

WHITHER JACKKNIFING IN STEIN-RULE ESTIMATION?

by

Pranab Kumar Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1486

May 1985

WHITHER JACKKNIFING IN STEIN-RULE ESTIMATION?

Pranab Kumar Sen

Department of Biostatistics 201H
University of North Carolina
Chapel Hill, NC 27514

Key Words and Phrases: asymptotic risk; bias; inadmissibility; jackknifing; jackknifed estimator of dispersion matrix; Pitman-alternatives; risk-estimation; shrinkage estimators.

ABSTRACT

In multi-parameter estimation, the Stein rule provides minimax and admissible estimators, compromising on their unbiasedness. The jackknifing, on the other hand, primarily aims to reduce the bias of an estimator (without necessarily compromising on its efficacy), providing, at the same time, an estimator of the sampling variance of the estimator as well. In shrinkage estimation (where minimization of the risk is the basic goal), one may wonder, how far the bias-reduction objective of jackknifing incorporates the dual objectives of minimaxity (or admissibility) and estimating the risk of the estimator? A critical review of this basic role of jackknifing in shrinkage estimation is made here. Restricted, semi-restricted and the usual versions of jackknifed shrinkage estimators are considered and their performance characteristics are studied.

1. INTRODUCTION

Let X_1, \dots, X_n be $n (\geq 1)$ independent and identically distributed random vectors (i.i.d.r.v.) with a $p (\geq 1)$ -variate distribution function (d.f.) F having finite (but unknown) mean vector θ and dispersion matrix Σ (assumed to be positive definite). When F is multi-normal, the maximum likelihood estimator (m.l.e.) of θ is the sample mean $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$. However, for $p \geq 3$, Stein (1956) showed that \bar{X}_n is not admissible, and estimators of θ dominating over \bar{X}_n , in risk, known as the *Stein-rule estimators*, have been considered by James and Stein (1961) and a host of other workers; we may refer to Berger (1980) for a nice account of the Stein-rule estimation theory for normal F .

For F not necessarily multi-normal, but admitting finite moments up to an adequate order, the dominance of a Stein-rule estimator has been studied in a more general setup (of U-statistics) by Sen (1984), where the usual Pitman-type (local) alternatives have been adapted to provide a meaningful and clear picture of this asymptotic dominance; for any fixed departure of θ from the pivot, asymptotically, a Stein-rule estimator loses its dominance over the conventional one.

A Stein-rule estimator is generally biased; this bias crops up mainly because of the shrinkage factor. For an arbitrary (and, possibly, biased) estimator, *jackknifing* not only reduces the bias, but also, provides an estimator of the *risk* of the estimator, when the latter is conveniently defined in terms of the sampling dispersion matrix of the estimator. An inherent *reversed martingale* structure of jackknifing (explored systematically by Sen (1977)) provides a clear picture of the asymptotics of jackknifing in the classical case. However, the picture may be quite different in the case of shrinkage estimation, as will be dealt with here. For any fixed alternative, asymptotically, a Stein-rule estimator and its jackknifed version are both equivalent, and neither dominates over the conventional one. However, for Pitman-type (local) alternatives, the situation is quite different. In such a case, the conventional

estimator \bar{X}_n , the related Stein-rule estimator and its jackknifed versions are generally not asymptotically equivalent, and their relative risk pictures clearly convey some asymptotic dominance relations, which will be studied here.

Since the Stein-rule estimator, generally, involves a test statistic (for testing the pivot), in addition to the sample mean vector and covariance matrix, jackknifing may only be applied to the mean vector, or to the test statistic and mean vector, or to all the three sets of random elements. These relate to *restricted*, *semi-restricted* and *unrestricted* jackknifed shrinkage estimators. These versions are all introduced in Section 2. Section 3 deals with their asymptotic risk properties for fixed alternatives. The main results on their asymptotic relative risk properties for Pitman-type (local) alternatives are considered in Section 4. The concluding section is devoted to some general discussions on the role of jackknifing in shrinkage estimation.

2. JACKKNIFED STEIN-RULE ESTIMATORS OF θ

For an estimator $\hat{\theta}_n$ of θ , consider a *quadratic loss function*

$$L(\hat{\theta}_n; \theta) = n(\hat{\theta}_n - \theta)'Q(\hat{\theta}_n - \theta) \quad (2.1)$$

where Q is some given positive-definite (p.d.) matrix. The *risk* of $\hat{\theta}_n$ is then given by

$$\rho_n(\hat{\theta}_n, \theta) = E_{\theta}L(\hat{\theta}_n; \theta) = T_r(Q\Gamma_n), \quad (2.2)$$

where $\Gamma_n = nE\{(\hat{\theta}_n - \theta)(\hat{\theta}_n - \theta)'\}$. A Stein-rule estimator involves a (known) pivot θ_0 and shrinkage of $\hat{\theta}_n$ is made towards θ_0 . Because of the translation-invariance of \bar{X}_n , we may take, without any loss of generality, $\theta_0 = 0$. Then, following Berger et al. (1977), we may consider the following Stein-rule estimator:

$$\hat{\theta}_n^S = (I - c_n d_n T_n^{-2} Q^{-1} S^{-1}) \bar{X}_n, \quad (2.3)$$

where c_n is a positive number (converging to a limit $c: 0 < c < 2(p-2)$), $T_n^2 = n\bar{X}_n' S_n^{-1} \bar{X}_n$ is the test statistic (for the null hypothesis $H_0: \theta = 0$), $d_n = ch_p(QS_n)$ is the smallest characteristic root of

