The Annals of Statistics
1981, Vol. 9, No. 5, 923–944

SHEFFER POLYNOMIALS FOR COMPUTING EXACT KOLMOGOROV-SMIRNOV AND RÉNYI TYPE DISTRIBUTIONS[1]

BY HEINRICH NIEDERHAUSEN

University of Toronto

Sheffer polynomials are used for proving algorithms and closed forms for a broad class of nonlinear one-sided order and rank distributions. For two-sided tests a more general concept results in a representation theorem for piecewise Sheffer polynomial functions.

1. Introduction.

Let $X_1, \cdots, X_M$ be i.i.d. random variables with continuous distribution function F and empirical distribution function

$$F_X(x) = M^{-1} \sum_{i=1}^{M} 1_{(-\infty,x]}(X_i).$$

We call the pair

$$\sup_{\alpha \le F(x) \le \beta} \frac{F_X(x) - \gamma F(x)}{\delta - \kappa F(x)}, \inf_{\alpha \le F(x) \le \beta} \frac{F_X(x) - \gamma F(x)}{\delta - \kappa F(x)}, \tag{1.1}$$

a bivariate Rényi statistic (α, β, γ, δ and κ have to be suitable constants). More generally, we call

$$P\{f(F_X(x)) \le F(x) \le g(F_X(x)) \quad \text{for all} \quad \alpha \le F(x) \le \beta\} \tag{1.2}$$

a Rényi type distribution. Of course, f and g are determined by the test statistic. For instance, $\alpha = \kappa = 0$ and $\gamma = \delta = 1$ yield the Kolmogorov-Smirnov one-sample test in (1.1). In that case $f(u) = u - s$ and $g(u) = u + r$ are the corresponding functions in (1.2), if $(s, -r)$ is the value of this test statistic. For two-sample tests replace $F(x)$ by $F_V(x)$ in (1.1) and (1.2), where V is the combined sample.

It is well known in the theory of order statistics that the probability (1.2) can be written as a solution of a system of differential equations with boundary conditions. A system of difference equations solves the two-sample case. Both systems are piecewise solvable by polynomials. But because of the complicated boundary conditions a general approach was never undertaken to find these polynomials directly from their defining equations. We show in this paper, however, that there are algebraic tools to solve the problem in a unified and straightforward way.

The "Finite Operator Calculus" of Rota, Kahaner and Odlyzko (1973) is the basis of our two simple representation theorems (Theorems A.1 and A.2) for the solutions of such equations. To our knowledge, all known closed forms and recursions for the exact probability (1.2) can be derived by means of Sheffer polynomials.

For a single problem this unified approach does not mean substantially less work. The benefit of our method is that this work has to be done only once for many different results previously scattered in the literature. Equally important, the present theory makes it transparent how these results are connected. For instance, specializing Theorem A.2 to the

Received March 15, 1979; revised December 4, 1980.

[1] Research supported by the German Research Council (DFG) and National Science Foundation Grant MCS 78-01108.

AMS 1980 *subject classification*. Primary 62G10; secondary 05A15.

Key words and phrases. Rényi type distributions, Sheffer sequences.

Kolmogorov-Smirnov one-sample test yields Durbin's (1973, (2.4.8)) algorithm, Kemperman's (1961, (5.40)) generating function identity, and Steck's (1971) determinant. Now change only the parameters and the operator in the very same theorem and obtain all the two sample analogs!

Sections 2 and 3 discuss respectively the one- and two-sample tests. The algebraic part of the method is summarized in the Appendix. Section 4 contains various examples, including the Kolmogorov-Smirnov statistics with weight factor $[F(x)\{1-F(x)\}]^{-1/2}$, and the Butler test for symmetry. Other important applications of (1.2) arise in power computation as has been investigated already by Steck (1969, 1971, 1974).

The following notations are used throughout the paper. $\mathbb{Z}$ stands for the set of all integers, and $\mathbb{R}$ for the real numbers; $\mathbb{N}_0 = \{n \in \mathbb{Z}, n \geq 0\}$, $\mathbb{N}_1 = \{n \in \mathbb{Z}, n \geq 1\}$, $x \wedge y = \min\{x, y\}$, $x \vee y = \max\{x, y\}$, $(x)_+ = \max\{0, x\}$, $(x)_- = \min\{0, x\}$, $\lceil x \rceil = \min\{i \in \mathbb{Z} \mid i \geq x\}$, $\lfloor x \rfloor = \max\{i \in \mathbb{Z} \mid i \leq x\}$,

$$\binom{x}{n} = \frac{x(x-1)\cdots(x-n+1)}{n!} \quad \text{for all} \quad n \in \mathbb{N}_1; \binom{x}{0} = 1; \binom{x}{z} = 0 \quad \text{for all} \quad z \notin \mathbb{N}_0.$$

For the values of a function $\nu: \mathbb{N}_0 \to \mathbb{R}$ we use both notations $\nu(i)$ and ν_i.

2. One-sample tests.

2.1. *Sheffer polynomials for D.*

Let μ and ν be monotone non-decreasing functions from $\mathbb{N}_0$ into $\mathbb{R}$, satisfying $0 \leq \nu_0 \leq \mu_0$ and $\nu_i < \mu_{i-1}$ for all $i \in \mathbb{N}_1$. The following functions define a μ-Sheffer sequence (see (A.12)) for the derivative operator $D: p(x) \mapsto \frac{d}{dx} p(x)$,

$$f_0(x) = \begin{cases} 1 & \text{if } x \leq 0 \\ 0 & \text{if } x > 0, \end{cases}$$

and

$$f_n(x) = \begin{cases} \displaystyle\int_{\nu(n)}^{x \wedge \mu(n-1)} \int_{\nu(n-1)}^{u_n \wedge \mu(n-2)} \cdots \int_{\nu(1)}^{u_2 \wedge \mu(0)} 1 \, du_1 \cdots du_n & \text{if } x \leq \mu_n \\ 0 & \text{if } x > \mu_n, \end{cases}$$

for all $n \in \mathbb{N}_1$. Obviously $f_n(\nu_n) = \delta_{0,n}$; hence (f_n) has roots in ν (see (A.14)).

Denote by $U_{(i)}$ the ith order statistic of a size M random sample from $U(0, 1)$. If $\mu_{M-1} \leq 1$, then

$$(2.1) \qquad f_m(\mu_{M-1}) = P(\nu_i \leq U_{(i)} \leq \mu_{i-1} \quad \text{for all} \quad i = 1, \cdots, M)/M! = f_M(\mu_M).$$

2.2. *Closed forms.*

If $\mu \geq 1$ in (2.1), the μ-Sheffer sequence (f_n) for D with roots in ν equals, for all $x \leq 1$, the Sheffer sequence (s_n) for D with roots in ν and

$$(2.2) \qquad P(\nu_i \leq U_{(i)} \quad \text{for all} \quad i = 1, \cdots, M) = M! s_M(1).$$

If $\nu \equiv 0$, but μ is arbitrary, then a probability like (2.2) is obtained from the **"dual form"**

$$\begin{aligned} &P(\nu_i \leq U_{(i)} \leq \mu_{i-1} \quad \text{for all} \quad i = 1, \cdots, M) \\ &= P(1 - \mu(M-i) \leq U_{(i)} \leq 1 - \nu(M+1-i) \quad \text{for all} \quad i = 1, \cdots, M) \\ &= M! f'_M(1), \end{aligned}$$

if (f'_n) is the μ'-Sheffer sequence for D with roots in ν', where

$$(2.4) \qquad \nu'_i = (1 - \mu(M-i))_+ \quad \text{and} \quad \mu'_i = 1 - \nu(M-i) \quad \text{for all} \quad i = 0, \cdots, M.$$

Therefore, (2.2) applies to both one-sided cases. For general ν, only recursions are available for computing $s_M(1)$ (see Section 2.3).

But there are well-known special cases, where Theorem A.1 allows us to derive closed forms for $s_M(1)$. With regard to the examples in Section 4, we list two formulas where the first one is only a special case of the second. If

$$\nu_i = \begin{cases} 0 & \text{for all} \quad i = 0, \cdots, L \\ ic + d & \text{for all} \quad i > L \end{cases} \tag{2.5}$$

then (see (A.7) and (A.8))

$$\begin{aligned} M! s_M(1) &= \sum_{i=0}^{L} \binom{M}{i} (ic + d)^i (1 - Mc - d)(1 - ic - d)^{M-i-1} \\ &= 1 - \sum_{i=L+1}^{M} \binom{M}{i} (ic + d)^i (1 - Mc - d)(i - ic - d)^{M-i-1}. \end{aligned}$$

If

$$\nu_i = \begin{cases} 0 & \text{for all} \quad i = 0, \cdots, L \\ ic + d & \text{for all} \quad i = L + 1, \cdots, K \\ \delta & \text{for all} \quad i > K, \end{cases} \tag{2.6}$$

then (see (A.9))

$$M! s_M(1) = \sum_{i=0}^{K} \binom{M}{i} (1 - \delta)^{M-i} \sum_{j=0}^{L} \binom{i}{j} (\delta - ic - d)(\delta - jc - d)^{i-j-1} (jc + d)^j. \tag{2.7}$$

Remember that the inner sum equals δ^i for all $i = 0, \cdots, L$. The outside method (A.8) can be used for the inner and/or the outer sum in (2.7). This explains the broad variety of formulas occurring in the literature for the examples in Section 4.

