A JAMES-STEIN TYPE OF DETOUR OF U-STATISTICS

by

Pranab Kumar Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1434

March 1983

A JAMES-STEIN TYPE OF DETOUR OF U-STATISTICS *

By PRANAB KUMAR SEN

University of North Carolina, Chapel Hill.

SUMMARY. For a vector of estimable parameters, a modified version of the James-Stein rule (incorporating the idea of preliminary test estimators) is utilized in formulating some estimators based on U-statistics and their jackknifed estimator of dispersion matrix. Asymptotic admissibility properties of the classical U-statistics, their preliminary test version and the proposed estimators are studied.

1. INTRODUCTION

For the multivariate normal mean (vector) estimation problem, Stein (1956) was able to establish the inadmissibility of the sample mean vector, the maximum likelihood estimator (MLE), under a total squared error risk measure. Later, James and Stein (1961) were able to specify a simple non-linear estimator which dominates the MLE for the case of three or more dimensional vectors. This work has generated a lot of interest , and the multinormal distributional problems have been studied in greater generality by a number of workers. An excellent account of these works is given in Berger(1980b). For some related non-normal problems, we may refer to Ghosh (1983) and Ghosh, Hwang and Tsui(1983).

The object of the present investigation is to take a detour of the nonparametric estimation problem (in the vector case) based on Hoeffding's(1948) U-statistics and to consider some smooth shrinkage estimators along the lines of James and Stein (1961), with adaptations from the preliminary test estimation

AMS Subject Classification : 62C15, 62G99.

Key Words & Phrases : Asymptotic risk, inadmissibility, jackknifing, kernel, local alternatives, shrinkage estimators, U-statistics.

* Work partially supported by the National Heart, Lung and Blood Institute, Contract No. NIH-NHLBI-71-2243-L from the National Institutes of Health.

(PTE) theory. Whereas in the multinormal distributional problems, the (exact) normality of the distribution of the sample mean vector and the Wishart property of the sample dispersion matrix provide some unique facilities in the study of the finite sample admissibility results, for U-statistics and its jackknifed estimator of dispersion matrix, such exact distributional properties are , in general, not available. Further, in the nonparametric case, the form of the underlying distribution is also not assumed to be specified. In view of these, a direct adaptation of the James-Stein rule for the case of a vector of U-statistics may generally stumble into problems with the study of the finite sample admissibility results; there are problems too with the asymptotic case. We may remark that the James-Stein estimator may be regarded as a smoother version of some PTE. In a PTE rule, under some conditional specifications, both the unconstrained and the constrained estimators are formulated, and a preliminary test (on the constraints) dictates the final choice of the estimator. In the current context, we incorporate the idea of PTE, but, not to the fuller extent, so that we do not have a fully smooth shrinkage estimator, but a PTE version quite close to it. This adaptation makes it possible to study the desired asymptotic admissibility properties with the help of the existing results on U-statistics and its jackknifed estimator of dispersion matrix (which are extensively available in the literature).

Along with the preliminary notions, the U-statistics, their jackknifed estimator of dispersion matix and the propsed estimators are all introduced in Section 2. Some general properties of the estimators are considered in Section 3. The principal results on the admissibility of the estimators are presented in Section 4. Some general remarks are made in the last section.

2. THE PROPOSED ESTIMATORS

Let $\{X_i, i \geq 1\}$ be a sequence of independent and identically distributed random vectors (i.i.d.r.v.) with a distribution function (d.f.) F, defined on the Euclidean space E^q, for some $q \geq 1$. Let $\mathfrak{F}$ be the space of all d.f.'s belonging to a class, and for every $F \in \mathfrak{F}$, consider a vector of functionals

$$\underset{\sim}{\theta} = \underset{\sim}{\theta}(F) = (\theta_1(F), \ldots, \theta_p(F))', \text{ for some } p \geq 1, \tag{2.1}$$

whose domain is $\mathfrak{F}$. If there exists a kernel $\phi_j(x_1, \ldots, x_{m_j})$, symmetric in its arguments, of degree $m_j (\geq 1)$, such that

$$\theta_j(F) = E_F\{\phi_j(X_1, \ldots, X_{m_j})\}, \ \forall F \in \mathfrak{F}, \ j=1,\ldots,p, \tag{2.2}$$

then, $\underset{\sim}{\theta}$ is an estimable parameter (vector). For $n \geq m^* = \max(m_1, \ldots, m_p)$, we may then define the vector of U-statistics $\underset{\sim}{U}_n = (U_{n1}, \ldots, U_{np})'$ by letting

$$U_{nj} = \binom{n}{m_j}^{-1} \sum_{1 \leq i_1 < \ldots < i_{m_j} \leq n} \phi_j(X_{i_1}, \ldots, X_{i_m}), \ j=1,\ldots,p. \tag{2.3}$$

$\underset{\sim}{U}_n$ is a symmetric and unbiased estimator of $\underset{\sim}{\theta}$ and it possesses some other optimality properties too [see Halmos(1946)]. In fact, for any unbiased estimator $\underset{\sim}{T}_n$ of $\underset{\sim}{\theta}$, the corresponding $\underset{\sim}{U}_n$ has a risk (using any convex loss function) smaller than or equal to that of $\underset{\sim}{T}_n$.

We assume that the kernels ϕ_j are all square integrable. Let then

$$\phi_{j,c}(x_1, \ldots, x_c) = E\phi_j(x_1, \ldots, x_c, X_{c+1}, \ldots, X_{m_j}), \ c=0,\ldots,m_j; \tag{2.4}$$

$$\zeta_{j\ell,c}(F) = E_F\{\phi_{j,c}(X_1, \ldots, X_c)\, \phi_{\ell,c}(X_1, \ldots, X_c)\} - \theta_j(F)\theta_\ell(F), \tag{2.5}$$

for $j, \ell = 1,\ldots,p$ and $c=0,\ldots,\min(m_j, m_\ell)$. Then [see Hoeffding(1948)]

$$nE[(\underset{\sim}{U}_n - \underset{\sim}{\theta})(\underset{\sim}{U}_n - \underset{\sim}{\theta})'] = n\left(\left(\binom{n}{m_j}^{-1} \sum_{c=1}^{m_j} \binom{m_\ell}{c}\binom{n-m_\ell}{m_j-c} \zeta_{j\ell,c}(F) \right)\right)$$

$$= \underset{\sim}{\Gamma} + \underset{\sim}{O}(n^{-1}) \tag{2.6}$$

where

$$\underset{\sim}{\Gamma} = ((\gamma_{j\ell})) = ((m_j m_\ell \zeta_{j\ell,1}(F))). \tag{2.7}$$

For an estimator $\underset{\sim}{\delta}_n = (\delta_{n1}, \ldots, \delta_{np})'$ of $\underset{\sim}{\theta}$, we consider a quadratic loss function

$$L_n(\underset{\sim}{\delta}_n, \underset{\sim}{\theta}) = n(\underset{\sim}{\delta}_n - \underset{\sim}{\theta})' \underset{\sim}{Q} (\underset{\sim}{\delta}_n - \underset{\sim}{\theta}) / \mathrm{Tr}(\underset{\sim}{Q}\underset{\sim}{\Gamma}), \tag{2.8}$$

where $\underset{\sim}{Q}$ is a given positive definite matrix and the denominator $\mathrm{Tr}(\underset{\sim}{Q}\underset{\sim}{\Gamma})$ is a convenient standardization factor. Then, according to the loss functuon in (2.8), the risk (expected loss) for $\underset{\sim}{U}_n$ is given by

$$\rho_n(\underset{\sim}{U}_n,\underset{\sim}{\theta}) = EL_n(\underset{\sim}{U}_n,\underset{\sim}{\theta}) = nE[(\underset{\sim}{U}_n-\underset{\sim}{\theta})'\underset{\sim}{Q}(\underset{\sim}{U}_n-\underset{\sim}{\theta})]/Tr(\underset{\sim}{Q}\underset{\sim}{\Gamma})$$
$$= 1 + O(n^{-1}), \quad \text{by (2.6)}. \tag{2.9}$$

Note that $\underset{\sim}{U}_n$ will be termed an inadmissible estimator of $\underset{\sim}{\theta}$, if there exists an alternative estimator $\underset{\sim}{T}_n$, such that

$$\rho_n(\underset{\sim}{T}_n,\underset{\sim}{\theta}) \leq \rho_n(\underset{\sim}{U}_n,\underset{\sim}{\theta}) \quad \text{for all } F \in \mathfrak{F}, \tag{2.10}$$

with strict inequality holding for some $F \in \mathfrak{F}$. If instead of (2.10) holding for every n, we have

$$\lim_{n\to\infty} \{\rho_n(\underset{\sim}{T}_n,\underset{\sim}{\theta})\} \leq \lim_{n\to\infty}\{\rho_n(\underset{\sim}{U}_n,\underset{\sim}{\theta})\} = 1, \quad F \in \mathfrak{F}, \tag{2.11}$$

with strict inequality for some $F \in \mathfrak{F}$, then, $\underset{\sim}{U}_n$ is termed asymptotically inadmissible. Our main concern is to study the asymptotic inadmissibility of $\underset{\sim}{U}_n$ in a meaningful context.

