

STEIN-JAMES ESTIMATORS OF A MULTIVARIATE LOCATION PARAMETER¹

BY SERGE L. WIND²

Columbia University

Bounds on the risks, under squared error loss, of a family of estimators of a multivariate location parameter are given for both fixed and random unknown location parameters when the covariance matrix of the observed random variable is unknown. The class of estimators considered in this paper contains Cogburn's [2].

1. Introduction and summary. Properties of the risks of a family of estimators of a multivariate location parameter are presented. These estimators are generalizations of those proposed by James and Stein [3] and Baranchik [1] for estimating the mean of a multivariate Normal distribution, but here no specific parametric distribution is assumed on either the underlying distribution or on any prior distribution. Cogburn [2] discussed mean squared error properties of the family, under the assumption that the covariance matrix $\sigma^2 I$ of the observed random variable was known. We extend his results in two directions: we assume σ^2 is not known (Corollary 2.1) and we assume this variance is a random variable with an unknown prior distribution (Corollary 2.4). Components of these "Stein-James estimators" may be employed as restricted asymptotically optimal solutions to estimation problems considered in an Empirical Bayes context [5].

2. Mean squared error properties of a family of Stein-James estimators. In this section, we will obtain four corollaries to Cogburn's Theorem ([2] page 25).

The following structure is assumed: Let R be r -dimensional Euclidean space; S , a subspace of R of dimension s , and T the t -dimensional orthogonal complement of S . The notation of [2] is used where X_S denotes the projection of X on S , and XY denotes the usual inner product of two vectors. Structure H : Let X, Y be random variables in R such that $X = Y + \xi$, with ξ a point in R , such that $EY = 0$ and $EY_T^2 = \tau^2$, with $0 < \tau^2 < \infty$. Suppose $\theta^2 = EY_S^2 < \infty$. We wish to estimate ξ subject to squared-error loss $(\delta - \xi)^2$. Bounds on the mean squared error (MSE) of

$$(2.1) \quad \delta^*(X) = X_S + \Lambda^* X_T,$$

where $\Lambda^* = (X_T^2 - \hat{\tau}^2)^+ / X_T^2$, $\hat{\tau}^2$ an estimator of τ^2 , and $Z^+ \equiv \max(0, Z)$, are determined.

Received December 15, 1970.

¹ Research supported in part by National Science Foundation Grant NSF-GP-9640 at Columbia University, Department of Mathematical Statistics.

² Now at A.T. & T., Management Sciences Division, New York.

REMARK. The general type of problem to which the canonical form results below apply can be stated as follows: Let Z be a random vector in q dimensional coordinate space Ω_0 distributed with mean θ and covariance matrix $\sigma^2 I$, with σ^2 unknown and I the $q \times q$ identity matrix. It is assumed that mean θ lies in a given subspace Ω of Ω_0 , and we let Ω^* denote the orthogonal complement of Ω . The corollaries then hold with $X = Z_\Omega$, $n =$ dimension of Ω^* , $\hat{\sigma}^2 = (n + 2)^{-1}Z_{\Omega^*}^2$ and $r = q - n =$ dimension of Ω . If S is the null space, $\hat{\tau}^2 = r\hat{\sigma}^2$; if S is the space generated by the equiangular line, $\hat{\tau}^2 = (r - 1)\hat{\sigma}^2$.

COROLLARY 2.1. $MSE(\hat{\delta}^*) = \theta^2 + \tau^2\{\lambda + o(1)\}$,

given assumption H , where $\lambda = \hat{\xi}_T^2(\tau^2 + \hat{\xi}_T^2)^{-1}$, if:

- (i) components Y_i of Y are independent with distributions drawn from a family of uniformly square integrable distributions;
- (ii) $\min_i EY_i^2 \geq \varepsilon$ for some fixed $\varepsilon > 0$;
- (iii) $\max_i EY_i^2 = o(t)$ as $t \rightarrow \infty$;
- (iv) the sequence $\{\hat{\tau}^2\tau^{-2}\}$ converges in mean of order 1 to 1, for $E\hat{\tau}^4 \leq K < \infty$.

PROOF. Let δ , the estimator used by Coghurn [2], be (2.1) with τ^2 the true (but unknown) parameter in place of $\hat{\tau}^2$.

$$\tau^{-2}(MSE(\hat{\delta}^*) - MSE(\delta)) \leq \tau^{-2}E(\hat{\delta}^* - \delta)^2 + 2\tau^{-2}\{E(\hat{\delta}^* - \delta)^2 MSE(\delta)\}^{\frac{1}{2}}.$$

Moreover,

$$\begin{aligned} \tau^{-2}E(\hat{\delta}^* - \delta)^2 &= \tau^{-2}E\{(\min(\hat{\tau}^2 X_T^{-2}, 1) - \min(\tau^2 X_T^{-2}, 1))X_T\}^2 \\ &\leq \tau^{-2}E\{(\hat{\tau}^2 - \tau^2)^2 X_T^{-2}\} \quad \text{for } X_T^2 \geq \min(\tau^2, \hat{\tau}^2) \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

$$\tau^{-2}E\{(\hat{\tau}^2 - \tau^2)^2 X_T^{-2}\} \rightarrow \tau^{-2}(\hat{\xi}_T^2 + \tau^2)^{-1}E(\hat{\tau}^2 - \tau^2)^2 \leq E(\hat{\tau}^2\tau^{-2} - 1)^2 \rightarrow 0.$$

The first limit follows because $X_T^2 \rightarrow \{\hat{\xi}_T^2 + \tau^2\}$ w.p. 1; the second, from hypothesis (iv). The result is proved, since, from [2], $MSE(\delta) = \theta^2 + \tau^2 \times \{\lambda + o(1)\}$ under (i)-(iii).

