EXACT POWER OF GENERALIZED KOLMOGOROV GOODNESS-OF-FIT TESTS

by

Leon Jay Gleser[1]

Mimeograph Series #81-13

Department of Statistics
Division of Mathematical Sciences

May 1981

[1]Research supported by NSF grant MCS-79-05815.

SUMMARY

EXACT POWER OF GENERALIZED KOLMOGOROV GOODNESS-OF-FIT TESTS

Running Title: EXACT POWER OF KOLMOGOROV TESTS

by

Leon Jay Gleser
Purdue University

Generalized Kolmogorov goodness-of-fit tests of a specified c.d.f. $F^*(x)$ are defined to reject H_0: $F = F^*$ when the sample c.d.f. $F_n(x)$ either strictly exceeds a function $G_1(x)$, or is strictly less than a function $G_2(x)$, for some x. Under the sole condition that the function $\inf\{G_1(z): z \geq x\}$ is right-continuous in x, it is shown that every such test is equivalent to a test which rejects H_0 when $X_{(i)} < a_i$ or $X_{(i)} > b_i$, some $i = 1,2,\ldots,n$, where a_i, b_i are constants, $-\infty \leq a_i, b_i \leq \infty$, $i = 1,2,\ldots,n$, and $X_{(i)}$ is the ith order statistic, $1 \leq i \leq n$. Conversely, every test of this latter type is equivalent to a generalized Kolmogorov test based on nondecreasing right-continuous step-functions $G_1(x), G_2(x)$. It is shown that even when the true c.d.f. $F(x)$ is discontinuous, the power functions of such tests can be obtained from the joint c.d.f. of the order statistics $U_{(1)} \leq \cdots \leq U_{(n)}$ from a sample of i.i.d. $U_{[0,1]}$ random variables. Consequently, all one-sided generalized Kolmogorov tests are unbiased tests of H_0 versus appropriate one-sided alternatives. Finally, it is shown that no additional generality is introduced by defining generalized Kolmogorov tests to reject H_0 when $\psi_1(F_n(x)) < W_1(x)$ or $\psi_2(F_n(x)) > W_2(x)$, some x, where $\psi_1(u)$, $\psi_2(u)$ are arbitrary nondecreasing functions of u, $0 \leq u \leq 1$, and $W_1(x)$, $W_2(x)$ are arbitrary functions of x.

EXACT POWER OF GENERALIZED KOLMOGOROV GOODNESS-OF-FIT TESTS[1]

by

Leon Jay Gleser
Purdue University

1. Introduction and Summary. Let $X_1, X_2, \ldots, X_n$ be n independent observations from the population of a random variable X. Let

$$(1.1) \qquad F(x) = P\{X \leq x\}, \quad F_n(x) = \frac{\#X_i\text{'s} \leq x}{n}, \quad -\infty < x < \infty,$$

be the population cumulative distribution function (c.d.f.) and sample c.d.f., respectively, of X. It is desired to test the null hypothesis

$$(1.2) \qquad H_0: F(x) = F^*(x), \quad \text{all } x, \quad -\infty < x < \infty,$$

where the null c.d.f., $F^*(x)$, is completely specified.

A generalized Kolmogorov goodness-of-fit test of H_0 rejects H_0 whenever

$$(1.3) \quad F_n(x) > G_1(x) \quad \text{or} \quad F_n(x) < G_2(x), \text{ some } x, -\infty < x < \infty,$$

where $G_1(x)$ and $G_2(x)$ are arbitrary functions. Note that if, say, $G_2(x) = -1$, all x, then (1.3) is equivalent to the rejection region

$$(1.4a) \qquad F_n(x) > G_1(x), \text{ some } x, -\infty < x < \infty,$$

while if, say, $G_1(x) = 2$, all x, then (1.3) is equivalent to the rejection region

$$(1.4b) \qquad F_n(x) < G_2(x), \text{ some } x, -\infty < x < \infty.$$

That is, one-sided rejection regions of the form (1.4a) or (1.4b) are special cases of the general rejection region (1.3).

