RANK SUM TESTS OF FIT

By Chia Kuei Tsao

Wayne University

Summary. This paper suggests several 'goodness of fit' test criteria, all having a linear form. The moment generating function and the limiting distribution of this linear form are obtained in Section 2. The best test criterion of this form for testing a simple hypothesis H_0 against a simple alternative hypothesis H_1 is shown, in Section 3, to be in general not independent of H_1.

The remainder of this paper deals with a special case of the linear form, that is, the rank sum test criterion. The distribution of this test criterion is derived in Section 4, its consistency is proved in Section 5, and some numerical asymptotic efficiencies are calculated in Section 6. Within a certain class of tests, the present test is shown, in Section 7, to be uniformly most powerful for a special family of alternatives.

1. Introduction. To test whether a sample $X_1, X_2, \cdots, X_m$ of a random variable X was obtained from a population having a completely specified continuous cdf $F(x)$, a number of tests (e.g., [8], [9]) have been based on a procedure by which the domain of the variable is divided into k sets of sizes so determined that the probability of each set under the null hypothesis is equal to $1/k$.

Evidently, for any given k, there are many possible ways of dividing the domain of the variable into sets with equal probabilities under the null hypothesis. For any given alternative hypothesis and given test procedure, one division might be better than another. On the other hand, for any given alternative and given division of the domain, one test might be better than another. We shall concern ourselves mainly with the latter problem. We assume that some knowledge of the alternative hypothesis is available.

The division used in this paper is as follows. Suppose (i) $f_0(x)$ and $f_1(x)$ are two completely specified continuous pdf's over a space R (which is either R_n or a subspace of R_n), and (ii) for every real number c the probability of the set $\{x;\ f_1(x)/f_0(x) = c\}$ is zero when the distribution is $f_0(x)$. Let H_i denote the hypothesis that the sample was obtained from $f_i(x)$, for $i = 0, 1$. Then R is to be divided into k disjoint sets $S_1, S_2, \cdots, S_k$ such that

$$S_j = \{x;\ c_{j-1} > f_1(x)/f_0(x) \geqq c_j\}, \qquad j = 1, 2, \cdots, k, \tag{1.1}$$

where the c_j, with $\infty = c_0 > c_1 > \cdots > c_k = 0$, are so determined that

$$p_{01} = p_{02} = \cdots = p_{0k} = 1/k, \tag{1.2}$$

$$p_{ij} = \int_{S_j} f_i(x)\, dx, \qquad j = 1, 2, \cdots, k; \quad i = 0, 1. \tag{1.3}$$

Received July 27, 1953, revised August 16, 1954.

Denote by m_j the number of the sample values $X_1, X_2, \cdots, X_m$ falling in the set S_j for $j = 1, 2, \cdots, k$, and $\sum m_j = m$. Several tests based on a linear form of $m_1, m_2, \cdots, m_k$ will be proposed and their properties studied.

2. Distribution of a general linear form and its asymptotic normality. In this section we shall derive the moment generating function (mgf) of the linear form

$$L = \sum_{j=1}^{k} a_j m_j, \tag{2.1}$$

where $a_1, a_2, \cdots a_k$ are real constants (not all equal). We shall further show that its asymptotic distribution is normal.

THEOREM 1. *Under the hypothesis* H_i for $i = 0, 1$, *the* mgf *of* L *is given by*

$$M_i(t) = \left[\sum_{j=1}^{k} p_{ij} \exp(a_j t)\right]^m. \tag{2.2}$$

PROOF. Under hypothesis H_i, the distribution of $m_1, m_2, \cdots, m_k$ is known to be given by the multinomial distribution

$$w_i(m_1, m_2, \cdots, m_k) = m! \prod_{j=1}^{k} (p_{ij})^{m_j}/m_j!. \tag{2.3}$$

Hence, by application of the definition, the mgf of L in (2.1), under H_i, is given by $M_i(t)$ in (2.2).

