OPTIMAL INVARIANT RANK TESTS FOR THE k-SAMPLE PROBLEM[1]

By T. K. Matthes and D. R. Truax

University of Oregon

0. Introduction. Suppose that we have k random samples of size n from populations with distribution functions $F_1, F_2, \cdots, F_k$ all belonging to the same class Ω of distribution functions. That is, $X_1, X_2, \cdots, X_{kn}$ are independent random variables, and X_i has distribution function F_α if $(\alpha - 1)n < i \leqq \alpha n$. The k-sample problem is to test the hypothesis that $F_1 = F_2 = \cdots = F_k$. The hypothesis, as well as the alternative that the distribution functions are not identical, remains invariant under relabelling of the distribution functions, and it is natural to ask that a test also be invariant under a relabelling of the samples. In this paper we will consider only tests which are invariant under all permutations of the k samples.

In Section 1, one-parameter families of distributions are introduced, and the locally most powerful invariant rank tests are found. These tests are all based on a statistic of the form

$$\sum_{i=1}^{kn} \sum_{j=1}^{kn} (W_{ij} - \bar{W})a_{ij}, \tag{0.1}$$

where the a_{ij} are constants depending on the one-parameter family, and W_{ij} is a random variable which is one if the ith and the jth order statistic from the combined sample both come from the same sample, and $W_{ij} = 0$ otherwise. We take $W_{ii} = 1$ and $\bar{W} = 1/k$.

In Section 2 the limiting distribution of this statistic is found under the null hypothesis, and under a sequence of alternatives in Section 3. In Section 4 the locally best invariant rank test statistic is shown to be asymptotically equivalent to a quadratic form in certain statistics which arise in the two-sample problem. In Section 5 it is proved for many families of alternatives which include those of the translation type, that the locally best invariant rank test is asymptotically equivalent to the test which maximizes, among all tests, the average power over spheres in the parameter space. This optimality property is an asymptotic analogue of the well-known F test and seems not to have been previously discussed. In Section 6 the general results are applied in a special case showing that the Kruskal-Wallis test possesses the cited optimality properties.

The point of departure in the present paper is the investigation of invariant rank tests. Previously, locally most powerful rank tests against one-sided parametric alternatives $\theta > 0$ have been obtained. For the two-sample problem such tests reject the hypothesis $\theta = 0$ when $L = \sum b_i Z_i >$ constant, where $\{Z_i\}$ is the rank order vector, $Z_i = 1$, or 0, according as the ith ordered observation is, or is not, from the first sample, and where the b_i's are constants depending on the

Received 29 July 1964; revised 4 February 1965.

[1] This research was supported by the National Science Foundation under grant GP-1643.

one-parameter family. Dwass [6] found the limiting distribution of certain approximating L statistics of polynomial type under a sequence of alternatives, and conjectured the limiting distribution of the actual locally most powerful rank statistic. Chernoff and Savage [4] derive asymptotic normality of L statistics under alternatives as well as under the hypothesis, although their conditions appear to be difficult to check except in special cases. In 1961 Hájek [7] obtains asymptotic normality of L statistics under very weak conditions when the hypothesis is true. In 1962 Hájek [8] uses elegant methods, based on the concept of contiguity for probability distributions developed by LeCam, in order to find the asymptotic distribution under local alternatives. Regarding the k-sample problem, Andrews [1] investigated the asymptotic power of the Kruskal-Wallis test and a few other non-parametric tests under local parametric alternatives, and discussed asymptotic relative efficiency. More recently, Puri [10] has adopted the approach of Chernoff and Savage in order to obtain the limiting distribution of quadratic forms in L statistics involving k samples. For other papers in similar areas, the reader is referred to [2], [3], [9], and [11].

1. Locally most powerful invariant rank tests. We suppose that X_i is distributed according to a probability density $f_\alpha(x;\theta)$, $(\alpha - 1)n < i \leqq \alpha n$, where θ is a real parameter whose range contains an open interval about the origin, and when $\theta = 0$ all the densities are equal to $f_0(x)$. That is,

$$f_1(x;0) = f_2(x;0) = \cdots = f_k(x;0) = f_0(x).$$

In addition, it will be assumed that

$$h_\alpha(x) = (\partial/\partial\theta) \log f_\alpha(x;\theta)|_{\theta=0}$$

exists for $\alpha = 1, 2, \cdots, k$ and that

$$\int h_\alpha^2(x) f_0(x)\,dx < \infty, \qquad \alpha = 1, 2, \cdots, k. \tag{1.1}$$

A further regularity condition will be needed for the derivation of locally best rank tests. We therefore will assume

$$\lim_{\theta\to 0} E_0\{|[f_\alpha(X;\theta) - f_0(X)]/\theta f_0(X) - h_\alpha(X)|\} = 0 \tag{1.2}$$

for $\alpha = 1, 2, \cdots, k$. Hereafter E_0 always denotes expectation when $\theta = 0$. Thus we have the expansion

$$f_\alpha(x;\theta)/f_0(x) = 1 + \theta h_\alpha(x) + R_\alpha(x;\theta), \tag{1.3}$$
$$\lim_{\theta\to 0} E_0\{|R_\alpha(X;\theta)/\theta|\} = 0.$$

Let $N = kn$ denote the total sample size, and let $X_{(1)}, X_{(2)}, \cdots, X_{(N)}$ denote the order statistics of the combined sample. The main result of this section is the following:

THEOREM 1. *Assuming conditions* (1.1) *and* (1.2), *the locally most powerful invariant rank test rejects the hypothesis* $\theta = 0$ *when*

$$S_n = (1/n) \textstyle\sum_{i,j=1}^{N} (W_{ij} - \bar{W}) E_0\varphi(X_{(i)}, X_{(j)}) > \textit{constant}, \tag{1.4}$$

where

$$\varphi(u, v) = \textstyle\sum_{\alpha=1}^{k} [h_\alpha(u) - \bar{h}(u)][h_\alpha(v) - \bar{h}(v)], \tag{1.5}$$

and

$$\bar{h}(u) = (1/k) \textstyle\sum_{\alpha=1}^{k} h_\alpha(u), \bar{W} = 1/k.$$

PROOF. Let $Z_{i\alpha}$ be 1 or 0 according as the ith ordered observation in the combined sample is, or is not, from the αth sample. An equivalent random vector is, of course, the vector of ranks. It is easily seen that the probability distribution of $Z = (Z_{i\alpha})$ is given by

$$p(z; \theta) = [(n!)^k/N!]E_0[\textstyle\prod_{i=1}^{N} \sum_{\alpha=1}^{k} z_{i\alpha} f_\alpha(X_{(i)} ; \theta)/f_0(X_{(i)})]. \tag{1.6}$$

Inserting (1.3) into (1.6) and expanding the product easily leads to

$$p(z; \theta) = [(n!)^k/N!]\{1 + A_1(z)\theta + A_2(z)\theta^2 + A_3(z; \theta) + o(\theta^2)\} \tag{1.7}$$

where

$$\begin{aligned} A_1(z) &= \textstyle\sum_{i=1}^{N} \sum_{\alpha=1}^{k} z_{i\alpha} E_0[h_\alpha(X_{(i)})], \\ A_2(z) &= \textstyle\sum_{i\neq j} E_0[\sum_{\alpha=1}^{k} z_{i\alpha} h_\alpha(X_{(i)})][\sum_{\beta=1}^{k} z_{j\beta} h_\beta(X_{(j)})], \\ A_3(z; \theta) &= \textstyle\sum_{i=1}^{N} \sum_{\alpha=1}^{k} z_{i\alpha} E_0[R_\alpha(X_{(i)} ; \theta)]. \end{aligned} \tag{1.8}$$