It is not hard to show that the rejection regions of the one-sided Kolmogorov tests D^+, D^-, the (two-sided) Kolmogorov-Smirnov test D, both the one-sided and two-sided weighted Kolmogorov tests (such as the

Anderson-Darling test), and Pyke's modifications of the tests D^+, D^- and D are all special cases of the general region (1.3). For example, the Kolmogorov-Smirnov test has rejection region of the form (1.3) with

$$G_i(x) = (-1)^i\lambda + F^*(x), \quad i = 1,2, \lambda > 0,$$

while the two-sided Anderson-Darling test has

$$G_i(x) = (-1)^i[F^*(x)(1-F^*(x))]^{\frac{1}{2}} + F^*(x), \quad i = 1,2, \lambda > 0.$$

There is an extensive literature which deals with exact (finite sample) power calculations for weighted and unweighted Kolmogorov tests. Two useful reviews of this literature are Durbin (1973) and Kendall and Stuart (1979; Chapter 30). Except for a paper by Conover (1972), methods for calculating the power of such tests against a c.d.f. F(x) are given only for problems in which F(x) is a continuous function of x. Conover (1972) gives a method for calculating the exact power function of the Kolmogorov tests D^+, D^- against possibly discontinuous alternatives F(x), but does not indicate how to extend his method to apply to other Kolmogorov-type tests.

Let

(1.5) $\rho(F) = P\{F_n(x) > G_1(x) \text{ or } F_n(x) < G_2(x), \text{ some } x, -\infty < x < \infty\}$

be the power of the test with rejection region (1.3) when the true c.d.f. for X is F(x). Note that if

$$\inf_{-\infty<x<\infty} G_1(x) < 0 \quad \text{or} \quad \sup_{-\infty<x<\infty} G_2(x) > 1,$$

the region (1.3) is the entire sample space, and $\rho(F) = 1$, all F. Thus, to avoid such trivial cases, it is assumed in the remainder of this paper that the functions $G_1(x)$, $G_2(x)$ in (1.3) satisfy:

(1.6) $$\inf_{-\infty<x<\infty} G_1(x) \geq 0, \qquad \sup_{-\infty<x<\infty} G_2(x) \leq 1.$$

In consequence, $G_1(x)$ is bounded below, while $G_2(x)$ is bounded above.

The main result of this paper is the following theorem.

Theorem 1.1. Assume that

(1.7) $$\inf_{x\leq z<\infty} G_1(z) = \underline{G}_1(x) \text{ is a right-continuous function of } x.$$

Define the extended real-valued constants

(1.8) $$a_i = \inf\{x: \inf_{z\geq x} G_1(z) \geq \frac{i}{n}\}, \quad b_i = \sup\{x: \sup_{z<x} G_2(z) < \frac{i-1}{n}\},$$

for $i = 1,2,\ldots,n$. Then for any c.d.f. $F(x)$, whether continuous or discontinuous,

(1.9) $$\rho(F) = P\{U_{(i)} < F(a_i-) \text{ or } U_{(i)} > F(b_i), \text{ some } i = 1,2,\ldots,n\},$$

where

$$F(x-) = \lim_{\substack{z<x \\ z\to x}} F(z) = P\{X < x\}$$

and $0 \leq U_{(1)} \leq \cdots \leq U_{(n)} \leq 1$ are distributed as the order statistics from a sample of n independent $U_{[0,1]}$ random variables.

In the special case where $G_1(x)$, $G_2(x)$ are continuous nondecreasing functions, and $F(x)$ is a continuous c.d.f., the representation (1.9) is well known. Consequently, a variety of methods for calculating

(1.10) $$R(s_1,\ldots,s_n;t_1,\ldots,t_n) = P\{s_i \leq U_{(i)} \leq t_i, \ 1 \leq i \leq n\}$$

appear in the literature. Theorem 1.1 says that any one of these methods can be used to obtain

$$\rho(F) = 1 - R(F(a_1-),\ldots,F(a_n-);F(b_1),\ldots,F(b_n))$$

even when $F(x)$ is discontinuous and $G_1(x)$, $G_2(x)$ are arbitrary functions restricted only by the conditions (1.6) and (1.7).

Using the equivalence between the one-sided regions (1.4a) and (1.4b) and the general region (1.3), the following useful consequence of Theorem 1.1 can be straightforwardly obtained.

Theorem 1.2. (i) Any test with rejection region of the form (1.4a), where $G_1(x)$ satisfies (1.7), has power function

$$\rho(F) = 1 - R(F(a_1-),\ldots,F(a_n-);1,1,\ldots,1),$$

where a_i, $1 \leq i \leq n$, is defined by (1.8). Any such test is unbiased for testing H_0 vs. one-sided alternatives

$$H_1^+: F(x) \geq F^*(x), \text{ all } x.$$

(ii) Any test with rejection region of the form (1.4b) has power function

$$\rho(F) = 1 - R(0,0,\ldots,0;F(b_1),\ldots,F(b_n)),$$

where b_i, $1 \leq i \leq n$, is defined by (1.8). Any such test is unbiased for testing H_0 vs. one-sided alternatives

$$H_1^-: F(x) \leq F^*(x), \text{ all } x.$$

Theorem 1.1 is proved in two steps. In Section 2, it is shown (Theorem 2.5) that, subject to the regularity conditions (1.6), (1.7), every test of H_0 with rejection region of the form (1.3) is equivalent to a test of H_0 with rejection region

$$X_{(i)} < a_i \text{ or } X_{(i)} > b_i, \text{ some } i = 1,2,\ldots,n, \tag{1.11}$$

where $a_1,\ldots,a_n,b_1,\ldots,b_n$ are defined by (1.8), and where

$$-\infty < X_{(1)} \leq \cdots \leq X_{(n)} < \infty$$

are the order statistics from the sample $X_1,\ldots,X_n$. This equivalence is well known in the case where $G_1(x), G_2(x)$ are both continuous nondecreasing functions, but does not seem to have been proved in the generality given here.

