BU-594-M
Rev. 7/1977

ON NULL-HYPOTHESIS LIMITING DISTRIBUTIONS OF KOLMOGOROV-SMIRNOV TYPE STATISTICS WITH ESTIMATED LOCATION AND SCALE PARAMETERS

By Constance L. Wood

Research supported in part by Hatch Project 402, Biometrics Unit, Cornell University.

Key words and phrases. Limit distributions, Kolmogorov-Smirnov statistics, estimated parameters, normality, completely randomized designs.

Summary

ON NULL-HYPOTHESIS LIMITING DISTRIBUTIONS OF KOLMOGOROV-SMIRNOV TYPE STATISTICS WITH ESTIMATED LOCATION AND SCALE PARAMETERS

We treat the "goodness of fit" problem of testing whether the distribution F of a random sample $X_1, \cdots, X_n$ belongs to a specified location and scale family, with the particular values (α, β) of the location and scale parameters not both specified. A foundation for testing such a composite hypotheis is provided by an associated "empirical" stochastic process, based upon the sample distribution function corresponding to $X_1, \cdots, X_n$ and upon estimates $(\hat{\alpha}_n, \hat{\beta}_n)$ of (α, β). In particular, statistics of Kolmogorov-Smirnov type may be represented as well-behaved functionals of the empirical process. It follows that the limit distributions of such test statistics are given by applying the corresponding functionals to the limit in distribution of the empirical process, provided that such a convergence result holds. We prove such a convergence theorem for the behavior of the empirical process under the null hypothesis. The result overlaps with theorems of Durbin (1973) and Neuhaus (1976), but presents conditions which are more readily verified. Regarding the test statistics, the asymptotic distributions so obtained serve as a basis for Monte Carlo studies for determination of appropriate critical points. These methods are also used to derive a "large-sample" test for normality for Completely Randomized Designs.

1. Introduction. Testing "goodness of fit" is a standard statistical problem having broad application. Here we consider the case of testing whether the distribution F of a random sample $X_1, \cdots, X_n$ belongs to a specified location and scale family, with the particular values of the location and scale parameters *not both known*. More precisely, denoting by G the class of c.d.f.'s G^* of the form

$$G^*(x) = G\left(\frac{x - \alpha}{\beta}\right), \quad \text{all } x,$$

where G is a specified *continuous* c.d.f. and Θ is a specified set of possible pairs $\theta = (\alpha,\beta)$, the "null-hypothesis" to be tested is

$$H_0\text{: } F(x) = G\left(\frac{x - \alpha}{\beta}\right), \quad \text{all } x, \text{ for some } (\alpha,\beta) \in \Theta .$$

It is assumed that Θ contains more than one element (α,β), so that H_0 is a *composite* hypothesis.

A foundation for testing H_0 is provided by an associated "empirical" stochastic process based upon the sample distribution function corresponding to $X_1, \cdots, X_n$ and upon estimates $(\hat{\alpha}_n,\hat{\beta}_n)$ of (α,β). A variety of relevant test statistics may be represented as well-behaved functionals of this empirical process. In particular, we shall consider statistics of Kolmogorov-Smirnov type. It will be shown that under H_0 the empirical process converges in distribution (in the sense of weak convergence of probability measures on a suitable metric space of functions) to a certain limit stochastic process, of Gaussian type. It follows that the asymptotic distributions of the relevant test statistics may be obtained as the distributions of the corresponding functionals of the limit process. The actual computation of such limit distributions is another problem, not dealt with here.

The results just described in connection with the composite hypothesis H_0 extend well-known results for the case of H_0 *simple*, i.e., where Θ consists of a single element (α,β). We shall review this case as a preliminary to precise formulation of the extended results.

For H_0 simple, and thus (α,β) specified, it is equivalent to test the hypothesis that the independent random variables

$$G\left[(X_i - \alpha)/\beta\right], \quad 1 \leq i \leq n,$$

have common uniform (0,1) distribution. The popular Kolmogorov-Smirnov test is based on the statistic

$$n^{\frac{1}{2}} \sup_{0\leq t\leq 1} |G_n(t) - t|, \tag{1.1}$$

where

$$G_n(t) = \frac{1}{n} \sum_{i=1}^{n} I\left(G\left[(X_i - \alpha)/\beta\right] \leq t\right), \quad 0 \leq t \leq 1, \tag{1.2}$$

and $I(E)$ denotes the indicator of the event E. In terms of the "empirical" stochastic process

$$W_n(t) = n^{\frac{1}{2}}[G_n(t) - t], \quad 0 \leq t \leq 1, \tag{1.3}$$

the statistic (1.1) is given by the functional $\sup_{0\leq t\leq 1}|x(t)|$ applied to $x(\cdot) = W_n(\cdot)$. Under H_0 the empirical process $W_n(\cdot)$ satisfies

$$W_n \xrightarrow{d} W^\circ \quad \text{in } \mathcal{D}[0,1], \tag{1.4}$$

where $\mathcal{D}[0,1]$ denotes the space of functions $x(\cdot)$ on $[0,1]$ which are right-continuous and have left-hand limits, $\xrightarrow{d}$ denotes convergence in distribution, and W° denotes the "tied-down Wiener" process on $[0,1]$, namely, the Gaussian process determined by

$$E[W^\circ(t)] = 0, \quad 0 \leq t \leq 1, \tag{1.5}$$

and

$$E[W^\circ(s)W^\circ(t)] = \min(s,t) - st, \quad 0 \leq s,t \leq 1. \tag{1.6}$$

(The convergence (1.4) is proved as Theorem 13.1 of Billingsley (1968).)