The concept of invariance under relabelling of samples corresponds formally to groups of transformations of the sample space and the space of distributions onto themselves. In the present case let π be a permutation of $\{1, 2, \cdots, k\}$. We are interested in the transformation of Z to Z_π given by $(Z_\pi)_{i\alpha} = Z_{i\pi\alpha}$. If ϕ is any invariant test we then have $\phi(z) = \phi(z_\pi)$ for all permutations π. The power function of any invariant rank test is given by

$$\beta_\phi(\theta) = \textstyle\sum_z \phi(z)p(z; \theta) = \sum_z \phi(z)(1/k!) \sum_\pi p(z_\pi ; \theta). \tag{1.9}$$

In order to evaluate the power in a neighborhood of $\theta = 0$, observe that

$$\textstyle\sum_\pi z_{i\pi\alpha} = (k - 1)!, \; i = 1, 2, \cdots, N; \; \alpha = 1, 2, \cdots, k, \tag{1.10}$$

and

$$\textstyle\sum_i E_0[h_\alpha(X_{(i)})] = E_0[\sum_i h_\alpha(X_i)] = 0 \tag{1.11}$$

by (1.2). Also

$$\textstyle\sum_i E_0[R_\alpha(X_{(i)} ; \theta)] = E_0[\sum_i R_\alpha(X_i ; \theta)] = 0 \tag{1.12}$$

by (1.3) and (1.11). (Here, and henceforth, we will suppress the ranges of summation, consistently letting i and j range from 1 to N, while α, β, and γ range from 1 to k.) It follows from (1.10) and (1.11) that

$$\textstyle\sum_\pi A_1(z_\pi) = 0, \tag{1.13}$$

and for exactly the same reasons, and (1.12)

$$(1.14) \qquad \sum_\pi A_3(z_\pi\,;\theta) = 0 \text{ for all } \theta.$$

A little computation shows that

$$(1.15) \qquad (1/k!) \sum_\pi z_{i\pi\alpha} z_{j\pi\beta} = W_{ij}/k, \ \alpha = \beta$$
$$= (1 - W_{ij})/k(k-1), \ \alpha \neq \beta.$$

For example, if $X_{(i)}$ and $X_{(j)}$ come from the α_0th sample and $\alpha = \beta$, there are exactly $(k-1)!$ permutations π such that $\pi(\alpha) = \pi(\beta) = \alpha_0$.

By means of (1.15) we obtain

$$(1.16) \quad (1/k!) \sum_\pi A_2(z_\pi) = \sum_{i\neq j} \{(W_{ij}/k) \sum_\alpha E_0[h_\alpha(X_{(i)})h_\alpha(X_{(j)})]$$
$$+ [(1 - W_{ij})/k(k-1)] \sum_{\alpha\neq\beta} E_0[h_\alpha(X_{(i)})h_\beta(X_{(j)})]\}.$$

As in the case of (1.11)

$$(1.17) \qquad \sum_{i\neq j} E_0[h_\alpha(X_{(i)})h_\alpha(X_{(j)})] = 0.$$

Further algebraic manipulation and use of (1.17) simplifies (1.16) to

$$(1.18) \quad (1/k!) \sum_\pi A_2(z_\pi) =$$
$$[1/(k-1)] \sum_{i\neq j} (W_{ij} - \bar{W}) E_0\varphi(X_{(i)}, X_{(j)}) = [n/(k-1)]S_n',$$

where the prime on S_n will indicate that the summation in (1.4) is only for $i \neq j$. Together, (1.13), (1.14) and (1.18) show

$$(1/k!) \sum_\pi p(z_\pi\,;\theta) = [(n!)^k/N!]\{1 + [n/(k-1)]\theta^2 S_n' + o(\theta^2)\}.$$

Therefore, from (1.9)

$$(1.19) \quad \beta_\phi(\theta) = [(n!)^k/N!] \sum_z \phi(z)\{1 + [n/(k-1)]\theta^2 S_n' + o(\theta^2)\}.$$

It follows from the Neyman-Pearson lemma that choosing a test which maximizes the second derivative of the power function at $\theta = 0$ gives a test which is uniformly most powerful for all θ in some open interval about $\theta = 0$. It is clear from (1.19) that the second derivative of the power function is maximized by choosing $\phi(z) = 1$ or 0 according as S_n' is greater or less than some constant. The proof is then completed by noting the constancy of

$$(1.20) \qquad S_n - S_n' = [(k-1)/kn] \sum_i E_0[\varphi(X_{(i)}, X_{(i)})]$$
$$= (k-1)E_0[\varphi(X_1, X_1)] < \infty.$$

2. Limiting distribution under the hypothesis. In this section we show that when $\theta = 0$ the limiting distribution of S_n is that of a positive linear combination of independent chi-square random variables. First we will need a lemma which is a slight generalization of a result of Hájek ([7], Lemma 6.1).

LEMMA 2.1. *Let $X_1, X_2, \cdots,$ be a sequence of independent random variables, each having the same continuous distribuion function, and let R_{Ni}, $1 \leqq i \leqq N$, be* *the rank of* X_i *among* $X_1, X_2, \cdots, X_N$. *Let* $1 \leqq p < \infty$ *and* φ *be a Borel measurable function of* k *variables such that*

$$E\{|\varphi(X_1, X_2, \cdots, X_k)|^p\} < \infty; \tag{2.1}$$

Define

$$Y_N = E\{\varphi(X_1, X_2, \cdots, X_k) \mid R_{N1}, R_{N2}, \cdots, R_{Nk}\}, \quad N \geqq k.$$

Then

$$\lim_{N\to\infty} E\{|Y_N - \varphi(X_1, X_2, \cdots, X_k)|^p\} = 0. \tag{2.2}$$

Proof. Let $\mathfrak{F}_N$ be the sigma-field generated by $R_{N1}, R_{N2}, \cdots, R_{NN}$. Clearly, $\mathfrak{F}_N \subset \mathfrak{F}_{N+1}$ for each N. Let $\mathfrak{F}_\infty$ denote the smallest sigma-field containing $\bigcup_{N=1}^{\infty}\mathfrak{F}_N$. Then (see [5], p. 293), $\{E\{\varphi(X_1, X_2, \cdots, X_k) \mid \mathfrak{F}_N\}, k \leqq N \leqq \infty\}$ is a martingale. Also,

$$E\{\varphi(X_1, X_2, \cdots, X_k) \mid \mathfrak{F}_N\} = E\{\varphi(X_1, X_2, \cdots, X_k) \mid R_{N1}, \cdots, R_{Nk}\}$$

with probability one since the conditional distribution of $X_1, X_2, \cdots, X_k$ given the ranks of $X_1, X_2, \cdots, X_N$ depends only on the ranks of $X_1, X_2, \cdots, X_k$. Thus, $\{Y_N; k \leqq N \leqq \infty\}$ is a martingale, if we define $Y_\infty = E\{\varphi(X_1, X_2, \cdots, X_k) \mid \mathfrak{F}_\infty\}$. Jensen's inequality for conditional expectations yields $E\{|Y_N|^p\} \leqq E\{|\varphi(X_1, X_2, \cdots, X_k)|^p\}$ which is finite by assumption. Doob's Theorem 4.1 (iii) applies to give $\lim_{N\to\infty} Y_N = Y_\infty'$ exists with probability one, and if $p > 1$

