The Annals of Statistics
1977, Vol. 5, No. 1, 110-123

LARGE SAMPLE THEORY FOR *U*-STATISTICS AND TESTS OF FIT

BY GAVIN G. GREGORY[1]

University of Arkansas

Let X_{ni}, $i = 1, \cdots, n$ be i.i.d. random variables on an arbitrary measurable space $(\mathscr{X}, B)$. Suppose $\mathscr{L}(X_{ni}) = Q_{n1}$, $i = 1, \cdots, n$ and let P_0 be a fixed probability measure on $(\mathscr{X}, B)$. We consider limiting distribution theory for U-statistics $T_n = n^{-1} \sum_{i \neq j} Q(X_{ni}, X_{nj})$ (1) under conditions which imply the product measures $Q_n = Q_{n1} \times \cdots \times Q_{n1}$, n times, are contiguous to the product measures $P_n = P_0 \times \cdots \times P_0$, n times, and (2) for kernels Q which are symmetric, square-integrable ($\int Q^2(\cdot, \cdot)\, dP_0 \times P_0 < \infty$) and degenerate in a certain sense ($\int Q(\cdot, t) P_0(dt) = 0$ a.e. (P_0)). Applications to chi-square and Cramér–von Mises tests for a simple hypothesis and Cramér–von Mises tests for the case when parameters have to be estimated, are given. A tail sensitive test for normality is introduced.

1. Introduction. Under nonrestrictive assumptions, Hoeffding (1948a) showed limiting normality for U-statistics in what may be called the nondegenerate case, when the variance of the limiting normal distribution is positive. When the variance is zero, we say the degenerate case obtains and nonnormal limiting distributions result. An example of the latter is given in Hoeffding (1948b) for a kernel of degree five. In this paper asymptotic distribution theory is presented under nonrestrictive assumptions for the degenerate case when the kernel is of degree two.

Our main result here is Theorem 2.1 which treats U-statistics $T_n = n^{-1} \sum_{i \neq j} Q(X_{ni}, X_{nj})$, where X_{ni}, $i = 1, \cdots, n$ are i.i.d. random variables on an arbitrary measurable space and Q is a symmetric kernel degenerate in the sense of (2.2). Under the conditions of the theorem the distributions of X_{n1}, $n = 1, 2, \cdots$ need not be identical but are contiguous to a fixed distribution.

The applications to simple hypotheses considered in Section 3 concern chi-square and Cramér–von Mises test statistics that have what is called the usual form $T_n^* = n^{-1} \sum_{i,j} Q(X_{ni}, X_{nj})$. For certain weight functions allowing sensitive comparisons in the tails of distributions the Cramér–von Mises statistic of form T_n^* will have a limiting distribution if centering is effected by subtracting constants μ_n tending to infinity with n. It is better to use the naturally centered statistic T_n which produces an asymptotically equivalent test. (Such tests are also called Cramér–von Mises tests.)

In Section 4 we deal with Cramér–von Mises tests when parameters have to

Received June 1975; revised May 1976.

[1] Now at University of Texas at El Paso.

AMS 1970 *subject classifications*. Primary 62E15, 62E20; Secondary 62G10.

Key words and phrases. Limiting distributions, U-statistics, chi-square test, Cramér–von Mises test.

be estimated. The results are illustrated in Section 5 with a tail sensitive test for normality. Section 6 contains a discussion on the comparison of tests and Section 7 contains the large sample results for uniform variables used in Section 4.

From the work on differentiable statistical functions of von Mises (1947) and Filippova (1961) one sees that a very large class of statistics will be asymptotically of the form T_n^* with a degenerate kernel Q. This work has parallels with the well-known stochastic process approach to limiting distributions (see Anderson and Darling (1952) for such a treatment of Cramér–von Mises statistics). Thus our Theorem 2.3 which treats statistics of the form T_n^* will have wide application beyond the examples we consider.

For related problems in large sample distribution theory see Varberg (1966) and Schach (1969). Pertinent references appear during the development.

2. The limiting distribution of U-statistics under contiguous alternatives. For each $n = 1, 2, \cdots$ let X_{ni}, $i = 1, \cdots, n$ be random variables which are independent under the null hypothesis H_0: $\mathscr{L}(X_{ni}) = P_0$, $i = 1, \cdots, n$ for some fixed probability measure P_0 on the measurable space $(\mathscr{X}, B)$, as well as under the alternatives H_1: $\mathscr{L}(X_{ni}) = Q_{n1}$, $i = 1, \cdots, n$. Our conditions will imply the sequence $\{Q_n\}$ of product measures $Q_n = Q_{n1} \times \cdots \times Q_{n1}$, n times, is contiguous (in the sense of Hájek and Sidák (1967), page 202) to the sequence $\{P_n\}$ of product measures $P_n = P_0 \times \cdots \times P_0$, n times. The concern of this section is the limiting distribution of U-statistics,

$$T_n = n^{-1} \textstyle\sum_{i \neq j} Q(X_{ni}, X_{nj}) \tag{2.1}$$

where $Q(s, t)$, $s, t \in \mathscr{X}$, is a symmetric, nonzero kernel on $\mathscr{X} \times \mathscr{X}$ with $\int Q^2(\cdot, \cdot)\, dP_0 \times P_0 < \infty$. In this paper only the degenerate case

$$\textstyle\int Q(\cdot, t) P_0(dt) = 0 \qquad \text{a.e.} \quad (P_0) \tag{2.2}$$

is considered. Let $\{\lambda_k, k \geqq 0\}$ denote the finite or infinite collection of eigenvalues of Q corresponding to orthonormal eigenfunctions $\{f_k, k \geqq 0\}$, i.e., for all k and j $\int Q(\cdot, t) f_k(t) P_0(dt) = \lambda_k f_k$ a.e. (P_0), $\int f_k f_j\, dP_0 = 0$ if $k \neq j$, and $\int f_k^2\, dP_0 = 1$. In view of (2.2) we may let $f_0 \equiv 1$ correspond to the eigenvalue $\lambda_0 = 0$. Our main result of this section is

THEOREM 2.1. *Let Q_{n1} be dominated by P_0 with Radon–Nikodym derivative $dQ_{n1}/dP_0 = 1 + n^{-\frac{1}{2}} h_n$, for some sequence $\{h_n\}$ in $L_2 = L_2(\mathscr{X}, B, P_0)$ converging to $h \in L_2$, say. Then*

$$\lim_n Q_n\{T_n \leqq x\} = P\{\textstyle\sum \lambda_k[(Z_k + a_k)^2 - 1] \leqq x\} \tag{2.3}$$

where $a_k = \int h f_k\, dP_0$ and $Z_1, Z_2, \cdots$ are i.i.d. *standard normal variables.*

COROLLARY 2.2. *Let $(\mathscr{X}, B)$ be the unit interval $(0, 1)$ with the Borel σ-algebra and P_0 the Lebesgue measure on $(0, 1)$. Let $dQ_{n1}/dP_0 = f_{\theta_n}$ where $\theta_n = bn^{-\frac{1}{2}}$ and $f_\theta(u)$, $0 < u < 1$, is a density in u for $|\theta| < \Theta$, $\Theta > 0$, with f_0 the uniform density on $(0, 1)$. Suppose*

A1. *For all u, $f_\theta(u)$ is absolutely continuous in θ.*

A2. *For all θ, $(\partial/\partial\theta)f_\theta(u)$ exists for almost all (Lebesgue measure) u.*

A3. *The Fisher information*

$$I(\theta) = \int_0^1 \{\partial/\partial\theta)f_\theta(u)\}^2 f_\theta(u)\,du$$

exists for all θ, is continuous at $\theta = 0$ and $I(0) > 0$.