THEOREM 2. *Under the hypothesis* H_i for $i = 0, 1$, *as* $m \to \infty$ *the distribution of* L *in* (2.1) *approaches the normal distribution with mean and variance*

$$\mu_i = m \sum_{j=1}^{k} a_j p_{ij}, \qquad \sigma_i^2 = m\left[\sum_{j=1}^{k} a_j^2 p_{ij} - \left(\sum_{j=1}^{k} a_j p_{ij}\right)^2\right]. \tag{2.4}$$

PROOF. Let $Y_1, Y_2, \cdots, Y_m$ be independent and identically distributed random variables such that $\Pr(Y_1 = a_j) = p_{ij}$ for $j = 1, 2, \cdots, k$ and $i = 0, 1$. Then L is distributed as $Y_1 + Y_2 + \cdots + Y_m$. Hence, by the central limit theorem, Theorem 2 is proved.

3. Most powerful test for testing H_0 against H_1. The main purpose of this section is to show that among the tests which depend only on $m_1, m_2, \cdots, m_k$, the most powerful test criterion for testing the simple null hypothesis H_0 against the simple alternative hypothesis H_1 is a linear function of $m_1, m_2, \cdots, m_k$, with coefficients a_j as functions of p_{1j} for $j = 1, 2, \cdots, k$.

THEOREM 3. *The best critical region for testing* H_0: $f(x) = f_0(x)$ *against* H_1: $f(x) = f_1(x)$ *is given by the subset of* $m_1, m_2, \cdots, m_k$:

$$\sum_{j=1}^{k} m_j \log(1/p_{1j}) \leqq c, \tag{3.1}$$

where c *is so determined that the level of significance is a preassigned size* α.

PROOF. By Neyman-Pearson's lemma and (2.3), it is easily seen that the best critical region for testing H_0 against H_1 is given by those sets of $m_1, m_2, \cdots, m_k$ where

$$\frac{w_1(m_1, m_2, \cdots, m_k)}{w_0(m_1, m_2, \cdots, m_k)} = k^m \prod_{j=1}^{k} (p_{1j})^{m_j} \geqq c' \tag{3.2}$$

for some constant c'. Taking logarithms of both sides, we obtain the equivalent best critical region determined by (3.1).

It is seen from Theorem 3 that in order to find the best critical region, one has to find first the distribution of the linear function on the left side of (3.1). For large samples, the distribution can be approximated by the normal distribution, according to Theorem 2. For small samples, however, the determination of the distribution is involved.

4. Distribution of the rank sum criterion. In cases where the alternative hypothesis is not specified, the statistic on the left side of (3.1) is not applicable, since it involves p_{1j}'s. Furthermore, if the alternative hypothesis is stated in the composite form, no uniformly most powerful test can be found for most of the known distributions.

However, a class of alternatives may be such that they give rise to the same division of R (i.e., same sets $S_1, S_2, \cdots, S_k$). Then a rank sum test criterion as defined in (4.1) below seems to be a reasonable one to use, since in this case all the most powerful test criteria have the form (2.1) with the same property: $a_1 \leqq a_2 \leqq \cdots \leqq a_k$.

The word "rank" is here used in the sense that an observation falling in the set S_j is given the rank j, for $j = 1, 2, \cdots, k$. This is different from the usual usage of the word in two-sample problems. In the latter case, ranks of observations are determined by their relative positions, while here, ranks are determined by a given division of the space R.

THEOREM 4. *Let*

$$s = \sum_{j=1}^{k} j m_j. \tag{4.1}$$

Then, under H_0, *the* pdf *of* s *may be written as*

$$g(s; k, m) = k^{-m} \sum_{r=0}^{m} (-1)^r \binom{m}{r} \binom{s - 1 - rk}{m - 1}, \quad s = m, m + 1, \cdots, km. \tag{4.2}$$

Furthermore, $g(s; k, m)$ *is symmetrical,* that is,

$$g(m + v; k, m) = g(km - v; k, m), \quad v = 0, 1, \cdots, (k - 1)m. \tag{4.3}$$

PROOF. The pdf $g(s; k, m)$ is the coefficient of t^s in the power expansion of the distribution generating function

$$D(t; k, m) = k^{-m} \left(\sum_{j=1}^{k} t^j \right)^m. \tag{4.4}$$

This may be written as

$$\begin{aligned} D(t; k, m) &= k^{-m}t^m(1 - t^k)^m(1 - t)^{-m} \\ &= k^{-m} \sum_{p=0}^{m} \sum_{q=0}^{\infty} (-1)^{m-p} \binom{m}{p} \binom{m - 1 + q}{m - 1} t^{m+q+k(m-p)}. \end{aligned} \tag{4.5}$$

Letting $s = m + q + k(m - p)$, we obtain the pdf

$$g(s; k, m) = k^{-m} \sum_{p=0}^{m} (-1)^{m-p} \binom{m}{p} \binom{s - 1 - k(m - p)}{m - 1}. \tag{4.6}$$

Now, setting $r = m - p$, we get the pdf (4.2).