$Q_{\tilde{n}}^S$, $p \geq 3$, and $S_{\tilde{n}} = (n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X}_{\tilde{n}})(X_i - \bar{X}_{\tilde{n}})'$ is an unbiased estimator of Σ . For multi-normal F and suitable $\{c_n\}$, the minimaxity of $\hat{\theta}_{\tilde{n}}^S$ has been established in Berger *et al.* (1977). In our study, we do not assume that F is necessarily multi-normal. Under suitable moment-condition on F , the dominance of $\hat{\theta}_{\tilde{n}}^S$ over $\bar{X}_{\tilde{n}}$, follows from Sen (1984). We rewrite (2.3) as

$$\hat{\theta}_{\tilde{n}}^S = (I - T_n^{-2} A_n) \bar{X}_{\tilde{n}}; A_n = c_n d_n Q_n^{-1} S_n^{-1}, \quad (2.4)$$

and note that three sets of random elements (viz., T_n^2 , A_n and $\bar{X}_{\tilde{n}}$), all based on X_1, \dots, X_n , appear in $\hat{\theta}_{\tilde{n}}^S$. As jackknifing may be used on a part of these random elements (or all), we consider the following versions:

(I) *Restricted jackknifed estimator.* Here, jackknifing is applied only to the part $\bar{X}_{\tilde{n}}$, leading to

$$\hat{\theta}_{\tilde{n}}^{SRJ} = n^{-1} \sum_{i=1}^n \hat{\theta}_{\tilde{n},i}^{SR}, \quad (2.5)$$

where $\hat{\theta}_{\tilde{n},i}^{SR} = n \hat{\theta}_{\tilde{n}}^S - (n-1) \hat{\theta}_{\tilde{n}-1(i)}^{SR}$, and

$$\hat{\theta}_{\tilde{n}-1(i)}^{SR} = (I - T_n^{-2} A_n) \bar{X}_{\tilde{n}-1(i)}, \quad \bar{X}_{\tilde{n}-1(i)} = \frac{n \bar{X}_{\tilde{n}} - X_i}{(n-1)}, \quad (2.6)$$

for $i=1, \dots, n$. It is easy to verify that

$$\hat{\theta}_{\tilde{n}}^{SRJ} = \hat{\theta}_{\tilde{n}}^S, \quad \text{for every } n \geq 2. \quad (2.7)$$

(II) *Semi-restricted jackknifed estimator.* Here, jackknifing is applied to both $\bar{X}_{\tilde{n}}$ and T_n^2 (but not on A_n), so that

$$\hat{\theta}_{\tilde{n}}^{SSJ} = n^{-1} \sum_{i=1}^n \hat{\theta}_{\tilde{n},i}^{SS}, \quad (2.8)$$

where $\hat{\theta}_{\tilde{n},i}^{SS} = n \hat{\theta}_{\tilde{n}}^S - (n-1) \hat{\theta}_{\tilde{n}-1(i)}^{SS}$, and

$$\hat{\theta}_{\tilde{n}-1(i)}^{SS} = (I - T_{n-1(i)}^{-2} A_n) \bar{X}_{\tilde{n}-1(i)}, \quad i=1, \dots, n, \quad (2.9)$$

where $T_{n-1(i)}^2$ is defined for $(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$, for $i=1, \dots, n$. In general, $\hat{\theta}_{\tilde{n}}^{SSJ}$ and $\hat{\theta}_{\tilde{n}}^S$ are not the same.

(III) *Unrestricted jackknifed estimator:*

$$\hat{\theta}_{\tilde{n}}^{SUJ} = n^{-1} \sum_{i=1}^n \hat{\theta}_{\tilde{n},i}^{SU}, \quad (2.10)$$

where $\hat{\theta}_{n,i}^{SU} = n\hat{\theta}_n^S - (n-1)\hat{\theta}_{n-1(i)}^{SU}$, and

$$\hat{\theta}_{n-1(i)}^{SU} = (I - T_{n-1(i)}^{-2})\bar{X}_{n-1(i)}, \quad i \leq i \leq n, \quad (2.11)$$

with the $A_{n-1(i)}$ defined on $(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$ for $1 \leq i \leq n$.

Again, $\hat{\theta}_n^S$ and $\hat{\theta}_n^{SUJ}$ are not, generally, the same.

Side by side, we introduce the jackknifed dispersion matrix estimators:

$$V_{\sim n}^{SRJ} = n^{-1} \sum_{i=1}^n (\hat{\theta}_{n,i}^{SR} - \hat{\theta}_n^{SRJ})(\hat{\theta}_{n,i}^{SR} - \hat{\theta}_n^{SRJ})', \quad (2.12)$$

$$V_{\sim n}^{SSJ} = n^{-1} \sum_{i=1}^n (\hat{\theta}_{n,i}^{SS} - \hat{\theta}_n^{SSJ})(\hat{\theta}_{n,i}^{SS} - \hat{\theta}_n^{SSJ})', \quad (2.13)$$

$$V_{\sim n}^{SUJ} = n^{-1} \sum_{i=1}^n (\hat{\theta}_{n,i}^{SU} - \hat{\theta}_n^{SUJ})(\hat{\theta}_{n,i}^{SU} - \hat{\theta}_n^{SUJ})'. \quad (2.14)$$

Our main interest lies in the study of these jackknifed shrinkage estimators and their estimated dispersion matrices.

3. FIXED ALTERNATIVE ASYMPTOTICS FOR JACKKNIFING

It follows from Sen (1984) that for any $\theta \neq 0$, the estimators $\bar{X}_{\sim n}$ and $\hat{\theta}_{\sim n}^S$ are asymptotically risk-equivalent in the sense that

$$E\{T_n^{-2} \mid \theta \neq 0\} \rightarrow 0 \text{ as } n \rightarrow \infty$$

$$\Rightarrow \lim_{n \rightarrow \infty} \{\rho_n(\hat{\theta}_{\sim n}^S, \theta) / \rho_n(\bar{X}_{\sim n}, \theta)\} = 1. \quad (3.1)$$

Even without the negative-moment condition on T_n^2 , whenever F has finite 4th order moments, one has

$$n^{\frac{1}{2}} \|\hat{\theta}_{\sim n}^S - \bar{X}_{\sim n}\| \rightarrow 0 \text{ a.s., as } n \rightarrow \infty \quad (\forall \theta \neq 0). \quad (3.2)$$

