STATISTICA, anno LIX, n. 2, 1999

ESTIMATION OF THE TRACE OF THE SCALE MATRIX OF A MULTIVARIATE *T*-MODEL UNDER A SQUARED ERROR LOSS

A.H. Joarder, G.K. Beg

1. INTRODUCTION

We consider the estimation of the trace of the scale matrix of the multivariate t-distribution. The trace of the scale matrix in this case gives the total variation of the component variables present in the population, and hence is important in many statistical analyses including principal component analysis.

Unlike the normal distribution, the multivariate t-distribution is fat tailed and hence important in modelling many real world data. It may be mentioned that many authors have observed that the empirical distribution of rates of return of common stocks have relatively thicker tails than those of the normal distribution. Blattberg and Gonedes (1974) assessed the suitability of independent t-distributions for stock return data. After a thorough investigation, Kelejian and Prucha (1985) proved that uncorrelated t-distributions are better able to capture heavy-tailed behavior than independent t-distributions.

The estimation of the trace of the covariance matrix (scale matrix) of the multivariate normal distribution was considered by Olkin and Selliah (1977) under a weighted squared error loss function. The present work is motivated primarily by the work of Dey (1988) who considered the estimation of the trace of the covariance matrix of the multivariate normal distribution under a squared error loss function. He developed estimation strategies by shrinking eigenvalues towards their geometric mean.

In particular, we assume N p-dimensional ($p \geq 2$) random vectors (not necessarily independent) $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_N$ having a joint p.d.f. (probability density function) given by

$$f(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_N) = \frac{|\underline{\Sigma}|^{-N/2}}{C(\nu, Np)\pi^{Np/2}} \left(1 + \frac{1}{\nu}\sum_{j=1}^{N}(\underline{x}_j - \mu)'\underline{\Sigma}^{-1}(\underline{x}_j - \mu)\right)^{-(\nu+Np)/2} \tag{1.1}$$

where the normalizing constant $C(\nu, Np)$ is defined by

The paper is based on a conference paper published by the first author in *Econometrics Conference at Monash University*, 1995, Australia, pp. 467-474.

$$C(\nu, p) = \frac{\nu^{p/2}\Gamma(\nu/2)}{\Gamma((\nu+p)/2)} \tag{1.2}$$

where $\underline{x}_j = (\underline{x}_{1j}, \underline{x}_{2j}, \ldots, \underline{x}_{pj})'$, μ is an unknown $p \times 1$ vector of location parameters, $\underline{\Sigma}$ is a $p \times p$ unknown positive definite matrix of scale parameters and ν (> 4) is assumed to be a known positive constant. Each p-dimensional random vector $\underline{X}_j$ $(j = 1, 2, \ldots, N)$ has a multivariate t-distribution with mean vector μ and covariance matrix $\nu\underline{\Sigma}/(\nu - 2)$, $(\nu > 4)$ and may be denoted by

$$\underline{X}_j \sim T_p\left(\mu, \frac{\nu}{\nu-2}\underline{\Sigma}\right), \qquad \nu > 4, \qquad j = 1, 2, \ldots, N.$$

The random vectors $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_N$ in the model (1.1) are uncorrelated for any $\nu > 4$, but not independent unless ν approaches infinity. The joint p.d.f. in (1.1) represents the multivariate t-model; it has been considered, among others, by Zellner (1976) in the context of stock market problems, and Lange, Little and Taylor (1989).

We now propose a class of estimators for the trace of the scale matrix of the multivariate t-distribution. It may be remarked that we are estimating the trace of the scale matrix instead of the trace of the covariance matrix since the scale matrix $\underline{\Sigma}$ determines the covariance matrix up to a known constant $\nu/(\nu - 2)$. Let $\delta = \mathrm{tr}(\underline{\Sigma})$ be the trace of the scale matrix $\underline{\Sigma}$. In estimating δ by an estimator $\hat{\delta}$, we consider the loss function

$$L(\hat{\delta}, \delta) = (\hat{\delta} - \delta)^2 \tag{1.3}$$

and the risk function

$$R(\hat{\delta}, \delta) = \mathrm{E}[L(\hat{\delta} - \delta)] \tag{1.4}$$

where $\hat{\delta}$ is any estimator of δ. In particular, consider the following two classes of positive estimators of $\delta = \mathrm{tr}(\underline{\Sigma})$:

