ON SOME ASYMPTOTICALLY NONPARAMETRIC COMPETITORS OF HOTELLING'S T^2 [1]

By Peter J. Bickel

University of California, Berkeley

1. Summary. In a previous paper [2] the author investigated some alternative estimates of shift in the p-variate one-sample problem. This paper examines the properties of tests for shift similar to Hotelling's T^2, based on (I) asymptotically normal estimates, in particular those of the type considered in [2], and (II) the originating univariate test statistics of the latter group. The notation, similar to that used in [2], and the tests are introduced in Section 2. In the third section the asymptotic distribution of such test statistics for sequences of alternatives tending to $\mathbf{0}$ as $n^{-\frac{1}{2}}$ is found to be noncentral χ^2. The tests of type I and, often simpler, tests of type II have the same asymptotic distribution. In Section 4 we find that the Pitman efficiency of two such tests depends, in general, on the direction in which the origin is approached. The efficiency, in terms of "generalized variance," of the estimates of [2] lies between the maximum and minimum Pitman efficiencies of the corresponding tests of type I (maximum and minimum being taken over direction).

This "generalized variance" efficiency is found to be equal to the efficiency of the tests as defined in terms of a criterion of goodness introduced by Isaacson [9] (D-optimality). In case the coordinates are identically distributed, *if correlation effects alone are taken into account*, it is shown that whatever two tests are considered, there always exists a direction in which one improves the other.

Section 5 continues with a discussion of the tests based on the estimates introduced in [2] in relation to Hotelling's T^2. Their desirable properties and pathologies are found to be similar to those of the parent estimates. The remaining sections deal with the case $p = 2$. Under normality the minimum Pitman efficiency with respect to T^2 of the tests mentioned above behaves like the efficiency of the parent estimates with respect to the mean.

2. Introduction and definitions. As in [2] let $\mathbf{X}_i = (X_{1i}, \cdots, X_{pi})$, $1 \leqq i \leqq n$, be a sample from $F(x_1 - \theta_1, \cdots, x_p - \theta_p)$ where F is symmetric about $\mathbf{0}$, the zero vector, $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_p)$ is unknown, and F has absolutely continuous marginals $F_1, \cdots, F_p$ with corresponding densities $f_1, \cdots, f_p$.

We are, in this paper, concerned with the properties of test statistics for the hypothesis $H: \boldsymbol{\theta} = \mathbf{0}$ versus the alternatives $\boldsymbol{\theta} \neq \mathbf{0}$.

$\dot{\mathbf{X}}_n$, $\mathbf{M}_n$ and $\mathbf{W}_n$ denote, as before, the vector mean, median and median of averages of pairs. In addition, $\hat{\boldsymbol{\theta}}_n$ will be used as a generic notation for an estimate

Received 14 April 1964.

[1] This research was based in part on the author's doctoral dissertation submitted to the University of California, Berkeley, and was written with the partial support of the National Science Foundation, Grant GP-10.

of $\boldsymbol{\theta}$ of the form $(\hat{\theta}_n(X_{11}, \cdots, X_{1n}), \cdots, \hat{\theta}_n(X_{p1}, \cdots, X_{pn}))$ where $\hat{\theta}_n(X_{i1}, \cdots, X_{in})$ is a univariate translation invariant estimate of θ_i.

Let $\mathcal{C}$ denote the class of (sequences of) estimates $\hat{\theta}_n$ of the above type which are asymptotically nonsingular p-variate normal under the hypothesis $\boldsymbol{\theta} = \mathbf{0}$, i.e., $\mathcal{L}(\mathbf{0}, n^{\frac{1}{2}}\hat{\boldsymbol{\theta}}_n) \to \Phi[\boldsymbol{\theta}, \|\beta_{ij}\|]$ as $n \to \infty$. Here $\Phi[\boldsymbol{\theta}, \|\beta_{ij}\|]$ denotes the p-variate normal distribution with mean $\boldsymbol{\theta}$ and covariance matrix $\|\beta_{ij}\|$, and $\mathcal{L}(\boldsymbol{\theta}, \mathbf{Z})$ the law of $\mathbf{Z}$ under $P_{\boldsymbol{\theta}}$. This class includes $\hat{\theta}_n$ which are generated as maximum likelihood estimates (under suitable regularity conditions), the very wide class of estimates introduced by Huber ([8], Lemma 5) and those estimates of the type introduced by Hodges and Lehmann in [7], whose associated univariate test statistics are asymptotically normal. By this we mean, following Hodges and Lehmann, that,

$$(2.1)\qquad \begin{aligned} 2\hat{\theta}_n(X_{i1}, \cdots, X_{in}) &= \sup \{\theta: h(X_{i1} - \theta, \cdots, X_{in} - \theta) > \mu_n\} \\ &\quad + \inf \{\theta: h(X_{i1} - \theta, \cdots, X_{in} - \theta) < \mu_n\} \end{aligned}$$

and that, if $\mathbf{h}(\mathbf{X}_1, \cdots, \mathbf{X}_n)$ denotes the vector whose ith component is $h(X_{i1}, \cdots, X_{in})$ and $\boldsymbol{\theta}_n = n^{-\frac{1}{2}}\mathbf{a}$, then

$$(2.2)\qquad \mathcal{L}\{\boldsymbol{\theta}_n, n^{\frac{1}{2}}[\mathbf{h}(\mathbf{X}_1, \cdots, \mathbf{X}_n) - \boldsymbol{\mu}_n]\} \to \Phi_{[(c_1a_1, \cdots, c_pa_p), \|\pi_{ij}\|]}.$$

Here, $\boldsymbol{\mu}_n = (\mu_n, \cdots, \mu_n) = E_{\mathbf{0}}[\mathbf{h}(\mathbf{X}_1, \cdots, \mathbf{X}_n)]$, $\mathbf{a} = (a_1, \cdots, a_p)$, and the c_i and $\|\pi_{ij}\|$ are fixed. $E_{\boldsymbol{\theta}}$ denotes that the expectation is being taken under $P_{\boldsymbol{\theta}}$. $\hat{\boldsymbol{\theta}}_n$ of the above type will be said to constitute class $\mathcal{D}$.

A simple extension of Theorem 5 of [7], to be stated in the next section, shows that $\mathcal{D} \subset \mathcal{C}$.

Finally we remark that if $\hat{\theta}_n \varepsilon \mathcal{C}$ and $n^{\frac{1}{2}}\boldsymbol{\theta}_n \to \boldsymbol{\delta}$ then $\mathcal{L}(\boldsymbol{\theta}_n, n^{\frac{1}{2}}\hat{\boldsymbol{\theta}}_n) \to \Phi_{[\boldsymbol{\delta}, \|\beta_{ij}\|]}$.

Under the supplementary assumption that F is p-variate nonsingular normal with unknown covariance matrix $\|\sigma_{ij}\|$, the test that is usually employed in the testing situation we are interested in is Hotelling's T^2. This test takes the form: Reject H if $T_n^2 = n(\dot{\mathbf{X}}_n\|\hat{\sigma}^{ij}\|\dot{\mathbf{X}}_n') > c$, where $\|\hat{\sigma}^{ij}\| = \|\hat{\sigma}_{ij}\|^{-1}$,

$$\hat{\sigma}_{ij} = (1/n - 1)\textstyle\sum_{\alpha=1}^{n} (X_{i\alpha} - X_{i\cdot})(X_{j\alpha} - X_{j\cdot})$$

and $X_{k\cdot} = (1/n)\sum_{\alpha=1}^{n} X_{k\alpha}$.

The above expressions are well defined, since, when F is nonsingular normal, $\|\hat{\sigma}_{ij}\|$ is known to be nonsingular a.s., and in fact has a Wishart distribution. In general, $\|\hat{\sigma}_{ij}\|$ is invertible a.s. if and only if $\|\sigma_{ij}\|$ the covariance matrix of $\mathbf{X}_1$ is nonsingular. We may then define T_n^2 as a test statistic for H in the framework of our model if $\|\sigma_{ij}\|$ is invertible.