Returning to the case of H_0 composite, we consider an analogous approach based on estimates $(\hat{\alpha}_n, \hat{\beta}_n)$ for (α,β). Put

$$Y_{ni} = \frac{X_i - \hat{\alpha}_n}{\hat{\beta}_n}, \quad 1 \leq i \leq n,$$

and, by analogy with $G_n(\cdot)$ and $W_n(\cdot)$, define

$$H_n(t) = \frac{1}{n} \sum_{i=1}^{n} I\big(G(Y_{ni}) \leq t\big), \quad 0 \leq t \leq 1, \tag{1.7}$$

and

$$V_n(t) = n^{\frac{1}{2}}[H_n(t) - t], \quad 0 \leq t \leq 1. \tag{1.8}$$

We shall generalize (1.4) by proving that under H_0 the process V_n satisfies $V_n \xrightarrow{d} V^\circ$ in $\mathcal{D}[0,1]$, where V° is a Gaussian process that depends upon G and upon properties of the estimates $(\hat{\alpha}_n, \hat{\beta}_n)$. The precise formulation of V° is given in Section 2, along with assumptions on G and $(\hat{\alpha}_n, \hat{\beta}_n)$, and with a formal theorem stating this convergence result. The proof of the theorem will be given in Section 3.

A similar convergence theorem has been established by Durbin (1973), under narrower restrictions on the parameter set Θ and the estimates $(\hat{\alpha}_n, \hat{\beta}_n)$, but allowing parameters other than location and scale, and under local parametric alternatives to H_0 as well as under H_0. (These results were extended by Neuhaus (1976) to a wider class of local alternatives.) Such extensions of our results will be considered elsewhere. In particular, both authors require that an estimator $(\hat{\theta}_n)$ of an unknown parameter (θ) be of the form

$$n^{\frac{1}{2}}(\hat{\theta}_n - \theta) = n^{-\frac{1}{2}} \sum_{i=1}^{n} \ell(X_i, \theta) + o_p(1)$$

where $E\left[\ell(X,\theta)^2\right] < \infty$. In fact, this representation is required to evaluate the covariance function of the limiting Gaussian process. Even though many estimators can be given such a representation, $\ell(\cdot,\cdot)$ is often hard to determine in practice. In order to obtain the limiting Gaussian process we only require that, jointly with the sample c.d.f., the estimators have an asymptotic normal distribution, a condition usually known or readily verified; in which case only the asymptotic covariance matrix need be derived.

More importantly, the class of estimators considered previously exclude estimators which use auxillary information on the unknown parameters. In particular, considering the problem of testing for normality in the Completely Randomized Design, previous results would not allow mean square for error to be used as an estimator of experimental error for one particular treatment. An application of our results to this problem is given in the last section.

Some statistics of interest in connection with H_0 are the *one-sided Kolmogorov-Smirnov statistics*

$$D_n^+ = \sup_{0\le t\le 1} V_n(t) \tag{1.9}$$

and

$$D_n^- = \inf_{0\le t\le 1} V_n(t), \tag{1.10}$$

the *Kolmogorov-Smirnov statistic*

$$D_n = \max(D_n^+, -D_n^-), \tag{1.11}$$

and the *Kuiper statistic*

$$D_n^{\pm} = D_n^+ - D_n^- . \tag{1.12}$$

The corresponding functionals $\sup x(t)$, $\inf x(t)$, $\sup|x(t)|$, and $\sup x(t) - \inf x(t)$ are *continuous* with respect to the Skorokhod metric on $\mathcal{D}[0,1]$ (see Billingsley (1968), Section 18). Hence the limit laws of D_n^+, D_n^-, D_n and $D_n^{\pm}$ under H_0, and provided that G and $(\hat{\alpha}_n, \hat{\beta}_n)$ are such that $V_n \xrightarrow{d} V°$, are given, respectively, by the laws of the random variables

$$D^+ = \sup_{0\le t\le 1} V°(t), \tag{1.13}$$

$$D^- = \inf_{0\le t\le 1} V°(t), \tag{1.14}$$

(1.15) $$D = \max(D^+, -D^-),$$

and

(1.16) $$D^{\pm} = D^+ - D^- .$$

(These results are consequences of $V_n \xrightarrow{d} V^\circ$ by the continuous mapping theorem. See Billingsley (1968), Section 5.) The results provide a basis for Monte Carlo studies of the null hypothesis asymptotic distributions of the Kolmogorov-Smirnov type statistics D_n^+, D_n^-, D_n and $D_n^{\pm}$.

