int v(B-\xi)G(d\xi) < \frac{Ca}{B}.$$

Then $a_2 < c\varepsilon_1^2 a$ by (21a). Hence $a_1' \leqslant a_2 < c\varepsilon_1^2 a$. On the other hand, $a_1/a < 1 - \varepsilon_1 = (a_1 + a_2 + a_3)/a - \varepsilon_1$ by (21b) so that $a_2'/a \geqslant a_3/a > \varepsilon_1 - c\varepsilon_1^2$. Hence $a_1'/a_2' \leqslant c\varepsilon_1/(1-c\varepsilon_1)$. Applying (21b) at B' yields $a_3' < (1-\varepsilon_1)(a_1' + a_2' + a_3')$ so that $\varepsilon_1 a_3'/a_2' < (1-\varepsilon_1) + (1-\varepsilon_1)a_1'/a_2' \leqslant 1 - \varepsilon_1 + (1-\varepsilon_1)c\varepsilon_1/(1-c\varepsilon_1) = (1-\varepsilon_1)/(1-c\varepsilon_1)$. Thus $a_2'/a' > [c\varepsilon_1/(1-c\varepsilon_1) + 1 + (1-\varepsilon_1)\varepsilon_1(1-c\varepsilon_1)]^{-1} = (\varepsilon_1 - c\varepsilon_1^2) > c\varepsilon_1^2 > C/cB > C/cB'$. Hence either (21c) is false or

(21d) $$\int v(B'-\xi)G(d\xi) > \frac{CG([B'-\bar{K}, B'-\underline{K}])}{B'}$$

for all B' as above. It follows that for some (actually, most) large B, σ

(22) $$\int_0^{B-\underline{K}}(-\mathbf{D}s_{0,\sigma}(\xi) - \mathbf{E}_\gamma s_{0,\sigma}(\xi))G(d\xi) < 0.$$

Similarly B, σ can be chosen so that also

(23) $$\int_{-B-\bar{K}}^{0}(-\mathbf{D}s_{0,\sigma}(\xi) - \mathbf{E}_\gamma s_{0,\sigma}(\xi))G(d\xi) < 0.$$

The combination of (18), (22), (23) contradicts (17). Hence the assumptions made at the start of this argument (including (11)) imply that (12) is false. This proves the inadmissibility of any δ satisfying (4) under the additional assumptions (i)–(iv).

(We conjecture that a proof of inadmissibility without the special assumptions (i), (ii), (iv) can be constructed following the above general outline, although clearly such a proof will require a number of additional error bounds to replace (ii) and will be algebraically and analytically more complex if (i) ($k = 1$) is to be relaxed. In particular if $k \geqslant 2$ then the efficient argument at (14) which uses integration by parts (and all similar arguments above) may have to be replaced by arguments like those described in the treatment of error terms in Section 1.2. Also condition (iii) may have to be replaced by something more complex such as: define

$$v_\zeta(t) = \int_{\zeta\cdot x > t}\zeta^\tau \cdot \nabla W(x)p(x)\,dx$$

for $\|\zeta\| = 1$ and assume $v_\zeta(t) > 0$ whenever $0 < \int_{\zeta\cdot x > t}p(x)\,dx < 1$.)

Admissibility. Suppose γ is a generalized Bayes procedure of the form described above (4) and $b + k - 2 \leqslant 0$. Let

$$\begin{aligned} \zeta_r(\xi) &= 1 \qquad \|\xi\| \leqslant 1 \\ &= \left(1 - \frac{\ln\|\xi\|}{\ln r}\right) \qquad 1 < \|\xi\| < r \\ &= 0 \qquad \|\xi\| \geqslant r \end{aligned} \tag{24}$$

and

$$h_i(\xi) = \left(i^{-1}\textstyle\int_2^\infty \zeta_r(\xi) e^{(2-r)/i}\, dr\right)^2. \tag{25}$$

Equation 1.3(10) for the approximate difference of (generalized) Bayes estimators is

$$\tilde{\gamma}_g(x) - \tilde{\gamma}_{hg}(x) = -N^{-1}M^\tau \nabla h(x)/h(x). \tag{26}$$

In particular when $M = 2\sigma^2 I$ and $N = 2I$ the above becomes $\tilde{\gamma}_g - \tilde{\gamma}_{hg} = -\sigma^2 \nabla h/h$. Substitution in the multivariate analog of 1.4(4) and direct computation, including a change to polar coordinates, yields (with $s \vee 2 = \max(s, 2)$)

$$\begin{aligned} \tilde{\Lambda}_i &\leqslant B \text{ max-eig}(MN^{-2}M^\tau)\textstyle\int_1^\infty\left(\int_{s\vee 2}^\infty \frac{i^{-1}e^{(2-r)/i}}{s \ln r}\, dr\right)^2 (cs^b)s^{k-1}\, ds \\ &\leqslant B'\textstyle\int_1^\infty \int_{s\vee 2}^\infty \frac{i^{-1}e^{(2-r)/i}}{\ln^2 r} s^{b+k-3}\, dr\, ds \\ &= B'\textstyle\int_2^\infty \frac{i^{-1}e^{(2-r)/i}}{\ln^2 r} \int_1^r s^{b+k-3}\, ds\, dr \\ &\leqslant B'\textstyle\int_2^\infty \frac{i^{-1}e^{(2-r)/i}}{\ln r}\, dr \to 0 \qquad \text{as} \quad i \to \infty. \end{aligned} \tag{27}$$

Note how the condition $b + k - 3 \leqslant -1$ was used in the last inequality. This indicates the result that the estimator described by γ_g is admissible.

(We note that a formal computation of $\tilde{\Lambda}_i$ based on the functions $h_i = \zeta_i^2$ would also yield the result $\tilde{\Lambda}_i \to 0$. However the functions ζ_i^2 are not sufficiently smooth at $\|x\| = i$, and so we use the smoothed version, h_i, as defined in (25).)

The above choice of h_i—while satisfactory for all values $b + k - 2 \leqslant 0$— is most appropriate when $b + k - 2 = 0$. Also, the spherically symmetric form of h is most suited to the case $M = 2\sigma^2 I$, $n_i = 2I$. In situations other than these a different choice of h than that above will generally yield faster convergence of $\tilde{\Lambda}_i$ to zero. (It is possible to construct the best functions ζ_r as solutions of an appropriate variational equation. For the location problem at hand, one may construct this variational equation by proceeding by analogy with the construction in Brown (1971) where the equation is given for a normal problem. In this way one obtains the connection between admissibility and global solubility of the variational differential equation. Admissibility is, therefore, also related to recurrence of the diffusion whose generator is the above mentioned variational equation.)

2.3. *Estimating several Poisson means* (*without nuisance parameters*). (This section was added to the paper only after we read the manuscript by Clevenson and Zidek (1974) which gives minimax and inadmissibility results for this problem.) Suppose $z_1, \cdots, z_k$ are observations of independent Poisson variables with parameters $\alpha_1, \cdots, \alpha_k$. It is desired to estimate $\alpha_1, \cdots, \alpha_k$ by $a_1, \cdots, a_k$ with loss function $\mathcal{L}(\alpha, a) = \Sigma_{i=1}^k(\alpha_i - a_i)^2/\alpha_i$.