The symmetrical property (4.3) may be seen from the following argument. Since $g(m + v; k, m)$ is the coefficient of t^{m+v} in $D(t; k, m)$, it is the coefficient of $t^{-(m+v)}$ in $D(1/t; k, m)$, and hence the coefficient of t^{km-v} in $t^{(k+1)m} D(1/t; k, m)$. Thus (4.3) is proved, since

$$t^{(k+1)m}D(1/t; k, m) = D(t; k, m).$$

5. Consistency of the rank sum test. We shall now show that for any given k the rank sum test is consistent for a specified alternative or a class of alternatives.

THEOREM 5. *Let $k \geqq 2$ be a preassigned integer. Let the simple hypothesis H_0: $f(x) = f_0(x)$ be tested against the simple alternative hypothesis H_1: $f(x) = f_1(x)$, where $f_0(x)$ and $f_1(x)$ satisfy conditions* (i) *and* (ii) *in Section* 1.

Then, for any preassigned level of significance α, the power $u(f_1)$ of the rank sum test with index k approaches unity as the sample size m becomes indefinitely large.

PROOF. By Theorem 2, the statistic s defined in (4.1), under H_i for $i = 0, 1$, is asymptotically normally distributed. Let $N(x; \mu, \sigma)$ denote the cumulative normal distribution with mean μ and standard deviation σ. Then, for large m, the level of significance and the power of the test will be given approximately by

$$\alpha = N(\mu_0 - z_\alpha\sigma_0; \quad \mu_0, \sigma_0), \tag{5.1}$$

$$u(f_1) = N[(\mu_0 - \mu_1)/\sigma_1 - z_\alpha\sigma_0/\sigma_1; \quad 0, 1], \tag{5.2}$$

where z_α depends only on α, and μ_i and σ_i^2 for $i = 0, 1$, are given by (2.4) with $a_j = j$.

Since the ratio σ_0/σ_1 is constant, then to prove Theorem 5 it is sufficient to prove that $(\mu_0 - \mu_1)/\sigma_1$ increases as m increases, for any $k \geqq 2$, or equivalently,

$$\sum_{j=1}^{k} j(1/k - p_{1j}) > 0, \qquad k \geqq 2. \tag{5.3}$$

We shall first prove that, for any $k \geqq 2$,

$$p_{11} > 1/k, \qquad p_{1k} < 1/k. \tag{5.4}$$

In what follows, we shall denote by $S_j^{(k)}$ the sets defined in (1.1) and by $p_{ij}^{(k)}$ the quantities defined in (1.3), that is,

$$S_j^{(k)} = S_j, \qquad p_{ij}^{(k)} = p_{ij}, \qquad j = 1, 2, \cdots, k; i = 0, 1.$$

Since, by assumption, the probability of the set $\{x; f_1(x)/f_0(x) > 1\}$ is positive when the distribution is $f_1(x)$, then by (1.1), (1.2), and (1.3), we have $p_{11}^{(2)} > \frac{1}{2}$ and $p_{12}^{(2)} < \frac{1}{2}$. Thus, the inequalities (5.4) are true for $k = 2$.

