Ann. Inst. Statist. Math.
Vol. 44, No. 2, 299–311 (1992)

ASYMPTOTIC RISK BEHAVIOR OF MEAN VECTOR AND VARIANCE ESTIMATORS AND THE PROBLEM OF POSITIVE NORMAL MEAN*

ANDREW L. RUKHIN

Department of Mathematics and Statistics, The University of Maryland, Baltimore County Campus, Baltimore, MD 21228-5398, U.S.A. and University of Münster

(Received August 20, 1990; revised February 20, 1991)

Abstract. Asymptotic risk behavior of estimators of the unknown variance and of the unknown mean vector in a multivariate normal distribution is considered for a general loss. It is shown that in both problems this characteristic is related to the risk in an estimation problem of a positive normal mean under quadratic loss function. A curious property of the Brewster-Zidek variance estimator of the normal variance is also noticed.

Key words and phrases: Bowl-shaped loss function, Brewster-Zidek estimator of normal variance, James-Stein estimator of normal mean, relative risk reduction, positive normal mean, Stein estimator of normal variance.

1. Introduction

In this paper we consider the asymptotic estimation of the unknown variance σ^2 and of the unknown mean of a multivariate normal distribution with a covariance matrix $\sigma^2 I$.

This is a classical problem of multivariate analysis. The inadmissibility of the traditional estimator of the mean for dimensions larger than three is known since 1955, when C. Stein discovered this phenomenon, and this field has been an active area of research since (cf. for example James and Stein (1961), Baranchik (1970), Efron and Morris (1976)).

The traditional estimator of the variance is also known to be inadmissible (Stein (1964)). Although somewhat similar these two results are different. For quadratic loss in the normal mean case one can use the by now popular integration by parts technique to derive an unbiased estimate of the risk difference between the traditional and an alternative estimator. It is possible to find a procedure which

* Research supported by NSF Grant DMS 9000999 and by Alexander von Humboldt Foundation Senior Distinguished Scientist Award.

makes this risk difference estimate nonnegative. The same technique applies in the normal variance case but no nonnegative risk difference estimate exists. Perhaps related is the fact that in the mean vector estimation problem the relative risk reduction tends to 1 as the dimension increases. In the univariate problem of estimating a normal variance the savings do not exceed 4% (Rukhin (1987)).

The goal of this paper is to explore these estimation problems when both the dimension and the sample size tend to infinity. We show that these problems are intimately related to the estimation problem of a positive normal mean on the basis of one observation with unit variance. In particular the largest possible risk improvement in variance estimation is determined by the corresponding quantity in the positive mean problem, which also enters the asymptotic expansion of relative risk reduction of a multivariate normal mean estimator.

The history of normal variance estimation is reviewed by Maatta and Casella (1990). We mention only a paper by Brewster and Zidek (1974), where an admissible improvement over the traditional estimator is derived. We study this estimator in Section 2 and show that the risk function of the Brewster-Zidek estimator has a maximum at the origin. This is surprising because this estimator is generalized Bayes with respect to an (improper) prior density with the mode at the origin.

In Section 3 the counterparts of this estimator and of the original Stein estimator in the positive mean estimation problem are found. They turn out to be the generalized Bayes estimator against the uniform distribution over the positive half-line, and the maximum likelihood estimator. In Section 4 a similar result is obtained for the classical mean vector estimators of Stein and of James-Stein.

2. The risk function of Brewster-Zidek estimator

Let X be a normal random vector with the distribution $N_k(\mu, \sigma^2 I)$ and let S^2 be independent of X with S^2/σ^2 having the chi-squared distribution with $m-1$ degrees of freedom.

This is a canonical form of classical problems of multivariate statistical analysis.

