Ann. Inst. Statist. Math.
38 (1986), Part A, 399-410

COMPONENT RISK IN MULTIPARAMETER ESTIMATION

KHURSHEED ALAM AND AMITAVA MITRA

(Received Mar. 18, 1985; revised Oct. 11, 1985)

Summary

For estimating the mean of a p-variate normal distribution under a quadratic loss, a class of estimators, known as Stein's estimators, is known to dominate the maximum likelihood estimator (MLE) for $p\geqq 3$. But, whereas the risk of the MLE has the same value, equal to a constant, for each component, the maximum component risk of Stein's estimator is large for large values of p. Certain modification of Stein's rule has been proposed in the literature for reducing the maximum component risk. In this paper, a new rule is given for reducing the maximum component risk. The new rule yields larger reduction in the maximum component risk, compared to its competitor.

1. Introduction

Stein [6] obtained the surprising result that for estimating p independent normal means simultaneously, the maximum likelihood estimator (MLE) was inadmissible under the sum of squared errors as the loss function, when $p\geqq 3$. Subsequently, an explicit estimator dominating the MLE was given by James and Stein [3]. However, the James-Stein estimator suffers from a serious deficiency in that whereas its total risk is smaller than the total risk of the MLE, the estimate of an individual mean may be subject to much larger error than the MLE, especially when p is large. As Lehmann ([4], p. 308) puts it, "no one wants his or her blood test or Pap smear subjected to the possibility of large errors in order to improve a laboratory's average performance". In this paper we consider the problem of controlling the risk due to an individual component of the estimator, in order to limit the maximum component risk. For this purpose we suggest a simple modification of the James-Stein rule. Efron and Morris [2] have proposed a general class of estimators for the same purpose. They have considered a

Key words and phrases: Multivariate normal distribution, maximum likelihood, risk, Stein's rule.

family of such estimators which have been called "limited translation estimators". However, the class of estimators, we suggest, is more tractable mathematically and gives better performance than the limited translation estimators of Efron and Morris, with respect to the maximum component risk.

The proposed estimator (family of estimators) is denoted by η and its improvement by η^+. They are introduced in Section 2. Formulas for the component risk function of η and η^+ are given in the Appendix. Table 1 gives the maximum component risk of η^+ for various values of an indexing parameter α. Some numerical results on the performances of η^+ are given in Section 3. It is shown in Sections 4 and 5 that η is minimax for sufficiently small values of α and that η^+ dominates η.

2. Estimators η and η^+

Let the p-component vector $X=(X_1,\cdots,X_p)'$ be normally distributed with mean $\Theta=(\Theta_1,\cdots,\Theta_p)'$ and covariance $\sigma^2 I$, where I denotes the identity matrix. Let the loss function for estimating Θ be given by

$$L(\delta,\Theta)=\sum_{i=1}^{p}(\delta_i-\Theta_i)^2$$

where $\delta_i=\delta_i(X)$, denoting the i-th component of δ, is an estimate of Θ_i. It is known that the MLE, which is equal to X, is a minimax estimator. Its risk is equal to $p\sigma^2$. The James-Stein estimator which dominates the MLE, is given by

$$\delta^*(X)=\left(1-\frac{\nu(p-2)\sigma^2}{S}\right)X\,, \tag{2.1}$$

when σ is known, $S=\sum_{1}^{p}X_i^2$ and ν is a constant, such that $0<\nu<2$. When σ is not known, we replace σ by its estimate in the expression for δ^*. We shall deal mainly with the case when σ is known. Therefore, we let $\sigma=1$ in (2.1) without loss of generality. The risk of δ^* is minimized for $\nu=1$. Therefore, we let $\nu=1$ throughout the following discussion.

The James-Stein estimator provides a poor estimate of some of the components of Θ with unusually large values. It is because the factor multiplying X in (2.1) unduly shrinks the corresponding components of X towards the origin or the normal value. The modification of the James-Stein estimator proposed by Efron and Morris is designed to constrain the estimator towards the MLE. The modified estimator is given by

$$\delta_i=\left(1-\frac{p-2}{S}\rho((p-2)X_i^2/S)\right)X_i\,, \tag{2.2}$$

where $\rho(u)$ is a suitably chosen function, which is ordinarily decreasing on its domain $0\leqq u\leqq p-2$. Specifically, the authors have considered the function $\rho=\rho_D$, depending on a parameter D, where $0\leqq D\leqq(p-2)^{1/2}$. It is given by

$$\rho_D(u)=\min(1, D/\sqrt{u})\;. \tag{2.3}$$

The estimator (2.2) with ρ given by (2.3) is denoted by δ^D.

