DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

ADMISSIBILITY IN FOUR BUT NOT IN FIVE DIMENSIONS OF THE DIFFUSE PRIOR BAYES RULE FOR THE CONTROL PROBLEM

Asad Zaman
California Institute of Technology
University of Pennsylvania

Charles Stein
Stanford University

SOCIAL SCIENCE WORKING PAPER 391

June 1981

ABSTRACT

It is shown that the Bayes decision rule corresponding to the prior of Lebesgue measure is admissible in dimension four but not in five for the control problem.

KEY WORDS

Admissibility, Control Problem, Four Dimensions, Five Dimensions, Diffuse Prior, Bayes Rules

ADMISSIBILITY IN FOUR BUT NOT IN FIVE DIMENSIONS OF THE DIFFUSE PRIOR BAYES RULE FOR THE CONTROL PROBLEM*

Asad Zaman
California Institute of Technology
University of Pennsylvania

Charles Stein
Stanford University

1. INTRODUCTION.

Suppose a random variable is generated by a regression relationship of the form $Z_t = \gamma' X_t + \varepsilon_t$, where ε_t is a sequence of independent and identically distributed scalar normal variates, γ is a p x 1 vector of unknown parameters, and X_t is a p x 1 vector of non-stochastic variables, presumed to be under direct control of the experimenter. The control problem to be discussed here arises if the experimenter wishes to choose values of the variables X_t so as to produce some desired value of Z_t, say $Z_t = Z^*$. It is assumed that the experimenter has an estimate $\hat{\gamma} \sim N(\gamma, \Omega)$ of γ, and is faced with a quadratic loss $(Z^* - Z_t)^2$.

As described in Zaman (1981), this problem can be reduced to the following canonical form. We observe $Y \sim N(\beta, I_p)$, and choose action $\delta \in R^p$ with the loss $L(\beta,\delta) = (\beta'\delta - 1)^2$. The formal Bayes rule against Lebesgue measure as a prior was obtained by Zellner to be $\delta^1(y)$, where, for $c \geq 0$, $\delta^c(y) = (c + |y|^2)^{-1}y$. Zellner (1978) showed that δ^1 is admissible for p = 1 by proving that the integral of the risk of δ^1 with respect to Lebesgue measure was finite in this case. Zaman (1981) proves that δ^1 is admissible for p = 2 and for p = 3. Kei Takeuchi (1979; personal communication) has shown that for $p \geq 6$, δ^1 is inadmissible and, in fact, that the risk $\rho(\beta,\delta^c)$ is an increasing function of c on [o,1] for $p \geq 6$. Here, for $\delta : R^p \rightarrow R^p$, $\rho(\beta,\delta) = E_\beta(\beta'\delta(y) - 1)^2$, which is the customary definition of the risk function.

In this paper we complete the study of the admissibility of δ^1 by showing it is admissible for p = 4 but not for p = 5, using direct methods.

Subsequent to the completion of this work, Berger and Zaman (1979) obtained a very general result characterizing inadmissible estimators, and Berliner (1980) obtained a general characterization of admissible estimators. These results use sophisticated indirect techniques, and include our results as a special case.

2. ADMISSIBILITY OF THE DIFFUSE PRIOR RULE IN FOUR DIMENSIONS.

In dimensions two and three, the admissibility of δ^1 is established by Zaman (1981) by picking a sequence of normal prior measures converging in an appropriate sense to the ''diffuse prior'' (Lebesgue measure), and showing that the excess average risk of δ^1 over the Bayes risk converges sufficiently rapidly to 0. For four dimensions we use essentially the same method, with a more complicated sequence of priors.

---

* We are indebted to Professor Kei Takeuchi for his encouragement and for some helpful conversations. This work was supported in part by NSF Grant SES79-13976 at the Institute for Mathematical Studies in the Social Sciences, Stanford University.

<u>Theorem 1</u>. For p = 4, the decision procedure $\delta^1(y) = (1 + |y|^2)^{-1}y$ is admissible.

