SOME NEW DISTRIBUTION-FREE STATISTICS

By C. B. Bell and K. A. Doksum[1]

San Diego State College and *University of California, Berkeley*

1. Introduction and summary. For the two-sample problem, Wilcoxon [21], Fisher and Yates [6], Terry [19], Hoeffding [10], Hodges and Lehmann [8], Savage [17], Chernoff and Savage [3], Lehmann [13], Capon [2] and others have considered rank-sum statistics equivalent to $S_N(H) = m^{-1} \sum E(Z(R(X_i)) \mid H) - n^{-1} \sum E(Z(R(Y_i)) \mid H)$, where $E(Z(j) \mid H)$ is the expectation of the jth order statistic of a sample of size $N = m + n$ from a population with cpf (cumulative probability function) H and $R(X_i)$ $[R(Y_j)]$ is the rank of $X_i[Y_j]$ in the combined sample of X's and Y's.

In order to perform tests based on these statistics one needs special tables of expected values as well as tables of the hypothesis distribution. Further, in general, exact desired significance levels can only be achieved through randomization.

The object of this note is to introduce rank-sum statistics in which one randomizes initially and circumvents the necessity of two special tables. These new randomized statistics, which are formed by deleting the expectation signs "E" in $S_N(H)$, generally satisfy the same asymptotic goodness criteria as their non-randomized counterparts. Moreover, they have the added advantage that for appropriate choices of the parameters they have null hypothesis distributions which are continuous, known and tabulated (e.g., normal, χ^2, etc.) In particular, one of the new two-sample statistics has an exact normal distribution and is asymptotically uniformly more efficient than the t-test for translation alternatives.

This idea is extended to obtain randomized rank-sum statistics for the independence, randomness, k-sample and two-factor problems analogous to the statistics of Friedman [7], Puri [15], Stuart [18] and others. As in the references listed above, prime interest will be in those cases for which H is normal, uniform or exponential. However, the methodology is equally applicable to other continuous cpf's H.

2. Two-sample tests. Basic to the obtaining of the aforementioned desirable null hypothesis distributions is the following lemma.

Lemma 2.1. *Let F be a continuous* cpf *and let H be any* cpf. *If $W_1, W_2, \cdots, W_N$, and $Z_1, Z_2, \cdots, Z_N$ are independent random samples with* cpf's *F and H, respectively, if $R(W_i)$ denotes the rank of W_i among $W_1, W_2, \cdots, W_N$, and if $Z(i)$ is the* ith *order statistic of $Z_1, Z_2, \cdots, Z_N$; then $Z(R(W_1)), Z(R(W_2)), \cdots, Z(R(W_N))$ have the same joint distribution as the random sample $Z_1, Z_2, \cdots, Z_N$.*

Received 16 January 1964; revised 6 June 1964.

[1] This research was supported by the National Science Foundation through grant NSFG-25220.

PROOF. Let A_N be a Borel set in N dimensional Euclidean space. First compute

$$\begin{aligned}P\{[Z(R(W_1)), \cdots, Z(R(W_N))] \varepsilon A_N\} &= \sum P\{[Z(r_1), \cdots, Z(r_N)] \varepsilon A_N \mid R(W_1) = r_1, \cdots, R(W_N) = r_N\} \\ &\quad \cdot P\{R(W_1) = r_1, \cdots, R(W_N) = r_N\} \\ &= (\sum P\{[Z(r_1), \cdots, Z(r_N)] \varepsilon A_N\})/N!\end{aligned}$$

where the summation is over all possible permutations $\{r_1, \cdots, r_N\}$ of the ranks $\{1, \cdots, N\}$.

Next compute

$$\begin{aligned}P\{[Z_1, \cdots, Z_N] \varepsilon A_N\} &= \sum P\{[Z(r_1), \cdots, Z(r_N)] \varepsilon A_N \mid R(Z_1) = r_1, \cdots, R(Z_N) = r_N\} \\ &\quad \cdot P\{R(Z_1) = r_1, \cdots, R(Z_N) = r_N\} \\ &= (\sum P\{[Z(r_1), \cdots, Z(r_N)] \varepsilon A_N\})/N!.\end{aligned}$$

The proof is complete since it follows that

$$P\{[Z(R(W_1)), \cdots, Z(R(W_N))] \varepsilon A_N\} = P\{[Z_1, \cdots, Z_N] \varepsilon A_N\}$$

for all Borel sets A_N.

Let $X_1, X_2, \cdots, X_m$ and $Y_1, Y_2, \cdots, Y_n$ be two independent random samples from populations with continuous cpf's F and G, respectively. The hypothesis considered is H_0: $F = G$ against one-sided location alternatives. Let $N = m + n$ and let $Z(1), Z(2), \cdots, Z(N)$ be the order statistics of a random sample $Z_1, Z_2, \cdots, Z_N$ from a population with cpf H (where $Z_1, Z_2, \cdots, Z_N$ is independent of $X_1, \cdots, X_m$; $Y_1, \cdots, Y_n$). Further, let $R(X_i)[R(Y_i)]$ denote the rank of $X_i[Y_i]$ in the combined sample of X's and Y's.

The randomized rank-sum statistic $T_N(H)$ is now defined by $T_N(H) = m^{-1} \sum_{i=1}^{m} Z(R(X_i)) - n^{-1} \sum_{i=1}^{n} Z(R(Y_i))$.