In Section 3, the proof of Theorem 1.1 is completed by showing (Theorem 3.1) that every test with rejection region (1.11) has power $\rho(F)$ given by (1.9). When the c.d.f. $F(x)$ is continuous, this assertion is known to be a direct consequence of the probability integral transformation. Interestingly, for general c.d.f.'s, the result (1.9) is a fairly straightforward consequence of the <u>inverse</u> probability integral transformation.

Finally, Section 4 demonstrates that no new generality can be introduced by expanding the class of generalized Kolmogorov tests to include tests which reject H_0 when

(1.12) $\psi_1(F_n(x)) > W_1(x)$ or $\psi_2(F_n(x)) < W_2(x)$, some x, $-\infty < x < \infty$,

where $\psi_1(u)$, $\psi_2(u)$ are nondecreasing functions of u, $0 \leq u \leq 1$, and $W_1(x)$, $W_2(x)$ are arbitrary functions of x, $-\infty < x < \infty$.

2. <u>Equivalence of Generalized Kolmogorov Tests and Tests Based on Order Statistics</u>. For any two real-valued functions $T_1(x), T_2(x)$ on the real line, let

(2.1)
$$\Delta(T_1, T_2) = \sup_{-\infty<x<\infty} \{T_1(x) - T_2(x)\}.$$

In terms of $\Delta(T_1, T_2)$, the rejection region (1.3) can be equivalently restated in the form: Reject H_0 if

(2.2)
$$\Delta(F_n, G_1) > 0 \text{ or } \Delta(G_2, F_n) > 0.$$

Lemma 2.1. Let U(x) be any nondecreasing function of x. For any function V(x) define

(2.3) $$\underline{V}(x) = \inf_{x\le z<\infty} V(z)\ ,\quad \bar{V} = \sup_{-\infty<z\le x} V(z),\ -\infty < x < \infty .$$

If V(x) is bounded below, $\underline{V}(x)$ is a real-valued function of x, while if V(x) is bounded above, $\bar{V}(x)$ is real-valued. Then

(2.4) $$\Delta(U,V) = \Delta(U,\underline{V})\ ,\quad \Delta(V,U) = \Delta(\bar{V},U).$$

Proof. Note that $\underline{V}(x) \le V(x) \le \bar{V}(x)$. It follows from (2.1) that

(2.5) $$\Delta(U,V) \le \Delta(U,\underline{V})\ ,\quad \Delta(V,U) \le \Delta(\bar{V},U).$$

On the other hand, since U(x) is nondecreasing in x,

$$\Delta(U,\underline{V}) = \sup_{-\infty<x<\infty}\{U(x) - \inf_{x\le z<\infty} V(z)\} = \sup_{-\infty<x\le z<\infty}\{U(x) - V(z)\}$$

$$\le \sup_{-\infty<x\le z<\infty}\{U(z) - V(z)\} = \Delta(U,V),$$

and

$$\Delta(\bar{V},U) = \sup_{-\infty<x<\infty}\{\sup_{-\infty<z\le x} V(z) - U(x)\} \le \sup_{-\infty<z\le x<\infty}\{V(z) - U(z)\} = \Delta(V,U).$$

From these two inequalities and (2.5), the result (2.4) follows. □

Corollary 2.2. Every generalized Kolmogorov test with rejection region (1.3) defined by arbitrary functions $G_1(x),G_2(x)$ is equivalent to a generalized Kolmogorov test with the rejection region

(2.6) $$F_n(x) > \underline{G}_1(x) \text{ or } F_n(x) < \bar{G}_2(x), \text{ some } x,\ -\infty < x < \infty ,$$

defined by the nondecreasing, real-valued functions

(2.7) $$\underline{G}_1(x) = \inf_{x\le z<\infty} G_1(z)\ ,\quad \bar{G}_2(x) = \sup_{-\infty<z\le x} G_2(z).$$

Proof. The functions $\underline{G}_1(x),\bar{G}_2(x)$ are real-valued because by (1.6), $G_1(x)$ is bounded below and $G_2(x)$ is bounded above. Note that the rejection region (2.6) is equivalent to