<u>Proof</u>. Consider the family of (proper) prior densities, indexed by $\sigma > 100$, defined by

$$\pi_\sigma(\beta) = c_\sigma\ \sigma^{-4+2\varepsilon}\ |\beta|^{-2\varepsilon}\ \exp\ (-|\beta|^2/2\sigma^2) \tag{2.1}$$

where $\varepsilon = \varepsilon_\sigma = (\log\ \sigma)^{-1/2}$, and $c_\sigma = 2^{\varepsilon-1}\ (2 - \varepsilon)$, which is the appropriate normalizing constant. Let δ_σ be the Bayes rule associated with π_σ. For an arbitrary decision rule $\delta: R^4 \rightarrow R^4$, and an arbitrary prior probability density π, let $\rho^*(\pi,\delta) = \int \rho(\beta,\delta)\pi(\beta)d\beta$. The proof is completed by the following lemmas.

<u>Lemma 1</u>. In order to prove the admissibility of δ^1, it suffices to show that

$$\lim_{\sigma\to\infty}\ \sigma^{4-2\varepsilon}[\rho^*(\pi_\sigma,\delta^1) - \rho^*(\pi_\sigma,\delta_\sigma)] = 0. \tag{2.2}$$

<u>Proof</u>. Assuming (2.2), suppose contrary to the assertion of the lemma, there exists a $\tilde{\delta}$ such that, for all $\beta \in R^4$, $\rho(\beta,\tilde{\delta}) \leq \rho(\beta,\delta^1)$ with strict inequality at $\beta = \beta_0$. Then, because of continuity of the risk functions in this problem, there exists a positive number a, such that for all β for which $|\beta - \beta_0| \leq a$, we have $\rho(\beta,\tilde{\delta}) \leq \rho(\beta,\delta^1) - a$. Now observe that $0 \leq \sigma^{4-2\varepsilon}[\rho^*(\pi_\sigma,\tilde{\delta}) - \rho^*(\pi_\sigma,\delta_\sigma)] = \sigma^{4-2\varepsilon}[\rho^*(\pi_\sigma,\tilde{\delta}) - \rho^*(\pi_\sigma,\delta^1)] + \sigma^{4-2\varepsilon}[\rho^*(\pi_\sigma,\delta^1) - \rho^*(\pi_\sigma,\delta_\sigma)]$. Taking limits as $\sigma \to \infty$ and observing that the second term approaches 0, we obtain from (2.2) (all integrals are over R^4 unless otherwise specified, in this

section)

$$0 \leq \overline{\lim_{\sigma\to\infty}}\ c_\sigma \int [\rho(\beta,\tilde{\delta}) - \rho(\beta,\delta^1)]|\beta|^{-2\varepsilon}\ \exp\ (-|\beta|^2/2\sigma^2)d\beta$$

$$\leq \overline{\lim_{\sigma\to\infty}}\ -ac_\sigma \int_{\{\beta:|\beta-\beta_0|\leq a\}} |\beta|^{-2\varepsilon}\ \exp\ -(|\beta|^2/2\sigma^2)d\beta$$

$$\leq (-a)(\overline{\lim_{\sigma\to\infty}}\ c_\sigma) \int_{\{\beta:|\beta-\beta_0|\leq a\}} d\beta < 0$$

This contradiction proves the lemma.

We now obtain a more convenient form for the difference between the average risks of δ^1 and δ_σ under π_σ. First observe that, because of the orthogonal invariance of π_σ, we can express δ_σ in the form $\delta_\sigma(y) = \phi_\sigma(|y|)\ |y|^{-1}y$ for some scalar function ϕ_σ. See Lemma 1 of Zaman (1981). An explicit formula for ϕ_σ will be given below in Lemma 4. Let $\phi^1(|y|) = |y|(|y|^2 + 1)^{-1}$, so that $\delta^1(y) = \phi^1(|y|)\ |y|^{-1}y$.

<u>Lemma 2</u>.

$$\rho^*(\pi_\sigma,\delta^1) - \rho^*(\pi_\sigma,\delta_\sigma) \tag{2.3}$$

$$\leq \int (\phi^1(y) - \phi_\sigma(y))^2\ [\int |\beta|^2\ \exp\ (-\tfrac{1}{2}|\beta - y|^2)\pi_\sigma(\beta)d\beta]dy$$

<u>Proof</u>. This follows easily from Lemma 2 of Zaman (1981) by an application of Cauchy's inequality $(\beta'y)^2 \leq |\beta|^2|y|^2$).

<u>Lemma 3</u>. The inner integral on the right hand side of (2.3) satisfies

the inequality

$$\int |\beta|^2 \exp\left(-\tfrac{1}{2}|\beta - y|^2\right) \pi_\sigma(\beta)d\beta \tag{2.4}$$

$$\leq C\sigma^{2\varepsilon-4}(|y|^2 + 4)^{1-\varepsilon} \exp\left[-\tfrac{1}{2}|y|^2(\sigma^2 + 1)^{-1}\right],$$

for some numerical constant C.