In the following we make the conjecture that for generalized priors of the form $g(\alpha) \sim (\Sigma\alpha_i)^b$ as $\Sigma\alpha_i \to \infty$ the generalized Bayes estimator for g is inadmissible if $b + k - 1 > 0$ and admissible if $b + k - 1 \leqslant 0$. The conjecture also describes the asymptotic form of the generalized Bayes estimator (see (2) and (7).) The difference between this conjecture and that for the location problem, where admissibility depended on $\operatorname{sgn}(b + k - 2)$, may be explained by an examination of the relevant differential equations.

It is not appropriate to identify z as the x of the general theory. This is because $\operatorname{Var}_{\alpha_i}(z_i) = \alpha_i \to \infty$ as $\alpha_i \to \infty$ which means that the basic assumptions of Section 1 are violated. In order to apply our general results one must first transform the problem, as described below.

Let $x_i^2 = z_i$, $\xi_i^2 = \alpha_i$, $d_i^2 = a_i$, $i = 1, \cdots, k$, and $L(\xi, d) = \Sigma_{i=1}^k(\xi_i^2 - d_i^2)^2/\xi_i^2$. This transformation was chosen because it is asymptotically variance stabilizing. This transformed problem is isomorphic to the original one, and so a procedure is admissible here if and only if the corresponding procedure is admissible in the original problem. We are particularly interested in the usual minimax estimator, $\hat{a} = z$. This estimator is generalized Bayes in the original problem with respect to the uniform prior on the parameter space $\{\alpha : \alpha_i > 0, i = 1, \cdots, k\}$. This estimator, of course, corresponds to $\delta(x) = x$ in the transformed problem, which is generalized Bayes with respect to the prior, having density $g(\xi) = \Pi_{i=1}^k\xi_i$ over the new parameter space, $\{\xi : \xi_i > 0\}$.

The appropriate version of 1.2(6) is

$$(1) \qquad \Delta_{\text{est}} = -\Sigma_{i=1}^k\big(m_1^{(i)} + \psi_i\big)\lambda_i - \Sigma_{i=0}^k\Sigma_{j=0}^k m_2^{(i,j)}\frac{\partial}{\partial x_j}\lambda_i - \left(\tfrac{1}{2}\right)\Sigma n_i\lambda_i^2 \geqslant 0$$

where

$$m_1^{(i)}(\xi) = m_1^{(i)}(\xi_i) = E_\xi\left(\frac{\partial}{\partial d_i}L(\xi, d)\bigg|_{d=x}\right) \geqslant 4/(1 + 2\xi_i)$$

$$m_2^{(i)}(\xi) = m_2^{(i,i)}(\xi) = E_\xi\left((x_i - \xi_i)\frac{\partial}{\partial d_i}L(\xi, d)\bigg|_{d=x}\right) = 2 + 0(1/(1 + \xi_i))$$

$$m_2^{(i,j)}(\xi) = E_\xi\left((x_j - \xi_j)\frac{\partial}{\partial d_i}L(\xi, d)\bigg|_{d=x}\right) = 0\big((1 + \xi_i)^{-1}(1 + \xi_j)^{-1}\big), i \neq j$$

$$n_i = E_\xi\left(\frac{\partial^2}{\partial d_i^2}L(\xi, d)\bigg|_{d=x}\right) = 8.$$

Also, $m_2^{(i)}$ is decreasing in ξ_i. (As an example of the above calculations observe $E_\alpha(z^{\frac{3}{2}}) = \Sigma_{z=1}^{\infty} z^{\frac{3}{2}}\alpha^z e^{-\alpha}/z! = \Sigma_{z=0}^{\infty}\alpha(z+1)^{\frac{1}{2}}\alpha^z e^{-\alpha}/z! = E_\alpha(\alpha(z+1)^{\frac{1}{2}})$. Hence $m_1^{(i)}(t) = 4E_{t^2}(z^{\frac{1}{2}}(z - t^2)/t^2) = 4E_{t^2}((z+1)^{\frac{1}{2}} - z^{\frac{1}{2}}) \geqslant 4[(t^2+1)^{\frac{1}{2}} - t] > 4/(1 + 2t)$, by Jensen's inequality.)

Inadmissibility results. Suppose that $\gamma(x)$ satisfies

$$\gamma_i(x) = \frac{2\beta x_i}{8\|x\|^2} + 0(\|x\|^{-2}), \text{ and } \gamma_i(x) = 0 \text{ if } x_i = 0, \tag{2}$$

$$\text{and } \frac{\partial \gamma_i}{\partial x_j} = 0(\|x\|^{-2}).$$

This form of γ is motivated following (7).

Suppose $\beta/2 + k - 1 > 0$. Let $0 < c < k - 1 + \beta/2$, $\theta = (1, 1, \cdots, 1)^\tau$ and define $\lambda = (\lambda_1, \cdots, \lambda_k)^\tau$ by

$$(\lambda(x))_i = -\frac{c(x_i+1)}{\|x+\theta\|^2} \qquad x_i > 0 \tag{3}$$
$$= 0 \qquad\qquad x_i = 0.$$

Then

$$\Delta_{\text{est}} = \frac{c}{\|\xi+\theta\|^2}\left(\Sigma\left(\frac{4(\xi_i+1)}{1+2\xi_i} + \psi_i(\xi)(\xi_i+1)\right)\right. \tag{4}$$
$$+\Sigma m_2^{(i)}(\xi)\left(1 - \frac{2(\xi_i+1)^2}{\|\xi+\theta\|^2}\right)$$
$$\left. + 0(\|\xi+\theta\|^{-1}) - 4c\right)$$
$$\geqslant \frac{4c}{\|\xi+\theta\|^2}(k + \beta/2 - 1 - c) + 0(\|\xi\|^{-3})$$

since $4(\xi_i + 1)/(1 + 2\xi_i) \geqslant 2$, $\psi_i(\xi) = 8\gamma_i(\xi) + 0(\|\xi\|^{-2})$ and $m_2^{(i)} \searrow 2$ with $(1 - (2(\xi_i + 1)^2/\|\xi + \theta\|^2))$ also decreasing as $\xi_i \nearrow \infty$. Hence $\Delta_{\text{est}} > 0$ for $\|\xi\|$ sufficiently large.

We think that a detailed treatment of error terms will show that $\Delta \geqslant \Delta_{\text{est}} + 0(\|\xi\|^{-3})$ so that $\Delta > 0$ for $\|\xi\|$ sufficiently large, and in fact

$$\liminf_{\|\xi\|\to\infty} H(\xi)\Delta(\xi) = k + \beta/2 - 1 - c > 0 \tag{5}$$

for $H(\xi) = 1 + \|\xi\|^2$. (Over most of its domain $\Delta = \Delta_{\text{est}} + 0(\|\xi\|^{-3})$. However, when $\min \xi_i$ is not large, then Δ is less well approximated by Δ_{est}. This is due to the fact that $(\lambda(x))_i$ is defined to be zero if $x_i = 0$ (so that $x_i + \gamma_i(x) = 0 = x_i + \gamma_i(x) + \lambda_i(x)$) instead of $-c/\|x + \theta\|^2$. However, it is easy to see that truncating λ as we have done can only increase Δ, and so leads to possible strict inequality in the statement $\Delta \geqslant \Delta_{\text{est}} + 0(\|\xi\|^{-3})$. (On the other hand the truncation of λ is necessary in order to guarantee the accuracy of the approximation as $\xi_1 \to 0$, say, for fixed $\xi_2, \cdots, \xi_k$ (large).))