Then (2.3) *is true with $a_k = b \int hf_k\,dP_0$ and $h(u) = (\partial/\partial\theta)\log f_\theta(u)|_{\theta=0}$, $0 < u < 1$.*

A companion result to Theorem 2.1 holds under a further restriction on the eigenvalues.

THEOREM 2.3. *If in addition to the conditions of Theorem* 2.1, $\sum \lambda_k$ *is finite then the modification of T_n*(2.1) *to include the terms for $i = j$, $T_n^* = T_n + n^{-1}\sum_i Q(X_{ni}, X_{ni})$, has an asymptotic distribution given by*

$$\lim_n Q_n\{T_n^* \leqq x\} = P\{\sum \lambda_k(Z_k + a_k)^2 \leqq x\}\,. \tag{2.4}$$

PROOF OF THEOREM 2.1. Let x be fixed throughout. Let

$$Q^K(s, t) = \sum_{k=1}^K \lambda_k f_k(s)f_k(t)$$

and

$$T_{n,K} = n^{-1}\sum_{i\neq j} Q^K(X_{ni}, X_{nj}) = \sum_{k=1}^K \lambda_k(\zeta_{nk}^2 - c_{nk})$$

with $\zeta_{nk} = n^{-\frac{1}{2}}\sum_{i=1}^n f_k(X_{ni})$, $c_{nk} = n^{-1}\sum_{i=1}^n f_k^2(X_{ni})$. It is well known that under the conditions of the theorem the sequences $\{Q_n\}$ are contiguous to $\{P_n\}$. From Behnen and Neuhaus (1975) it follows that for fixed k, there is a sequence f_{nk}, $n = 1, 2, \cdots$ in L_2 with $\int f_{nk}\,dP_0 = 0$ and $\lim_n \int (f_{nk} - f_k)^2\,dP_0 = 0$ such that $\bar{\zeta}_{nk} - \zeta_{nk} \to 0$ in probability ($n \to \infty$) under $\{Q_n\}$ and $\bar{\zeta}_{nk} - \int f_{nk}h_n\,dP_0$ satisfies the Lindeberg condition, where $\bar{\zeta}_{nk} = n^{-\frac{1}{2}}\sum_{i=1}^n f_{nk}(X_{ni})$. From the Cramér–Wold device and the fact that the resulting linear combinations satisfy the Lindeberg condition since the components do, it follows that under Q_n probability $(\zeta_{n1}, \cdots, \zeta_{nK})$ converges in distribution to $(Z_1 + a_1, \cdots, Z_K + a_K)$. Since $c_{nk} \to_n 1$ in Q_n probability, $k = 1, \cdots, K$, it follows that

$$D(n, K) = |Q_n\{T_{n,K} \leqq x\} - P\{\sum_{k=1}^K \lambda_k[(Z_k + a_k)^2 - 1] \leqq x\}| \to 0 \tag{2.5}$$

as $n \to \infty$, K fixed.

Furthermore we have for $\varepsilon > 0$

$$\begin{aligned} P\{\textstyle\sum_{k=1}^K \lambda_k[(Z_k + a_k)^2 - 1]\varepsilon(x, x' + \varepsilon]\} \\ \leqq \sup_x P\{\lambda_1[(Z_1 + a_1)^2 - 1]\varepsilon(x, x + \varepsilon]\} \equiv E(\varepsilon)\,. \end{aligned} \tag{2.6}$$

Clearly $E(\varepsilon) \to 0$ for $\varepsilon \to 0$.

Now for every integer K and $\varepsilon > 0$

$$\begin{aligned} Q_n\{T_n \leqq x\} \leqq P\{\textstyle\sum_{k=1}^K \lambda_k[(Z_k + a_k)^2 - 1] \leqq x\} \\ + Q_n\{|T_n - T_{n,K}| \geqq \varepsilon\} + D(n, K) + E(\varepsilon)\,. \end{aligned} \tag{2.7}$$

From $P_n\{|T_n - T_{n,K}| \geqq \varepsilon\} \leqq \varepsilon^{-2}\int (T_n - T_{n,K})^2\,dP_n = \varepsilon^{-2}(n-1)n^{-1}||Q - Q^K||^2$, with $||Q - Q^K||^2 = \int (Q - Q^K)^2\,dP_0 \times P_0$ and $\lim_{K\to\infty} ||Q - Q^K|| = 0$ (see, e.g., Dunford and Schwartz (1963), page 1087), we have

$$\lim_n Q_n\{|T_n - T_{n,K(n)}| \geqq \varepsilon\} = 0 \quad \text{for all } \varepsilon > 0 \text{ and every sequence } \{K(n)\} \text{ of integers with } K(n) \to \infty . \tag{2.8}$$

Now (2.5) shows there exists such a sequence $\{K(n)\}$ with

$$\lim_n D(n, K(n)) = 0 . \tag{2.9}$$

If we show the mean square convergence of $T_{\infty,K} = \sum_{k=1}^{K} \lambda_k[(Z_k + a_k)^2 - 1]$ for $K \to \infty$ then (2.6), (2.7), (2.8) and (2.9) will establish $\lim\sup_n Q_n\{T_n \leqq x\} \leqq P\{\sum \lambda_k[(Z_k + a_k)^2 - 1] \leqq x\}$. Since an inequality similar to (2.7) but in the other direction holds as well this would prove (2.3).

Some calculation yields

$$E(T_{\infty,K_2} - T_{\infty,K_1})^2 = [\textstyle\sum_{k=K_1+1}^{K_2} \lambda_k a_k^2]^2 + 4 \sum_{k=K_1+1}^{K_2} \lambda_k^2 a_k^2 + 2 \sum_{k=K_1+1}^{K_2} \lambda_k^2 . \tag{2.10}$$

From $||Q - Q^K||^2 = ||Q||^2 - \sum_{k=1}^{K} \lambda_k^2$ it follows that $\sum_{k=1}^{\infty} \lambda_k^2 = ||Q||^2 < \infty$. Also

$$|\textstyle\int\int Q(s, t)h(s)h(t)\, dP_0 \times P_0 - \sum_{k=1}^{K} \lambda_k a_k^2|^2 = |\int\int [Q(s, t) - Q^K(s, t)]h(s)h(t)\, dP_0 \times P_0|^2 \leqq ||Q - Q^K||^2||h||^4$$

which implies $\sum_{k=1}^{\infty} \lambda_k a_k^2 = \int\int Q(s, t)h(s)h(t)\, dP_0 \times P_0 < \infty$. This shows that the RHS in (2.10) tends to zero for $K_1 \to \infty$, $K_2 \to \infty$. □