By virtue of (2.7), (3.1)-(3.2) also hold for $\{\hat{\theta}_{\sim n}^{SRJ}\}$. Note that $\bar{X}_{\sim n} = n^{-1} \sum_{i=1}^n \bar{X}_{n-1(i)}$. Hence, by (2.3), (2.8) and (2.9),

$$\begin{aligned} \hat{\theta}_{\sim n}^{SSJ} &= \hat{\theta}_{\sim n}^S - (n-1)A_{\sim n} \{T_n^{-2} \bar{X}_{\sim n} - n^{-1} \sum_{i=1}^n T_{n-1(i)}^{-2} \bar{X}_{n-1(i)}\} \\ &= \hat{\theta}_{\sim n}^S - (n-1)T_n^{-2} \{A_{\sim n} (n^{-1} \sum_{i=1}^n \{1 - T_{n-1(i)}^{-2}\} \bar{X}_{n-1(i)})\}. \end{aligned} \quad (3.3)$$

Let c_n be the sigma-field generated by the unordered collection $\{X_1, \dots, X_n\}$ and X_{n+1}, X_{n+2}, \dots , $n \geq 1$, so that c_n is monotone

nonincreasing. Also, let $F_n = F(\bar{X}_k, S_k; k \geq n)$ be the σ -field generated by $\{\bar{X}_k, S_k; k \geq n\}$, so that $F_n \subset C_n$ and F_n is also monotone nonincreasing. Then

$$\begin{aligned} n^{-1} \sum_{i=1}^n \{1 - T_n^2 T_{n-1}^{-2}(i)\} \bar{X}_{n-1}(i) \\ = E\{(1 - T_{n-1}^{-2} T_n^2) \bar{X}_{n-1} \mid C_n\}, \quad n \geq 2 \end{aligned} \quad (3.4)$$

Thus, we have

$$\begin{aligned} (n-1) T_n^{-2} \{A_n (n^{-1} \sum_{i=1}^n \{1 - T_n^2 T_{n-1}^{-2}(i)\} \bar{X}_{n-1}(i))\} \\ = (n-1) T_n^{-2} \{A_n \bar{X}_n E\{(1 - T_{n-1}^{-2} T_n^2) \mid C_n\} \\ + A_n E\{(1 - T_{n-1}^{-2} T_n^2) (\bar{X}_{n-1} - \bar{X}_n) \mid C_n\}\}. \end{aligned} \quad (3.5)$$

Now, for every $\theta \neq 0$, as $n \rightarrow \infty$,

$$\begin{aligned} (n-1) T_n^{-2} &\rightarrow \{\theta' \Sigma^{-1} \theta\}^{-1} (< \infty) \text{ a.s.}, \\ \bar{X}_n &\rightarrow \theta \neq 0 \text{ a.s.}, \end{aligned} \quad (3.6)$$

while, independently of θ ,

$$A_n \rightarrow c \delta Q^{-1} \Sigma^{-1} \text{ a.s.}, \text{ as } n \rightarrow \infty, \quad (3.7)$$

where $\delta = \text{ch}_p(Q\Sigma)$. Further, note that

$$T_n^2 = T(\bar{X}_n, S_n), \quad n \geq 2 \quad (3.8)$$

where \bar{X}_n and S_n are both U-statistics and $T(\cdot)$ is a smooth function. Hence, proceeding as in Sen (1977) [See his (3.20)], it follows that as $n \rightarrow \infty$,

$$E\{(1 - T_n^2 T_{n-1}^{-2}) \mid C_n\} = 0 (n^{-1+\epsilon}) \text{ a.s.}, \quad (3.9)$$

for every $\epsilon > 0$ (and we let $0 < \epsilon < \frac{1}{2}$). Similarly,

$$E\{(1 - T_n^2 T_{n-1}^{-2}) (\bar{X}_{n-1} - \bar{X}_n) \mid C_n\} = 0 (n^{-1+\epsilon}) \text{ a.s.}, \text{ as } n \rightarrow \infty. \quad (3.10)$$

Thus, the right hand side (rhs) of (3.5) is $O(n^{-\frac{1}{2}+\epsilon})$ a.s., as $n \rightarrow \infty$, and this ensures that for every (fixed) $\theta \neq 0$,

$$n^{\frac{1}{2}} \|\hat{\theta}_{n-1}^{SSJ} - \hat{\theta}_n^S\| \rightarrow 0 \text{ a.s.}, \text{ as } n \rightarrow \infty. \quad (3.11)$$

Finally, note that $d_n = \text{ch}_p(Q^{-1} S_n^{-1})$ is a 'smooth' function of S_n ,

so that by appeal to Sen (1977), we have

$$\max_{1 \leq i \leq n} ||A_{\sim n-1}(i) - A_{\sim n}|| = O(n^{-\frac{1}{2}}) \text{ a.s., as } n \rightarrow \infty, \quad (3.12)$$

and therefore, proceeding as in (3.4) through (3.10) and using (2.4), (2.10) and (2.11), we obtain that for every (fixed) $\theta \neq 0$,

$$n^{\frac{1}{2}} ||\hat{\theta}_{\sim n}^{\text{SUJ}} - \hat{\theta}_{\sim n}^{\text{S}}|| \rightarrow 0 \text{ a.s., as } n \rightarrow \infty. \quad (3.13)$$

It may be noted that (3.6) plays a basic role in this context. If $\theta = 0$ (or is "close to" 0 in the Pitman-sense), then (3.6) may not hold, and, as a result, (3.11) or (3.13) may not hold. We shall comment more on it in the next section.