$$\lim_{N\to\infty} E\{|Y_N - Y_\infty'|^p\} = 0. \tag{2.3}$$

Also, by Theorem 4.3 of Doob, $Y_\infty' = Y_\infty$ with probability one. In the case $p = 1$, (2.3) follows from the fact that $\{Y_1, Y_2, \cdots, Y_\infty\}$ is a martingale so that the Y_N are uniformly integrable. The proof is completed by showing $\varphi(X_1, X_2, \cdots, X_k) = Y_\infty$ with probability one. That is, by showing $\varphi(X_1, X_2, \cdots, X_k)$ is $\mathfrak{F}_\infty$ measurable. Let F denote the common distribution function of the X_i. Then

$$E\{|F(X_i) - R_{Ni}/(N+1)|^2\} = E\{|F(X_1) - R_{N1}/(N+1)|^2\}$$
$$= \left[\sum_{r=1}^{N} E\{|F(X_{Nr}) - r/(N+1)|^2\}\right]/N$$

where X_{Nr} is the rth smallest X_i among $X_1, X_2, \cdots, X_N$. The right hand side approaches zero as $N \to \infty$, so that $F(X_i)$ is the mean square limit of $\mathfrak{F}_\infty$ measurable functions. It must then follow that $X_1, X_2, \cdots, X_k$ are equal almost everywhere to $\mathfrak{F}_\infty$ measurable functions, and hence $\varphi(X_1, X_2, \cdots, X_k)$ is equal almost everywhere to an $\mathfrak{F}_\infty$ measurable function.

After making a change of variables in the index of summation in (1.4) by letting $i \to R_i$ we can express S_n as

$$S_n = (1/n)\sum_{i,j} (a_{ij} - \bar{a})E_0\{\varphi(X_i, X_j) \mid R_{Ni}, R_{Nj}\}, \tag{2.4}$$

where $a_{ij} = 1$, or 0 according as i and j are indices corresponding to the same sample or not. That is, if we let $I(\alpha) = \{(\alpha - 1)n + 1, \cdots \alpha n\}$, then $a_{ij} = 1$, or 0, according as i and j do or do not both belong to the same set $I(\alpha)$ for some α. Of course, $\bar{a} = 1/k$. Now, define

$$T_n = (1/n)\textstyle\sum_{i,j} (a_{ij} - \bar{a})\varphi(X_i, X_j), \tag{2.5}$$
$$T_n' = (1/n)\textstyle\sum_{i\neq j} (a_{ij} - \bar{a})\varphi(X_i, X_j).$$

LEMMA 2.2. *Under assumption* (1.1), $S_n - T_n$ *approaches zero in probability under the hypothesis as* $n \to \infty$.

PROOF. First, notice that

$$T_n - T_n' = [(k-1)/N]\textstyle\sum_{i=1}^{N} \varphi(X_i, X_i) \tag{2.6}$$

and this converges with probability one to $(k-1)E_0\varphi(X_1, X_1)$. Taking (1.20) into account, it is clear that it is sufficient to prove

$$\lim_{n\to\infty} E_0\{(S_n' - T_n')^2\} = 0. \tag{2.7}$$

It is convenient to put $c_{ij} = (a_{ij} - \bar{a})/n$. For the sake of notational simplicity, the dependence of c_{ij} on n will be suppressed. Note that

$$\textstyle\sum_i c_{ij} = \sum_j c_{ij} = 0. \tag{2.8}$$

Define

$$Y_{ij} = E_0\{\varphi(X_i, X_j) \mid R_{Ni}, R_{Nj}\} - \varphi(X_i, X_j). \tag{2.9}$$

A simple calculation shows

$$E\{(S_n' - T_n')^2\} = \textstyle\sum_{i\neq j} c_{ij}^2 \operatorname{Var}_0(Y_{ij}) + \sum_{i\neq j}\sum_{k\neq l} c_{ij}c_{kl} \operatorname{Cov}_0(Y_{ij}, Y_{kl}). \tag{2.10}$$

By symmetry we see that $\operatorname{Var}_0(Y_{ij})$ is the same for all $i \neq j$, and

$$\operatorname{Cov}_0(Y_{ij}, Y_{kl}) = \Delta_p, \qquad p = 0, 1, 2 \tag{2.11}$$

depends only on the number of equal indices among i, j, k and l. It is easily shown that $\sum_{i\neq j}\sum_{k\neq l}^{(p)} c_{ij}c_{kl} = O(1)$, $n \to \infty$, where the superscript (p) denotes summation over only those indices corresponding to p, and that $\Delta_p \leqq E(Y_{12}^2) \to 0$, $p = 0, 1, 2$ by Lemma 2.1. Thus (2.7) holds and the lemma is established.

THEOREM 2. *Under assumption* (1.1), *the limiting distribution of* S_n, *when* $\theta = 0$, *is the distribution of* $\sum_{\alpha=1}^{k} \lambda_\alpha Z_\alpha$, *where* $Z_1, Z_2, \cdots, Z_k$ *are independent random variables each having the chi-square distribution with* $k-1$ *degrees of freedom, and* $\lambda_1, \lambda_2, \cdots, \lambda_k$ *are the eigenvalues of* $\mathbf{\Sigma} = (\sigma_{\alpha\beta})$,

$$\sigma_{\alpha\beta} = \operatorname{Cov}_0[h_\alpha(X) - \bar{h}(X), h_\beta(X) - \bar{h}(X)]. \tag{2.12}$$

PROOF. By virtue of Lemma 2.2, it suffices to show that T_n has the asserted limit distribution.

We again denote by $I(\beta)$ the βth block of n consecutive integers, $\{(\beta-1)n + 1, \cdots, \beta n\}$. Then define the random variables

$$U_{\alpha\beta}^{(n)} = (1/n^{\frac{1}{2}})\textstyle\sum_{i\in I(\beta)} [h_\alpha(X_i) - \bar{h}(X_i)], \tag{2.13}$$
$$\bar{U}_\alpha^{(n)} = (1/k)\textstyle\sum_{\beta=1}^{k} U_{\alpha\beta}^{(n)}.$$

Introduce the vector notation

$$\mathbf{U}_\beta^{(n)} = (U_{1\beta}^{(n)}, U_{2\beta}^{(n)}, \cdots, U_{k\beta}^{(n)})', \qquad (2.14)$$
$$\bar{\mathbf{U}}^{(n)} = (1/k)\textstyle\sum_{\beta=1}^k \mathbf{U}_\beta^{(n)}.$$

The expression for T_n in (2.5) becomes

$$T_n = \textstyle\sum_{\alpha=1}^k \{[(k-1)/k]\sum_\beta (U_{\alpha\beta}^{(n)})^2 - (1/k)\sum_{\beta\neq\gamma} U_{\alpha\beta}^{(n)}U_{\alpha\gamma}^{(n)}\} \qquad (2.15)$$
$$= \textstyle\sum_\alpha \sum_\beta (U_{\alpha\beta}^{(n)} - \bar{U}_\alpha)^2 = \operatorname{tr}[\sum_\beta (\mathbf{U}_\beta^{(n)} - \bar{\mathbf{U}}^{(n)})(\mathbf{U}_\beta^{(n)} - \bar{\mathbf{U}}^{(n)})'].$$

