The Annals of Statistics
1979, Vol. 7, No. 5, 1144–1147

A FAMILY OF MINIMAX ESTIMATORS IN SOME MULTIPLE REGRESSION PROBLEMS

BY Y. TAKADA

University of Tsukuba

A family of estimators which dominate the maximum likelihood estimators of regression coefficients is given when the dependent variable and (3 or more) independent variables have a joint normal distribution.

1. Introduction. Let $X_1, \cdots, X_n$ be independently normally distributed $(p+1)$-dimensional random vectors with unknown mean θ and unknown nonsingular covariance matrix Σ.

The following partitions are used in the sequel:

$$X_i = \begin{pmatrix} Y_i \\ Z_i \end{pmatrix}, \qquad \theta = \begin{pmatrix} \eta \\ \xi \end{pmatrix}, \qquad i = 1, 2, \cdots, n \tag{1.1}$$

and

$$\Sigma = \begin{pmatrix} A & B' \\ B & \Gamma \end{pmatrix} \tag{1.2}$$

where Y_i, η and A are 1×1, z_i, ξ and B are $p \times 1$. Then it is well known that for $\alpha = \eta - \beta'\xi$ and $\beta = \Gamma^{-1}B$,

$$E(Y_i|Z_i) = \alpha + \beta' Z_i.$$

The problem in this paper is to estimate the regression coefficients (α, β) of Y_i on Z_i with respect to the loss function given by Stein [4]:

$$\begin{aligned} L((\theta, \Sigma); (\hat\alpha, \hat\beta)) \\ = \left[\{(\hat\alpha - \alpha) + (\hat\beta - \beta)'\xi\}^2 + (\hat\beta - \beta)'\Gamma(\hat\beta - \beta)\right] / (A - B'\Gamma^{-1}B). \end{aligned} \tag{1.3}$$

The maximum likelihood estimators $(\hat\alpha_M, \hat\beta_M)$ of (α, β) are given by

$$\hat\alpha_M = \bar{Y} - \hat\beta_M'\bar{Z}, \qquad \hat\beta_M = V^{-1}U \tag{1.4}$$

where

$$U = \Sigma_{i=1}^n Z_i Y_i - n\bar{Z}\bar{Y} \text{ and } V = \Sigma_{i=1}^n Z_i Z_i' - n\bar{Z}\bar{Z}'.$$

For this problem, Stein [4] first showed that the maximum likelihood estimators are minimax but inadmissible for $p \geqslant 3$. Baranchik [2] proves that each member of a family of specific estimators suggested by Stein [4] dominates the maximum likelihood estimators (1.4).

Received March 1978; revised June 1978.
AMS 1970 *subject classifications.* Primary 62C99; secondary 62F10, 62H99.
Key words and phrases. Minimax estimator, invariance structure, regression coefficients.

To facilitate the search for practical alternatives to the Stein rule described above, a family of minimax estimators containing Stein's is derived below.

2. A family of minimax estimators. Consider the following estimators:

$$\hat{\alpha} = \bar{Y} - \hat{\beta}'\bar{Z}, \qquad \hat{\beta} = f(R^2)\hat{\beta}_M. \tag{2.1}$$

Here $f(R^2)$ is any measurable function of the sample multiple correlation coefficient

$$R^2 = U'V^{-1}U/T, \tag{2.2}$$

where

$$T = \Sigma_{i=1}^n Y_i^2 - n\bar{Y}^2.$$

The invariant structure for this problem (see [2] and [4]) implies we may assume without loss of generality that

$$(\xi, \Gamma, A - B'\Gamma^{-1}B) = (0, I_p, 1) \quad \text{and} \quad \beta' = (\|\beta\|, 0, \cdots, 0). \tag{2.3}$$

The main result follows from the next lemma, the proof of which is similar to that of Lemmas 3 and 4 of [2].