Thus, if θ differs from the pivot (here, 0) by any fixed amount, jackknifing may not lead to any significant change to $\hat{\theta}_{\sim n}^{\text{S}}$, which is also asymptotically equivalent to $\bar{X}_{\sim n}$. In such a case, it also follows (on the same line as in Sen (1977)) that as $n \rightarrow \infty$,

$$V_{\sim n}^{\text{SRJ}} \xrightarrow{P} \Sigma, \quad V_{\sim n}^{\text{SSJ}} \xrightarrow{P} \Sigma \quad \text{and} \quad V_{\sim n}^{\text{SUJ}} \xrightarrow{P} \Sigma, \quad (3.14)$$

so that $T_r(QV_{\sim n})$ converges to $T_r(Q\Sigma)$, in probability, for each of the three jackknifed versions. In view of the fact that $S_{\sim n} \xrightarrow{P} \Sigma$ (even under the second moment condition only), and, there is no gain in the asymptotic risk-efficiency due to shrinkage estimation or jackknifing on the top of that, there seems to be very little ground in advocating jackknifing in shrinkage estimation in the case where $\theta = 0$ and n is large. This asymptotic picture also provides finite sample justifications when θ is away from the pivot. However, for θ "close to" the pivot, we may derive more meaningful results through considerations of Pitman-type alternatives, as will be done in Section 4.

4. ASYMPTOTICS FOR PITMAN-TYPE ALTERNATIVES

For Pitman-type (local) alternatives, the dominance of the Stein-rule estimator over the classical one follows directly from Sen (1984). It is therefore of interest to study the performance characteristics of the jackknifed versions under the same setup, and, thereby, to examine the role of jackknifing in shrinkage

estimation.

We conceive of a triangular array $\{X_{n1}, \dots, X_{nn}; n \geq 1\}$ of row-wise i.i.d.r.v.'s, where X_{ni} has the d.f. $F(x - n^{-\frac{1}{2}}\lambda)$, for some (fixed) $\lambda \in E^p$, $p \geq 3$, and F has the mean vector 0 and a p.d. dispersion matrix Σ ; both λ and Σ are unknown. We denote this sequence of Pitman-type alternatives by $\{K_n\}$. It follows from Sen (1984) that under $\{K_n\}$,

$$\rho(\lambda) = \lim_{n \rightarrow \infty} \rho_n(\bar{X}; \lambda) = T_r(Q\Sigma), \quad \forall \lambda \in E^p. \quad (4.1)$$

Further, if we assume that under $\{K_n\}$

$$d_n T_n^{-2} \text{ converges in } L_1\text{-norm to } \delta \chi_{p,\Delta}^{-2} \quad (4.2)$$

where $\chi_{p,\Delta}^2$ stands for a positive r.v. having the noncentral chi-square distribution with p degrees of freedom (DF) and noncentrality parameters $\Delta = \lambda' \Sigma^{-1} \lambda$ (and $\delta = \text{ch}_p(Q\Sigma)$), then under $\{K_n\}$,

$$\begin{aligned} \rho^S(\lambda) &= \lim_{n \rightarrow \infty} \rho_n(\hat{\theta}_n^S, \lambda) \\ &= T_r(Q\Sigma) - 2c\delta E\{\chi_{p+2,\Delta}^{-2}\} \\ &\quad + c^2 \delta^2 \{\Delta^* E(\chi_{p+4,\Delta}^{-4}) + T_r(Q^{-1} \Sigma^{-1}) E(\chi_{p+2,\Delta}^{-4})\}, \end{aligned} \quad (4.3)$$

where $\Delta^* = \lambda' \Sigma^{-1} Q^{-1} \Sigma^{-1} \lambda$ and $c = \lim_{n \rightarrow \infty} c_n$; $0 < c < 2(p-2)$ [See Section 3 of Sen and Saleh (1985)]. Even without the L_1 -convergence result in (4.2), one may consider the asymptotic distribution of $n^{\frac{1}{2}}(\hat{\theta}_n^S - n^{-\frac{1}{2}}\lambda)$ (under K_n) and compute the risk from this asymptotic distribution; if this is termed the *asymptotic distributional risk* (ADR), then we conclude that (4.3) holds for the ADR of $\hat{\theta}_n^S$, without requiring (4.2). For this, finite 4th order moments of F suffice. Now, the rhs of (4.3) is less than $T_r(Q\Sigma) = \rho(\lambda)$, for every $c \in (0, 2(p-2))$ and every finite λ : $\Delta < \infty$. As Δ increases, $\rho^S(\lambda)$ converges to $\rho(\lambda)$.

Given this picture, it is of interest to study the effect of jackknifing on the ADR (or the asymptotic risk). Also, there is a more important question: Can $\rho^S(\lambda)$ (or the other risks) be

estimated consistently by jackknifing?

First, consider $\hat{\theta}_n^{SRJ}$. By (2.7), we conclude that (4.3) holds for this restricted jackknifed version as well. Further, by (2.5), (2.6) and (2.12), we have

$$\begin{aligned} V_n^{SRJ} &= \frac{n}{n-1} (I - T_n^{-2} A_n) S_n (I - T_n^{-2} A_n)' \\ &= \frac{n}{n-1} \{ S_n - 2c_n d_n T_n^{-2} Q^{-1} + c_n^2 d_n^2 T_n^{-4} Q^{-1} S_n^{-1} Q^{-1} \} \end{aligned} \quad (4.4)$$

Note that under $\{K_n\}$,

$$S_n \xrightarrow{p} \Sigma, \quad d_n \xrightarrow{p} \delta; \quad T_n^2 \xrightarrow{D} \chi_{p,\Delta}^2; \quad c_n \rightarrow c. \quad (4.5)$$

Therefore, by (4.4) and (4.5)

$$V_n^{SRJ} \xrightarrow{D} \Sigma - 2c\delta \chi_{p,\Delta}^{-2} Q^{-1} + c^2 \delta^2 \chi_{p,\Delta}^{-4} Q^{-1} \Sigma^{-1} Q^{-1}, \quad (4.6)$$

and, as a result,

$$T_r(QV_n^{SRJ}) \xrightarrow{D} T_r(Q\Sigma) - 2pc\delta \chi_{p,\Delta}^{-2} + c^2 \delta^2 \chi_{p,\Delta}^{-4} T_r(\Sigma^{-1} Q^{-1}), \quad (4.7)$$

and the rhs of (4.7) is a nondegenerate r.v. for every $\lambda: \Delta < \infty$.