The multivariate central limit theorem asserts that $\mathbf{U}_\beta^{(n)}$ has a limiting multivariate normal distribution with mean vector $\mathbf{0}$, and covariance matrix $\mathbf{\Sigma}$ given by (2.12). Moreover, $\mathbf{U}_1^{(n)}, \mathbf{U}_2^{(n)}, \cdots, \mathbf{U}_k^{(n)}$ are independent. As a trace of a sample covariance matrix involving normally distributed vectors, the limiting distribution in (2.15) is well known and easily derived. Indeed, choose an orthogonal matrix $\mathbf{P}$ which diagonalizes $\mathbf{\Sigma}$,

$$\mathbf{P\Sigma P}' = \mathbf{\Lambda} \qquad (2.16)$$

and define

$$\mathbf{V}_\beta^{(n)} = \mathbf{P}\mathbf{U}_\beta^{(n)}, \qquad \beta = 1, 2, \cdots, k. \qquad (2.17)$$

The αth component, $V_{\alpha\beta}^{(n)}$, of this vector has a limiting normal distribution with mean zero and variance λ_α under the hypothesis, and all $\{V_{\alpha\beta}^{(n)}\}$ are independent. Expressed in terms of V's, we can now represent

$$T_n = \textstyle\sum_\alpha \lambda_\alpha \sum_\beta[(V_{\alpha\beta}^{(n)} - \bar{V}_\alpha^{(n)})^2/\lambda_\alpha] \qquad (2.18)$$

as the asserted linear combination of independent random variables which, in the limit, are chi-square with $k-1$ degrees of freedom. This completes the proof of Theorem 2.

3. Limiting distribution under local alternatives. In this section we investigate the limiting distribution of S_n as θ tends to zero at a rate of $n^{-\frac{1}{2}}$. Under an additional condition it will be shown that the limit distribution of S_n is a weighted sum of independent non-central chi-square random variables. The methods we use depend crucially on the concept of contiguity for sequences of probability distributions due to LeCam, and developed by LeCam and Hájek in [8]. In particular, we apply LeCam's important Lemma 4.2 of [8] to the problem at hand. The application is somewhat along the lines of Hájek, although the alternatives we consider are more general in one sense than the translation type treated there (and less general in the sense that we do not allow the possibility of a scale parameter). However, for the purposes of Section 5, we need a stronger form of LeCam's lemma, one entailing a uniformity with respect to a parameter.

We recall the definition of contiguity. Let $\{P_n; n \geqq 1\}$ and $\{Q_n; n \geqq 1\}$ be two sequences of probability measures defined on measurable spaces $\{\mathfrak{X}_n, \mathfrak{A}_n; n \geqq 1\}$. The sequence $\{Q_n\}$ is said to be contiguous to $\{P_n\}$ if for any sets $A_n \varepsilon \mathfrak{A}_n$, $\lim_{n\to\infty} P_n(A_n) = 0$ implies $\lim_{n\to\infty} Q_n(A_n) = 0$. One important consequence is the following. Let X_n and Y_n be random variables in $\mathfrak{X}_n$ whose difference approaches zero in P_n probability, i.e., for each $\epsilon > 0$,

$$\lim_{n\to\infty} P_n\{|X_n - Y_n| > \epsilon\} = 0.$$

Then their difference also approaches zero in Q_n probability.

As a preliminary we consider in Lemma 3.1 below the following situation, not directly related to the k sample problem. Suppose $X_1, X_2, \cdots$ is a sequence of independent, identically distributed random variables with density $f(x;\theta)$. Let us designate the distribution of $X_1, X_2, \cdots, X_n$ by P_n when $\theta = 0$, and by Q_n when $\theta = \theta_n = \theta_0/n^{\frac{1}{2}}$. Put $f_0(x) = f(x; 0)$, and as before,

$$h(x) = (\partial/\partial\theta) \log f(x;\theta) \mid_{\theta=0}.$$

Borrowing the terminology in [8], define

$$\begin{aligned} W_n(\theta_0) &= 2\textstyle\sum_{i=1}^n [f^{\frac{1}{2}}(X_i;\theta_n)/f_0^{\frac{1}{2}}(X_i) - 1], \\ L_n(\theta_0) &= \textstyle\sum_{i=1}^n \log[f(X_i;\theta_n)/f_0(X_i)], \\ T_n(\theta_0) &= \theta_n \textstyle\sum_{i=1}^n h(X_i). \end{aligned} \tag{3.1}$$

The reader is cautioned not to confuse $W_n(\theta_0)$ and $T_n(\theta_0)$ with W_{ij} and T_n previously defined. Here, the dependence of these statistics on θ_0 is made explicit since θ_0 will not be fixed, but will vary in some finite interval which we may as well take to be $|\theta_0| \leqq 1$. It should also be kept in mind that θ_n depends on θ_0.

Let $\mathbf{u}(x) = (u_1(x), u_2(x), \cdots, u_k(x))'$ be a vector of functions which satisfy

$$\begin{aligned} &E_0(u_\alpha(X)) = 0, \\ &\mathrm{Var}_0\,(u_\alpha(X)) < \infty, \qquad \alpha = 1, 2, \cdots, k. \end{aligned} \tag{3.2}$$

Denote the covariance matrix of $(h(X); u_1(X), \cdots, u_k(X))$, partitioned in the indicated way, by

$$\begin{pmatrix} \sigma^2 & \boldsymbol{\gamma}' \\ \boldsymbol{\gamma} & \boldsymbol{\Sigma} \end{pmatrix}. \tag{3.3}$$

Of interest is the distribution of the vector

$$\mathbf{U}^{(n)}(\theta_0) = \theta_n \textstyle\sum_{i=1}^n \mathbf{u}(X_i). \tag{3.4}$$

LEMMA 3.1. *Suppose* $\sigma^2 = E_0(h^2(X)) < \infty$ *and*

$$\lim_{\theta\to 0} E_0\{[f^{\frac{1}{2}}(X;\theta) - f_0^{\frac{1}{2}}(X)]/\theta f_0^{\frac{1}{2}}(X) - h(X)/2\}^2 = 0. \tag{3.5}$$

Then,

(i) *for each* θ_0, $\{Q_n(\theta_0)\}$ *is contiguous to* $\{P_n\}$;

(ii) $\mathcal{L}(\mathbf{U}^{(n)}(\theta_0) \mid Q_n) \to N(\theta_0\boldsymbol{\gamma}, \theta_0^2\boldsymbol{\Sigma})$;

(iii) $L_n(\theta_0) - T_n(\theta_0) \to -\theta_0^2\sigma^2/2$, P_n*-uniformly in* θ_0, *i.e., for each* $\epsilon > 0$, $\lim_{n\to\infty} P_n\{|L_n(\theta_0) - T_n(\theta_0) + \theta_0^2\sigma^2/2| > \epsilon\} = 0$ *uniformly in* $|\theta_0| \leqq 1$.

σ^2, $\boldsymbol{\gamma}$, $\boldsymbol{\Sigma}$ *are as defined by* (3.3).