Proof of Corollary 2.2. The conditions A1, A2 and A3 appear in the appendix of Hajék (1970). Then P_n approximation to $\log dQ_n/dP_n$ given there establishes contiguity of $\{Q_n\}$ to $\{P_n\}$. The proof above would apply if we could establish (2.5). With the approximation to the log likelihood ratio as given this may be accomplished more easily by using Le Cam's third lemma (see Hajék and Sidák (1967), page 208) instead of the results of Behnen and Neuhaus (1975). □

3. Application to tests of fit: the simple hypothesis. For a degenerate kernel satisfying (2.2) we have $EQ(X_{n1}, X_{n2}) = 0$ under the hypothesis. If $EQ(X_{n1}, X_{n2}) > 0$ under a class of alternatives then the statistic T_n(2.1) may be used to test the hypothesis since $ET_n = (n - 1)EQ(X_{n1}, X_{n2})$. The test would reject for large values of T_n. Let us identify the form T_n as the pure form of the test statistic and the form of Theorem 2.3, $T_n^* = T_n + n^{-1} \sum_{i=1}^{n} Q(X_{ni}, X_{ni})$, as the usual form. If $\mu = \int Q(\cdot, \cdot)\, dP_0$ exists then $T_n^* - \mu$ and T_n are asymptotically equivalent under contiguous alternatives and asymptotically the tests based on T_n and T_n^* would be equivalent. Under certain conditions that do not appear to be restrictive (see the treatment of Cramér–von Mises tests below), when $\int Q(\cdot, \cdot)\, dP_0$ does not exist we may also assert the equivalence of T_n and T_n^*. By the weak law of large numbers (see Feller (1966), page 232) if

$$\lim_{t\to\infty} tP_0\{|Q(X_{n1}, X_{n1})| > t\} = 0 , \tag{3.1}$$

there is a sequence $\{\mu_n\}$ such that $n^{-1}\sum_{i=1}^n Q(X_{ni}, X_{ni}) - \mu_n \to 0$ in P_n probability. Then $T_n{}^* - \mu_n$ and T_n would be asymptotically equivalent under contiguous alternatives.

Consider now some specific tests of the hypothesis of the previous section when $\mathscr{X} = (0, 1)$ and P_0 is the Lebesgue measure on $(0, 1)$. Let us first consider chi-square tests. For a division of the unit interval $d_0 = 0 < d_1 < \cdots < d_{c-1} < d_c = 1$, let

$$
\begin{aligned}
&Q(s, t) = \textstyle\sum_{k=1}^c g_k(s)g_k(t) \quad \text{where} \\
(3.2)\qquad &g_k(s) = \{I[d_{k-1} < s \leqq d_k] - p_k{}^*\}(p_k{}^*)^{-\frac{1}{2}}, \quad 1 \leqq k \leqq c, \qquad \text{and} \\
&p_k{}^* = d_k - d_{k-1}.
\end{aligned}
$$

Then the $p_k{}^*$ are null hypothesis category probabilities and the usual chi-square test statistic is of the form $T_n{}^*$ with the kernel (3.2). It is easy to show that the kernel (3.2) has $c - 1$ nonzero eigenvalues equal to unity. Then Theorem 2.3 gives the familiar limiting chi-square distribution under the hypothesis. The limiting form under the alternatives is considered in Section 6.

Now consider the modified Cramér–von Mises tests (see Anderson and Darling (1952) and Darling (1957)). For a nonnegative weight function $w(u)$, $0 < u < 1$, define the kernel

$$
(3.3)\qquad Q(s, t) = \textstyle\int_0^1 \{I[s \leqq u] - u\}\{I[t \leqq u] - u\}w(u)\,du, \qquad 0 < s, t < 1.
$$

With this kernel the form $T_n{}^*$ is a modified Cramér–von Mises statistic.

Desired are conditions on w ensuring the square integrability of the symmetric kernel (3.3). For $s \leqq t$ we have

$$
(3.4)\qquad Q(s, t) = \textstyle\int_0^s u^2 w(u)\,du - \int_s^t u(1-u)w(u)\,du + \int_t^1 (1-u)^2 w(u)\,du.
$$

It may be demonstrated that

$$
(3.5)\qquad \textstyle\int_0^1\int_0^1 Q^2(s, t)\,ds\,dt = \int_0^1\int_0^1 \{\min\{u, v\} - uv\}^2 w(u)w(v)\,du\,dv
$$

by squaring (3.4) and integrating over $\{(s, t)\colon 0 < s \leqq t < 1\}$.

Interchanging the order of integrations on nonnegative components may be justified by Tonelli's theorem if the right-hand side of (3.5) is finite. Some conditions ensuring that Q is square-integrable (the right-hand side of (3.5) is finite) are the following:

$$
\begin{aligned}
&\text{(i)}\quad w\ (\geqq 0) \ \text{ is bounded on } \ [\varepsilon, 1-\varepsilon] \ \text{ for all } \ \varepsilon > 0. \\
(3.6)\qquad &\text{(ii)}\quad \lim_{u\to 0} u^2 w(u) = \lim_{u\to 1}(1-u)^2 w(u) = 0. \\
&\text{(iii)}\quad u^2 w(u)[(1-u)^2 w(u)] \ \text{ is monotone near } \ u = 0[u = 1]. \\
&\text{(iv)}\quad \textstyle\int_0^1 u^3(1-u)^3 w^2(u)\,du < \infty.
\end{aligned}
$$

These may be applied by noting that

$$
\textstyle\iint_{u<v<\varepsilon} u^2(1-v)^2 w(u)w(v)\,du\,dv = \int_0^\varepsilon (1-v)^2 w(v) \int_0^v u^2 w(u)\,du\,dv
$$

and

$$
\textstyle\iint_{1-\varepsilon<u<v} u^2(1-v)^2 w(u)w(v)\,du\,dv = \int_{1-\varepsilon}^1 u^2 w(u) \int_u^1 (1-v)^2 w(v)\,dv\,du.
$$

We assert that under the conditions (3.6) the kernel (3.3) satisfies (3.1) thus ensuring the equivalence of tests based on T_n and T_n^*.

Now Q has the same form as the kernel of Section 4 of de Wet and Venter (1973). There the eigenvalues and eigenfunctions are shown to satisfy a certain Sturm–Liouville type differential equation. For certain choices of the weight function w, the eigenfunctions of Q are related to classical orthogonal polynomials. For example, when

$$(3.7) \qquad w(u) = 1/\phi^2(\Phi^{-1}(u)) \,, \qquad 0 < u < 1 \,,$$

where ϕ and Φ are the standard normal density and cdf, the eigenvalues are $\lambda_k = 1/k$, $k \geqq 1$, corresponding to orthonormal eigenfunctions $h_k(\Phi^{-1}(u))$, where h_k is the kth Hermitian polynomial suitably normalized. It may be checked that w (3.7) satisfies (3.6) to ensure the square integrability of Q. Useful here is that the tails of (3.7) are asymptotic to those of $u^{-2}(1-u)^{-2}(-2\log u(1-u))^{-1}$, $0 < u < 1$. It may also be noted that Theorem 2.3 does not apply here since $\sum \lambda_k = \infty$.