Suppose now that $\hat{\boldsymbol{\theta}}_n \varepsilon \mathcal{C}$ and there exists a consistent sequence $\|\hat{\beta}^{ij}\|$ of estimates of $\|\beta^{ij}\| = \|\beta_{ij}\|^{-1}(\|\beta_{ij}\|$ nonsingular) i.e., $P_{\boldsymbol{\theta}} \lim_n \|\hat{\beta}^{ij}\| = \|\beta^{ij}\|$ as $\boldsymbol{\theta} \to \mathbf{0}$.

Then a natural generalization of Hotelling's idea leads us to consider tests of the form "Reject if $n\hat{\boldsymbol{\theta}}_n\|\hat{\beta}^{ij}\|\hat{\boldsymbol{\theta}}_n' > c$." We shall refer to such tests as Hotelling tests of type I. For $\hat{\boldsymbol{\theta}}_n \varepsilon \mathcal{D}$ a different class of test statistics readily presents itself. Let $\mathbf{h}(\mathbf{X}_1, \cdots, \mathbf{X}_n)$ be the vector of test statistics associated with $\hat{\boldsymbol{\theta}}_n$ by (2.1). Then, (2.2) suggests, if $\|\hat{\pi}^{ij}\|$ is a sequence of consistent estimates of $\|\pi_{ij}\|^{-1} =$

$\|\pi^{ij}\|$, ($\|\pi_{ij}\|$ nonsingular) using as a critical region "$n[\mathbf{h}(\mathbf{X}_1, \cdots, \mathbf{X}_n) - \boldsymbol{\mu}_n] \cdot \|\hat{\pi}^{ij}\|[\mathbf{h}(\mathbf{X}_1, \cdots, \mathbf{X}_n) - \boldsymbol{\mu}_n]' > c$." Such sequences of tests will be referred to as Hotelling tests of type II. There is, of course, a natural relation between Hotelling tests of type I and II whose corresponding bases are related by (2.1), which will be given in the next section.

From [2] it follows that (a) $\dot{\mathbf{X}}_n$, (b) $\mathbf{W}_n$, (c) $\mathbf{M}_n$ are members of $\mathfrak{D}$ with covariance matrices $\|\beta_{ij}\|$ given respectively by

$$
(2.3)\quad
\begin{aligned}
&\text{(a)} && \beta_{ij}^{(1)} = \sigma_{ij} = E_0(X_{i1} X_{j1}),\\
&\text{(b)} && \beta_{ij}^{(2)} = \frac{E_0[F_i(X_{i1}) - \frac{1}{2})(F_j(X_{j1}) - \frac{1}{2})]}{\int_{-\infty}^{\infty} f_i^2(x)\,dx \int_{-\infty}^{\infty} f_j^2(x)\,dx},\\
&\text{(c)} && \beta_{ij}^{(3)} = \frac{E_0[(I^+(X_{i1}) - \frac{1}{2})(I^+(X_{j1}) - \frac{1}{2})]}{f_i(0)f_j(0)},
\end{aligned}
$$

where $I^+(x) = 1$ for $x \geqq 0$, 0 otherwise. The corresponding test statistics (a) $\mathbf{h}_1$, (b) $\mathbf{h}_2$, (c) $\mathbf{h}_3$ are given by (see [7]),

$$
(2.4)\quad
\begin{aligned}
&\text{(a)} && \mathbf{h}_1(\mathbf{X}_1, \cdots, \mathbf{X}_n) = \dot{\mathbf{X}}_n,\\
&\text{(b)} && [\mathbf{h}_2(\mathbf{X}_1, \cdots, \mathbf{X}_n)]_k = [2/n(n+1)] \textstyle\sum_{1 \leqq i \leqq j \leqq n} I^+(X_{ki} + X_{kj}),\\
&\text{(c)} && [\mathbf{h}_3(\mathbf{X}_1, \cdots, \mathbf{X}_n)]_k = n^{-1} \textstyle\sum_{i=1}^{n} I^+(X_{ki}),
\end{aligned}
$$

where $[\mathbf{h}]_k$ denotes the kth component of $\mathbf{h}$. These statistics are asymptotically equivalent to the t, Wilcoxon, and sign statistics, respectively.

In the derivation of (2.3) (see [1]) one obtains the asymptotic covariance matrices $\|\pi_{ij}\|$ of $\mathbf{h}_1$, $\mathbf{h}_2$, and $\mathbf{h}_3$ which are given by;

$$
(2.5)\quad
\begin{aligned}
&\text{(a)} && \pi_{ij}^{(1)} = \sigma_{ij},\\
&\text{(b)} && \pi_{ij}^{(2)} = \beta_{ij}^{(2)} \int_{-\infty}^{\infty} f_i^2(x)\,dx \int_{-\infty}^{\infty} f_j^2(x)\,dx,\\
&\text{(c)} && \pi_{ij}^{(3)} = \beta_{ij}^{(3)} f_i(0) f_j(0).
\end{aligned}
$$

The Hotelling tests of type I and II for (a), if the natural unbiased estimate of σ^{ij} is used, is Hotelling's T^2. We shall use as generic notation for the tests of type I corresponding to (b) and (c) the symbols $\mathcal{W}_n^2$ and $\mathfrak{M}_n^2$ and for those of type II, $\hat{\mathcal{W}}_n^2$ and $\hat{\mathfrak{M}}_n^2$. The vagueness in our definition rests on the choice of $\|\hat{\beta}^{ij}\|$. As we shall see, however, for $\hat{\mathfrak{M}}_n^2$ and $\hat{\mathcal{W}}_n^2$ natural candidates present themselves, and in any case asymptotically the choice is of no importance.

3. Asymptotic theory. We first note the following lemma which is an immediate consequence of Theorem 4.2 of [1].

LEMMA 3.1. *Suppose* $\mathcal{L}[\boldsymbol{\theta}_n, n^{\frac{1}{2}}(\mathbf{h}_n - \boldsymbol{\mu}_n)] \to \Phi[(c_1\delta_1, \cdots, c_p\delta_p), \|\pi_{ij}\|]$ *as* $\boldsymbol{\theta}_n = n^{-\frac{1}{2}}\boldsymbol{\delta} \to \mathbf{0}$. *Then if* $\hat{\boldsymbol{\theta}}_n$ *and* $\mathbf{h}_n$ *are related by* (2.1), $\mathcal{L}(\boldsymbol{\theta}_n, n^{\frac{1}{2}}\hat{\boldsymbol{\theta}}_n) \to \Phi_{[\boldsymbol{\delta}, \|\pi_{ij}c_ic_j\|]}$.

REMARK. Hence the asymptotic means of $\hat{\mathcal{W}}_n^2$ and $\hat{\mathfrak{M}}_n^2$ are given respectively by $(\delta_1 \int_{-\infty}^{\infty} f_1^2(x)\,dx, \cdots, \delta_p \int_{-\infty}^{\infty} f_p^2(x)\,dx)$ and $(\delta_1 f_1(0), \cdots, \delta_p f_p(0))$.

Let $\chi_p^2(\lambda^2)$ denote the distribution function of the noncentral χ^2 distribution with noncentrality parameter λ^2 and p degrees of freedom.

Then we have

THEOREM 3.1. (a) *Let* $\hat{\boldsymbol{\theta}}_n \varepsilon \mathcal{C}$ *with nonsingular asymptotic covariance* $\|\beta_{ij}\|$, $\boldsymbol{\theta}_n = n^{-\frac{1}{2}}\boldsymbol{\delta} \to \mathbf{0}$. *Then* $\mathcal{L}(\boldsymbol{\theta}_n, n\hat{\boldsymbol{\theta}}_n\|\hat{\beta}^{ij}\|\hat{\boldsymbol{\theta}}_n') \to \chi_p^2(\boldsymbol{\delta}\|\beta^{ij}\|\boldsymbol{\delta}')$.