Assume that the unknown variance σ^2 is to be estimated under a nonnegative bowl-shaped smooth loss function $W(\delta/\sigma^2)$ with a unique minimum at 1, $W'(1) = W(1) = 0$. The estimator δ_{BZ} due to Brewster and Zidek (1974) has the form

$$\delta_{\mathrm{BZ}}(X, S) = S^2 \phi_{\mathrm{BZ}}(V),$$

where

$$V = S(\|X\|^2 + S^2)^{-1/2}$$

and the function ϕ_{BZ} is found from the condition

$$E_{01}[W\{S^2\phi_{\mathrm{BZ}}(v)\} \mid V > v] = \min_{\phi} E_{01}\{W(S^2\phi) \mid V > v\}.$$

Here $E_{\mu\sigma}$ refers to the expected value under parameters μ and σ.

We assume that

$$E_{01}|W'\{S^2\phi_{\mathrm{BZ}}(V)\}| < \infty,$$

so that

$$E_{01}[W'\{S^2\phi_{\mathrm{BZ}}(v)\}S^2 \mid V > v] = 0. \tag{2.1}$$

It is known (Brewster and Zidek (1974)) that $\phi_{\mathrm{BZ}}(0) = c_0$, where $c_0 S^2$ is the best multiple of S^2 (the best equivariant estimator), i.e.

$$E_{01}W(c_0 S^2) = \min_c E_{01}W(cS^2). \tag{2.2}$$

Using the explicit form of the joint distribution of S and V, one obtains from (2.1)

$$\begin{aligned}\int_0^\infty \int_v^1 W'\{s^2\phi_{\mathrm{BZ}}(u)\} \exp\left(-\frac{1}{2}s^2/u^2\right)& \\ \times\, s^{m+k-2}(1-u^2)^{(k-2)/2}u^{-k}\,duds &= 0.\end{aligned} \tag{2.3}$$

Multiplying both parts of (2.3) by $\phi'_{\mathrm{BZ}}(v)$ and integrating by parts we see that

$$\begin{aligned} 0 &= \int_0^\infty \left\{\int_v^1 \exp\left(-\frac{1}{2}s^2/u^2\right)(1-u^2)^{(k-2)/2}u^{-k}du\right\} s^{m+k-4}dW(s^2\phi_{\mathrm{BZ}}(v)) \\ &= -\int_0^\infty\int_0^1 W\{s^2\phi_{\mathrm{BZ}}(0)\}\exp\left(-\frac{1}{2}s^2/u^2\right)(1-u^2)^{(k-2)/2}u^{-k}s^{m+k-4}duds \\ &\quad + \int_0^\infty\int_0^1 W\{s^2\phi_{\mathrm{BZ}}(v)\}\exp\left(-\frac{1}{2}s^2/v^2\right)(1-v^2)^{(k-2)/2}v^{-k}s^{m+k-4}dvds \\ &= -E_{01}W(c_0S^2) + E_{01}W\{S^2\phi_{\mathrm{BZ}}(V)\}. \end{aligned}$$

Thus we have proved

PROPOSITION 2.1. *Let $S^2\phi_{\mathrm{BZ}}(V)$ be the Brewster-Zidek estimator of the normal variance under differentiable bowl-shaped loss function W. Then*

$$E_{01}W\{S^2\phi_{\mathrm{BZ}}(V)\} = E_{01}W(c_0S^2), \tag{2.4}$$

i.e. at the origin $\mu = 0$ the risk function of the Brewster-Zidek estimator equals the risk of the best equivariant estimator.

Formula (2.4) is surprising for the following reason: the Brewster-Zidek estimator is known to be minimax,

$$E_{\mu\sigma}W(S^2\phi_{\mathrm{BZ}}(V)/\sigma^2) \le E_{\mu\sigma}W(c_0S^2/\sigma^2) = E_{01}W(c_0S^2).$$

Also it is the generalized Bayes rule with respect to a prior of the form

$$\frac{1}{\sigma^k}\lambda\left(\frac{\mu}{\sigma}\right)d\mu\frac{d\sigma}{\sigma},$$

where

$$\lambda(\eta) = \int_0^\infty \exp\left(-\frac{t\|\eta\|^2}{2}\right) t^{k/2-1}(1+t)^{-1}dt.$$

This (improper) density has a unique mode at $\eta = 0$, and yet the corresponding Bayes estimator has its frequentist risk taking the largest value at $\eta = 0$. This curious fact, noticed first for the quadratic loss and $k = 1$ by Rukhin (1987), shows a difficulty with the traditional interpretation of prior distribution as a parametric weight assignment which reflects the relative importance of different parameter values.