One sees from the formula for $T_N(H)$ that one essentially replaces the original data $X_1, X_2, \cdots, X_m$; $Y_1, Y_2, \cdots, Y_n$, by a random sample $Z_1, Z_2, \cdots, Z_N$ known to have cpf H, and forms the difference of the sample means of the transformed X- and Y-values, where the transformation depends only on the ranks of the X's and Y's in the combined sample. This randomized statistic $T_N(H)$ has the advantage over its nonrandomized rank-sum counterpart $S_N(H) = m^{-1} \sum E(Z(R(X_i)) \mid H) - n^{-1} \sum E(Z(R(Y_i)) \mid H)$ as defined in Section 1, of having a continuous distribution for which all significance levels are exactly available, and of having a well-tabulated hypothesis distribution when one chooses H properly. This is reflected in the following theorem and example. The theorem is an immediate consequence of the lemma above.

THEOREM 2.2. *When $F = G$, then $T_N(H)$ has the distribution of the difference of means of independent samples of sizes m and n, respectively, from populations with* cpf's H.

In particular,

EXAMPLE 2.1.

(i) If $H = \Phi$, the standard normal cpf, then $(mn/N)^{\frac{1}{2}}T_N(\Phi) = (mn/N)^{\frac{1}{2}}\{m^{-1}\sum Z(R(X_i)) - n^{-1}\sum Z(R(Y_i))\}$ has a standard normal cpf;

(ii) if $H = U$, the standard uniform cpf, and if $n = m$, then the cpf of $\frac{1}{2} + \frac{1}{2}T_N(U) = \frac{1}{2} + \frac{1}{2}\{n^{-1}\sum Z(R(X_i)) - n^{-1}\sum Z(R(Y_i))\}$ is the same as the cpf of the mean of a sample of N standard uniform variables, i.e., its density is $N^N[(N-1)!]^{-1}\sum(-1)^r\binom{N}{r}(x - r/N)^{N-1}$, where the summation is over $r \leqq Nx$;

(iii) if $H = K$, the standard exponential cpf on $[0, \infty)$, and if $n = m$, then the cpf of $NT_N(K) = 2\{\sum Z(R(X_i)) - \sum Z(R(Y_i))\}$ is the same as the cpf of the difference of two independent chi-square variables, each with N degrees of freedom. This cpf is discussed in Section 5; see Table 5.1.

On applying $T_N(H)$, one will need to construct the random sample $Z_1, \cdots, Z_N$. This can be done by referring to tables such as [16]. Note that from $Z(R(X_1)), \cdots, Z(R(X_m)); Z(R(Y_1)), \cdots, Z(R(Y_n))$ one can obtain the ranks $R(X_1), \cdots, R(X_m); R(Y_1), \cdots, R(Y_n)$. Thus no "information" is lost on passing from the ranks to $Z(R(X_1)), \cdots, Z(R(X_m)); Z(R(Y_1)), \cdots, Z(R(Y_n))$, and on applying the rank-sum statistic $S_N(H)$ to $Z(R(X_1)), \cdots, Z(R(X_m)); Z(R(Y_1)), \cdots, Z(R(Y_n))$ one obtains the same result as on applying $S_N(H)$ to $X_1, \cdots, X_m; Y_1, \cdots, Y_n$.

On comparing the computations involved in $S_N(H)$ and $T_N(H)$, note that $S_N(H)$ can only be applied when the expected values $E(Z(1) \mid H), \cdots, E(Z(N) \mid H)$ are easily computable or tables of them are available. Further, the significance levels of $S_N(H)$ must be obtained by computing the ordering induced by $S_N(H)$ on the different permutations of the ranks, or by referring to special tables. In the latter case, the true significance levels (which are of the form $k\binom{N}{n}^{-1}$) are usually smaller than the ones listed in the table.

In the next section on the k-sample problem, randomized statistics will be even more useful since for that problem the significance levels of the usual rank tests are hard to compute and tables are not available.

In view of the construction of the randomized statistic $T_N(H)$, one might expect that it is in some sense asymptotically equivalent to its rank-sum counterpart $S_N(H)$. This will be reflected in theorems to follow, but first one needs two lemmas.

LEMMA 2.3. *If one defines $a_{Ni} = 1$ if the ith observation in the ordered combined sample of X's and Y's is an X, $a_{Ni} = 0$ otherwise, and if we define $\lambda_1 = m/N$, $\lambda_2 = n/N$; then one may write $T_N(H) = (\lambda_1\lambda_2N)^{-1}\sum_{i=1}^N Z(i)(a_{Ni} - \lambda_1)$ and $S_N(H) = (\lambda_1\lambda_2N)^{-1}\sum_{i=1}^N E(Z(i) \mid H)(a_{Ni} - \lambda_1)$ where the Z's and a_{Ni}'s are independent.*

PROOF. Follows from straightforward calculations.

LEMMA 2.4. *If the random variable with* cpf *H has second moments, then $N^{-1}\sum_{i=1}^N \mathrm{Var}\,(Z(i) \mid H) \to \infty$.*

PROOF. Hoeffding [11] has essentially shown that $N^{-1}\sum E^2(Z(i) \mid H) \to$

$\int x^2\, dH(x)$ as $N \to \infty$. It remains to note that $N^{-1} \sum \operatorname{Var}(Z(i) \mid H) = N^{-1} \sum E(Z^2(i) \mid H) - N^{-1} \sum E^2(Z(i) \mid H)$ and that $N^{-1} \sum E(Z^2(i) \mid H) = N^{-1}E(\sum Z^2(i) \mid H) = N^{-1}E(\sum Z_i^2 \mid H) = \int x^2\, dH(x)$.