Towards this study, we motivate the estimators (to be proposed) through the special case of the multivariate mean vector problem where the $\underset{\sim}{X}_i$ are all p-vectors, $m_1=\ldots=m_p=1$ and $\phi_j(\underset{\sim}{X}_i)$ is the jth coordinate of $\underset{\sim}{X}_i$, for $j=1,\ldots,p$. Then, $\underset{\sim}{U}_n = \underset{\sim}{\overline{X}}_n = n^{-1}\Sigma_{i=1}^n \underset{\sim}{X}_i$. If $\underset{\sim}{S}_n = (n-1)^{-1}\Sigma_{i=1}^n(\underset{\sim}{X}_i - \underset{\sim}{\overline{X}}_n)(\underset{\sim}{X}_i-\underset{\sim}{\overline{X}}_n)'$ be the sample covariance matrix, then for normal F with an unknown dispersion matrix $\underset{\sim}{\Sigma}$, Berger, Bock, Brown, Casella and Gleser (1977) considered the James-Stein type estimator

$$\underset{\sim}{\overline{X}}_n^* = (\underset{\sim}{I} - cd_n(n\underset{\sim}{\overline{X}}_n'\underset{\sim}{S}_n^{-1}\underset{\sim}{\overline{X}}_n)^{-1}\underset{\sim}{Q}^{-1}\underset{\sim}{S}_n^{-1})\,\underset{\sim}{\overline{X}}_n, \tag{2.12}$$

where $d_n = ch_p(\underset{\sim}{Q}\underset{\sim}{S}_n)$, the minimum characteristic root of $\underset{\sim}{Q}\underset{\sim}{S}_n$, and $c(0\leq c\leq c_{n,p})$ is a positive constant. For some computed values of $c_{n,p}$, they have shown that $\rho_n(\underset{\sim}{\overline{X}}_n^*,\underset{\sim}{\theta}) \leq \rho_n(\underset{\sim}{\overline{X}}_n,\underset{\sim}{\theta})$. In this context, it may be remarked [see Berger(1980a, p.125)] that for any $\underset{\sim}{\mu} = E\underset{\sim}{X} \neq \underset{\sim}{0}$, as $n \to \infty$, $\rho_n(\underset{\sim}{\overline{X}}_n^*,\underset{\sim}{\theta}) \to 1 = \rho_n(\underset{\sim}{\overline{X}}_n,\underset{\sim}{\theta})$. Thus, the James-Stein type estimator in (2.12), for normal F, is only of practical importance when either n is not very large or when the noncentrality parameter $n\underset{\sim}{\mu}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{\mu}$ is not very large. As a matter of fact, to emphasize the last point, if we consider a sequence $\{K_n\}$ of local translation alternatives, where under K_n, $X_{n1},\dots,X_{nn}$ are i.i.d.r.v. with mean vector $\underset{\sim}{\mu} = \underset{\sim}{\mu}_n = n^{-\frac{1}{2}}\underset{\sim}{\lambda}$, for some fixed $\underset{\sim}{\lambda}$,and dispersion matrix $\underset{\sim}{\Sigma}$. Note that under $\{K_n\}$, the non-centrality parameter $n\underset{\sim}{\mu}_n'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{\mu}_n = \underset{\sim}{\lambda}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{\lambda} = \Delta$, say, is a finite number. Then, it can be shown that for every (fixed) $\underset{\sim}{\lambda} \in E^p$,

$$\lim_{n\to\infty}\{\rho_n(\overline{\underset{\sim}{X}}_n^*,\underset{\sim}{\mu}_n)|K_n\} < 1 \text{ , for every } c \in (0,2(p-2)),\ p>2. \qquad (2.13)$$

This may be interpreted as the asymptotic inadmissibility of the MLE $\overline{X}_n$ under local alternatives. Our main objective is to generalize this asymptotic inadmissibility result under local alternatives to the general class of U-statistics under consideration, without making any explicit assumption on the form of the underlying d.f. F.

For U-statistics, a convenient estimator of $\underset{\sim}{\Gamma}$ is obtained through the jackknifing technique [see Sen(1960, 1981)]. For this, we write $\underset{\sim}{U}_n = \underset{\sim}{U}(X_1,\dots,X_n)$, and, for every i(=1,...,n), we let

$$\underset{\sim}{U}_{n-1}^{(i)} = U(X_1,\dots,X_{i-1},X_{i+1},\dots,X_n) \text{ , } \underset{\sim}{U}_{n,i} = n\underset{\sim}{U}_n - (n-1)\underset{\sim}{U}_{n-1}^{(i)} . \qquad (2.14)$$

Then, the jackknifed estimator of $\underset{\sim}{\Gamma}$ is given by

$$\hat{\underset{\sim}{\Gamma}}_n = (n-1)^{-1}\Sigma_{i=1}^n[\underset{\sim}{U}_{n,i} - \underset{\sim}{U}_n][\underset{\sim}{U}_{n,i} - \underset{\sim}{U}_n]' . \qquad (2.15)$$

By comparison with (2.12), we now introduce the James-Stein type U-statistics as

$$\underset{\sim}{U}_n^{JS} = (\underset{\sim}{I} - cd_n(n\underset{\sim}{U}_n'\hat{\underset{\sim}{\Gamma}}_n^{-1}\underset{\sim}{U}_n)^{-1}\underset{\approx}{Q}^{-1}\hat{\underset{\sim}{\Gamma}}_n^{-1})\underset{\sim}{U}_n , \qquad (2.16)$$

where c is a positive constant and $d_n = ch_p(\underset{\sim}{Q}\hat{\underset{\sim}{\Gamma}}_n)$. Whereas in the normal F case, the Wishart property of $\underset{\sim}{S}_n$ and the stochastic independence of $\overline{\underset{\sim}{X}}_n$ and $\underset{\sim}{S}_n$ enable one to compute the negative moments of $\mathcal{L}_n^o = n\overline{\underset{\sim}{X}}_n'\underset{\sim}{S}_n^{-1}\overline{\underset{\sim}{X}}_n$, needed for the risk computation of the estimator in (2.12), for possibly non-normal F and for general kernels, though $\mathcal{L}_n = n\underset{\sim}{U}_n'\hat{\underset{\sim}{\Gamma}}_n^{-1}\underset{\sim}{U}_n$ may converge in distribution to a non-degenerate random variable, the convergence of its negative moments is not generally insured. Indeed, when $\underset{\sim}{U}_n$ may assume the null value $(\underset{\sim}{0})$ with a positive probability, however small, $E\mathcal{L}_n^{-t} = +\infty$, for every $t > 0$. This may create some difficulties in the computation of the risk of the estimator in (2.16). In this context, we may also remark that a natural competing estimator is the preliminary test estimator (PTE) $\underset{\sim}{U}_n^{PT}$ given by

$$\underset{\sim}{U}_n^{PT} = \begin{cases} \underset{\sim}{U}_n & \text{if } \mathcal{L}_n \geq \mathcal{L}_{n,\alpha} , \\ \underset{\sim}{0}, & \text{if } \mathcal{L}_n < \mathcal{L}_{n,\alpha} , \end{cases} \tag{2.17}$$

where α is the level of significance of the preliminary test (for $\underset{\sim}{\theta} = \underset{\sim}{0}$) based on $\mathcal{L}_n$, $\mathcal{L}_{n,\alpha}$ is the critical level of $\mathcal{L}_n$, and for large n, one may also replace $\mathcal{L}_{n,\alpha}$ by $\chi_p^2(\alpha)$, the upper $100\alpha\%$ point of the central chi square d.f. with p degrees of freedom. Of course, if $\underset{\sim}{\theta} \neq \underset{\sim}{0}$, then, by virtue of the consistency of the test based on $\mathcal{L}_n$, $\mathcal{L}_n$ goes to $+\infty$, as $n \to \infty$, with probability one, so that both $\underset{\sim}{U}_n^{JS}$ and $\underset{\sim}{U}_n^{PT}$ become asymptotically equivalent to $\underset{\sim}{U}_n$. Thus, both $\underset{\sim}{U}_n^{JS}$ and $\underset{\sim}{U}_n^{PT}$ may be regarded as adjusted U-statistics, $\underset{\sim}{U}_n^{JS}$ being a smoother version than $\underset{\sim}{U}_n^{PT}$, and both are designed to yield smaller risk when $\underset{\sim}{\theta}$ is close to $\underset{\sim}{0}$. First, we propose an estimator which combines the feature of the smoothness of $\underset{\sim}{U}_n^{JS}$ and the stochastic smallness of $\underset{\sim}{U}_n^{PT}$ for small values of $\underset{\sim}{U}_n$. We define