Generalized Bayes estimators. Let $g(\xi)$ be a generalized prior density in the transformed problem. Suppose $|(\partial/\partial\xi_i)g|/g = O(\xi_i^{-1})$ and higher derivatives of g are similarly well behaved. The multidimensional version of 1.3(5) is

$$(\gamma_g)_i \approx -m_1^{(i)}/n + \left(\frac{\partial}{\partial\xi_i} m_2^{(i)} g\bigg|_{\xi=x}\right)/ng = (\tilde{\gamma}_g)_i.$$

Computations involving the Poisson distribution show that $(\partial m_2^{(i)}/\partial\xi_i) = O(1/(1 + \xi_i^2))$. Hence

$$(6) \qquad (\gamma_g(x))_i \approx (\tilde{\gamma}_g(x))_i = -1/(2 + 4x_i) + \left(\frac{\partial g}{\partial\xi_i}\bigg|_{\xi=x}\right)/4g(x) + O(x_i^{-2}).$$

Note that the uniform generalized prior in the original problem leads to the minimax estimator $(a(x))_i = x_i$ (direct computation). This prior corresponds to the prior $g_0(\xi) = \Pi_{i=1}^k \xi_i$ in the transformed problem. Hence the generalized Bayes estimator for this prior is $\delta_{g_0}(x) = x$, or $\gamma_{g_0}(x) \equiv 0$. The formula (6) yields $(\tilde{\gamma}_{g_0}(x))_i = (4x_i(1 + 2x_i))^{-1} + O(x_i^{-2})$. Hence there is an error in the approximation $\gamma_{g_0} \approx \tilde{\gamma}_{g_0}$, and the error is $O(1/x_i^2) = O(1/x_i)$.

For a prior which behaves like g_0 the approximation (6) is therefore satisfactory $(O(\|x\|^{-1}))$ only on the region where $\min x_i$ is large.

A more satisfactory approximation to γ_g may be obtained by using 1.3(10) instead of 1.3(5). Suppose $g = hg_0$ where h is smooth and $\|\nabla h\|/h = O(\|x\|^{-1})$. Then from the appropriate multivariate version of 1.3(10) and the fact that $\gamma_{g_0} \equiv 0$,

$$(7) \qquad (\gamma_{hg_0}(x))_i = m_2^{(i)}(x)\left(\frac{\partial}{\partial\xi_i} h\bigg|_{\xi=x}\right)/8h(x) + O(\|x\|^{-2}).$$

(In both (6) and (7) we have neglected to include certain terms which involve $m_2^{(i,j)}$ for $i \neq j$. These terms do not affect the qualitative conclusions made above, and in fact can be absorbed in the $O(\|x\|^{-2})$ terms of (7) if $h(\xi) \sim \|\xi\|^\beta$.)

Suppose that $h(\xi) \sim \|\xi\|^\beta$ and that the derivatives of h behave asymptotically to those of $\|\xi\|^\beta$. Then by (7) $(\gamma_{hg_0}(x))_i = (m_2^{(i)}(x)/8)\beta x_i/\|x\|^2 + O(\|x\|^{-2})$. It can be checked that $m_2^{(i)}(x)x_i - 2x_i = O(1)$. Hence this is an estimator of the form (2). (A more direct calculation shows that for $h(\xi) \sim \|\xi\|^\beta$ $(\gamma_{hg_0}(x))_i \sim 4\beta x_i/\|x\|^2$ uniformly as $\|x\|^2 \to \infty$. Hence (2) is actually a closer approximation than (7).)

How to prove inadmissibility. The expansion (7) indicates that if γ is generalized Bayes and satisfies (2) then the prior (density) is of the form $g = hg_0$ with $h(\xi) \sim \|\xi\|^\beta$. If so, then

$$\int H^{-1}(\xi)g(\xi)\, d\xi = \int (1 + \|\xi\|^2)^{-1}\|\xi\|^\beta \Pi\xi_i\, d\xi_i = \infty$$

since $\beta/2 > 1 - k$. This indicates that $\delta = x + \gamma$ is inadmissible, and that the inadmissibility can be proved by Method II.

The usual minimax estimator $\delta(x) = x$ is therefore inadmissible for dimensions $k \geqslant 2$. This situation differs from the case of location parameters where inadmissibility first appears in dimension 3. However, this apparent difference may be explained by the fact that the differential inequality for the Poisson problem in k dimensions is qualitatively similar to the differential inequality 2.2(1) for the location problem in $2k$ dimensions. This similarity can best be seen by comparing 2.2(5) to (4), above. It can thus be said that the Poisson inadmissibility in dimension 2 follows for the same cause as the location inadmissibility in dimension 4. The similarity of the differential equations indicates that the minimax estimator here should be admissible when $k = 1$, a result which we discuss later.

We also note that the generalized prior densities $g = hg_0 \sim \|\xi\|^{\beta} g_0$ in this problem correspond to generalized prior densities of the asymptotic form $(\Sigma\alpha_i)^{\beta/2}$ in the original Poisson problem. The densities $m_\beta(\Lambda)$ and $m_\beta^0(\Lambda)$ of Clevenson and Zidek (1974) with $\beta' = -(\beta/2 + k - 1)$ are densities of this form. Hence the generalized priors $m_{\beta'}$ and $m_{\beta'}^0$ of Clevenson and Zidek (1974) apparently lead to inadmissible procedures if $\beta' < 0$.

Actually, a proof of the indicated inadmissibility would involve a complication not mentioned above. Since $\Omega = \{\xi : \xi_i > 0\}$ is not closed the expression (5) does not suffice to establish the truth of the basic hypothesis 1.2(7) of the theorem for Method II. It is also necessary to examine $H(\xi)\Delta(\xi)$ as one (or more) of the coordinates, ξ_i, tends to zero. This can be done as follows:

Make the additional assumption that $\delta_i(x) = x_i + \gamma_i(x) > \varepsilon > 0$ whenever $x_i > 0$, $1 \leqslant i \leqslant k$. Modify the definition of λ by replacing the expression $\|x + \theta\|^2$ in the definition (3) by $\sigma + \|x + \theta\|^2$, $\sigma > 0$. This modification does not effect the truth of (5). Also a direct computation shows that if $\xi_i' = 0$, then

$$\liminf_{\xi \to \xi'} E_\xi \left[\frac{\left((x_i + \gamma_i(x))^2 - \xi_i^2\right)^2}{\xi_i^2} - \frac{\left((x_i + \gamma_i(x) + \lambda_i(x))^2 - \xi_i^2\right)^2}{\xi_i^2} \right] \geqslant 4c\varepsilon^3 E_{\xi'}\left(\frac{1}{\sigma + \|x + \theta\|^2} \right).$$

On the other hand our heuristic analysis shows that if $\xi_j' > \varepsilon/2$ then

$$E_{\xi'}\left[\frac{\left((x_j + \gamma_j(x))^2 - \xi_j'^2\right)^2}{\xi_j'^2} - \frac{(x_j + \gamma_j(x) + \lambda_j(x))^2 - \xi_j'^2)}{\xi_j'^2} \right] = 0\left(\frac{c}{(\sigma + \|\xi' + \theta\|^2)^2} \right)$$

uniformly for $\xi_j' > \varepsilon/2$. As above, if $\xi_j' \leqslant \varepsilon/2$ then the left side of the preceding expression is positive for σ large enough. Hence if σ is chosen large enough then

$$\liminf_{\xi \to \xi'} H(\xi)\Delta(\xi) > 0 \tag{8}$$

whenever $\xi_i' = 0$ for some $i = 1, \cdots, k$. The basic hypothesis 1.2(7) is implied by (5) and (8) combined.