We consider an alternative specification of ρ in (2.2), given by

$$\rho(u)=\frac{(p-2)}{p-2+(\alpha-1)u}\,, \tag{2.4}$$

where $\alpha\geqq 1$ is a constant. The corresponding estimator η is given component-wise by

$$\eta_i(X)=\left(1-\frac{(p-2)}{T_i}\right)X_i\,, \tag{2.5}$$

where $T_i=(\alpha-1)X_i^2+S$. Clearly, $\eta=\delta^*$ for $\alpha=1$.

Let $(x)^+$ denote the positive part of x. Substituting for the factor multiplying X_i in (2.5), its positive part, we obtain the estimator η^+, given component-wise by

$$\eta_i^+=\left(1-\frac{(p-2)}{T_i}\right)^+X_i\;. \tag{2.6}$$

We shall see below that η^+ dominates η, component-wise. Therefore, η^+ should be used in place of η, in practice, even though η^+ is less tractable than η, mathematically.

A derivation of δ^* is given as follows: Suppose that $\theta_1,\cdots,\theta_p$ are independently and identically distributed à priori according to a normal distribution with mean 0 and variance τ^2. A Bayes estimator of θ is given component-wise by

$$\hat{\theta}_i=\left(1-\frac{\sigma^2}{\sigma^2+\tau^2}\right)X_i\;. \tag{2.7}$$

Now, X_i is distributed marginally according to the normal distribution $N(0,\sigma^2+\tau^2)$. Therefore, S is distributed as $(\sigma^2+\tau^2)\chi_p^2$ (chi-square with p degrees of freedom). Hence

$$E\frac{p-2}{S}=\frac{1}{\sigma^2+\tau^2}\;.$$

The James-Stein estimator (δ^*) is obtained by substituting $(p-2)/S$ for $(\sigma^2+\tau^2)^{-1}$ in (2.7).

A motivation for the choice of the proposed estimator η is given as follows:

We observe that the same shrinkage factor is applied to each component of X in the expression for δ^*, given by (2.1). If all the θ's are fairly close to 0, then the components of X are substantially reduced in absolute value, leading to an improvement in the estimated values. If there is a good number of moderate or large values of θ_i's, then the factor multiplying X will be close to 1 and therefore δ^* will not be very different from X. On the other hand, if most of the θ_i's are close to 0 but there are a few large θ_i's, the estimated values of these θ_i's are heavily shrunk towards the norm. Therefore, the corresponding component risks become large. The presence of the additional term $(\alpha-1)X_i^2$ in the expression for T_i alleviates this problem by bringing the shrinkage factor of X_i closer to 1 when X_i is large. The mathematical tractability of η is also an important consideration in the choice of the proposed estimator.

3. Maximum component risk of η^+

From (2.1), putting $\nu=1$, we obtain the risk of δ^*, given by

$$R(\delta^*, \theta)=p-(p-2)^2E\frac{1}{S}\,. \tag{3.1}$$

The above formula shows that δ^* dominates the MLE. Let $R(\delta_i^*, \theta)=E(\delta_i^*-\theta_i)^2$ denote the i-th component risk of δ^*, and let $R_p^*(\theta)=\max_i R(\delta_i^*, \theta)$ denote the largest component risk. Given $\lambda=\sum_{i=1}^{p}\theta_i^2$, it can be shown (Baranchik [1]) that $R_p^*(\theta)$ is maximum when all but one of the components of θ are equal to zero and the remaining component is equal to $\sqrt{\lambda}$. The maximum risk as a function of λ increases from a minimum value of $2/p$ at $\lambda=0$ to a maximum R_p^*, say, and then decreases, tending to 1 as $\lambda\to\infty$. The value of R_p^* is approximately equal to $p/4$ for large p.