Proof. Here, and subseqently, C will be a generic numerical constant, possibly different on subseqent appearances, whose exact value is not relevant to the calculations at hand.

Let $\lambda = \sigma^2(\sigma^2 + 1)^{-1}$. Writing out $\pi_\sigma(\beta)$ and completing squares in the resulting exponent in (2.4) leads to

$$\int |\beta|^2 \exp\left(-\tfrac{1}{2}|\beta - y|^2\right)\pi_\sigma(\beta)d\beta \tag{2.5}$$

$$= c_\sigma(2\pi\lambda)^2\sigma^{2\varepsilon-4} \exp\left[-\tfrac{1}{2}|y|^2(\sigma^2 + 1)^{-1}\right]$$

$$\times \int |\beta|^{2-2\varepsilon}(2\pi\lambda)^{-2} \exp\left(-\tfrac{1}{2\lambda}(\beta - \lambda|y|^2\right)dy$$

Let Z denote a four-dimensional normal random vector with mean λy and covariance matrix λI_4.

Then the integral on the right hand side of (2.5) can be written as

$$E|Z|^{2-2\varepsilon} = (\lambda^2|y|^2 + 4\lambda)^{1-\varepsilon} E[Z^2(\lambda^2|y|^2 + 4\lambda)^{-1}]^{1-\varepsilon}$$

$$\leq (\lambda^2|y|^2 + 4\lambda)^{1-\varepsilon}[EZ^2(\lambda^2|y|^2 + 4\lambda)^{-1}]^{1-\varepsilon}$$

$$= (\lambda^2|y|^2 + 4\lambda)^{1-\varepsilon}$$

$$\leq (|y|^2 + 4)^{1-\varepsilon}$$

This implies the lemma.

Next, we need a bound for $(\phi^1(|y|) - \phi_\sigma(|y|))^2$. First we obtain a convenient expression for ϕ_σ.

Lemma 4. Let γ and X be independent real random variables, $\gamma \sim N(\lambda|y|,\lambda)$ and $X \sim \chi^2_3$. Then

$$\phi_\sigma(|y|) = E\gamma(\gamma^2 + \lambda X)^{-\varepsilon} \,/\, E\gamma^2(\gamma^2 + \lambda X)^{-\varepsilon}.$$

Proof. From the proof of Lemma 1 of Zaman (1981) we obtain

$$\phi_\sigma(|y|) = \left[\int_{-\infty}^{\infty} \gamma e^{\gamma|y|}\, \tilde{\pi}_\sigma(\gamma)d\gamma\right]\left[\int_{-\infty}^{\infty} \gamma^2 e^{\gamma|y|}\, \tilde{\pi}_\sigma(\gamma)d\gamma\right]^{-1} \tag{2.6}$$

where $\tilde{\pi}_\sigma(\beta_1) = \int_{R^3} \exp\left(-\tfrac{1}{2}|\beta|^2\right) \pi_\sigma(\beta)d\beta_2 d\beta_3 d\beta_4$.

This can be computed to be

$$\tilde{\pi}_\sigma(\beta_1) = C\sigma^{2\varepsilon-4}(2\pi\lambda)^{3/2} \int_{R^3} (2\pi\lambda)^{-3/2}|\beta|^{-2\varepsilon} \exp\left(-\tfrac{1}{2\lambda}|\beta|^2\right)d\beta_2 d\beta_3 d\beta_4 \tag{2.7}$$

$$= C\sigma^{2\varepsilon-4}\,(2\pi\lambda)^{3/2} \exp -\beta_1^2/2\lambda^2\; E\,(\beta_1^2 + \lambda X)^{-\varepsilon}$$

where $X \sim \chi^2_3$. Substituting (2.7) into (2.6) gives the lemma.

It is convenient to introduce $\phi_\sigma^*(|y|) = E\gamma/E\gamma^2 = |y|(\lambda|y|^2 + 1)^{-1}$, and note that $(\phi^1 - \phi_\sigma)^2 \le 2(\phi^1 - \phi_\sigma^*)^2 + 2(\phi_\sigma^* - \phi_\sigma)^2$. The difference $\phi^1 - \phi_\sigma^*$ is easily calculated:

$$|\phi^1(|y|) - \phi_\sigma^*(|y|)| = (\sigma^2 + 1)^{-1}|y|^3(|y|^2 + 1)^{-1}(\lambda|y|^2 + 1)^{-1}.$$

It requires more work to bound $\phi_\sigma^* - \phi_\sigma$.