usual estimator $\tilde{\delta} = c_0 \ \mathrm{tr}(\underline{A})$, (1.5)

proposed estimator $\hat{\delta} = c_0 \ \mathrm{tr}(\underline{A}) - cp \mid \underline{A} \mid^{1/p}$ (1.6)

where c_0 is a known positive constant, c is a constant so that the proposed estimator $\hat{\delta}$ is positive, and $\underline{A} = \sum_{j=1}^{N}(\underline{X}_j - \underline{\overline{X}})(\underline{X}_j - \underline{\overline{X}})'$ is the sample sum of product matrix (Wishart matrix) where $\underline{\overline{X}} = (\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_p)', \overline{X}_i = \sum_{j=1}^{N} X_{ij}/N$, $i = 1, 2, \ldots, p$.

In this paper we prove dominance theorem that the proposed estimator $\hat{\delta}$ dominates the usual estimator $\tilde{\delta}$ of $\delta = \mathrm{tr}(\underline{\Sigma})$ in the sense of having smaller risk i.e.

$$R(\hat{\delta}, \delta) = \mathrm{E}(\hat{\delta} - \delta)^2 < R(\tilde{\delta}, \delta) = \mathrm{E}(\tilde{\delta} - \delta)^2 .$$

Exact expressions for the risk functions of the estimators are derived. Relative Risks of the estimators are compared with some numerical examples.

2. SOME PRELIMINARIES

We need some lemmas on the expectation of the Wishart matrix based on a sample (not necessarily independent observations) from the multivariate t-distribution which will be required in the sequel. The proofs of these lemmas due to Joarder (1998) are adapted here for the sake of completeness.

The p-dimensional random vector $\underline{X}_j$ in model (1.1) can be represented by the scale mixture of the multivariate normal distribution $N_p(\mu, \tau^2 \underline{\Sigma})$ and the distribution of a univariate random variable τ where τ^{-2} has a gamma distribution with mean 1 and variance $2/\nu$ i.e. $\underline{X}_j \mid \tau \sim N(\mu, \tau^2 \underline{\Sigma})$. It follows that given τ, the Wishart matrix having p.d.f in (2.1) has an usual Wishart distribution *i.e.*

$$\underline{A} \mid \tau \sim \mathcal{W}(n,\ \tau^2 \underline{\Sigma}), \qquad n = N - 1. \tag{2.1}$$

The p.d.f. of $\underline{A} \mid \tau$ is given by

$$f(\underline{A} \mid \tau) = \frac{|\tau^2 \underline{\Sigma}|^{-n/2} |\underline{A}|^{(n-p-1)/2}}{2^{np/2}\Gamma_p(n/2)} \exp\left(-\frac{1}{2}\operatorname{tr}(\underline{\Sigma}^{-1}\underline{A}/\tau^2)\right),$$

where $A > 0$, $n = N - 1 \geq p$ and $\Gamma_p(n/2)$ is the generalized gamma function defined by

$$\Gamma_p(\alpha) = \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma((2\alpha - i + 1)/2), \qquad \alpha > (p-1)/2. \tag{2.2}$$

It can be easily shown that

$$\mathrm{E}(|\underline{A}|^k \underline{A} \mid \tau) = 2^k (n+2k) \frac{\Gamma_p(n/2+k)}{\Gamma_p(n/2)} |\tau^2 \underline{\Sigma}|^k \tau^2 \underline{\Sigma}, \qquad n + 2k > 0. \tag{2.3}$$

In view of the mixture representation given by (2.1), it then follows that

$$\begin{aligned}
\mathrm{E}(|\underline{A}|^k \underline{A}) &= \mathrm{E}[\mathrm{E}(|\underline{A}|^k \underline{A} \mid \tau)] \\
&= \mathrm{E}\left[2^k (n+2k) \frac{\Gamma_p(n/2+k)}{\Gamma_p(n/2)} |\tau^2 \underline{\Sigma}|^k \tau^2 \underline{\Sigma}\right] \\
&= 2^k (n+2k) \frac{\Gamma_p(n/2+k)}{\Gamma_p(n/2)} |\underline{\Sigma}|^k \underline{\Sigma}\, \mathrm{E}(\tau^{2kp+2}).
\end{aligned}$$

An exact expression of the above expectation can then be found by noting that for any integer r

$$\mathrm{E}(\tau^r) = (\nu/2)^{r/2} \frac{\Gamma(\nu/2 - r/2)}{\Gamma(\nu/2)}, \qquad \nu > r. \tag{2.4}$$

The result is summarized in the following lemma.