Whereas $R(\delta_i^*, \theta)$ is maximized, given λ, for $\theta_i^2=\lambda$ and $\theta_j=0$ $(j\neq i)$, to maximize $R(\eta_i, \theta)$, given by (A.4) of Appendix we maximize this expression first with respect to $\lambda_i=\theta_i^2$ for fixed λ and then with respect to λ. We maximize $R(\eta_i^+, \theta)$ similarly. Let $(\lambda_i^\circ, \lambda^\circ)$ denote the maximizing value of (λ_i, λ) for η^+. We have examined the ratio $\lambda_i^\circ/\lambda^\circ$. For small and moderate values of p. The ratio is close to one. But for large values of p the ratio varies considerably with the value of s, as defined by (3.2). The figures for $p=20$ are shown below for illustration:

s	0.50	0.60	0.75	0.80	0.90	0.95	1.00
$\lambda_i^\circ/\lambda^\circ$	0.40	0.84	0.48	0.60	0.91	0.97	1.00

To compare the component risk of η with δ^* we use the index

$$s=(1-R(\eta_i, 0))/(1-R(\delta_i^*, 0)) , \tag{3.2}$$

which is a measure of the relative savings in the component risk of η at the origin, compared with δ^*. We have found numerically that s is a decreasing function of α with $s=1$ for $\alpha=1$ when $\eta=\delta^*$. Thus a given value of s gives a corresponding value of α. The type of indexing given by (3.2) has been considered by Efron and Morris [2].

Numerical results. Formulas for the component risks of η and η^+ are derived in the Appendix. Table 1 below gives the values of the maximum component risk (MCR) of η^+ with the associated values of α for $p=3(1)6(2)12, 16, 20, 30$ and $s=.5, .6, .75, .8, .9, .95, 1.00$, and $\nu=1$. The upper and lower figures in each entry represent the values of MCR and α, respectively. The figures for the MCR in the last column of the table, corresponding to $s=1$ or equivalently $\alpha=1$, represent the MCR of Stein's estimator δ^*, given by (2.1), with the factor multiplying X replaced by its positive part. It is seen from the table that a small reduction in the value of s from $s=1$ to $s=.95$, say, leads to a considerable reduction in the value of MCR for η^+, especially when p is large. Thus for $p=30$, the value of the MCR is reduced from 7.891 for $s=1$ to 2.078 for $s=.95$.

It is seen from Table 1 for the specified values of p that the value of MCR is decreasing in s. We can use the table for application in a simple fashion, as follows: Let $p=12$, for example. Suppose that we can tolerate a maximum component risk not more than two times the standard deviation of an individual component of X. We see that the given condition is satisfied for a value of α equal to or smaller than 2.25. Therefore, we should use η^+ to estimate Θ with the value of $\alpha=2.25$ approximately. On the other hand, the maximum component risk for the James-Stein rule ($\alpha=1$) is seen to be as large as 3.392, approximately.

Let us compare the MCR values of η^+ with the MCR values of δ^D, the limited translation estimator of Efron and Morris [2], given in Table 1 of their paper. We find that the MCR value of η^+ is smaller than the MCR value of δ^D for each of the specified values of p and s. The difference between the two values tends to increase with increasing values of p. For illustration, consider $p=10$. The MCR values of η^+ (δ^D) are 1.059 (1.14) and 1.603 (1.66) for $s=.5$ and .9, respectively. Next, consider $p=30$. The MCR values of η^+ (δ^D) are 1.000 (1.16) and 1.533 (2.02) for $s=.5$ and .9, respectively.

Table 1. Maximum component risk of η^+ and values of α

p \ s	0.5	0.6	0.75	0.8	0.9	0.95	1.00
3	1.049 5.61	1.066 4.16	1.094 2.74	1.106 2.38	1.134 1.76	1.156 1.46	1.215 1.00
4	1.071 6.73	1.101 4.90	1.157 3.12	1.180 2.67	1.242 1.91	1.292 1.55	1.439 1.00
5	1.080 7.86	1.110 5.64	1.197 3.49	1.234 2.97	1.334 2.06	1.413 1.64	1.672 1.00
6	1.072 9.00	1.126 6.39	1.232 3.88	1.280 3.26	1.413 2.21	1.527 1.72	1.911 1.00
8	1.080 11.28	1.119 7.88	1.259 4.64	1.332 3.85	1.540 2.51	1.720 1.90	2.399 1.00
10	1.059 13.57	1.109 9.39	1.238 5.41	1.340 4.44	1.603 2.82	1.878 2.08	2.894 1.00
12	1.044 15.86	1.099 10.89	1.257 6.18	1.331 5.04	1.673 3.12	1.990 2.25	3.392 1.00
16	1.010 20.45	1.065 13.91	1.210 7.73	1.310 6.23	1.675 3.73	2.136 2.61	4.392 1.00
20	1.008 25.05	1.008 16.92	1.171 9.26	1.265 7.42	1.654 4.36	2.172 2.96	5.392 1.00
30	1.000 31.53	1.001 25.46	1.074 13.12	1.141 10.40	1.533 5.87	2.078 3.84	7.891 1.00