For $k > 2$, we will prove (5.4) by considering the k sets $\{S_j^{(k)}\}$, for $j = 1, 2, \cdots, k$; the $2k$ sets $\{S_j^{(2k)}\}$ for $j = 1, 2, \cdots, 2k$; and the two sets $S_1^{(2)}$ and $S_2^{(2)}$. From (1.1), (1.2), and (1.3), it is easily seen that these sets satisfy the following relations

$$S_1^{(k)} = S_1^{(2k)} \cup S_2^{(2k)}, \qquad \bigcup_{j=1}^{k} S_j^{(2k)} = S_1^{(2)}. \tag{5.5}$$

Consequently, we find

$$p_{11}^{(k)} = p_{11}^{(2k)} + p_{11}^{(2k)}, \qquad \sum_{j=1}^{k} p_{1j}^{(2k)} = p_{11}^{(2)}. \tag{5.6}$$

Since $p_{11}^{(2k)} \geqq p_{12}^{(2k)} \geqq \cdots \geqq p_{1k}^{(2k)} \cdots$, we have

$$p_{11}^{(k)} = p_{11}^{(2k)} + p_{12}^{(2k)} \geqq \frac{2}{k} \sum_{j=1}^{k} p_{1j}^{(2k)} = \frac{2p_{11}^{(2)}}{k} > \frac{1}{k}. \tag{5.7}$$

Similarly, we can prove that $p_{1k}^{(k)} < 1/k$. Hence, the inequalities (5.4) are true for any $k \geqq 2$.

Now, it follows from a special case of the Tchebycheff's inequality (when $r = 1$ in Theorem 43 of [5]) that for any $k \geqq 2$, the inequality (5.3) is true, since $\{j\}$ and $\{1/k - p_{1j}^{(k)}\}$ are similarly ordered and $\sum(1/k - p_{1j}^{(k)}) = 0$. Thus, Theorem 5 is proved.

Theorem 5 implies that a one-sided rank sum test is consistent for the case where the null hypothesis H_0: $f(x) = f_0(x)$ is simple and the alternative hypothesis H_1^* consists of distributions $\{f_1^*(x)\}$ such that the ratios $\{r(x) = f_1^*(x)/f_0(x)\}$ are monotonic increasing (or decreasing) functions of x, provided that conditions (i) and (ii) in Section 1 are satisfied for each $f_1^*(x)$, when $f_1(x)$ is replaced by $f_1^*(x)$ there. On the other hand, a two-sided rank sum test also may be shown to be consistent under the following assumptions.

Suppose it is required to test the simple hypothesis H_0: $f(x) = f_0(x)$ against the composite hypothesis H_1^{**} which consists of two and only two classes C_1 and C_2 of those distributions such that the ratio $f_1^{**}(x)/f_0(x)$ is a monotonic decreasing function of x when $f_1^{**}(x)$ is in C_1 and a monotonic increasing function of x when $f_1^{**}(x)$ is in C_2. Suppose, also, that for each $f_1^{**}(x)$, conditions (i) and (ii) are satisfied, when $f_1(x)$ is replaced by $f_1^{**}(x)$. Then, we obtain

THEOREM 6. *Let $k \geqq 2$ be a preassigned integer, and α_1, α_2, and α be three preassigned positive numbers such that $\alpha_1 + \alpha_2 = \alpha$ where $0 < \alpha < 1$. Let h_1 and h_2, with $m \leqq h_1 < h_2 \leqq km$ be the values such that*

$$\sum_{s=m}^{h_1} g(s; k, m) = \alpha_1, \qquad \sum_{s=h_2}^{km} g(s; k, m) = \alpha_2. \tag{5.8}$$

Then, the test consisting of rejecting H_0 whenever $\sum jm_j \leqq h_1$ or $\sum jm_j \geqq h_2$ is consistent for testing H_0 against H_1^{**}.

PROOF. Theorem 6 is a corollary of Theorems 4 and 5, considering the two classes C_1 and C_2 separately.

Theorem 6 implies that the two-sided rank sum test is unbiased for testing H_0 against H_1^{**} for large samples. For small samples, it may or may not be unbiased.

6. Asymptotic efficiency of the rank sum test. In this section, we shall compare the rank sum test with one of the standard parametric tests, and obtain the power efficiency for large samples.

Suppose a sample is to be drawn from a population having a normal distribution with known variance σ^2 and unknown mean θ. Let the hypothesis H_0: $\theta = \theta_0$ be tested against the alternative hypothesis H_1: $\theta = \theta_1$, where $\theta_1 > \theta_0$. Further, let both the level of significance $\alpha = u(\theta_0)$ and the power of the test $u = u(\theta_1)$ be specified in advance, for example, $\alpha = .05$ and $u = .95$.