(2.8) $$\Delta(F_n,\underline{G}_1) > 0 \text{ or } \Delta(\bar{G}_2,F_n) > 0.$$

Since $F_n(x)$ is a nondecreasing function, Lemma 2.1 applies to show that $\Delta(F_n,G_1) = \Delta(F_n,\underline{G}_1), \Delta(G_2,F_n) = \Delta(\bar{G}_2,F_n)$. Thus, (2.2) and (2.8) are equivalent regions, proving that (1.3) and (2.6) are equivalent. □

Recall that for any (real-valued) nondecreasing function $V(x)$, the limits

(2.9) $$V(x-) = \lim_{\substack{z<x \\ z\to x}} V(z) = \sup_{z<x} V(z),\quad V(x+) = \lim_{\substack{z>x \\ z\to x}} V(z) = \inf_{z>x} V(z)$$

are well defined. The function $V(x)$ is right-continuous if $V(x+) = V(x)$, all x. For any c, $-\infty < c < \infty$, define

(2.10) $$V^{-1}(c) = \inf\{x: V(x) \geq c\}.$$

Note that $V^{-1}(c) = -\infty$ if and only if $V(x) \geq c$, all x, or, equivalently, if and only if $\inf_{-\infty<x<\infty} V(x) \geq c$. Also $V^{-1}(c) = \infty$ if and only if $\lim_{x\to\infty} V(x) = \sup_{-\infty<x<\infty} V(x) < c$.

<u>Lemma 2.3</u>. For any (real-valued) nondecreasing function $V(x)$, the <u>inverse function</u> $V^{-1}(c)$, defined for $-\infty < c < \infty$ by (2.10), has the following properties:

(i) $V^{-1}(c)$ is non-decreasing in c,

(ii) $V^{-1}(c+) = \sup\{x: V(x-) \leq c\} = \inf\{x: V(x) > c\}$,

(iii) $x > V^{-1}(c+) \Leftrightarrow V(x-) > c$,

(iv) $x < V^{-1}(c) \Rightarrow V(x) < c \Rightarrow x \leq V^{-1}(c)$,

while if $V(x)$ is right continuous,

(v) $x < V^{-1}(c) \Leftrightarrow V(x) < c$, and

$$\text{(vi)}\quad V(V^{-1}(c)) \geq c, \text{ all } c, -\infty < c < \infty .$$

Proof. Properties (i) and (ii) follow directly from the definition (2.10) of $V^{-1}(c)$. To prove property (ii), let

$$T_1(c) = \sup\{x: V(x-) \leq c\},\ T_2(c) = \inf\{x: V(x) > c\}.$$

Since $V(x)$ is nondecreasing in x,

$$x < T_2(c) \Rightarrow V(x-) \leq V(x) \leq c \Rightarrow x \leq T_1(c),$$

while

$$x < T_1(c) \Rightarrow V(x-) = \sup_{y<x} V(y) \leq c \Rightarrow y < T_2(c), \text{ all } y < x \Rightarrow x \leq T_2(c) .$$

This shows that $T_1(c) = T_2(c)$. However,

$$x < T_2(c) \Rightarrow V(x) \leq c \Leftrightarrow V(x) < d, \text{ all } d > c \Leftrightarrow x \leq \inf_{d>c} V^{-1}(c) = V^{-1}(c+),$$

and also

$$x < V^{-1}(c+) \Rightarrow V(x) < d, \text{ all } d > c \Rightarrow V(x) \leq c \Rightarrow x \leq T_2(c).$$

Thus, $V^{-1}(c+) = T_2(c) = T_1(c)$, proving property (ii). It directly follows from property (ii) that

$$x > V^{-1}(c+) = T_1(c) \Leftrightarrow V(x-) > c,$$

since

$$V(T_1(c)-) = \sup_{x<T_1(c)} V(x) \leq c$$

by definition of $T_1(c)$. This verifies property (iii). Finally, property (v) follows from properties (iv) and (vi), while property (vi) is a consequence of the inequality

$$V(V^{-1}(c)+) = \inf_{x>V^{-1}(c)} V(x) \geq c$$

and the right-continuity of $V(x)$ at $x = V^{-1}(c)$. □

Lemma 2.4. Assume that

$$U(x) = \begin{cases} u_0, & \text{if } -\infty < x < x_1, \\ u_i, & \text{if } x_i \leq x < x_{i+1},\ 1 \leq i \leq k-1, \\ u_k, & \text{if } x_k \leq x < \infty, \end{cases}$$

is a real-valued, right-continuous step function. Then for any real-valued nondecreasing function $V(x)$