The upper and lower figures in each entry represent the MCR value of η^+ and the value of α, respectively.

4. Minimax property of η

The i-th component risk of η is given by

$$(4.1)\qquad \begin{aligned} R(\eta_i, \Theta) &= E(\eta_i(X)-\Theta_i)^2 \\ &= 1-2(p-2)E(X_i(X_i-\Theta_i)/T_i)+(p-2)^2E(X_i^2/T_i^2) \\ &= 1+(p-2)(4\alpha+(p-2))E(X_i^2/T_i^2)-2(p-2)E(T_i^{-1})\,. \end{aligned}$$

As $X_i^2/T_i=y/((\alpha-1)y+S)$ is a concave function of $y=X_i^2$, from Jensen's inequality we have that

$$(4.2)\qquad \sum_{i=1}^{p}\frac{X_i^2}{T_i}\leqq S\Big/\left(\frac{(\alpha-1)S}{p}+S\right)=\frac{p}{\alpha+p-1}\,.$$

Since X_i^2/T_i is increasing in y and $1/T_i=1/((\alpha-1)y+S)$ is decreasing in y, we have

$$(4.3)\qquad \sum_{i=1}^{p}\frac{X_i^2}{T_i^2}\leqq\frac{1}{p}\left(\sum_{i=1}^{p}\frac{X_i^2}{T_i}\right)\left(\sum_{i=1}^{p}\frac{1}{T_i}\right)\leqq\frac{1}{(\alpha+p-1)}\sum_{i=1}^{p}\frac{1}{T_i}\,,\qquad \text{by (4.2)}\,.$$

From (4.1) and (4.3) we have

$$(4.4)\qquad R(\eta, \Theta)\leqq p+(p-2)\left(\frac{4\alpha+(p-2)}{\alpha+p-1}-2\right)E\sum_{i=1}^{p}\frac{1}{T_i}\leqq p$$

for

$$\alpha \leqq \frac{p}{2} . \tag{4.5}$$

Therefore, η is minimax for all values of α, satisfying the inequality (4.5).

Numerical values of the risk of η (not shown here) show that the inequality (4.5), giving the values of α for which η is minimax, is fairly sharp. For example, we find that for $p=12$ and $\alpha=6.18$, which is slightly larger than the upper bound $p/2=6$, the maximum risk is equal to 12.001, which is slightly larger than the minimax value 12. We find also that even when η is not minimax, the maximum risk of η exceeds only slightly from the maximum risk of the MLE or δ^* which are minimax.

When σ^2 is not known but an estimate of σ^2 is given by W which is distributed as $\sigma^2\chi_m^2$ (chi-square with m degrees of freedom) independent of X, we let η_i be given by

$$\eta_i=\left(1-\frac{(p-2)W}{(m+2)T_i}\right)X_i .$$

It can be shown that η is again minimax for $\alpha \leqq p/2$.

5. Dominating property of η^+

Lehmann [4] gives an interesting proof of the result (Theorem 6.2) that the James-Stein estimator δ^* is improved by replacing the factor of X in (2.1) by its positive part. Following his argument, it is straightforward to show that η^+ dominates η, component-wise. We state this result but omit the proof.

THEOREM 5.1. $R(\eta_i^+, \theta)<R(\eta_i, \theta)$, $\forall \theta$, $i=1,\cdots,p$.

We have seen from Table 1 that the maximum component risk of η^+ is smaller than the maximum component risk of the Efron-Morris estimator δ^D. It does not imply that η^+ dominates δ^D. We would like to see when is one better or worse than the other. However, we do not have figures for the risk functions to make the comparison, at this stage.