In principle, each function ν can be decomposed into affine pieces, and (A.9) can be applied.

2.3. *Recursions.*

For this section we assume $\mu_M = 1$ without loss of generality. With $q_{n-k}(x) = x^{n-k}/(n - k)!$ in (A.13), the Algorithm A.1 was found by Nóe (1972). Given that μ_k is constant for $k = 0, \cdots, K$, say, it may be possible to use an explicit formula for $f_i(\mu_K)$ $(i = 0, \cdots, K)$. The same is true for the values p_i in the following application of (A.16). Define $p_0 = 1$ and

$$p_i = \sum_{k=0}^{i-1} \binom{i}{k} (-1)^{i-k-1} (\mu_k - \nu_i)_+^{i-k} p_k. \tag{2.8}$$

Then

$$P(\nu_i \le U_{(i)} \le \mu_{i-1} \quad \text{for all} \quad i = 1, \cdots, M) = p_M.$$

Alternatively, we get from Corollary A.2:

$$p_M = \det\left(\binom{i}{j-1} (\mu_{j-1} - \nu_i)_+^{i-j+1} \right)_{i,j=1,\cdots,M}.$$

Epanechnikov (1968) found recursion (2.8) and Steck (1971) independently derived this determinant and many applications. See also Pitman (1972) for another proof.

REMARK. Depending on the accuracy of the computer, (2.8) should be used only for small M because of the alternating summation. In Algorithm A.1 the summation does not alternate, but compared with (2.8) the amount of computation is approximately squared! In the computation of significance points, it often occurs that $\nu_M \le \mu_0$. The same is true, of course, in both one sided cases. Theorem A.2 (with $i = 0$) yields for any real function σ on $\mathbb{N}_0$,

$$p_j = j!t_{j,0}(\sigma_j) - \sum_{k=0}^{j-1} \binom{j}{k} (\sigma_j - \mu_k)^{j-k} p_k \quad \text{for all} \quad j = 1, \cdots, M. \tag{2.9}$$

From $\nu_M \le \mu_0$ we see that $t_{j,0}$ equals for $j = 0, \cdots, M$ the Sheffer sequence for D with roots in ν. Given that $\sigma_j \ge \mu_{j-1}$ for all $j = 1, \cdots, M$, the summation in (2.9) is non-alternating. But how to compute $t_{j,0}(\sigma_j)$? In the simple case $\nu \equiv 0$ (one sided tests) we get $j!t_{j,0}(\sigma_j) = \sigma_j^j$. With $\sigma_j = \mu_{j-1}$ Steck's formula (1971, (2.3)) is obtained. See the previous section for other closed forms. We suggest the following procedure for general ν (with $\nu_M \le \mu_0$).

Choose $\sigma \equiv 1$. Thus, for all $= 1, \cdots, M$,

$$j!t_{j,0}(1) = P(\nu_i \le U_{(i)} \quad \text{for all} \quad i = 1, \cdots, j) = j!f_j^{(j)}(1),$$

if $(f_n^{(j)})$ is the $\mu^{(j)}$-Sheffer sequence for D with roots in 0, where $\mu_i^{(j)} = 1 - \nu(j-i)$ for all $i = 0, \cdots, j$. Hence, (2.9) can be applied to compute $t_{j,0}(1)$ (we choose again $\sigma \equiv 1$) using

$$p_0^{(j)} = 1 \quad \text{and} \quad p_i^{(j)} = 1 - \sum_{k=0}^{i-1} \binom{i}{k} \nu(j-k)^{i-k} p_k^{(j)} \quad \text{for all} \quad i = 1, \cdots, j.$$

Thus, $j!t_{j,0}(1) = p_j^{(j)}$ for all $j = 1, \cdots, M$. Finally, enter again (2.9) and compute p_M from $p_0 = 1$ and

$$p_j = p_j^{(j)} - \sum_{k=0}^{j-1} \binom{j}{k} (1 - \mu_k)^{j-k} p_k \quad \text{for all} \quad j = 1, \cdots, M.$$

2.4. *Rényi type distributions.*

In applications, the test distributions seldom occur in the form of (2.1). But if our method is applicable at all, they are easily transformed so that one of the following two lemmas can be used.

LEMMA 2.1. *Let f and g be monotone non-decreasing functions from* $[0, 1]$ *into itself such that $f < g$ and*

$$f(0) \le a/M < b/M \le g(1) = 1$$

for two fixed integers a and b. Then

$$P\{f(F_U(x)) \le x \le g(F_U(x)) \quad \text{for all} \quad a/M \le F_U(x) \le b/M\}$$
$$= P(\nu_i \le U_{(i)} \le \mu_{i-1} \quad \text{for all} \quad i = 1, \cdots, M),$$

where

$$\nu_i = \begin{cases} 0 & \text{for all} \quad i = 0, \cdots, a-1 \\ f(i/M) & \text{for all} \quad i = a, \cdots, b \\ f(b/M) & \text{for all} \quad i > b \end{cases}$$

and

$$\mu_i = \begin{cases} g(a/M) & \text{for all} \quad i = 0, \cdots, a-1 \\ g(i/M) & \text{for all} \quad i = a, \cdots, b \\ 1 & \text{for all} \quad i > b. \end{cases}$$

The proof is obvious. The situation in the following lemma is more complicated.

LEMMA 2.2. *Replace a/M and b/M in Lemma 2.1 by any real α and β such that $0 \le \alpha < \beta \le 1$. Then, under the same assumptions about f and g,*

$$P\{f(F_U(x)) \le x \le g(F_U(x)) \quad \text{for all} \quad \alpha \le x \le \beta\}$$
$$= P(\nu_i \le U_{(i)} \le \mu_{i-1} \quad \text{for all} \quad i = 1, \cdots, M)$$

if

$$\nu_i = \begin{cases} 0 & \text{for all} \quad i = 0, \cdots, \alpha_f \\ f(i/M) & \text{for all} \quad i = \alpha_f + 1, \cdots, \beta_f \\ \beta & \text{for all} \quad i > \beta_f \end{cases}$$

and

$$\mu_i = \begin{cases} \alpha & \textit{for all} \quad i = 0, \cdots, \alpha_g - 1 \\ g(i/M) & \textit{for all} \quad i = \alpha_g, \cdots, \beta_g - 1, \\ 1 & \textit{for all} \quad i \geq \beta_g \end{cases}$$

where

$$\alpha_f = \max\{k \leq M \,|\, f(k/M) \leq \alpha\},\ \beta_f = \max\{k \leq M \,|\, f(k/M) \leq \beta\}$$

$$\alpha_g = \min\{k \geq 0 \,|\, g(k/m) \geq \alpha\},\ \beta_g = \min\{k \geq 0 \,|\, g(k/M) \geq \beta\}.$$

PROOF. Denote by $[0, 1]^{(M)}$ the set of all monotone non-decreasing ordered vectors $u \in [0, 1]^M$, i.e. vectors

$$u = (u_1, \cdots, u_M) \text{ such that } 0 \leq u_1 \leq \cdots \leq u_M \leq 1.$$

Define the subset A of $[0, 1]^{(M)}$ by $A = \{f(i/M) \leq x$ holds for all $i = 0, \cdots, M$ and $x \in [u_i, u_{i+1}) \cap [\alpha, \beta]\}$ ($u_0 = 0$, $u_{M+1} = 1$). Then

$$\begin{aligned} A &= \{f(i/M) \leq x \quad \text{for all} \quad i = \alpha_f + 1, \cdots, M \text{ and } x \in [u_i, u_{i+1}) \cap [\alpha, \beta]\} \\ &= \{f(i/M) \leq u_i \quad \text{for all} \quad i = \alpha_f + 1, \cdots, M \text{ such that } u_i \leq \beta\} \\ &= \{u_{\beta_f+1} > \beta, \text{ and } f(i/M) \leq u_i \quad \text{for all} \quad i = \alpha_f + 1, \cdots, \beta_f \text{ such that } u_i \leq \beta\} \\ &= \{u_{\beta_f+1} > \beta, \text{ and } f(i/M) \leq u_i \quad \text{for all} \quad i = \alpha_f + 1, \cdots, \beta_f\}. \end{aligned}$$

By interchanging the roles of f and g it follows analogously that

$$\begin{aligned} B &= \{x \leq g(i/M) \quad \text{for all} \quad i = 0, \cdots, M \text{ and } x \in [u_i, u_{i+1}) \cap [\alpha, \beta]\} \\ &= \{u_{\alpha_g} < \alpha, \text{ and } u_i \leq g((i-1)/M) \quad \text{for all} \quad i = \alpha_g + 1, \cdots, \beta_g\}. \end{aligned}$$

$P(A \cap B) = P\{f(F_U(x)) \leq x \leq g(F_U(x)) \quad \text{for all} \quad \alpha \leq x \leq \beta\}$ finishes the proof. □

REMARK. If $\nu_{i+1} \leq \mu_i$ for all $i = 0, \cdots, M - 1$ in the lemmas above, look for the best applicable method in Section 2.2 or 2.3. The probability is zero otherwise.