Comparing (4.3) and (4.7), we conclude that whereas the restricted jackknifed estimator $\hat{\theta}_n^{SRJ}$ agrees with $\hat{\theta}_n^S$, the allied jackknifed dispersion matrix, V_n^{SRJ} , fails to provide a consistent estimator of the asymptotic risk (or ADR) $\rho^S(\lambda) = \rho^{SRJ}(\lambda)$. In passing, we may remark that for $Q = \Sigma^{-1}$, ($\delta = 1$), we have the rhs of (4.7) equal to $p(1 - c\chi_{p,\Delta}^{-2})^2$, and this may even exceed p whenever $c\chi_{p,\Delta}^{-2} > 2$ (which has a positive probability).

Consider next the case of $\hat{\theta}_n^{SSJ}$. We rewrite (3.3) as

$$\begin{aligned} \hat{\theta}_n^{SSJ} &= \hat{\theta}_n^S + (n-1) A_n E\{(T_{n-1}^{-2} \bar{X}_{n-1} - T_n^{-2} \bar{X}_n) \mid c_n\} \\ &= \hat{\theta}_n^S + (n-1) E\{(T_{n-1}^{-2} - T_n^{-2}) \mid c_n\} A_n \bar{X}_n \\ &\quad - A_n E\{(T_{n-1}^{-2} - T_n^{-2})(X_n - \bar{X}_n) \mid c_n\}. \end{aligned} \quad (4.8)$$

Note that $S_{n-1} = (n-2)^{-1}(n-1)[S_n - \frac{n}{(n-1)^2} (X_n - \bar{X}_n)(X_n - \bar{X}_n)']$, so that

$$S_{n-1}^{-1} = \frac{n-2}{n-1} S_n^{-1} \left\{ I + \frac{(n-1)^{-2} n (\tilde{X}_n - \bar{X}_n) (\tilde{X}_n - \bar{X}_n)' S_n^{-1}}{1 - (n-1)^{-2} n (\tilde{X}_n - \bar{X}_n)' S_n^{-1} (\tilde{X}_n - \bar{X}_n)} \right\} \quad (4.9)$$

Also, $\bar{X}_{n-1} = \bar{X}_n - (n-1)^{-1} (\tilde{X}_n - \bar{X}_n)$. Hence, we have

$$\begin{aligned} T_{n-1}^2 &= (n-1) T_{n-1} (S_{n-1}^{-1} \bar{X}_{n-1} \bar{X}_{n-1}') \\ &= \frac{n-2}{n} T_n^2 - \frac{2(n-2)}{(n-1)} \bar{X}_n' S_n^{-1} (\tilde{X}_n - \bar{X}_n) + \frac{n-2}{(n-1)^2} (\tilde{X}_n - \bar{X}_n)' S_n^{-1} (\tilde{X}_n - \bar{X}_n) \\ &\quad + \frac{n(n-2)}{(n-1)^2} \left\{ 1 - \frac{n}{(n-1)^2} (\tilde{X}_n - \bar{X}_n)' S_n^{-1} (\tilde{X}_n - \bar{X}_n) \right\}^{-1} \\ &\quad \cdot \left\{ (\bar{X}_n' S_n^{-1} (\tilde{X}_n - \bar{X}_n))^2 + \frac{1}{(n-1)^2} ((\tilde{X}_n - \bar{X}_n)' S_n^{-1} (\tilde{X}_n - \bar{X}_n))^2 \right. \\ &\quad \left. - \frac{2}{n-1} ((\bar{X}_n' S_n^{-1} (\tilde{X}_n - \bar{X}_n)) ((\tilde{X}_n - \bar{X}_n)' S_n^{-1} (\tilde{X}_n - \bar{X}_n))) \right\}. \quad (4.10) \end{aligned}$$

At this stage, we note that as $n \rightarrow \infty$,

$$\|A_{\tilde{X}_n}\| = o_p(1) \text{ a.s.}, \quad \|S_{\tilde{X}_n}\| = o_p(1) \text{ a.s.}, \quad (4.11)$$

$$\max_{1 \leq i \leq n} |\tilde{X}_i - \bar{X}_n| = o_p(n^{\frac{1}{4}}) \text{ a.s.}, \quad (4.12)$$

(where we assume the finiteness of 4th order moments of F), and under $\{K_n\}$ (as well as $H_0: \lambda = 0$), $n^{\frac{1}{2}} \|A_{\tilde{X}_n}\| = o_p(1)$. Therefore, we obtain (through some standard steps) that

$$\begin{aligned} T_{n-1}^{-2} &= \frac{n}{n-2} T_n^{-2} \left[1 + \frac{2n}{(n-1)T_n^2} (\tilde{X}_n - \bar{X}_n)' S_n^{-1} \bar{X}_n - \right. \\ &\quad \left. \frac{n}{(n-1)^2 T_n^2} (\tilde{X}_n - \bar{X}_n)' S_n^{-1} (\tilde{X}_n - \bar{X}_n) + \frac{4n^2}{(n-1)^2 T_n^4} ((\tilde{X}_n - \bar{X}_n)' S_n^{-1} \bar{X}_n)^2 \right. \\ &\quad \left. - \frac{n^2}{(n-1)^2 T_n^2} \left\{ 1 + \frac{n}{(n-1)^2} (\tilde{X}_n - \bar{X}_n)' S_n^{-1} (\tilde{X}_n - \bar{X}_n) \right\} \right. \\ &\quad \left. \left\{ (\bar{X}_n' S_n^{-1} (\tilde{X}_n - \bar{X}_n))^2 + \frac{1}{(n-1)^2} ((\tilde{X}_n - \bar{X}_n)' S_n^{-1} (\tilde{X}_n - \bar{X}_n))^2 \right. \right. \\ &\quad \left. \left. - \frac{2}{n-1} (\bar{X}_n' S_n^{-1} (\tilde{X}_n - \bar{X}_n)) ((\tilde{X}_n - \bar{X}_n)' S_n^{-1} (\tilde{X}_n - \bar{X}_n)) \right\} \right] + o_p\left(\frac{1}{n}\right). \quad (4.13) \end{aligned}$$