It is now necessary to check that if γ is generalized Bayes and satisfies (2) with $\beta/2 > 1 - k$ then $\int H^{-1}(\xi)G(d\xi) = \infty$. For this purpose it is easiest to return to the original form of the problem. Then (2) becomes $a(z) = z + \varphi(z)$ where

$$\varphi(z) = \frac{4\beta z}{8\Sigma z_i} + 0\big((\Sigma z_i)^{-\frac{1}{2}}\big) \tag{9}$$

with $\beta/2 > 1 - k$. Let J be any prior measure on the space of original parameters, $\{\alpha : \alpha_i > 0\}$. Suppose $a(\cdot)$ is generalized Bayes for J. It is necessary to show that

$$\int(1 + \Sigma\alpha_i)^{-1}J(d\alpha) = \infty. \tag{10}$$

Direct computation shows that

$$a_i(z) = \frac{\int\Pi(e^{-\alpha_j}\alpha_j^{z_j})J(d\alpha)}{\int\Pi\big(e^{-\alpha_j}\alpha_j^{(z_j-\delta_{ij})}\big)J(d\alpha)} \tag{11}$$

for $z_i \geq 1$, where δ_{ij} is the Kronecker δ function. It can be seen that the function $a(\cdot)$ uniquely determines the moments of the finite measure $K(d\alpha) = (\Pi e^{-\alpha_j})J(d\alpha)$. The growth of the moments of K is restricted by the combination of (9) and (11). We think this restriction should suffice to show that there is a 1-1 relationship between prior measures J and their generalized Bayes procedures satisfying (9). From this it follows that *if* $a(\cdot)$ (*satisfying* (9)) *is generalized Bayes for some J which satisfies* (10), *then* $a(\cdot)$ *is not admissible* (assuming the heuristic analysis leading to (5) and (8) is correct).

It seems as if one should be able to exploit (11) in order to show that for any generalized Bayes $a(\cdot)$ satisfying (9) the corresponding prior, J, satisfies (10). We have been able to do this in dimension $k = 1$. Here is the argument: there is some Z_0 such that $a(z) > z$ for $z \geq Z_0$. Let $c_0 = a(Z_0)/Z_0!$. Then, (11) implies that $\int\alpha^j K(d\alpha) \geq c_0 j!$ for $j \geq Z_0$ where $K(d\alpha) = e^{-\alpha}J(d\alpha)$. Now,

$$\begin{aligned} e\int(1 + \alpha)^{-1}J(d\alpha) &= \int e^{1+\alpha}(1 + \alpha)^{-1}K(d\alpha) \\ &= \int\Sigma_{j=0}^{\infty}\frac{(1 + \alpha)^{j-1}}{j!}K(d\alpha) \\ &\geq \Sigma_{j\geq Z_0}\int\frac{\alpha^{j-1}}{j!}K(d\alpha) \\ &\geq c_0\Sigma_{j\geq Z_0}\frac{1}{j} = \infty, \end{aligned}$$

which is the desired result.

Admissibility. Formula (7) and the algebraic computations in Section 2.2 make it easy to decide which of the above procedures should be admissible. Suppose $g = hg_0 \sim \|\xi\|^{\beta}g_0$, as above, with $\beta \leq 2 - 2k$. Let $h_i(\xi)$ be defined as in 2.2(25). Then $\tilde{\Lambda}_i \to 0$ as $i \to \infty$ as in Section 2.2. This indicates that the estimator described by γ_g is admissible. (Since the distribution of X is discrete, the expression for Λ_j is actually a sum over the possible values of X. Nevertheless this sum should be adequately bounded by a multiple of $\tilde{\Lambda}_j$, where

$$\tilde{\Lambda}_j = \int\int_{\{x : x_i > 0\}} \Sigma \left[\frac{m_2^{(i)}(x) \frac{\partial}{\partial \xi_i} h_j(x)}{8 h_j(x)} \right]^2 h_j(x) \|x\|^{\beta} \Pi x_i \, dx_i$$

with $h_j(\cdot)$ as in 2.2(25). It is particularly important to check the relation between the expressions for Λ_j and $\tilde{\Lambda}_j$ on the region when min x_i is small; and according to our calculations the relation is as desired so that Λ_j does seem to be uniformly bounded by a multiple of $\tilde{\Lambda}_j$.)

As noted, the priors $m_{\beta'}$ and $m_{\beta'}^0$ of Clevenson and Zidek (1974) correspond to priors $g = hg_0 \sim \|\xi\|^{\beta} g_0$ with $\beta \leqslant 2 - 2k$ if (and only if) $\beta' \geqslant 0$. Hence, the corresponding (generalized) Bayes procedures are apparently admissible in the original Poisson problem. For $m_{\beta'}^0$ these generalized Bayes procedures are

$$\delta_{\beta'}^0(z_1, \cdots, z_k) = \left(1 - \frac{\beta' + k - 1}{\Sigma_{i=1}^k z_i + \beta' + k - 1}\right) z, \qquad \beta' \geqslant 0.$$

Clevenson and Zidek show that these procedures are minimax for $0 \leqslant \beta' \leqslant k - 1$; so the above formula apparently yields a class of admissible minimax procedures when $k \geqslant 2$, each of which dominates the usual minimax procedure. (Clevenson and Zidek demonstrate the unconditional admissibility of $\delta_{\beta'}^0$ only for $\beta' > 1$, when $m_{\beta'}^0$ is a proper prior density.)

Other loss functions. The above methods can be applied to Poisson problems involving a variety of other loss functions. For example, it is relatively straightforward to generalize the above to apply to loss functions of the form $\mathcal{L}(\alpha, a) = \Sigma w(a_i / \alpha_i^{\frac{1}{2}} - \alpha_i^{\frac{1}{2}})$ where w is any appropriate function.

On the other hand, the loss function $\Sigma(\alpha_i - a_i)^2$ behaves differently. Consider, instead, the loss function $\Sigma(\alpha_i - a_i)^2/(\Sigma\alpha_i)$, which is equivalent for admissibility consideration (and has bounded risk). In the transformed problem,

$$\text{(12)} \quad \Delta_{\text{est}} = \Sigma\big(m_1^{(i)} + \psi_i\big)\xi_i^2\lambda_i/\|\xi\|^2 - \Sigma\Sigma\big(m_2^{(i,j)}\xi_i^2/\|\xi\|^2\big)\frac{\partial}{\partial x_j}\lambda_i - \Sigma n_i \xi_i^2 \lambda_i^2/\|\xi\|^2$$

where $m_1^{(i)}$, $m_2^{(i,j)}$, and $n^{(i)}$ are defined in (1). Again, a direct calculation yields that $\gamma_g(x) = 0$ when $g(\xi) = \|\xi\|^2$. The analog of (7) is

$$\text{(13)} \qquad \big(\gamma_{hg}(x)\big)_i = \big(m_2^{(i)}(x)x_i^2/8h(x)\|x\|^2\big)\left(\frac{\partial h}{\partial \xi_i}\bigg|_{\xi = x}\right) + 0\big(\|x\|^{-2}\big)$$

for "smooth" functions, h. Several interesting results for this problem appear in Peng (1976) and Hudson (1977).