THEOREM 2.5. *If H is any* cpf, *then*

(i) $E(T_N(H) \mid F, G) = E(S_N(H) \mid F, G)$. *If H has second moments and if λ_1 is bounded away from* 0 *and* 1, *then*

(ii) $\operatorname{Var}(N^{\frac{1}{2}}[(T_N(H) - S_N(H)] \mid F, F) \to 0$ *as* $N \to \infty$, *and*

(iii) $\operatorname{Var}(N^{\frac{1}{2}}(T_N(H) - S_N(H)) \mid F, G) \to 0$ *as* $N \to \infty$ *whenever one of the following is true*

(a) $\sum_{i<j} \operatorname{Cov}(Z(i), Z(j))E(a_{Ni} - \lambda_1)(a_{Nj} - \lambda_1) = o(N)$ as $N \to \infty$,

(b) $\sum_{i<j} \operatorname{Cov}(Z(i), Z(j))E(a_{Ni} - \lambda_1)(a_{Nj} - \lambda_1) \leqq 0$.

PROOF. (i) follows at once from Lemma 2.3. For the remaining parts of the theorem, observe that Lemma 2.3 and a few computations yield

$$\begin{aligned}
&\operatorname{Var}(N^{\frac{1}{2}}[T_N(H) - S_N(H)] \mid F, G)\\
&\quad = [\lambda_1\lambda_2]^{-2}N^{-1}\textstyle\sum_{i,j} \operatorname{Cov}(Z(i), Z(j))E(a_{Ni} - \lambda_1)(a_{Nj} - \lambda_1)\\
&\quad = 2[\lambda_1\lambda_2]^{-2}N^{-1}\textstyle\sum_{i<j} \operatorname{Cov}(Z(i), Z(j))E(a_{Ni} - \lambda_1)(a_{Nj} - \lambda_1)\\
&\qquad + [\lambda_1\lambda_2]^{-2}N^{-1}\textstyle\sum \operatorname{Var}(Z(i) \mid H)E(a_{Ni} - \lambda_1)^2.
\end{aligned}$$

Using Lemma 2.4, (iii) now follows at once. To obtain (ii), observe that when $F = G$, then $E(a_{Ni} - \lambda_1)(a_{Nj} - \lambda_1) = \operatorname{Cov}(a_{Ni}, a_{Nj}) = -\lambda_1\lambda_2(N-1)^{-1}$ for $i \neq j$. (ii) now follows from (iii) (b).

EXAMPLE 2.1. In particular, if $H = U$, the standard uniform cpf (so that $S_N(U)$ is the Wilcoxon statistic), then

$$\operatorname{Var}(N^{\frac{1}{2}}[T_N(U) - S_N(U)] \mid F, F) = [12\lambda_1\lambda_2(N-1)]^{-1}.$$

This follows since when $H = U$, then

$$\operatorname{Cov}(Z(i), Z(j)) = i(N + 1 - j)(N+1)^{-2}(N+2)^{-1} \qquad \text{for } i \leqq j;$$

and when $F = G$, then $\operatorname{Var}(a_{Ni}) = \lambda_1\lambda_2$.

From Theorem 2.5, one can obtain results both about $S_N(H)$ and $T_N(H)$. For instance, (ii) implies that $S_N(H)$ and $T_N(H)$ have the same asymptotic distribution under the hypothesis $F = G$ whenever H has second moments. Since $T_N(H)$ is asymptotically normally distributed by Theorem 2.2, it follows that

COROLLARY 2.6. *If H has second moments, if $F = G$, and if $\lambda_1 = m/N$ is bounded away from* 0 *and* 1, *then the distribution of $(mn/N)^{\frac{1}{2}}\sigma_H^{-1}S_N(H)$ tends to the standard normal distribution as $N \to \infty$.*

Similar results have been obtained by Dwass [5] using essentially the same approach; moreover the result was obtained by Chernoff and Savage [3] using a different approach and the "smoothness" condition

$$|(d^i/du^i)J(u)| \leqq K[u(1-u)]^{-i-\frac{1}{2}+\delta}, \tag{2.1}$$

for $i = 0, 1, 2$ and some $\delta > 0$ where $J = H^{-1}$ and K is a constant.

The result of Chernoff and Savage holds when $F \neq G$, thus

COROLLARY 2.7. *If either* (a) *or* (b) *of Theorem* 2.5 *hold, if* (2.1) *is satisfied and if* λ_1 *is bounded away from* 0 *and* 1, *then the* cpf *of* $[T_N(H) - \mu_N]\sigma_N^{-1}$ *tends to a standard normal* cpf *as* $N \to \infty$, *where* μ_N *and* σ_N^2 *are the mean and variance of* $S_N(H)$ *respectively* (*see Theorem* 1 *of Chernoff and Savage*).

One can now compute the asymptotic relative efficiency (ARE) of $T_N(H)$ with respect to $S_N(H)$. The concept of ARE is due to Pitman and is derived as follows. Let T and T' be two test statistics for $H_0 : \theta = \theta_0$ vs $H_1 : \theta > \theta_0$. Then the relative efficiency of T' with respect to T is the ratio n/n' where n and n' are the number of observations required to give T and T' the same power β for a given level of significance α. The ARE of T' with respect to T is written $A(T', T)$ and is obtained by considering the limit of n/n' for a sequence on alternatives depending on the sample size and converging to H_0 in such a way that the power of both tests tends to a limit <1.

The following results hold for asymptotic relative efficiencies.

THEOREM 2.8. *If H has second moments, then*

(i) $A(T(H), S(H)) = 1$ *for each class of continuous alternatives* $\{G_N\}$ *such that* $G_N \geqq F$ *and* $\int G_N \, dF \to \frac{1}{2}$ *as* $N \to \infty$.