$$\underset{\sim}{U}_n^{S} = \begin{cases} \underset{\sim}{U}_n^{JS}, & \text{if } \mathcal{L}_n \geq \varepsilon , \\ \underset{\sim}{0}, & \text{if } \mathcal{L}_n < \varepsilon , \end{cases} \tag{2.18}$$

where ε (> 0) is some arbitrarily small number. This may be termed an adaptive version of the James-Stein type estimator. We also propose a second estimator which is a smoother version of (2.18), where for small values of $\mathcal{L}_n$, we have a non-null counterpart. This is defined by

$$\underset{\sim}{U}_n^{S*} = \begin{cases} \underset{\sim}{U}_n^{JS} , & \text{if } \mathcal{L}_n \geq \varepsilon , \\ (\underset{\sim}{I} - \varepsilon^{-\frac{1}{2}} c d_n \mathcal{L}_n^{-\frac{1}{2}} \underset{\sim}{Q}^{-1} \hat{\underset{\sim}{\Gamma}}_n^{-1}) \underset{\sim}{U}_n , & \text{if } \mathcal{L}_n < \varepsilon , \end{cases} \tag{2.19}$$

where c, d_n etc. are all defined as in after (2.16). Note that in this way, the discontinuity of the function at $\mathcal{L}_n = \varepsilon$ is avoided, and also, for $\mathcal{L}_n < \varepsilon$, instead of taking the estimator as $\underset{\sim}{0}$, we have a non-trivial estimator. Our main interest centers around the study of the (in-)admissibility results of the classical estimator $\underset{\sim}{U}_n$ with respect to the estimators $\underset{\sim}{U}_n^{JS}$, $\underset{\sim}{U}_n^{PT}$, $\underset{\sim}{U}_n^{S}$ and $\underset{\sim}{U}_n^{S*}$. In this respect, both the fixed and local alternative asymptotic setups will be considered. In passing, we may remark that for local alternatives, we have considered the pivot $\underset{\sim}{\theta} = \underset{\sim}{0}$. This may be taken as $\underset{\sim}{\theta} = \underset{\sim}{\theta}_o$, for any specified $\underset{\sim}{\theta}_o$, and then, working with the adjusted kernels as $\underset{\sim}{\phi} - \underset{\sim}{\theta}_o$, we may always reduce the case to the null pivot.

3. ASYMPTOTIC ADMISSIBILITY RESULTS FOR FIXED ALTERNATIVES

Note that by (2.17), we have

$$\underset{\sim}{U}_n^{PT} - \underset{\sim}{U}_n = -\underset{\sim}{U}_n I(\mathcal{L}_n < \mathcal{L}_{n,\alpha}). \tag{3.1}$$

Therefore, by (3.1), we have

$$\begin{aligned} n(\underset{\sim}{U}_n^{PT} - \underset{\sim}{U}_n)'\underset{\sim}{Q}(\underset{\sim}{U}_n^{PT} - \underset{\sim}{U}_n) &= I(\mathcal{L}_n < \mathcal{L}_{n,\alpha})n\underset{\sim}{U}_n'\underset{\sim}{Q}\underset{\sim}{U}_n \\ &= I(\mathcal{L}_n < \mathcal{L}_{n,\alpha})\mathcal{L}_n \{n\underset{\sim}{U}_n'\underset{\sim}{Q}\underset{\sim}{U}_n / n\underset{\sim}{U}_n'\hat{\underset{\sim}{\Gamma}}_n^{-1}\underset{\sim}{U}_n\} \\ &\leq I(\mathcal{L}_n < \mathcal{L}_{n,\alpha})\mathcal{L}_n \, ch_1(\underset{\sim}{Q}\hat{\underset{\sim}{\Gamma}}_n) \\ &\leq \mathcal{L}_{n,\alpha} I(\mathcal{L}_n < \mathcal{L}_{n,\alpha}) ch_1(\underset{\sim}{Q}\hat{\underset{\sim}{\Gamma}}_n) \\ &\leq \mathcal{L}_{n,\alpha} I(\mathcal{L}_n < \mathcal{L}_{n,\alpha}) Tr(\underset{\sim}{Q}\hat{\underset{\sim}{\Gamma}}_n) . \end{aligned} \tag{3.2}$$

Let us now assume that for some $r > 1$, $\underset{\sim}{\phi} \in L^{2r}$, i.e.,

$$E_F||\underset{\sim}{\phi}||^{2r} < \infty \text{ , for some } r > 1 \text{ (not necessarily an integer).} \tag{3.3}$$

Then, using the Hölder inequality, we have

$$E_F I(\mathcal{L}_n < \mathcal{L}_{n,\alpha}) Tr(\underset{\sim}{Q}\hat{\underset{\sim}{\Gamma}}_n) \leq \{E_F I(\mathcal{L}_n < \mathcal{L}_{n,\alpha})\}^{1/s} \{E_F(Tr(\underset{\sim}{Q}\hat{\underset{\sim}{\Gamma}}_n))^r\}^{1/r}, \tag{3.4}$$

where $1/s + 1/r = 1$. Note that $Tr(\underset{\sim}{Q}\hat{\underset{\sim}{\Gamma}}_n)$ is a linear combination of the elements of $\hat{\underset{\sim}{\Gamma}}_n$ which are expressible as a linear combination of U-statistics of degree $2m^*$ (or less) and whose moments of order $r(>1)$ exist under (3.3) [see Sen and Ghosh(1981) in this respect]. Therefore, under (3.3), $E(Tr(\underset{\sim}{Q}\hat{\underset{\sim}{\Gamma}}_n))^r$ exists and converges to a finite limit as $n \to \infty$. Also, $E_F I(\mathcal{L}_n < \mathcal{L}_{n,\alpha}) = P\{\mathcal{L}_n < \mathcal{L}_{n,\alpha}\}$ where $\mathcal{L}_{n,\alpha}$ converges to a finite limit $\chi_p^2(\alpha)$, as $n \to \infty$. Finally, under (3.3), $\hat{\underset{\sim}{\Gamma}}_n$ converges a.s. to $\underset{\sim}{\Gamma}$, as $n \to \infty$, and $\underset{\sim}{U}_n$ converges a.s. to $\underset{\sim}{\theta}$, as $n \to \infty$, so that $n^{-1}\mathcal{L}_n$ converges a.s. to $\underset{\sim}{\theta}'\underset{\sim}{\Gamma}^{-1}\underset{\sim}{\theta} = \Delta$ as $n \to \infty$. Therefore, for every $\underset{\sim}{\theta} \neq \underset{\sim}{0}$, i.e., $\Delta > 0$, $P\{\mathcal{L}_n < \mathcal{L}_{n,\alpha}\} \to 0$, as $n \to \infty$. Hence, the right hand side of (3.2) converges to 0 as $n \to \infty$. This leads us to the following.

<u>Theorem 3.1. Under</u> (3.3), <u>for every fixed</u> $\theta \neq 0$, <u>with respect to the loss function in</u> (2.8), U_n <u>and</u> U_n^{PT} <u>are asymptotically risk equivalent.</u>

Let us next consider the estimator U_n^{JS} in (2.16). Then, we have

$$U_n^{JS} - U_n = c\, d_n \mathcal{L}_n^{-1} Q^{-1} \hat{\Gamma}_n^{-1} U_n \; ; \; \mathcal{L}_n = n U_n' \hat{\Gamma}_n^{-1} U_n , \tag{3.5}$$

so that

$$\begin{aligned} n(U_n^{JS} - U_n)' Q (U_n^{JS} - U_n) &= c^2 d_n^2 \mathcal{L}_n^{-2} \{ n U_n' \hat{\Gamma}_n^{-1} Q^{-1} \hat{\Gamma}_n^{-1} U_n \} \\ &= \frac{c^2 \{ ch_p(Q\hat{\Gamma}_n) \}^2}{(n U_n' \hat{\Gamma}_n^{-1} Q^{-1} \hat{\Gamma}_n^{-1} U_n)} \left\{ \frac{n U_n' \hat{\Gamma}_n^{-1} Q^{-1} \hat{\Gamma}_n^{-1} U_n}{n U_n' \hat{\Gamma}_n^{-1} U_n} \right\}^2 \\ &\leq c^2 \{ ch_p(Q\hat{\Gamma}_n) \}^2 \{ ch_1(Q^{-1}\hat{\Gamma}_n^{-1}) \}^2 (n U_n' \hat{\Gamma}_n^{-1} Q^{-1} \hat{\Gamma}_n^{-1} U_n)^{-1} \\ &= c^2 \left(n U_n' \hat{\Gamma}_n^{-1} Q^{-1} \hat{\Gamma}_n^{-1} U_n \right)^{-1} . \end{aligned} \tag{3.6}$$