LEMMA. *If $\phi(\cdot)$ is any measurable function on* $[0, \infty)$, *then under the condition* (2.3),

$$E\left[\phi\left(\frac{R^2}{1-R^2}\right)\hat{\beta}_M'\beta\right] = h(\|\beta\|, n)\Sigma_{k=0}^{\infty}\Gamma\frac{((n-1)/2+k-1)2k}{k!}r^k E\left[\phi\left(\frac{\chi_{p+2k}^2}{\chi_{n-p-1}^2}\right)\right], \tag{2.4}$$

and

$$E\left[\phi\left(\frac{R^2}{1-R^2}\right)\hat{\beta}_M'\hat{\beta}_M\right] = h(\|\beta\|, n)\Sigma_{k=0}^{\infty}\frac{\Gamma((n-1)/2+k-1)}{k!}r^k \times\left[\frac{n-3}{n-p-2} - r\frac{(n+2k-3)(p-1)}{(n-p-2)(p+2k)}\right]E\left[\phi\left(\frac{\chi_{p+2k}^2}{\chi_{n-p-1}^2}\right)\chi_{p+2k}^2\right], \tag{2.5}$$

where

$$h(\|\beta\|, n) = \left[2\Gamma((n-1)/2)(1+\|\beta\|^2)^{(n-1)/2-1}\right]^{-1}, r = \|\beta\|^2/(1+\|\beta\|^2),$$

and χ_{p+2k}^2 is a chi-squared random variable with $p + 2k$ degrees of freedom independent of χ_{n-p-1}^2.

THEOREM. *Relative to the loss function* (1.3) *the estimator*

$$\left(\hat{\alpha} = \bar{Y} - \hat{\beta}'\bar{Z}, \hat{\beta} = \left(1 - \tau\left(R^2(1-R^2)^{-1}\right)(1-R^2)R^{-2}\right)\hat{\beta}_M\right) \tag{2.6}$$

dominates the maximum likelihood estimator $(\hat{\alpha}_M, \hat{\beta}_M)$ *if*

(i) $n > p + 2$,

(ii) $\tau(\cdot)$ *is nondecreasing*,

(iii) $0 \leqslant \tau(\cdot) \leqslant 2(p-2)(n-p+1)^{-1}$.

PROOF. Lemma 2 of Baranchik [2] implies that the estimators (2.6) dominate the maximum likelihood estimator (1.4) under the loss function (1.3) if they do so under the loss function $\|\hat{\beta} - \beta\|^2$.

Let $g(R^2(1-R^2)^{-1}) = \tau(R^2(1-R^2)^{-1})(1-R^2)R^{-2}$. Then, from the above lemma,

$$(2.7)\quad \begin{aligned} &E\left[\|\hat{\beta}-\beta\|^2\right] - E\left[\|\hat{\beta}_M - \beta\|^2\right] \\ &= h(\|\beta\|, n)\Sigma_{k=0}^{\infty}\frac{\Gamma((n-1)/2+k-1)}{k!}r^k \\ &\times\Bigg[\left\{\frac{n-3}{n-p-2} - r\frac{(n+2k-3)(p-1)}{(n-p-2)(p+2k)}\right\} \\ &\times E\left\{\left[g^2\left[\frac{\chi^2_{p+2k}}{\chi^2_{n-p-1}}\right] - 2g\left[\frac{\chi^2_{p+2k}}{\chi^2_{n-p-1}}\right]\right]\chi^2_{p+2k}\right\} + 4kE\left\{g\left[\frac{\chi^2_{p+2k}}{\chi^2_{n-p-1}}\right]\right\}\Bigg]. \end{aligned}$$

Baranchik [1] showed in his proof of the theorem that the following inequality holds under assumptions (ii) and (iii):

$$(2.8)\quad E\left[\left[g^2\left[\frac{\chi^2_{p+2k}}{\chi^2_{n-p-1}}\right] - 2g\left[\frac{\chi^2_{p+2k}}{\chi^2_{n-p-1}}\right]\right]\chi^2_{p+2k} + 4kg\left[\frac{\chi^2_{p+2k}}{\chi^2_{n-p-1}}\right]\right] \leqslant 0.$$