Inadmissibility results. We consider here only the minimax procedure $\delta(x) = x$; so that $\gamma = 0 = \psi$ in the following. It can be seen that (12) is fundamentally different from (1) because of the additional factors, $\xi_i^2/\|\xi\|^2$, which appear in (12).

For, if λ is defined by (3) and $\xi_1 \to \infty$ for fixed (but large) $\xi_2, \cdots, \xi_k$ the formula (12) yields

$$\Delta_{\text{est}} \sim m_1^{(i)}\lambda_i - m_2^{(i)}\frac{\partial}{\partial x_i}\lambda_i - n_i\lambda_i^2$$
$$\sim \frac{c}{\|\xi\|^2}(2 + 2(-1) - c) < 0.$$

This indicates that an estimator of the form $\delta + \lambda$ can never beat δ for all ξ no matter what is the dimension k.

In order to demonstrate inadmissibility choose

$$(\lambda(x))_i = -\frac{c \ln x_i}{x_i \Sigma \ln^2 x_i} \qquad \min x_i > 0$$
$$= 0 \qquad \text{otherwise,}$$

for $c > 0$. This corresponds to an estimator in the original problem of the form $z + Q(z)$ with

$$(Q(z))_i \sim -\frac{4c \ln z_i}{\Sigma \ln^2 z_i} \qquad \text{as} \quad z_i \to \infty.$$

Then

$$\Delta_{\text{est}} = \frac{c}{\|\xi\|^2 \Sigma \ln^2 \xi_i} \Sigma\Bigg[\left(\frac{m_1^{(i)}\ln \xi_i}{\xi_i} - \frac{m_2^{(i)}\ln \xi_i}{\xi_i^2}\right)\xi_i^2$$
$$+ m_2^{(i)}\left(1 - \frac{2 \ln^2 \xi_i}{\Sigma \ln^2 \xi_i}\right) - \frac{nc \ln^2 \xi_i}{2\Sigma \ln^2 \xi_i} + 0(\xi_i^{-1})\Bigg]$$
$$= \frac{c}{\|\xi\|^2 \Sigma \ln^2 \xi_i}\left[2(k-2) - 4c + 0\big((\min \xi_i)^{-1}\big)\right].$$

Hence $\Delta_{\text{est}} > 0$ when $\min \xi_i$ is large whenever $0 < c < (k-2)/2$. The indicated result is that $\delta(x) = x$ *is inadmissible in dimension* $k \geq 3$. Of course a proof of this fact—or even a careful heuristic demonstration—requires considerably more of an argument. However, as Peng (1976) and Hudson (1977) have both proved this fact (by direct computations) using competitors $z + Q(z)$ with Q of the above form we will not proceed further here.

Admissibility. When $k = 1$ the two loss functions are identical, so that the estimator $\delta(x) = x$ is admissible here as before. The following argument indicates that now the estimator $\delta(x) = x$ is also admissible when $k = 2$. This estimator is generalized Bayes with respect to the prior density $g(\xi) = \|\xi\|^2/\xi_1\xi_2$ on $\{\xi : \xi_i > 0\}$.

Define

$$h_j(\xi) = \left(1 - \frac{\ln \Sigma(\ln^+ \xi_i)^2}{\ln j}\right)^2 \qquad 1 \leq \Sigma(\ln^+ \xi_i)^2 \leq j$$
$$1 \qquad \Sigma(\ln^+ \xi_i)^2 \leq 1$$
$$0 \qquad \Sigma(\ln^+ \xi_i)^2 > j$$

Then

$$\tilde{\Lambda}_j = \int\int \Sigma\left(\tilde{\gamma}_{h_j g}(x) - \tilde{\gamma}_g(x)\right)_i^2 n_i h_j(x) g(x)\, dx_1\, dx_2$$

$$\leqslant B \int\int_{\{x\,:\,x_i \geqslant 1,\ 1 \leqslant \Sigma \ln^2 x_i \leqslant j\}} \Sigma \left[\frac{\frac{\partial}{\partial x_i} h_j(x)}{h_j(x)} \right]^2 \frac{x_i^2}{\|x\|^2} h_j(x) \frac{\|x\|^2}{x_1 x_2}\, dx_1\, dx_2$$

$$= B' \int\int_{\{x\,:\,x_i \geqslant 1,\ 1 \leqslant \Sigma \ln^2 x_i \leqslant j\}} \left[\Sigma \frac{\ln^2 x_i}{\ln^2 j (\Sigma \ln^2 x_l)^2} \right] \frac{dx_1\, dx_2}{x_1 x_2}$$

$$= B'' \int_1^j \frac{1}{\rho^2 \ln^2 j} \rho\, d\rho \to 0 \qquad \text{as } j \to \infty.$$

As in Section 2.2 the bound of Λ_j by a multiple of $\tilde{\Lambda}_j$ may be easier to establish if a smoother-randomized-version of h_j is used in place of the above. See the parenthetic discussion following 2.2(27). Also, as noted before, Λ_j is actually a summation rather than an integral as in $\tilde{\Lambda}_j$. It is particularly important to check the adequacy of the bounds implied above when min x_i is small. This must be done (by someone) before the above assertion can be considered anything more than a conjecture.

2.4. *Estimating a location parameter with nuisance parameters: setup.* Let $X \times Z \in R^1 \times R^l$ with density $p(x - \xi, z - \zeta)$, $\xi, \zeta \in R^1 \times R^l$. Let L be location invariant $-L((\xi, \zeta), d) = W(d - \xi)$ as in Section 1.1. This problem has been treated in detail in Berger (1976b, c) following the methods outlined in our manuscript. His results cover a wide variety of possible situations, but do leave a few unsettled questions. We will not repeat his results here. The next two sections describe some other problems which involve nuisance parameters. See Portnoy (1975, 1976) for a different, but related, treatment of a special case of this problem.

2.5. *Estimating the largest mean: setting.* Suppose $Z_1, \cdots, Z_N$ are independent normal random variables with $E(Z_i) = \theta_i$, $\mathrm{Var}(Z_i) = \sigma_i^2$, $i = 1, \cdots, N$. It is desired to estimate $\xi = \theta^{(1)} = \max\{\theta_i : 1 \leqslant i \leqslant N\}$. For simplicity of presentation, take $\sigma_i^2 \equiv 1$ and $\mathcal{L}(\theta, a) = (\xi - a)^2$. There are several problems depending on what (if any) a priori knowledge there is concerning the ordering of the θ_i's. Treat first the most definite situation in which the entire ordering is assumed known. Take—without loss of generality—$\theta^{(1)} = \theta_1 \geqslant \theta_2 \geqslant \cdots \geqslant \theta_N$.

A reasonable procedure to propose using in this situation is the generalized Bayes procedure for the uniform prior on the parameter space. For simplicity, we will concentrate on the question of admissibility of this estimator, but note that the following considerations apply almost without change to a variety of other generalized Bayes estimators. Various aspects of this problem and related problems are considered in Cohen and Sackrowitz (1970) and in Blumenthal and Cohen

(1968a, b). In particular, Cohen and Sackrowitz (1970) prove minimaxity of the procedure defined above. The following considerations indicate that this procedure is admissible for any value of N.