(b) *Let* $\hat{\boldsymbol{\theta}}_n \varepsilon \mathfrak{D}$, $\hat{\boldsymbol{\theta}}_n \leftrightarrow \mathbf{h}$ *with nonsingular asymptotic covariance* $\|\pi_{ij}\| = \|c_ic_j\beta_{ij}\|$ *and mean* $(c_1\delta_1, \cdots, c_p\delta_p)$.

Then,

$\mathcal{L}\{\boldsymbol{\theta}_n, n[\mathbf{h}(\mathbf{X}_1, \cdots, \mathbf{X}_n) - \boldsymbol{\mu}_n]\|\hat{\pi}^{ij}\|[\mathbf{h}(\mathbf{X}_1, \cdots, \mathbf{X}_n) - \boldsymbol{\mu}_n]'\} \to \chi_p^2(\boldsymbol{\delta}\|\beta^{ij}\|\boldsymbol{\delta}')$.

PROOF. Let $\|r_{ij}\|$ be the asymptotic correlation matrix of $\hat{\boldsymbol{\theta}}_n$. Then $\|\beta^{ij}\| = \|d_{ij}\|\|r_{ij}\|^{-1}\|d_{ij}\|$ where

$$d_{ij} = 0, \qquad i \neq j,$$

$$= 1/\beta_{ii} = c_i/\pi_{ii}, \qquad i = j.$$

Similarly, $\|\pi^{ij}\| = \|d_{ij}^*\|\|r_{ij}\|^{-1}\|d_{ij}^*\|$ where $d_{ij}^* = d_{ij}/c_i$. Hence $(c_1\delta_1, \cdots, c_p\delta_p)\cdot\|\pi^{ij}\|(c_1\delta_1, \cdots, c_p\delta_p)' = \boldsymbol{\delta}\|\beta^{ij}\|\boldsymbol{\delta}'$ and (a) $\Rightarrow$ (b). To prove (a) we remark first that

$$(3.1) \qquad n\hat{\boldsymbol{\theta}}_n\|\hat{\beta}^{ij}\|\hat{\boldsymbol{\theta}}_n' - n\hat{\boldsymbol{\theta}}_n\|\beta^{ij}\|\hat{\boldsymbol{\theta}}_n' = \sum_{i,j=1}^{p}(\hat{\beta}^{ij} - \beta^{ij})n[\hat{\boldsymbol{\theta}}_n]_i[\hat{\boldsymbol{\theta}}_n]_j,$$

which converges to 0 in probability as $n^{\frac{1}{2}}\boldsymbol{\theta}_n \to \boldsymbol{\delta}$ since $\hat{\beta}^{ij} - \beta^{ij}$ converges to 0 in probability and $[n^{\frac{1}{2}}\hat{\boldsymbol{\theta}}_n]_i[n^{\frac{1}{2}}\hat{\boldsymbol{\theta}}_n]_j$ converges in distribution.

We now employ the following standard lemma on convergence in law, to $n^{\frac{1}{2}}\hat{\boldsymbol{\theta}}_n$.

LEMMA. *Let* $\mathbf{X}_n$ *b a sequence of random vectors which converge in law to a random vector* $\mathbf{X}$. *Let* g *be a continuous function on* R^p *to* R^q. *Then* $g(\mathbf{X}_n)$ *converges to* $g(\mathbf{X})$ *in law.*

Hence,

$$(3.2) \qquad \lim_n P(n\hat{\boldsymbol{\theta}}_n\|\beta^{ij}\|\hat{\boldsymbol{\theta}}_n' \leqq x) = P(\mathbf{X}\|\beta^{ij}\|\mathbf{X}' < x),$$

where $\mathbf{X}$ has a $\Phi[\boldsymbol{\delta}, \|\beta_{ij}\|]$ distribution. Since $\mathbf{X}\|\beta^{ij}\|\mathbf{X}'$ has a $\chi_p^2(\boldsymbol{\delta}\|\beta^{ij}\|\boldsymbol{\delta}')$ distribution the result follows from (3.1) and (3.2).

Two important conclusions may be drawn from this theorem. First, the choice of $\hat{\beta}^{ij}$ is of no importance *in the limit* and, secondly, the tests of type I and II are equivalent.

4. Efficiency considerations. We employ a measure of the efficiency of two test statistics with respect to each other due to Pitman. Since there is a slight generalization involved in considering a vector hypothesis we give a definition of this quantity.

DEFINITION 4.1. Let $\{\phi_n\}$, $\{\phi_n^*\}$ be two sequences of test statistics for $H: \boldsymbol{\theta} = \boldsymbol{\theta}_0$ versus $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ which are of asymptotic size α. Let $\boldsymbol{\theta}_n$ be a sequence of alternatives converging to $\boldsymbol{\theta}_0$ and $\beta_n(\boldsymbol{\theta}_n)$, $\beta_n^*(\boldsymbol{\theta}_n)$ be the power of ϕ_n, ϕ_n^* respectively against $\boldsymbol{\theta}_n$. Then if for any two sequences of integers $\{n_i\}$ and $\{N_i\}$ such that $\lim_{i\to\infty}\beta_{n_i}(\boldsymbol{\theta}_i) = \lim_{i\to\infty}\beta_{N_i}^*(\boldsymbol{\theta}_i) = \beta$, where $\beta \neq 0$ or 1, $\lim_{i\to\infty} N_i/n_i$ exists and is unique, we call this limit the Pitman efficiency of ϕ_n with respect to ϕ_n^* for level α, power β and sequence $\boldsymbol{\theta}_n$. It is denoted by $e(\phi_n, \phi_n^*)$ where the dependence on θ_n, α and β is understood.

A well-known theorem of Hannan [5] states that,

If ϕ_n, ϕ_n^* tend in law to $\chi_p^2(\lambda_1^2)$, $\chi_p^2(\lambda_2^2)$ for a given sequence $\{\boldsymbol{\theta}_n\}$, then $e(\phi_n, \phi_n^*) = \lambda_1^2/\lambda_2^2$, which is independent of α but depends on β and the sequence $\boldsymbol{\theta}_n$ through λ_1^2 and λ_2^2.

Now let ϕ_n, ϕ_n^* be two Hotelling tests of type I or II, with associated estimates $\hat{\boldsymbol{\theta}}_n$, $\hat{\boldsymbol{\theta}}_n^*$, such that if $n^{\frac{1}{2}}\boldsymbol{\theta}_n \to \boldsymbol{\delta}$, $\mathcal{L}(\boldsymbol{\theta}_n, \phi_n) \to \chi_p^2(\boldsymbol{\delta}\|\gamma^{ij}\|\boldsymbol{\delta}')$, $\mathcal{L}(\boldsymbol{\theta}_n, \phi_n^*) \to \chi_p^2(\boldsymbol{\delta}\|\lambda^{ij}\|\boldsymbol{\delta}')$.

Then by Hannan's theorem,

$$e(\phi_n, \phi_n^*)(\boldsymbol{\delta}) = \boldsymbol{\delta}\|\gamma^{ij}\|\boldsymbol{\delta}'/\boldsymbol{\delta}\|\lambda^{ij}\|\boldsymbol{\delta}', \tag{4.1}$$

which depends on the direction $\boldsymbol{\delta}$ in which the origin is approached.

On the other hand in [2], we gave as a reasonable definition of the asymptotic efficiency of two consistent asymptotically normal estimates $\hat{\boldsymbol{\theta}}_n$, $\hat{\boldsymbol{\theta}}_n^*$ the ratio of sample sizes required to reach equal asymptotic generalized variances. This quantity will be denoted by $e(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\theta}}_n^*)$. In [2] we found,

$$e(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\theta}}_n^*) = [\det \|\gamma^{ij}\|/\det \|\lambda^{ij}\|]^{p^{-1}} = [\det \|\gamma^{ij}\|\|\lambda^{ij}\|^{-1}]^{p^{-1}}, \tag{4.2}$$

where det denotes determinant.