PROOF. In order to apply the aforementioned lemma of LeCam it is necessary to find the limiting distribution of $W_n(\theta_0)$, under P_n. Under assumption (3.5) we evaluate

$$\begin{aligned} 2E_0[f^{\frac{1}{2}}(X;\theta)/f_0^{\frac{1}{2}}(x) - 1] &= 2[\int f_0^{\frac{1}{2}}(x) f^{\frac{1}{2}}(x;\theta)\,dx - 1] \\ &= -E_0\{[f^{\frac{1}{2}}(X;\theta) - f_0^{\frac{1}{2}}(X)]/f_0^{\frac{1}{2}}(X)\}^2 \\ &\approx -\theta^2\sigma^2/4. \end{aligned}$$

Hence,

$$E_0[W_n(\theta_0)] \approx -n\theta_n^2\sigma^2/4 = -\theta_0^2\sigma^2/4. \tag{3.6}$$

Furthermore,

$$\begin{aligned} &\operatorname{Var}_0 [W_n(\theta_0) - T_n(\theta_0)] \\ &\qquad \leqq 4n\theta_n^2 E_0\{[f^{\frac{1}{2}}(X;\theta_n) - f_0^{\frac{1}{2}}(X)]/\theta_n f_0^{\frac{1}{2}}(X) - h(X)/2\}^2 \end{aligned} \tag{3.7}$$

and the right hand side approaches zero uniformly in θ_0 as n tends to infinity, again by (3.5). The limiting distribution of $T_n(\theta_0)$ is, of course, normal with mean zero and variance $\theta_0^2\sigma^2$. From (3.6) and (3.7) we can then conclude that the limiting distribution of $W_n(\theta_0)$, under P_n, is normal with mean $-\theta_0^2\sigma^2/4$ and variance $\theta_0^2\sigma^2$. In addition to this result, LeCam's lemma requires that

$$\lim_{n\to\infty} \max_{1\leqq i\leqq n} P_n[|f(X_i;\theta_n)/f_0(X_i) - 1| > \epsilon] = 0$$

for every $\epsilon > 0$. Since $f(X_i;\theta_n)/f_0(X_i)$, $i = 1, 2, \cdots, n$ are identically distributed, we need only show that $f(X_1;\theta_n)/f_0(X_1) \to 1$ in P_n probability. However, (3.5) clearly shows $f^{\frac{1}{2}}(X_1;\theta_n)/f^{\frac{1}{2}}(X_1) \to 1$ in P_n probability which implies the desired result. The basic conditions of LeCam's lemma are now met. This lemma states, first of all, that for each θ_0, $\{Q_n(\theta_0)\}$ is contiguous to $\{P_n\}$ proving part (i) of our lemma.

LeCam's lemma states, moreover, that under these circumstances

$$W_n(\theta_0) - L_n(\theta_0) \to \theta_0^2\sigma^2/4, \quad \text{in } P_n\text{-probability}. \tag{3.8}$$

In fact, this convergence is P_n-uniform in θ_0, although this stronger fact is not required at the moment. Let $\mathbf{c}$ be an arbitrary k-vector. It follows from the multidimensional central limit theorem that $\mathcal{L}(\mathbf{c}'\mathbf{U}^{(n)}(\theta_0) \mid P_n) \to N(0, b^2)$, where $b^2 = \theta_0^2\mathbf{c}'\mathbf{\Sigma}\mathbf{c}$, and that $L_n(\theta_0)$ and $\mathbf{c}'\mathbf{U}^{(n)}(\theta_0)$ have a limiting bivariate normal distribution under P_n. The correlation coefficient is then seen to be the same as the correlation coefficient between $T_n(\theta_0)$ and $\mathbf{c}'\mathbf{U}^{(n)}(\theta_0)$, namely,

$$\rho = \mathbf{c}'\boldsymbol{\gamma}\theta_0^2/(\sigma^2\theta_0^2 b^2)^{\frac{1}{2}}.$$

In this situation, LeCam's lemma asserts, finally

$$\mathcal{L}(\mathbf{c}'\mathbf{U}^{(n)}(\theta_0) \mid Q_n(\theta_0)) \to N(\rho\sigma b, b^2) = N(\theta_0\mathbf{c}'\boldsymbol{\gamma}, \theta_0^2\mathbf{c}'\mathbf{\Sigma}\mathbf{c}).$$

This proves part (ii) of our lemma in view of the arbitrariness of $\mathbf{c}$.

Turning to assertion (iii), note that $W_n(\theta_0) - T_n(\theta_0) \to 0$, P_n-uniformly in θ_0 as a consequence of (3.7). The proof will be completed by establishing the uniformity assertion made following (3.8). This, however, follows by a closer examination of the proof of LeCam's lemma given in [8]. The only essential modification needed in that proof is the easily shown result that

$$\text{(3.9)} \qquad \lim_{n\to\infty} P_n\{\max_{1\leqq i\leqq n} |f^{\frac{1}{2}}(X_i\,;\theta_n)/f_0^{\frac{1}{2}}(X_i) - 1| > \epsilon\} = 0$$

uniformly in $|\theta_0| \leqq 1$ for each $\epsilon > 0$.

Let us return now to the context of the k-sample problem. With a slight change in notation we now designate the distribution of $\{X_i\,;\, i \varepsilon I(\beta)\}$ by $P_{n\beta}$ when $\theta = 0$, and by $Q_{n\beta}(\theta_0)$ when $\theta = \theta_0/n^{\frac{1}{2}}$. Since we consider only a single value of θ_0 in the remainder of this section, it is unnecessary to always indicate how quantities like $Q_{n\beta}$ depend on θ_0. Finally, denote the distribution of the entire sample by $P_n = \prod_{\beta=1}^k P_{n\beta}$ and $Q_n = \prod_{\beta=1}^k Q_{n\beta}$ under the hypothesis and alternative respectively.

Lemma 3.2. *Let $\{Q_{n\beta}\}$ be contiguous to $\{P_{n\beta}\}$ for $\beta = 1, 2, \cdots, k$. Then $\{Q_n\}$ is contiguous to $\{P_n\}$.*

This lemma can be proved quite generally. For the distributions at hand, however, it suffices to define a statistic for the whole sample like W_n in (3.1) and show that it is asymptotically normal under P_n with mean equal to minus one-fourth its variance. LeCam's lemma applies as above and yields the contiguity. We omit the details.

Proceeding to the limiting distribution of S_n under local alternatives $\theta = \theta_0/n^{\frac{1}{2}}$, put $u_\alpha(x) = h_\alpha(x) - \bar{h}(x)$, $\alpha = 1, 2, \cdots, k$. Lemma 3.1 will be applied to each of the samples where in each case the vector $\mathbf{u}(x)$ has these components. In (2.13) and (2.14) we may write $\mathbf{U}_\beta^{(n)} = n^{-\frac{1}{2}}\sum_{i\varepsilon I(\beta)} \mathbf{u}(X_i)$. The covariance matrix of $\mathbf{u}(X)$ is $\mathbf{\Sigma}$ given in (2.12). In addition, the vector of covariances between $\mathbf{u}(X)$ and $h_\beta(X)$ is denoted by

$$\text{(3.10)} \qquad \gamma_\beta = E_0[h_\beta(X)\mathbf{u}(X)] = \boldsymbol{\sigma}^{(\beta)} + \mathbf{c},$$

where $\boldsymbol{\sigma}^{(\beta)}$ is the βth column of $\mathbf{\Sigma}$ and $c = \mathrm{Cov}_0(\bar{h}(X), \mathbf{u}(X))$. The main theorem pertaining to the asymptotic distribution of S_n under local alternatives (including, of course, $\theta_0 = 0$) follows.