Limiting null hypothesis distribution theory for the modified Cramér–von Mises statistics was developed by Anderson and Darling (1952) with no provision for $\sum |\lambda_k| = \infty$. They tabulated the limiting distribution of Theorem 2.3 for the weight function $w \equiv 1$. In Anderson and Darling (1954) a few points of the limiting distribution are given for the weight function $w(u) = 1/u(1-u)$. De Wet and Venter (1973) consider a statistic similar to the modified Cramér–von Mises statistic. For the weight functions giving $\sum |\lambda_k| = \infty$, a limiting distribution was obtained by subtracting the expected value of the statistic which approaches infinity as the sample size approaches infinity. We avoid this by considering the pure form T_n.

For the weight function (3.7) with eigenvalues $\lambda_k = 1/k$, $k \geqq 1$, the limiting null hypothesis distribution of $\sum_{k=1}^{\infty} (1/k)(Z_k^2 - 1)$ has been tabulated in de Wet and Venter (1972).

4. Applications to the Cramér–von Mises test when parameters are estimated. In this section we suppose that i.i.d. random variables $x_1, \cdots, x_n$ are observed and that it is desired to test the composite hypothesis that the common cdf is of a given form $F(x, \theta_1, \theta_2)$ for some unspecified parameters θ_1 and θ_2. We restrict ourselves to the two parameter case though the methods are general. The hypothesis here is composite and the hypothesis of Section 2 was simple but we shall blur the distinction by referring to a null hypothesis probability $P_n = P_0 \times \cdots \times P_0$, n times, generated by letting P_0 be the measure associated with $F(x, \theta_{10}, \theta_{20})$ where θ_{10} and θ_{20} represent the true but unknown parameter values. Further $X_{n1}, \cdots, X_{nn}$ of Section 2 may be associated with $x_1, \cdots, x_n$ and the notation will not be deficient in this section since only null hypothesis probabilities will be used. Our program is to obtain P_n approximations to various Cramér–von Mises statistics in a form that allows the results of Section

2 to be used. Then limiting distributions under contiguous alternatives will be available as well.

For testing goodness of fit to the two parameter family $F(x, \theta_1, \theta_2)$ the modified Cramér–von Mises statistic with weight function w is

$$n \int [F_n(x) - F(x, \hat{\theta}_1, \hat{\theta}_2)]^2 w[F(x, \hat{\theta}_1, \hat{\theta}_2)]\, dF(x, \hat{\theta}_1, \hat{\theta}_2)\,, \tag{4.1}$$

where F_n is the sample cdf and $\hat{\theta}_1$ and $\hat{\theta}_2$ are estimates of the unknown parameters. Previous work relating to the limiting distribution of (4.1) for various special cases is contained in Darling (1955), Darling (1957), Kac, Kiefer and Wolfowitz (1955) and Sukhatme (1972).

Following the work of Section 3 we wish to consider tests embodying the basic notion behind (4.1) but allowing weight functions w which would have such heavy tails that (4.1) would not possess a limiting distribution. This is accomplished by considering statistics

$$\begin{aligned}\hat{T}_n = n^{-1} \textstyle\sum_{i\neq j} \int &\{I[x_i \leqq x] - F(x, \hat{\theta}_1, \hat{\theta}_2)\} \\ &\times \{I[x_j \leqq x] - F(x, \hat{\theta}_1, \hat{\theta}_2)\} w[F(x, \hat{\theta}_1, \hat{\theta}_2)]\, dF(x, \hat{\theta}_1, \hat{\theta}_2)\,.\end{aligned} \tag{4.2}$$

Theorem 4.3 states that $\hat{T}_n$ is asymptotically equivalent under P_n probability to $T_{n2} + \gamma$ where $T_{n2} = n^{-1} \sum_{i\neq j} Q(x_i, x_j)$ and Q and γ involve no random components. The results of Section 2 may be applied to T_{n2} to find the limiting distribution under $\{P_n\}$ or contiguous alternatives $\{Q_n\}$.

In Section 5 the theorems are illustrated by a test for normality using the weight function (3.7). In Section 7 are the lemmas which create the sharpness in the results.

Let $f^*(x, \theta_1, \theta_2) = w[F(x, \theta_1, \theta_2)]f(x, \theta_1, \theta_2)$ where f is the density of F. For any function g let $g_{ij} = (\partial/\partial\theta_1)^i(\partial/\partial\theta_2)^j g$. Thus $F_{10}(x, \theta_1, \theta_2) = (\partial/\partial\theta_1)F(x, \theta_1, \theta_2)$ and $f_{11}^*(x, \theta_1, \theta_2) = (\partial^2/\partial\theta_1\partial\theta_2)f^*(x, \theta_1, \theta_2)$. If we drop the dependency on the parameters θ_1 and θ_2, the values θ_{10} and θ_{20} are implied. Thus $F(x) = F(x, \theta_{10}, \theta_{20})$. Let

$$\begin{aligned}T_{n1} = n^{-1} \textstyle\sum_{i\neq j} \int &\{I[x_i \leqq x] - F(x) - (\hat{\theta}_1 - \theta_{10})F_{10}(x) \\ &- (\hat{\theta}_2 - \theta_{20})F_{01}(x)\}\{I[x_j \leqq x] - F(x) - (\hat{\theta}_1 - \theta_{10})F_{10}(x) \\ &- (\hat{\theta}_2 - \theta_{20})F_{01}(x)\} f^*(x)\, dx\,.\end{aligned} \tag{4.3}$$

The following theorem carries the approximation notion of Lemma 3.1 of Sukhatme (1972) to the statistics $\hat{T}_n$ (4.2) of our concern.

THEOREM 4.1. *Suppose that*

(1) $n^{\frac{1}{2}}(\hat{\theta}_i - \theta_{i0}) = O_{P_n}(1)$, $i = 1, 2$.

(2) *For almost all x and all (θ_1, θ_2) in a neighborhood of $(\theta_{10}, \theta_{20})$, $F_{10}(x, \theta_1, \theta_{20})$ $[F_{01}(x, \theta_{10}, \theta_2)]$ is continuous in θ_1 $[\theta_2]$ and $F_{11}(x, \theta_1, \theta_2)$ and $f_{11}^*(x, \theta_1, \theta_2)$ exist and are bounded for each fixed x.*

(3) *For some $\alpha_1 < \frac{3}{2}$, some $\alpha_2 < 2$, and any $\alpha_3 < 1$ there is a neighborhood N*

of $(\theta_{10}, \theta_{20})$ *such that for almost all* x *and all* (θ_1, θ_2) *in* N

$$\max\{|f^*(x)|, |f_{10}^*(x, \theta_1, \theta_{20})|, |f_{01}^*(x, \theta_{10}, \theta_2)|\} \leqq q(x)$$
$$\max\{|f_{11}^*(x, \theta_1, \theta_2)|, q(x)\} \leqq q^*(x)$$
$$\max\{|F_{10}(x, \theta_1, \theta_{20})|, |F_{01}(x, \theta_{10}, \theta_2)|, |F_{11}(x, \theta_1, \theta_2)|\} \leqq m(x)$$

where

(a) $\int [F(1-F)]^{\alpha_1} q < \infty$
(b) $\int [F(1-F)]^{\alpha_2} q^* < \infty$
(c) $m \leqq c[F(1-F)]^{\alpha_3}$ *with* c *constant.*

Then $\hat{T}_n$ (4.2) *and* T_{n1} (4.3) *satisfy* $\hat{T}_n - T_{n1} \to_{P_n} 0$.