In the case $p = 1$, Hodges and Lehmann found [7] that $e(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\theta}}_n^*) = e(\phi_n, \phi_n^*)$ is independent of $\boldsymbol{\delta}$.

In general we have

THEOREM 4.1. $\inf_{\boldsymbol{\delta}\neq\mathbf{0}} e(\phi_n, \phi_n^*)(\boldsymbol{\delta}) \leqq e(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\theta}}_n^*) \leqq \sup_{\boldsymbol{\delta}\neq\mathbf{0}} e(\phi_n, \phi_n^*)(\boldsymbol{\delta})$ *with strict inequality holding unless* $\|\lambda_{ij}\| = c\|\gamma_{ij}\|$ *where c is a scalar.*

PROOF. We employ the following classical theorem of Courant (see Bodewig [4]).

THEOREM. *The maximal and minimal values of $\mathbf{x}A\mathbf{x}'/\mathbf{x}B\mathbf{x}'$, where A and B are nonnegative definite, and B is nonsingular, are given by the maximal and minimal eigenvalues of AB^{-1}.*

By (4.2) $e(\hat{\boldsymbol{\theta}}_n, \hat{\boldsymbol{\theta}}_n^*)$ is equal to the geometric mean of the eigenvalues of $\|\gamma^{ij}\|\|\lambda^{ij}\|^{-1}$ which certainly lies between the maximum and minimum eigenvalues, with equality holding if and only if all of the eigenvalues are equal, i.e., $\Leftrightarrow c\|\gamma^{ij}\|\|\lambda_{ij}\| = I$ (the identity) $\Leftrightarrow \|\gamma_{ij}\| = c\|\lambda_{ij}\|$. The theorem is established.

REMARK. If equality holds the efficiency is independent of both $\boldsymbol{\delta}$ and p. A simple sufficient condition for equality is that $\hat{\boldsymbol{\theta}}_n$, $\hat{\boldsymbol{\theta}}_n^*$ have totally symmetric distributions under the hypothesis [see (2)] and that $\mathbf{X}_1$ has identically distributed components.

In the latter case (identically distributed components) we can now state a curious theorem whose gist is that there is always a direction in which one can lose by going to higher dimensions. More specifically, in this case, let $\|r_{ij}\|$ be the asymptotic correlation matrix of $\hat{\boldsymbol{\theta}}_n$, $\|\tilde{r}_{ij}\|$ that of $\hat{\boldsymbol{\theta}}_n^*$, and $\|r^{ij}\|$, $\|\tilde{r}^{ij}\|$ their respective inverses. Then,

$$e(\phi_n, \phi_n^*, \boldsymbol{\delta}) = \boldsymbol{\delta}\|r^{ij}\|\boldsymbol{\delta}'/\boldsymbol{\delta}\|\tilde{r}^{ij}\|\boldsymbol{\delta}' \cdot \lambda_{11}^2/\gamma_{11}^2. \tag{4.3}$$

The first factor of this product may be interpreted as the "multivariate portion" of the efficiency. We have

THEOREM 4.2. *For F with identically distributed components if $e(\phi_n, \phi_n^*, \mathbf{\delta})$ is not constant,*

$$\inf_{\mathbf{\delta}} e(\phi_n, \phi_n^*, \mathbf{\delta}) < \lambda_{11}^2/\gamma_{11}^2 < \sup_{\mathbf{\delta}} e(\phi_n, \phi_n^*, \mathbf{\delta}).$$

PROOF. It suffices to prove one of the inequalities since the other follows by reversing ϕ_n and ϕ_n^*. Suppose that $\inf_{\mathbf{\delta}\neq \mathbf{0}} \mathbf{\delta}\|r^{ij}\|\mathbf{\delta}'/\mathbf{\delta}\|r^{ij}\|\mathbf{\delta}' \geqq 1$. Now the eigenvalues of $\|r^{ij}\|\|\tilde{r}^{ij}\|^{-1}$ are the same as those of $\|\tilde{r}_{ij}\|\|r_{ij}\|^{-1}$ and by the classical theorem of Courant, $1 \leqq \inf_{\mathbf{\delta}\neq\mathbf{0}} \mathbf{\delta}\|r^{ij}\|\mathbf{\delta}'/\mathbf{\delta}\|\tilde{r}^{ij}\|\mathbf{\delta}' = \inf_{\mathbf{\delta}\neq\mathbf{0}} \mathbf{\delta}\|\tilde{r}_{ij}\|\mathbf{\delta}'/\mathbf{\delta}\|r_{ij}\|\mathbf{\delta}'$. This holds, if and only if, $\mathbf{\delta}\,\|r_{ij} - \tilde{r}_{ij}\|\mathbf{\delta}' \geqq 0$ for all $\mathbf{\delta} \neq \mathbf{0}$, that is, if and only if $\|r_{ij} - \tilde{r}_{ij}\|$ is nonnegative definite. But, $\|r_{ij} - \tilde{r}_{ij}\|$ has trace equal to 0. The theorem follows.

In other words, *as far as the "multivariate portion" of the efficiency is concerned*, there always exists a direction in which one does worse than in 1 dimension whatever estimate one chooses (and necessarily one in which one does better). Of course, the efficiency as a whole can still remain above or below 1 for all directions. Illustrations of what can happen are provided by the behavior of $\mathfrak{M}_n^2$, $\mathfrak{W}_n^2$, and T_n^2 given in Sections 5 and 6.

We also remark that if the argument of the theorem is applied to $\|\lambda_{ij}\|$ and $\|\gamma_{ij}\|$ we find, as might be expected, that if $\lambda_{ii} > \gamma_{ii}$ for all i, then

$$\sup_{\mathbf{\delta}\neq\mathbf{0}} e(\phi_n, \phi_n^*, \mathbf{\delta}) > 1,$$

and similarly the inf is <1 if $\lambda_{ii} < \gamma_{ii}$ for all i. It is also interesting to remark that if $\hat{\mathbf{\theta}}_n, \hat{\mathbf{\theta}}_n^*$ are (1) invariant under the nonsingular group of transformations and (2) give uncorrelated estimates for F with uncorrelated components, we may obtain a simple characterization of the maximum and minimum Pitman efficiencies. In this case, it may readily be shown that $\min_{\mathbf{\delta}\neq\mathbf{0}} e(\phi_n, \phi_n^*, \mathbf{\delta})$, $e(\hat{\mathbf{\theta}}_n, \hat{\mathbf{\theta}}_n^*)$, $\max_{\mathbf{\delta}\neq\mathbf{0}} e(\phi_n, \phi_n^*, \mathbf{\delta})$ are constant on the pencil of distributions obtained from a given F by nonsingular transformations of $\mathbf{X}_1$ when $\mathbf{\theta} = \mathbf{0}$. In particular, condition (2) is satisfied on the p-variate normal pencil by any estimates of the structure proposed which satisfy condition (1) and it follows that in this situation, the univariate efficiency is independent of σ^2, and coincides with all three measures of multivariate efficiency whatever be $\|\sigma_{ij}\|$.

There is an interesting connection between $e(\hat{\mathbf{\theta}}_n, \hat{\mathbf{\theta}}_n^*)$ and a criterion of test optimality introduced by Isaacson [9]. We first introduce a reformulation of the structure surrounding asymptotic efficiency considerations in our context. The extension to the general hypothesis testing problem in R^p is immediate.