Theorem 3. *Suppose, in addition to the assumptions of Theorem 2, that*

$$\text{(3.11)} \quad \lim_{\theta\to 0} E_0\{[f_\alpha^{\frac{1}{2}}(X;\theta) - f_0^{\frac{1}{2}}(X)]/\theta f_0^{\frac{1}{2}}(X) - h_\alpha(X)/2\}^2 = 0, \qquad \alpha = 1, 2, \cdots, k.$$

Then $S_n - T_n \to 0$ in Q_n-probability, and both have the limiting distribution of

$$\sum_{\alpha=1}^k \lambda_\alpha Z_\alpha,$$

where the λ_α are the eigenvalues of $\mathbf{\Sigma}$, and $Z_1, Z_2, \cdots, Z_k$ are independent non-central chi-square random variables with $k - 1$ degrees of freedom in which $\theta_0^2\lambda_\alpha$ is the non-centrality parameter associated with Z_α.

Proof. Lemma 3.1 (ii) shows that

$$\text{(3.12)} \qquad \mathcal{L}(\mathbf{U}_\beta^{(n)} \mid Q_n) \to N(\theta_0\gamma_\beta, \theta_0^2\mathbf{\Sigma}).$$

Employing the same transformation $\mathbf{P}$ as in (2.17), it is clear from (2.18) that the limiting distribution is the asserted linear combination of non-central chi-square random variables. It remains only to identify the non-centrality parameters. Let $U_1, U_2, \cdots, U_k$ be independent multivariate normal vectors with common covariance matrix $\theta_0^2\mathbf{\Sigma}$ and mean vectors $\theta_0\gamma_1, \theta_0\gamma_2, \cdots, \theta_0\gamma_k$. If we let $\mathbf{P}$ be as in (2.17) and define $\mathbf{V}_\beta$ as $\mathbf{V}_\beta = \mathbf{P}\mathbf{U}_\beta$, then the non-centrality parameter for Z_α is

$$\omega_\alpha^2 = \sum_{\beta=1}^k E_{\theta_0}^2(V_{\alpha\beta} - \bar{V}_\alpha)/\lambda_\alpha,$$

where $V_{\alpha\beta}$ is the αth component of $\mathbf{V}_\beta$, and $\bar{V}_\alpha = (1/k)\sum_{\beta=1}^k V_{\alpha\beta}$. But, if we let $\bar{\mathbf{V}} = (1/k)\sum_{\beta=1}^k \mathbf{V}_\beta$,

$$E_{\theta_0}(\mathbf{V}_\beta - \bar{\mathbf{V}}) = E_{\theta_0}[\mathbf{P}(\mathbf{U}_\beta - \bar{\mathbf{U}})] = \theta_0\mathbf{P}[\mathbf{d}^{(\beta)} - (1/k)\sum_{j=1}^k \mathbf{d}^{(j)}] = \theta_0\mathbf{P}\mathbf{d}^{(\beta)}.$$

The last equality follows because $\sum_\beta [h_\beta(x) - \bar{h}(x)] = 0$. If $\mathbf{P}_{(\alpha)}$ denotes the αth row of $\mathbf{P} = (p_{\alpha\beta})$,

$$E_{\theta_0}(V_{\alpha\beta} - \bar{V}_\alpha) = \theta_0\mathbf{P}_{(\alpha)}\mathbf{d}^{(\beta)} = \theta_0\lambda_\alpha p_{\alpha\beta}$$

taking (2.16) into account. Finally

$$\omega_\alpha^2 = \theta_0^2\lambda_\alpha\sum_\beta p_{\alpha\beta}^2 = \theta_0^2\lambda_\alpha.$$

It has been shown in Lemma 2.2 that $S_n - T_n$ approaches zero in P_n-probability. Now, Lemmas 3.1 (i) and 3.2 show that $\{Q_n\}$ is contiguous to $\{P_n\}$. It follows from the remark made at the beginning of this section that in this situation $S_n - T_n$ approaches zero in Q_n probability as well. The limiting distribution of S_n under Q_n is therefore the same as that just found for T_n. This proves the theorem.

It should be remarked, perhaps, that although only the asymptotic equivalence of S_n and T_n has been established, the asymptotic power of the tests defined in terms of them is the same.

4. Special cases and quadratic forms in L-statistics. Until now our problem has been phrased in quite general terms, and consideration of special cases would be of interest at this point.

For the special case of the two-sample problem ($k = 2$), the covariance matrix $\mathbf{\Sigma}$ is of rank one. The locally most powerful invariant rank statistic reduces to

$$(1/n)\sum_{i,j}(W_{ij} - \bar{W})E_0[h(X_{(i)})h(X_{(j)})], \tag{4.1}$$

where,

$$h(x) = (\partial/\partial\theta)\log[f_1(x;\theta)/f_2(x;\theta)]\,|_{\theta=0}. \tag{4.2}$$

The limit distribution of (4.1) under local alternatives is that of a constant times a non-central chi-square variable with one degree of freedom. In considering locally most powerful rank tests against one sided alternatives $\theta > 0$ certain statistics, known as L-statistics, arise. These are linear combinations of Z_i, $i = 1, 2, \cdots, N$, where Z_i is one or zero according as the ith ordered observation in the combined sample is from the first or the second sample. For the two sample problems at hand the locally best test rejects when $L = n^{-\frac{1}{2}}\sum_{i=1}^{N} E_0[h(X_{(i)})]Z_i$ is large [4]. It seems plausible, then, that a test based on L^2 would have good power for two sided alternatives. Although, L^2 and the statistic given by (4.1) are not the same, the result stated later in this section shows that they are asymptotically equivalent.

In the k-sample problem for $k \geqq 2$, if we set $f_\alpha(x;\theta) = f(x; c_\alpha\theta)$, where $f(x;\theta)$ is a given family of densities, then $h_\alpha(x) = c_\alpha h(x)$, where $h(x)$ is the derivative, with respect to θ, of the logarithm of $f(x;\theta)$ at $\theta = 0$. Thus, $h_\alpha(x) - \bar{h}(x) = (c_\alpha - \bar{c})h(x)$, and the covariance matrix $\mathbf{\Sigma}$ is again of rank one. The locally best invariant rank test statistic S_n then becomes,

$$S_n = (1/n)\textstyle\sum_{\alpha=1}^{k}(c_\alpha - \bar{c})^2\sum_{i,j}(W_{ij} - \bar{W})E_0[h(X_{(i)})h(X_{(j)})] \tag{4.3}$$

so that the locally best invariant test is equivalent to rejecting for large values of

$$S_n^* = (1/n)\textstyle\sum_{i,j}(W_{ij} - \bar{W})E_0[h(X_{(i)})h(X_{(j)})]. \tag{4.4}$$

This statistic has a limiting distribution under $\{Q_n\}$ of a constant times a noncentral chi-square variable with $k - 1$ degrees of freedom. Notice that this test does not depend on $c_1, c_2, \cdots, c_k$. Again, Theorem 4 shows that (4.4) is asymptotically equivalent to a quadratic form in L-statistics.