Let us denote by $\mathcal{L}_n^* = n U_n' \hat{\Gamma}_n^{-1} Q^{-1} \hat{\Gamma}_n^{-1} U_n$. Then, by the a.s. convergence of U_n and $\hat{\Gamma}_n$ (as were discussed earlier), we have for any fixed F,

$$n^{-1} \mathcal{L}_n^* \rightarrow \theta' \Gamma^{-1} Q^{-1} \Gamma^{-1} \theta = \Delta^* \text{ a.s., as } n \rightarrow \infty . \tag{3.7}$$

Hence, for every (fixed) $\theta \neq 0$, as $n \rightarrow \infty$, $(\mathcal{L}_n^*)^{-1} \rightarrow 0$ a.s., but this does not necessarily imply that $(\mathcal{L}_n^*)^{-1} \rightarrow 0$ in the first moment too. In fact, the behaviour of $\mathcal{L}_n^*$ for U_n close to 0 may negate this moment convergence. However, if we assume that for every (fixed) $\theta \neq 0$,

$$E(\mathcal{L}_n^*)^{-1} \rightarrow 0 \text{ as } n \rightarrow \infty , \tag{3.8}$$

then, the right hand side of (3.6) converges to 0 in the first mean when $n \rightarrow \infty$. This leads us to the following.

<u>Theorem 3.2. Under</u> (3.8), <u>for every fixed</u> $\theta \neq 0$, <u>with respect to the loss function in</u> (2.8), U_n <u>and</u> U_n^{JS} <u>are asymptotically risk equivalent.</u>

Let us next note that by (2.16) and (2.18),

$$U_n^S - U_n = -U_n I(\mathcal{L}_n < \varepsilon) + I(\mathcal{L}_n \geq \varepsilon)(U_n^{JS} - U_n), \tag{3.9}$$

so that

$$n(\underset{\sim}{U}_n^S - \underset{\sim}{U}_n)'\underset{\sim}{Q}(\underset{\sim}{U}_n^S - \underset{\sim}{U}_n) = nI(\mathcal{L}_n < \varepsilon)\underset{\sim}{U}_n'\underset{\sim}{Q}\underset{\sim}{U}_n + I(\mathcal{L}_n \geq \varepsilon)c^2 d_n^2 \mathcal{L}_n^*/\mathcal{L}_n^2$$

$$= nI(\mathcal{L}_n < \varepsilon)(\underset{\sim}{U}_n'\underset{\sim}{Q}\underset{\sim}{U}_n/\underset{\sim}{U}_n'\hat{\underset{\sim}{\Gamma}}_n^{-1}\underset{\sim}{U}_n)\mathcal{L}_n + I(\mathcal{L}_n \geq \varepsilon)c^2 d_n^2(\mathcal{L}_n^*/\mathcal{L}_n)\mathcal{L}_n^{-1}$$

$$\leq \mathcal{L}_n I(\mathcal{L}_n < \varepsilon)[ch_1(\underset{\sim}{Q}\underset{\sim}{\Gamma}_n)] + \mathcal{L}_n^{-1} I(\mathcal{L}_n \geq \varepsilon)c^2 d_n^2 \, ch_1(\underset{\sim}{Q}^{-1}\underset{\sim}{\Gamma}_n^{-1})$$

$$= \mathcal{L}_n I(\mathcal{L}_n < \varepsilon)ch_1(\underset{\sim}{Q}\underset{\sim}{\Gamma}_n) + ch_p(\underset{\sim}{Q}\underset{\sim}{\Gamma}_n)c^2\mathcal{L}_n^{-1} I(\mathcal{L}_n \geq \varepsilon)$$

$$\leq [\mathrm{Tr.}(\underset{\sim}{Q}\underset{\sim}{\Gamma}_n)]\{\varepsilon I(\mathcal{L}_n < \varepsilon) + c^2\mathcal{L}_n^{-1} I(\mathcal{L}_n \geq \varepsilon)\}. \qquad (3.10)$$

Since $n^{-1}\mathcal{L}_n \to \Delta$ a.s. as $n \to \infty$, where $\Delta > 0$ for every $\underset{\sim}{\theta} \neq \underset{\sim}{0}$, and $\mathcal{L}_n^{-1} I(\mathcal{L}_n \geq \varepsilon)$ is bounded (by ε^{-1}), proceeding as in (3.4), it follows that under (3.3), the right hand side of (3.10) converges in the first mean to 0 as $n \to \infty$, whenever $\underset{\sim}{\theta} \neq \underset{\sim}{0}$. This leads us to the following.

Theorem 3.3. *Under* (3.3), *for every (fixed)* $\underset{\sim}{\theta} \neq \underset{\sim}{0}$, *with respect to the loss function* (2.8), $\underset{\sim}{U}_n$ *and* $\underset{\sim}{U}_n^S$ *are asymptotically risk equivalent.*

Note that (3.8) in general needs more stringent regularity conditions than (3.3), so that the asymptotic risk equivalence in Theorems 3.1 and 3.3 holds under less stringent regularity conditions than in Theorem 3.2.

Finally, we consider the case of $\underset{\sim}{U}_n^{S*}$. Note that by (2.18) and (2.19),

$$n(\underset{\sim}{U}_n^{S*} - \underset{\sim}{U}_n^S)'\underset{\sim}{Q}(\underset{\sim}{U}_n^{S*} - \underset{\sim}{U}_n^S) = I(\mathcal{L}_n < \varepsilon)\{n\underset{\sim}{U}_n'\underset{\sim}{Q}\underset{\sim}{U}_n - 2\varepsilon^{-\frac{1}{2}}cd_n\mathcal{L}_n^{\frac{1}{2}} + \varepsilon^{-1}c^2 d_n^2\mathcal{L}_n^*/\mathcal{L}_n\}$$

$$\leq I(\mathcal{L}_n < \varepsilon)\{ch_1(\underset{\sim}{Q}\hat{\underset{\sim}{\Gamma}}_n)\mathcal{L}_n - 2\varepsilon^{-\frac{1}{2}}cd_n\mathcal{L}_n^{\frac{1}{2}} + \varepsilon^{-1}c^2 d_n^2\mathcal{L}_n^*/\mathcal{L}_n\}$$

$$\leq I(\mathcal{L}_n < \varepsilon)\{\mathrm{Tr}(\underset{\sim}{Q}\hat{\underset{\sim}{\Gamma}}_n)\mathcal{L}_n + 2cd_n + \varepsilon^{-1}c^2 d_n\}$$

$$\leq \mathrm{Tr}(\underset{\sim}{Q}\hat{\underset{\sim}{\Gamma}}_n)\, I(\mathcal{L}_n < \varepsilon)\{\varepsilon + 2c + \varepsilon^{-1}c^2\}$$

$$= \mathrm{Tr}(\underset{\sim}{Q}\hat{\underset{\sim}{\Gamma}}_n) I(\mathcal{L}_n < \varepsilon)\varepsilon(1 + c/\varepsilon)^2. \qquad (3.11)$$

Under (3.3), proceeding as in (3.4), for every (fixed) $\underset{\sim}{\theta} \neq \underset{\sim}{0}$, the right hand side of (3.11) converges to 0 in the first mean, as $n \to \infty$. Hence, we have the following.

Theorem 3.4. *Under* (3.3), *for every (fixed)* $\underset{\sim}{\theta} \neq \underset{\sim}{0}$, *with respect to the loss function in* (2.8), $\underset{\sim}{U}_n$, $\underset{\sim}{U}_n^S$ *and* $\underset{\sim}{U}_n^{S*}$ *are all asymptotically risk equivalent.*

As we shall see in the next section, the situation is different for the case of local alternatives.