Let ϕ_n, ϕ_n^* be as above. Define $\beta(\mathbf{\theta}) = P[\chi_p^2(\mathbf{\theta}\|\lambda^{ij}\|\mathbf{\theta}') \geqq x_\alpha]$, $\beta^*(\mathbf{\theta}) = P[\chi_p^2(\mathbf{\theta}\|\gamma^{ij}\|\mathbf{\theta}') \geqq x_\alpha]$ where x_α is the $(1-\alpha)$ percentile of the χ_p^2 distribution. Then, by a slight extension of Theorem 3 it follows that $\beta_n(\mathbf{\theta}_n) \to \beta(\mathbf{\theta})$ if $n^{\frac{1}{2}}\mathbf{\theta}_n \to \mathbf{\theta}$. Equivalently, $\beta_n(n^{-\frac{1}{2}}\mathbf{\theta}) \to \beta(\mathbf{\theta})$ uniformly on compacta in $\mathbf{\theta}$, since $\beta(\mathbf{\theta})$ is continuous. Then, if $\hat{n}/n \to \rho$, $\beta_{\hat{n}}(n^{-\frac{1}{2}}\mathbf{\theta}) \to \beta(\rho^{-\frac{1}{2}}\mathbf{\theta})$. We shall call the function $\beta(\mathbf{\theta})$ the asymptotic power structure (APS) of $\{\phi_n\}$ to scale $n^{-\frac{1}{2}}$. Thus, the APS of $\{\phi_{\hat{n}}\}$ to scale $n^{-\frac{1}{2}}$ is $\beta(\rho^{-\frac{1}{2}}\mathbf{\theta})$.

In this context the Pitman efficiency for the alternative sequence $\{\mathbf{\delta}n^{-\frac{1}{2}}\}$ becomes the limit of $\hat{n}/n$ for those sequences $\{\hat{n}\}$ for which $\{\phi_{\hat{n}}\}$ and $\{\phi_n^*\}$ have equal APS at δ.

Then it may readily be derived from [9] that $e(\hat{\mathbf{\theta}}_n, \hat{\mathbf{\theta}}_n^*) = e(\phi_n, \phi_n^*)$ is the limit of $\hat{n}/n$ for those sequences $\{\hat{n}\}$ for which $\{\phi_{\hat{n}}\}$ and $\{\phi_n^*\}$ have APS with equal Gaussian curvature at $\mathbf{0}$.

Upon requiring that all partial derivatives up to the second order in components of $\mathbf{\theta}$ of $\beta_n(n^{-\frac{1}{2}}\mathbf{\theta})$ evaluated at $\mathbf{0}$ also converge to the corresponding partial derivatives of $\beta(\mathbf{\theta})$ evaluated at $\mathbf{0}$ and imposing similar conditions on β_n^* and β^* we obtain that this characteristic is indeed asymptotic, namely, that $\hat{n}/n$ is the limit of the ratios of sample sizes required to give $\beta_n(\mathbf{\theta})$ and $\beta_n^*(\mathbf{\theta})$ equal Gaussian curvature at $\mathbf{0}$.

5. Existence of $\mathcal{W}_n^2$, $\hat{\mathcal{W}}_n^2$, $\mathfrak{M}_n^2$, $\hat{\mathfrak{M}}_n^2$ and general efficiency properties. To establish existence of these test statistics we need only find consistent estimates of $\|\beta_{ij}^{(k)}\|^{-1}$, $\|\pi_{ij}^{(k)}\|^{-1}$, $k = 1, 2, 3$.

The following lemma from the theory of matrices which we state without proof shows that, in general, it suffices that $\|\beta_{ij}\|$ (and hence $\|\pi_{ij}\|$) be nonsingular and that consistent estimates $\|\hat{\beta}_{ij}\|$ exist.

LEMMA 5.1. *If $\|a_{ij}\|$ is any nonsingular matrix and $\epsilon > 0$, then there exists a constant $\delta(\|a_{ij}\|)$ such that if $|b_{ij} - a_{ij}| < \delta(\|a_{ij}\|)$ for every i and j, then $\|b_{ij}\|$ is invertible, and if $\|b_{ij}\|^{-1} = \|b^{ij}\|$, $\|a_{ij}\|^{-1} = \|a^{ij}\|$ then $|b^{ij} - a^{ij}| < \epsilon$ for every i and j.*

For if we now define $\|\hat{\beta}^{ij}\| = \|\hat{\beta}_{ij}\|^{-1}$ if $\|\hat{\beta}_{ij}\|$ is nonsingular and equal to the identity otherwise (or to a close nonsingular matrix), then $P_{\mathbf{0}}[|\hat{\beta}^{ij} - \beta^{ij}| \geqq \epsilon] \leqq P_{\mathbf{0}}[|\hat{\beta}_{ij} - \beta_{ij}| \geqq \delta(\|\beta_{ij}\|)$ for some i and $j]$ by Lemma 5.1 and the last term converges to 0 by hypothesis.

We remark that natural estimates of $\|\pi_{ij}^{(i)}\|$, $i = 1, 2, 3$, exist, namely,

$$\hat{\pi}_{ij}^{(1)} = \hat{\sigma}_{ij},$$

$$(5.1)\quad \pi_{ij}^{(2)} = [n^3/n(n-1)(n-2)]\int_{-\infty}^{\infty} \hat{F}_i(x)\hat{F}_j(y)\, d\hat{F}_{i,j}(x, y) - \tfrac{1}{4}, \quad i \neq j,$$

$$\pi_{ij}^{(3)} = \hat{F}_{i,j}(0, 0) - \tfrac{1}{4}, \qquad i \neq j,$$

where $\hat{F}_i$, $\hat{F}_j$, $\hat{F}_{i,j}$ are the sample cumulative distribution functions of X_{ik}, X_{jk}, and (X_{ik}, X_{jk}), respectively, $1 \leqq k \leqq n$. If $i = j$ the last two estimates trivially reduce to $\frac{1}{12}$ and $\frac{1}{4}$ respectively. These estimates may be shown to be U.M.V unbiased for the family of all distributions F. The second has an alternative form $1/n(n-1)(n-2)\sum_{r\neq s\neq t} I^+(X_{ir} - X_{is}, X_{jr} - X_{jt})$ where $I^+(x, y) = 1$ if $x > 0$ and $y > 0$, and 0 otherwise. The consistency of all three follows from the consistency of U-statistics (Lehmann [12]) (in the first and third case the law of large numbers). Hence, natural choices of $\hat{\mathfrak{M}}_n$ and $\hat{\mathcal{W}}_n$ exist.

In [16] for $p = 2$ Sen and Chatterjee have independently obtained interesting versions of the two-sample analogues of $\hat{\mathfrak{M}}_n$ and $\hat{\mathcal{W}}_n$ (corresponding to the two-sample Wilcoxon and Mood's median test respectively), which may be used in the construction of genuinely nonparametric tests of H. They are at present extending these results to the sign test and the c-sample problem. Essentially it follows from their work that, for instance, given the sign of $X_{1i}Y_{2i}$ for each i, the sign statistic h_3 has a nonparametric distribution under the hypothesis. The conditional covariance matrix is given by

$$\hat{\pi}_{12}^{(3)} = \tfrac{1}{2}\{1 + 2\hat{F}(0, 0) - \hat{F}_1(0) - \hat{F}_2(0)\} - \tfrac{1}{4},$$

which is the U.M.V. unbiased estimate of $\pi_{12}^{(3)}$ for the family of all symmetric F. The $\hat{\mathcal{W}}_n$ which is evolved by this procedure, however, involves a $\hat{\pi}_{12}^{(2)}$ which is not U.M.V. unbiased in either sense. The nonparametric application of this test is quite complicated.