PROOF. We may write

$$F(x, \hat{\theta}_1, \hat{\theta}_2) = F(x) + \int_{\theta_{10}}^{\hat{\theta}_1} F_{10}(x, \theta_1, \theta_{20})\, d\theta_1 + \int_{\theta_{20}}^{\hat{\theta}_2} F_{01}(x, \theta_{10}, \theta_2)\, d\theta_2 + \int_{\theta_{10}}^{\hat{\theta}_1} \int_{\theta_{20}}^{\hat{\theta}_2} F_{11}(x, \theta_1, \theta_2)\, d\theta_1\, d\theta_2 \quad .$$

and similarly for $f^*(x, \hat{\theta}_1, \hat{\theta}_2)$. These expansions may be substituted in $\hat{T}_n$ (4.2) to obtain

$$\hat{T}_n = n^{-1} \sum_{i \neq j} \int B(x_i, x) B(x_j, x) C(x)\, dx \quad \text{where}$$
$$\begin{aligned} B(x_i, x) = {} & I[x_i \leqq x] - F(x) - (\hat{\theta}_1 - \theta_{10}) F_{10}(x) - (\hat{\theta}_2 - \theta_{20}) F_{01}(x) \\ & - \int_{\theta_{10}}^{\hat{\theta}_1} [F_{10}(x, \theta_1, \theta_{20}) - F_{10}(x)]\, d\theta_1 \\ & - \int_{\theta_{20}}^{\hat{\theta}_2} [F_{01}(x, \theta_{10}, \theta_2) - F_{01}(x)]\, d\theta_2 \\ & - \int_{\theta_{10}}^{\hat{\theta}_1} \int_{\theta_{20}}^{\hat{\theta}_2} F_{11}(x, \theta_1, \theta_2)\, d\theta_1\, d\theta_2 \end{aligned} \tag{4.4}$$

and

$$C(x) = f^*(x) + \int_{\theta_{10}}^{\hat{\theta}_1} f_{10}^*(x, \theta_1, \theta_{20})\, d\theta_1 + \int_{\theta_{20}}^{\hat{\theta}_2} f_{01}^*(x, \theta_{10}, \theta_2)\, d\theta_2 + \int_{\theta_{10}}^{\hat{\theta}_1} \int_{\theta_{20}}^{\hat{\theta}_2} f_{11}^*(x, \theta_1, \theta_2)\, d\theta_1\, d\theta_2 \, .$$

The theorem asserts that cross product terms in (4.4) converge to zero in probability except those involving the first four terms of $B(x_i, x)$, the first four terms of $B(x_j, x)$ and the first term of $C(x)$. We treat two of the cross product terms to show the types of arguments used for the remainder; it is then tedious but straightforward to verify the theorem.

Consider the cross product term

$$n^{-1} \sum_{i \neq j} \int \{I[x_i \leqq x] - F(x)\}\{I[x_j' \leqq x] - F(x)\} \int_{\theta_{10}}^{\hat{\theta}_1} f_{10}^*(x, \theta_1, \theta_{20})\, d\theta_1\, dx \, . \tag{4.5}$$

Rewrite (4.5) as

$$n^{\frac{1}{2}} \int_{\theta_{10}}^{\hat{\theta}_1} \int \{[n^{-\frac{3}{4}} \sum_i \{I[x_i \leqq x] - F(x)\}]^2 - n^{-\frac{3}{2}} \sum_i \{I[x_i \leqq x] - F(x)\}^2\} f_{10}^*(x, \theta_1, \theta_{20})\, dx\, d\theta_1 \, . \tag{4.6}$$

Lemma 7.3 implies that

$$\sup_x n^{-\frac{3}{4}} |\sum \{I[x_i \leqq x] - F(x)\}| [F(x)(1 - F(x))]^{-\alpha_1} \to_{P_n} 0 \quad \text{for } \tfrac{1}{2} \leqq \alpha_1 < \tfrac{3}{4} \, . \tag{4.7}$$

Lemma 7.4 implies that

$$\sup_x n^{-\frac{3}{2}} \sum \{I[x_i \leqq x] - F(x)\}^2 [F(x)(1 - F(x))]^{-\alpha_2} \to_{P_n} 0 \quad \text{for} \quad 1 \leqq \alpha_2 < \tfrac{3}{2}. \tag{4.8}$$

Recall that $|f_{10}^*(x, \theta_1, \theta_{20})| \leqq q(x)$. Now (4.7), (4.8) and assumption 3a imply that (4.6) converges to zero in probability.

Consider another cross product term

$$n^{-1} \sum_{i \neq j} \int (\hat{\theta}_1 - \theta_{10}) F_{10}(x) \int_{\theta_{10}}^{\hat{\theta}_1} [F_{10}(x, \theta_1, \theta_{10}) - F_{10}(x)]\, d\theta_1\, f^*(x)\, dx \tag{4.9}$$

$$= (n-1)(\hat{\theta}_1 - \theta_{10}) \int_{\theta_{10}}^{\hat{\theta}_1} g(\theta_1)\, d\theta_1, \quad \text{where} \tag{4.10}$$

$$g(\theta_1) = \int F_{10}(x)[F_{10}(x, \theta_1, \theta_{10}) - F_{10}(x)] f^*(x)\, dx. \tag{4.11}$$

The integrand in (4.11) is bounded by $2m^2(x)q(x)$ which is integrable by the assumptions. Since the integrand in (4.11) converges to zero as $\theta_1 \to \theta_{10}$ for almost all x, we have that $\lim_{\theta_1 \to \theta_{10}} g(\theta_1) = 0$. This shows that (4.10) and hence (4.9) converges to zero in probability. □

COROLLARY 4.2. *In the definition of* $\hat{T}_n$ *and* T_{n1} *we may change* $\sum_{i \neq j}$ *to* $\sum_{i,j}$ *and Theorem* 4.1 *still holds.*

THEOREM 4.3. *Suppose the conditions of Theorem* 4.1 *hold and* $\hat{\theta}_k - \theta_{k0} = n^{-1} \sum_{i=1}^n g_k(x_i) + \varepsilon_{kn}$, *where* $n^{\frac{1}{2}} \varepsilon_{kn} = o_{P_n}(1)$ *as* $n \to \infty$ *and* $g_k(x_i)$ *has mean zero and finite variance,* $k = 1, 2$. *If the expectations below exist then* $\hat{T}_n - (T_{n2} + \gamma) \to_{P_n} 0$, *where*

$$T_{n2} = n^{-1} \sum_{i \neq j} Q(x_i, x_j), \quad \text{and} \tag{4.12}$$

$$\begin{aligned} Q(x_i, x_j) = \int \{ & I[x_i \leqq x] - F(x) - g_1(x_i)F_{10}(x) - g_2(x_i)F_{01}(x)\} \\ & \times \{I[x_j \leqq x] - F(x) - g_1(x_j)F_{10}(x) \\ & - g_2(x_j)F_{01}(x)\} f^*(x)\, dx \end{aligned} \tag{4.13}$$

and

$$\begin{aligned} \gamma = -2E \int & \{I[x_i \leqq x] - F(x)\}\{g_1(x_i)F_{10}(x) + g_2(x_i)F_{01}(x)\} f^*(x)\, dx \\ & + E[g_1(x_i)]^2 \int F_{10}^2 f^* + E[g_2(x_i)]^2 \int F_{01}^2 f^* \\ & + 2Eg_1(x_i)g_2(x_i) \int F_{10}F_{01} f^*. \end{aligned} \tag{4.14}$$

PROOF. With the techniques of Theorem 4.1 it is easy to show that $\hat{T}_n$ is asymptotically equivalent under P_n to (4.3) with $n^{-1} \sum_{i=1}^n g_k(x_i)$ in place of $\hat{\theta}_k - \theta_{k0}$, $k = 1, 2$. We will treat one of the cross product terms in this modification of (4.3). Treatment of the others is similar.