Monte Carlo simulations of these distributions with $G = \Phi$, the standard normal distribution function, are given in Section 4. In the last section, it is shown that these distributions also provide appropriate "large-sample" critical values for testing for normality in Completely Randomized Designs.

2. The convergence theorem. The various conditions to be imposed on G are given by

ASSUMPTIONS A.

(i) $G'(x)$ is continuous in x and positive on the support of G;

(ii) $xG'(x) \to 0$ as $|x| \to \infty$;

(iii) G'' is bounded.

We shall suppose that under H_0 the estimates $(\hat{\alpha}_n, \hat{\beta}_n)$ are asymptotically normal jointly with the empirical process W_n, i.e.,

ASSUMPTION B. Under H_0, for all $k = 1, 2, \cdots$ and $0 < t_1, \cdots, t_k < 1$, the vector

$$\left(\frac{\hat{\alpha}_n - \alpha}{\beta}, \frac{\hat{\beta}_n - \beta}{\beta}, W_n(t_1), \cdots, W_n(t_k)\right)$$

converges in distribution to multivariate normal with mean 0 and $(k + 2) \times (k + 2)$ covariance matrix $\Sigma(t_1, \cdots, t_k)$.

Under H_0 and Assumptions A and B, we may define the stochastic process

$$V^\circ(t) = W^\circ(t) + G'\left(G^{-1}(t)\right)Y + G'\left(G^{-1}(t)\right)G^{-1}(t)Z,$$

where (Y,Z) are jointly normal with the tied-down Wiener process W° in the sense of Assumption B. That is, V° is the Gaussian process determined by

$$E[V^\circ(t)] = 0, \quad 0 \leq t \leq 1, \tag{2.1}$$

and, for all $0 \leq s,t \leq 1$,

$$\begin{aligned} E[V^\circ(s)V^\circ(t)] = \min(s,t) - st &+ G'\left(G^{-1}(s)\right)G'\left(G^{-1}(t)\right)\sigma_{11} \\ &+ G^{-1}(s)G'\left(G^{-1}(s)\right)G^{-1}(t)G'\left(G^{-1}(t)\right)\sigma_{22} \\ &+ G'\left(G^{-1}(t)\right)\sigma_{13} + G'\left(G^{-1}(s)\right)\sigma_{14} \\ &+ G^{-1}(t)G'\left(G^{-1}(t)\right)\sigma_{23} + G^{-1}(s)G'\left(G^{-1}(s)\right)\sigma_{24} \\ &+ \left[G^{-1}(s)G'\left(G^{-1}(s)\right)G'\left(G^{-1}(t)\right) + G^{-1}(t)G'\left(G^{-1}(t)\right)G'\left(G^{-1}(s)\right)\right]\sigma_{12}, \end{aligned} \tag{2.2}$$

where $[\sigma_{ij}]_{4\times 4} = \Sigma(s,t)$.

THEOREM. *Assume that* G *satisfies Assumptions* A *and that* $(\hat{\alpha}_n, \hat{\beta}_n)$ *satisfies Assumption* B. *Then, under* H_0,

$$V_n \xrightarrow{d} V^\circ \quad \textit{in} \quad \mathcal{D}[0,1]. \tag{2.3}$$

REMARKS. (i) Note that the constants σ_{ij} appearing in (2.2) implicitly depend upon s,t.

(ii) Note that terms of (2.2) drop out if either α or β are known, making $\hat{\alpha}_n$ or $\hat{\beta}_n$ degenerate.

(iii) Note that V°, like W°, is continuous with probability 1.

3. The proof. Let all random variables $X_1, X_2, \cdots$ be defined on a probability space (Ω, A, P).

Note that $H_n(\cdot)$ may be written

$$(3.1) \qquad H_n(t) = G_n\left\{G[(\hat{\beta}_n/\beta)G^{-1}(t) + (\hat{\alpha}_n - \alpha)/\beta]\right\}, \quad 0 \le t \le 1,$$

where $G^{-1}(t) = \inf\{X : G(X) > t\}$ and thus

$$(3.2) \qquad V_n(t) = n^{\frac{1}{2}}\left[G_n\big(\phi_n(t)\big) - t\right], \quad 0 \le t \le 1,$$

where

$$(3.3) \qquad \phi_n(t) = G\left[(\hat{\beta}_n/\beta)G^{-1}(t) + (\hat{\alpha}_n - \alpha)/\beta\right], \quad 0 \le t \le 1.$$

Note that $\phi_n(t)$ is increasing in t and hence is a "random change of time" in the sense of Billingsley (1968), p. 144. With this notation, we have

$$(3.4) \qquad V_n(t) = \Delta_n \circ \phi_n(t), \quad 0 \le t \le 1,$$

where

$$(3.5) \qquad \Delta_n(t) = n^{\frac{1}{2}}[G_n(t) - \phi_n^{-1}(t)], \quad 0 \le t \le 1.$$