Let M and m be the sample sizes required by the likelihood ratio test and the rank sum test, respectively. We shall call $\mathcal{E} = M/m$ the *efficiency* of the rank sum test.

When $d = (\theta_1 - \theta_0)/\sigma$ is small, so that the sample size m is large, we can use the normal approximation (5.2) for the power function of the rank sum test. Hence, m is approximately given by

$$m = \left(\frac{z_\alpha \sqrt{(k^2 - 1)/12} + z_u \sqrt{\Sigma_1^k j^2 p_{1j} - (\Sigma_1^k j p_{1j})^2}}{\frac{1}{2}(k + 1) - \Sigma_1^k j p_{1j}}\right)^2, \tag{6.1}$$

while M is given by

$$M = (z_\alpha + z_u)^2 \sigma^2 / (\theta_1 - \theta_0)^2, \tag{6.2}$$

where z_α and z_u depend only on the specified values $\alpha = u(\theta_0)$ and $u = u(\theta_1)$, respectively. Consequently, the efficiency is given approximately by

$$\mathcal{E} = \frac{(z_\alpha + z_u)^2}{d^2}\left(\frac{\frac{1}{2}(k + 1) - \Sigma_1^k j p_{1j}}{z_\alpha \sqrt{(k^2 - 1)/12} + z_u \sqrt{\Sigma_1^k j^2 p_{1j} - (\Sigma_1^k j p_{1j})^2}}\right)^2. \tag{6.3}$$

The asymptotic efficiency may be found by evaluating the limit $\lim_{d \to 0} \mathcal{E}$. Thus, letting $n(x; \theta, \sigma)$ denote the normal density with mean θ and standard deviation σ, we obtain

THEOREM 7. *Let $k \geqq 2$ be a preassigned integer. Then, for any preassigned α and u with $u > \alpha$, the asymptotic efficiency (as $d \to 0$) of the rank sum test with index k is given by*

$$\mathcal{E}' = \frac{12}{k^2 - 1}\left[\sum_{j=1}^{k-1} n(y_j ; 0, 1)\right]^2, \tag{6.4}$$

where the y_j, with $\infty = y_0 > y_1 > \cdots > y_k = -\infty$, are so determined that

$$N(y_{j-1}; 0, 1) - N(y_j; 0, 1) = 1/k, \qquad j = 1, 2, \cdots, k. \tag{6.5}$$

PROOF. It follows from the definition of p_{ij} that

$$p_{ij} = N(x_{j-1}; \theta_i, \sigma) - N(x_j; \theta_i, \sigma), \qquad j = 1, 2, \cdots, k; \quad i = 0, 1, \tag{6.6}$$

where the x_i, with $\infty = x_0 > x_1 > \cdots > x_k = -\infty$, are so determined that Equation (1.2) is satisfied. Since $\lim_{d\to 0} p_{1j} = p_{0j}$, we have

$$\mathcal{E}' = \lim_{d\to 0} \mathcal{E} = \frac{12}{k^2 - 1}\left[\lim_{d\to 0}\left(\frac{k+1}{2} - \sum_{j=1}^{k} jp_{1j}\right) \Big/ d\right]^2. \tag{6.7}$$

The limit in square brackets is an indeterminate form. However, the numerator and denominator satisfy the assumptions of l'Hospital's rule. Consequently, we obtain

$$\mathcal{E}' = \frac{12}{k^2 - 1}\left[\sum_{j=1}^{k} \sigma j\{n(x_{j-1}; \theta_0, \sigma) - n(x_j; \theta_0, \sigma)\}\right]^2. \tag{6.8}$$

From (6.5) and (6.6), it is easily seen that $y_j = (x_j - \theta_0)/\sigma$ for $j = 0, 1, \cdots, k$. Thus, the asymptotic efficiency $\mathcal{E}'$ may be written as

$$\begin{aligned} \mathcal{E}' &= \frac{12}{k^2 - 1}\left[\sum_{j=1}^{k} j\{n(y_{j-1}; 0, 1) - n(y_j; 0, 1)\}\right]^2 \\ &= \frac{12}{k^2 - 1}\left[\sum_{j=1}^{k} n(y_{j-1}; 0, 1) - kn(y_k; 0, 1)\right]^2. \end{aligned} \tag{6.9}$$

By assumption, $n(y_0; 0, 1) = n(y_k; 0, 1) = 0$. Therefore (6.9) becomes (6.4). This completes the proof of Theorem 7.