4. INADMISSIBILITY OF U_n FOR LOCAL (PITMAN) ALTERNATIVES

We are mainly interested here in extending (2.13) to the class of U-statistics under consideration. Thus, we conceive of a triangular array $\{X_{n1},\ldots,X_{nn};\ n \geq 1\}$ of r.v.'s, where, for each n, the X_{ni} are i.i.d. with a common d.f. $F^{(n)}$ converging to a limit F as $n \to \infty$, in such a way that

$$\underset{\sim}{\theta}_n = \underset{\sim}{\theta}(F^{(n)}) = n^{-\frac{1}{2}}\underset{\sim}{\lambda}\ , \text{ for some (fixed) } \underset{\sim}{\lambda} \in E^p, \tag{4.1}$$

so that $\underset{\sim}{\theta}(F) = \lim_{n\to\infty} \underset{\sim}{\theta}(F^{(n)}) = \underset{\sim}{0}$. We denote this sequence of alternatives in (4.1) by $\{K_n\}$. With respect to this sequence of local (Pitman type) alternatives, we would like to study the asymptotic risk for the different estimators considered in Section 2, and this would provide information on the asymptotic (in-)admissibility results for these estimators. In this context, whenever needed, we shall assume that (3.3) holds uniformly in $n(\geq n_o)$.

First, consider the case of the PTE $\underset{\sim}{U}_n^{PT}$ in (2.17). Note that under (4.1),

$$n(\underset{\sim}{U}_n^{PT}-\underset{\sim}{\theta}_n)'\underset{\sim}{Q}(\underset{\sim}{U}_n^{PT}-\underset{\sim}{\theta}_n) = (\underset{\sim}{\lambda}'\underset{\sim}{Q}\underset{\sim}{\lambda})I(\mathcal{L}_n < \mathcal{L}_{n,\alpha}) + I(\mathcal{L}_n \geq \mathcal{L}_{n,\alpha})[n(\underset{\sim}{U}_n - \underset{\sim}{\theta}_n)'\underset{\sim}{Q}(\underset{\sim}{U}_n - \underset{\sim}{\theta}_n)]\ , \tag{4.2}$$

where $\mathcal{L}_n = n\underset{\sim}{U}_n'\hat{\underset{\sim}{\Gamma}}_n^{-1}\underset{\sim}{U}_n$ and $\mathcal{L}_{n,\alpha}$ converges to $\chi_p^2(\alpha)$, as $n \to \infty$. Therefore, by (2.8), (2.9) and (4.2), we have

$$\rho(\underset{\sim}{U}_n^{PT},\underset{\sim}{\theta}_n) = (\underset{\sim}{\lambda}'\underset{\sim}{Q}\underset{\sim}{\lambda})(\mathrm{Tr.}(\underset{\sim}{Q}\underset{\sim}{\Gamma}))^{-1}P\{\mathcal{L}_n < \mathcal{L}_{n,\alpha} \mid K_n\} + (\mathrm{Tr}(\underset{\sim}{Q}\underset{\sim}{\Gamma}))^{-1}E\{I(\mathcal{L}_n \geq \mathcal{L}_{n,\alpha})n(\underset{\sim}{U}_n-\underset{\sim}{\theta}_n)'\underset{\sim}{Q}(\underset{\sim}{U}_n-\underset{\sim}{\theta}_n) \mid K_n\}. \tag{4.3}$$

Now, under (3.3), $\forall\ x \in [0,\infty)$

$$\lim_{n\to\infty} P\{\mathcal{L}_n \leq x \mid K_n\} = H_p(x\ ;\ \Delta)\ ;\ \ \Delta = \underset{\sim}{\lambda}'\underset{\sim}{\Gamma}^{-1}\underset{\sim}{\lambda} = \mathrm{Tr}(\underset{\sim}{\lambda}\underset{\sim}{\lambda}'\underset{\sim}{Q}), \tag{4.4}$$

where $H_q(x;\delta)$ stands for the noncentral chi square d.f. with q degrees of freedom and noncentrality parameter δ. Note that $H_q(\chi_q^2(\alpha);0) = 1-\alpha$, $H_q(x;\delta)$ is $\downarrow$ in δ (≥ 0) and it converges to 0 as $\delta \to \infty$; further, $H_q(x;\delta)$ is a nonincreasing function of q when δ is held fixed. We may also note that if $G_p(x;0,\underset{\sim}{\Sigma})$ stands for the p-variate normal d.f. with mean 0 and dispersion matrix $\underset{\sim}{\Sigma}$, and if $\underset{\sim}{B}$ is a positive semi-definite matrix of rank q, then, for

for every $\underset{\sim}{a} \in E^p$, $c \geq 0$,

$$\int_{\{(\underset{\sim}{x}+\underset{\sim}{a})'\underset{\sim}{B}(\underset{\sim}{x}+\underset{\sim}{a}) \geq c\}} \underset{\sim}{x}\underset{\sim}{x}' dG_p(\underset{\sim}{x};\underset{\sim}{0}, \underset{\sim}{\Sigma}) = \{1-H_q(c;\delta)\}\underset{\sim}{\Sigma}+ \tag{4.5}$$

$$[H_q(c;\delta) - H_{q+2}(c;\delta)]\underset{\sim}{\Sigma}\underset{\sim}{B}\underset{\sim}{\Sigma} - [H_q(c;\delta)-2H_{q+2}(c;\delta)+H_{q+4}(c;\delta)]\underset{\sim}{\Sigma}\underset{\sim}{B}\underset{\sim}{a}\underset{\sim}{a}'\underset{\sim}{B}\underset{\sim}{\Sigma} \quad ,$$

where $\delta = \underset{\sim}{a}'\underset{\sim}{B}\underset{\sim}{a}$. Further, note that under $\{K_n\}$ and (3.3), $\hat{\underset{\sim}{\Gamma}}_n$ converges to in the rth mean and $n(\underset{\sim}{U}_n-\underset{\sim}{\theta}_n)'\underset{\sim}{Q}(\underset{\sim}{U}_n-\underset{\sim}{\theta}_n)$ is uniformly (in n) integrable [see for example, Sen and Ghosh (1981)]. Also, $n^{\frac{1}{2}}(\underset{\sim}{U}_n - \underset{\sim}{\theta}_n)$ is asymptotically normal with null mean vector and dispersion matrix $\underset{\sim}{\Gamma}$. Hence, noting that here $\underset{\sim}{\Sigma} = \underset{\sim}{\Gamma}$, $\underset{\sim}{B} = \underset{\sim}{\Gamma}^{-1}$, $\underset{\sim}{a} = \underset{\sim}{\lambda}$, $q= p$ and $\delta = \Delta$, we obtain from (4.5) and some standard manipulations that the second term on the right hand side of (4.3) converges to

$$[\mathrm{Tr}(\underset{\sim}{Q}\underset{\sim}{\Gamma})]^{-1}\{ (1-H_p(\chi_p^2(\alpha);\Delta))\mathrm{Tr}(\underset{\sim}{Q}\underset{\sim}{\Gamma}) + (H_p(\chi_p^2(\alpha);\Delta)-H_{p+2}(\chi_p^2(\alpha);\Delta))\mathrm{Tr}(\underset{\sim}{Q}\underset{\sim}{\Gamma})$$

$$- (H_p(\chi_p^2(\alpha);\Delta)-2H_{p+2}(\chi_p^2(\alpha);\Delta)+H_{p+4}(\chi_p^2(\alpha);\Delta))\underset{\sim}{\lambda}'\underset{\sim}{Q}\underset{\sim}{\lambda} \} \tag{4.6}$$

$$= 1-H_{p+2}(\chi_p^2(\alpha);\Delta) - [(\underset{\sim}{\lambda}'\underset{\sim}{Q}\underset{\sim}{\lambda})/\mathrm{Tr}(\underset{\sim}{Q}\underset{\sim}{\Gamma})](H_p(\chi_p^2(\alpha);\Delta)-2H_{p+2}(\chi_p^2(\alpha);\Delta)+H_{p+4}(\chi_p^2(\alpha);\Delta)).$$

Therefore, from (4.3) through (4.6), we obtain that

$$\rho^{PT}(\underset{\sim}{\lambda}) = \lim_{n\to\infty}\{ \rho(\underset{\sim}{U}_n^{PT}, \underset{\sim}{\theta}_n) | K_n \} \tag{4.7}$$

$$= \{1- H_{p+2}(\chi_p^2(\alpha);\Delta)\} +[(\underset{\sim}{\lambda}'\underset{\sim}{Q}\underset{\sim}{\lambda})/\mathrm{Tr}(\underset{\sim}{Q}\underset{\sim}{\Gamma})]\{2H_{p+2}(\chi_p^2(\alpha);\Delta)-H_{p+4}(\chi_p^2(\alpha);\Delta)\}.$$