3. Two-sample tests.

3.1. *Sheffer polynomials for* ∇.

Denote by $\Upsilon(i, j)$ the set of all vectors T consisting of exactly i ones and j zeros. For each $T = (T_1, \cdots, T_{i+j}) \in \Upsilon(i, j)$ define the **path** T'_ℓ of T by $T'_0 = 0$ and $T'_\ell = \sum_{k=1}^{\ell} T_k$ for all $\ell = 1, \cdots, i + j$. The set $\Upsilon(i, j)$ is closely related to empirical distribution functions: Let $X_1, \cdots, X_M, Y_1, \cdots, Y_N$ be $M + N$ continuous and i.i.d. random variables. Denote the monotone non-decreasing ordered combined sample by $V_1, \cdots, V_{M+N}$. Define

$$T_\ell = \begin{cases} 1 \text{ if } V_\ell = X_i & \text{for some} \quad i, 1 \leq i \leq M \\ 0 \text{ if } V_\ell = Y_j & \text{for some} \quad j, 1 \leq j \leq N. \end{cases} \tag{3.1}$$

Then $T'_\ell = MF_X(V_\ell)$ and $\ell - T'_\ell = NF_Y(V_\ell)$. Let μ and ν be integer valued functions on $\mathbb{N}_0$, $-1 = \nu_0 \leq \mu_0$ and

$$\nu_{i-1} - 1 \leq \nu_i < \mu_{i-1} \leq \mu_i \quad \text{for all} \quad i \in \mathbb{N}_1. \tag{3.2}$$

Then $f_i(j) = \#\Upsilon(i, j \,|\, \nu(T'_\ell) < \ell - T'_\ell \leq \mu(T'_\ell)$ for all $\ell = 0, \cdots, i + j)$, if (f_n) is the μ-Sheffer sequence (with variables in $\mathbb{Z}$) for the backwards difference operator ∇ (see (A.6)) with roots in ν. Hence,

$$P\{\nu(T'_\ell) < \ell - T'_\ell \leq \mu(T'_\ell) \quad \text{for all} \quad \ell = 0, \cdots, M + N\} = \binom{M+N}{M}^{-1} f_M(N). \tag{3.3}$$

REMARKS. (a) Denote by R_ι the relative rank of the ith order statistic $X_{(i)}$ in V. Then $R_\iota = (M+N)F_V(X_{(\iota)})$, and $\nu(T'_\ell) < \ell - T'_\ell \le \mu(T'_\ell)$ for all $\ell = 0, \cdots, M+N$ is equivalent to $\nu(i) + i < R_\iota \le \mu(i-1) + i$ for all $i = 1, \cdots, M$ (if $\mu_M \ge N$).

(b) It is the restriction to integer valued functions ν that makes closed forms so rare. Sometimes the following elementary identities can be helpful ($c \in \mathbb{N}_1$, $d \in \mathbb{Z}$),

$$\begin{aligned} &\#\Upsilon(i,j \mid T'_\ell/c + d \le \ell - T'_\ell \quad \text{for all} \quad \ell = 0, \cdots, i+j) \\ &= \#\Upsilon(j,i \mid \ell - T'_\ell \le cT'_\ell - cd \quad \text{for all} \quad \ell = 0, \cdots, i+j) \\ &= \#\Upsilon(j,i \mid c(T'_\ell + d - j) + i \le \ell - T'_\ell \quad \text{for all} \quad \ell = 0, \cdots, i+j). \end{aligned}$$

3.2. *Closed forms.*

If $\mu \ge N$ in (3.3), we obtain

$$P\{\nu(T'_\ell) < \ell - T'_\ell \quad \text{for all} \quad \ell = 0, \cdots, M+N\} = s_M(N) \Big/ \binom{M+N}{M}, \tag{3.4}$$

where (s_n) is the Sheffer sequence for ∇ with roots in ν. The dual form of (3.3) is now

$$\begin{aligned} P\{N - 1 - \mu(M - T'_\ell) < \ell - T'_\ell \\ \le N - 1 - \nu(M - T'_\ell) \quad \text{for all} \quad \ell = 0, \cdots, M+N\}. \end{aligned} \tag{3.5}$$

The delta operator ∇ allows two special cases. First, if

$$s_n(j) = \binom{j+n}{n} - \binom{j+n}{n-L-1} \quad \text{for any} \quad L \in \mathbb{N}_0, \tag{3.6}$$

then (s_n) is the Sheffer sequence for ∇ with roots in $\nu_i = (i-L)_+ - 1$ ($i \in \mathbb{N}_0$). Second, with the same s_n, we get from

$$\begin{aligned} f_n(j) = \textstyle\sum_{k\ge 0}\{ &s_{n-k(K+L+2)}\,(j + k(K+L+2)) \\ &- s_{j-k(K+L+2)-K-1}(n + k(K+L+2) + K + 1)\} \end{aligned} \tag{3.7}$$

for all $\nu_n \le j \le \mu_n + 1$, the μ-Sheffer sequence for ∇ with root in ν, where $\mu_i = i + K$ for all $i \in \mathbb{N}_0$, and ν as above. The proof is in both cases just verification of recursion and side conditions. See Fray and Roselle (1971) for another proof. The case $K = L$ was derived by Koroljuk (1955), using reflection arguments.

Now assume in (3.4) that

$$\nu_\iota = \begin{cases} -1 & \text{for all} \quad i = 0, \cdots, L \\ ic + d & \text{for all} \quad i > L, \end{cases}$$

where the integer constants c, d and L may not contradict (3.2). Then we get from (A.10) and (A.8)

$$\begin{aligned} s_M(N) &= \textstyle\sum_{i=0}^{L} \binom{ic+d+i}{i} \dfrac{N - Mc - d}{N - ic - d} \binom{M+N-ic-i-d-1}{M-i} \\ &= \binom{M+N}{M} - \textstyle\sum_{i=L+1}^{M} \binom{ic+d+i}{i} \\ &\quad \cdot \dfrac{N - Mc - d}{N - ic - d} \binom{M+N-ic-i-d-1}{M-i}. \end{aligned} \tag{3.8}$$

If

$$\nu_i = \begin{cases} -1 & \text{for all} \quad i = 0, \cdots, L \\ ic + d & \text{for all} \quad i = L+1, \cdots, K \\ ia + b & \text{for all} \quad i > K, \end{cases}$$

where $c, a, L, K \in \mathbb{N}_0$ and $d, b \in \mathbb{Z}$, then

$$s_M(N) = \sum_{k=0}^{K} \frac{N - Ma - b}{N - ka - b} \binom{M + N - ka - b - k - 1}{M - k}$$

(3.9)

$$\cdot \sum_{i=0}^{L} \binom{ic + d + i}{i} \frac{ka + b - kc - d}{ka + b - ic - d} \binom{k + ka + b - ic - i - d - 1}{k - i}$$

(see (A.9)). Again, the outside method (A.8) can be used for the inner and/or the outer sum in (3.9). Observe that the inner sum equals $\binom{i + ka + b}{i}$ for all $k = 0, \cdots, L$.

3.3. *Recursions.*

We assume $\mu(M) = N$ throughout this section. From the definition of a μ-Sheffer sequence (f_n) for ∇ with roots in ν we get the following two-dimensional recursion

$$f_i(j) = \begin{cases} f_i(j-1) + f_{i-1}(j) & \text{for all } \nu_i < j \le \mu_i \\ 0 & \text{otherwise,} \end{cases} \tag{3.10}$$

with initial values $f_0(j) = 1$ for all $j \le \mu_0$, and $f_i(\nu_i) = 0$ for all $i \in \mathbb{N}_1$. On a computer with unlimited integer precision, this algorithm may be slow but absolutely accurate!