(ii) *In particular,* $A(T(H), S(H)) = 1$ *for H normal, uniform or exponential, and for translation alternatives* $G(x) = F(x + \theta_N)$, $\theta_N \to 0$ *as* $N \to \infty$.

PROOF. Write $P_N = P(Y < X) = \int G_N \, dF$. When $P_N = \frac{1}{2}$, then

$$E(a_{Ni} - \lambda_1)(a_{Nj} - \lambda_1) < 0,$$

thus there exists a positive number δ_N such that whenever $|P_N - \frac{1}{2}| < \delta_N$, then $E(a_{Ni} - \lambda_1)(a_{Nj} - \lambda_1) \leqq 0$. From Theorem 2.5 (iii) (b), it follows that whenever $|P_N - \frac{1}{2}| < \delta_N$, $T_N(H)$ and $S_N(H)$ have the same asymptotic distributions, thus they have the same ARE.

It follows from Theorem 2.8 that the ARE properties of $T_N(H)$ are the same as the well-known ARE properties of the rank-sum statistic $S_N(H)$. In particular one has

COROLLARY 2.9. *Let* Φ, U *and* K *denote the standard normal* cpf, *the standard uniform* cpf *and the standard exponential* cpf *on* $[0, \infty)$ *respectively. Then*

(i) $T_N(\Phi)$, $T_N(U)$ *and* $T_N(K)$ *are asymptotically efficient and asymptotically locally most powerful rank tests for normal translation alternatives, logistic translation alternatives and exponential alternatives respectively, where exponential alternatives are of the form*

$$F(x) = 1 - e^{-\theta_0 x}, \qquad G(x) = 1 - e^{-\theta_1 x}, \qquad x \geqq 0, \theta_1 > \theta_0 .$$

(ii) *If F has a density and finite second moments and if t is Student's statistic, then* $A(T(\Phi), t) \geqq 1$ *for translation alternatives with equality if and only if F is a normal* cpf.

PROOF. (i) follows at once from Theorem 2.8 since the corresponding results for $S_N(\Phi)$, $S_N(U)$ and $S_N(K)$ are known (see for instance Lehmann [13] and Capon [2]). (ii) follows similarly from Chernoff and Savage [3].

Although the finite sample size distribution of $T_N(H)$ is simple when $F = G$, this distribution has not been obtained when $F \neq G$. Considerations of probabilities of rank orderings for normal translation alternatives indicates, however, that whenever the significance level α is of the form $k\binom{N}{n}^{-1}$, then $S_N(H)$ is slightly more powerful than $T_N(H)$ for normal translation alternatives.

The following finite sample size result was obtained by Lehmann in [13], p. 238.

THEOREM 2.10. *If H is continuous, then $T_N(H)$ is unbiased for one-sided alternatives of the type $G < F$.*

3. k-sample tests. The idea introduced in the preceding section will be extended to the k-sample problem described below.

$X_{11}, \cdots, X_{1n_1}; X_{21}, \cdots, X_{2n_2}, \cdots, X_{k1}, \cdots, X_{kn_k}$ are k independent random samples with continuous cpf's $F_1, F_2, \cdots, F_k$ respectively. The hypothesis considered is $H_0 : F_1 = F_2 \cdots = F_k$ against location alternatives.

Let $N = \sum n_i$ and let $Z(1), Z(2), \cdots, Z(N)$ be the order statistics of a random sample of size N from a population with cpf H. Now reject H_0 for large values of $V(H) = \sum_{i=1}^{k} n_i(\bar{Z}_{i\cdot} - \bar{Z}_{\cdot\cdot})^2$ where $\bar{Z}_{i\cdot} = n_i^{-1}\sum_{j=1}^{n_i} Z(R(X_{ij}))$, $\bar{Z}_{\cdot\cdot} = N^{-1}\sum_{i=1}^{N} Z(i)$ and $R(X_{ij})$ is the rank of X_{ij} in the combined sample of X's.

As indicated by the formula, the procedure is to

(1) rank the combined sample of X's,

(2) replace each X_{ij} by $Z(R(X_{ij}))$, i.e., the Z value with the same rank in the Z-sample as X_{ij} has in the X-sample.

(3) compute the numerator of the usual $\mathfrak{F}$-statistic for the Z's obtained in (2) above.

One recalls that the denominator of the usual $\mathfrak{F}$-statistic is the pooled variance, which is an estimate of the unknown common variance in the k-sample case. This estimate is unnecessary in the randomized statistic $V(H)$ above since a random sample of known variance is imposed by the procedure.

The distribution of $V(H)$ is given by the following theorem, which makes use of the basic Lemma (2.1) and a well-known result ([12], p. 504).

LEMMA 3.1. *Under H_0,*

(i) *if $H = \Phi$, the standard normal* cpf, *then $V(H)$ has exactly a chi-square distribution with $(k - 1)$ degrees of freedom,*

(ii) *if the variance $\sigma_H^2 = \mathrm{Var}\,(Z \mid H)$ exists, then $\sigma_H^{-2}\, V(H)$ is asymptotically a chi-square variate with $(k - 1)$ degrees of freedom,*

(iii) *in all cases, $V(H)$ has the distribution of $\sum_{i=1}^{k} n_i(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2$, where $\{Y_{ij}\}$ are k independent random samples of sizes n_i from a population with* cpf's H.

PROOF. Set $F_1 = F_2 = \cdots = F_k = F$, then $X_{11}, X_{12}, \cdots, X_{kn_k}$ is a random sample from a population with cpf F, thus (iii) follows from Lemma 2.1. (i) and (ii) now follow at once from (iii).