Lemma 2.1. Let $\underline{A}$ have the mixture representation given by (2.1). Then for any real number k and any positive number ν satisfying the conditions $n+2k>0$ and $\nu>2(kp+1)$, the following result holds:

$$\mathrm{E}(|\underline{A}|^k \underline{A}) = \nu^{kp+1}(n/2+k)\frac{\Gamma(\nu/2-kp-1)}{\Gamma(\nu/2)}\frac{\Gamma_p(n/2+k)}{\Gamma_p(n/2)}|\underline{\Sigma}|^k\underline{\Sigma}.$$

It follows from Anderson (1958, p. 161) that for the Wishart matrix $\underline{A}=((a_{ik}))$ satisfying (2.1) we have

$$\mathrm{E}(a_{ii}a_{kk}\mid\tau) = n^2(\tau^2\sigma_{ii})(\tau^2\sigma_{kk})+2n(\tau^2\sigma_{ik})^2 \qquad (i,k=1,\ 2,\ \ldots,\ p)$$

so that

$$\mathrm{E}[(\mathrm{tr}\,\underline{A})^2\mid\tau] = \sum_{i=1}^{p}\mathrm{E}(a_{ii}^2)+2\sum_{i(<k)=1}^{p}\sum_{k=1}^{p}\mathrm{E}(a_{ii}a_{kk})$$

$$= \tau^4\left[n^2\sum_{i=1}^{p}\sigma_{ii}^2+2n\sum_{i=1}^{p}\sigma_{ii}^2+2n^2\sum_{i(<k)=1}^{p}\sum_{k=1}^{p}\sigma_{ii}\sigma_{kk}+4n\sum_{i(<k)=1}^{p}\sum_{k=1}^{p}\sigma_{ik}^2\right].$$

Rearranging, we have

$$\mathrm{E}[(\mathrm{tr}\,\underline{A})^2\mid\tau] = n^2(\mathrm{tr}\ \tau^2\ \underline{\Sigma})^2+2n\,\mathrm{tr}(\tau^2\ \underline{\Sigma})^2\,. \tag{2.5}$$

Recalling the mixture representation of the Wishart matrix given by (2.1), we have

$$\mathrm{E}[(\mathrm{tr}\,\underline{A})^2] = \mathrm{E}[\mathrm{E}[(\mathrm{tr}\,\underline{A})^2\mid\tau]] = \mathrm{E}[n^2(\mathrm{tr}\ \tau^2\ \underline{\Sigma})^2+2n\ \mathrm{tr}(\tau^2\ \underline{\Sigma})^2]$$

and then by the use of (2.4), we have the following lemma.

Lemma 2.2. Let $\underline{A}$ have the mixture representation given by (2.1). Then for $\nu>4$, we have

$$\mathrm{E}[(\mathrm{tr}\,\underline{A})^2] = \frac{n}{(1-2/\nu)(1-4/\nu)}[n(\mathrm{tr}\,\underline{\Sigma})^2+2\,\mathrm{tr}(\underline{\Sigma}^2)].$$

The reader may be referred to Joarder and Ali (1992) for many other useful expectations on Wishart matrix based on the multivariate t-model.

3. THE MAIN RESULTS

The main results are presented in this section in the form of some theorems.

Theorem 3.1. Consider the multivariate t-model given by (1.1). Then the proposed estimator $\hat{\delta}$ defined by (1.6), dominates the usual estimator $\tilde{\delta}=c_0\ \mathrm{tr}(\underline{A})$ in the sense of having smaller risk under the risk function given by (1.4) for any c satisfying

$$d < c < 0 \tag{3.1}$$

$$\text{where } d = \left(c_0 \frac{np+2}{p} - \frac{\nu - 4}{\nu} \right) \frac{\Gamma_p(n/2 + 1/p)}{\Gamma_p(n/2 + 2/p)}, \tag{3.2}$$

$n = N - 1 \geq p$ with $c_0 < (1 - 4/\nu)/(n + 2/p)$,

$$\text{or, } 0 < c < d \tag{3.3}$$

where d is given by (3.2) with $c_0 > (1 - 4/\nu)/(n + 2/p)$.