To conclude this section we give an explicit form of the Brewster-Zidek estimator for the quadratic loss:

$$\phi_{\mathrm{BZ}}(v) = (m+1)^{-1}\left\{1 - \frac{v^{m-1}(1-v^2)^{k/2}}{(m+k+1)\int_v^1 t^m(1-t^2)^{k/2-1}dt}\right\} \tag{2.5}$$

and

$$c_0 = (m+1)^{-1}.$$

In the next section we shall use the following analogue of the original Stein (1964) estimator for a general loss function W:

$$\phi_s(v) = \min(c_0, c_1 v^{-2}). \tag{2.6}$$

Here c_0 is defined by (2.2) and c_1 is determined by the condition

$$E_{01}\{W'(c_1 v^{-2}S^2)S^2 \mid V = v\} = 0,$$

which means that

$$\int_0^\infty W'(c_1 s^2)e^{-s^2/2}s^{m+k-2}ds = 0. \tag{2.7}$$

In particular for quadratic loss

$$\phi_s(v) = \min\{(m+1)^{-1}, (m+k+1)^{-1}v^{-2}\}. \tag{2.8}$$

Some numerical results for the risk of these estimators in the case of quadratic loss and entropy loss are reported by Rukhin and Ananda (1992).

3. Asymptotic risk behavior of scale-equivariant variance estimators

Keeping the notation of the previous section, we consider here scale-equivariant variance estimators $\delta(X, S)$ written in the form

$$\delta(X, S) = c_0 S^2\{1 - \Phi(V)\}, \tag{3.1}$$

where c_0 is the constant determined by (2.2) and Φ is a continuous function.

It is easy to see that the corresponding risk function depends only on $\|\mu\|/\sigma = \eta$, so that in risk evaluations one can put $\sigma = 1$.

Since the exact form of the risk is not tractable, we study its asymptotic behavior for large dimension k and large "sample size" m. Their rates and the following limiting formulae (3.4) and (3.5) are suggested by the behavior of estimators (2.5) and (2.8).

Notice first of all that $c_0 m \to 1$ as $m \to \infty$. Indeed

$$E_{01} W'(c_0 S^2) S^2 = 0 \tag{3.2}$$

and with probability one

$$S^2/m \to 1.$$

Since W is bowl-shaped this fact and (3.2) imply that

$$\lim_{m\to\infty} c_0 m = 1.$$

In fact

$$c_0 m = 1 - W''(1)/m + o(m^{-1}). \tag{3.3}$$

Indeed, because of the central limit theorem the distribution of

$$m^{1/2} 2^{-1/2}(S^2 - m)$$

converges to the standard normal one. Let

$$c_0 m = 1 + r_m.$$

Then (3.2) shows that

$$\begin{aligned} 0 &= \lim_{m\to\infty} E_{01} W'(c_0 S^2) S^2 \\ &= \lim_{m\to\infty} mEW'\{(1+r_m)(1+2^{1/2}Zm^{-1/2})\}(1+2^{1/2}Zm^{-1/2}) \\ &= W''(1)EZ^2 + \lim_{m\to\infty} r_m m. \end{aligned}$$

Here Z is a standard normal random variable. It follows that

$$\lim_{m\to\infty} r_m m = -W''(1).$$

Now we can analyze the asymptotic risk behavior of the best equivariant estimator $c_0 S^2$ as m increases.