Our plan of attack is to establish, under appropriate conditions, that

$$(3.6) \qquad \phi_n \xrightarrow{P} I \quad \text{in} \quad \mathcal{D}[0,1],$$

where I is the identity function $I(t) \equiv t$ on $[0,1]$; that

(3.7) $$\Delta_n - \Delta_n^* \xrightarrow{p} 0 \quad \text{in} \quad \mathcal{D}[0,1],$$

where

(3.8) $$\Delta_n^*(t) = n^{\frac{1}{2}}\left\{G_n(t) - t + G'\left(G^{-1}(t)\right)\left[G^{-1}(t)\left(\frac{\hat{\beta}_n - \beta}{\beta}\right) + \frac{\hat{\alpha}_n - \alpha}{\beta}\right]\right\}, \quad 0 \le t \le 1;$$

and that

(3.9) $$\Delta_n^* \xrightarrow{d} V^\circ \quad \text{in} \quad \mathcal{D}[0,1].$$

Given (3.6), (3.7), and (3.9), it follows by Billingsley (1968), Theorem 4.4, that

(3.10) $$(\Delta_n, \phi_n) \xrightarrow{d} (V^\circ, I) \quad \text{in} \quad \mathcal{D}[0,1].$$

Consequently, since $P(V^\circ \in \mathcal{C}[0,1]) = 1$, where $\mathcal{C}[0,1]$ denotes the space of continuous functions on $[0,1]$, we have by Section 17.1 of Billingsley (1968) that

(3.11) $$\Delta_n \circ \phi_n \xrightarrow{d} V^\circ \circ I = V^\circ \quad \text{in} \quad \mathcal{D}[0,1].$$

By (3.4), we thus have (2.3).

LEMMA 1. *Under Assumptions* A (ii), (iii), *and if* $n^{\frac{1}{2}}\hat{\theta}_{1n}$ *and* $n^{\frac{1}{2}}(\hat{\theta}_{2n} - 1)$ *are both* $O_p(1)$, $n \to \infty$, *we have*

(3.12) $$n^{\frac{1}{2}}\sup_x |G(\hat{\theta}_{2n}x + \hat{\theta}_{1n}) - G(x) - G'(x)[(\hat{\theta}_{2n} - 1)x + \hat{\theta}_{1n}]| = o_p(1), \quad n \to \infty.$$

PROOF. Write

(3.13) $$G(ax + b) - G(x) - G'(x)[(a - 1)x + b] = A(a,b,x) + B(a,x) + C(a,b,x),$$

where

$$A(a,b,x) = G(ax + b) - G(ax) - bG'(ax),$$

$$B(a,x) \quad = G(ax) - G(x) - (a - 1)xG'(x)$$

and

$$C(a,b,x) = b[G'(ax) - G'(x)].$$

First, by Taylor expansion, $\sup_x |A(a,b,x)| \le \frac{1}{2}b^2 \sup_x |G''(x)|$. By A (iii) and since $n^{\frac{1}{2}}\hat{\theta}_{1n}$ is $O_p(1)$, we thus have

$$(3.14) \qquad n^{\frac{1}{2}}\sup_x |A(\hat{\theta}_{2n},\hat{\theta}_{1n},x)| \le \tfrac{1}{2}n^{\frac{1}{2}}\hat{\theta}_{1n}^2 \sup_x |G''(x)| = o_p(1), \quad n \to \infty.$$

Next we establish

$$(3.15) \qquad n^{\frac{1}{2}}\sup_x |B(\hat{\theta}_{2n},x)| = o_p(1), \quad n \to \infty.$$

Let $0 < \epsilon < 1$ be given. Then there exist M_ϵ and N_ϵ such that

$$(3.16) \qquad P[n^{\frac{1}{2}}|\hat{\theta}_{2n} - 1| > M_\epsilon] < \epsilon, \quad n > N_\epsilon.$$

Also, by A (ii), there exists M'_ϵ such that $|xG'(x)| < \epsilon(1 - \epsilon)/M_\epsilon$ for $|x| > M'_\epsilon$. Now, for $|a - 1| < \epsilon$ and $|x| > M'_\epsilon/(1 - \epsilon) \equiv M''_\epsilon$, and using

$$|B(a,x)| \le |(a - 1)xG'(x^*)| + |(a - 1)xG'(x)|,$$

where x^* lies between x and ax, we have $M'_\epsilon < (1 - \epsilon)|x| < |x^*|$ and thus

$$|B(a,x)| \le |a - 1|[|x^*G'(x^*)|/(1 - \epsilon) + |xG'(x)|]$$

$$\le |a - 1|[\epsilon/M_\epsilon + \epsilon(1 - \epsilon)/M_\epsilon].$$

Consequently, for $n > N'_\epsilon$, we have

$$(3.17) \qquad P[\sup_{|x|>M''_\epsilon} n^{\frac{1}{2}}|B(\hat{\theta}_{2n},x)| > 2\epsilon] < \epsilon.$$