Proof. Let the risk functions of the two estimators be defined by $R(\tilde{\delta}, \delta; c) = \mathrm{E}(\tilde{\delta} - \delta)^2$ and $R(\hat{\delta}, \delta) = \mathrm{E}(\hat{\delta} - \delta)^2$ respectively. The risk difference of the estimators can be written as

$$R(\hat{\delta}, \delta; c) - R(\tilde{\delta}, \delta) = \mathrm{E}[\hat{\delta}^2 - \tilde{\delta}^2 - 2(\hat{\delta} - \tilde{\delta})\delta]$$

so that by virtue of $\hat{\delta}^2 = \tilde{\delta}^2 - 2cp\tilde{\delta}\,|\underline{A}|^{1/p} + (cp)^2\,|\underline{A}|^{2/p}$, we have

$$R(\hat{\delta}, \delta; c) - R(\tilde{\delta}, \delta) = \mathrm{E}[-2cp\tilde{\delta}\,|\underline{A}|^{1/p} + (cp)^2\,|\underline{A}|^{2/p} - 2(-cp\,|\underline{A}|^{1/p})\delta]$$

$$= p\mathrm{E}[-2c_0c\,|\underline{A}|^{1/p}\,\mathrm{tr}(\underline{A}) + pc^2\,|\underline{A}|^{2/p} + 2c\bar{\xi}\,|\underline{A}|^{1/p}]$$

where $\bar{\xi} = \delta/p$ is the arithmetic mean of the eigenvalues $\xi_1, \xi_2, \ldots, \xi_p$ of the scale matrix $\underline{\Sigma}$.

The k-th moment of the generalized sample variance $|\underline{A}|$ is given by

$$\mathrm{E}(|\underline{A}|^k) = \frac{\Gamma(\nu/2 - kp)}{\nu^{-kp}\Gamma(\nu/2)} \frac{\Gamma_p(n/2 + k)}{\Gamma_p(n/2)} |\underline{\Sigma}|^k, \qquad \nu > 2kp$$

(see Joarder and Ali, 1992). Then by virtue of Lemma 2.1, we have

$$R(\hat{\delta}, \delta; c) - R(\tilde{\delta}, \delta) = -2c_0cp\left[\frac{2(np+2)}{(1 - 2/\nu)(1 - 4/\nu)p} \frac{\Gamma_p(n/2 + 1/p)}{\Gamma_p(n/2)} p\,\bar{\xi}\,\ddot{\xi} \right]$$

$$+ 2p\left[\frac{2}{(1 - 2/\nu)} \frac{\Gamma_p(n/2 + 2/p)}{\Gamma_p(n/2)} \ddot{\xi} \right] p\,\bar{\xi}$$

$$+ (cp)^2\left[\frac{4}{(1 - 2/\nu)(1 - 4/\nu)} \frac{\Gamma_p(n/2 + 2/p)}{\Gamma_p(n/2)} \ddot{\xi}^2 \right]$$

which, after simple algebraic manipulation, reduces to

$$R(\hat{\delta}, \delta; c) - R(\tilde{\delta}, \delta) = \frac{4(p\,\ddot{\xi})^2}{(1 - 2/\nu)(1 - 4/\nu)} \frac{\Gamma_p(n/2 + 2/p)}{\Gamma_p(n/2)} (c^2 - dc\,\bar{\xi}/\ddot{\xi}) \tag{3.4}$$

where $\ddot{\xi}$ is the geometric mean of the eigenvalues of the scale matrix $\underline{\Sigma}$ and d

is given by (3.2). In order that $\hat{\delta}$ dominates $\tilde{\delta}$, it is sufficient to prove that the risk difference $R(\hat{\delta}, \delta; c) - R(\tilde{\delta}, \delta)$ is negative which is true if

$$d(\bar{\xi}/\ddot{\xi}) < c < 0, \quad \text{or} \quad 0 < c < d(\bar{\xi}/\ddot{\xi}).$$

The above conditions involve $\bar{\xi}$ and $\ddot{\xi}$ which are unknown quantities. By the well-known arithmetic mean and geometric mean inequality $\bar{\xi} \geq \ddot{\xi}$, it then follows from (3.4) that $\hat{\delta}$ dominates $\tilde{\delta}$ if $d < c < 0$, or $0 < c < d$ where d is given by (3.2). However,

$$d < 0 \qquad \text{if and only if } c_0 < (1 - 4/\nu)/(n + 2/p)\,,$$

$$\text{while } d > 0 \qquad \text{if and only if } c_0 > (1 - 4/\nu)/(n + 2/p)\,.$$

Hence the proposed estimator $\hat{\delta}$ dominates the usual estimator $\tilde{\delta}$ if c satisfies the conditions mentioned in the theorem.