Appendix

Component risk of η

First we give certain results which will be used in the sequel. Let $F_m(x)$ denote the chi-square distribution function with m degrees of freedom and let $F_{m,\lambda}(x)$ denote the non-central chi-square distribution function with m degrees of freedom and non-centrality parameter λ. For $\alpha \geqq 1$, let $c_{j,m} = c_{j,m}(\alpha)$ denote the coefficient of x^j in the power series expansion of the expression

$$\alpha^{-m/2}\left(1-\left(1-\frac{1}{\alpha}\right)x\right)^{-m/2}.$$

LEMMA 1. *Let χ^2_m and $\chi^2_{n,\lambda}$ be independent chi-square random variables. Then*

$$E(\alpha\chi^2_m+\chi^2_{n,\lambda})^{-1}=\sum_{j=0}^{\infty} c_{j,m}E(\chi^2_{m+n+2j,\lambda})^{-1}.$$

PROOF. By Theorem 1 of Robbins and Pitman [5] we have that

$$P\{\alpha\chi^2_m+\chi^2_n \leqq x\}=\sum_{j=0}^{\infty} c_{j,m}F_{m+n+2j}(x)$$

for each $x \geqq 0$. From the representation of non-central chi-squre distribution as a mixture of central chi-square distributions, the above result is generalized as follows:

$$P\{\alpha\chi^2_m+\chi^2_{n,\lambda} \leqq x\}=\sum_{j=0}^{\infty} c_{j,m}F_{m+n+2j,\lambda}(x). \tag{A.1}$$

The conclusion of the lemma follows from (A.1). The formula (A.1) is essentially given in Robbins and Pitman [5], but it is not explicitly stated.

The component risk of η is derived as follows: Let $\lambda_i=\Theta_i^2$, $\lambda=\Theta'\Theta$, let $\chi^2_{m,\lambda}$ denote a non-central chi-square random variable with m degrees of freedom and non-centrality parameter λ, and let

$$\Phi(a, b; x)=1+\frac{a}{b}x+\frac{a(a+1)}{b(b+1)}\frac{x^2}{2!}+\cdots$$

denote the confluent hypergeometric function. We have

$$\begin{aligned} \text{(A.2)}\quad E(T_i)^{-1} &= e^{-\lambda_i/2}\sum_{r=0}^{\infty}\frac{(\lambda_i/2)^r}{r!}E(\alpha\chi^2_{2r+1}+\chi^2_{p-1,\lambda-\lambda_i})^{-1} \\ &= e^{-\lambda_i/2}\sum_{r=0}^{\infty}\frac{(\lambda_i/2)^r}{r!}\sum_{m=0}^{\infty}c_{m,2r+1}E(\chi^2_{p+2m+2r,\lambda-\lambda_i})^{-1} \qquad \text{(by Lemma 1)} \end{aligned}$$

$$=e^{-\lambda_i/2}\sum_{r=0}^{\infty}\frac{(\lambda_i/2)^r}{r!}\sum_{m=0}^{\infty}\frac{c_{m,2r+1}}{p+2m+2r-2}$$
$$\times\Phi\left(1,\frac{p}{2}+m+r;\ -(\lambda-\lambda_i)/2\right)$$
$$=\frac{1}{2}e^{-\lambda_i/2}\sum_{r=0}^{\infty}\frac{(\lambda_i/2)^r}{r!}\sum_{m=0}^{\infty}\int_0^1 c_{m,2r+1}(1-u)^{p/2+m+r-2}e^{-(\lambda-\lambda_i)u/2}du$$
$$=\frac{1}{2}e^{-\lambda_i/2}\sum_{r=0}^{\infty}\frac{(\lambda_i/2)^r}{r!}\int_0^1 \alpha^{-r-1/2}(1-u)^{p/2+r-2}$$
$$\times\left(1-\left(1-\frac{1}{\alpha}\right)(1-u)\right)^{-r-1/2}e^{-(\lambda-\lambda_i)u/2}du$$
$$=\frac{1}{2}e^{-\lambda_i/2}\sum_{r=0}^{\infty}\frac{(\lambda_i/2)^r}{r!}\int_0^1(1-u)^{p/2+r-2}$$
$$\times(1+(\alpha-1)u)^{-r-1/2}e^{-(\lambda-\lambda_i)u/2}du$$
$$=\frac{1}{2}\int_0^1(1-u)^{p/2-2}(1+(\alpha-1)u)^{-1/2}$$
$$\times\exp\left(-\frac{\lambda u}{2}-\frac{\lambda_i(\alpha-1)u(1-u)}{2(1+(\alpha-1)u)}\right)du\ .$$