(2.11) $$\Delta(U,V) = \max\{(u_0 - \inf_{-\infty<x<\infty} V(x)), \max_{1\leq i\leq k} (u_i - V(x_i))\}$$

(2.12) $$\Delta(V,U) = \max\{\max_{0\leq i\leq k-1} (V(x_{i+1}-) - u_i), (\sup_{-\infty<x<\infty} V(x) - u_k)\}.$$

Proof. Straightforward, taking the supremum in the definition (2.1) over the intervals $(-\infty,x_1),[x_1,x_2),\ldots,[x_k,\infty)$, and using the fact that $V(x)$ is nondecreasing. □

Theorem 2.5. Consider any rejection region for H_0 of the form: Reject H_0 if

(2.13) $$F_n(x) > G_1(x) \text{ or } F_n(x) < G_2(x), \text{ some } x, -\infty < x < \infty .$$

Assume that $G_1(x)$ and $G_2(x)$ satisfy (1.6), and also that

(2.14) $\underline{G}_1(x) = \inf_{x\leq z<\infty} G_1(z)$ is right-continuous in x, $-\infty < x < \infty$.

Define

$$a_i = \underline{G}_1^{-1}(\tfrac{i}{n}) = \inf\{x: \inf_{z\geq x} G_1(z) \geq \tfrac{i}{n}\},$$

$$b_i = \bar{G}_2^{-1}(\tfrac{i-1}{n}+) = \sup\{x: \sup_{z<x} G_2(z) \leq \tfrac{i-1}{n}\},\ i = 1,2,\ldots,n.$$

Then the rejection region (2.13) is equivalent to the rejection region

(2.16) $$X_{(i)} < a_i \text{ or } X_{(i)} > b_i, \text{ some } i = 1,2,\ldots,n,$$

where $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ are the order statistics based on the sample $X_1, X_2, \ldots, X_n$.

Proof. Use the fact that the sample c.d.f. $F_n(x)$ is a nondecreasing, right-continuous step function with jumps at the order statistics $X_{(i)}$, $1 \leq i \leq n$. Apply Corollary 2.2, Lemma 2.4 (with $u_0 = 0$, $u_k = 1$) and (1.6) to show that the region (2.13) is equivalent to the region

(2.17) $$\max_{1 \leq i \leq n} \left(\frac{i}{n} - \underline{G}_1(X_{(i)})\right) > 0 \text{ or } \max_{1 \leq i \leq n} \left(\bar{G}_2(X_{(i)}-) - \frac{i-1}{n}\right) > 0.$$

Next, apply Lemma 2.3 (v) and the definition of a_i, $1 \leq i \leq n$, to show that

$$\max_{1 \leq i \leq n} \left(\frac{i}{n} - \underline{G}_1(X_{(i)})\right) > 0 \Leftrightarrow X_{(i)} < a_i, \text{ some } i = 1,2,\ldots,n,$$

and then apply (2.14), Lemma 2.3(ii),(iii) and the definition of b_i, $1 \leq i \leq n$, to show that

$$\max_{1 \leq i \leq n} \left(\bar{G}_2(X_{(i)}-) - \frac{i-1}{n}\right) > 0 \Leftrightarrow X_{(i)} > b_i, \text{ some } i = 1,2,\ldots,n.$$

Thus, the region (2.16) is equivalent to the region (2.17), which in turn is equivalent to (2.13). This completes the proof. □

Remark 1. It can be shown that (2.14) holds if and only if

(2.18) $$G_1(x) \geq G_1(x+), \text{ all } x, -\infty < x < \infty .$$

The condition (2.18) holds if $G_1(x)$ is right-continuous. Of course, if $G_1(x)$ is nondecreasing, then $\underline{G}_1(x) = G_1(x)$, and (2.14) and (2.18) hold if and only if $G_1(x)$ is right-continuous.

Remark 2. The constants a_i, b_i defined by (2.15) can equal ∞ or $-\infty$. It is easily shown from Lemma 2.3(i) and (2.15) that

(2.19) $$-\infty \leq a_1 \leq a_2 \leq \cdots \leq a_n \leq \infty, -\infty \leq b_1 \leq b_2 \leq \cdots \leq b_n \leq \infty .$$

It should also be clear that the region (2.16), and thus the region (2.13), is the entire sample space if

$$b_i < a_i, \text{ some } i = 1,2,\ldots,n.$$

Theorem 2.5 shows that under some very mild conditions, (1.6) and (2.14), on the functions $G_1(x)$ and $G_2(x)$, every generalized Kolmogorov goodness-of-fit test (1.3) is equivalent to a test based on the order statistics $X_{(1)} \leq \cdots \leq X_{(n)}$ with rejection region of form (2.16). The converse of this result is also true. Indeed, let $a_1,a_2,\ldots,a_n,b_1,b_2,\ldots,b_n$ be any extended real-valued constants satisfying (2.19). [See Remark 2 above.] Define