Consider

$$\begin{aligned} n^{-1} \sum_{i \neq j} & \int \{I[x_i \leqq x] - F(x)\}\{-n^{-1} \textstyle\sum_{j=1}^n g_1(x_j)F_{10}(x)\} f^*(x)\, dx \\ & = -(n-1)n^{-2} \textstyle\sum_{i,j} \int \{I[x_i \leqq x] - F(x)\} g_1(x_j) F_{10}(x) f^*(x)\, dx \\ & = -n^{-1} \textstyle\sum_{i \neq j} \int \{I[x_i \leqq x] - F(x)\} g_1(x_j) F_{10}(x) f^*(x)\, dx \\ & \quad - n^{-1} \textstyle\sum_i \int \{I[x_i \leqq x] - F(x)\} g_1(x_i) F_{10}(x) f^*(x)\, dx \\ & \quad + [n^{-1} \textstyle\sum_j g_1(x_j)][n^{-1} \textstyle\sum_i \int \{I[x_i \leqq x] - F(x)\} F_{10}(x) f^*(x)\, dx]. \end{aligned} \tag{4.15}$$

The first term of (4.15) is a component of Q (4.13). The second term of (4.15) converges to a component of γ (4.14). The third term of (4.15) converges to zero by assumptions on g_1 and the conditions of Theorem 4.1. □

5. A test for normality. Consider testing the hypothesis of normality with unknown mean and variance, θ_{10} and θ_{20} respectively. Let us test the hypothesis by the statistic $\hat{T}_n$ (4.2), (i) with weight function w given by (3.7) and (ii) using maximum likelihood estimates $\hat{\theta}_1 = n^{-1} \sum x_i$ and $\hat{\theta}_2 = n^{-1} \sum (x_i - \bar{x})^2$. We consider below the satisfaction of the conditions of Theorem 4.3 in this case. Then applying the work of Section 3 the limiting distribution of $\hat{T}_n$ is obtained.

Writing $y = \theta_2^{-\frac{1}{2}}(x - \theta_1)$ we find $f^*(x, \theta_1, \theta_2) = \theta_2^{-\frac{1}{2}}/\phi(y)$ and $F(x, \theta_1, \theta_2) = \Phi(y)$. It is straightforward to obtain the derivatives needed in Theorem 4.1.

For (θ_1, θ_2) in a neighborhood of $(\theta_{10}, \theta_{20})$ where $1 - \varepsilon < \theta_{20}/\theta_2 < 1 + \varepsilon$, the following may be used as bounds in condition 3 of Theorem 4.1:

$$\begin{aligned} m(x) &= K_1[1 + x^2] \exp\{-(1 - \varepsilon)x^2/2\theta_{20}\} \\ q(x) &= K_2[1 + |x|] \exp\{(1 + \varepsilon)x^2/2\theta_{20}\} \quad \text{and} \\ q^*(x) &= K_3[1 + |x|^3] \exp\{(1 + \varepsilon)x^2/2\theta_{20}\}, \end{aligned}$$

where K_1, K_2 and K_3 are constants. From the fact that $\log \Phi(x) \sim -x^2/2$ as $x \to -\infty$, it follows that for any $\rho > 0$ there is a constant K_4 such that

$$F(x)[1 - F(x)] \leqq K_4 \exp\{-(1 - \rho)x^2/2\theta_{20}\}.$$

It is now easy to verify the conditions of Theorem 4.1.

For the application of Theorem 4.3 take $g_1(x_i) = x_i - \theta_{10}$ and $g_2(x_i) = (x_i - \theta_{10})^2 - \theta_{20}$. We proceed to calculate γ (4.14).

In what follows it is convenient to realize that the original statistic $\hat{T}_n$ is location and scale invariant so that without loss of generality one may assume $\theta_{10} = 0$ and $\theta_{20} = 1$. Then $g_1(x_i) = x_i$ and $g_2(x_i) = x_i^2 - 1$, and also $F(x) = \Phi(x)$, $F_{10}(x) = -\phi(x)$, $F_{01}(x) = -x\phi(x)/2$, and $f^*(x) = 1/\phi(x)$.

Entering into γ (4.14) are the functions

$$\begin{aligned} H_1(x_i) &= \int \{I[x_i \leqq x] - F(x)\}F_{10}(x)f^*(x)\,dx \\ &= -\int \{I[x_i \leqq x] - \Phi(x)\}\,dx = x_i, \end{aligned} \tag{5.1}$$

and

$$\begin{aligned} H_2(x_i) &= \int \{I[x_i \leqq x] - F(x)\}F_{01}(x)f^*(x)\,dx \\ &= -(\tfrac{1}{2}) \int \{I[x_i \leqq x] - \Phi(x)\}x\,dx = (x_i^2 - 1)/4. \end{aligned} \tag{5.2}$$

Substitution in (4.14) yields $\gamma = -\frac{3}{2}$. Using (5.1) and (5.2) the kernel Q (4.13) may be written

$$\begin{aligned} Q(x_i, x_j) &= Q_0(x_i, x_j) - x_i x_j - (x_i^2 - 1)(x_j^2 - 1)/4, \quad \text{where} \\ Q_0(x_i, x_j) &= \int \{I[x_i \leqq x] - \Phi(x)\}\{I[x_j \leqq x] - \Phi(x)\}(1/\phi(x))\,dx. \end{aligned} \tag{5.3}$$

THEOREM 5.1. *Let $x_1, \cdots, x_n$ be a sample from a normal* cdf. *Then the limiting* *distribution of*

$$n^{-1} \sum_{i \neq j} Q_0\left(\frac{x_i - \hat{\theta}_1}{\hat{\theta}_2^{\frac{1}{2}}}, \frac{x_j - \hat{\theta}_1}{\hat{\theta}_2^{\frac{1}{2}}}\right) \tag{5.4}$$

is that of $\sum_{k=3}^{\infty} (Z_k^2 - 1)/k - (\frac{3}{2})$ *where* $Z_1, Z_2, \cdots$ *are* i.i.d. *standard normal variables.*

Proof. The expression (5.4) is $\hat{T}_n$ with the weight function (3.7). Theorem 4.3 has given us that (5.4) is equivalent to $T_{n2} - \frac{3}{2}$, where $T_{n2} = n^{-1} \sum_{i \neq j} Q(x_i, x_j)$ and Q is given by (5.3). Theorem 2.1 shows that the limiting distribution of T_{n2} is that of $\sum \lambda_k(Z_k^2 - 1)$, where $\{\lambda_k\}$ are the eigenvalues of Q (5.3).