On the other hand, no one simple consistent estimate of either $f_i(0)$ or $\int_{-\infty}^{\infty} f_i^2(x)\,dx$ is known. The reader is referred to Rosenblatt [15] for consistent estimates of the former and Lehmann [13] for the latter. Thus, although $\mathcal{W}_n$, $\mathfrak{M}_n$, $\hat{\mathfrak{M}}_n$, $\hat{\mathcal{W}}_n$, can all be satisfactorily defined asymptotically, the latter seem simpler and more usable. Moreover, for a unified theory of testing and estimation in this problem, it would seem reasonable to wish to have our tests and estimates involve the same statistics, a criterion fulfilled by $\mathcal{W}_n^2$ and $\mathfrak{M}_n^2$ but not $\hat{\mathfrak{M}}_n^2$ and $\hat{\mathcal{W}}_n^2$.

We might also remark that the error in estimating $\|\pi_{ij}^{(2)}\|^{-1}$, $\|\pi_{ij}^{(3)}\|^{-1}$ should be small unless $\|\pi_{ij}^{(k)}\|$, $k = 2, 3$ are practically degenerate (i.e., unless det $\|\pi_{ij}\|$ is very close to 0) since $\pi^{ij} = 1/\det \|\pi_{ij}\|(\|\pi_{ij}\|]_{ij}$, where $[\|a_{ij}\|]_{kl}$ is the cofactor of a_{kl} in $\|a_{ij}\|$.

We now turn to the efficiency of $\mathcal{W}_n^2$ and T_n^2.

By (3.3),

$$e(\mathcal{W}_n^2, T_n^2, \boldsymbol{\delta}) = \boldsymbol{\delta}\|\beta_{ij}^{(2)}\|^{-1}\boldsymbol{\delta}'/\boldsymbol{\delta}\|\sigma^{ij}\|\boldsymbol{\delta}',$$
$$e(\mathfrak{M}_n^2, T_n^2, \boldsymbol{\delta}) = \boldsymbol{\delta}\|\beta_{ij}^{(1)}\|^{-1}\boldsymbol{\delta}'/\boldsymbol{\delta}\|\sigma^{ij}\|\boldsymbol{\delta}'. \tag{5.2}$$

In the event that $\|\beta_{ij}\|$ for $k = 1, 2, 3$ are diagonal matrices, for instance in the totally symmetric case, our expressions simplify considerably and we obtain,

$$e(\mathcal{W}_n^2, T_n^2, \boldsymbol{\delta}) = \frac{12\sum_{i=1}^{p}\left(\int_{-\infty}^{\infty} f_i^2(x)\,dx\right)^2 \delta_i^2}{\sum_{i=1}^{p}\frac{\delta_i^2}{\sigma_i^2}} \tag{5.3}$$

and a similar expression for $e(\mathfrak{M}_n^2, T_n^2, \boldsymbol{\delta})$. Thus, usually only in the case when the components of F are identically distributed as well as totally symmetric do (5.2) and (5.3) become independent of $\boldsymbol{\delta}$.

Nonetheless in the totally symmetric case we do have two results analogous to (4.7) and (4.8) of [2]:

$$\inf_{F\in\mathfrak{F}} \inf_{\boldsymbol{\delta}} e(\mathcal{W}_n^2, T_n^2, \boldsymbol{\delta}) = .86, \tag{5.4}$$

$$\inf_{F\in\mathfrak{F}^*} \inf_{\boldsymbol{\delta}} e(\mathfrak{M}_n^2, T_n^2, \boldsymbol{\delta}) = .33, \tag{5.5}$$

where $\mathfrak{F}$ is the family of all totally symmetric p-variate distributions whose marginal densities exist, and $\mathfrak{F}^*$ is the family of all totally symmetric p-variate distributions with unimodal densities continuous at 0.

These follow from the previously quoted theorem of Courant and the Hodges-Lehmann univariate results [7]. Behavior in this situation is thus good.

We also have, for $p \geqq 3$,

$$\inf_{\Phi\varepsilon\varphi} \inf_{\boldsymbol{\delta}} e(\mathcal{W}_n^2, T_n^2, \boldsymbol{\delta}) = 0, \tag{5.6}$$

and for $p \geqq 2$,

$$\inf_{\Phi\varepsilon\varphi} \inf_{\boldsymbol{\delta}} e(\mathfrak{M}_n^2, T_n^2, \boldsymbol{\delta}) = 0, \tag{5.7}$$

where φ is the family of all p-variate nonsingular normal distributions.

These relations follow upon applying Theorem 3.1 to the family of distributions defined in Proposition 4.1 of [2] for (5.6) and to the bivariate normal family for (5.7). Thus the pathologies of [2] carry over, including

$$\inf_{F\varepsilon\mathcal{Y}} \inf_{\boldsymbol{\delta}} e(\mathcal{W}_n^2, T_n^2, \boldsymbol{\delta}) = 0, \tag{5.8}$$

where $\mathcal{Y}$ is the family of all nonsingular bivariate distributions. This too follows upon applying Theorem 3.1 to the results of [2]. However, the desirable properties also carry over.

Define

$$F_{\epsilon,\tau}(x_1, \cdots, x_p) = (1 - \epsilon)\Lambda(x_1, \cdots, x_p + \epsilon\psi(x_1/\tau_1, \cdots, x_p/\tau_p),$$

where $\boldsymbol{\tau} = (\tau_1, \cdots, \tau_p)$, and Λ and ψ are symmetric nonsingular distributions. Call $e(\mathcal{W}_n^2, T_n^2, \boldsymbol{\delta})$, $e_1(\boldsymbol{\varepsilon}, \boldsymbol{\tau})$ when $F_{\epsilon,\tau}$ is the underlying distribution under the null hypothesis, and define $e_2(\boldsymbol{\varepsilon}, \boldsymbol{\tau})$ similarly for $\mathfrak{M}_n^2$. Then, we have, for $\epsilon > 0$,

$$\inf_{\boldsymbol{\delta}\neq\mathbf{0}} e_i(\epsilon, \boldsymbol{\tau}) \to \infty \tag{5.9}$$

as any coordinate of $\boldsymbol{\tau}$ tends to ∞ for $i = 1, 2$. The proof is similar to that of Theorem 6.1 in [2]. Thus, these tests are much better than Hotelling's T^2 in cases where heavy tails are to be expected. Of course, this is also true whenever $\mathbf{W}_n$, $\mathbf{M}_n$ are close to a degenerate distribution but $\dot{\mathbf{X}}_n$ is not, for instance, if the distributions are close to ones satisfying (3.6) but not (3.5) of [2].

Let us now consider the special case in which the distribution of (X_{i1}, X_{j1}) is independent of i and j. In this case,

$$\begin{aligned} \beta_{ij}^{(k)} &= \rho_k\beta_{11}^{(k)}, & i \neq j, \\ &= \beta_{11}^{(k)}, & i = j, \end{aligned} \tag{5.10}$$

for $k = 1, 2, 3$. ρ_k is defined by the above relation. It then follows from the theorem of Courant and the discussion for Theorem 3.2 that $\inf_{\boldsymbol{\delta}} e(\mathcal{W}_n^2, T_n^2, \boldsymbol{\delta})$ and $\inf_{\boldsymbol{\delta}} e(\mathfrak{M}_n^2, T_n^2, \boldsymbol{\delta})$ are given by the smallest λ satisfying

$$\det \|\beta_{ij}^{(1)} - \lambda\beta_{ij}^{(2)}\| = 0, \tag{5.11}$$

and hence in this case,

$$\inf_{\boldsymbol{\delta}\neq 0} e(\mathcal{W}_n^2, T_n^2, \boldsymbol{\delta}) \tag{5.12}$$
$$= 12\sigma_{11}\left(\int_{-\infty}^{\infty} f_1^2(x)\,dx\right)^2 \min\left(\frac{1-\rho_1}{1-\rho_2}, \frac{1+(p-1)\rho_1}{1+(p-1)\rho_2}\right)$$

and similarly for $\sup_{\boldsymbol{\delta}\neq 0}$, and $\mathfrak{M}_n^2$.