Let us again put $u_\alpha(x) = h_\alpha(x) - \bar{h}(x)$. Define the L-statistics

$$L_{\alpha\beta}^{(n)} = n^{-\frac{1}{2}}\textstyle\sum_{i\epsilon I(\beta)} E_0\{u_\alpha(X_i) \mid R_{Ni}\}, \qquad 1 \leqq \alpha, \beta \leqq k \tag{4.5}$$

THEOREM 4. *Under the conditions of Theorem* 3,

$$S_n - \textstyle\sum_\alpha\sum_\beta (L_{\alpha\beta}^{(n)})^2 \rightarrow 0 \text{ in } Q_n\text{-probability.}$$

We include here the case $\theta_0 = 0$ *so that, in particular, we have convergence to zero in* P_n *probability.*

This theorem is a consequence of Lemma 2.1 and the contiguity of $\{Q_n\}$ to $\{P_n\}$ and its proof will be omitted.

5. Best average power. In this section we specialize the densities to be of the form $f_\alpha(x:\theta) = f(x; c_\alpha\theta)$, where $c_1, c_2, \cdots, c_k$ are parameters. Of particular interest is the translation case $f_\alpha(x;\theta) = f(x - c_\alpha\theta)$. We have seen in Section 4 that the locally best invariant rank test rejects when S_n, given by (4.3), or S_n^*, given by (4.4), is large. According to Theorem 3, S_n has the limiting distribution of $\lambda_1 Z$ where $\lambda_1 = \sum_\alpha (c_\alpha - \bar{c})^2\sigma^2$ and Z has the non-central chi-square distribution with $k - 1$ degrees of freedom and non-centrality parameter $\theta_0^2\lambda_1$ so that the limiting power depends only on $\theta_0^2\sum(c_\alpha - \bar{c})^2$. This shows that as far as power is concerned we may as well take $\theta_0 = 1$ and suppose that the hypothesis to be tested is $c_1 = c_2 = \cdots = c_k = \bar{c}$. In fact, there is no loss in supposing, as we shall, that

$$\textstyle\sum_\alpha c_\alpha = 0. \tag{5.1}$$

Suppose for the moment that n is fixed. Set $\mathbf{X}_n = (X_1, X_2, \cdots, X_N)$ and $\gamma = (c_1, c_2, \cdots, c_k)$. We shall designate the joint density of the sample by $P_{n\gamma}(\mathbf{X}_n)$ depending on γ and $\theta_n = n^{-\frac{1}{2}}$. Also, let $\Gamma = \{\gamma : \|\gamma\| = r, \sum c_\alpha = 0\}$ be the r-sphere in the $k - 1$ space determined by (5.1). Finally, put Haar measure on Γ.

The average power over Γ of any test ϕ is given by

$$\beta_\phi(r) = \text{const.} \int_\Gamma \int_{\mathfrak{X}_N^n} \phi(\mathbf{X}_n) P_{n\gamma}(\mathbf{X}_n)\, d\mathbf{X}_n\, d\gamma.$$

Of course, we must assume that $P_{n\gamma}(\mathbf{X}_n)$ is measurable in $(\mathbf{X}_n, \gamma)$. According to the generalized Neyman-Pearson lemma, $\beta_\phi(r)$ is maximized within the class of all tests of the same size by the test which rejects when

$$Y_n = \int_\Gamma [P_{n\gamma}(\mathbf{X}_n)/P_{n0}(\mathbf{X}_n)]\, d\gamma > \text{const.} \tag{5.2}$$

THEOREM 5. *Suppose*

$$\lim_{\theta \to 0} E_0\{[f^{\frac{1}{2}}(X;\theta) - f_0^{\frac{1}{2}}(X)]/\theta f_0^{\frac{1}{2}}(X) - h(X)/2\}^2 = 0. \tag{5.3}$$

Then the test given by (5.2) *which maximizes the average power is asymptotically equivalent to the best invariant rank test given in* (1.4).

PROOF. Let us put $U_\beta^{(n)} = n^{-\frac{1}{2}} \sum_{i \varepsilon I(\beta)} h(X_i)$. This is not quite $U_{\alpha\beta}^{(n)}$ defined in (2.13), but $U_{\alpha\beta}^{(n)} = c_\alpha U_\beta^{(n)}$. An evaluation of (2.15) gives

$$T_n = r^2 \sum_\beta (U_\beta^{(n)} - \bar{U}^{(n)})^2. \tag{5.4}$$

In order to prove the theorem, it suffices to show that the test in (5.2) is asymptotically equivalent to the test which rejects when T_n is large.

Denote the logarithm of the liklihood ratio for the βth sample by $L_{n\beta}(c_\beta)$ and apply Lemma 3.1 (iii) where now c_β takes the place of θ_0. One concludes

$$L_{n\beta}(c_\beta) - c_\beta U_\beta^{(n)} \to -c_\beta^2 \sigma^2/2 \tag{5.5}$$

P_n-uniformly in $|c_\beta| \leqq r$. Adding (5.5) over samples yields

$$L_n(\gamma) - \sum_\beta c_\beta U_\beta^{(n)} \to -r^2\sigma^2/2 \tag{5.6}$$

P_n-uniformly in $\|\gamma\| = r$, where we have set

$$L_n(\gamma) = \sum_\beta L_{n\beta}(c_\beta) = \log [P_{n\gamma}(\mathbf{X}_n)/P_{n0}(\mathbf{X}_n)].$$

Since $\sum_\beta c_\beta = 0$ and each $U_\beta^{(n)}$ has a limiting distribution, (5.6) is, moreover, equivalent to

$$Z_n(\gamma) = |\exp [L_n(\gamma)] - c \exp (\gamma' \xi_n)| \to 0 \tag{5.7}$$

P_n-uniformly in $\|\gamma\| = r$, where $\xi_{n\beta} = U_\beta^{(n)} - \bar{U}^{(n)}$, $\xi_n = (\xi_{n1}, \xi_{n2}, \cdots, \xi_{nk})$, and $c = \exp(-r^2\sigma^2/2)$.

We will prove below that it is permissible to integrate (5.7) over Γ. Granting the possibility of this for the moment, we would obtain

$$I_n = \int_\Gamma Z_n(\gamma)\, d\gamma \to 0 \quad \text{in } P_n\text{-probability}. \tag{5.8}$$

As a consequence,

$$\int_\Gamma \exp\,(L_n(\gamma)]\,d\gamma - c\int_\Gamma \exp\,(\gamma'\xi_n)\,d\gamma \to 0 \quad \text{in } P_n\text{-probability}.$$

The contiguity condition (5.3) would then guarantee that the difference approaches zero in Q_n-probability as well. It is well known that $c\int_\Gamma \exp\,(\gamma'\xi_n)\,d\gamma$ is an increasing function of $\|\xi_n\|$, and from (5.4), $T_n = r^2\|\xi_n\|^2$. These observations show that the test defined by (5.2) is asymptotically equivalent to the test which rejects when T_n is large. The theorem will be proved by establishing (5.8).