It may be remarked that if $c_0 = (1 - 4/\nu)/(n + 2/p)$, then $d = 0$. In this case it is seen from (3.4) that the risk difference is positive so that there exists no proposed estimator $\hat{\delta}$ dominating the usual estimator $\tilde{\delta}$. The risk difference vanishes only if $c = 0$ in which case the two estimators coincide.

The risk function of the usual estimator is given by

$$R(\tilde{\delta}, \delta) = \mathrm{E}(\tilde{\delta} - \delta)^2 = c_0^2\, \mathrm{E}[(\mathrm{tr}\, \underline{A})^2] - 2\delta c_0\, \mathrm{tr}(\mathrm{E}(\underline{A})) + \delta^2\,.$$

The expected value of the Wishart matrix is given by

$$\mathrm{E}(\underline{A}) = \mathrm{E}[\mathrm{E}(\underline{A} \mid \tau)] = \mathrm{E}(n\tau^2 \underline{\Sigma}) = n\underline{\Sigma}/(1 - 2/\nu)\,.$$

An exact expression of the risk function of the usual estimator can then be found by using Lemma 2.2. Consequently, the risk function of the proposed estimator follows from (3.4). These results are summarized in the following theorem.

Theorem 3.2. For $\nu > 4$, the risk functions of the usual estimator and the proposed estimator are given by

$$R(\tilde{\delta}, \delta) = \left[\frac{nc_0}{1 - 2/\nu}\left(\frac{nc_0}{1 - 4/\nu} - 2\right) + 1\right](\mathrm{tr}\, \underline{\Sigma})^2 + \frac{2nc_0^2}{(1 - 2/\nu)(1 - 4/\nu)}\,\mathrm{tr}(\underline{\Sigma}^2)$$

and

$$R(\hat{\delta}, \delta; c) = \frac{4p^2 |\underline{\Sigma}|^{2/p}}{(1 - 2/\nu)(1 - 4/\nu)}\,\frac{\Gamma_p(n/2 + 2/p)}{\Gamma_p(n/2)}\left(c^2 - \frac{cd\,\mathrm{tr}(\underline{\Sigma})}{p|\underline{\Sigma}|^{1/p}}\right) + R(\tilde{\delta}, \delta)$$

respectively, where c_0 and c are defined in Theorem 3.1.

4. RELATIVE RISK ANALYSIS WITH NUMERICAL EXAMPLES

To compare the risk of the two classes of estimators $\hat{\delta}$ and $\tilde{\delta}$, we use the Relative Risk (RR) defined by

$$RR(\hat{\delta}:\tilde{\delta};c) = \frac{R(\hat{\delta},\delta;c)}{R(\tilde{\delta},\delta)} \tag{4.1}$$

where $0 \le RR(\hat{\delta}:\ \tilde{\delta};\ c) < 1$ for the choices of c given by Theorem 3.1. The RR in (4.1) is a parabola in c. Theorem 3.1 provides range of values of c where the proposed estimator $\hat{\delta}$ dominates the usual estimator $\tilde{\delta} = c_0\ \mathrm{tr}(\underline{A})$. The Maximum Likelihood Estimator (MLE) and the Unbiased Estimator (UE) of δ are given by $\tilde{\delta} = c_0\ \mathrm{tr}(\underline{A})$ for $c_0 = 1/(n+1)$ and $c_0 = 1/n$ respectively (Fang and Anderson, 1990, p. 208). Thus for a fixed value of n (equivalently of c_0), RR in (4.1) poses a problem of minimization of a parabola in c on a restricted set $\{c : d < c < 0\}$ or $\{c : 0 < c < d\}$. However, neither of these two sets is closed and consequently we may not have an optimal solution for c. Note that the unrestricted minimization occurs at

$$c_m = \frac{\bar{\xi}}{\ddot{\xi}}\frac{d}{2}$$

but this is not usable in practice since neither $\bar{\xi}$ nor $\ddot{\xi}$ is known.