Note that at $\underset{\sim}{\lambda} = \underset{\sim}{0}$, (4.7) reduces to $1-H_{p+2}(\chi_p^2(\alpha);0) = \alpha+ h_{p+2}(\chi_p^2(\alpha))$, where $h_q(x)$ stands for the density function corresponding to $H_q(x;0)$, and hence, at $\underset{\sim}{\lambda} = \underset{\sim}{0}$, (4.7) is strictly less than 1 for every $\alpha \in (0,1)$ and $p \geq 1$. Also, note that $2H_{p+2}(x;\Delta)-H_{p+4}(x;\Delta)$ is a positive finite quantity (bounded by 1) which converges to 0 if x or Δ goes to $+\infty$. Hence, there exists a closed elliptical region E^* with centre $\underset{\sim}{0}$, such that for every $\underset{\sim}{\lambda} \in E^*$, $\rho^{PT}(\underset{\sim}{\lambda}) < 1$. Further, as $\underset{\sim}{\lambda}$ moves away from $\underset{\sim}{0}$ (i.e., $\Delta \to \infty$), (4.7) converges to 1. Finally, for $\underset{\sim}{\lambda} \notin E^*$, though (4.7) may be greater than 1, it is, nevertheless, bounded and depending on p, $\underset{\sim}{Q}$ and $\underset{\sim}{\Gamma}$, this upper bound (i.e., $\sup_{\underset{\sim}{\lambda}\in E} \rho^{PT}(\underset{\sim}{\lambda})$) is usually quite close to 1. Some numerical values tabulated in Sen and Saleh (1979), in a different context , may throw additional light on this for specific values of p.

However, it may be noted that for all λ such that $\delta = \lambda' Q \lambda \geq \text{Tr}(Q\Gamma)$, $\rho^{PT}(\lambda) \geq 1 - H_{p+2}(\chi_p^2(\alpha);\Delta) + [2H_{p+2}(\chi_p^2(\alpha);\Delta) - H_{p+4}(\chi_p^2(\alpha);\Delta)] = 1 + H_{p+2}(\chi_p^2(\alpha);\Delta) - H_{p+2}(\chi_p^2(\alpha);\Delta) \geq 1$, though the excess over 1 may be quite small for moderate values of $\lambda' Q \lambda$. Hence, we may conclude from the above discussion that

$$\rho^{PT}(\lambda) \begin{smallmatrix} < \\ = \\ > \end{smallmatrix} 1 \text{ according as } \lambda' Q \lambda \begin{smallmatrix} < \\ = \\ > \end{smallmatrix} \delta^{o}, \tag{4.8}$$

where $0 < \delta^{o} < \text{Tr}(Q\Gamma)$. On the other hand, by (2.9),

$$\rho(\lambda) = \lim_{n\to\infty}\{\rho(U_n, \theta_n) \mid K_n\} = 1, \forall \lambda \in E^p. \tag{4.9}$$

Hence, we conclude that <u>for local alternatives in</u> (4.1), <u>with respect to the PTE</u>, U_n <u>is not asymptotically inadmissible in the light of the criterion in</u> (2.13). Each one performs better than the other for a subspace of E^p for λ.

Let us next consider the estimator U_n^{JS} in (2.16), where we may need a more stringent regularity condition for the valid computation of the asymptotic risk. Note that under $\{K_n\}$,

$$\begin{aligned} n(U_n^{JS} - \theta_n)' Q (U_n^{JS} - \theta_n) &= n(U_n - \theta_n)' Q (U_n - \theta_n) \\ &\quad - 2cd_n + 2cd_n \mathcal{L}_n^{-1}(n\, U_n' \hat{\Gamma}_n^{-1} \lambda) + c^2 d_n^2 \mathcal{L}_n^{-2} \mathcal{L}_n^{*}, \end{aligned} \tag{4.10}$$

where $\mathcal{L}_n$ is defined in (3.5) and $\mathcal{L}_n^*$ after (3.6). Therefore, whenever the expectations on the right hand side exist, we would have

$$\begin{aligned} \rho^{JS}(\lambda) &= \lim_{n\to\infty}\{\rho(U_n^{JS}, \theta_n) \mid K_n\} \\ &= \rho(\lambda) - 2c[\text{Tr}(Q\Gamma)]^{-1} E(d_n \mid K_n) + 2[\text{Tr}(Q\Gamma)]^{-1} E(d_n \mathcal{L}_n^{-1} n^{\frac{1}{2}} U_n' \hat{\Gamma}_n^{-1} \lambda \mid K_n) \\ &\quad + c^2 [\text{Tr}(Q\Gamma)]^{-1} E(d_n^2 \mathcal{L}_n^{-1} (\mathcal{L}_n^* / \mathcal{L}_n) \mid K_n). \end{aligned} \tag{4.11}$$

Now, under (3.3) and (4.1),

$$E(d_n \mid K_n) = E(ch_p(Q\hat{\Gamma}_n) \mid K_n) \to ch_p(Q\Gamma), \text{ as } n \to \infty. \tag{4.12}$$

Also, by the Cauchy-Schwarz inequality

$$\begin{aligned} (d_n \mathcal{L}_n^{-1} n^{\frac{1}{2}} U_n' \hat{\Gamma}_n^{-1} \lambda)^2 &\leq d_n^2 \mathcal{L}_n^{-2} (n U_n' \hat{\Gamma}_n^{-1} U_n)(\lambda' \Gamma_n^{-1} \lambda) \\ &= d_n \mathcal{L}_n^{-1} [(\lambda' \hat{\Gamma}_n^{-1} \lambda)/(\lambda' Q \lambda)] d_n (\lambda' Q \lambda) \leq d_n (\lambda' Q \lambda) \mathcal{L}_n^{-1}. \end{aligned} \tag{4.13}$$

Finally,

$$d_n(\mathcal{L}_n^* / \mathcal{L}_n) \leq [ch_p(Q\hat{\Gamma}_n)][ch_1(Q^{-1}\hat{\Gamma}_n^{-1})] = 1. \tag{4.14}$$

We may note that under (3.3) and $\{K_n\}$ in (4.1), by virtue of the asymptotic normality of $n^{\frac{1}{2}}(\mathbf{U}_n - \boldsymbol{\theta}_n)$ and consistency of $\hat{\mathbf{\Gamma}}_n$,

$$d_n \mathcal{L}_n^{-1} \text{ conveeges in law to } ch_p(\mathbf{Q\Gamma})\, \chi_{p,\Delta}^{-2} \quad ; \; \Delta = \boldsymbol{\lambda}'\mathbf{\Gamma}^{-1}\boldsymbol{\lambda} \; , \tag{4.15}$$

where $\chi_{p,\Delta}^{2}$ is a random variable having the noncentral chi square d.f. with p degrees of freedom and noncentrality parameter Δ ,i.e., $P\{\chi_{p,\Delta}^{2} \leq x\} = H_p(x;\, \Delta)$ for every $x \in [0,\infty)$. We now make a more stringent assumption that under $\{K_n\}$,

$$d_n \mathcal{L}_n^{-1} \text{ converges in } L_1\text{-norm to } ch_p(\mathbf{Q\Gamma})\, \chi_{p,\Delta}^{-2} \; . \tag{4.16}$$

Under (3.3) and (4.16), we obtain from (4.11) through (4.14) that

$$\rho^{JS}(\boldsymbol{\lambda}) = 1 - 2c[Tr(\mathbf{Q\Gamma})]^{-1} ch_p(\mathbf{Q\Gamma}) + 2c[ch_p(\mathbf{Q\Gamma})/Tr(\mathbf{Q\Gamma})]E\{(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\boldsymbol{\omega}\}$$
$$+ c^2[ch_p(\mathbf{Q\Gamma})]^2[Tr(\mathbf{Q\Gamma})]^{-1}E\{(\mathbf{W}'\mathbf{W})^{-2}\mathbf{W}'\mathbf{A}\mathbf{W}\} \; , \tag{4.17}$$

where $\mathbf{W}$ has the normal distribution with mean vector $\boldsymbol{\omega} = \mathbf{\Gamma}^{-\frac{1}{2}}\boldsymbol{\lambda}$ and dispersion matrix $\mathbf{I}$, and $\mathbf{A} = \mathbf{\Gamma}^{-\frac{1}{2}}\mathbf{Q}^{-1}\mathbf{\Gamma}^{-\frac{1}{2}}$. Using the Stein-type results in Section 2 of Sclove, Morris and Radhakrishnan (1972), we may simplify (4.17) as

$$\rho^{JS}(\boldsymbol{\lambda}) = 1 - 2c[ch_p(\mathbf{Q\Gamma})/Tr(\mathbf{Q\Gamma})] + 2c[ch_p(\mathbf{Q\Gamma})/Tr(\mathbf{Q\Gamma})]\, E(\chi_{p+2,\Delta}^{-2})$$
$$+ c^2[ch_p(\mathbf{Q\Gamma})]^2[Tr(\mathbf{Q\Gamma})]^{-1}\{Tr(\mathbf{Q\Gamma})^{-1}E\chi_{p+2,\Delta}^{-4} + \Delta^* E\chi_{p+4,\Delta}^{-4}\}, \tag{4.18}$$

where Δ and Δ^* are defined in (3.7) and (4.4). Now, from the results of Berger et al (1977), it follows that the right hand side of (4.18) is smaller than one whenever $0 < c < 2(p-2)$, $p > 2$. Therefore, we conclude that for $p > 2$, under (3.3) and (4.16), for local alternatives in (4.1), U_n is asymptotically inadmissible whenever in (2.16), c is contained in $(0,\, 2(p-2))$.