Also, for $|x| \leq M''_\epsilon$ we have by Taylor expansion

$$|B(a,x)| \leq \tfrac{1}{2}(a - 1)^2(M''_\epsilon)^2 \sup_x |G''(x)|,$$

so that, for n sufficiently large, say $n \geq N''_\epsilon \geq N'_\epsilon$, we have

$$P[\sup_{|x|\leq M''_\epsilon} n^{\frac{1}{2}}|B(\hat{\theta}_{2n},x)| > 3\epsilon] < 2\epsilon. \tag{3.18}$$

By (3.17) and (3.18), (3.15) follows. In similar fashion we may establish

$$n^{\frac{1}{2}}\sup_x |C(\hat{\theta}_{2n},\hat{\theta}_{1n},x)| = o_p(1), \quad n \to \infty. \tag{3.19}$$

By (3.13), (3.14), (3.15), and (3.19), we have (3.12). □

COROLLARY. *Under the conditions of Lemma* 1,

$$n^{\frac{1}{2}}\sup_x |G(\hat{\theta}_{2n}x + \hat{\theta}_{1n}) - G(x)| = O_p(1), \quad n \to \infty. \tag{3.20}$$

PROOF. Clearly we have

$$n^{\frac{1}{2}}\sup_x G'(x)[\hat{\theta}_{2n} - 1)x + \hat{\theta}_{1n}] = O_p(1), \quad n \to \infty,$$

which immediately, via (3.12), yields (3.20). □

We are now ready to prove (3.6) and (3.7).

LEMMA 2. *Under Assumptions* A (ii), (iii), *and if* $n^{\frac{1}{2}}(\hat{\alpha}_n - \alpha)$ *and* $n^{\frac{1}{2}}(\hat{\beta}_n - \beta)$, $\beta > 0$, *both have nondegenerate limit laws, we have* (3.6) *and* (3.7).

PROOF. Recalling (3.3) and making the substitutions $\hat{\theta}_{1n} = (\hat{\alpha}_n - \alpha)/\beta$ and $\hat{\theta}_{2n} = \hat{\beta}_n/\beta$ in the preceding Corollary, we have

$$(3.21) \qquad n^{\frac{1}{2}} \sup_t |\phi_n(t) - I(t)| = o_p(1), \quad n \to \infty.$$

Therefore, since convergence in the uniform topology implies convergence in the Skorokhod topology in $\mathcal{D}[0,1]$ (see Billingsley (1968), Section 18), (3.6) follows.

To obtain (3.7), we first consider Δ_n^{**} defined by

$$(3.22) \quad \Delta_n^{**}(t) = n^{\frac{1}{2}}\Big\{G_n(t) - t + G'\big(G^{-1}(t)\big)\Big[G^{-1}(t)\Big(\frac{\hat{\beta}_n - \beta}{\beta}\Big) + \frac{\hat{\alpha}_n - \alpha}{\beta}\Big]\Big\}, \quad 0 \le t \le 1.$$

Setting $\hat{\theta}_{1n} = -(\hat{\alpha}_n - \alpha)/\hat{\beta}_n$ and $\hat{\theta}_{2n} = \beta/\hat{\beta}_n$, we have

$$\Delta_n^{**}(t) - \Delta_n^{*}(t) = n^{\frac{1}{2}}\Big\{G(\hat{\theta}_{2n}G^{-1}(t) + \hat{\theta}_{1n}) - t - G'\big(G^{-1}(t)\big)\Big[(\hat{\theta}_{2n} - 1)G^{-1}(t) + \hat{\theta}_{2n}\Big]\Big\},$$

$$0 \le t \le 1,$$

and it follows by Lemma 1 that

$$(3.23) \qquad \sup_{0 \le t \le 1} |\Delta_n^{**}(t) - \Delta_n(t)| = o_p(1), \quad n \to \infty.$$

But

$$\sup_{0 \le t \le 1} |\Delta_n^{*}(t) - \Delta_n^{**}(t)| = \sup_{-\infty < x < \infty} n^{\frac{1}{2}} G'(x) \Big| x \frac{(\hat{\beta}_n - \beta)^2}{\hat{\beta}_n \beta} + \frac{(\hat{\alpha}_n - \alpha)(\hat{\beta}_n - \beta)}{\hat{\beta}_n \beta}\Big|$$

$$(3.24) \qquad \le \sup_x n^{\frac{1}{2}} |xG'(x)| \Big|\frac{(\hat{\beta}_n - \beta)^2}{\hat{\beta}_n \beta}\Big| + \sup_x n^{\frac{1}{2}} G'(x) \Big|\frac{(\hat{\alpha}_n - \alpha)(\hat{\beta}_n - \beta)}{\hat{\beta}_n \beta}\Big|$$

$$= o_p(1), \quad n \to \infty. \ \square$$

We now establish (3.9).

LEMMA 3. *Under Assumptions* A (i), (ii) *on* G *and Assumption* B *on* $(\hat{\alpha}_n, \hat{\beta}_n)$, *we have* (3.9).