It is interesting to know the behavior of RR for different values of $\bar{\xi}/\ddot{\xi}$. For the choice of $c_0 = 1/(n+1)$, the Relative Risk given by (4.1) compares the proposed estimator with MLE. Whenever $\bar{\xi}/\ddot{\xi} \ge 2$ and $d > 0$ we have $0 < c < d \le c_m$. For example, if $p = 4$, $\nu = 5$, $n = 10$ (so that $d = 0.1841907 > 0$) and

$$\underline{\Sigma} = \begin{pmatrix} 30 & 12 & 7 & 1 \\ 12 & 8 & 3 & 2 \\ 7 & 3 & 4 & -1 \\ 1 & 2 & -1 & 5 \end{pmatrix},$$

then $\bar{\xi}/\ddot{\xi} = 2.229844$ and $c_m = 0.2053583$. The graph of RR versus c for this case is shown in figure 1.

If $1 \le \bar{\xi}/\ddot{\xi} < 2$ and $d > 0$, then $0 < c_m < d$. For example, if $p = 4$, $\nu = 5$, $n = 10$ (so that $d = 0.1841907 > 0$) and

$$\underline{\Sigma} = \begin{pmatrix} 30 & 2 & 2 & 1 \\ 2 & 8 & 3 & 2 \\ 2 & 3 & 4 & 1 \\ 1 & 2 & 1 & 5 \end{pmatrix}, \tag{4.2}$$

then $\bar{\xi}/\ddot{\xi} = 1.590451$ and $c_m = 0.1464731$. The graph of RR versus c for this case is shown in figure 2.

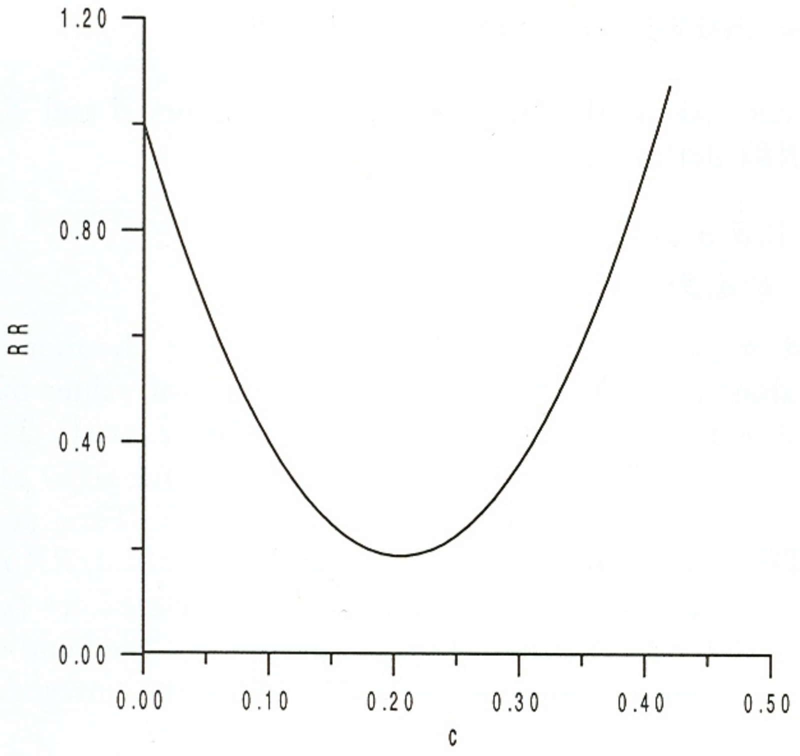

Figure 1 – Relative Risk (RR) versus c ($p = 4$, $\nu = 5$, $n = 10$ and $\bar{\xi}/\ddot{\xi} = 2.229844$).

The behavior of RR versus c for $\underline{\Sigma}$ given by (4.2) with different values of n is shown in figure 3. The proposed estimator performs better, as expected, for increasing n.