The one-dimensional recursion (A.16) is left to the reader. From Corollary A.2 one gets the determinantal solution

$$P(\nu(T'_\ell) < \ell - T'_\ell \le \mu(T'_\ell) \text{ for all } \ell = 0, \cdots, M + N)$$
$$= \binom{M + N}{N}^{-1} \det\left(\left(\binom{(\mu_{j-1} - \nu_i)_+}{i - j + 1}\right)\right)_{i, j = 1, \cdots, M.}$$

This determinant has been found independently by Kreweras (1965) and Steck (1969). See also Mohanty (1971) and Pitman (1972) for other proofs.

A close look on ν and μ may save some recursion steps. If $\nu(M) < \mu(0)$ the outside method allows non-alternating summation as described in Section 2.3.

3.4. *Rényi type distributions.*

LEMMA 3.1. *Define f, g, a and b as in Lemma 2.1. Then*

$$P\{f(F_X(x)) \le F_V(x) \le g(F_X(x)) \quad \text{for all} \quad a/M \le F_X(x) \le b/M\}$$
$$= P(\nu(T'_\ell) < \ell - T'_\ell \le \mu(T'_\ell) \quad \text{for all} \quad \ell = 0, \cdots, M + N),$$

if

$$\nu_i = \begin{cases} -1 & \text{for all } i = 0, \cdots, a - 1 \\ \lceil (M + N) f(i/M) \rceil - i - 1, & \text{for all } i = a, \cdots, b \\ \nu_b & \text{for all } i > b, \end{cases}$$

and

$$\mu_i = \begin{cases} \mu_a & \text{for all } i = 0, \cdots, a - 1 \\ \lfloor (M + N) g(i/M) \rfloor - i & \text{for all } i = a, \cdots, b \\ N & \text{for all } i > b. \end{cases}$$

The proof is obvious from Section 3.1.

LEMMA 3.2. *Define f, g, a and b as in Lemma 2.1, and α_f, β_f, α_g and β_g as in Lemma 2.2 with $\alpha = a/(M + N)$ and $\beta = b/(M + N)$. Then*

$$P\{f(F_X(x)) \le F_V(x) \le g(F_X(x)) \quad \text{for all} \quad a/(M + N) \le F_V(x) \le b/(M + N)\}$$
$$= P\{\nu(T'_\ell) < \ell - T'_\ell \le \mu(T'_\ell) \quad \text{for all} \quad \ell = 0, \cdots, M + N\},$$

if

$$\nu_i = \begin{cases} -1 & \textit{for all} \quad i = 0, \cdots, \alpha_f \\ \lceil (M+N) f(i/M) \rceil - i - 1 & \textit{for all} \quad i = \alpha_f + 1, \cdots, \beta_f \\ b - \beta_f - 1 & \textit{for all} \quad i > \beta_f, \end{cases}$$

and

$$\mu_i = \begin{cases} a - \alpha_g & \textit{for all} \quad i = 0, \cdots, \alpha_g - 1 \\ \lfloor (M+N) g(i/M) \rfloor - i & \textit{for all} \quad i = \alpha_g, \cdots, \beta_g - 1 \\ N & \textit{for all} \quad i \geq \beta_g. \end{cases}$$

The proof follows the same pattern as the proof of Lemma 2.2 and is therefore omitted.

REMARK. It may happen that ν or μ in Lemma 3.1 or 3.2 violates the monotonicity conditions (3.2). In this case define the "monotone hulls" $\tilde{\nu}$ and $\hat{\mu}$ by

$$(3.11) \qquad \begin{aligned} &\tilde{\nu}_0 = -1 \\ &\tilde{\nu}_i = \max\{\nu_i, \tilde{\nu}_{i-1}\} \qquad \text{for all} \quad i = 1, \cdots, M, \end{aligned}$$

and

$$(3.12) \qquad \begin{aligned} &\hat{\mu}_M = N \\ &\hat{\mu}_i = \min\{\mu_i, \hat{\mu}_{i+1}\} \qquad \text{for all} \quad i = 0, \cdots, M-1. \end{aligned}$$

If $\tilde{\nu}_{i+1} < \hat{\mu}_i$ for all $i = 0, \cdots, M-1$, look for the best applicable method in Section 3.2 or 3.3. The probability is zero otherwise.

4. Examples.

As in the preceding sections we assume that $X_1, \cdots, X_M, Y_1, \cdots, Y_N$ are $M + N$ i.i.d. random variables with continuous distribution function F. We will work out some details in the first (and easiest) example. This task is left to the reader for the other examples, where we will derive only the middle parts of the functions ν and μ.

Example 1. The test statistic $\sup_{0<F(x)\leq 1} F_X(x)/F(x)$ has for $s \geq 1$ the distribution function

$$\begin{aligned} G(s) &= P\{F_X(x)/F(x) \leq s \quad \text{for all} \quad 0 < F(x) \leq 1\} \\ &= P\{F_X(x)/s \leq F(x) \quad \text{for all} \quad x \in \mathbb{R}\}. \end{aligned}$$

We find from Lemma 2.1 and (2.2) that $G(s) = M!s_M(1)$, if (s_n) is the Sheffer sequence for D with roots in $\nu_i = i/(Ms)$ for all $i \in \mathbb{N}_0$. Thus, ν is even easier than in (2.5), and we get directly from Lemma A.3 that

$$s_n(x) = (x - n/(sM))x^{n-1}/n!.$$

Hence, $G(s) = 1 - 1/s$. For at least equally elegant proofs see Robbins (1954) and Rényi (1968). The result was proved first by Daniels (1945, page 415), and was rediscovered also by Chang Li-Chien (1955). I thank the referee for pointing out these references.

Our method also computes immediately

$$G_\beta(s) = P\{F_X(x)/F(x) \leq s \quad \text{for all} \quad 0 \leq F(x) \leq \beta\}$$

for $0 < \beta < 1$ and $s > 0$. From Lemma 2.2 we obtain

$$\nu_i = \begin{cases} i/(sM) & \text{for all} \quad i = 0, \cdots, \beta_f \\ \beta & \text{for all} \quad i > \beta_f, \end{cases}$$

where $\beta_f = \max\{k \leq M \mid k/(sM) \leq \beta\} = \lfloor \beta sM \rfloor$. Therefore, choose $L = 0$, $c = (sM)^{-1}$, $d = 0$, $K = \lfloor \beta sM \rfloor$ and $\delta = \beta$ in (2.6). Equation (2.7) yields

$$G_\beta(s) = \sum_{i=0}^{K} \binom{M}{i} (1-\beta)^{M-i} \left(\beta - \frac{i}{sM}\right) \beta^{i-1}. \tag{4.1}$$

If $s\beta > 1$, we get $K > M$ and $G_\beta(s) = 1 - 1/s$. Chang Li-Chien (1955) found for all $s > 0$

$$G_\beta(s) = \sum_{i=0}^{K} \binom{M}{i} (1-\beta)^{M-i} \beta^i$$
$$- \sum_{k=1}^{K} \binom{M}{k} \frac{k^{k-1}(Ms-k)^{M-k}}{(Ms)^M} \sum_{i=k}^{K} \binom{M-k}{i-k} \left(\frac{Ms\beta - k}{Ms-k}\right)^{i-k} \left(1 - \frac{Ms\beta - k}{Ms-k}\right)^{M-i}.$$

That this sum really reduces to (4.1) may be shown by induction or again by using Sheffer polynomials.

A double summation becomes necessary if we compute

$$G_{\alpha,\beta}(s) = P\{F_X(x)/F(x) \le s \quad \text{for all} \quad \alpha \le F(x) \le \beta\}. \tag{4.2}$$

Then

$$\nu_i = \begin{cases} 0 & \text{for all} \quad i = 0, \cdots, \lfloor \alpha s M \rfloor \\ i/(sM) & \text{for all} \quad i = \lfloor \alpha s M \rfloor + 1, \cdots, \lfloor \beta s M \rfloor \\ \beta & \text{for all} \quad i > \lfloor \beta s M \rfloor, \end{cases}$$

and $G_{\alpha,\beta}(s)$ is obtained from (2.7) with $L = \lfloor \alpha s M \rfloor$ and K, c, d, δ as before. See Eicker (1970) for another proof of this general result.