The kernel $Q_0(\Phi^{-1}(s), \Phi^{-1}(t))$, $0 < s, t < 1$, is the one discussed earlier in (3.3) and (3.7). The first two Hermitian polynomials are $h_1(x) = x$ and $h_2(x) = 2^{-\frac{1}{2}}(x^2 - 1)$. Comparison with (5.3) shows that Q has eigenvalues $1/k$, $k = 3, 4, \cdots$. □

When either the mean or variance is known a result analogous to Theorem 5.1 may easily be derived.

The distribution of $\sum_{k=3}^{\infty} (Z_k^2 - 1)/k$ is tabulated in de Wet and Venter (1972). They find this to be the limiting distribution of the test statistic $L_n - a_n$ where $L_n = \sum [(x_{in} - \hat{\theta}_1)/\hat{\theta}_2^{\frac{1}{2}} - \Phi^{-1}(i/(n+1))]^2$, with $x_{1n} \leqq \cdots \leqq x_{nn}$ the order statistics of $x_1, \cdots, x_n$ and $\{a_n\}$ a sequence of constants approaching infinity approximating the mean of L_n. We expect the test based on (5.4) to be asymptotically equivalent under contiguous alternatives to the one based on L_n. Subtracting the constants a_n from L_n corresponds to our removing the $i = j$ terms from (4.1) to obtain (4.2).

6. The asymptotic performance of the tests. In Sections 3, 4 and 5 have been examples of tests whose asymptotic power properties may be studied by means of Theorem 2.1. We identify the following asymptotic testing situation depending on a probability space $(\mathscr{X}, B, P_0)$: (1) Corresponding to each kernel Q in a given class one observes a null hypothesis "test statistic" $\sum \lambda_k(Z_k^2 - 1)$ where $\{\lambda_k\}$ are the eigenvalues of Q. (2) Under alternatives defined by $h \in L_2(\mathscr{X}, B, P_0)$ one observes the "test statistic" $\sum \lambda_k[(Z_k + a_k)^2 - 1]$ where $a_k = \int h f_k \, dP_0$ with $\{f_k\}$ the orthonormal eigenfunctions of Q.

Consider the following simple index of performance for a test based on Q against the alternative h:

$$\begin{aligned} e_{Q,h} &= \frac{\text{alternative hypothesis mean}}{\text{null hypothesis standard deviation}} \\ &= \{\sum \lambda_k a_k^2\}\{2 \sum \lambda_k^2\}^{-\frac{1}{2}} \\ &= \{\int\int Q(s, t)h(s)h(t)\, dP_0(s)\, dP_0(t)\}\{2 \int\int Q^2(s, t)\, dP_0(s)\, dP_0(t)\}^{-\frac{1}{2}}\,. \end{aligned} \tag{6.1}$$

For fixed alternative h and two tests of the same limiting chi-square type the denominators of the indices (6.1) are the same and the ratio of the indices is the ratio of the noncentrality parameters. This ratio of noncentrality parameters is the true asymptotic relative efficiency, i.e., the limiting ratio of sample sizes to give the same limiting power under the alternatives. In general the ratio of indices (6.1) may be interpreted as the limiting ratio of sample sizes to give the same value of the index (6.1).

When Q (the test) is fixed and the alternatives h are varied some power results may be obtained from Theorem 2.2 of Neuhaus (1976). The only change needed in the arguments presented there to cover our case involves centering the random variables and this change is easily made. (We would now require only $\sum \lambda_k^2 < \infty$ instead of $\sum \lambda_k < \infty$.) The basic result is that for alternatives h defined by eigenfunctions f_i and f_j, the asymptotic powers have the same order relation as the eigenvalues λ_i and λ_j associated with the eigenfunctions. It is interesting to note that for fixed Q the indices $e_{Q,h}$ enjoy the same property (see Dunford and Schwartz (1963), pages 907, 908).

We now give the forms of (6.1) that result from chi-square tests and modified Cramér–von Mises tests of Section 3. For the kernel (3.2) of the chi-square test

$$e_{Q,h} = \{\textstyle\sum_{k=1}^{c} (\int_{d_{k-1}}^{d_k} h)^2 (d_k - d_{k-1})^{-1}\}\{2(c-1)\}^{-\frac{1}{2}} . \tag{6.2}$$

The first factor of (6.2) is the noncentrality parameter under the alternatives which was first given in a different setting by Eisenhart (1938). As the number of cells c approaches infinity while the maximum cell width approaches zero the noncentrality parameter approaches the finite value $\int h^2$. Thus the limiting power will approach zero as $c \to \infty$ for a particular alternative sequence. However with increased c the test has some power over a wider class of alternatives. This decrease in power with increasing c is noted in Kendall and Stuart (1961), page 436.

For the kernel (3.3) of the Cramer–von Mises test

$$e_{Q,h} = \{\textstyle\int (\int_0^u h)^2 w(u)\, du\}\{2 \iint [\min\{u, v\} - uv]^2 w(u) w(v)\, du\, dv\}^{-\frac{1}{2}} .$$

7. Some large sample results for uniform variables. Let $u_1, \cdots, u_n$ be i.i.d. uniform (0, 1) variables and $u_{n1} \leqq \cdots \leqq u_{nn}$ the associated order statistics. In this section we obtain bounds on various functions of the u_{ni}'s where a symmetry exists between considerations for $1 \leqq i \leqq [n/2]$ and considerations for $[n/2] \leqq i \leqq n$. For brevity we state the result for the complete set of u_{ni}'s only when it is notationally simple, otherwise we consider u_{ni}, $1 \leqq i \leqq [n/2]$.

LEMMA 7.1. *For any* $\Delta < \frac{1}{2}$,

$$\lim_{k\to\infty} \liminf_n P\{n^{-\frac{1}{2}} |\textstyle\sum_{i=1}^n \{I[u_i \leqq u] - u\}| \leqq k u^{\Delta}(1-u)^{\Delta}, 0 < u < 1\} = 1 .$$

PROOF. Follows from Lemma 2.2 of Pyke and Shorack (1968).

LEMMA 7.2. $\lim_{k\to\infty} \liminf_n P\{i/kn \leqq u_{ni} \leqq ki/n,\ 1 \leqq i \leqq [n/2]\} = 1.$

PROOF. Follows from Lemma 8 of Govindarajulu, Le Cam and Raghavachari (1967).

LEMMA 7.3. *If* $\frac{1}{2} \leqq \alpha \leqq 1$ *and* $\beta > \alpha$

$$\sup_{0<u<1} n^{-\beta}|\sum_{i=1}^n \{I[u_i \leqq u] - u\}|u^{-\alpha}(1-u)^{-\alpha} \rightarrow_P 0 .$$

PROOF. From symmetry considerations it is sufficient to show that

$$\sup_{0<u\leqq u_{n,[n/2]}} n^{-\beta}|\sum_{i=1}^n \{I[u_i \leqq u] - u\}|u^{-\alpha} \tag{7.1}$$

converges to zero in probability. Let

$$B_{nk} = \{|\sum_{i=1}^n \{I[u_i \leqq u] - u\}| \leqq n^{\frac{1}{2}}ku^{\Delta}, 0 < u \leqq u_{n,[n/2]}, \\ i/kn \leqq u_{ni} \leqq ki/n, 1 \leqq i \leqq [n/2]\} ,$$

where $\Delta < \frac{1}{2}$ is chosen so that $\frac{1}{2} - \Delta < \beta - \alpha$. From Lemmas 7.1 and 7.2 we have $\lim_{k\to\infty} \liminf_n P(B_{nk}) = 1$. The proof will be finished if (7.1) is bounded on B_{nk} and the bound approaches zero in probability as $n \to \infty$ for fixed k.