Then the left hand side of (2.7) is bounded above by $h(\|\beta\|, n)$ times

$$(2.9)\quad \begin{aligned} &\Sigma_{k=0}^{\infty}\frac{\Gamma((n-1)/2+k-1)4k}{k!}r^k \\ &\left[1 - \left\{\frac{n-3}{n-p-2} - r\frac{(n+2k-3)(p-1)}{(n-p-2)(p+2k)}\right\}\right]E\left[g\left[\frac{\chi^2_{p+2k}}{\chi^2_{n-p-1}}\right]\right] \\ &= \frac{4(p-1)}{(n-p-2)}\Sigma_{k=0}^{\infty}\frac{\Gamma((n-1)/2+k-1)k}{k!}r^k\left[r\frac{(n+2k-3)}{p+2k} - 1\right] \\ &\times E\left[\tau\left[\frac{\chi^2_{p+2k}}{\chi^2_{n-p-1}}\right]\frac{\chi^2_{n-p-1}}{\chi^2_{p+2k}}\right]. \end{aligned}$$

This can be simplified by noticing that $E[\psi(\chi^2_m)] = mE[\psi(\chi^2_{m+2})/\chi^2_{m+2}]$ for any function $\psi(\cdot)$. The simplified version of equation (2.9) is

$$(2.10)\quad \Sigma_{k=0}^{\infty}t_k\left[r(n+2k-3) - (p+2k)\right],$$

where

$$t_k = \frac{4(p-1)}{(n-p-2)} \frac{\Gamma((n-1)/2+k-1)k}{k!\,(p+2k)(p+2k-2)} r^k E\left[\tau\left[\frac{\chi^2_{p+2k-2}}{\chi^2_{n-p-1}}\right]\chi^2_{n-p-1}\right].$$

The upper bound (2.10) is nonpositive as shown below.

As $t_0 = 0$, equation (2.10) can be expressed as

$$\Sigma_{k=1}^{\infty} t_{k-1} r(n+2k-5) - \Sigma_{k=1}^{\infty} t_k(p+2k). \tag{2.11}$$

Using the inequality

$$E\left[\tau(\chi^2_{p+2k-4}/\chi^2_{n-p-1})\chi^2_{n-p-1}\right] \leqslant E\left[\tau(\chi^2_{p+2k-2}/\chi^2_{n-p-1})\chi^2_{n-p-1}\right],$$

we get

$$t_{k-1} r \leqslant t_k 2(k-1)(p+2k)(p+2k-4)^{-1}(n+2k-5)^{-1},$$

which can be applied to the first term of (2.11), giving

$$\Sigma_{k=1}^{\infty} t_k 2(k-1)(p+2k)(p+2k-4)^{-1} - \Sigma_{k=1}^{\infty} t_k(p+2k)$$
$$= \Sigma_{k=1}^{\infty} t_k(p+2k)(p+2k-4)^{-1}(2-p).$$

Since $p \geqslant 3$, each term of the above infinite series is negative which completes the proof of the theorem.

EXAMPLE 1. Setting $\tau(\cdot)$ in the theorem equal to a constant c satisfying $0 < c \leqslant 2(p-2)(n-p+1)^{-1}$, we have the estimators obtained by Baranchik [2],

$$\hat{\beta}_c = \left(1 - c(1-R)^2 R^{-2}\right)\hat{\beta}_M.$$

EXAMPLE 2. Setting $\tau(R^2(1-R^2)^{-1}) = c/[1 + c(1-R^2)R^{-2}]$ for $0 < c \leqslant 2(p-2)(n-p+1)^{-1}$, we have a new family of estimators $\hat{\beta}$ given by

$$\hat{\beta} = \left[R^2/\left(R^2 + c(1-R^2)\right)\right]\hat{\beta}_M,$$

which contain Narula's estimate (Narula [3], page 17) as a special case, namely $c = p(n-p-2)^{-1}$.

Acknowledgment. The author wishes to thank Professor N. Sugiura for his suggestions. The author also wishes to thank the referee and associate editor for their suggestions.

REFERENCES

[1] BARANCHIK, A. J. (1970). A family of minimax estimators of the mean of a multivariate normal distribution. *Ann. Math. Statist.* **41** 642–645.

[2] BARANCHIK, A. J. (1973). Inadmissibility of maximum likelihood estimators in some multiple regression problems with three or more independent variables. *Ann. Statist.* **1** 312–321.

[3] NARULA, S. C. (1974). Predictive mean square error and stochastic regressor variables. *Appl. Statist.* **23** 11–17.

[4] STEIN, C. (1960). Multiple regression. In *Probability and Statistics*. Stanford Univ. Press.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TSUKUBA
SAKURA-MURA, NIIHARI-GUN
IBARAKI 300-31
JAPAN