Basic equations. To begin to reformulate this problem in the desired format let $x' = Z_1$, $\xi = \theta_1$, $y_i = Z_{i+1} - Z_1$, and $\eta_i = \theta_{i+1} - \theta_1$, $i = 1, \cdots, N-1$. (y and η are $(N-1)$ dimensional vectors.) Then $L((\xi, \eta), d) = (\xi - d)^2$.

The uniform prior on the original parameter space, $\{\theta\}$, corresponds to the uniform prior on the parameter space $\Omega = \{\xi, \eta : 0 \geqslant \eta_1 \geqslant \cdots \geqslant \eta_{N-1}\}$. For this development, let u' denote the density of this prior.

In the above formulation the moment coefficients in the basic equation are $m_{20} = 2E((x' - \xi)^2) = 2$, $n = 2$, and all other coefficients are zero. Thus the appropriate version of formula 1.3(5) yields the approximation $\tilde{\gamma}_{u'} = 0$. Actually $\gamma_{u'} = \Gamma(y)$ where Γ is a continuous function and $\Gamma(0, y_2, \cdots, y_{N-1}) > \varepsilon > 0$ for all $y_2, \cdots, y_{N-1}$. Hence the approximation is not a very good one for values x', y with y_1 near 0. In particular, the approximation is not uniformly good as $\|(x', y)\| \to \infty$, and it is not true that $\gamma_{u'} \to 0$ as $\|(x', y)\| \to \infty$. Hence the approximations to the risk in Section 1.2 and to those used to prove admissibility in Section 1.4 will not be accurate uniformly as $\|(x', y)\| \to \infty$. The methods of Sections 1.2 and 1.4 cannot be used with the form of the basic equation indicated above to prove admissibility or inadmissibility. (The poor accuracy of 1.3(5) here may be traced to the fact that u' is not continuous on $R^1 \times R^{N-1} = R^N$ at the boundaries of Ω.)

In order to use the methods of Sections 1.2 and 1.4 the above inaccuracy must be avoided. Let $x = x' + \gamma_{u'}(x', y_1, \cdots, y_{N-1})$. The transformation $x', y_1, \cdots, y_{N-1} \to x, y_1, \cdots, y_{N-1}$ is one-one (see Brown (1971, Lemma 3.1.2)), so the problem with observations (x, y) is equivalent to that with observations (x', y). In this new problem the generalized Bayes procedure for u, the transform of u', is of course given by $\gamma_u \equiv 0$.

The appropriate form of the basic equation is now

$$(1) \qquad \Delta_{\text{est}} = -m_{10}\lambda - m_{20}\frac{\partial\lambda}{\partial x} - \Sigma_{i=1}^{N-1} m_{11}^{(i)} \frac{\partial}{\partial y_i}\lambda - \left(\tfrac{1}{2}\right)\Sigma_{i,j=1}^{N-1} m_{12}^{(ij)} \frac{\partial^2}{\partial y_i \partial y_j}\lambda - n\lambda^2/2$$

where

$$\begin{aligned} m_{10} &= 0(e^{\eta_1}), \\ m_{20} &= 2 + 0(e^{\eta_1}), \\ m_{11}^{(i)} &= 0(e^{\eta_i}) \qquad & i = 1, \cdots, N-1, \\ m_{12}^{(ij)} &= 0(e^{\eta_i + \eta_j}) \qquad & i, j = 1, \cdots, N-1, \\ n &= 2. \end{aligned}$$

(The order terms above are conservative; smaller bounds are also valid. For example,

$$\Gamma(y_1, \cdots, y_{N-1}) = \frac{\int_S \xi \exp\left(-\left(\xi^2 + \Sigma(\xi - y_i + \eta_i)^2\right)/2\right) d\xi \Pi\, d\eta_i}{\int_S \exp\left(-\left(\xi^2 + \Sigma(\xi - y_i + \eta_i)^2\right)/2\right) d\xi \Pi\, d\eta_i}$$

$$= 0(e^{-\Sigma y_i^2/4}) = 0(e^{\Sigma y_i}) \qquad \text{as} \quad y_i \to -\infty$$

on S so that $m_{10} = E_{0,\eta}(x_1 + \Gamma(y_1, \cdots, y_{N-1})) = 0(e^{\Sigma \eta_i}) = 0(e^{\eta_1})$ as $\eta_1 \to -\infty$.)

Admissibility. The appropriate form of 1.3(10) is

$$(2) \qquad \gamma_{hu} - \gamma_u \approx \left(m_{20} \frac{\partial h}{\partial \xi} + \Sigma_{i=1}^{N-1} m_{11}^{(i)} \frac{\partial}{\partial \eta_i} h \right.$$

$$\left. - \Sigma_{i,j=1}^{N-1} \left(\frac{\partial}{\partial y_i} m_{12}^{(ij)} \right) \left(\frac{\partial}{\partial \eta_j} h \right) - \left(\tfrac{1}{2}\right) \Sigma_{i,j=1}^{N-1} m_{12}^{(ij)} \frac{\partial^2}{\partial \eta_i \partial \eta_j} h \right) / nh$$

where—in addition to the bounds in (1)—

$$\frac{\partial}{\partial y_i} m_{12}^{(ij)} = 0(e^{\eta_i + \eta_j}) \qquad i, j = 1, \cdots, N-1.$$

The essential features of (2) may be summarized as

$$(3) \qquad |\tilde{\gamma}_{hu} - \tilde{\gamma}_u| \leq B\left(a \frac{\partial}{\partial \xi} h + \Sigma_{i=1}^{N-1} b_i \frac{\partial}{\partial \eta_i} h + \Sigma_{i,j=1}^{N-1} b_i b_j \frac{\partial^2 h}{\partial \eta_i \partial \eta_j} \right) / nh$$

where

$$a = 2 + 0(e^{\eta_1})$$
$$b_i = 0(e^{\eta_i}).$$

Define ζ_R on Ω by

$$\zeta_R(\xi, \eta_1, \cdots, \eta_{N-1}) = \left(1 - \frac{|\xi|}{R^N}\right) + \Pi_{j=1}^{N-1}(1 - e^{-\eta_j}/e^R)^+ .$$

Then

$$(\tilde{\gamma}_{\zeta_R^2 u} - \tilde{\gamma}_u)^2 \zeta_R^2 \leq B'(R^{-2N} + (N-1)e^{-2R}$$
$$+ (N-1)^2 e^{-2R}) \chi_{\{|\xi| < R^N, 0 > \eta_j > -R\}}(\xi, \eta)$$

so that

$$\int_\Omega (\tilde{\gamma}_{\zeta_R^2 u} - \tilde{\gamma}_u)^2 \zeta_R^2 \to 0 \qquad \text{as} \quad R \to \infty.$$

The indicated result is that the uniform prior, u, yields an admissible generalized Bayes procedure. As elsewhere, a proof of this conjecture would require a verification that the relevant error terms are insignificant. For this purpose, it would probably be more suitable to proceed as indicated in Sections 2.2 and 2.3 and to use a smoother version of ζ_R^2 such as

$$h_i = \left(i^{-1} \int_1^\infty \zeta_R e^{-(R-1)/i}\, dR\right)^2.$$

The form of ζ_R was motivated by ignoring the second order terms in (3) and then solving a variational problem as in Brown (1971). In closing we note that there remain error terms here resulting from the truncated nature of Ω which have no analogy in the admissibility considerations in Sections 2.2 and 2.3. We believe that these error terms can be shown to be unimportant by methods similar to those used in Brown (1971) to handle truncated parameter spaces.