Consequently, we have, under $\{K_n\}$, as $n \rightarrow \infty$

$$(n-1) E\{(T_{n-1}^{-2} - T_n^{-2}) \mid c_n\} = T_n^{-2} - (p-4)T_n^{-4} + o_p(1), \quad (4.14)$$

$$E\{(T_{n-1}^{-2} - T_n^{-2})(\tilde{x}_n - \bar{\tilde{x}}_n) \mid c_n\} = 2T_n^{-4} \bar{\tilde{x}}_n + o_p(n^{-\frac{1}{2}}). \quad (4.15)$$

By (4.8), (4.14) and (4.15), we obtain that under $\{K_n\}$,

$$\begin{aligned} \hat{\tilde{\theta}}^{SSJ} &= \hat{\tilde{\theta}}^S + T_n^{-2}(1 - (p-2)T_n^{-2}) A_n \bar{\tilde{x}}_n + o_p(n^{-\frac{1}{2}}) \\ &= [I - (p-2)T_n^{-4} A_n] \bar{\tilde{x}}_n + o_p(n^{-\frac{1}{2}}). \end{aligned} \quad (4.16)$$

Thus, here $\hat{\tilde{\theta}}^S$ and $\hat{\tilde{\theta}}^{SSJ}$ are not asymptotically equivalent, although $\hat{\tilde{\theta}}^{SSJ}$ acts like a shrinkage estimator (with $A_n T_n^{-2}$ being replaced by $(p-2)T_n^{-4} A_n$). If we strengthen (4.2) to

$$d_n T_n^{-4} \text{ converges in } L_1\text{-norm to } \delta_{\chi_{p,\Delta}^{-4}}, \quad (4.17)$$

then proceeding as in Sen (1984), we have, under $\{K_n\}$,

$$\begin{aligned} \rho^{SSJ}(\lambda) &= \lim_{n \rightarrow \infty} \rho_n(\hat{\tilde{\theta}}^{SSJ}, \lambda) \\ &= \text{Tr}(\underline{Q}\underline{\Sigma}) - 2(p-2)c\delta E\{\chi_{p,\Delta}^{-2}\} + 2(p-2)c\delta E\{(\underline{W}'\underline{W})^{-2}\underline{W}'\underline{W}\} \\ &\quad + c^2\delta^2(p-2)^2 E\{(\underline{W}'\underline{W})^{-4}\underline{W}'\underline{A}\underline{W}\}. \end{aligned} \quad (4.18)$$

where $c = \lim_{n \rightarrow \infty} c_n$, $\delta = \text{ch}_p(\underline{Q}\underline{\Sigma})$, $\underline{W} \sim N(\underline{W}, I)$, $\underline{w} = \underline{\Sigma}^{-\frac{1}{2}}\lambda$, and

$\underline{A} = \underline{\Sigma}^{-\frac{1}{2}}\underline{Q}^{-1}\underline{\Sigma}^{-\frac{1}{2}}$. For the ADR result, (4.17) is not needed. Now, by the Stein-identity [viz., Appendix B of Judge and Bock (1978)],

$$\begin{aligned} E\{(\underline{W}'\underline{W})^{-2}\underline{W}'\underline{W}\} &= \Delta E(\chi_{p+2,\Delta}^{-4}) \\ &= E(\chi_{p,\Delta}^{-2}) - (p-4) E(\chi_{p,\Delta}^{-4}); \end{aligned} \quad (4.19)$$

$$\Delta E(\chi_{p+4,\Delta}^{-8}) = E(\chi_{p,\Delta}^{-6}) - pE(\chi_{p+2,\Delta}^{-8}); \quad (4.20)$$

$$E\{(\underline{W}'\underline{W})^{-4}\underline{W}'\underline{A}\underline{W}\} = \text{Tr}(\underline{Q}^{-1}\underline{\Sigma}^{-1}) E(\chi_{p+2,\Delta}^{-8}) + \Delta^* E(\chi_{p+4,\Delta}^{-8}), \quad (4.21)$$

where $\Delta^* = \lambda' \underline{\Sigma}^{-1} \underline{Q}^{-1} \underline{\Sigma}^{-1} \lambda$, and in order that (4.21) is well defined,

we need $p \geq 7$ (compared to $p \geq 3$ for $\hat{\tilde{\theta}}^S$). Note that $\delta\Delta^*/\Delta \leq \delta \text{ch}_1(\underline{Q}^{-1}\underline{\Sigma}^{-1}) = 1$ and $\delta \text{Tr}(\underline{Q}^{-1}\underline{\Sigma}^{-1}) = \text{Tr}(\underline{Q}^{-1}\underline{\Sigma}^{-1})/\text{ch}_1(\underline{Q}^{-1}\underline{\Sigma}^{-1}) \leq p$.