We remark that although the asymptotic efficiency $\mathcal{E}'$ was obtained under the assumption that α and u are preassigned, it is actually independent of them.

TABLE I

Power Efficiencies of the Rank Sum Tests

k \ d	0	.1	.2	.3	.4	.5
2	.637	.637	.636	.636	.635	.634
3	.793	.793	.793	.794	.794	.794
4	.856	.856	.857	.857	.858	.859
5	.888	.888	.889	.890	.891	.892
6	.906	.906	.906	.908	.909	.911
7	.918	.918	.919	.920	.921	.923
8	.926	.926	.927	.928	.929	.931
9	.931	.931	.932	.933	.934	.936
10	.935	.936	.937	.939	.940	.942

Table I gives the asymptotic efficiencies ($d = 0$) and certain approximate efficiencies ($d = .1, .2, .3, .4, .5$) for $k = 2, 3, \cdots, 10$. The approximate efficiencies were obtained under the assumption that $\alpha = 1 - u$ (where α is small).

For $0 < d \leqq .5$, the approximate efficiencies in Table I are very close to the exact ones. We note also that $.637 = 2/\pi$ is just the power efficiency of the large sample binomial test, since the rank sum test reduces to the binomial test when $k = 2$ (see also [2], [4]).

7. Power of a uniformly most powerful rank sum test. We have shown that the rank sum tests are consistent, and the numerical calculations show that their asymptotic efficiencies are approximately over 90% for the normal alternatives, when $k \geqq 6$. However, these are only the limiting behavior of the tests. Naturally, the power of the rank sum tests for small samples is also desirable. It is the purpose of this section to investigate the power of the rank sum tests for a special family of alternatives. This family is so chosen that among the tests which depend only on $m_1, m_2, \cdots, m_k$, the rank sum test is uniformly most powerful for testing the uniform distribution against such a family.

Let a family of cdf's be given by

$$F(x; A) = \begin{cases} x, & A = 1, \\ (A^x - 1) / (A - 1), & A > 1, \end{cases} \qquad 0 \leqq x \leqq 1. \tag{7.1}$$

Let the hypothesis H_0': $A = 1$ be tested against the alternative hypothesis H_1': $A > 1$. Then,

THEOREM 8. *Among the tests which depend only on $m_1, m_2, \cdots, m_k$, the rank sum tests are uniformly most powerful for testing H_0' against H_1'. The power functions are given by*

$$u(A; k, m) = \frac{k^m A^m (A^{1/k} - 1)^m}{(A - 1)^m} \sum_{s=m}^{c^*} g(s; k, m) A^{-s/k}, \tag{7.2}$$

where the level of significance is

$$\sum_{s=m}^{c^*} g(s; k, m) = \alpha. \tag{7.3}$$

PROOF. Under H_0', $F(x; A)$ is just the cdf for the uniform distribution on (0, 1). Therefore, according to the assumptions in Section 1, we have

$$S_j = \left(\frac{k-j}{k}, \frac{k-j+1}{k}\right), \qquad j = 1, 2, \cdots, k. \tag{7.4}$$

Hence

$$p_{1j} = \frac{A^{(k-j+1)/k} - A^{(k-j)/k}}{A - 1} = \frac{A^{-j/k}(A^{(k+1)/k} - A)}{A - 1}, \qquad j = 1, 2, \cdots, k. \tag{7.5}$$

Thus, by (3.1), the best critical region is given by

$$\begin{aligned} &\sum_{j=1}^{k} m_j \log \left(\frac{1}{p_{1j}}\right) \\ &\qquad = \log A^{1/k} \sum_{j=1}^{k} j m_j + m \,[\log (A - 1) - \log (A^{(k+1)/k} - A)] \leqq c. \end{aligned} \tag{7.6}$$

Since $A > 1$ and the second term is constant, we have the equivalent best critical region determined by

$$\sum_{j=1}^{k} j m_j \leqq c^*. \tag{7.7}$$

It is obvious that for any given level of significance α the best critical region determined by (7.7) is independent of the parameter A. Consequently, the rank sum test is uniformly most powerful for all $A > 1$. Thus, we have proved the first assertion of Theorem 8.