(2.20) $$Q_1(x) = \frac{\#a_i\text{'s} \leq x}{n}, \quad Q_2(x) = \frac{\#b_i\text{'s} \leq x}{n}, \quad -\infty < x < \infty .$$

Note that $Q_1(x)$ and $Q_2(x)$ are real-valued, nondecreasing, right-continuous step-functions for which

$$a_i = Q_1^{-1}\left(\frac{i}{n}\right), \quad b_i = Q_2^{-1}\left(\frac{i-1}{n}+\right), \quad 1 \leq i \leq n,$$

holds. [These results hold even when some of the a_i's, or b_i's, are equal, or when some a_i's or b_i's equal ∞ or $-\infty$.] The following converse to Theorem 2.5 (see also Remark 2) has thus been established.

Theorem 2.6. For every region of the form (2.16), defined by constants $a_i, b_i, 1 \leq i \leq n$, satisfying (2.19), there exists a region

$$F_n(x) > Q_1(x) \text{ or } F_n(x) < Q_2(x), \text{ some } x, -\infty < x < \infty$$

of the form (2.13), where $Q_1(x), Q_2(x)$ are nondecreasing, right-continuous step-functions satisfying (1.6).

3. Representation for the Power Function (1.5). It is clear from Theorem 2.5 that Theorem 1.1 will be established once the following result is shown to hold.

Theorem 3.1. Let $X_{(1)} \leq \cdots \leq X_{(n)}$ be the order statistics from a sample, $X_1,\ldots,X_n$, of i.i.d. random variables having a common c.d.f. $F(x)$. Let

$$0 \leq U_{(1)} \leq \cdots \leq U_{(n)} \leq 1$$

be the order statistics from a sample $U_1, U_2, \ldots, U_n$ of independent random variables uniformly distributed on [0,1]. Then, for any extended real-valued constants $a_i, b_i, -\infty \leq a_i, b_i \leq \infty, 1 \leq i \leq n$,

$$(3.1) \qquad P\{a_i \leq X_{(i)} \leq b_i, 1 \leq i \leq n\} = P\{F(a_i-) \leq U_{(i)} \leq F(b_i), 1 \leq i \leq n\} ,$$

where it is understood that $F((-\infty)-) = F(-\infty) = 0$, $F(\infty-) = F(\infty) = 1$.

Proof. When $F(x)$ is continuous, (3.1) is a direct consequence of the probability integral transformation. In general, define the inverse probability integral transformation:

$$(3.2) \qquad F^{-1}(u) = \inf\{x: F(x) \geq u\}, \; 0 \leq u \leq 1.$$

It is well known that the random variables

$$\tilde{X}_i = F^{-1}(U_i), \; 1 \leq i \leq n,$$

have the same joint distribution as $X_1,\ldots,X_n$. Further, since $F^{-1}(u)$ is, by Lemma 2.3(i), nondecreasing in u,

$$\tilde{X}_{(i)} = F^{-1}(U_{(i)}), \; 1 \leq i \leq n,$$

where $\tilde{X}_{(1)} \leq \cdots \leq \tilde{X}_{(n)}$ are the order statistics formed from $\tilde{X}_1,\ldots,\tilde{X}_n$.

Consequently,

$$(3.3)\qquad P\{a_i \le X_{(i)} \le b_i, 1 \le i \le n\} = P\{a_i \le \tilde{X}_{(i)} = F^{-1}(U_{(i)}) \le b_i, 1 \le i \le n\}.$$

From Lemma 2.3(v), for every i = 1,2,...,n,

$$b_i < F^{-1}(U_{(i)}) \Leftrightarrow F(b_i) < U_{(i)},$$

or, equivalently,

$$(3.4)\qquad F^{-1}(U_{(i)}) \le b_i \Leftrightarrow U_{(i)} \le b_i.$$

By Lemma 2.3(i), $F^{-1}(u+) \ge F^{-1}(u)$, all u. Thus, by Lemma 2.3(iii),

$$(3.5)\qquad a_i \le F^{-1}(U_{(i)}) \Rightarrow F^{-1}(U_{(i)}+) \ge a_i \Rightarrow U_{(i)} \ge F(a_i-)$$

while by the definitions of $F(x-)$ and $F^{-1}(u)$,

$$(3.6)\qquad \begin{aligned} U_{(i)} > F(a_i-) &\Rightarrow U_{(i)} > F(y), \text{ all } y < a_i \\ &\Rightarrow y \le F^{-1}(U_{(i)}), \text{ all } y < a_i \\ &\Rightarrow a_i \le F^{-1}(U_{(i)}), \end{aligned}$$