Thus, from (4.18) through (4.21), we obtain that

$$\begin{aligned} \rho^{SSJ}(\lambda) &\leq \text{Tr}(\underline{Q}\underline{\Sigma}) - (p-2)c\delta\{2(p-4)E(\chi_{p,\Delta}^{-4}) - c(p-2)E(\chi_{p,\Delta}^{-6})\} \\ &\leq \text{Tr}(\underline{Q}\underline{\Sigma}) \text{ if } c < \frac{2(p-4)}{(p-2)} \frac{E(\chi_{p,\Delta}^{-4})}{E(\chi_{p,\Delta}^{-6})} \end{aligned} \quad (4.22)$$

Since this is to hold for all $\Delta \geq 0$, we have

$$0 < c < \frac{2(p-4)(p-6)}{(p-2)}, p \geq 7 \implies \rho^{SSJ}(\underline{\lambda}) \leq \rho(\underline{\lambda}), \forall \underline{\lambda}. \quad (4.23)$$

Thus, for the (asymptotic) dominance of the semi-restricted jackknifed estimator, we not only need p to be ≥ 7 , but also, c has to be restricted to a smaller domain (as $(p-4)(p-6)/(p-2) < (p-6) < (p-2)$). Also, in (4.3), ideally, we take $c = p-2$, while in (4.22) we have a smaller value for the optimal c ; this leads to (generally) $\rho^{SSJ}(\underline{\lambda}) > \rho^S(\underline{\lambda}), \forall \underline{\lambda} \in E^p$. There also remains the question: Is it worthy to reduce the bias of a shrinkage estimator at the cost of its dominance?

Let us denote by $\eta_{ni} = T_{n-1}^{-2}(i) - T_n^{-2}$, $\delta_{ni} = \bar{X}_{n-1}(i) - \bar{X}_n$, $i=1, \dots, n$, and let $\bar{\eta}_n = n^{-1} \sum_{i=1}^n \eta_{ni}$. Note that $\sum_{i=1}^n \delta_{ni} = 0$. Then, using (4.11) through (4.15), it follows by some standard steps that under $\{K_n\}$ and the assumed regularity conditions,

$$n^{-1} \sum_{i=1}^n (\eta_{ni} - \bar{\eta}_n)^2 \stackrel{p}{\sim} 4n^{-1} T_n^{-6} + o_p(n^{-1}) \quad (4.24)$$

$$n^{-1} \sum_{i=1}^n (\eta_{ni} - \bar{\eta}_n)(X_{ni} - \bar{X}_n) \stackrel{p}{\sim} 2T_n^{-4} \bar{X}_n + o_p(1). \quad (4.25)$$

Therefore, using (2.13), (2.8), (2.9), (3.3), and (4.13)-(4.15), along with (4.24) and (4.25), we obtain by some routine steps that under $\{K_n\}$ and the assumed regularity conditions,

$$\begin{aligned} V_n^{SSJ} &\stackrel{p}{\sim} V_n^{SRJ} + 4T_n^{-6} c^2 \delta^2 Q^{-1} S_n^{-1} \bar{X}_n \bar{X}_n' S_n^{-1} Q^{-1} \\ &\quad + 4T_n^{-4} c \delta Q^{-1} S_n^{-1} \bar{X}_n \bar{X}_n' \\ &\quad - 4T_n^{-6} c^2 \delta^2 Q^{-1} S_n^{-1} \bar{X}_n \bar{X}_n' S_n^{-1} Q^{-1} + o_p(1) \\ &\stackrel{p}{\sim} S_n + 4c \delta T_n^{-4} Q^{-1} S_n^{-1} \bar{X}_n \bar{X}_n' - 2c \delta T_n^{-2} Q^{-1} \\ &\quad + c^2 \delta^2 T_n^{-4} Q^{-1} S_n^{-1} Q^{-1}. \end{aligned} \quad (4.26)$$

As a result, we have

$$\begin{aligned}
\text{Tr}(\underline{Q}\underline{V}^{\text{SSJ}}) &\stackrel{p}{\sim} \text{Tr}(\underline{Q}\underline{S}_n) + 4c\delta T_n^{-2} - 2c\delta p T_n^{-2} \\
&\quad + c^2\delta^2 T_n^{-4} \text{Tr}(\underline{Q}^{-1}\underline{S}_n^{-1}) \\
\stackrel{D}{\rightarrow} &\text{Tr}(\underline{Q}\underline{\Sigma}) - 2c\delta(p-2)\chi_{p,\Delta}^{-2} + c^2\delta^2 \text{Tr}(\underline{Q}^{-1}\underline{\Sigma}^{-1})\chi_{p,\Delta}^{-4}. \quad (4.27)
\end{aligned}$$

Again, the rhs of (4.27) is a non-degenerate r.v., and hence, here also, jackknifing fails to provide a consistent estimator of the asymptotic risk $\rho^{\text{SSJ}}(\lambda)$.

Finally, let us consider the case of $\hat{\theta}_n^{\text{SUJ}}$. By (2.8) through (2.11), we have on defining η_{ni} , δ_{ni} as in before,

$$\begin{aligned}
\hat{\theta}_n^{\text{SUJ}} - \hat{\theta}_n^{\text{SSJ}} &= (n-1)n^{-1} \sum_{i=1}^n T_{n-1}^{-2}(i) (A_{n-1}(i) - A_n)\bar{X}_{n-1}(i) \\
&= (n-1)T_n^{-2} (n^{-1} \sum_{i=1}^n (A_{n-1}(i) - A_n))\bar{X}_n \\
&\quad + (n-1)n^{-1} \sum_{i=1}^n \eta_{ni} (A_{n-1}(i) - A_n)\bar{X}_n \\
&\quad + (n-1)n^{-1} \sum_{i=1}^n (A_{n-1}(i) - A_n)\delta_{ni} T_n^{-2} \\
&\quad + (n-1)n^{-1} \sum_{i=1}^n \eta_{ni} (A_{n-1}(i) - A_n)\delta_{ni} \quad (4.28)
\end{aligned}$$