Therefore, $e(\mathcal{W}_n^2, T_n^2, \boldsymbol{\delta}) \geqq 12\sigma_{11}[\int_{-\infty}^{\infty} f_1^2(x)\,dx]^2 \min(1-\rho_1, 1/p)$. The bound is not sharp. A similar statement holds for $e(\mathfrak{M}_n^2, T_n^2, \boldsymbol{\delta})$.

A particular case of the above is $X_{i1} = Z_{i+1} - Z_1$ where the Z_i's are independent, identically distributed random variables for $\boldsymbol{\theta} = 0$, for $i = 1, \cdots, p$.

This admittedly very artificial situation might arise for instance in a testing for trend situation where the first reading of an instrument is unavailable and subsequent readings are possible only as differences from the first.

In this situation we have,

$$3f_1^2(0)\sigma_{11} \leqq e(\mathfrak{M}_n^2, T_n^2, \boldsymbol{\delta}) \leqq 4.5 f_1^2(0)\sigma_{11}, \tag{5.13}$$

$$6\sigma_{11}[\textstyle\int_{-\infty}^{\infty} f_1^2(x)\,dx]^2 < e(\mathcal{W}_n^2, T_n^2, \boldsymbol{\delta}) \leqq 12\sigma_{11}[\int_{-\infty}^{\infty} f_1^2(x)\,dx]^2. \tag{5.14}$$

The last statement is a consequence of Lehmann [14] (Theorem 2). Typical lower bound values are .925 for Z_1 normal, .88 for Z_1 rectangular.

Clearly if Z_1 is available this test is foolish, but so is Hotelling's T^2. Lehmann in [12] has obtained a nonparametric analogue to the usual analysis of variance test in this situation, whose efficiency to it is somewhat better than the usual Wilcoxon univariate efficiency.

In a subsequent paper we intend to explore models related to the last example more fully.

We remark that all the preceding, by Theorem 3.1, also applies to $\hat{\mathcal{W}}_n^2$ and $\hat{\mathfrak{M}}_n^2$.

As in [2] we continue with the case $p = 2$.

6. Efficiency in the Normal case ($p = 2$). In the case $p = 2$, the Pitman efficiency of $\mathcal{W}_n^2$ with respect to T_n^2, for the sequence $\boldsymbol{\theta}_n = \boldsymbol{\delta} n^{-\frac{1}{2}}$ and distribution $F(x, y)$, which is the same as the efficiency of $\hat{\mathcal{W}}_n^2$ with respect to T_n^2, will be denoted by $e_1(\boldsymbol{\delta}, F)$ and is given by the following specialization of (5.2):

$$e_1(\boldsymbol{\delta}, F) = [Q_2(\delta_1/\sigma_1^*, \delta_2/\sigma_2^*)]/[Q_1(\delta_1/\sigma_1, \delta_2/\sigma_2)], \tag{6.1}$$

where $\sigma_i = (\sigma_{ii})^{\frac{1}{2}}$, $\sigma_i^* = [\beta_{ii}^{(2)}]^{\frac{1}{2}}$,

$$Q_1(x, y) = (1-\rho_1^2)^{-1}[x^2 + y^2 - 2\rho_1 xy]$$

and

$$Q_2(x, y) = (1-\rho_2^2)^{-1}[x^2 + y^2 - 2\rho_2 xy]$$

where ρ_k is given in (5.10).

Similarly, the Pitman efficiency of $\mathfrak{M}_n^2$ with respect to T_n^2, will be denoted by $e_2(\boldsymbol{\delta}, F)$ and is given by

$$e_2(\delta, F) = [Q_3(x/\sigma_1^{**}, y/\sigma_2^{**})]/[Q_1(\delta_1/\sigma_1, \delta_2/\sigma_2)], \tag{6.2}$$

where $\sigma_i^{**} = (\beta_{ii}^{(3)})^{\frac{1}{2}}$ and, $Q_3(x, y) = (1 - \rho_3^2)^{-1}[x^2 + y^2 - 2\rho_3 xy]$.

Let $e_1^M(F)$ equal $\max_{\boldsymbol{\delta}} e_1(\boldsymbol{\delta}, F)$, $e_1^m(F) = \min_{\boldsymbol{\delta}} e_1(\boldsymbol{\delta}, F)$ and similarly define $e_2^M(F)$ and $e_2^m(F)$.

The bivariate normal situation is described in the following two theorems analogous to Theorems 5.1 and 5.2 of [2].

THEOREM 6.1. *If* $F = \boldsymbol{\Phi}(\mathbf{0}, \sigma_1, \sigma_2, \rho)$ *then,*

$$e_2(\boldsymbol{\delta}, F) = \frac{(1 - \rho^2)}{2\cos^{-1}\rho(1 - (\cos^{-1}\rho)/\pi)} \cdot \left[\frac{t_1^2 + t_2^2 - 2(1 - (2/\pi)\cos^{-1}\rho)t_1 t_2}{t_1^2 + t_2^2 - 2\rho t_1 t_2}\right], \tag{6.3}$$

where $t_i = \delta_i/\sigma_i$, $i = 1, 2$, *and*

$$\begin{aligned} e_2^M(F) &= (1 + \rho)/(\pi - \cos^{-1}\rho), & 0 \leqq \rho < 1, \\ &= (1 - \rho)/(\cos^{-1}\rho), & -1 < \rho \leqq 0, \end{aligned} \tag{6.4}$$

$$\begin{aligned} e_2^m(F) &= (1 - \rho)/(\cos^{-1}\rho), & 0 \leqq \rho < 1, \\ &= (1 + \rho)/(\pi - \cos^{-1}\rho), & -1 < \rho \leqq 0. \end{aligned} \tag{6.5}$$

The following two relations also hold:

(1) $2/\pi \leqq e_2^M(F) \leqq .65 < 1$, where the minimum is reached for $\rho = 0$, and approached as $|\rho| \to 1$, and the maximum of .65 is reached for $\rho = .7$. e_2^M is symmetric about $\rho = 0$, and concave, as a function of ρ, between 0 and 1 and 0 and -1.

(2) $e_2^m(F)$, considered as a function of ρ is symmetric about $\rho = 0$ and monotone decreasing to its minimum of 0 as ρ ranges from 0 to 1.

PROOF. (6.3) follows upon employing Sheppard's relation as in Theorem 5.1 of [2]. (6.4) and (6.5) follow either from the classical theorem of Courant, or directly by partial differentiation with respect to t_1 and t_2. The extreme points are found to be given by $t_1 = \pm t_2$. Upon substituting $p = (\cos^{-1}\rho)/\pi$ in (5.3) we can differentiate between the extremes by using the inequalities

$$\begin{aligned} \cos p\pi &\geqq 1 - 2p, & 0 < p \leqq \tfrac{1}{2} \\ &\leqq 1 - 2p, & \tfrac{1}{2} \leqq p < 1. \end{aligned} \tag{6.6}$$

Upon making the above substitution we obtain the analytically simpler form

$$\begin{aligned} e_2^M(p) &= (1 + \cos p\pi)/(\pi(1 - p)), & 0 < p \leqq \tfrac{1}{2} \\ &= (1 - \cos p\pi)/\pi p, & \tfrac{1}{2} \leqq p < 1, \end{aligned} \tag{6.7}$$

and a similar expression for $e_2^m(p)$.

We now give the proof of statement (1). (2) may be proved similarly. The first part of (1) is a consequence of the second, where the value and location of the maximum are obtained by the numerical solution of the equations $[e_2^M(p)]' = 0$. Hence we need prove only the second statement. Since,

$$(6.8)\quad [e_2^M(p)]' = (-\pi^2(1-p)\sin p\pi + \pi(1+\cos p\pi))/\pi^2(1-p)^2$$

for $0 < p \leqq \frac{1}{2}$, and $[e_2^M(0)]' > 0$, and $e_2^M(\frac{1}{2}) = \lim_{p\to 0} e_2^M(p) = 2/\pi$, it follows that it suffices to show that

$$(6.9)\qquad v(p) = -\pi(1-p) + \pi\csc p\pi + \pi\cot p\pi$$

changes sign exactly once for $0 < p \leqq \frac{1}{2}$. But we readily find $v''(p) < 0$ in this range. Hence v is concave and changes sign at most twice. Our result follows since $v(0) = +\infty$, $v(\frac{1}{2}) = 0$. We remark that Theorem 6.1 implies (5.7).