In the above context, the L_1-convergence result in (4.16) plays a vital role. This may, however, be not really needed if we consider the other estimators $\mathbf{U}_n^S$ and $\mathbf{U}_n^{S*}$. Basically, we would like to establish the asymptotic inadmissibility of $\mathbf{U}_n$ for local alternatives under (3.3) alone. Towards this, we may note that by (2.16) and (2.18),

$$n(\underset{\sim n}{U}^S_n - \underset{\sim n}{\theta})'\underset{\sim}{Q}(\underset{\sim n}{U}^S - \underset{\sim n}{\theta}) = (\underset{\sim}{\lambda}'\underset{\sim\sim}{Q\lambda})I(\mathcal{L}_n < \varepsilon) + \tag{4.19}$$

$$I(\mathcal{L}_n \geq \varepsilon)\{ n(\underset{\sim n}{U}-\underset{\sim n}{\theta})'\underset{\sim}{Q}(\underset{\sim n}{U}-\underset{\sim n}{\theta}) - 2cd_n + 2cd_n\mathcal{L}_n^{-1}n^{\frac{1}{2}}\underset{\sim n}{U}'\underset{\sim n}{\hat{\Gamma}}^{-1}\underset{\sim}{\lambda} + c^2d_n^2\mathcal{L}_n^*/\mathcal{L}_n\},$$

so that by (2.8) and (4.19), under (3.3) and $\{K_n\}$ in (4.1),

$$\rho^S(\underset{\sim}{\lambda}) = \lim_{n\to\infty}\{\rho(\underset{\sim n}{U}^S, \underset{\sim n}{\theta}) \mid K_n\}$$
$$= [(\underset{\sim}{\lambda}'\underset{\sim\sim}{Q\lambda})/Tr(\underset{\sim\sim}{Q\Gamma})]H_p(\varepsilon;\Delta) + \{1- H_{p+2}(\varepsilon;\Delta)\} - [(\underset{\sim}{\lambda}'\underset{\sim\sim}{Q\lambda})/Tr(\underset{\sim\sim}{Q\Gamma})][H_p(\varepsilon;\Delta) -$$
$$2H_{p+2}(\varepsilon;\Delta) + H_{p+4}(\varepsilon;\Delta)] + 2c\lim_{n\to\infty} E\{ I(\mathcal{L}_n \geq \varepsilon)d_n\mathcal{L}_n^{-1}n^{\frac{1}{2}}\underset{\sim n}{U}'\underset{\sim n}{\hat{\Gamma}}^{-1}\underset{\sim}{\lambda} \mid K_n\}/Tr(\underset{\sim\sim}{Q\Gamma})$$
$$+ c^2\lim_{n\to\infty} E\{ d_n^2 I(\mathcal{L}_n \geq \varepsilon)\mathcal{L}_n^{-2}\mathcal{L}_n^* \mid K_n\}/Tr(\underset{\sim\sim}{Q\Gamma}), \tag{4.20}$$

where results similar to the ones in (4.6) were used for $\underset{\sim n}{U}$. Note that for $\mathcal{L}_n \geq \varepsilon$, $\mathcal{L}_n^{-1}$ is bounded from above by $\varepsilon^{-1}(< \infty)$, while (4.15) continues to hold (without requiring (4.16)) for this region too. Hence, we obtain that under (3.3) and $\{K_n\}$ in (4.1),

$$\lim_{n\to\infty} E\{ I(\mathcal{L}_n \geq \varepsilon)d_n\mathcal{L}_n^{-1}n^{\frac{1}{2}}\underset{\sim n}{U}'\underset{\sim n}{\hat{\Gamma}}^{-1}\underset{\sim}{\lambda} \mid K_n\}/Tr(\underset{\sim\sim}{Q\Gamma})$$
$$= [ch_p(\underset{\sim\sim}{Q\Gamma})/Tr(\underset{\sim\sim}{Q\Gamma})]\{ E(\chi^{-2}_{p+2,\Delta}) - E(I(\underset{\sim}{W}'\underset{\sim}{W} < \varepsilon)(\underset{\sim}{W}'\underset{\sim}{W})^{-1}\underset{\sim}{W}'\underset{\sim}{\omega}) \} \tag{4.21}$$

and

$$\lim_{n\to\infty} E\{ I(\mathcal{L}_n \geq \varepsilon)\, d_n^2\mathcal{L}_n^{-2}\mathcal{L}_n^* \mid K_n\}/Tr(\underset{\sim\sim}{Q\Gamma})$$
$$= [ch_p(\underset{\sim\sim}{Q\Gamma})]^2[tr(\underset{\sim\sim}{Q\Gamma})]^{-1}\{ Tr(\underset{\sim}{Q}^{-1}\underset{\sim}{\Gamma}^{-1})[E(\chi^{-4}_{p+2,\Delta})] + \Delta^* E(\chi^{-4}_{p+4,\Delta})$$
$$- E(I(\underset{\sim}{W}'\underset{\sim}{W} < \varepsilon)(\underset{\sim\sim}{W'W})^{-2}\underset{\sim}{W}'\underset{\sim\sim}{AW}) \} \tag{4.22}$$

where $\underset{\sim}{W}$, $\underset{\sim}{\omega}$ and $\underset{\sim}{A}$ are all defined as in (4.17). Therefore, from (4.20), (4.21) and (4.22), we obtain that

$$\rho^S(\underset{\sim}{\lambda}) = \rho^{JS}(\underset{\sim}{\lambda}) + [(\underset{\sim}{\lambda}'\underset{\sim\sim}{Q\lambda})/Tr(\underset{\sim\sim}{Q\Gamma})][2H_{p+2}(\varepsilon;\Delta) - H_{p+4}(\varepsilon;\Delta)] +$$
$$2c[ch_p(\underset{\sim\sim}{Q\Gamma})/Tr(\underset{\sim\sim}{Q\Gamma})][H_p(\varepsilon;\Delta) - E(I(\underset{\sim}{W}'\underset{\sim}{W} < \varepsilon)(\underset{\sim}{W}'\underset{\sim}{W})^{-1}\underset{\sim}{W}'\underset{\sim}{\omega})]$$
$$-c^2[(ch_p(\underset{\sim\sim}{Q\Gamma}))^2/Tr(\underset{\sim\sim}{Q\Gamma})]E(I(\underset{\sim}{W}'\underset{\sim}{W} < \varepsilon)(\underset{\sim}{W}'\underset{\sim}{W})^{-2}\underset{\sim}{W}'\underset{\sim\sim}{AW}), \tag{4.23}$$

where $\rho^{JS}(\underset{\sim}{\lambda})$ is defined by (4.18). Also, note that

$$(\underset{\sim}{\lambda}'\underset{\sim\sim}{Q\lambda})/Tr(\underset{\sim\sim}{Q\Gamma}) = [(\underset{\sim}{\lambda}'\underset{\sim\sim}{Q\lambda})/(\underset{\sim}{\lambda}'\underset{\sim}{\Gamma}^{-1}\underset{\sim}{\lambda})][Tr(\underset{\sim\sim}{Q\Gamma})]^{-1}(\underset{\sim}{\lambda}'\underset{\sim}{\Gamma}^{-1}\underset{\sim}{\lambda}) \leq \underset{\sim}{\lambda}'\underset{\sim}{\Gamma}^{-1}\underset{\sim}{\lambda} = \Delta. \tag{4.24}$$

Further, for every $q \geq 1$, $\varepsilon > 0$ and $\delta \geq 0$,

$$\delta H_q(\varepsilon;\delta) = \delta e^{-\delta/2} \Sigma_{r=0}^{\infty} (\delta/2)^r(1/r!)H_{q+2r}(\varepsilon;0)$$
$$\leq \delta e^{-\delta/2} \Sigma_{r=0}^{\infty} (\delta/2)^r(1/r!)(\varepsilon/2)^{q/2+r}(\Gamma(q/2 + r))^{-1}$$
$$\leq \delta(\varepsilon/2)^{q/2}.\delta \exp(-\delta/2)/(\Gamma(q/2)). \qquad (4.25)$$