Let A_n, $n = 1, 2, \cdots$ be any sequence of sets with $A_n \subset \mathfrak{X}_N$ and

$$\lim_n P_n(A_n) = 0.$$

It is claimed that

$$(5.9)\qquad \int_{A_n} \exp\,[L_n(\gamma)]\,dP_n = \int_{A_n} P_{n\gamma}(\mathbf{X}_n)\,dX_n = \theta_n(A_n\,;\gamma) \to 0$$

uniformly in $\|\gamma\| = 1$. Suppose to the contrary that there exist sequences A_n and $\gamma_n = (c_{n1}, c_{n2}, \cdots, c_{nk})$ for which $P_n(A_n) \to 0$, $\|\gamma_n\| = 1$, but

$$(5.10)\qquad Q_n(A_n\,;\gamma_n) \geqq \epsilon > 0.$$

We may suppose without loss that $\gamma_n \to \gamma_0$. Define $W_{n\beta}(c_{n\beta})$ for the βth sample as in (3.1). The evaluations (3.6) and (3.7) show that $W_{n\beta}(c_{n\beta})$ has a limiting normal distribution under P_n whose mean is equal to minus one-fourth its variance. This being true for each sample then shows that the particular sequence $\{Q_n(\gamma_n)\}$ considered is contiguous to $\{P_n\}$, thereby contradicting (5.10).

Fix $\epsilon > 0$. Since ξ_n has a limiting distribution, one can choose K so that, putting $M_n = \{\mathbf{X}_n : \|\xi_n\| < K\}$, $P_n(M_n) > 1 - \epsilon$ for all n. Then $P_n(A_n) \to 0$ implies

$$(5.11)\qquad \int_{M_n \cap A_n} \exp\,(\gamma'\xi_n)\,dP_n \leqq e^{Kr}P_n(A_n) \to 0.$$

Together, (5.9) and (5.11) show that

$$(5.12)\qquad \int_{M_n \cap A_n} Z_n(\gamma)\,dP_n \to 0$$

uniformly in γ and $\{A_n\}$ whenever $P_n(A_n) \to 0$.

Let $\delta > 0$ be arbitrary and define the sets $A_n(\gamma) = \{\mathbf{X}_n : Z_n(\gamma) > \delta\}$. Now (5.7) states that $P_n(A_n(\gamma)) \to 0$ uniformly in γ. Consequently, in view of (5.12)

$$\int_{M_n} Z_n(\gamma)\,dP_n = \int_{M_n \cap A_n(\gamma)} Z_n(\gamma)\,dP_n + \int_{M_n \cap A_n'(\gamma)} Z_n(\gamma)\,dP_n < 2\delta$$

uniformly in γ, provided only n is sufficiently large. Integrating this over Γ gives

$$(5.13)\qquad \int_{M_n} I_n\,dP_n = \int_\Gamma \int_{M_n} Z_n(\gamma)\,dP_n\,d\gamma < (\text{const.})\,\delta.$$

One concludes from (5.13) since δ was arbitrary $\lim_n P_n\{M_n \cap (I_n > \epsilon)\} = 0$. This shows that $\lim_n P_n(I_n > \epsilon) < \epsilon$ because $P_n(M_n) > 1 - \epsilon$, and the fact that ϵ was arbitrary establishes (5.8). The proof of the theorem is now complete.

6. Some classical tests for the k-sample problem. Here we relate the general theory to the important Kruskal-Wallis and Mood-Brown k-sample tests. Of course, the sample sizes are assumed to be equal in this discussion.

From Theorem 4 and remarks made in Section 4, it follows for translation type alternatives $f_\alpha(x; \theta) = f(x - c_\alpha\theta)$ that the locally most powerful invariant rank test is asymptotically equivalent to a test which rejects when $\tilde{S}_n = \sum_\beta (L_\beta^{(n)})^2$ is large, where

$$L_\beta^{(n)} = n^{-\frac{1}{2}} \sum_{i \epsilon I(\beta)} E_0\{h(X_i) | R_{Ni}\}.$$

The limiting distribution of this test statistic was shown to be $\sigma^2\chi^2_{k-1}(\omega^2)$ where $\sigma^2 = E_0[h^2(X)]$ and the non-certrality parameter $\omega^2 = \sum_\alpha (c_\alpha - \bar{c})^2\sigma^2$.

In the case of the logistic density $f(x) = e^{-x}(1 + e^{-x})^{-2}$, it is easily checked that our basic assumptions (1.2) and (3.12) hold. As a matter of fact, in order to verify (3.12) in the translation parameter case it suffices to show that $(d/dx)f^{\frac{1}{2}}(x)$ is square integrable ([8], Lemma 4.3). Now, $L_\beta^{(n)} = 2n^{\frac{1}{2}}[\bar{R}_\beta - (N + 1)/2]/(N + 1)$, where $\bar{R}_\beta$ is the average of the ranks of the observations in the βth sample. The test statistic $\tilde{S}_n$ is seen to be asymptotically equivalent to the Kruskal-Wallis test statistic $H = [12n/N(N + 1)] \sum_\beta [\bar{R}_\beta - (N + 1)/2]^2$ up to a factor of $\sigma^2 = \frac{1}{3}$. (See [1]). Therefore, it can be stated that the Kruskal-Wallis test is asymptotically equivalent to the locally most powerful invariant rank test, and also to the test which maximizes, among all tests, the average power over spheres $\sum_\alpha(c_\alpha - \bar{c})^2 = r^2$.

A similar argument can be applied to the double exponential density $f(x) = \exp(-\frac{1}{2}|x|)$ to prove the analogous optimal properties for the Mood-Brown test which is based on the statistic $(4/n) \sum_\beta (M_\beta - (n/2))^2$, where M_β is the number of observations from the βth sample which exceed the median of the combined sample.

Finally, we state that the k-sample analogue of the Fisher-Yates test is also asymptotically equivalent to the locally best invariant rank test as well as the test which maximizes the average power over spheres, when the alternatives are normal. It might be remarked, however, that none of the three tests considered is actually locally best, but only equivalent in the limit to the locally best test specified by (1.4).

REFERENCES

[1] ANDREWS, F. C. (1954). Asymptotic behavior of some rank tests for analysis of variance. *Ann. Math. Statist.* **29** 724–735.

[2] ANDREWS, F. C. and TRUAX, D. R. (1964). Locally most powerful rank tests for several sample problems. *Metrika.* **8** 16–24.

[3] CAPON, J. (1961). Asymptotic efficiency of certain locally most powerful rank tests. *Ann. Math. Statist.* **32** 88–100.

[4] CHERNOFF, H. and SAVAGE, I. R. (1958). Asymptotic normality and efficiency of certain non-parametric test statistics. *Ann. Math. Statist.* **29** 972–994.

[5] DOOB, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[6] DWASS, M. (1956). The large sample power of rank order tests in the two sample case. *Ann. Math. Statist.* **27** 352–374.

[7] Hájek, J. (1961). Some extensions of the Wald-Wolfowitz-Noether theorem. *Ann. Math. Statist.* **32** 506–523.

[8] Hájek, J. (1962). Asymptotically most powerful rank-order tests. *Ann. Math. Statist.* **33** 1124–1147.

[9] Lehmann, E. L. (1953). The power of rank tests. *Ann. Math. Statist.* **24** 23–43.

[10] Puri, M. L. (1964). Asymptotic efficiency of a class of c-sample tests. *Ann. Math. Statist.* **35** 102–121.

[11] Uzawa, H. (1960). Locally most powerful rank tests for two sample problems. *Ann. Math. Statist.* **31** 685–702.