The power function of the test is derived as follows:

$$
\begin{aligned}
u(A; k, m) &= \sum m! \prod_{j=1}^{k} \frac{(p_{1j})^{m_j}}{m_j!} \\
&= \sum m! \prod_{j=1}^{k} \frac{A^{-jm_j/k}(A^{(k+1)/k} - A)^{m_j}}{(A-1)^{m_j} m_j!} \\
&= \frac{(A^{(k+1)/k} - A)^m}{(A-1)^m} \sum m! \prod_{j=1}^{k} \frac{A^{-jm_j/k}}{m_j!} \qquad (7.8) \\
&= \frac{(A^{(k+1)/k} - A)^m k^m}{(A-1)^m} \sum m! \prod_{j=1}^{k} \frac{A^{-jm_j/k}}{k^m m_j!} \\
&= \frac{A^m (A^{1/k} - 1)^m k^m}{(A-1)^m} \sum_{s=m}^{c^*} g(s; k, m) A^{-s/k},
\end{aligned}
$$

where the summation $\sum$ is extended over all possible combinations of $m_1, m_2, \cdots, m_k$ such that $m \leqq \sum j m_j \leqq c^*$ and $\sum m_j = m$. This completes the proof of Theorem 8.

In Figure 1 are plotted six power curves $u(A; k, m)$ for $k = 2, 3, 6$ and $m = 4, 6$, comparing the power curves of the rank sum tests with indices $k = 3, 6$ with that of the binomial tests; when $k = 2$, the rank sum test reduces to the binomial test. The levels of significance (nonrandomized) are

$$
\begin{aligned}
&u(1; 2, 4) = .0625, \quad u(1; 3, 4) = .0617, \quad u(1; 6, 4) = .0540, \\
&u(1; 2, 6) = .1094, \quad u(1; 3, 6) = .1070, \quad u(1; 6, 6) = .0965.
\end{aligned}
$$

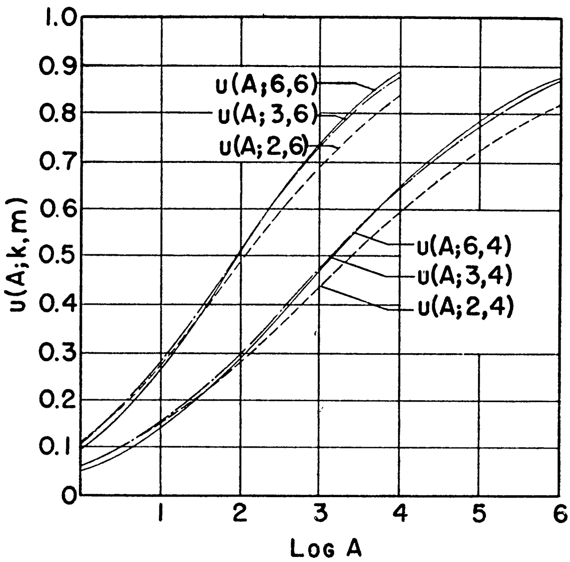

Fig. 1

The family of distributions $F(x;\ A)$ not only yields nice power functions for the rank sum tests, but also possesses the nice property that A plays the same role as a location parameter in the sense that the mean $\mu = A/(A-1) - 1/\log A$ is a monotonic increasing function of A. For, letting $y = \log A$, we have

$$\mu = e^y/(e^y - 1) - 1/y, \tag{7.9}$$

and hence

$$\frac{d\mu}{dy} = \frac{(e^y-1)^2 - y^2e^y}{y^2(e^y-1)^2} = \frac{(e^y - 1 + ye^{y/2})(e^y - 1 - ye^{y/2})}{y^2(e^y-1)^2}. \tag{7.10}$$

Now, when $y > 0$, we have

$$e^y - 1 - ye^{y/2} = \left(\frac{1}{3!} - \frac{1}{2^2 2!}\right)y^3 + \left(\frac{1}{4!} - \frac{1}{2^3 3!}\right)y^4 + \cdots > 0. \tag{7.11}$$

Consequently, μ is a monotonic increasing function of y, for $y > 0$, and hence of A, for $A > 1$.