Indeed one has

$$\lim_{m\to\infty} mE_{01} W(c_0 S^2) = \lim_{m\to\infty} mEW\{(1+r_m)(1+2^{1/2}Zm^{-1/2})\} = W''(1).$$

Also for any estimator (3.1)

$$\lim_{m\to\infty} mE_\eta W\{\delta(X,S)\} = W''(1),$$

so that the first order asymptotic behavior of $\delta(X,S)$ coincides with that of c_0S^2

Therefore the relative risk improvement

$$r(\eta) = [E_\eta W(c_0S^2) - E_\eta W\{\delta(X,S)\}]/E_\eta W(c_0S^2)$$

tends to zero as $m \to \infty$.

Let us consider now the case when $k \to \infty$ and

$$\|\eta\|^2/k \to \eta_1.$$

We also assume that

$$\Phi(vk^{-1/2}) \to \Phi_0(v). \tag{3.4}$$

Then with probability one

$$\frac{\|x\|^2}{k} = \frac{\sum x_j^2}{k} \to 1+\eta_1,$$

and for a fixed m, as $k \to \infty$

$$E_\eta W\{\delta(X,S)\} \to E_0 W[c_0S^2\{1 - \Phi_0(S(1+\eta_1)^{-1/2})\}].$$

Suppose now that for $m \to \infty$

$$m^{1/2}\Phi_0(m^{1/2} - 2^{-1/2}Z) \to 2^{1/2}\chi(Z) \tag{3.5}$$

and

$$m^{1/2}\eta_1 \to 2^{1/2}\Theta$$

with a nonegative finite Θ.

Then with a standard normal Z

$$\begin{aligned}
&\lim_{m\to\infty} mE_0W[c_0S^2\{1-\Phi_0(S(1+\eta_1)^{-1/2})\}] \\
&\quad = \lim_{m\to\infty} mEW[(1+r_m)(1+2^{1/2}Zm^{-1/2}) \\
&\qquad\qquad \times \{1-\Phi_0(m^{1/2}+(Z-\Theta)\cdot(2m)^{-1/2})\}] \\
&\quad = \lim_{m\to\infty} mEW[1+2^{1/2}\{Z+\chi(\Theta-Z)\}m^{-1/2}] \\
&\quad = W''(1)E\{Z-\chi(\Theta-Z)\}.
\end{aligned}$$

Therefore, under our assumptions

$$\lim_{m\to\infty}\lim_{k\to\infty} r(\eta) = 1 - E\{Y+\chi(Y)-\Theta\}^2. \tag{3.6}$$

Here $Y = \Theta - Z$ is a normal random variable with the positive mean Θ and the unit variance.

In other terms the relative risk reduction in the variance estimation problem converges to the risk improvement over the estimator Y in the estimation problem of a positive normal mean.

We formulate results obtained so far.

THEOREM 3.1. *Let $\delta(X, S)$ be a scale-equivariant estimator of the normal variance under bowl-shaped smooth loss function W such that $W(1) = W'(1) = 0$. Assume that as $k \to \infty$ and $m \to \infty$ conditions* (3.4) *and* (3.5) *hold, and that*

$$\lim_{m\to\infty} \lim_{k\to\infty} \|\mu\|^2 \sigma^{-2} m^{1/2} k^{-1} 2^{-1/2} = \Theta. \tag{3.7}$$

Then the asymptotic formula (3.6) *for the relative risk reduction is valid.*

Clearly in the limiting problem of the positive normal mean estimation, the estimator Y, which corresponds to the best equivariant estimator $c_0 S^2$ of the variance, is not a good procedure. (It does not make any sense to estimate a nonnegative parameter by a negative number which can happen with Y.)

However, as we show now, both Stein estimator and Brewster-Zidek estimator have limiting forms which are important estimators of the positive normal mean.