PROOF. First we note that the finite-dimensional distributions of Δ_n^* converge to those of V°. This is easily established using (3.24) and the "Cramér-Wold device" (see Billingsley (1968), Theorem 7.7). We omit the details.

Secondly, we turn to the question of tightness of the process Δ_n^* in $\mathcal{D}[0,1]$. By A (ii), we have $\Delta_n^*(0) = 0$. Therefore, by Theorem 15.5 of Billingsley (1968), it suffices to show that, given $\epsilon > 0$ and $\eta > 0$, there exist a δ, $0 < \delta < 1$, and an integer N_0 such that

$$P[\sup_{|s-t|<\delta} |\Delta_n^*(s) - \Delta_n^*(t)| > \epsilon] < \eta, \quad n \geq N_0. \tag{3.25}$$

To obtain (3.25), we show the existence of δ_0, δ_1, δ_2, N_1 and N_2 for which

$$P[\sup_{|s-t|<\delta_0} |W_n(s) - W_n(t)| > \epsilon/3] < \eta/3, \quad n \geq N_1, \tag{3.26}$$

$$P[n^{\frac{1}{2}}\left|\frac{\hat{\alpha}_n - \alpha}{\beta}\right| \sup_{|s-t|<\delta_1} |a(s) - a(t)| > \epsilon/3] < \eta/3, \quad n \geq N_2, \tag{3.27}$$

and

$$P[n^{\frac{1}{2}}\left|\frac{\hat{\beta}_n - \beta}{\beta}\right| \sup_{|s-t|<\delta_2} |b(s) - b(t)| > \epsilon/3] < \eta/3, \quad n \geq N_2, \tag{3.28}$$

where

$$a(t) = G'\big(G^{-1}(t)\big), \quad b(t) = G'\big(G^{-1}(t)\big)G^{-1}(t), \quad 0 \leq t \leq 1. \tag{3.29}$$

Following Billingsley (1968), Section 13, we introduce a continuous process $\tilde{W}_n(\cdot)$ which is the analogue of $W_n(\cdot)$ produced by replacing, in the definition (1.3), the step function $G_n(\cdot)$ by a continuous version. For this process we have

$$\sup_{0 \le t \le 1} |W_n(t) - \tilde{W}_n(t)| \le \frac{1}{\sqrt{n}} . \tag{3.30}$$

Further, by Billingsley's Theorems 13.1 and 8.2, there exist $0 < \delta_0 < 1$ and integer N_1^* such that

$$P[\sup_{|s-t|<\delta_0} |\tilde{W}_n(s) - \tilde{W}_n(t)| > \epsilon/6] < \eta/3, \quad n_1 \ge N_1^*. \tag{3.31}$$

By (3.30), for $n > (12/\epsilon)^2$, we have

$$\sup_{|s-t|<\delta_0} |W_n(s) - W_n(t)| \le \sup_{|s-t|<\delta_0} |\tilde{W}_n(s) - \tilde{W}_n(t)| + \epsilon/6. \tag{3.32}$$

Consequently, (3.26) holds with $N_1 = \max\{N_1^*, (12/\epsilon)^2\}$.

Since $n^{\frac{1}{2}}(\hat{\alpha}_n - \alpha)$ and $n^{\frac{1}{2}}(\hat{\beta}_n - \beta)$ have limiting distributions, there exist M_1, M_2 and $N_2 \ge N_1$ such that

$$P[n^{\frac{1}{2}}|\frac{\hat{\alpha}_n - \alpha}{\beta}| > M_1] < \eta/3, \quad n \ge N_2, \tag{3.33}$$

and

$$P[n^{\frac{1}{2}}|\frac{\hat{\beta}_n - \beta}{\beta}| > M_2] < \eta/3, \quad n \ge N_2. \tag{3.34}$$

Now let ϵ_1 and ϵ_2 satisfy $\epsilon/3\epsilon_i > M_i$, $i = 1, 2$. Then, since by A (i) the functions $a(\cdot)$ and $b(\cdot)$ are uniformly continuous on $[0,1]$, there exist δ_1 and δ_2 such that

$$\sup_{|s-t|<\delta_1} |a(s) - a(t)| < \epsilon_1 \tag{3.35}$$

and

$$\sup_{|s-t|<\delta_2} |b(s) - b(t)| < \epsilon_2.$$

The relations (3.27) and (3.28) now follow easily. $\square$

4. Applications to tests for normality. As an example we consider the Kolmogorov-Smirnov statistics given in (1.9) - (1.12) for testing normality, i.e., $G(x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$, $-\infty < x < \infty$, for some μ and $\sigma > 0$. The unknown parameters, μ and σ, will be estimated by the sample mean ($\bar{X}_n$) and sample standard deviation (S_n), respectively. Monte Carlo methods were used to simulate the distributions of the limiting r.v.'s given in (1.13) - (1.16). The procedure was to approximate the Gaussian Process $V^\circ = V^\circ(\Phi)$ by its finite-dimensional distributions, corresponding to evaluation of the process at 119 equally spaced points in the unit interval. One thousand multivariate normal random vectors with this covariance structure were generated using a program from the International Mathematical and Statistical Library. The empirical distributions for the supremum, the infimum, and the difference between the supremum and the infimum of the resulting multivariate normal vectors were then tabulated, thus approximating the limit laws of D_n^+, D_n^-, D_n, and $D_n^\pm$.