Use Lemma 2.1 for the distribution

$$P\{F_X(x)/F(x) \le s \quad \text{for all} \quad a/M \le F_X(x) \le b/M\}.$$

Then choose $L = a - 1$, $K = b$, $\delta = b/(sM)$ and c, d as before in formula (2.7). For $b = M$, the result was found by Ishii (1959).

The statistic

$$\inf_{a/M \le F_X(x) \le b/M} F_X(x)/F(x)$$

has for $s \ge a/M > 0$ the distribution

$$P\{F(x) \le F_X(x)/s \quad \text{for all} \quad a/M \le F_X(x) \le b/M\} = M! f_M(1)$$

(see Lemma 2.1) if $\nu \equiv 0$ and

$$\mu_i = \begin{cases} a/(Ms) & \text{for all} \quad i = 0, \cdots, a-1 \\ 1 \wedge \dfrac{i}{Ms} & \text{for all} \quad i = a, \cdots, b \\ 1 & \text{for all} \quad i > b. \end{cases}$$

Now we use the dual version (see (2.3)) and get $\mu' \equiv 1$ and

$$\nu_i' = \begin{cases} 0 & \text{for all} \quad i = 0, \cdots, M+1-b \\ \left(\dfrac{i}{Ms} + 1 - \dfrac{1}{s}\right)_+ & \text{for all} \quad i = M+2-b, \cdots, M-a \\ 1 - \dfrac{a}{Ms} & \text{for all} \quad i > M-a. \end{cases}$$

Again, (2.7) is applicable. If $a = 1$, use the outside method for the outer sum in (2.7). Then only a single sum is left (see Chang Li-Chien (1955) and Rényi (1968, (2.20), misprint: the upper boundary of the summation range is $\lceil n/t \rceil - 2$.)

See the next example for two-sided and/or two-sample analogs.

Example 2. The following distribution contains (4.2) as a special case

$$P\left\{-r \leq \frac{F_X(x) - \gamma F(x)}{F(x)} \leq s \qquad \text{for all} \quad \alpha \leq F(x) \leq \beta\right\}.$$

If $\gamma + s$ and $\gamma - r$ are both positive, then

$$\nu_i = \frac{i}{M(\gamma + s)} \qquad \text{for all} \quad i = \lfloor \alpha M(\gamma + s) \rfloor \wedge M + 1, \cdots, \lfloor \beta M(\gamma + s) \rfloor \wedge M$$

$$\mu_i = \frac{i}{M(\gamma - r)} \qquad \text{for all} \quad i = \lceil \alpha M(\gamma - r) \rceil, \cdots, \lceil \beta M(\gamma - r) \rceil - 1.$$

Again, the one-sided case does not need more than a double summation (use 2.7)). See Takács (1967, page 177), Birnbaum and Lientz (1969), and Eicker (1970) for other proofs. In the two-sample version

$$P\left\{-r \leq \frac{N}{M+N} \cdot \frac{F_X(x) - \gamma F_Y(x)}{F_V(x)} \leq s,\ a/(M+N) \leq F_V(x) \leq b/(M+N)\right\}$$

we get from Lemma 3.2 that

$$\nu_i = \left\lceil i \frac{N/M - s}{\gamma + s} \right\rceil - 1$$

$$\text{for all} \quad i = \lfloor aM(\gamma + s)/(N + \gamma M) \rfloor \wedge M + 1, \cdots, \lfloor bM(\gamma + s)/(N + \gamma M) \rfloor \wedge M,$$

$$\mu_i = \left\lceil i \frac{N/M + r}{\gamma - r} \right\rceil$$

$$\text{for all} \quad i = \lceil aM(\gamma - r)/(N + \gamma M) \rceil, \cdots, \lceil bM(\gamma - r)/(N + \gamma M) \rceil - 1.$$

Example 3.

$$P\left\{-r \leq \frac{F_X(x) - \gamma F(x)}{1 - \kappa F(x)} \leq s \qquad \text{for all} \quad \alpha \leq F(x) \leq \beta\right]$$

yields

$$\nu_i = \frac{i/M - s}{\gamma - s\kappa}$$

$$\text{for all} \quad i = \lfloor \alpha M(\gamma - s\kappa) + Ms \rfloor \wedge M + 1, \cdots, \lfloor \beta M(\gamma - s\kappa) + Ms \rfloor \wedge M,$$

$$\mu_i = \frac{i/M + r}{\gamma + r\kappa} \tag{4.3}$$

$$\text{for all} \quad i = \lceil \alpha M(\gamma + r\kappa) - Mr \rceil_+, \cdots, \lceil \beta M(\gamma + r\kappa) - Mr \rceil_+ - 1,$$

where $\gamma - s\kappa$ and $\gamma + r\kappa$ have to be positive. The one-sided case for $\gamma = \kappa = 1$ has been studied by Birnbaum and Lientz (1969). If $\alpha = \kappa = 0$ and $\beta = \gamma = 1$, the well known Kolmogorov-Smirnov distribution is obtained; namely

$$P\{-\rho/M \leq F_X(x) - F(x) \leq \sigma/M \qquad \text{for all} \quad x \in \mathbb{R}\} = M! f_M(1),$$

if (f_n) is the μ-Sheffer sequence for D with roots in ν, where

$$\nu_i = (i - \sigma)_+/M \text{ and } \mu_i = (i + \rho)/M \qquad \text{for all} \quad i \in \mathbb{N}_0. \tag{4.4}$$

These simple functions ν and μ allow the following simplified application of Algorithm A.1.

First, observe that the probability is zero if $\nu_M \geq \mu_{M-i}$. Therefore, we assume that $\Delta =$

$\sigma + \rho$ is greater than 1. Let K be the largest integer such that $\nu_K < \mu_0$, i.e., $K = \lfloor \Delta \rfloor$. From (A.8) we get

$$(4.5)\quad j!f_j(\mu_0) = (\rho/M)^j - \frac{\Delta - j}{M^j} \sum_{i=\lfloor\sigma\rfloor+1}^{j} \binom{j}{i} (i-\sigma)^i (\Delta - i)^{j-i-1} \qquad \text{for all} \quad j = 0, \cdots, K.$$

These are our $K + 1$ starting values. We will use Algorithm A.1 to show how to compute $f_j(\mu_{i+1})$ for all $j = i + 1, \cdots, K + i + 1$, if $f_j(\mu_i)$ is already known for all $j = i, \cdots, K + i$. Assume for the moment that Δ is not an integer. Now $\mu_i < \nu(K + i + 1) < \mu_{i+1}$. Hence, we get from step (a) in Algorithm A.1

$$f_j(\mu_{i+1}) = \sum_{k=i+1}^{j} f_k[\nu(K + i + 1)][(\Delta - K)/M]^{j-k}/(j-k)!.$$

The unknown $f_k(\nu(K + i + 1))$ in this formula is also obtained from step (a):

$$\begin{aligned} f_j(\mu_{i+1}) &= \sum_{k=i+1}^{j} \left[\sum_{\ell=i+1}^{k} f_\ell(\mu_i) \left(\frac{K+1-\Delta}{M}\right)^{k-\ell} \Big/ (k-\ell)! \right] \left(\frac{\Delta - k}{M}\right)^{j-k} \Big/ (j-k)! \\ &= \sum_{\ell=i+1}^{j} \frac{M^{\ell-j}}{(j-\ell)!} f_\ell(\mu_i) \qquad \text{for all} \quad j = i+1, \cdots, K+i, \end{aligned}$$

and

$$f_{K+i+1}(\mu_{i+1}) = \sum_{\ell=i+1}^{K+i} \frac{M^{\ell-K-i-1}}{(K+i+1-\ell)!} f_\ell(\mu_i)[1 - (K+1-\Delta)^{K+i+1-\ell}].$$

(If Δ is an integer, then $\nu(K + i) = \mu_i$, and the right hand side in the equation above equals zero, as it has to.)

As soon as $K + i$ reaches M, we can omit the computation of $f_j(\mu_{i+1})$ for all $j > M$. The algorithm stops when $i = \lfloor M - \rho \rfloor$. Now $\mu_i < 1 < \mu_{i+1}$, and a final step yields the desired probability, namely

$$M!f_M(1) = M! \sum_{\ell=i+1}^{M} f_\ell(\mu_i)(1-\mu_i)^{M-\ell}/(M-\ell)!.$$

No alternating summation occurs in this algorithm, which can be seen as a compact form of Durbin's (1973, (2.4.4)) matrix multiplication method. If ρ is an integer, the final step is not necessary, because we get $f_M(1) = f_M(\mu_{M-\rho})$. If σ and ρ are both integer and equal, the algorithm of Massey (1950, page 117) is obtained.