[At this stage, we assume that $|c_{n-1} - c_n| = o(n^{-\frac{1}{2}})$; in view of the fact that $\lim_{n \rightarrow \infty} c_n = c$ exists, we may always choose the sequence $\{c_n\}$ in such a way that this holds. Suppose that in (2.4), we only replace c_n by c_{n-1} and define the resulting quantity by A_n^0 . Then $\|A_n - A_n^0\| = o(n^{-\frac{1}{2}})$ a.s., and this replacement will make no asymptotic difference. As such, for simplicity of proof, we let $c_n = c_{n-1} = c$.] Then, by an appeal to [(3.13) of] Sen (1977), we obtain that as $n \rightarrow \infty$,

$$\begin{aligned}
&\|n^{-1} \sum_{i=1}^n (A_{n-1}(i) - A_n)\| \\
&= \|E(A_{n-1} - A_n \mid \underline{c}_n)\| = o(n^{-2}) \text{ a.s.}, \quad (4.29)
\end{aligned}$$

where we make use of the fact that the elements of \underline{S}_n are all U-statistics and $d_n = \text{Ch}_p(\underline{Q}\underline{S}_n)$ is a smooth function of \underline{S}_n . Further, note that $(n-1)\delta_{ni} = -(X_i - \bar{X}_n)$, $1 \leq i \leq n$, so that

$\max\{|(n-1)\delta_{ni}| : 1 \leq i \leq n\} = o(n^{\frac{1}{4}})$ a.s., as $n \rightarrow \infty$. Also, by using (3.7) and (3.8) of Sen (1977), we obtain that

$$\left(\sum_{i=1}^n \|A_{n-1(i)} - A_n\|\right)^2 \leq n \sum_{i=1}^n \|A_{n-1(i)} - A_n\|^2 = o(1) \text{ a.s.}, \quad (4.30)$$

as $n \rightarrow \infty$. Also, under $\{K_n\}$, $n^{\frac{1}{2}} \|X_n\| = O_p(1)$. Finally, using (4.13), we obtain that

$$\max_{1 \leq i \leq n} |T_{n-1}^{-2} - T_n^{-2}| = o(n^{-\frac{1}{4}}) \text{ a.s.}, \text{ as } n \rightarrow \infty. \quad (4.31)$$

Consequently, from (4.28) through (4.31), we have under $\{K_n\}$

$$n^{\frac{1}{2}} \|\hat{\theta}_{\sim n}^{\text{SUJ}} - \hat{\theta}_{\sim n}^{\text{SSJ}}\| \rightarrow 0, \text{ in probability, as } n \rightarrow \infty. \quad (4.32)$$

As a result, (4.16) through (4.23) also pertain to $\hat{\theta}_{\sim n}^{\text{SUJ}}$. Again, we may use (4.24)-(4.28) and conclude that (4.26) and (4.27) also hold for $\hat{\theta}_{\sim n}^{\text{SUJ}}$.

5. SOME GENERAL REMARKS

The results in Sections 3 and 4 raise some interesting questions on the role of jackknifing in shrinkage estimation. In the conventional fixed alternative case (treated in Section 3), like the Stein-rule estimator, its jackknifed versions are of no special use if n is large. This asymptotic picture also indicates the performance characteristics when θ is not very close to the pivot (here, 0). If F is itself a multinormal d.f., then, of course, such a finite-sample study can be made relatively more precisely, but, at the cost of tremendous mathematical complications.

In the case of Pitman-type alternatives (which is quite appropriate in the asymptotic case), it appears that whereas $\hat{\theta}_{\sim n}^{\text{SRJ}}$ agrees with $\hat{\theta}_{\sim n}^{\text{S}}$, but the other two versions, $\hat{\theta}_{\sim n}^{\text{SSJ}}$ and $\hat{\theta}_{\sim n}^{\text{SUJ}}$, are not asymptotically equivalent to $\hat{\theta}_{\sim n}^{\text{S}}$. Moreover, the other two versions have dominance (over \bar{X}_n) for larger values of p and for the case where $c < 2(p-4)(p-6)/(p-2) \ll 2(p-2)$. Thus, here, jackknifing has a reducing effect on the dominance of the Stein-rule estimator. On the top of that, for either of the three jackknifed versions, the corresponding jackknifed estimator of the dispersion matrices

fails to provide a consistent estimator of the ADR. This asymptotic picture also provides a good indication of the finite sample case when $\tilde{\theta}$ is "close to" the pivot. Thus, in either case, we have insufficient ground to advocate the use of jackknifing in the Stein-rule estimation. If our goal is also to estimate $\rho^S(\lambda)$, S_n may be used for $\tilde{\Sigma}$ and a consistent estimator of $\Delta(=n\tilde{\theta}'\tilde{\Sigma}^{-1}\tilde{\theta})$ and Δ^* may then be used in (4.3) to provide the desired answer. Perhaps, jackknifed estimators of Δ and Δ^* will have better performance in this context.

ACKNOWLEDGEMENT

This work was partially supported by the National Heart, Lung, and Blood Institute, Contract NIH-NHLBI-71-2243-L from the National Institutes of Health.

BIBLIOGRAPHY

- Berger, J.O. (1980). A robust generalized Bayes estimator and confidence region for a multivariate normal mean. *Ann Statist* 8, 716-761.
- Berger, J.O., Bock, M.E., Brown, L.D., Casella, G., and Glesser, L. (1977). Minimax estimation of a normal mean vector for arbitrary quadratic loss and unknown covariance matrix. *Ann Statist* 5, 763-771.
- James, W. and Stein, C. (1961). Estimation with quadratic loss. *Proc 4th Berkeley Symp Math Statist Prob* 1, 361-379.
- Judge, G.G. and Bock, M.E. (1978). *The Statistical Implications of Pretest and Stein-rule estimators in Econometrics*. North Holland, Amsterdam.
- Sen, P.K. (1977). Some invariance principles relating to jackknifing and their role in sequential analysis. *Ann Statist* 5, 316-329.
- Sen, P.K. (1984). A James-Stein detour of U-statistics. *Commun Statist Theor Meth* A13, 2725-2748.
- Sen, P.K. and Saleh, A.K.Md.E. (1985). On some shrinkage estimators of multivariate location. *Ann Statist* 13, 272-281.
- Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proc Third Berkeley Symp Math Statist Prob* 1, 197-206.