From the above theorem it is clear that by a judicious choice of the cpf H, one can get a desirable null hypothesis distribution. The next step is to give some indication of the goodness properties of the test statistic $V(H)$. An immediate extension of the nonrandomized two-sample rank-sum statistic $S_N(H)$ is $L(H) = \sum n_i(\bar{E}_{i\cdot} - \mu_H)^2$ where $\bar{E}_{i\cdot} = n_i^{-1}\sum_{j=1}^{n_i} E(Z(R(X_{ij})) \mid H)$ and $\mu_H = \int x\, dH(x)$.

Statistics equivalent to $L(H)$ have been studied by Puri [15] and others. Puri has shown that for alternatives of the form $F_i(x) = F(x + \theta_i/N^{\frac{1}{2}})$, the ARE of $L(H)$ with respect to the classical $\mathfrak{F}$ test is under appropriate conditions independent of the number of samples k and thus equals $A(S(H), t)$.

The following theorem will enable us to show that $V(H)$ has the same properties as Puri shows $L(H)$ to have. The following notation is needed.

Let $\lambda_i = n_i/N$ and define $a_{Nj}^{(i)} = 1$ if the jth observation in the ordered combined sample of N X's is from the ith sample, $a_{Nj}^{(i)} = 0$ otherwise. Further, let $T_{Ni} = n_i^{-\frac{1}{2}}\sum_{j=1}^{N} Z(j)(a_{Nj}^{(i)} - \lambda_i)$ and $S_{Ni} = n_i^{-\frac{1}{2}}\sum_{j=1}^{N} E(Z(j) \mid H)(a_{Nj}^{(i)} - \lambda_i)$; then it is easy to see that one may write $V(H) = \sum_{i=1}^{k} T_{Ni}^2$ and $L(H) = \sum_{i=1}^{k} S_{Ni}^2$. It follows from Lemma 2.1 that whenever $\sigma_H^2 < \infty$ then T_{Ni} converges in probability to a normal variable T_i under H_0. Moreover, Puri has given conditions under which S_{Ni} converges in probability to a normal variable S_i. One can now easily prove

THEOREM 3.2. *If* $\int |x|\, dH(x) < \infty$, *then*

(i) $\sum_{i=1}^{k} E^2(T_{Ni}) = \sum_{i=1}^{k} E^2(S_{Ni})$ *under the null hypothesis and under any alternative,* i.e., $V(H)$ *and* $L(H)$ *have the same noncentrality parameter. If* $\int x^2\, dH(x) < \infty$ *and if* λ_i *is bounded away from* 0 *and* 1 *for each* $i = 1, \cdots, k$, *then*

(ii) $[V(H) - L(H)]$ *tends to zero in probability as* $N \to \infty$ *under* H_0. *If further, either* S_{Ni} *or* T_{Ni} *converges in probability to a random variable for each* $i = 1, \cdots, k$, *then*

(iii) $[V(H) - L(H)]$ *tends to zero in probability whenever one of the following is true for each* $i = 1, \cdots, k$.

(a) $\sum_{s<j} \operatorname{Cov}(Z(s), Z(j))E(a_{Ns}^{(i)} - \lambda_i)(a_{Nj}^{(i)} - \lambda_i) = o(N)$ *as* $N \to \infty$.

(b) $\sum_{s<j} \operatorname{Cov}(Z(s), Z(j))E(a_{Ns}^{(i)} - \lambda_i)(a_{Nj}^{(i)} - \lambda_i) \leqq 0$.

PROOF. (i) is immediate since the Z's and a_{Ni}'s are independent. Suppose T_{Ni} converges in probability to T_i (say) for each $i = 1, \cdots, k$. Then by the arguments of Theorem 2.6, S_{Ni} converges in probability to T_i. Thus both $V(H)$ and $L(H)$ converge in probability to $\sum_{i=1}^{k} T_i^2$. This completes the proof of (ii) and (iii).

From Lemma 3.1(ii) and Theorem 3.2(ii) one obtains

COROLLARY 3.3. *If* $\int x^2\, dH(x) = \sigma_H^2 < \infty$ *and if* λ_i *is bounded away from* 0 *and* 1 *then* $\sigma_H^{-2}L(H)$ *is asymptotically a chi-square variate with* $(k - 1)$ *degrees of freedom under* H_0.

Further properties of $V(H)$ can now be obtained from Theorem 3.2 and Puri [15].

COROLLARY 3.4.

(i) *If* λ_i *is bounded away from* 0 *and* 1, *if* $\int x^2\, dH(x) < \infty$ *and if either* T_{Ni} *or* S_{Ni} *converge in probability for each* $i = 1, \cdots, k$, *then* $A(V(H), L(H)) = 1$ *for each class of alternatives* $\{F_{iN}\}$ *such that for some* F, $|F_{iN} - F| \to 0$ *as* $N \to \infty$ *for each* $i = 1, \cdots, k$.

(ii) *In particular,* $A(V(H), \mathrm{L}(H)) = 1$ *for* H *normal, uniform or exponential and for translation alternatives* $F_i(x) = F(x + \theta_{iN})$, $\theta_{iN} \to 0$ *as* $N \to \infty$.

(iii) *If* F *of* (ii) *satisfies the regularity conditions of Theorem* 8.2 *of Puri* [15], *then* $A(V(\Phi), \mathfrak{F}) \geqq 1$ *for translation alternatives with equality* iff F *is a normal* cpf *where* $\mathfrak{F}$ *is the usual* $\mathfrak{F}$ *ratio test statistic.*

4. 2-factor tests. In view of the distribution and ARE results of the preceding sections it is reasonable to attempt to extend the idea of randomized rank-sum statistics to the case of two-factor experiments with one observation per cell.