Parameter ordering unknown. An equally interesting variant of the preceding situation is the case where the parameter ordering is unknown. The variable Z, with $E(Z) = \theta$, is as before, but the parameter space, in terms of θ, is all of R^n. $\mathcal{L}(\theta, a) = (\theta^{(1)} - a)^2$, as before. A generalized prior for which one might want to check admissibility of the corresponding procedure is the uniform prior on all of R^n. Let δ^* denote this procedure. (This procedure is probably no longer minimax; at least it is not in the case $n = 2$, which is fully treated in Blumenthal and Cohen (1968a)).

Let $Z^{(1)} \geqslant Z^{(2)} \geqslant \cdots \geqslant Z^{(n)}$ denote the ordered values of $Z_1, \cdots, Z_n$. Define $x = \delta^*(z_1, \cdots, z_n)$, $\xi = \theta^{(1)}$, $y_i = z^{(i+1)} - z^{(1)}$, and $\eta_i = E(y_i)$ $i = 1, \cdots, n-1$. (This definition of η_i, rather than the more obvious $\eta_i = \theta^{(i+1)} - \theta^{(1)}$, seems to be necessary in order to make sufficiently small the error in 1.3(10).) We believe that the transformation $\theta^{(1)}, \cdots, \theta^{(n)} \to \xi, \eta_1, \cdots, \eta_{n-1}$ is one-one, so that the transformed problem is isomorphic to the original one. The uniform prior on the original problem transforms to a prior which is bounded and smooth on the set of possible parameter points.

To check admissibility of the procedure $\delta(x, y) = x$, which corresponds to the procedure δ^*, begin by writing down equation 1.3(10). The resulting equation is accurately summarized by (3) above. Hence, using the sequence of prior distributions defined in (4) yields the conjecture that $\delta(x, y) = x$ is admissible for this problem, and that δ^* is admissible for the original problem.

Similar admissibility results should be valid for any prior density in the original problem which is bounded and smooth on the parameter space, R^n. Of particular interest are priors of this type whose density depends only on the differences $\theta^{(1)} - \theta^{(j)}$ $j = 2, \cdots, n$ since these priors result in procedures which are invariant under the transformations $Z_1, \cdots, Z_n \to Z_{\pi(1)} + c, \cdots, Z_{\pi(n)} + c$ for any permutation, π. By the Hunt-Stein theorem (see e.g. Kiefer (1957)) at least one invariant procedure must be minimax and admissible. It is reasonable to suppose that there is a smooth, bounded, invariant generalized prior which yields this minimax and admissible procedure.

2.6 *Estimating a normal variance.* Suppose one observes $\bar{z} \sim N(\mu, \sigma^2/N)$ and $s^2 \sim \sigma^2\chi_p^2$ where μ, σ^2 are unknown. (In the usual applications $p = N - 1$.) It is desired to estimate σ^2. For now, assume the loss function is $\mathcal{L}(\sigma^2, a) = (\ln(a/\sigma^2))^2$. Other loss functions will be mentioned below. As background, we remind the reader that the best fully invariant estimator (which is $a = \alpha s^2$) is not admissible for this problem. See Stein (1964) and Brown (1968).

In order to apply our methods it does not suffice to identify $(s^2, \bar{Z})$ as (x, y) since, for example, $\mathrm{Var}_{\sigma^2}\, s^2 = 2\sigma^2$ as $\sigma^2 \to \infty$; and thus the basic condition 1.1(2) is violated.

Instead, let

$$\begin{aligned}
x &= \ln s^2 + \ln \alpha \text{ where } \alpha = \exp\{-E_{\sigma^2=1}(\ln s^2)\},\\
\xi &= \ln \sigma^2,\ d = \ln a,\ W(d-\xi) = (d-\xi)^2,\\
y &= lg(\bar{z}/\alpha^{\frac{1}{2}} s), \text{ and} \\
\eta &= lg(\mu/\sigma) \text{ where}\\
lgt &= (\operatorname{sgn} t)\ln(1+|t|).
\end{aligned} \tag{1}$$

It is easy to check that this new problem is isomorphic to the original one. In particular, admissibility of an estimator $\delta(x, y)$ in the new problem is equivalent to admissibility of the estimator $\exp \delta(x(s^2), y(s^2, \bar{z}))$ in the original problem.

Computations yield that

$$\begin{aligned}
&m_{10} \equiv 0,\ m_{20} \equiv \text{const.} > 0,\ n \equiv 2,\\
&m_{11}(\xi, \eta) = M_1(\eta) = -(\operatorname{sgn} \eta) m_{20}/2 + 0(e^{-|\eta|})\\
&\qquad (M_1 \text{ is an odd function, } M_1(\eta) < 0 \text{ for } \eta > 0),\\
&m_{12}(\xi, \eta) = M_2(\eta) = 0(1) \text{ where } M_2(\eta) = M_2(-\eta),\ M_2(0) < 0,\\
&\text{and also } M_i'(\eta) = 0(e^{-|t|}) \qquad i = 1, 2.
\end{aligned} \tag{2}$$

Inadmissibility results. The best invariant estimator $a = \alpha s^2$ corresponds to the estimator $\delta(x, y) = x$ in the transformed problem. One can check from the conditions (2) that there are smooth solutions to the ordinary differential inequality

$$\Delta_{\text{est}} = M_1(\eta)h'(\eta) - \left(\tfrac{1}{2}\right)M_2(\eta)h''(\eta) - h^2(\eta) \geqslant 0.$$

(The crucial facts about the M_i are that they are continuous, and $M_2(0) < 0$, and $\operatorname{sgn} M_1(\eta) = -\operatorname{sgn} \eta$.) It follows from the discussion in Section 1.2 that $\delta(x, y) = x$ can be improved on by an estimator of the form $\delta'(x, y) = x + h(y)$ with h as above. This indicates that $a = \alpha s^2$ should be inadmissible. (This result was proved in Brown (1968)). A more careful computation of M_1, M_2 could enable an explicit determination of an appropriate function, h. The above alternative estimator corresponds to a scale invariant estimator in the original problem.

Admissibility results. Let $g(\xi, \eta) = G(\eta)$ be a generalized Bayes prior density. Suppose $G = 0(1)$, $G'/G = 0(1)$ and $G''/G = 0(1)$. Suppose also that G'/G and $G''/G \to 0$ as $\eta \to \pm\infty$. (Note that the resulting generalized Bayes estimator corresponds to a scale invariant, but not location-scale invariant estimator in the original problem.)

From 1.3(10), and the discussion following it,

$$\tilde{\gamma}_g - \tilde{\gamma}_{hg} \leqslant K(|h_1| + |h_2| + |h_{22}|)/h.$$

As in Section 2.2 let $h_i = (1 - (\ln(\xi^2 + \eta^2)/\ln i^2))^2$ for $1 < \xi^2 + \eta^2 < i^2$, $= 1$ for $\xi^2 + \eta^2 \leqslant 1$, $= 0$ for $\xi^2 + \eta^2 \geqslant i^2$. Computation yields that $\tilde{\Lambda}_i \to 0$. Hence the indicated result is that the procedure δ_g is admissible for g as above. (Technical note: again, as in Section 2.2 it is more convenient to check negligence of error terms if an everywhere smooth and positive kernel such as $(i^{-1}\int_{v>2} h_v e^{-(v-2)/i}\, dv)^2$ is used in place of the h's described above.)