Similarly

$$\text{(A.3)}\qquad E(X_i^2/T_i^2)=-\frac{\partial}{\partial\alpha}E\frac{1}{T_i}$$
$$=\frac{1}{4}\int_0^1 u(1-u)^{p/2-2}(1+(\alpha-1)u)^{-3/2}\left(1+\frac{\lambda_i(1-u)}{1+(\alpha-1)u}\right)$$
$$\times\exp\left(-\frac{\lambda u}{2}-\frac{\lambda_i(\alpha-1)u(1-u)}{2(1+(\alpha-1)u)}\right)du\ .$$

From (A.1), (A.2) and (A.3) we get

$$\text{(A.4)}\qquad R(\eta_i,\Theta)=1+\frac{1}{4}(p-2)(4\alpha+(p-2))\left[\int_0^1(1-u)^{p/2-2}(1+(\alpha-1)u)^{-1/2}\right.$$
$$\times\left(\frac{u}{1+(\alpha-1)u}\left(1+\frac{1-u}{1+(\alpha-1)u}\lambda_i\right)-\frac{4}{4\alpha+(p-2)}\right)$$
$$\left.\times\exp\left(-\frac{\lambda}{2}u-\frac{(\alpha-1)u(1-u)}{2(1+(\alpha-1)u)}\lambda_i\right)du\right]\ .$$

Putting $\Theta=0$ in (A.4) we get the component risk at the origin, equal to

$$\text{(A.5)}\qquad R(\eta_i,0)=1+\frac{1}{4}(p-2)(4\alpha+(p-2))\int_0^1(1-u)^{p/2-2}(1+(\alpha-1)u)^{-1/2}$$
$$\times\left(\frac{u}{1+(\alpha-1)u}-\frac{4}{4\alpha+(p-2)}\right)du$$

$$=1+\frac{1}{p}(4\alpha+(p-2))\alpha^{-3/2}\left[F\left(\frac{3}{2},\frac{p}{2}-1;\frac{p}{2}+1;\frac{\alpha-1}{\alpha}\right)\right.$$
$$\left.-\frac{2\alpha p}{4\alpha+(p-2)}F\left(\frac{1}{2},\frac{p}{2}-1;\frac{p}{2};\frac{\alpha-1}{\alpha}\right)\right],$$

where $F(a,b;c;x)=1+\frac{ab}{c}x+\frac{a(a+1)b(b+1)}{c(c+1)}\frac{x^2}{2!}+\cdots$ denotes the hypergeometric function.

The value of $R(\delta_i^*,\Theta)$, the component risk of δ^*, is obtained by putting $\alpha=1$ in (A.4). Thus we have

$$\text{(A.6)}\qquad R(\delta_i^*,\Theta)=1+\frac{1}{4}(p-2)(4+(p-2))\int_0^1(1-u)^{p/2-2}\left[u(1+\lambda_i(1-u))\right.$$
$$\left.-\frac{4}{4+(p-2)}\right]e^{-\lambda u/2}du\ .$$

Putting $\Theta=0$ we get

$$\text{(A.7)}\qquad R(\delta_i^*,0)=2/p\ .$$

The maximum component risk of δ^* is obtained by substituting λ for λ_i in the square bracket on the right side of (A.6) and maximizing the corresponing expression for $R(\delta_i^*,\Theta)$, with respect to λ.

Component risk of η^+

We have

$$\eta_i^+(X)=X_i+g_i(X)$$

where

$$g_i(X)=\begin{cases}-\dfrac{(p-2)}{T_i}X_i & \text{for } T_i>(p-2)\\ -X_i & \text{for } T_i\leq(p-2)\,.\end{cases}$$