2nd factor	1st factor			
	1	2	$\cdots$	c
1	X_{11}	X_{12}	$\cdots$	X_{1c}
2	X_{21}	X_{22}	$\cdots$	X_{2c}
$\vdots$				
r	X_{r1}	X_{r2}	$\cdots$	X_{rc}

It is assumed that all the X's are independent and that each X_{ij} has a continuous cpf F_{ij}.

First one notes that if it is assumed that there is no row effect, then testing the hypothesis "H_0 : No column effect," becomes a k-sample problem. The analogous result holds when it is assumed that there is no row effect. For that reason, attention will be restricted to the hypothesis "H_0 : No column effect" under the assumption that there is row effect. More precisely, the hypothesis to be considered is $H_0 : F_{i1} = F_{i2} = \cdots = F_{ic}$ for each $i = 1, 2, \cdots, r$, against the alternative that there is a shift in the location of the columns.

Perhaps the most natural extension of the procedure of the preceding sections consists of

(1) ranking the collection $\{X_{ij}\}$ of X's;

(2) replacing each X_{ij} by $Z(R(X_{ij}))$, i.e., the Z value in the independent random sample $Z_1, Z_2, \cdots, Z_{rc}$ from H with the same rank among the Z's as X_{ij} has among the X's; and

(3) computing the numerator of the appropriate $\mathfrak{F}$-statistic.

This procedure leads to the calculation of the statistic

$$Q'(H) = r\textstyle\sum_{j=1}^{c} (\bar{Z}_{.j} - \bar{Z}_{..})^2,$$

where $\bar{Z}_{.j} = r^{-1}\sum_{i=1}^{r} Z(R(X_{ij}))$ and $\bar{Z}_{..} = (rc)^{-1}\sum Z_i$. In view of the preceding results, one might suspect that this statistic has corresponding good properties. However, neither the hypothesis distribution nor any finite-sample-size power or ARE's of this statistic are known (when there is row effect).

To see that the null distribution of $Q'(H)$ depends on the row effect, consider the extreme cases when (a) there is no row effect, and (b) the row effect is such that all the X's in the ith row are smaller than the X's in the $(i + 1)$st row for $i = 1, 2, \cdots, (r - 1)$.

This leads one to employ an idea used by Friedman [7], Mood [14], Durbin [4] and others to eliminate row effect. The new procedure consists of

(1) getting r independent ordered random samples each of size c, i.e., $Z_1(1), \cdots, Z_1(c); Z_2(1), \cdots, Z_2(c); \cdots; Z_r(1), \cdots, Z_r(c)$.

(2) Ranking the X's in each row to obtain the ranks $R_1(X_{11}), \cdots, R_1(X_{1c}); R_2(X_{21}), \cdots, R_2(X_{2c}); \cdots; R_r(X_{r1}), \cdots, R_r(X_{rc})$.

(3) Replacing each X_{ij} by $Z_i(R_i(X_{ij}))$, i.e., that Z value of the ith sample of Z's that has the same row-rank as X_{ij}.

(4) Rejecting the hypothesis for large values of the statistic $Q(H) = r\sum_{j=1}^{c} (\bar{Z}_{\cdot j} - \bar{Z}_{\cdot\cdot})^2$ where $\bar{Z}_{\cdot j} = r^{-1}\sum_{i=1}^{r} Z_i(R_i(X_{ij}))$ and $\bar{Z}_{\cdot\cdot} = (rc)^{-1}\sum_{i,j} Z_i(j)$. For this statistic one obtains from the basic Lemma 2.1

THEOREM 4.1. *Under* H_0,

(i) *If* $H = \Phi$, *the standard normal* cpf, *then* $Q(\Phi)$ *has an exact chi-square distribution with* $(c - 1)$ *degrees of freedom,*

(ii) *if the variance* $\sigma_H^2 = \text{Var}(Z \mid H)$ *exists, then* $Q(H)$ *is asymptotically a chi-square variate with* $(c - 1)$ *degrees of freedom,*

(iii) *in all cases* $Q(H)$ *has the distribution of* $r\sum_{j=1}^{c} (\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot})^2$ *where* $\{Y_{ij}\}$ *are* r *independent random samples from populations with* cpf's H.

PROOF. (iii) follows at once by applying Lemma 2.1 to each row. (i) and (ii) follow from (iii).

Note that the above results are valid not only if the rows differ in location but also if they differ in scale, i.e., if $\text{Var}(X_{ik}) \neq \text{Var}(X_{jk})$ for $i \neq j$.

On comparing $Q(H)$ with its rank counterpart $D(H)$ obtained from $Q(H)$ by replacing $Z_i(R_i(X_{ij}))$ by $E(Z_i(R_i(X_{ij})) \mid H)$, one finds that $D(H)$ and $Q(H)$ are not asymptotically equivalent in general. The rank statistics used in [7], [14] and [5] are essentially obtained from $D(H)$ by letting $H = U$, the uniform cpf. In this case one can compute (on letting the number of columns c be fixed and the number r of rows tend to infinity) $A(Q(U), D(U)) = c/(c + 1)$ for alternatives such that this efficiency exists. A similar result holds in general, thus the ARE is in favor of $D(H)$ when the number of rows is "much larger" than the number of columns.

5. Bivariate independence tests. Continuing the pattern of the preceding sections, one considers randomized tests for the following bivariate problem. $(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ is a random sample from a bivariate population with continuous cpf F. X and Y have the continuous marginal cpf's F_1 and F_2, respectively. The hypothesis tested is $H_0 : F(x, y) = F_1(x)F_2(y)$ for all x, y.