<u>Lemma 5</u>. There exists a positive constant C such that for all y

$$[\phi_\sigma^*(|y|) - \phi_\sigma(|y|)]^2 \le C\varepsilon^2(|y|^2 + 1)^{-3} \tag{2.8}$$

(Recall that $\varepsilon = (\log \sigma)^{-1/2}$).

<u>Proof</u>. We have

$$\begin{aligned}\phi_\sigma^*(|y|) - \phi_\sigma(|y|) &= [E\gamma^2(\gamma^2 + \lambda X)^{-\varepsilon}]^{-1} \\ &\quad \times E[(\gamma^2|y|(\lambda|y|^2 + 1)^{-1} - \gamma)(\gamma^2 + \lambda X)^{-\varepsilon}] \\ &= (\lambda^2|y|^2 + 4\lambda)^{-\varepsilon}[E\gamma^2(\gamma^2 + \lambda X)^{-\varepsilon}]^{-1} \\ &\quad \times E[(\gamma^2|y|(\lambda|y|^2 + 1)^{-1}-\gamma)((\gamma^2 + \lambda X)^{-\varepsilon}(\lambda^2|y|^2 + 4\lambda)^{\varepsilon}-1)].\end{aligned}$$

The second equality uses the fact that $E[\gamma^2|y|(\lambda|y|^2 + 1)^{-1}-\gamma] = 0$. It follows from Schwarz's inequality that

$$\begin{aligned}|\phi_\sigma^*(|y|) - \phi_\sigma(|y|)|^2 &\le E[\gamma^2|y|(\lambda|y|^2 + 1)^{-1}-\gamma]^2 \\ &\quad \times E[(\gamma^2 + \lambda X)^{-\varepsilon}(\lambda^2|y|^2 + 4\lambda)^{\varepsilon}-1]^2 \\ &\quad \times E[(\gamma^2(\gamma^2 + \lambda X)^{-\varepsilon}(\lambda^2|y|^2 + 4\lambda)^{\varepsilon})]^{-2}\end{aligned}$$

The first factor can be computed explicitly and is clearly bounded. We can find an upper bound for the last factor by applying Jensen's inequality conditionally on γ to obtain

$$E\gamma^2(\lambda^2|y|^2 + 4\lambda)^{\varepsilon}(\gamma^2 + \lambda X)^{-\varepsilon} \ge E\gamma^2(\lambda^2|y|^2 + 4\lambda)^{\varepsilon}(\gamma^2 + 3\lambda)^{-\varepsilon}. \tag{2.9}$$

For sufficiently small ε, the right hand side of (2.9) is convex in $|\gamma|$, so a second application of Jensen's inequality yields

$$\begin{aligned}&E\gamma^2(\lambda^2|y|^2 + 4\lambda)^{\varepsilon}(\gamma^2 + \lambda X)^{-\varepsilon} \\ &\quad \ge (E|\gamma|)^2((E|\gamma|)^2 + 3\gamma)^{-\varepsilon}(\lambda^2|y|^2 + 4\lambda)^{\varepsilon} \\ &\quad \ge C(|y|^2 + 1)\end{aligned}$$

Thus, in order to complete the proof of (2.8), we need only verify that there exists a constant C such that

$$M \equiv E[\gamma^2 + \lambda X)^{-\varepsilon}(\lambda^2|y|^2 + 4\lambda)^{\varepsilon}-1]^2 \le C\varepsilon^2(|y|^2 + 1)^{-1} \tag{2.10}$$

From the fact that for all real t (with a v b = min (a,b)), $|e^t - 1| \leq |t|(1 \vee e^t)$, it follows that

$$M \leq \varepsilon^2 E \log^2 [(\gamma^2 + \lambda X)(\lambda^2|y|^2 + 4\lambda)^{-1} \tag{2.11}$$

$$\times (1 \vee (\gamma^2 + \lambda X)^{-2\varepsilon}(\lambda^2|y|^2 + 4\lambda)^{2\varepsilon})]$$