THEOREM 3.2. *Conditions* (3.4) *and* (3.5) *are satisfied for both Stein estimator* (2.6) *and Brewster-Zidek estimator* (2.1). *In the first case*

$$\Phi_0(v) = \max(0, 1 - c_0^{-1} v^{-2}), \qquad \chi_S(z) = \max(-z, 0), \tag{3.8}$$

in the second case $\Phi_0(c)$ is found from the equation

$$\int_v^\infty W'[c_0 u^2 \{1 - \Phi_0(v)\}] u^{m+4} e^{-u^2/2} du = 0, \tag{3.9}$$

$$\chi_{BZ}(z) = e^{-z^2/2} \bigg/ \int_{-\infty}^{z} \exp\left(-\frac{1}{2} t^2\right) dt. \tag{3.10}$$

PROOF. If one writes the Stein estimator (2.6) in the form (3.1), then

$$\Phi(v) = \max(1 - c_1 c_0^{-1} v^{-2}, 0),$$

where c_1 is defined by (2.7).

As in the case of c_0, for a fixed m

$$c_1 = (m + k)^{-1} + o(k^{-1}).$$

Therefore

$$\Phi(v k^{-1/2}) \to \max(1 - c_0^{-1} v^{-2}, 0) = \Phi_0(v).$$

Also

$$\begin{aligned}
&\lim_{m\to\infty} m^{1/2}\max\{1 - c_0^{-1}(m^{1/2} - 2^{-1/2}z)^{-2}, 0\} \\
&\quad = \lim_{m\to\infty} \max\left\{\left(-2^{1/2}zm + \frac{1}{2}z^2m^{1/2}\right)(m^{1/2} - 2^{-1/2}z)^{-2}, 0\right\} \\
&\quad = 2^{1/2}\max(-z, 0),
\end{aligned}$$

and (3.8) follows.

In the case of the Brewster-Zidek estimator make a transformation of variables in (2.3) to see that

$$\begin{aligned}
\int_0^\infty\int_0^{k^{1/2}} & W'[c_0 s^2\{1 - \Phi_0(v)\}]\exp\left(-\frac{1}{2}ks^2u^{-2}\right) \\
& \times s^{m+k+2}(1 - u^2/k)^{k/2-1}u^{-k}duds = 0.
\end{aligned}$$

Notice that for fixed u, v

$$\begin{aligned}
I &= \int_0^\infty W'[c_0s^2\{1 - \Phi_0(v)\}]\exp\left(-\frac{1}{2}ks^2u^{-2}\right)s^{k+m+2}ds \\
&= \int_0^\infty W'[c_0s^2\{1 - \Phi_0(v)\}]\exp\left\{-k\left(\frac{1}{2}s^2u^{-2} - \log s\right)\right\}s^{m+2}ds.
\end{aligned}$$

Since the maximum of the function $s^2u^{-2}/2 - \log s$ occurs at $s = u$, Laplace's method shows that as $k \to \infty$,

$$I \sim W'[c_0u^2\{1 - \Phi_0(v)\}]u^{m+k+2}e^{-k/2}(2\pi)^{1/2}k^{-1},$$

so that (3.9) obtains.

This formula also shows that

$$\begin{aligned}
\int_{-\infty}^z & W'[c_0(m^{1/2} - 2^{-1/2}t)^2\{1 - 2^{1/2}\chi(z)m^{-1/2}\}] \\
& \times \exp\left\{-\frac{1}{2}(m^{1/2} - 2^{-1/2}t)^2\right\}(1 - 2^{-1/2}tm^{-1/2})^{m+4}dt = 0.
\end{aligned}$$

It is easy to verify that

$$\begin{aligned}
&-\frac{1}{2}(m^{1/2} - 2^{-1/2}t)^2 + m\log(1 - 2^{-1/2}tm^{-1/2}) \\
&\quad = -\frac{1}{2}m + tm^{1/2}2^{-1/2} - t^2/4 - tm^{1/2}2^{-1/2} - t^2/4 + o(1) \\
&\quad = -\frac{1}{2}m - t^2/2 + o(1),
\end{aligned}$$

so that

$$\begin{aligned}
0 &= \lim_{m\to\infty} m^{1/2}\int_{-\infty}^z W'[(1 + r_m)(1 - 2^{1/2}tm^{-1/2})\{1 - 2^{1/2}\chi(z)m^{-1/2}\}]e^{-t^2/2}dt \\
&= -W''(1)2^{1/2}\int_{-\infty}^z \{\chi(z) + t\}e^{-t^2/2}dt.
\end{aligned}$$

Formula (3.10) follows now immediately.