The usual rank-sum statistic for this hypothesis is the rank correlation coefficient τ, which is equivalent to $d = \sum_{i=1}^{n} R(X_i)R(Y_i)$ where $R(X_i)$ denotes the rank of X_i among $X_1, X_2, \cdots, X_n$ and $R(Y_i)$ denotes the rank of Y_i among $Y_1, Y_2, \cdots, Y_n$. Kendall and Stuart ([12], p. 486), suggest the statistic $c = \sum R(X_i)E(Z(R(Y_i)) \mid \Phi)$ where $E(Z(j) \mid \Phi)$ is the expected value of the jth order statistic in a random sample of size n from a standard normal distribution.

The statistic considered here is $W(H) = n^{-1}\sum Z(R(X_i))Z'(R(Y_i))$ where

$Z(1), Z(2), \cdots, Z(n)$ and $Z'(1), Z'(2), \cdots, Z'(n)$ are two independent ordered random samples of sizes n from a population with cpf H.

This randomized statistic has the following desirable properties under independence.

THEOREM 5.1. *Under* H_0

(i) *if* $H = \Phi$, *the standard normal* cpf, *then the distribution of* $2nW(\Phi)$ *is the same as the distribution of the difference of two independent chi-square variables each with n degrees of freedom.*

(ii) *If the moment* $\mu_2(H) = \int x^2 \, dH(x)$ *exists and if we write* $\mu_H = \int x \, dH(x)$, *then* $n^{\frac{1}{2}}(\mu_2^2(H) - \mu_H^4)^{-\frac{1}{2}} [W(H) - \mu_H^2]$ *has asymptotically the distribution of a standard normal variable.*

(iii) *In all cases,* $W(H)$ *has exactly the distribution of* $n^{-1}\sum_{i=1}^{n} Z_i Z_i'$ *where*

TABLE 5.1

γ	Exact	Asympt.	Exact	Asympt.	Exact	Asympt.
	n					
	2	2	4	4	6	6
1.282	.9184	.90	.9123	.90	.9092	.90
1.645	.9512	.95	.9507	.95	.9507	.95

$Z_1, Z_2, \cdots, Z_n$ *and* $Z_1', Z_2', \cdots, Z_n'$ *are two independent random samples from populations with* cpf H.

PROOF. (iii) follows directly by applying Lemma 2.1 separately to the sample of X's and Y's and by using the fact that $Z(R(X_i))$ and $Z'(R(Y_i))$ are independent whenever X_i and Y_i are. (ii) follows from (iii) and the Central Limit Theorem. (i) follows by writing $2nW(\Phi) = \sum_{i=1}^{n} \{[Z(R(X_i)) + Z'(R(Y_i))]/2^{\frac{1}{2}}\}^2 - \{[Z(R(X_i)) - Z'(R(Y_i))]/2^{\frac{1}{2}}\}^2$ and using (iii).

The rapidity of convergence of the distribution of $n^{\frac{1}{2}}W(\Phi)$ to a standard normal distribution is indicated by the coincidence of the three first moments of $n^{\frac{1}{2}}W(\Phi)$ with that of a standard normal variable, and by the difference of the fourth moments being $6/n$. Moreover, consider Table 5.1 giving the exact and asymptotic values of $P(n^{\frac{1}{2}}W(\Phi) \leqq \gamma)$ for $\gamma = 1.282$, $\gamma = 1.645$, $n = 2$, 4 and 6. The table was obtained by computing the cpf of $W(\Phi)$ from its characteristic function.

Kendall and Stuart ([12], p. 486), consider a rank test of independence to be optimum if it is obtained from the usual correlation coefficient r by replacing observations by functions of the ranks that are asymptotically perfectly correlated with the observations they replace, whenever the observations are normally distributed. The following result is therefore of interest.

LEMMA 5.2. *If* X_i *is normally distributed, then* X_i *and* $Z(R(X_i))$ *are asymptotically perfectly correlated. A similar result holds for* Y_i *and* $Z(R(Y_i))$.

PROOF. Without loss of generality, let X_i have a standard normal distribution. Let $X(1), X(2), \cdots, X(n)$ be the order statistics of $X_1, X_2, \cdots, X_n$, then

$$\rho(X_i, Z(R(X_i))) = E(X_i Z(R(X_i)))$$
$$= \sum_{j=1}^{n} E(X(j)Z(j) \mid R(X_i) = j)P(R(X_i) = j)$$
$$= n^{-1}\sum E^2(X(j)) \to \int x^2\, d\Phi(x) \quad \text{as} \quad n \to \infty,$$

where the last step follows from Hoeffding [11].

It follows that if we replace (X_i, Y_i) in r by $(Z(R(X_i)), Z(R(Y_i)))$ we obtain an optimum rank test in the sense of Kendall and Stuart. Since r is asymptotically equivalent to $\sum X_i Y_i$ and since a random sample of known variances and means is imposed by the procedure, the statistic $n^{-1}\sum Z(R(X_i))Z'(R(Y_i))$ may be used.

Finally note that

THEOREM 5.3.

(i) *If X_i and Y_i are normally distributed and independent, then*

$$\rho(n^{-1}\sum X_i Y_i, W(\Phi)) \to 1$$

as $n \to \infty$.