all i = 1,2,...,n. From (3.4), (3.5) and (3.6),

$$(3.7)\qquad \begin{aligned} P\{F(a_i-) < U_{(i)} \le F(b_i), 1 \le i \le n\} &\le P\{a_i \le F^{-1}(U_{(i)}) \le b_i, 1 \le i \le n\} \\ &\le P\{F(a_i-) \le U_{(i)} \le F(b_i), 1 \le i \le n\}, \end{aligned}$$

but since the random variables $U_{(i)}$, $1 \le i \le n$, have a continuous joint distribution, the left-hand and right-hand sides of (3.7) are equal, proving (3.1). □

Although the representation provided by (3.1) is clearly useful, it should be noted that the method of proof of Theorem 3.1 is also of considerable applicability. In many cases, the inverse probability integral transformation, used in conjunction with Lemma 2.3, can provide distributional results (or inequalities) for "nonparametric" procedures in contexts where the c.d.f. $F(x)$ of the data is discontinuous. In comparison to the "projection" method described in Noether (1967; Section 3.3), the method here applies to mixed discrete-continuous cases, as well as to purely discrete populations.

4. <u>Generalizations</u>. Consider the test of H_0 with rejection region: Reject H_0 if

$$(4.1) \qquad \sin^{-1}[(F_n(x))^{\frac{1}{2}}] > \sin^{-1}[(F^*(x))^{\frac{1}{2}}] + \lambda, \text{ some } x, -\infty < x < \infty ,$$

where $\lambda > 0$. Such a test is appropriate for testing H_0 vs. the one-sided alternatives H_1^+ described in Theorem 1.2. Recall that for each x, $Y_n(x) = nF_n(x)$ has a binomial distribution with parameters n and $p = F(x)$. The test defined by (4.1) corresponds to simultaneously testing H_{0x}: $F(x) = F^*(x)$ for all x, using the large sample tests based on the variance-stabilizing transformation of the binomial distribution. The use of the variance-stabilizing transformation here is an attempt to make the large-sample null distributions of the test statistics

$$\sin^{-1}[(F_n(x))^{\frac{1}{2}}] - \sin^{-1}[(F^*(x))^{\frac{1}{2}}]$$

equal for all x.

A test which appears to be equivalent to the test defined by (4.1) is the test with rejection region: Reject H_0 if

$$(4.2) \qquad F_n(x) > \sin^2\{\lambda + \sin^{-1}[(F^*(x))^{\frac{1}{2}}]\}, \text{ some } x, -\infty < x < \infty .$$

Knott (1970) attributes this last test to J. W. Tukey. However, it is more likely that Tukey proposed the test based on (4.1). Indeed, using Corollary 2.2 it can be shown that (4.2) is equivalent to

$$F_n(x) > \min\{\sin^2(\lambda + \sin^{-1}[(F^*(x))^{\frac{1}{2}}]), \sin^2(\frac{\pi}{2}+\lambda)\}, \text{ some } x.$$

Since for $\lambda > 0$, $\sin^2(\frac{1}{2}\pi+\lambda) < 1$, while $F_n(x) \to 1$, as $x \to \infty$, it is apparent that the region (4.2) is the entire sample space, and thus does not define a reasonable rejection region for H_0.

There is a test with rejection region of the form (1.4a) which is equivalent to the test based on (4.1). However, the equivalent test is not that based on (4.2), but instead is based on the rejection region: Reject H_0 if

(4.3) $$F_n(x) > \sin^2\{\lambda + \sin^{-1}[(F^*(x))^{\frac{1}{2}}]\}, \text{ some } x, -\infty < x \leq x^*,$$

where $x^* = F^{*-1}(\sin^2(\frac{1}{2}\pi+\lambda))$.

The equivalence between the rejection regions (4.1) and (4.3) is a special case of the following general result.

<u>Theorem 4.1</u>. Let $\psi_1(u),\psi_2(u)$ be nondecreasing functions of u, $0 \leq u \leq 1$, and let $W_1(x),W_2(x)$ be arbitrary functions of x, $-\infty < x < \infty$. For every test of H_0 with rejection region

(4.4) $$\psi_1(F_n(x)) > W_1(x) \text{ or } \psi_2(F_n(x)) < W_2(x), \text{ some } x, -\infty < x < \infty,$$

there is an equivalent test with rejection region of the form (1.3).