Next, we use the fact that $(\gamma^2 + \lambda X)/\lambda$ is a non-central χ^2 with four degrees of freedom and noncentrality parameter $\lambda|y|^2$ so that, by the method of mixtures, which will also be used in Section 3, we can introduce a Poisson random variable K with mean $\frac{1}{2}\lambda|y|^2$ and take the conditional distribution of $(\gamma^2 + \lambda X)/\lambda$ given K to be that of a central χ^2_{4+2K}. Thus (2.11) yields

$$M \leq \varepsilon^2 E \log^2 \Big[\chi^2_{4+2K}(\lambda|y|^2 + 4)^{-1} (1 \vee (\lambda|y|^2 + 4)^{2\varepsilon}(\chi^2_{4+2K})^{-2\varepsilon})\Big] \tag{2.12}$$

$$\leq \varepsilon^2 E\{I_{\{K \leq \frac{1}{4}\lambda|y|^2 + 2\}} E^K \log^2 \Big[\chi^2_{4+2K} (\lambda|y|^2 + 4)^{-1}$$

$$\times (1 \vee (\lambda|y|^2 + 4)^{2\varepsilon}(\chi^2_{4+2K})^{-2\varepsilon}\Big]$$

$$+ 2I_{\{K > \frac{1}{4}\lambda|y|^2 + 2\}} E^K\Big[(\chi_{4+2K}(\lambda|y|^2 + 4)^{-1/2} - \chi^{-1}_{4+2K} (\lambda|y|^2 + 4)^{1/2})^2$$

$$\times (1 + \chi^{-1}_{4+2K} (\lambda|y|^2 + 4))\Big]$$

$$\equiv \varepsilon^2 E(I_{\{K \leq \frac{1}{4}\lambda|y|^2 + 2\}} E^K_1 + 2I_{\{K > \frac{1}{4}\lambda|y|^2 + 2\}} E^K_2)$$

Here I is an indicator function, and E^K denotes conditional expectation with respect to K. We have assumed $\varepsilon < \frac{1}{2}$ and used the fact that $|\log t| \leq 2|t^{1/2} - t^{-1/2}|$. For the first contribution to the expected value in (2.12) we have

$$EI_{\{K \leq \frac{1}{4}\lambda|y|^2 + 2\}} E^K_1$$

$$\leq P\{K \leq \tfrac{1}{4}\lambda|y|^2 + 2\} C[\log^2 (\lambda|y|^2 + 4)](\lambda|y|^2 + 4)$$

$$\leq C(|y|^2 + 1)^{-1}$$

In order to evaluate the second part of the expectation in (2.8), we observe that for any q > 4 and positive a,

$$E\Big[\chi_q a^{-1/2} - \chi^{-1}_q a^{1/2})^2(1 + \chi^{-2}_q a)\Big]$$

$$= (q/a) - 1 - (a/(q - 2)) + (a^2/((q - 2)(q - 4)))$$

$$= [(q + a)(q - a)^2 + (8 - 6q)(q - a) + 4a^2] [a(q - 2)(q - 4)]^{-1}$$

Applying this with $q = 4 + 2K$ and $a = \lambda|y|^2 + 4$, we obtain

$$EI_{\{K > \frac{1}{4}\lambda|y|^2+2\}} E^K_2$$

$$\leq CE(|y|^2 + 1)^{-3}[(2K + 8 + \lambda|y|^2)(2K - \lambda|y|^2)^2 + 4(\lambda|y|^2 + 4)^2]$$

$$\leq C(|y|^2 + 1)^{-1}$$

This completes the verification of (2.10) and thus also of the lemma.

By Lemma 1, in order to prove Theorem 1 we need only verify (2.2). By Lemmas 2,3, and 5, we have

$$\sigma^{4-2\varepsilon}[\rho^*(\pi_\sigma,\delta^1) - \rho^*(\pi_\sigma,\delta_\sigma)]$$

$$\leq \sigma^{4-2\varepsilon} \int (\phi^1(|y|) - \phi_\sigma(|y|))^2 \Big[\int |\beta|^2 \exp\,(-\tfrac{1}{2}|\beta - y|^2)\ \pi_\sigma(\beta)d\beta \Big] dy$$

$$\leq C \int \Big[\sigma^{-4}(|y|^2 + 1)^{-1} + \varepsilon^2(|y|^2 + 1)^{-3}\Big](|y|^2 + 4)^{1-\varepsilon} \exp\,(-\ |y|^2/2\sigma^2)\,dy$$