Therefore

$$\text{(A.8)}\qquad R(\eta_i^+,\Theta)=1+(p-2)(4\alpha+(p-2))\int_{T_i>(p-2)}(X_i^2/T_i^2)dG(X_i,T_i)$$
$$-2(p-2)\int_{T_i>(p-2)}T_i^{-1}dG(T_i)+\int_{T_i\leq(p-2)}(X_i^2-2)dG(X_i,T_i)$$

where G is a generic notation for the cdf. With the help of Lemma 1 we obtain after simplification

$$\text{(A.9)}\qquad \int_{T_i>(p-2)}(X_i^2/T_i^2)dG(X_i,T_i)$$

$$
\begin{aligned}
&= e^{-\lambda/2} \sum_{r=0}^{\infty} \frac{(\lambda_i/2)^r (1+2r)}{r!} \sum_{s=0}^{\infty} c_{s,3+2r} \\
&\times \sum_{t=0}^{\infty} \left(\frac{((\lambda-\lambda_i)/2)^t}{t!(p+2r+2s+2t-2)(p+2r+2s+2t)} \right. \\
&\times \left[1 - \frac{((p-2)/2)^{p/2+r+s+t-1}}{\Gamma(p/2+r+s+t)} e^{-(p-2)/2} \right. \\
&\times \left. \Phi\left(1, \frac{p}{2}+r+s+t; \frac{(p-2)}{2}\right) \right],
\end{aligned}
$$

(A.10)
$$
\begin{aligned}
&\int_{T_i>(p-2)} T_i^{-1} dG(T_i) \\
&= e^{-\lambda/2} \sum_{r=0}^{\infty} \frac{(\lambda_i/2)^r}{r!} \sum_{s=0}^{\infty} c_{s,1+2r} \sum_{t=0}^{\infty} \frac{((\lambda-\lambda_i)/2)^t}{t!(p+2r+2s+2t-2)} \\
&\times \left[1 - e^{-(p-2)/2} \frac{((p-2)/2)^{p/2+r+s+t-1}}{\Gamma(p/2+r+s+t)} \right. \\
&\times \left. \Phi\left(1, \frac{p}{2}+r+s+t; \frac{(p-2)}{2}\right) \right],
\end{aligned}
$$

(A.11)
$$
\begin{aligned}
&\int_{T_i \le (p-2)} (X_i^2-2) dG(X_i, T_i) \\
&= e^{-\lambda/2} \sum_{r=0}^{\infty} \frac{(1+2r)(\lambda_i/2)^r}{r!} \sum_{s=0}^{\infty} c_{s,3+2r} \sum_{t=0}^{\infty} \frac{((\lambda-\lambda_i)/2)^t}{t!} \\
&\times \frac{((p-2)/2)^{p/2+1+r+s+t}}{\Gamma(p/2+2+r+s+t)} e^{-(p-2)/2} \Phi\left(1, \frac{p}{2}+2+r+s+t; \frac{(p-2)}{2}\right) \\
&-2e^{-\lambda/2} \sum_{r=0}^{\infty} \frac{(\lambda_i/2)^r}{r!} \sum_{s=0}^{\infty} c_{s,1+2r} \sum_{t=0}^{\infty} \frac{((\lambda-\lambda_i)/2)^t}{t!} \\
&\times \frac{((p-2)/2)^{p/2+r+s+t}}{\Gamma(p/2+1+r+s+t)} e^{-(p-2)/2} \Phi\left(1, \frac{p}{2}+r+s+t+1; \frac{(p-2)}{2}\right).
\end{aligned}
$$

We substitute for the integrals in (A.8) their corresponding values given by (A.9) through (A.11).

CLEMSON UNIVERSITY
AUBURN UNIVERSITY

REFERENCES

[1] Baranchik, A. J. (1964). Multiple regression and estimation of the mean of a multivariate normal distribution, Department of Statistics, Stanford University, TR **51**.

[2] Efron, B. and Morris, C. (1972). Limiting the risk of Bayes and empirical Bayes estimators—Part II: The empirical Bayes case, *J. Amer. Statist. Ass.*, **67**, 130-139.

[3] James, W. and Stein, C. (1961). Estimation with quadratic loss, *Proceedings of the Fourth Berkeley Symposium*, University of California Press, 361-379.

[4] Lehmann, E. L. (1983). *Theory of Point Estimation*, Wiley Series in Probability and Statistics.

[5] Robbins, H. and Pitman, E. J. G. (1949). Application of the method of mixtures to quadratic forms in normal variates, *Ann. Math. Statist.*, **50**, 552-560.

[6] Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, *Proc. 3rd Berkeley Symposium on Mathematical Statistics and Probability*, **1**, 197-206.