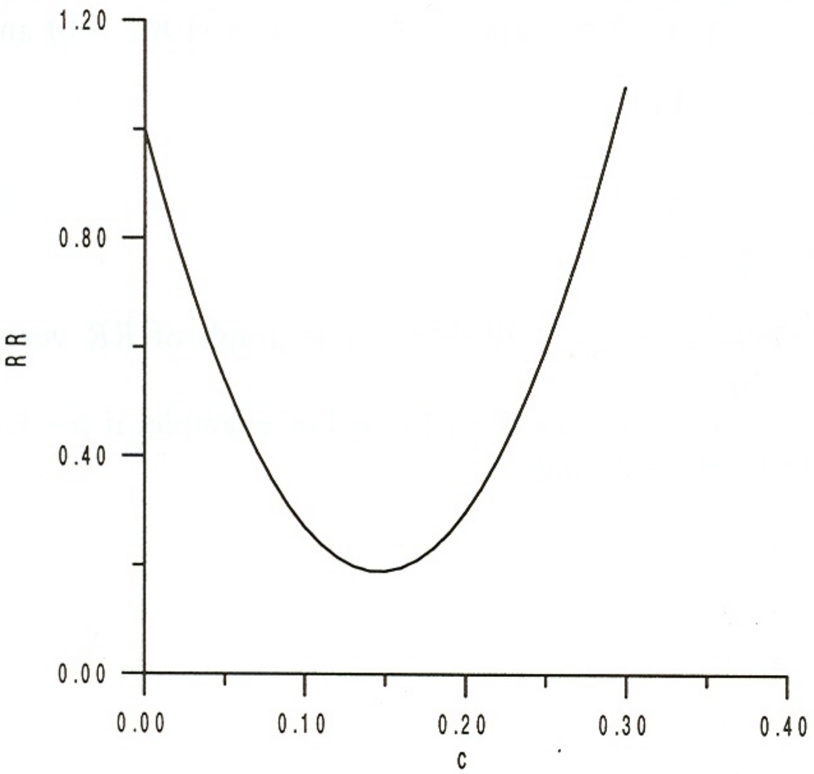

Figure 2 – Relative Risk (RR) versus c ($p = 4$, $\nu = 5$, $n = 10$ and $\bar{\xi}/\ddot{\xi} = 1.590451$).

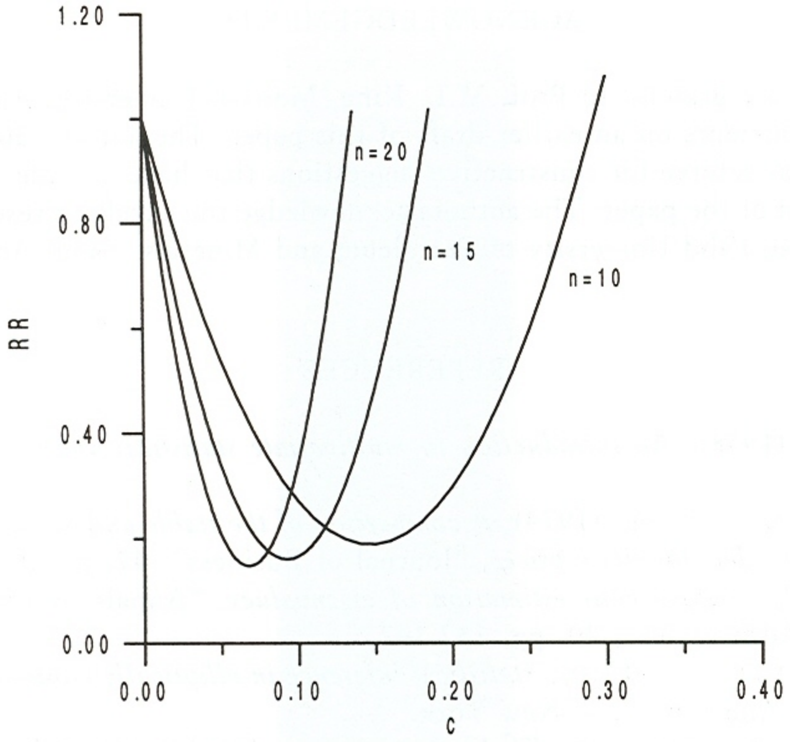

Figure 3 – Relative Risk (RR) versus c ($p = 4$, $\nu = 5$ and $n = 10, 15, 20$).