It is clear that the supremum of (7.1) over $0 < u \leqq u_{n1}$ converges to zero in probability. On B_{nk} the supremum of (7.1) over $u_{n1} \leqq u \leqq u_{n,[n/2]}$ is bounded by

$$\sup_{u_{n1}\leqq u\leqq u_{n,[n/2]}} n^{-\beta}n^{\frac{1}{2}}ku^{\Delta}u^{-\alpha} = kn^{\frac{1}{2}-\beta}(1/kn)^{\Delta-\alpha} = o(1) . \qquad \square$$

LEMMA 7.4. *If* $1 \leqq \alpha \leqq 2$ *and* $\beta > \alpha$

$$\sup_{0<u<1} n^{-\beta} \sum_{i=1}^n \{I[u_i \leqq u] - u\}^2u^{-\alpha}(1-u)^{-\alpha} \rightarrow_P 0 .$$

PROOF. As before, symmetry considerations make it sufficient to show

$$\sup_{0<u\leqq u_{n,[n/2]}} n^{-\beta} \sum_{i=1}^n \{I[u_i \leqq u] - u\}^2u^{-\alpha} \tag{7.2}$$

converges to zero in probability. We bound (7.2) on B_{nk} of the previous lemma.

It is easy to see that the supremum of (7.2) over $0 < u \leqq u_{n1}$ converges to zero in probability. Now for $u_{ni} \leqq u \leqq u_{n,i+1}$ we have

$$\begin{aligned} \sum_{i=1}^n \{I[u_i \leqq u] - u\}^2u^{-\alpha} &= [i(1-u)^2 + (n-i)u^2]u^{-\alpha} \\ &< n[i/n + u]u^{-\alpha} \\ &< n[(i/n)u_{ni}^{-\alpha} + u_{ni}^{1-\alpha}] . \end{aligned}$$

On B_{nk} the above is bounded by

$$n[k^{\alpha}(i/n)^{1-\alpha} + k^{1-\alpha}(i/n)^{1-\alpha}] < n2k^{\alpha}(i/n)^{1-\alpha} .$$

Therefore on B_{nk} the supremum of (7.2) over $u_{n1} \leqq u \leqq u_{n,[n/2]}$ is bounded by $\max_{1\leqq i\leqq[n/2]} n^{-\beta}n2k^{\alpha}(i/n)^{1-\alpha} = 2k^{\alpha}n^{1-\beta+\alpha-1}$. $\square$

Acknowledgment. The author would like to thank one of the referees for suggesting the generalization of Theorem 2.1 over the original Corollary 2.2 and for providing many details. Gratitude is expressed also to a second referee for providing the asymptotic equivalence of tests based on the pure and the usual form when (3.1) is satisfied.

REFERENCES

[1] ANDERSON, T. W. and DARLING, D. A. (1952). Asymptotic theory of certain goodness of fit criteria based on stochatic processes. *Ann. Math. Statist.* **23** 193–212.

[2] ANDERSON, T. W. and DARLING, D. A. (1954). A test for goodness of fit. *J. Amer. Statist. Assoc.* **49** 765–769.

[3] BEHNEN, K. and NEUHAUS, G. (1975). A central limit theorem under contiguous alternatives. *Ann. Statist.* **3** 1349–1353.

[4] DARLING, D. A. (1955). The Cramér-Smirnov test in the parametric case. *Ann. Math. Statist.* **26** 1-20.

[5] DARLING, D. A. (1957). The Kolmogorov-Smirnov, Cramér-von Mises tests. *Ann. Math. Statist.* **28** 823–838.

[6] DE WET, T. and VENTER, J. H. (1972). Asymptotic distributions of certain test criteria of normality. *South African Statist. J.* **6** 135–149.

[7] DE WET, T. and VENTER, J. H. (1973). Asymptotic distributions for quadratic forms with applications to tests of fit. *Ann. Statist.* **1** 380–387.

[8] DUNFORD, N. and SCHWARTZ, J. T. (1963). *Linear Operators Part II.* Interscience, New York.

[9] EISENHART, C. (1938). The power function of the χ^2 test. *Bull. Amer. Math. Soc.* **44** 32.

[10] FELLER, W. F. (1966). *An Introduction to Probability Theory and Its Applications* **2**. Wiley, New York.

[11] FILIPPOVA, A. A. (1961). Mises' theorem on the asymptotic behavior of functionals of empirical distribution functions and its statistical applications. *Theor. Probability Appl.* **7** 24–56.

[12] GOVINDARAJULU, Z., LE CAM, L. and RAGHAVACHARI, M. (1967). Generalizations of theorems of Chernoff-Savage on asymptotic normality of nonparametric test statistics. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* 609–638, Univ. of California Press.

[13] HÁJEK, J. and SIDÁK, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

[14] HÁJEK, J. (1970). Asymptotic minimax and admissibility in estimation. FSU Statistics Report M186. Florida State Univ., Tallahassee.

[15] HOEFFDING, W. (1948a). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.

[16] HOEFFDING, W. (1948b). A non-parametric test of independence. *Ann. Math. Statist.* **19** 546–557.

[17] KAC, M., KIEFER, J. and WOLFOWITZ, J. (1955). On tests of normality and other tests of goodness of fit based on distance methods. *Ann. Math. Statist.* **23** 189–211.

[18] KENDALL, M. G. and STUART, A. (1961). *The Advanced Theory of Statistics* **2**. Hafner, New York.

[19] NEUHAUS, G. (1974). Asymptotic properties of the Cramér-von Mises statistic when parameters are estimated. *Proc. Prague Symp. on Asymptotic Statist.* Sept. 3-6, 1973 (J. Hájek, ed.). *Universita Karlova Praha* **2** 257–297.

[20] NEUHAUS, G. (1976). Asymptotic power properties of the Cramér-von Mises test under contiguous alternatives. *J. Multivariate Analysis* **6** 95–110.

[21] PYKE, R. and SHORACK, G. R. (1968). Weak convergence of a two-sample empirical process and a new approach to Chernoff-Savage theorems. *Ann. Math. Statist.* **39** 755–771.

[22] SCHACH, S. (1969). The asymptotic distribution of some non-linear functions of the two-sample rank vector. *Ann. Math. Statist.* **40** 1011–1021.

[23] SUKHATME, S. (1972). Fredholm determinant of a positive definite kernel of a special type and its application. *Ann. Math. Statist.* **43** 1914–1925.

[24] VARBERG, D. E. (1966). Convergence of quadratic forms in independent random variables. *Ann. Math. Statist.* **37** 567–575.

[25] VON MISES, R. (1947). On the asymptotic distribution of differentiable statistical functions. *Ann. Math. Statist.* **18** 309–348.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ARKANSAS
FAYETTEVILLE, ARKANSAS 72701