Since A plays the role of a location parameter, the class of alternatives $F(x;A)$ may be regarded as representative of the principal types of deviation from the null hypothesis. To show the general shape we plot, in Figure 2, five curves of this family along with five curves of normal alternatives $N(x;\ \theta,\ \sigma)$, where the normal distributions are transformed by $\xi = N(x;\ \theta_0,\ \sigma)$. The family $F(x;\ A)$ is shown by solid curves, while the family $N(x;\ \theta,\ \sigma)$ is shown by dotted curves. In the latter case, the letter d represents the distance of the mean θ from the hypothetical mean θ_0 in terms of the standard deviation σ, that is, $d = (\theta - \theta_0)/\sigma$.

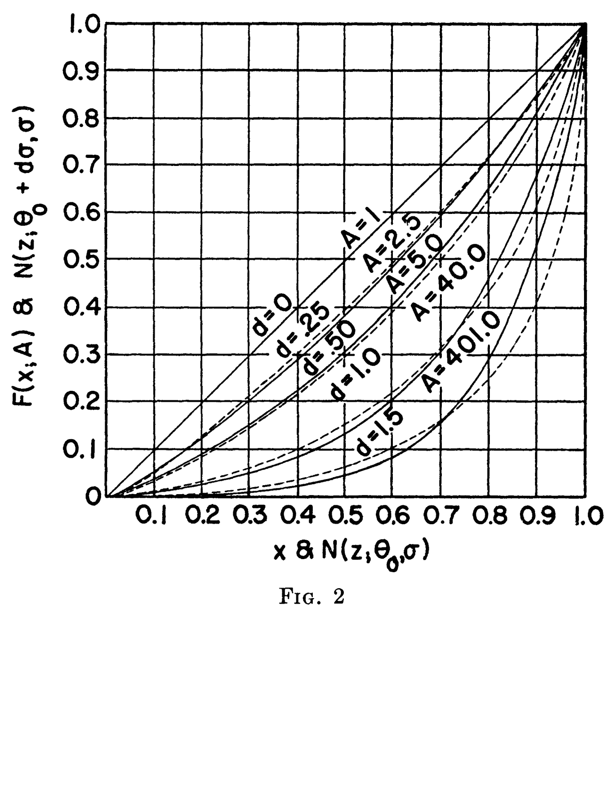

FIG. 2

Figure 2 shows that certain normal alternatives (when d is small) can be approximated by certain alternatives $F(x; A)$. Therefore, for a fixed level of significance their corresponding power curves can be used as approximations to each other.

The author wishes to express his appreciation to the referee for valuable comments and suggestions.

REFERENCES

[1] Z. W. Birnbaum, "Distribution-free tests of fit for continuous distribution functions," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 1–8.

[2] W. G. Cochran, "The efficiencies of the binomial test of significance of a mean and of a correlation coefficient," *J. Roy. Stat. Soc.*, Vol. 100 (1937), pp. 69–73.

[3] W. G. Cochran, "The χ^2 test of goodness of fit," *Ann. Math. Stat.*, Vol. 23 (1952), pp. 315–345.

[4] W. J. Dixon and A. M. Mood, "The statistical sign test," *J. Amer. Stat. Assn.*, Vol. 41 (1946), pp. 557–566.

[5] G. H. Hardy, J. E. Littlewood and G. Polya, *Inequalities*, Cambridge University Press, 1934.

[6] W. Hoeffding, "Optimum nonparametric tests," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1951, pp. 83–92.

[7] E. L. Lehmann, "Consistency and unbiasedness of certain nonparametric tests," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 165–179.

[8] A. B. Mann and A. Wald, "On the choice of the number of intervals in the application of the chi square test," *Ann. Math. Stat.*, Vol. 13 (1942), pp. 306–317.

[9] F. J. Massey, "A note on the estimation of a distribution function by confidence limits," *Ann. Math. Stat.*, Vol. 21 (1950), pp. 116–119.

[10] A. M. Mood, *Introduction to the Theory of Statistics*, McGraw-Hill Book Co., 1950.

[11] C. A. Williams, Jr., "On the choice of the number and width of classes for the chi square test of goodness of fit," *J. Amer. Stat. Assn.*, Vol. 45 (1950), pp. 77–86.