$$\leq C \int (\sigma^2|z|^2 + 1)^{-\varepsilon} \exp\,(-\tfrac{1}{2}|z|^2)dz + \varepsilon^2 \int (|y|^2 + 1)^{-2-\varepsilon}dy$$

$$\leq C \int_{|z|<\sigma^{-1/2}} \exp\,(-\tfrac{1}{2}|z|^2)dz + \int_{|z|>\sigma^{-1/2}} \sigma^{-\varepsilon} \exp\,(-\tfrac{1}{2}|z|^2)dz + \varepsilon$$

$$\leq C(\sigma^{-2} + \sigma^{-\varepsilon} + \varepsilon).$$

With $\varepsilon = (\log\sigma)^{-\frac{1}{2}}$ this approaches 0 as $\sigma \to \infty$.

3. INADMISSIBILITY FOR p = 5

In order to prove that the procedure δ, defined by (1.2) with $c = 1$ is inadmissible for $p \geq 5$, we first observe that the risk of any procedure δ of the form $\delta(y) = (|y|^2)y$ is given by $\rho(\beta,\delta)$ $= E(\ (|y|^2)\beta'y - 1)^2 = E[(\beta'y)^2\ ^2(|y|^2) - 2(\beta'y)\ (|y|^2) + 1]$. We

shall need the fact that

$$E(\beta'y)^2\psi^2(|y|^2) \tag{3.1}$$

$$= |\beta|^2 \exp\,(-\tfrac{1}{2}|\beta|^2) \sum_{k=0}^{\infty} (2^k k!)^{-1}|\beta|^{2k}(2k + 1)(2k + p)^{-1} E\ \chi^2_{2k+p}\ \psi^2(\chi^2_{2k+p}),$$

$$= |\beta|^2 \exp\,(-\tfrac{1}{2}|\beta|^2) \sum_{k=0}^{\infty} (2^k k!)^{-1}|\beta|^{2k}(2k + 1)\ E\ \psi^2(\chi^2_{2(k+1)+p})$$

and

$$E(\beta'y)\psi(|y|^2) \tag{3.2}$$

$$= |\beta|^2 \exp\,(-\tfrac{1}{2}|\beta|^2) \sum_{k=0}^{\infty} (2^k k!)^{-1}|\beta|^{2k}E\psi(\chi^2_{2(k+1)+p})$$

Proofs of (3.1) and (3.2) will be sketched later. Thus $\delta(y) = \psi(|y|^2)y$ is better than δ^1, which corresponds to $\psi(t) = (t + 1)^{-1}$, if for all $K \geq 0$

$$E\Big[(2K + 1)\psi^2(\chi^2_{2(K+1)+p})\ -2\psi(\chi^2_{2(K+1)+p})\Big] \tag{3.3}$$

$$\leq E\Big[(2K + 1)(\chi^2_{2(K+1)+p} + 1)^{-2}\ -2(\chi^2_{2(K+1)+p} + 1)^{-1}\Big].$$

Of course this condition is not necessary.

Let us verify inequality (3.3) with $\psi(t) = t^{-1}$, for $p = 5$. We shall need the fact that

$$E\chi_{q+2}^{-2}\,\psi(\chi_{q+2}^2) = q^{-1}E\psi(\chi_q^2). \qquad (3.4)$$

Then, writing χ^2 for $\chi^2_{2(K+1)+p}$ until it becomes necessary to vary the degrees of freedom, we obtain for the right hand side minus the left hand side

$$E\{[(2K + 1)(\chi^2 + 1)^{-2} - 2(\chi^2 + 1)^{-1}] - [(2K + 1)\chi^{-4} - 2\chi^{-2}]\} \qquad (3.5)$$

$$= E(\chi^2 + 1)^{-2}\chi^{-4}[-(2K + 1)(2\chi^2 + 1) + 2\chi^2(\chi^2 + 1)]$$

$$= E(\chi^2 + 1)^{-2}\chi^{-4}[2(\chi^4 + 2\chi^2 + 1) - 4(K + 1)(\chi^2 + (2K + 3)(4K + 4)^{-1})]$$