In particular, since

$$\begin{aligned}\sigma_{14} &= \mathrm{Cov}[H_n(t), \bar{X}_n] \\ &= \int_{-\infty}^{\Phi^{-1}(t)} x d\left(\Phi(x)\right) \\ &= -\Phi'\left(\Phi^{-1}(t)\right), \quad 0 \le t \le 1,\end{aligned}$$

$$\sigma_{24} = \mathrm{Cov}\Big(H_n(t),\ S_n\Big)$$

$$\cong (2)^{-\frac{1}{2}}\Big[\int_{-\infty}^{\Phi^{-1}(t)} x^2 d\Big(\Phi(x)\Big) - t\Big]$$

$$= -(2)^{-\frac{1}{2}}\Phi^{-1}(t)\Phi'[\Phi^{-1}(t)], \quad 0 \le t \le 1,$$

and $\bar{X}_n$ and S_n^2 are uncorrelated, from (2.2) we have for $0 \le s,t \le 1$

$$E\Big[V^\circ(s)V^\circ(t)\Big] = \min(s,t) - st$$
$$- \Big[1 + \tfrac{1}{2}\Phi^{-1}(s)\Phi^{-1}(t)\Big]\Phi'\Big[\Phi^{-1}(s)\Big]\Phi'\Big[\Phi^{-1}(t)\Big],$$

a result originally derived by Kac, Kiefer, and Wolfowitz (1955). Various sample quantiles for the generated frequency distributions are shown below:

[INSERT TABLE 4.1]

In order to investigate the validity of the above approximations, the empirical distributions were also generated for multivariate normal vectors corresponding to the finite-dimensional distributions of V° at 20, 30, 40, 60, and 180 equally spaced points. Since the differences in the observed quantiles diminished rapidly as the number of evaluation points increased and, in fact, some reversals were observed, the approximating procedure was terminated at 120 equally spaced intervals.

5. Model validation in completely randomized designs. A frequently occurring problem is that of testing the validity of the generally assumed linear model for data arising from a completely randomized design, i.e.,

$$(5.1) \qquad Y_{ij} = \mu + \alpha_i + \epsilon_{ij}, \quad j = 1, \cdots, n_i \text{ and } i = 1, \cdots, a,$$

where α_i is the non-stochastic effect of the i^{th} treatment with $\sum_{i=1}^{a} n_i \alpha_i = 0$ and ϵ_{ij}, the experimental error associated with the j^{th} experimental unit in the i^{th} treatment group, is normally distributed with zero mean and unknown variance σ^2.

The relevant departures from these assumptions are (i) a non-additive error structure resulting in heteroscedasticity of the experimental errors and (ii) non-normality of the experimental errors. If both (i) and (ii) are of concern, then the development in Section 1 suggests basing a test of the hypothesis that (5.1) is the correct model on the combined modified empirical c.d.f.

$$H_{n.}(t) = (n_.)^{-1} \sum_{i=1}^{a} \sum_{j=1}^{n_i} I\left[\Phi\left[\frac{Y_{ij} - \bar{Y}_{i.}}{s_p}\right] \leq t\right], \quad 0 \leq t \leq 1,$$

where $n_. = \sum_{i=1}^{a} n_i$ and s_p^2 is the usual pooled variance estimator of σ^2. (We will employ the usual dot notation in the remainder of this paper.) In turn, $V_{n.}$ and the statistics $D_{n.}$, $D^+_{n.}$, $D^-_{n.}$, and $D^{\pm}_{n.}$ can be defined analogously to (1.8) and (1.13) - (1.16), respectively.

We will now show that the null hypothesis asymptotic distributions of $D_{n.}$, $D^+_{n.}$, $D^-_{n.}$, and $D^{\pm}_{n.}$ are identical to the limiting distributions for the analogous one-sample statistics for testing normality given in Section 4. First define

$$H_{ni}(t) = n_i^{-1} \sum_{j=1}^{n_i} I\left[\Phi\left[\frac{Y_{ij} - \bar{Y}_{i.}}{s_p}\right] \le t\right], \quad 0 \le t \le 1,\ i = 1, \cdots, a.$$

Similarly define V_{ni} and Δ^*_{ni}, $i = 1, \cdots, a$. Then it follows that

$$H_{n.}(t) = \sum_{i=1}^{a} \left(\frac{n_i}{n_.}\right) G_{ni}(t), \quad 0 \le t \le 1,$$

$$V_{n.}(t) = \sum_{i=1}^{a} \left(\frac{n_i}{n_.}\right)^{\frac{1}{2}} V_{ni}(t), \quad 0 \le t \le 1,$$

and

$$\Delta^*_{n.}(t) = \sum_{i=1}^{a} \left(\frac{n_i}{n_.}\right)^{\frac{1}{2}} \Delta^*_{ni}(t), \quad 0 \le t \le 1.$$