Now consider ξ to be a (Bayesian) random variable in R . Assume H^* : Y is independent of ξ ; components Y_i are independent and drawn from a given family of uniformly square integrable distributions with mean 0 and variance in the interval $[m, m']$, for some $0 < m \leq m' < \infty$. The corollary below follows from Result 1 ([2] page 29) and Corollary 2.1.

COROLLARY 2.2 Assuming H and H^* , if $E\{|\hat{\tau}^2 - \tau^2|\tau^{-2}|\hat{\xi}\} \rightarrow 0$ for $E(\hat{\tau}^4|\hat{\xi}) \leq K < \infty$, then

$$MSE(\hat{\delta}^*) = \theta^2 + \tau^2\{\rho + o(1)\}, \quad \text{where } \rho = E\lambda.$$

Let us deal specifically with the empirical Bayes assumption. In addition to H^* , we postulate that H' : ξ_i are i.i.d., with variance Ψ^2 , and Y_i are i.i.d., with common second moment σ^2 . We suspect ξ lies on the equiangular line S in r -dimensional space, or $\xi = ce$, where e is the vector of r ones and c is

any constant. A reduction in $MSE(\delta^*)$ over $\theta^2 + \tau^2 = r\sigma^2$ for the estimator X , can be expected. With a proof paralleling that of Corollary 2.1, we obtain

COROLLARY 2.3. *Under the hypotheses of Corollary 2.2 and H' ,*

$$\frac{MSE(\delta^*)}{r} = \frac{\sigma^2\Psi^2}{\sigma^2 + \Psi^2} + o(1).$$

We now consider the case where $\sigma_i^2 = EY_i^2$ are random variables on $[m, m']$. Let

$$(2.2) \quad X_i = W_i V_i + \xi_i \quad (i = 1, \dots, r).$$

We introduce a joint prior on (ξ_i, W_i) which is independent of V_i :

$$(2.3) \quad \begin{aligned} EW_i^2 &= \sigma^*, & EV_i &= 0, & E\xi_i &= \mu, \\ EV_i^2 &= 1, & E\xi_i^2 &= \mu^2 + \Psi^2. \end{aligned}$$

Hence, $EX_i = \mu$ and $\text{Var } X_i = \sigma^* + \Psi^2$. Let $\eta = (\sigma_1, \dots, \sigma_r)$. Then $E(X_i|\xi, \eta) = \xi_i$ and $\text{Var}(X_i|\xi, \eta) = \sigma_i^2$, which is the structure assumed in H . Let $F_i = W_i V_i$. Then $X = F + \xi$, and now $\tau^2 = E(F_T^2|\xi, \eta)$.

COROLLARY 2.4. *Given (2.2) with ξ, W random variables with a joint distribution, with pairs (ξ_i, W_i) i.i.d., V_i i.i.d., with moments as specified in (2.3). Let S be the equiangular line. Also assume that, conditional on $\eta = (\sigma_1, \dots, \sigma_r)$, F_i are drawn from a uniformly square integrable distribution family with means 0 and variances in the interval $[m, m']$, and that $E\{t^2 - \tau^2|\tau^{-2}|\xi, \eta\} \rightarrow 0$ for $E(t^4|\xi, \eta) \leq K < \infty$. Then,*

$$\frac{MSE(\delta^*)}{r} = \frac{\sigma^*\Psi^2}{\sigma^* + \Psi^2} + o(1).$$

PROOF.

$$\begin{aligned} \frac{MSE(\delta^*)}{r} &= \frac{E\theta^2}{r} + \frac{E(\tau^2(\lambda + o(1)))}{r}, \\ \frac{\lambda\tau^2}{r} &\xrightarrow{\text{a.s.}} \frac{1}{r} \frac{E\tau^2 E\xi_T^2}{E\tau^2 + E\xi_T^2} = \frac{(r-1)\sigma^*\Psi^2}{r(\sigma^* + \Psi^2)}, \end{aligned}$$

and the result follows by applying the Dominated Convergence Theorem and then taking the limit as $r \rightarrow \infty$.

NOTE. If V_i is distributed Normally, and each (ξ_i, W_i^{-1}) is distributed as a Normal—gamma 2 distribution (Raiffa and Schlaifer [4] page 300), for $i = 1, \dots, r$, then the Bayes risk is precisely $r\sigma^*\Psi^2/(\sigma^* + \Psi^2)$.

Acknowledgment. I wish to acknowledge, with gratitude, the helpful comments of Professor A. J. Baranchik and the referee.

REFERENCES

[1] BARANCHIK, A. (1964). Multiple regression and estimation of the mean of a multivariate

- normal population. Technical Report No. 51, Dept. of Statistics, Stanford Univ.
- [2] COGBURN, R. (1965). On the estimation of a multivariate location parameter with squared error loss. *Bernoulli Bayes Laplace Anniversary Volume*. Springer-Verlag, Berlin, 24-29.
 - [3] JAMES, W. and STEIN, C. (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob. 1* Univ. of California Press, 361-379.
 - [4] RAIFFA, H. and SCHLAIFER, R. (1961). *Applied Statistical Decision Theory*. The M.I.T. Press.
 - [5] WIND, S. (1970). An empirical Bayes approach to the multiple linear regression problem. Ph.D. dissertation, Columbia Univ.