A similar result holds for $e_1(\boldsymbol{\delta}, F)$.

THEOREM 6.2.

$$(6.10)\quad e_1(\boldsymbol{\delta}, F) = \frac{3}{\pi}\frac{(1-\rho^2)}{1-9(1-(2/\pi)\cos^{-1}\rho/2)^2} \cdot \left[\frac{t_1^2 + t_2^2 - 6(1-(2/\pi)\cos^{-1}\rho)t_1 t_2}{t_1^2 + t_2^2 - 2\rho t_1 t_2}\right],$$

where $t_i = \delta_i/\sigma_i$, $i = 1, 2$.

$$(6.11)\quad \begin{aligned} e_1^M(F) &= \frac{3}{\pi}\frac{1+\rho}{2(2-(3/\pi)\cos^{-1}(\rho/2))}, & 0 \leqq \rho < 1\\ &= \frac{3}{\pi}\frac{1-\rho}{2((3/\pi)\cos^{-1}(\frac{1}{2}\rho - 1))}, \end{aligned}$$

$$(6.12)\quad \begin{aligned} e_1^m(F) &= \frac{3}{\pi}\frac{1-\rho}{2((3/\pi)\cos^{-1}(\frac{1}{2}\rho - 1))}, & 0 \leqq \rho < 1\\ &= \frac{3}{\pi}\frac{1+\rho}{2(2-(3/\pi)\cos^{-1}(\rho/2))}, & -1 < \rho \leqq 0. \end{aligned}$$

The following two relations also hold.

(1) $3/\pi \leqq e_1^M(F) \leqq .96 < 1$. As a function of ρ, $e_1^M(F)$ is symmetric about $\rho = 0$, concave for $0 \leqq \rho < 1$ and $-1 < \rho \leqq 0$, attains its minimum of $3/\pi$ at $\rho = 0$, approaches it as $|\rho|$ tends to 1, and attains its maximum of .96 at $\rho = .78$.

(2) $e_1^m(F)$ considered as a function of ρ is also symmetric about 0, and decreases monotonely from a maximum of $3/\pi$ at $\rho = 0$, to an infimum of $\sin \pi/3$ (.87) as ρ tends to 1.

PROOF. The proof is analogous to that of Theorem 6.1 and Theorem 7.2 of [2], upon observing the inequalities

$$(6.13)\quad \begin{aligned} 2\cos p\pi &\geqq 3(1-2p), & \tfrac{1}{3} < p \leqq \tfrac{1}{2},\\ 2\cos p\pi &\leqq 3(1-2p), & \tfrac{1}{2} \leqq p < \tfrac{2}{3}, \end{aligned}$$

and making the substitution $p = \cos^{-1}\rho/2/\pi$ in (6.7) and (6.8).

We remark that in both cases, e^M and e^m are independent of σ_1 and σ_2 but the values of ρ for which they are assumed depend on σ_1 and σ_2 only and are in fact given by $\delta_1 = \pm(\sigma_1/\sigma_2)\delta_2$.

7. Conclusions. Our general conclusions in the testing problem for $\mathcal{W}_n^2$ and $\mathfrak{M}_n^2$, $\hat{\mathcal{W}}_n^2$ and $\hat{\mathfrak{M}}_n^2$ as compared to T_n^2 are similar to those we reached in [2] for $\mathbf{W}_n$ and $\mathbf{M}_n$ as compared to $\bar{\mathbf{X}}_n$. They seem to be definitely better in the presence of gross errors than T_n^2, but they should be used with caution in a situation in which considerable degeneracy is present. Section 4 also points up the need for a satisfactory criterion of what is meant by better, independent of $\boldsymbol{\delta}$. If D-optimality, an attractive criterion in many ways (Kiefer [11]) is used, Theorem 6 of Hodges and Lehmann [7] giving the equality of the asymptotic efficiency for tests and their derived univariate estimates generalizes to give the equality of the efficiency of the estimates and their derived type I and II tests.

Comparison of these tests with the other nonparametric tests available in this context, Hodges' bivariate sign test [6] and Blumen's sign test [3], does not seem readily feasible since the asymptotic distribution of these tests for alternatives $n^{-\frac{1}{2}}\boldsymbol{\theta}$ are not known. The author has been able to show that the asymptotic distribution of the Hodges test does exist in this sense, but we can characterize it only in terms of the distribution of the maximum of certain Gaussian processes with nonlinear trend. This would suggest that the Pitman efficiency of the former with respect to T_n^2 and hence $\mathcal{W}_n^2$ and $\mathfrak{M}_n^2$ does not exist. The Bahadur efficiency of this test for normal F has also been computed by Klotz and Joffe [10] and shown to be independent of ρ and to tend to $2/\pi$ as $\boldsymbol{\theta} \to \mathbf{0}$. This would indicate that although for $|\rho|$ close to 1, $\mathfrak{M}_n^2$ may be worse than the Hodges test, $\mathcal{W}_n^2$ is always better. More precise information on these points would be of interest.

Acknowledgment. I should like to thank Professor Erich Lehmann for his great help and encouragement during the progress of the thesis, a portion of which formed the bulk of this work. I am also indebted to Doctors A. Maitra and S. W. Dharmadhikari for helpful discussions.

REFERENCES

[1] Bickel, P. J. (1963). Asymptotically nonparametric statistical inference. Unpublished dissertation, University of California, Berkeley.

[2] Bickel, P. J. (1964). On some alternative estimates for shift in the p-variate one sample problem. *Ann. Math. Statist.* **35** 1079–1090.

[3] Blumen, I. (1958). A new bivariate sign test. *J. Amer. Statist. Assoc.* **53** 448–456.

[4] Bodewig, E. (1956). *Matrix Calculus.* Interscience.

[5] Hannan, E. J. (1956). The asymptotic power of tests based upon multiple correlation. *J. Roy. Statist. Soc. Ser. B.* **18** 227–233.

[6] Hodges, J. L., Jr. (1955). A bivariate sign test. *Ann. Math. Statist.* **26** 523–527.

[7] Hodges, J. L., Jr. and Lehmann, E. L. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* **34** 598–611.

[8] Huber, P. J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.* **35** 73–101.

[9] Isaacson, S. L. (1951). On the theory of unbiased tests of simple hypotheses specifying the values of two or more parameters. *Ann. Math. Statist.* **22** 217–234.

[10] Joffe, A. and Klotz, J. (1962). Null distribution and Bahadur efficiency of the Hodges bivariate sign test. *Ann. Math. Statist.* **33** 803–807.

[11] Ki fer, J. (1959). Optimum experimental designs. *J. Roy. Statist. Soc. Ser. B.* **21** 272–319.

[12] Lehmann, E. L. (1963). Robust estimation in analysis of variance. *Ann. Math. Statist.* **34** 957–966.

[13] Lehmann, E. L. (1963). Nonparametric confidence intervals for a shift parameter. *Ann. Math. Statist.* **34** 1507–1512.

[14] Lehmann, E. L. (1964). Asymptotically nonparametric inference in linear models with one observation per cell. *Ann. Math. Statist.* **35** 726–734.

[15] Rosenblatt, M. (1956). Remarks on nonparametric estimates of a density function. *Ann. Math. Statist.* **27** 832–837.

[16] (a) Sen, P. K. and Chatterjee, S. K. (1964). Nonparametric tests for the bivariate two-sample location problem. *Calcutta Statist. Assn. Bulletin.* 13 Nos. 49–50. (b) Sen, P. K. Personal communication.