Again, let σ and ρ be positive real numbers. Define $g_n(x) = M^n f_n(x/M)$ for all $n \in \mathbb{N}_0$. It is easy to check that (g_n) is the $(i + \rho)$-Sheffer sequence for D with roots in $(i - \sigma)_+$. Hence, (g_n) is independent of M. The generating function identity (A.19) can be applied to (g_n) as follows.

$$(4.6)\quad \begin{aligned} &\sum_{i\ge0} \frac{(ti)^i}{i!} P\{-\rho/i \le F_{X,i}(x) - F(x) \le \sigma/i \text{ for all } x \in \mathbb{R}\} \\ &\qquad = \sum_{i\ge0} t^i g_i(i) = \varphi_\rho(t) = \frac{\gamma_\rho(t)\gamma_\sigma(t)}{\gamma_{\rho+\sigma}(t)}, \end{aligned}$$

where we made use of

$$\varphi_\Delta(t) = \sum_{i\ge0} g_i(i + \rho - \Delta)t^i = \sum_{i=0}^{\lfloor\sigma\rfloor} (i-\sigma)^i t^i = \gamma_\sigma(t)$$

(see A.18); ($F_{X,i}$ in (4.6) means of course the empirical distribution function of a sample X of size i.) Identity (4.6) is due to Kemperman (1961, (5.40)) and is useful for asymptotic results.

In the one sided case ($\mu \equiv 1$) and for general ν as in (4.3), it may be necessary, at worst, to work with a double summation in the closed expression (2.7). Smirnov (1944) solved the

Kolmogorov-Smirnov case. For more generating functions (not obtainable by our method) and asymptotic results see Lauwrier (1963). A detailed study of the exact distribution has been done by Alter, Govindarajulu, and Gragg (1975).

Replacing $F(x)$ by $F_V(x)$ in the general distribution, the two sample version is obtained. ν and μ are easily derived. We restrict ourselves to the Kolmogorov-Smirnov case

$$P\{-\rho/N \le F_X(x) - F_Y(x) \le \sigma/N \quad \text{for all} \quad x \in \mathbb{R}\} = f_M(N) \Big/ \binom{M+N}{N}, \tag{4.7}$$

where $\nu_i = \lceil ic - \sigma \rceil_+ - 1$ and $\mu_i = \lfloor ic + \rho \rfloor$ for all $i \in \mathbb{N}_0$, and $c = N/M$. Massey (1951) derived the recursion (3.10) for this case. If M divides N, i.e., if c is an integer, then σ and ρ can be chosen as integers without loss of generality. For this case the generating function identity (A.19) with $\Delta = \rho + \sigma + 1$ yields

$$\sum_{i\ge 0} t^i \binom{ic+i}{i} P\left\{-\frac{\rho}{ci} \le F_{X,i}(x) - F_{Y,ic}(x) \le \frac{\sigma}{ci} \quad \text{for all} \quad x \in \mathbb{R}\right\}$$

$$= \sum_{i\ge 0} f_i(i)\, t^i = \varphi_\rho(t) = \frac{\gamma_\rho(t)\gamma_\sigma(t)}{\gamma_{\rho+\sigma+1}(t)},$$

where we made use of

$$\varphi_{\rho+\sigma+1}(t) = \sum_{i\ge 0} f_i(ic - \sigma - 1)t^i = \sum_{i=0}^{\lfloor \sigma/c \rfloor} \binom{i(c+1) - \sigma - 1}{i} t^i = \gamma_\sigma(t)$$

(see (A.18); Kemperman (1961)).

If $N = M$ in (4.7), apply (3.7). In the one sample case ($\mu \ge N$) use (3.6). Use (3.8) if c is any integer greater than 1.

The condition $N/M \in \mathbb{N}_1$ is sufficient, but not necessary for obtaining a simple function ν. Obviously, if $\sigma \ge M - 2$, ν cannot contain more than two different affine pieces (between 0 and M, of course). If $N = cM + \epsilon r$ ($|\epsilon| = 1$; $r \le M/2$; $c, r \in \mathbb{N}_0$), then $r + 2 - \lceil (r(L+1) - s_\epsilon)/M \rceil$ is a sharp upper bound for the number of different affine pieces, where $L = \lfloor \sigma M/N \rfloor$,

$$s_\epsilon = \begin{cases} M\sigma - M\lfloor \sigma \rfloor & \text{if} \quad \epsilon = 1, \\ M - s_{-\epsilon} - 1 & \text{if} \quad \epsilon = -1, \end{cases}$$

and $\sigma M = aN - bM$ with $a, b \in \mathbb{N}_0$, $0 \le a \le M$, $0 \le b \le N$. (It is always possible to choose σ in this way without changing the probability; see Niederhausen (1978, Section 18) for details.) If $r = 1$, the number of different pieces cannot exceed three and the function ν equals

$$\nu(i) = \begin{cases} -1 & \text{for all} \quad 0 \le i \le L \\ ic - \lfloor \sigma \rfloor - 1 & \text{for all} \quad L < i \le s_\epsilon \\ ic - \lfloor \sigma \rfloor - 1 + \epsilon & \text{for all} \quad s_\epsilon < i \le M. \end{cases}$$

Compute the corresponding probability from (3.9). This case has been investigated for $c = 1$ by Hodges (1957), and for general $c \in \mathbb{N}_1$ by Steck (1969).

Example 4. Borokov and Sycheva (1968) found the exact and the asymptotic distribution of

$$W_M^+ = \sup_{\theta_1 \le F(x) \le \theta_2} \frac{F_X(x) - F(x)}{[F(x)(1 - F(x))]^{1/2}}.$$

We call the test statistic W_M if the supremum of the absolute value of the ratio is taken. Let

$$h^\pm(i) = \frac{2i + s \pm [s^2 + 4si(1-i)]^{1/2}}{2(1+s)}$$

and

$$c^{\pm}(\gamma) = M(\gamma \pm [s\gamma(1-\gamma)]^{1/2}).$$

We get from Lemma 2.2 that

$$P(W_M \leq s^{1/2}) = M! f_M(1),$$

if (f_n) is the μ-Sheffer sequence for D with roots in ν, where

(4.8) $$\nu_i = h^-(i/M) \quad \text{for all} \quad i = \lfloor c^+(\theta_1) \rfloor \wedge M + 1, \cdots, \lfloor c^+(\theta_2) \rfloor \wedge M,$$

and

(4.9) $$\mu_i = h^+(i/M) \quad \text{for all} \quad i = \lceil c^-(\theta_1)_+ \rceil, \cdots, \lceil c^-(\theta_2) \rceil_+ - 1.$$

The following short table of the percentage points of $M^{1/2}W_M$ was computed by Algorithm A.1 and by the outside method (2.9) if applicable. We chose always $\theta_1 = \theta = 1 - \theta_2$ for $\theta = 0$, .01, .05, .1 and .25. Let

$$P(z) = P(M^{1/2} W_M \leq z).$$

We consider the significance probabilities $\alpha = 1 - P(z_\alpha)$ for $\alpha = .1$, .05 and .01. Because of the discontinuities, these levels cannot always be attained. If the absolute difference between α and $1 - P(z_\alpha)$ is less than 0.000005, this small discontinuity is not noted in the tables, and z_α is rounded to 4 digits after the decimal point. If

$$.000005 \leq |\alpha - 1 + P(z_\alpha)| < .005,$$

and α is greater (smaller) than $1 - P(z_\alpha)$, then five digits are given and a bar is placed under (over) the last digit. This last digit is not rounded. Decreasing (increasing) it by one yields a probability greater (smaller) than α. Two bars indicate an absolute difference between .005 and .013. The asymptotic values of Borokov and Sycheva (1968, Theorem 3A) are given in the last row of Table 1(b)–(e).

Table 1(a) is a confirmation of Noé's (1972) computations. In Tables 1(b) and 1(c) the results of Canner's (1975) simulation study are given in parentheses. In Tables 1(d) and 1(e) the rows marked by F contain the percentage points of $M^{1/2}W_M$ as in 1(a)–(c). The rows marked by F_X refer to the corresponding statistic where the supremum is taken over $d/M \leq F_X(x) \leq 1 - d/M$. The integer d is chosen such that d/M is closest to the desired θ, i.e.

$$d = \begin{cases} \lfloor M\theta \rfloor & \text{if} \quad M\theta - \lfloor M\theta \rfloor < 0.5 \\ \lceil M\theta \rceil & \text{else.} \end{cases}$$

The F_X-row in Table 1(d) equals the F-row when $M = 10$ and is therefore omitted.