Also,

$$|E[I(\underset{\sim}{W}'\underset{\sim}{W} <\varepsilon)(\underset{\sim}{W}'\underset{\sim}{W})^{-1}\underset{\sim}{W}'\underset{\sim}{\omega}]| \leq E[I(\underset{\sim}{W}'\underset{\sim}{W}< \varepsilon)(\underset{\sim}{W}'\underset{\sim}{W})^{-\frac{1}{2}} (\underset{\sim}{\omega}'\underset{\sim}{\omega})^{\frac{1}{2}}]$$
$$\leq \varepsilon^{-\frac{1}{2}}\Delta^{\frac{1}{2}}H_p(\varepsilon;\Delta) \qquad (4.26)$$

and

$$E[I(\underset{\sim}{W}'\underset{\sim}{W} <\varepsilon)(\underset{\sim}{W}'\underset{\sim}{W})^{-2}\underset{\sim}{W}'\underset{\sim}{A}\underset{\sim}{W}] \leq \varepsilon^{-1}Tr(\underset{\sim}{A})H_p(\varepsilon;\Delta) = O(\varepsilon^{p/2-1}), \qquad (4.27)$$

uniformly in $\underset{\sim}{\omega}$ (i.e., inΔ). Hence, from (4.23) through (4.27), we conclude that for $\underset{\sim}{U}_n^{JS}$ with $c \varepsilon (0, 2(p-2))$, for every $\eta >0$, there exists a $\delta > 0$, such that for $\underset{\sim}{U}_n^{S}$ with $0 < \varepsilon < \delta$,

$$\rho^S(\underset{\sim}{\lambda}) \leq \rho^{JS}(\underset{\sim}{\lambda}) + \eta \leq 1 \text{ , uniformly in } \underset{\sim}{\lambda} \varepsilon E^p. \qquad (4.28)$$

This proves the asymptotic inadmissibility of $\underset{\sim}{U}_n$ for local alternatives $\{K_n\}$ in (4.1), and this does not need the L_1-convergence in (4.16). This also illustrate the applicability and utility of the adaptive estimator $\underset{\sim}{U}_n^S$; which may not need the stringent condition (4.16), needed for $\underset{\sim}{U}_n^{JS}$.

Finally, we may note that by definition in (2.19)

$$[n(\underset{\sim}{U}_n^{S*}-\underset{\sim}{\theta}_n)'\underset{\sim}{Q}(\underset{\sim}{U}_n^{S*}-\underset{\sim}{\theta}_n) - n(\underset{\sim}{U}_n^{S} -\underset{\sim}{\theta}_n)'\underset{\sim}{Q}(\underset{\sim}{U}_n'-\underset{\sim}{\theta}_n)]$$
$$= I(\mathcal{L}_n <\varepsilon)[c^2\varepsilon^{-1}d_n^2(\mathcal{L}_n^*/\mathcal{L}_n) - 2c \varepsilon^{-\frac{1}{2}} d_n\mathcal{L}_n^{\frac{1}{2}}] \qquad (4.29)$$

and this is bounded by

$$c(\varepsilon^{-1}c + 2) d_n I(\mathcal{L}_n < \varepsilon). \qquad (4.30)$$

Therefore, we obtain that under (3.3) and $\{K_n\}$ in (4.1)

$$\rho^{S*}(\underset{\sim}{\lambda}) = \lim_{n\to\infty} \{ \rho(\underset{\sim}{U}_n^{S*},\underset{\sim}{\theta}_n) \mid K_n \}$$
$$\leq \rho^S(\underset{\sim}{\lambda}) + c(\varepsilon^{-1}c+2) \lim_{n\to\infty} E[d_nI(\mathcal{L}_n < \varepsilon) \mid K_n]/Tr(\underset{\sim}{Q}\underset{\sim}{\Gamma}). \qquad (4.31)$$
$$\leq \rho^S(\underset{\sim}{\lambda}) + c(\varepsilon^{-1}c+2)H_p(\varepsilon;\Delta) = \rho^S(\underset{\sim}{\lambda}) + O(\varepsilon^{p/2-1}), \qquad (4.31)$$

uniformly in $\underset{\sim}{\lambda}$. Thus, for $p > 2$, choosing ε sufficiently small, $\underset{\sim}{U}_n^S$ and $\underset{\sim}{U}_n^{S*}$ can be made asymptotically risk equivalent for local alternatives.

5. SOME GENERAL REMARKS

The results on U-statistics treated here apply equally well to sample mean vectors. In that case, the jackknife estimator $\hat{\underset{\sim}{\Gamma}}_n$ reduces to the sample covariance matrix $\underset{\sim}{S}_n$. Hence, for the James-Stein type estimator in (2.12), the asymptotic admissibility results apply for possibly nonnormal F as well, provided (4.16) holds. For normal F, this condition may easily be verified by using the Wishart property of $\underset{\sim}{S}_n$ and the stochastic independence of $\bar{\underset{\sim}{X}}_n$ and $\underset{\sim}{S}_n$. For non-normal F, this may be quite involved. Also, if $\bar{\underset{\sim}{X}}_n$ assumes the null value $\underset{\sim}{0}$ with a positive probability, however small it may be, then (4.16) may not hold. This is particularly true for lattice distributions. Thus, the proposed estimators $\underset{\sim}{U}_n^S$ and $\underset{\sim}{U}_n^{S*}$ can not only be used to establish the inadmissibility of $\underset{\sim}{U}_n$ for local alternatives, but also they stand as robust estimators, where (4.16) is not needed. In this context, a natural question may arise : What is an optimal or desirable choice of ε for $\underset{\sim}{U}_n^S$ or $\underset{\sim}{U}_n^{S*}$? Ideally, we need to choose ε so small that (4.28) remains in tact. As such, this choice may also depend on the value of $c(>0)$ used in $\underset{\sim}{U}_n^{JS}$. Also, for very small values of $\varepsilon(>0)$, the convergence rates to the asymptotic limits may be slow. The ideal choice depends on c, F as well as the kernel $\underset{\sim}{\phi}$. However, for moderately large sample sizes, ε can be chosen so small that (4.28) is attained upto a certain margin of difference, and, at the same time, the estimator remains robust against the contribution of the small values of the statistics to the risk function. It is clear that the basic requirement of $p > 2$ (in the normal theory models) remains in tact in the nonparametric case as well. However, for very local alternatives [i.e., in (4.8) $\underset{\sim}{\lambda}'\underset{\sim}{Q}\underset{\sim}{\lambda} \leq \delta^o$] the PTE $\underset{\sim}{U}_n^{PT}$ may render $\underset{\sim}{U}_n$ as inadmissible, even for p=1,2. However, in general, the PTE is not admissible (even for local alternatives) whereas the proposed ones are so.

REFERENCES

Berger, J.O. (1980a). Statistical Decision Theory. Springer-Verlag, New York.

Berger, J.O. (1980b). 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 Gleser, L.(1977). Minimax estimation of a normal mean vector for arbitrary quadratic loss and unknown covariance matrix. Ann.Statist.5, 763-771.

Ghosh, M. (1983). Estimates of multiple Poisson means: Bayes and Empirical Bayes. Statistics & Decision . 1, 183-195.

Ghosh, M., Hwang,J. and Tsui, K.W. (1983). Construction of improved estimators for multiparameter estimation in discrete exponential families. Ann.Statist.11, to appear (with discussion).

Halmos, P.R. (1946). The theory of unbiased estimation. Ann.Math. Statist. 17, 34-44.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. Ann.Math. Statist.19, 293-325.

James, W. and Stein, C.(1961). Estimation with quadratic loss. Proc.Fourth Berkeley Symp. Math. Statist. Prob. 1, 361-379.

Sclove, S.L., Morris, C. and Radhakrishnan, R. (1972). Non-optimality of preliminary-test estimators for the mean of a multivariate normal distribution. Ann.Math. Statist. 43, 1481-1490.

Sen, P.K. (1960). On some convergence properties of U-statistics. Calcutta Statist. Asso. Bull.10,1-18.

Sen, P.K. (1981). Sequential Nonparametric : Invariance Pronciples and Statistical Inference. John Wiley , New York.

Sen, P.K. and Ghosh, M. (1981). Sequential point estimation of estimable parameters based on U-statistics. Sankhya, Ser.A 43, 331-344.

Sen, P.K. and Saleh, A.K.Md.E. (1979). Nonparametric estimation of location parameter after a preliminary test on regression in the multivariate case. Jour. Multivar. Anal. 9, 322-331.

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.