For fixed i, $i = 1, \cdots, a$, we may apply the results of Section 3. In fact, from (3.6) and (3.7) it follows that

$$\sup_{0\le t\le 1} |\Delta^*_{ni}(t) - V_{ni}(t)| = \sup_{0\le t\le 1} |\Delta^*_{ni}(t) - \Delta_n(\phi_n(t))|$$

$$\le \sup_{0\le t\le 1} |\Delta^*_{ni}(\phi_n(t)) - \Delta^*_{ni}(t)|$$

$$+ \sup_{0\le t\le 1} |\Delta_{ni}(t) - \Delta^*_{ni}(t)|$$

converges to zero in probability as $n_i \to \infty$. From (5.3) and (5.4) it is immediately obvious that

$$V_{n\cdot} - \Delta^*_{n\cdot} \xrightarrow{p} 0 \text{ as } \min_{1\le i\le a} (n_i) \to \infty \quad . \tag{5.5}$$

In particular, for $0 \le t \le 1$,

$$\Delta^*_{n\cdot}(t) = \sum_{i=1}^{a} \sqrt{\frac{n_i}{n_\cdot}} \left\{ \sqrt{n_i} \left[G_{ni}(t) - t + \Phi'(\Phi^{-1}(t)) \left[\frac{\bar{Y}_{i\cdot} - (\mu + \alpha_i)}{\sigma} + \Phi^{-1}(t) \left[\frac{s_p - \sigma}{\sigma} \right] \right] \right] \right\}$$

$$= \sqrt{n_\cdot} \left[G_{n\cdot}(t) - t + \Phi'(\Phi^{-1}(t) \left[\frac{(\bar{Y}_{i\cdot} - \mu)}{\sigma} + \Phi^{-1}(t) \left[\frac{s_p - \sigma}{\sigma} \right] \right] \right] ,$$

where $G_{n\cdot}$ is the combined empirical c.d.f. of the standardized i.i.d. r.v.'s $\left\{\frac{Y_{ij} - (\mu + \alpha_i)}{\sigma},\ j = 1, \cdots, n_i,\ i = 1, \cdots, a\right\}$. We may make similar interpretations of $(\bar{Y}_{\cdot\cdot} - \mu)$ and $\left(\frac{s_p - \sigma}{\sigma}\right)$, from which the desired result follows.

However, if we are concerned only about the non-normality of the errors, i.e., we wish to test H_0: $\{\epsilon_{ij}\colon\ j = 1, \cdots, n_i,\ i = 1, \cdots, a\}$ are independently normally distributed with zero mean and unknown variance σ_i^2, against the general alternative, it is appropriate to base the test on

$$H'_{n.}(t) = (n.)^{-1} \sum_{i=1} \sum_{j=1} I\left[\Phi\left(\frac{Y_{ij} - \bar{Y}_{i.}}{s_i}\right) \le t\right], \quad 0 \le t \le 1 ,$$

where s_i^2, $i = 1, \cdots, a$, is the sample variance of the i^{th} treatment group. Arguments similar to those given above show that the Kolmogorov-Smirnov statistics based on this modified empirical c.d.f. have the same limiting H_o-distributions as given before.

Remark. The result given in (5.5) remains valid under H_o if Φ is replaced by any G satisfying Assumptions A with any consistent estimator of (α_i, β_i), $i = 1, \cdots, a$, replacing $(\bar{Y}_{i.}, s_i)$.

REFERENCES

[1] Billingsley, Patrick (1968). *Convergence of Probability Measures*. Wiley, New York.

[2] Durbin, J. (1973). Weak convergence of the sample distribution function when parameters are estimated. *Ann. Statist.*, 1, 279-290.

[3] 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.*, 26, 189-211.

[4] Neuhaus, G. (1976). Weak convergence under contiguous alternatives of the empirical process when parameters are estimated: The D_K approach. *Lecture Notes in Mathematics*, 566, Springer-Verlag, New York.

TABLE 4.1

SAMPLING DISTRIBUTIONS OF APPROXIMATE KOLMOGOROV-SMIRNOV TYPE STATISTICS FOR TESTING NORMALITY WITH MEAN AND VARIANCE ESTIMATED $(\bar{X}_n, S_n^2)$*

p	p^{th} Quantile			
	D_n^+	D_n^-	D_n	$D_n^{\pm}$
.010	0.249	-0.926	0.327	0.614
.025	0.268	-0.841	0.351	0.670
.050	0.293	-0.768	0.381	0.704
.100	0.319	-0.698	0.409	0.770
.250	0.383	-0.583	0.469	0.880
.500	0.469	-0.478	0.555	1.033
.750	0.573	-0.394	0.659	1.221
.900	0.685	-0.333	0.752	1.414
.950	0.746	-0.300	0.835	1.576
.975	0.826	-0.272	0.910	1.698
.990	0.926	-0.237	0.991	1.853

* k = 120.