We denote by $W_{M,N}$ the two-sample version of W_M, that is

$$W_{M,N} = \sup_{\theta_1 \leq F_V(x) \leq \theta_2} \frac{|F_X(x) - F_Y(x)|}{[F_V(x)(1 - F_V(x))]^{1/2}},$$

where $\theta_1 = a/(M+N)$ and $\theta_2 = b/(M+N)$; a and b integer.

Now we get from Lemma 3.2

$$P\left(\frac{N}{M+N} W_{M,N} \leq s^{1/2}\right) = f_M(N) \Big/ \binom{M+N}{M},$$

if (f_n) is the $\hat{\mu}$-Sheffer sequence for ∇ with roots in $\tilde{\nu}$, where

$$\nu_i = \lceil (M+N)h^-(i/M) \rceil - i - 1 \quad \text{and} \quad \mu_i = \lfloor (M+N)h^+(i/M) \rfloor - i,$$

TABLE 1

Percentage points z_α of $M^{1/2}W_m$

M	$\alpha = .1$	$\alpha = .05$	$\alpha = .01$
10	4.6146	6.4257	14.1863
20	4.6423	6.4398	14.1908
50	4.6631	6.4488	14.1929
100	4.6719	6.4519	14.1931

(a) $\theta = 0.$

M	$\alpha = .1$	$\alpha = .05$	$\alpha = .01$
10	3.2900	3.9829	$6.0385\bar{9}$
		(3.33)	(5.70)
20	3.3962	$4.0451\bar{9}$	4.9094
		(3.76)	(4.77)
50	3.2029	$3.5533\underline{\underline{4}}$	4.5353
		(3.69)	(4.40)
100	3.0640	3.4379	4.1899
∞	3.05	3.30	3.79

(b) $\theta = .01$

M	$\alpha = .1$	$\alpha = .05$	$\alpha = .01$
10	2.9218	3.4216	4.1705
		(3.33)	(4.00)
20	2.9094	3.1831	$4.1039\bar{1}$
		(3.07)	(3.80)
50	2.8616	3.1525	3.8289
		(3.15)	(3.80)
100	2.8384	3.1417	3.7419
∞	2.89	3.15	3.67

(c) $\theta = .05$

M		$\alpha = .1$	$\alpha = .05$	$\alpha = .01$
10	F	2.7148	3.1336	3.8203
20	F	2.7130	2.9830	$3.7267\bar{7}$
	F_X	3.3938	4.0798	6.1725
50	F	2.7284	3.0071	3.6284
	F_X	2.9641	3.3533	4.2777
100	F	2.7362	3.0120	3.5960
	F_X	2.8562	3.1803	3.8839
∞		2.78	3.05	3.59

(d) $\theta = .1$

M		$\alpha = .1$	$\alpha = .05$	$\alpha = .01$
10	F	2.4383	2.6340	3.2863
	F_X	3.2747	3.9777	6.0714
20	F	2.4694	2.7236	3.2852
	F_X	2.7419	3.1507	4.0971
50	F	2.4890	$2.7760\underline{9}$	3.3414
	F_X	2.6223	2.9530	3.6498
100	F	2.5159	2.7929	3.3568
	F_X	2.5657	2.8727	3.4969
∞		2.53	2.83	3.40

(e) $\theta = .25$

TABLE 2

Percentage points for $\left(\frac{MN}{M+N}\right)^{1/2} W_{M,N}$

M = N	$\alpha = .1$	$\alpha = .05$	$\alpha = .01$
10	$2.344\underline{\underline{1}}$	$2.683\bar{2}$	$3.146\bar{2}$
		(2.71)	(3.17)
20	$2.581\bar{9}$	$2.760\underline{3}$	$3.227\bar{4}$
		(2.70)	(3.18)
50	$2.700\underline{7}$	$2.948\bar{8}$	$3.429\bar{9}$
		(2.93)	(3.44)
100	$2.791\bar{4}$	$3.024\bar{9}$	$3.502\underline{2}$
		(3.02)	(3.47)
500	$2.944\underline{1}$	$3.186\bar{3}$	$3.669\underline{4}$

(a) $\theta = .01$

M	N		$\alpha = .1$	$\alpha = .05$	$\alpha = .01$
100	**100**	F_V	$2.779\bar{5}$	$3.008\underline{9}$	$3.502\underline{2}$
		F_X	$2.736\underline{9}$	$2.982\underline{0}$	$3.472\underline{4}$
100	**99**	F_V	$2.774\bar{7}$	$3.015\bar{7}$	$3.505\underline{1}$
		F_X	$2.752\underline{4}$	$2.189\bar{3}$	$3.478\bar{1}$
100	**98**	F_V	$2.760\underline{7}$	$3.014\bar{6}$	$3.511\underline{2}$
		F_X	$2.747\underline{0}$	$2.992\underline{2}$	$3.468\bar{6}$
100	**95**	F_V	$2.766\bar{9}$	$3.023\bar{9}$	$3.505\bar{7}$
		F_X	$2.742\underline{6}$	$2.993\bar{9}$	$3.486\underline{3}$
500	**500**	F_V	$2.842\bar{3}$	3.0990	$3.613\underline{6}$

(b) $\theta = .05$

$M = N$	$\alpha = .1$	$\alpha = .05$	$\alpha = .01$
50	$2.666\underline{7}$	$2.896\bar{8}$	$3.429\bar{9}$
100	$2.700\underline{3}$	$2.971\bar{1}$	$3.472\bar{0}$
500	$2.741\bar{5}$	$3.013\underline{9}$	3.5468

(c) $\theta = .1$

$M = N$	$\alpha = .1$	$\alpha = .05$	$\alpha = .01$
20	$2.363\underline{1}$	$2.652\underline{0}$	$3.187\underline{0}$
50	$2.472\underline{3}$	$2.735\bar{7}$	$3.296\bar{3}$
100	$2.482\bar{9}$	$2.758\underline{0}$	$3.313\underline{1}$
500	$2.522\underline{7}$	2.8028	$3.362\underline{3}$

(d) $\theta = .25$

TABLE 3

$p_K = P\{\sup_{x \in R} \mid F_X(x) - F_{-X}(x) \mid \geq K/M\}$.

M \ K	4	5	6	7	8	9	10
4	.125						
6	.250	.063	.031				
8	.359	.141	.078	.016	.008		
10	.453	.219	.131	.043	.023	.004	.002

TABLE 4

Percentage points for $M^{1/2}A_M$.

M	$\alpha = .1$	$\alpha = .05$	$\alpha = .01$	M	$\alpha = .1$	$\alpha = .05$	$\alpha = .01$
10	$2.121\bar{3}$	$2.333\bar{3}$	$2.645\bar{8}$	10	$1.999\bar{9}$	$2.236\underline{1}$	$2.529\underline{9}$
20	$2.236\underline{1}$	$2.496\underline{2}$	$2.982\bar{4}$	20	$2.236\underline{0}$	$2.449\underline{5}$	$2.886\underline{8}$
50	$2.428\underline{6}$	$2.654\bar{9}$	$3.152\bar{9}$	50	$2.334\underline{9}$	$2.600\underline{0}$	$3.123\underline{6}$
100	$2.496\underline{2}$	$2.722\underline{2}$	3.2497	100	$2.449\underline{5}$	$2.710\underline{7}$	$3.207\underline{2}$
∞	2.74	3.02	3.56	∞	2.56	2.84	3.41

(*a*) $\theta = .01$ (*b*) $\theta = .05$

M	$\alpha = .1$	$\alpha = .05$	$\alpha = .01$	M	$\alpha = .1$	$\alpha = .05$	$\alpha = .01$
10	$1.999\bar{9}$	$2.236\underline{1}$	$2.529\underline{9}$	10	$1.897\underline{4}$	$2.121\bar{3}$	$2.449\underline{5}$
20	$2.182\underline{9}$	$2.449\bar{4}$	$2.840\underline{2}$	20	$2.110\bar{5}$	$2.309\bar{4}$	$2.713\underline{7}$
50	$2.333\bar{3}$	$2.558\underline{5}$	$3.086\bar{9}$	50	$2.142\underline{9}$	$2.428\underline{6}$	$3.000\underline{0}$
100	$2.363\bar{5}$	$2.666\bar{6}$	$3.183\bar{1}$	100	$2.165\bar{2}$	$2.465\underline{9}$	$3.037\underline{8}$
∞	2.44	2.74	3.31	∞	2.23	2.53	3.12

(*c*) $\theta = .1$ (*d*) $\theta = .25$

with i in the same range as in (4.8) and (4.9). (See (3.11) and (3.12) for $\tilde{\nu}$ and $\hat{\mu}$.) The preceding table of percentage points for $\left(\frac{MN}{M+N}\right)^{1/2} W_{M,N}$ was computed using only algorithm (3.10). Discontinuities occur at almost each entry. The bars are set following the same rules as above, but only four digits are given. The table for $\theta = 0$ is the same as Table 2(a) for $\theta = 0.01$ and is therefore omitted. The numbers in parentheses are taken from Canner's (1975) simulation study (computed for $\theta = 0$). For $\theta = 0.05$, the rows $M = N = 10$, 20 and 50 are equal to those in Table 2(a) and are omitted. Instead, we demonstrate the effect of slightly different, but large sample sizes. Again, the rows are marked by F_X, if the supremum is taken over all $a'/M \leq F_X(x) \leq b'/M$. In Tables 2(c) and 2(d) those rows are omitted which do not differ from Table 2(a). The asymptotic values of Tables 1(b)–1(e) may be used for comparison. For applications see Doksum and Sievers (1976).