$$> E[2\chi^{-4} - 4(K + 1)(\chi^2 + 1)^{-1}\chi^{-4}]$$

$$= 2(2K + p)^{-1}(2(K - 1) + p)^{-1}$$

$$- 4(K + 1)(2K + p)^{-1}(2(K - 1) + p)^{-1}E(\chi^2_{2(K-1)+p} + 1)^{-1}$$

$$> 2(2K + p)^{-1}(2(K - 1) + p)^{-1}[1 - 2(K + 1)E\chi^{-2}_{2K+p-1}$$

$$= 2(2K + p)^{-1}(2(K - 1) + p)^{-1}(1 - 2(K + 1)(2K + p - 3)^{-1})$$

$$\geq 0 \text{ for } p \geq 5.$$

At the third equality sign we have used (3.4) twice, and at the second inequality sign we have used Jensen's inequality. The

function $\psi(t) = (t + 1)^{-1}$ is convex on R^t. Thus if a standard normal random variable V is independent of χ^2_q, we have

$$E(\chi_q^2 + 1)^{-1} < E(\chi^2 + U^2)^{-1} = E\chi^{-1}_{q+1} = (q - 1)^{-1}.$$

Here Jensen's inequality was applied conditionally given χ^2_q. This was used with $q = 2(k - 1) + p$. Thus for $p \geq 5$ we have verified (3.3) with $\psi(t) = (t + 1)^{-1}$, and this shows that for $p \geq 5$, $\delta^0(y) = |y|^{-1}y$ is better than $\delta^1(y) = (|y| + 1)^{-1}y$.

The identity (3.3) is obtained by a simple special case of the method of mixtures of Pitman and Robbins (1948). We write $|y|^2 = (\beta'y)^2|\beta|^{-2} + |y - (\beta'y)|\beta|^{-2}\beta|^2 = (U + |\beta|^2)^2 + \chi^2_{p-1}$, where U is a standard normal random variable independent of χ^2_{p-1}. In the present case the method of mixtures is applied by introducing a Poisson random variable K with mean $\frac{1}{2}|\beta|^2$ independent of χ^2_{p-1} and taking the conditional distribution of $(U + |\beta|^2)$ given K and χ^2_{p-1} to be that of χ^2_{2K+1}. Then we have

$$E(\beta'y)^2\psi^2(|y|^2)$$

$$= |\beta|^2E(U + |\beta|)^2\psi^2((U + |\beta|)^2 + \chi^2_{p-1})$$

$$= |\beta|^2E(\chi^2_{2K+1} + \chi^2_{p-1})^{-1}\chi^2_{2K+1}(\chi^2_{2K+1} + \chi^2_{p-1})\psi^2(\chi^2_{2K+1} + \chi^2_{p-1})$$

$$= |\beta|^2E\Big[(E^K(\chi^2_{2K+1} + \chi^2_{p-1})^{-1}\chi^2_{2K+1})(E^K\chi^2_{2K+p}\psi^2(\chi^2_{2K+p}))\Big]$$

$$= |\beta|^2 E \frac{2K+1}{2K+1} E^K \chi^2_{2K+p} \psi^2(\chi^2_{2K+p})$$

This is the first form of the identity (3.1). The second form is obtained by use of (3.4). The identity (3.2) is proved in Section 2 of James and Stein (1961).

REFERENCES

Berger, J.O. and A. Zaman. ''Inadmissibility in the Control Problem.'' CARESS Working Paper No. 80-03 . University of Pennsylvania, 1979.

Berliner, Mark L. ''Admissibility of Generalized Bayes Rules in the Control Problem.'' Technical Report #214. Department of Statistics, Ohio State University, 1980.

James, W. and C. Stein. ''Estimation with Quadratic Loss.'' In Proceedings of the Fourth Berkeley Symposium I, pp. 361-379, 1961.

Pitman, E.J.G. and H. Robbins. ''Application of the Method of Mixtures to Quadratic Forms in Normal Variates.'' Annals of Mathematical Statistics 20 (1949):552-560.

Takeuchi, K. ''On the Problem of Fixing the Level of Independent Variables in a Linear Regression Function.'' IMM 367. New York: New York University, Courant Institute of Mathematical Science, 1968.

Zaman, A. ''A Complete Class Theorem for the Control Problem and Further Results on Admissibility and Inadmissibility.'' Annals of Statistics, forthcoming, 1981.

Zellner, A. An Introduction to Bayesian Inference in Econometrics. New York: John Wiley and Sons, Inc., 1971.

__________. ''Estimation of Functions of Population Means and Regression Coefficients Including Structural Coefficients: A Minimum Expected Loss (MELO) Approach.'' Journal of Econometrics 8 (1978).