(ii) $A(r, W(\Phi)) = 1$ *for bivariate normal alternatives.*

PROOF. (i) follows at once from Theorem 5.2, (ii) follows from (i) and a result by van Eeden [20] that essentially states that if T and T' are two statistics such that one of them is asymptotically locally most powerful, then under appropriate conditions, $A(T, T') = \lim_{n\to\infty} \rho(T, T' \mid H_0)$. Here r is most powerful.

Note that in this case, optimality in the sense of Kendall and Stuart means $A(r, W(\Phi)) = 1$ for normal alternatives.

6. Randomness tests. Kendall and Stuart ([12], pp. 483–487), have noted that in some sense "good" independence tests are "good" randomness tests, because the randomness hypothesis $H_0 : F_1 = F_2 = \cdots = F_n$ is equivalent to the hypothesis of independence of values and chronological order. More precisely, the univariate sample $X_1, X_2, \cdots, X_n$ (where X_i has cpf F_i) can be represented by the bivariate sample $(1, X_1), (2, X_2), \cdots, (n, X_n)$, which, under H_0, is from a bivariate population with X_i independent of i.

Following the results of the preceding section, one considers the statistic $\sum iZ(R(X_i))$, where $Z(1), Z(2), \cdots, Z(n)$ are the usual order statistics. It turns out that this straightforward extension has small efficiency when compared to the most powerful statistic, $b = \sum iX_i - \frac{1}{2}n(n+1)\bar{X}$, for normal trend alternatives (Kendall and Stuart ([12), p. 485)).

To obtain better efficiency, one considers the statistic

$$d'(H) = n^{-1}\sum iZ(R(X_i)) - [(n+1)/2]\bar{Z},$$

which has the following desirable properties.

THEOREM 6.1. *Under* H_0,

(i) *if* $\mu_H = \int x\, dH(x)$ *and* $\sigma_H^2 = \int (x - \mu_H)^2\, dH(x)$ *exist, then* $E(d'(H)) = 0$ *and* $\mathrm{Var}\,(d'(H)) = \sigma_H^2(n^2 - 1)(12n)^{-1}$,

(ii) *if* $H = \Phi$, *the standard normal* cpf, *then* $d'(\Phi)$ *is distributed as a normal variate with mean* 0 *and variance* $(n^2 - 1)(12n)^{-1}$.

PROOF. (i) follows by straightforward calculations, (ii) follows from Lemma 2.1 after writing $d'(\Phi) = n^{-1}\sum [i - \frac{1}{2}(n + 1)]Z(R(X_i))$.

From Lemma 5.2, one obtains at once

THEOREM 6.2. *If* X_i *is normally distributed, then under* H_0, $\rho(n^{-1}b, d'(\Phi)) \rightarrow 1$ *as* $N \rightarrow \infty$.

Also, $d'(\Phi)$ is an optimum rank test of randomness in the sense of Kendall and Stuart ([12], p. 486).

REFERENCES

[1] BELL, C. B., MOSER, J. M. and THOMPSON, R. Goodness criteria for two-sample distribution-free tests. In preparation.

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

[3] CHERNOFF, HERMAN and SAVAGE, I. RICHARD (1958). Asymptotic normality and efficiency of certain nonparametric tests. *Ann. Math. Statist.* **29** 972–994.

[4] DURBIN, J. (1951). Incomplete blocks in ranking experiments. *British J. Psych.* **4** 85–90.

[5] DWASS, M. (1953). On the asymptotic normality of some statistics used in non-parametric tests. *Ann. Math. Statist.* **26** 334–339.

[6] FISHER, R. A. and YATES, F. (1938). *Statistical Tables for Biological, Agricultural and Medical Research.* Oliver and Boyd, London.

[7] FRIEDMAN, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. Statist. Assoc.* **32** 675–698.

[8] HODGES, J. L., JR. and LEHMANN, E. L. (1956). The efficiency of some nonparametric competitors of the t-test. *Ann. Math. Statist.* **27** 324–335.

[9] HODGES, J. L., JR. and LEHMANN, E. L. (1961). Comparison of the normal scores and Wilcoxon tests. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1** 307–318. Univ. of California Press.

[10] HOEFFDING, W. (1950). "Optimum" nonparametric tests. *Proc. Second Berkeley Symp. Math. Statist. Prob.* Univ. of California Press.

[11] HOEFFDING, W. (1953). On the distribution of the expected values of the order statistics. *Ann. Math. Statist.* **24** 93–100.

[12] KENDALL, M. G. and STUART, A (1961) *The Advanced Theory of Statistics*, **2.** Griffin, London.

[13] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

[14] MOOD, A. M. (1950). *Introduction to the Theory of Statistics.* McGraw-Hill, New York.

[15] PURI, M. L. (1963). Asymptotic efficiency of a class of c-sample tests. *Ann. Math. Statist.* **35** 102–121.

[16] RAND CORPORATION (1955). *A Million Random Digits With 100,000 Normal Deviates.* Free Press, Glencoe.

[17] SAVAGE, I. R. (1956). Contributions to the theory of rank order statistics. Two sample case. *Ann. Math. Statist.* **27** 590–616.

[18] STUART, ALAN (1954). Asymptotic relative efficiencies of distribution-free tests of randomness against normal alternatives. *J. Amer. Statist. Assoc.* **49** 147–157.

[19] TERRY, M. E. (1952). Some rank order tests which are most powerful against specific parametric alternatives. *Ann. Math. Statist.* **23** 346–366.

[20] VAN EEDEN, C. (1963). The relation between Pitman's asymptotic relative efficiency of two tests and the correlation coefficient between their test statistics. *Ann. Math. Statist.* **34** 1442–1451.

[21] WILCOXON, F. (1945). Individual comparisons by ranking methods. *Biometrics* **1** 80–83.