Theorem 3.2 shows that the Stein variance estimator corresponds to the maximum likelihood estimator $\max(Y, 0)$ of the positive normal mean. The Brewster-Zidek variance estimator corresponds to the generalized Bayes estimator of the positive normal mean with respect to the "uniform" prior distribution over the positive half-line (cf. Katz (1961)). Both estimators of the positive mean are minimax, i.e. their quadratic risks are bounded by 1. The latter estimator is admissible (Lehmann (1983), pp. 267–268), and the second is not.

The form of the risk of the generalized Bayes estimator resembles this of the Brewster-Zidek estimator: it is a unimodal function which takes its largest (minimax) value at $\Theta = 0$ and tends to 1 as $\Theta \to \infty$. Its minimum equal to 0.584 is attained at $\Theta = 1.08$.

The risk of the maximum likelihood estimator has a different form: it is a monotonically increasing function which takes value 0.5 at $\Theta = 0$ and tends to 1 as $\Theta \to \infty$. It is curious that the inadmissible maximum likelihood estimator provides a larger degree of improvement than the admissible generalized Bayes estimator.

The problems of finding an explicit improvement over the maximum likelihood estimator and of determination of the smallest risk value at a point within the class of all minimax estimators of the positive normal mean apparently are very difficult. Yet they are not only related to the problem of the largest possible improvement over the best equivariant variance estimator, but also to that of the mean vector estimators, as we shall see in the next section.

4. Asymptotic risk behavior of scale-equivariant estimators of the mean vector

In this section we perform an asymptotic analysis of the risk of the mean vector estimators which have the form

$$\delta(X, S) = \{1 - \gamma(V)(m+1)^{-1}\}X \tag{4.1}$$

with a continuous function γ.

This form is motivated by the existing estimators of the multivariate normal mean, namely, Stein estimator δ_{S} with

$$\gamma_{\text{S}}(V) = (k-2)V^2(1-V^2)^{-1} \tag{4.2}$$

and the positive part of this estimator, the James-Stein estimator δ_{JS}, such that

$$1 - (m+1)^{-1}\gamma_{\text{JS}}(V) = \max\{1 - (m+1)^{-1}\gamma_{\text{S}}(V), 0\}. \tag{4.3}$$

Assume that the loss function has the form

$$W(\delta; \mu, \sigma) = w(\|\delta(X, S) - \mu\|^2 k^{-1}\sigma^{-2})$$

with a smooth nonnegative function w such that $w(0) = 0$, $w(1) = 1$.

The risk of estimators (4.1) depends only on $\eta = \mu/\sigma$, so that we can put $\sigma = 1$.

Clearly the risk of the estimator X does not depend on η and as $k \to \infty$

$$E_0 w(\|X\|^2 k^{-1}) \to w(1) = 1,$$

so that the relative risk reduction $\rho(\eta)$ has the form

$$\rho(\eta) = 1 - E_\eta w\{\|\delta(X,S) - \eta\|^2 k^{-1}\}.$$

Assume as in Section 3 that as $k \to \infty$

$$\|\eta\|^2 k^{-1} \to \eta_1$$

and

$$\gamma(vk^{-1/2}) \to \gamma_0(v). \tag{4.4}$$

Then

$$\lim_{k\to\infty} E_\eta w\{\|\delta(X,S) - \eta\|^2 k^{-1}\} = \lim_{k\to\infty} E_\eta w\{\|X - \eta - \tilde{\gamma}(S)X\|^2 k^{-1}\},$$