Brewster and Zidek (1974) studied the above problem. They found that the estimator $\delta(\bar{z}, s^2) = \varphi^{**}(\bar{z}/s)s^2$ where $\varphi^{**}(z) = \exp\{-E_{\mu=0,\sigma^2=1}(\ln s^2 \,|\, |\bar{z}/s| \leqslant z)\}$ is minimax and is generalized Bayes in the original problem. The generalized prior which leads to this procedure is given explicitly in their papers. When transformed to (ξ, η) coordinates it is of the type discussed above. Hence the Brewster-Zidek estimator should be an admissible minimax estimator.

The above conclusions concerning admissibility are based on only a few qualitative characteristics of the coefficients m, u in the basic equations. It is thus easy to see that they hold for a wide variety of loss functions. For example, if the original loss function is $\mathcal{L}(\sigma^2, a) = (a - \sigma^2)^2/\sigma^4$ then the corresponding loss function in the transformed problem is $W(d - \xi) = (e^{d-\xi} - 1)^2$. Just as above it is true that the generalized Bayes estimators for the above priors should be admissible. In particular, the Brewster-Zidek minimax estimator for this loss function is

$$\begin{aligned} \delta(s^2, \bar{z}) &= \varphi^{**}(\bar{z}/s)s^2 \text{ with } \varphi^{**}(t) \\ &= E_{\mu=0,\sigma^2=1}(s^2 \,|\, |\bar{z}/s| \leqslant z)/E_{\mu=0,\sigma^2=1}(s^4 \,|\, |\bar{z}/s| \leqslant t) \end{aligned}$$

and this should be an admissible estimator.

Other related problems. The independence of $\bar{z}$, s^2 for the normal distribution eases the computation of the moments in (2). But it does not appear to be essential. The types of generalized prior described above should yield admissible generalized Bayes estimators of the scale parameter for a variety of problems involving unknown location and scale parameters. For example, Zidek (1973) considers the estimation of the scale parameter of an exponential distribution with unknown location. Computations like those described above indicate that the generalized Bayes estimators $(T_{n,a,b})$ described by Zidek for this problem are admissible procedures.

Acknowledgments. I would like to thank the referee for his repeated, careful reading of the preliminary versions of this manuscript. And I would like to thank Jim Berger for his highly perceptive criticism of a variety of inadequate arguments contained in these versions. The arguments have been replaced by the material concerning Method II.

REFERENCES

BERGER, J. (1976a). Tail minimaxity in location vector problems and its applications. *Ann. Statist.* 4 33–50.

BERGER, J. (1976b). Inadmissibility results for generalized Bayes estimators of coordinates of a location vector. *Ann. Statist.* **4** 302–333.

BERGER, J. (1976c). Admissibility results for generalized Bayes estimators of coordinates of a location vector. *Ann. Statist.* **4** 334–356.

BERGER, J. (1976d). Inadmissibility results for the best invariant estimator of two coordinates of a location vector. *Ann. Statist.* **4** 1065–1076.

BERGER, J. and SRINIVASAN, C. (1977). Generalized Bayes estimators in multivariate problems. *Ann. Statist.* **6** 783–801.

BLUMENTHAL, S. and COHEN, A. (1968a). Estimation of the larger translation parameter. *Ann. Math. Statist.* **39** 502–516.

BLUMENTHAL, S. and COHEN, A. (1968b). Estimation of the larger of two normal means. *J. Amer. Statist. Assoc.* **63** 861–876.

BREWSTER, J. and ZIDEK, J. (1974). Improving on equivariant estimators. *Ann. Statist.* **2** 21–38.

BROWN, L. (1966). On the admissibility of estimators of one or more location parameters. *Ann. Math. Statist.* **37** 1087–1136.

BROWN, L. (1968). Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters. *Ann. Math. Statist.* **39** 29–48.

BROWN, L. (1971). Admissible estimators, recurrent diffusions, and insoluble boundary value problems. *Ann. Math. Statist.* **42** 855–904.

BROWN, L. (1975). Estimation with incompletely specified loss functions. *J. Amer. Statist. Assoc.* **70** 417–427.

BROWN, L. (1980). A minimax theorem and a necessary condition for admissibility. To appear in *Ann. Statist.*

CLEVENSON, M. and ZIDEK, J. (1977). Simultaneous estimation of the mean of independent Poisson laws. *J. Amer. Statist. Assoc.* **70** 698–705.

COHEN, A. and SACKROWITZ, H. (1970). Estimation of the last mean of a monotone sequence. *Ann. Math. Statist.* **41** 2021–2034.

FARRELL, R. (1964). Estimators of a location parameter in the absolutely continuous case. *Ann. Math. Statist.* **35** 949–998.

FARRELL, R. (1968). Towards a theory of generalized Bayes tests. *Ann. Math. Statist.* **39** 1–22.

FUJITA, H. (1966). On the blowing up of solutions to the Cauchy problem for $w_t = \Delta w + u^{1+\alpha}$. *J. Fac. Sci. Univ. Tokyo* **13** 109–124.

HUDSON, H. (1977). A natural identity for exponential families with applications in multiparameter estimation. *Ann. Statist.* **6** 473–484.

KIEFER, J. (1957). Invariance, minimax sequential estimation, and continuous time processes. *Ann. Math. Statist.* **28** 573–601.

MEEDEN, G. and ISAACSON, D. (1977). Approximate behavior of the posterior distribution for a large observation. *Ann. Statist.* **5** 899–908.

PENG, J. C. M. (1975). Simultaneous estimation of the parameters of independent Poisson distributions. Technical Report 78, Depart. Statist., Stanford Univ.

PORTNOY, S. (1971). Formal Bayes estimation with application to the random effects model. *Ann. Math. Statist.* **42** 1379–1402.

PORTNOY, S. (1974). Admissibility of the best invariant estimator of one coordinate of a location vector. *Ann. Statist.* **3** 448–450.

PORTNOY, S. (1975). Transience and solvability of a non-linear diffusion equation. *Ann. Probability* **3** 465–477.

PORTNOY, S. (1976). On solutions to $w_t = \Delta w + w^2$ in two dimensions. *J. Math. Anal. Appl.* **55** 291–294.

STEIN, C. (1955). A necessary and sufficient condition for admissibility. *Ann. Math. Statist.* **26** 518–522.

STEIN, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proc. Third Berkeley Symp. Math. Statist. Probability* **1** 197–206.

STEIN, C. (1959). The admissibility of Pitman's estimator for a single location parameter. *Ann. Math. Statist.* **30** 970–979.

STEIN, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Ann. Inst. Statist. Math.* **16** 155–160.

STEIN, C. (1965). Approximation of improper prior measures by prior probability measures. In *Bernoulli, Bayes, Laplace Anniversary Volume*. Springer-Verlag, New York. 217–241.

UMBACH, D. (1976). Approximate behavior of the posterior distribution for an extreme observation in a multivariate setting. Technical report, Univ. Oklahoma.

ZIDEK, J. (1970). Sufficient conditions for the admissibility under squared error loss of formal Bayes estimators. *Ann. Math. Statist.* **41** 446–456.

ZIDEK, J. (1973). Estimating the scale parameter of the exponential distribution with unknown location. *Ann. Statist.* **1** 264–278.

MATHEMATICS DEPARTMENT
CORNELL UNIVERSITY
ITHACA, NEW YORK 14850