<u>Proof</u>. A constructive proof will be given. First, observe that since the only possible values of $F_n(x)$ are the rational fractions i/n, $i = 0,1,2,\ldots,n$, the values of $\psi_j(u)$, $j = 1,2$, for $u \neq i/n$, some $i = 0,1,\ldots,n$, do not affect the occurrence or non-occurrence of the event (4.4). Hence, without loss of generality, replace $\psi_j(u)$, $j = 1,2$, in (4.4) by the

unique continuous, piecewise-linear functions $\tilde{\psi}_j(u)$ having the property that

$$\tilde{\psi}_j(\tfrac{i}{n}) = \psi_j(\tfrac{i}{n}),\ 0 \le i \le n,\ j = 1,2.$$

Since the $\psi_j(u)$'s are nondecreasing functions, so are the $\tilde{\psi}_j(u)$'s.

Next, note that unless

$$\text{(4.5)} \qquad \inf_{-\infty<x<\infty} W_1(x) \ge \psi_1(0),\ \sup_{-\infty<x<\infty} W_2(x) \le \psi_2(1),$$

the region (4.4) is the entire sample space (and is thus equivalent to any region (1.3) for which the functions $G_1(x), G_2(x)$ fail to satisfy (1.6)). Consequently, it can be assumed without loss of generality that (4.5) holds. Then, letting the functions $\tilde{W}_1(x), \tilde{W}_2(x)$ be defined by

$$\text{(4.6)} \qquad \tilde{W}_1(x) = \min[\psi_1(1), W_1(x)],\ \tilde{W}_2(x) = \max[\psi_2(0), W_2(x)],$$

it is easily seen that (4.4) is equivalent to

$$\text{(4.7)} \quad \tilde{\psi}_1(F_n(x)) > \tilde{W}_1(x) \text{ or } \tilde{\psi}_2(F_n(x)) < \tilde{W}_2(x),\ \text{some } x,\ -\infty < x < \infty .$$

However, since the range of $\tilde{W}_i(x)$ is included in the range of $\tilde{\psi}_i(u)$, $i = 1,2$, the following functions are well defined.

$$\text{(4.8)} \qquad \begin{aligned} \tilde{G}_1(x) &= \tilde{\psi}_1^{-1}(\tilde{W}_1(x)+) = \inf\{u\colon 0 \le u \le 1, \tilde{\psi}_1(u) > \tilde{W}_1(x)\}, \\ \tilde{G}_2(x) &= \tilde{\psi}_2^{-1}(\tilde{W}_2(x)) = \inf\{u\colon 0 \le u \le 1, \tilde{\psi}_2(u) \ge \tilde{W}_2(x)\} . \end{aligned}$$

It now follows from (4.8), from the fact that $\tilde{\psi}_1(u), \tilde{\psi}_2(u)$ are nondecreasing, continuous functions (so that $\tilde{\psi}_2$ is right-continuous, while $\tilde{\psi}_1(x-) = \tilde{\psi}_1(x)$, all x), and from Lemma 2.3, properties (iii) and (v), that (4.7) and

$$\text{(4.9)} \qquad F_n(x) > \tilde{G}_1(x) \text{ or } F_n(x) < \tilde{G}_2(x),\ \text{some } x,\ -\infty < x < \infty ,$$

are equivalent regions. Thus, (4.9) is equivalent to (4.4). Since (4.9) is of the form (1.3), the proof is complete. □

Acknowledgement. This paper was begun while the author was visiting The Sidney Farber Cancer Institute and the Department of Biostatistics, Harvard University School of Public Health. Thanks to Drs. Marvin Zelen and Frederick Mosteller for the financial support that made that visit possible. The author is also grateful to Drs. Prem Goel and Herman Rubin for comments that led to improvements in the proof of Theorem 3.1.

REFERENCES

[1] Conover, W. J. (1972). A Kolmogorov goodness-of-fit test for discontinuous distributions. J. Amer. Statist. Assoc. **67**, 591-596.

[2] Durbin, J. (1973). Distribution Theory for Tests Based on the Sample Distribution Function. SIAM, Philadelphia.

[3] Kendall, M. G. and Stuart, A. (1979). The Advanced Theory of Statistics, Vol. II, 4th ed., Macmillan, N.Y.

[4] Knott, M. (1970). The small sample power of one-sided Kolmogorov tests for a shift in location of the normal distribution. J. Amer. Statist. Assoc. **65**, 1384-1391.

[5] Noether, Gottfried E. (1967). Elements of Nonparametric Statistics. John Wiley & Sons, New York, London, Sydney.

FOOTNOTES

1. Research supported by the National Science Foundation under grant MCS-79-05815.

AMS 1970 subject classifications. Primary 62G10, 62E15.

Key words and phrases. Goodness-of-fit tests, Kolmogorov tests, discontinuous distributions, exact power function, order statistics, inverse probability integral transformation, one-sided tests, unbiased tests.