where

$$\tilde{\gamma}(S) = (m+1)^{-1}\gamma_0\{S(1+\eta_1)^{-1/2}\}.$$

One has

$$\begin{aligned}\|X - \eta - \tilde{\gamma}(S)X\|^2 k^{-1} \\ &= \{1-\tilde{\gamma}(S)\}^2\|X-\eta\|^2 k^{-1} \\ &\quad + \tilde{\gamma}^2(S)\|\eta\|^2 k^{-1} + 2\tilde{\gamma}(S)\{1-\tilde{\gamma}(S)\}\sum_1^k (X_j - \eta_j)\eta_j k^{-1}.\end{aligned}$$

By the strong law of large numbers with probability one

$$\|X - \eta\|^2 k^{-1} \to 1.$$

Also the normal random variable $\sum (X_j - \eta_j)\eta_j k^{-1}$ has zero mean and the variance $\|\eta\|^2 k^{-2}$.

Therefore

$$\lim_{k\to\infty} E_\eta w\{\|X - \eta - \tilde{\gamma}(S)X\|^2 k^{-1}\} = Ew[\{1-\tilde{\gamma}(S)\}^2 + \eta_1\tilde{\gamma}^2(S)].$$

Now we suppose that as $m \to \infty$

$$m^{1/2}\{1 - m^{-1}\gamma_0(m^{1/2} - z2^{-1/2})\} \to 2^{1/2}\kappa(z) \tag{4.5}$$

and

$$m^{1/2}\eta_1 \to 2^{1/2}\Theta$$

with a nonnegative finite Θ.

Then

$$\begin{aligned}&Ew[\{1-\tilde{\gamma}(S)\}^2+\eta_1\tilde{\gamma}^2(S)]\\&\quad\sim Ew[2\kappa^2(\Theta-Z)m^{-1}+2^{1/2}\Theta\{1-2^{1/2}\kappa(\Theta-Z)m^{-1/2}\}^2m^{-1/2}]\\&\quad=\frac{W'(0)2^{1/2}\Theta}{m^{1/2}}+\frac{2W'(0)}{m}E(\kappa-\Theta)^2+\frac{W''(0)-2W'(0)}{m}\Theta^2+o(m^{-1})\end{aligned}$$

with a standard normal Z.

We have proved the following result.

THEOREM 4.1. *Let $\delta(X,S)$ be a scale-equivariant estimator of the multivariate normal mean of the form (4.1) such that as $k\to\infty$ and $m\to\infty$ conditions (4.5) and (4.4) hold. If (3.7) is satisfied, then*

$$\begin{aligned}r(\eta)=1&-\frac{W'(0)2^{1/2}\Theta}{m^{1/2}}-\frac{2W'(0)E\{\kappa(Y)-\Theta\}^2}{m}\\&-\frac{W''(0)-2W'(0)}{m}\Theta^2+o(m^{-1}).\end{aligned}$$

Here Y is a normal random variable with the nonnegative mean Θ and the unit variance.

Theorem 4.1 shows that the estimation problem of the multivariate normal mean is also intimately related to that of scalar positive mean. However in the variance estimation problem the quadratic risk of a positive mean estimator enters the leading term of the asymptotic expansion of the relative risk reduction. For the vector mean estimation this quantity enters only the third term which has the order m^{-1}. This facts explains why better estimators of the multivariate normal mean offer more sizable savings than better variance estimators.

Notice that conditions (4.5) and (4.4) are satisfied for both Stein and James-Stein estimators (4.2) and (4.3).

Indeed in the first case

$$\tilde{\gamma}_S(S)=S^2/(m+1)$$

and

$$\begin{aligned}2^{1/2}\kappa_S(Y)&=\lim_{m\to\infty}m^{1/2}\{1-\tilde{\gamma}_S(m^{1/2}-Y2^{-1/2})\}\\&=\lim_{m\to\infty}m^{1/2}\{1-(m^{1/2}-Y2^{-1/2})^2m^{-1}\}=2^{1/2}Y.\end{aligned}$$

Thus the limiting form of this estimator is not quite reasonable (although unbiased) estimator which is Y itself.