If we have the prior information that $\bar{\xi}/\ddot{\xi}$ is closer to 1 then $c \approx d/2$ is an optimal choice while if $\bar{\xi}/\ddot{\xi}$ is closer to 2 then $c \approx d$ is an optimal choice. Theorem 3.1 provides choices of c whenever no prior information on $\bar{\xi}/\ddot{\xi}$ is available. It is observed that the behavior of RR for any dimension ($p \geq 2$) is similar. We also note that the behavior of RR for the unbiased case ($c_0 = 1/n$) is similar to that of MLE described above.

5. CONCLUSION

We remark that the gain we get from using the proposed estimator is a smaller mean square error. This is because of the modification made to the usual estimator by introducing a correction term which involves a choice of c depending on $n = N - 1$, p and ν. It is thus easy to calculate the value of c once the sample is selected, as the value of ν is assumed to be known. It may be remarked that as $\nu \to \infty$, Theorem 3.1 of the present paper specializes to Theorem 2.3 of Dey (1988). Since the multivariate t-model given by (1.1) converges to the joint p.d.f. of N independent $N_p(\mu, \Sigma)$ variables as ν approaches infinity, the present work may be viewed as a generalization of the corresponding work of Dey (1988).

Department of Mathematical Sciences
King Fahd University of Petroleum and Minerals
Dhahran, Saudi Arabia

ANWAR H. JOARDER
GHULAM KABIR BEG

ACKNOWLEDGEMENTS

The authors are grateful to Prof. M.L. King, Monash University, Australia for his constructive comments on an earlier draft of this paper. The authors are also grateful to an anonymous referee for constructive suggestions that have considerably improved the presentation of the paper. The authors acknowledge the excellent research facilities available at King Fahd University of Petroleum and Minerals, Saudi Arabia.

REFERENCES

T.W. ANDERSON (1958), *An introduction to multivariate statistical analysis*, Wiley, New York.

R.C. BLATTBERG, N.J. GONEDES (1974), *A comparison of the stable and Student distributions as statistical models for stock prices*, "Journal of Business", 47, pp. 224-280.

D.K. DEY (1988), *Simultaneous estimation of eigenvalues*, "Annals of the Institute of Statistical Mathematics", 40, pp. 137-147.

K.T. FANG, T.W. ANDERSON (1990), *Statistical inference in elliptically contoured and related distributions*, Allerton Press, New York.

A.H. JOARDER (1998), *Some useful Wishart expectations based on the multivariate t-model*, To appear in "Statistical Papers".

A.H. JOARDER, M.M. ALI (1992), *On some generalized Wishart expectations*, "Communications in Statistics - Theory and Methods", 21, pp. 283-294.

H.H. KELEJIAN, I.R. PRUCHA (1985), *Independent or uncorrelated disturbances in linear regression: an illustration of the difference*, "Economic Letters", 19, pp. 35-38.

K.L. LANGE, R.J.A. LITTLE, J.M.G. TAYLOR (1989), *Robust statistical modelling using the t-distribution*, "Journal of the American Statistical Association", 84, pp. 881-896.

I. OLKIN, J.B. SELLIAH (1977), *Estimating covariance in a multivariate normal distribution*, "Statistical Decision Theory and Related Topics II", (eds. S.S. GUPTA and D. MOORE), Academic Press, New York, pp. 313-326.

A. ZELLNER (1976), *Bayesian and non-Bayesian analysis of the regression model with multivariate Student-t error term*, "Journal of American Statistical Association", 71, pp. 400-405 (correction, 71, p. 1000).

RIASSUNTO

Stima della traccia della matrice di scala per il modello t multivariato in presenza di una funzione di perdita quadratica

Si considera il problema della stima della traccia della matrice di scala per il modello t multivariato. La tecnica di stima è sviluppata considerando una funzione di perdita quadratica. Vengono derivate le condizioni che rendono lo stimatore proposto migliore di quello solitamente utilizzato, ed inoltre viene calcolata l'espressione esatta per le funzioni di rischio associate agli stimatori. Sono presenti infine alcuni esempi numerici.

SUMMARY

Estimation of the trace of the scale matrix of a multivariate t-model under a squared error loss

The trace of the scale matrix of the multivariate t-distribution is considered for estimation. The estimation strategy is developed assuming a quadratic loss function. The conditions under which the proposed estimator outperforms the usual estimator are derived. Exact expressions for the risk functions of the estimators are also derived. Numerical examples are considered as well.