For James-Stein estimator

$$\tilde{\gamma}_{JS}(S)=\min\{1,\tilde{\gamma}_S(S)\}$$

and

$$\kappa_{\mathrm{JS}}(Y) = \max\{0, \kappa_{\mathrm{S}}(Y)\} = \max(0, Y),$$

i.e. the limit of this estimator corresponds to the familiar maximum likelihood estimator of the positive normal mean. Clearly $\max(0, Y)$ is better than Y. This corresponds to the fact that the positive part of the Stein estimator is better than the Stein estimator itself.

The results of the previous section suggest new generalized Bayes estimators of the multivariate normal mean which are analogous to the Brewster-Zidek estimator. However, because of their risk behavior (which is the worst around the origin) they are of less interest than the James-Stein estimator.

Notice that in the case of the known variance an asymptotic analysis of this estimator for large dimensions has been performed by Casella and Hwang (1982). For large sample sizes approximations to the risk functions of a normal covariance matrix were derived by Sugiura and Fujimoto (1982) and Sugiura and Konno (1987).

We conclude this paper with the following remark. Our results are true in a much broader setting than a normal vector X and a chi-squared distributed S^2. In fact, the only facts needed in the proofs of Theorems 3.2 and 4.1 were that with probability one as $k \to \infty$

$$\|X\|^2 k^{-1} \to 1 + \eta_1 \quad \text{if} \quad \|EX\|^2 k^{-1} \to \eta_1$$

and that the asymptotic distribution of $m^{1/2}(S^2 - m)$ as $m \to \infty$ is a normal one.

Acknowledgements

The author is grateful to the National Science Foundation and to the Alexander von Humboldt-Stiftung for their support and to the University of Münster for hospitality.

REFERENCES

Baranchik, A. J. (1970). A family of minimax estimators of the mean of a multivariate normal distribution, *Ann. Math. Statist.*, **41**, 642–656.

Brewster, J. F. and Zidek, J. V. (1974). Improving on equivariant estimators, *Ann. Statist.*, **2**, 21–38.

Casella, G. and Hwang, J. T. (1982). Limit expressions for the risk of James-Stein estimators, *Canad. J. Statist.*, **10**, 305–309.

Efron, B. and Morris, C. N. (1976). Families of minimax estimators of the mean of a multivariate normal distribution, *Ann. Statist.*, **4**, 11–21.

James, W. and Stein, C. (1961). Estimation with quadratic loss, *Proc. Fourth Berkeley Symp. on Math. Statist. Prob.*, Vol. 1, 361–379, Univ. of California Press, Berkeley.

Katz, M. W. (1961). Admissible and minimax estimates of parameters in truncated spaces, *Ann. Math. Statist.*, **32**, 136–142.

Lehmann, E. L. (1983). *Theory of Point Estimation*, Wiley, New York.

Maatta, J. M. and Casella, G. (1990). Developments in decision-theoretic variance estimation, *Statist. Sci.*, **5**, 90–120.

Rukhin, A. L. (1987). How much better are better estimators of a normal variance?, *J. Amer. Statist. Assoc.*, **82**, 925–928.

Rukhin, A. L. and Ananda, M. A. (1992). Risk behavior of variance estimators in multivariate normal distributions, *Statist. Probab. Lett.*, **13**, 159–166.

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.

Sugiura, N. and Fujimoto, M. (1982). Asymptotic risk comparison of improved estimators for normal covariance matrix, *Tsukuba J. Math.*, **6**, 107–126.

Sugiura, N. and Konno, Y. (1987). Risk of improved estimators for generalized variance and precision, *Advances in Multivariate Statistical Analysis* (ed. A. K. Gupta), 353–371, Reidel, Dordrecht.