Example 5. *Kuiper's statistic.*

$$\sup_{x \in R}(F_X(x) - F(x)) + \sup_{x \in R}(F(x) - F_X(x))$$

was originally suggested by Kuiper (1960) as a Kolmogorov-Smirnov test on the circle. Durbin (1973) proved that

$$\begin{aligned} &P\{\sup_{x \in R}(F_X(x) - F(x)) + \sup_{x \in R}(F(x) - F_X(x)) \leq \tau/M\} \\ &= M \cdot P\{Mx - 1 \leq (M-1)F_{U,M-1}(x) \leq Mx + \tau - 1 \quad \text{for all} \quad 0 \leq x \leq 1\}. \end{aligned} \tag{4.10}$$

Hence, ν and μ are similar to (4.4), specifically

$$\nu_i = (i - \tau + 1)_+/M \quad \text{and} \quad \mu_i = (i+1)/M \qquad \text{for all} \quad i = 0, \cdots, M-1,$$

and the probability (4.10) equals $M!f_{M-1}(1)$, where (f_n) is the μ-Sheffer sequence for D with roots in ν. For computing $f_{M-1}(1)$ the same methods are available as for the ordinary Kolmogorov-Smirnov test in Example 3. Choose $\sigma = \tau - 1$ and $\rho = 1$ in (4.4). The simplified version of Algorithm A.1 can be used (see (4.5) and what follows). The alternating recursion (2.8) is now basically the same as the recursion in Durbin (1973, page 35). Stephens (1965) derived explicit expressions under special assumptions about σ.

Durbin (1973) proved also the two sample analog of (4.10). We get from his result

$$P\{\sup_{x\in\mathbb{R}} (F_X(x) - F_Y(x)) + \sup_{x\in\mathbb{R}}(F_Y(x) - F_X(x)) \le \tau/M\} = Mf_{M-1}(N) \Big/ \binom{M+N-1}{M-1},$$

if (f_n) is the μ-Sheffer sequence for ∇ with roots in ν, where

$$\nu_i = \lceil (N-1)i/M - (N+M-1)(\tau-1)/M \rceil_+ - 1$$

and

$$\mu_i = \lfloor (N-1)i/M + (N+M-1)/M \rfloor \qquad \text{for all} \quad i = 0, \cdots, M-1.$$

The same considerations as for the Kolmogorov-Smirnov two sample test in Example 3 are valid. In particular, the generating function identity (A.19) can be formulated for the two sample and the one sample case.

Example 6. *The Butler test.*

$$P\{-K/M \le F_X(x) - F_{-X}(x) \le L/M \qquad \text{for all} \quad x \in \mathbb{R}\} = P_M(K, L),$$

with $K, L \in \mathbb{N}_0$; $K \le M$; $L \le M$, is the distribution of the Butler test (Butler, 1969) for symmetry about zero. Here F_{-X} denotes the empirical distribution function of $-X_1, \cdots, -X_M$. If the vector T is defined in (3.1) with $Y_j = -X_j (j = 1, \cdots, M)$, we obtain

$$T_\ell = 1 - T_{2M+1-\ell} \qquad \text{for all} \quad \ell = 1, \cdots, 2M.$$

Denote the set of all such vectors by $\Upsilon_s(M, M) \subset \Upsilon(M, M)$. Then $\#\Upsilon_s(M, M) = 2^M$ and

$$\begin{aligned} &\#\Upsilon_s(M, M \mid -K \le 2T'_\ell - \le L \qquad \text{for all} \quad \ell = 0, \cdots, 2M) \\ &= \#\Upsilon_s(M, M \mid -K \le 2T'_\ell - \ell \le L \qquad \text{for all} \quad \ell = 0, \cdots, M) \\ &= \textstyle\sum_{i=0}^{M} \#\Upsilon(\mathrm{i}, M - i \mid -K \le 2T'_\ell - \ell \le L \qquad \text{for all} \quad \ell = 0, \cdots, M). \end{aligned}$$

From (3.7) we obtain, for $J = K + L + 2$

$$\begin{aligned} P_M(K, L) &= 2^{-M} \textstyle\sum_{i=\lceil (M-K)/2 \rceil}^{\lfloor (M+L)/2 \rfloor} \sum_{\ell \ge 0} \left\{ \binom{M}{i + (\ell+1)J} + \binom{M}{i - \ell J} - \binom{M}{i + \ell J + K + 1} \right. \\ &\quad \left. - \binom{M}{i - \ell J - L - 1} \right\} = 2^{-M} \textstyle\sum_{\ell=-\infty}^{+\infty} \sum_{i=\lceil (M-K)/2 \rceil}^{\lfloor (M+L/2 \rfloor} \left\{ \binom{M}{i + \ell J} - \binom{M}{i + \ell J + K + 1} \right\} \end{aligned} \tag{4.11}$$

if $X_1, \cdots, X_M$ are i.i.d. random variables with a continuous distribution function, symmetric about zero. The one-sided case $(K = M)$ can be written in various different forms as follows.

$$\begin{aligned} P_M(M, L) &= 2^{-M} \textstyle\sum_{i=0}^{\lfloor (L+M)/2 \rfloor} \left\{ \binom{M}{i} - \binom{M}{i - L - 1} \right\} \\ &= 2^{-M} \textstyle\sum_{i=\lfloor (M-L)/2 \rfloor}^{\lfloor (M+L)/2 \rfloor} \binom{M}{i} \end{aligned}$$

$$= 1 - 2^{1-M} \sum_{i=0}^{\lfloor (M-L)/2 \rfloor - 1} \binom{M}{i} - 2^{-M} \binom{M}{(M-L-1)/2}$$

$$= 1 - B_{M,.5}((M-L-2)/2) - B_{M,.5}((M-L-1)/2)$$

$$= \begin{cases} P_{M+1}(M+1, L) & \text{if } M-L \text{ is odd} \\ P_{M-1}(M-1, L) & \text{if } M-L \text{ is even,} \end{cases}$$

where $B_{M,.5}(x) = 2^{-M} \sum_{i=0}^{\lfloor x \rfloor} \binom{M}{i}$ is the binomial distribution function with probability $p = 0.5$ of success. Thus, no extra tables are needed for the one-sided case. The case $K = L$ in (4.11) has been considered first by Smirnov (1947) (I owe this reference to W. R. Pirie) and was rediscovered by Butler (1969), the result being

$$P_M(K, K) = 2^{-M} \sum_{\ell=-\infty}^{\infty} (-1)^{\ell} \sum_{i=\lceil (M-K)/2 \rceil}^{\lfloor (M+K)/2 \rfloor} \binom{M}{i + \ell(K+1)}$$

$$= 2 \sum_{\ell \geq 0} (-1)^{\ell} P_M[M, (2\ell + 1)(L+1))].$$

For tables of $P_M(K, K)$ one should use

$$P_M(K, K) = \begin{cases} P_{M+1}(K, K) & \text{if } M-K \text{ is odd} \\ P_{M-1}(K, K) & \text{if } M-K \text{ is even.} \end{cases} \tag{4.12}$$

If the distribution functions F_i of X_i are not all equal, the distribution of the Butler statistic does not change. This result is due to Chatterjee and Sen (1973). Because of several misprints in their Table 2, we give here a small table of $P_K = 1 - P_M(K-1, K-1)$. Use (4.12) for odd M.

The test statistic

$$A_M = \sup_{a/M \leq F_X(x) \leq b/M} \frac{|F_X(x) - F_{-X}(x)|}{\{1 - |1 - F_X(x) - F_{-X}(x)|\}^{1/2}} \tag{4.13}$$

has been introduced by Aaberge, Doksum and Fenstad (1977), who found the asymptotic distribution for the one-sided case. The exact distribution can be computed recursively, if the functions ν and μ are defined by

$$\nu(i) = \begin{cases} -1 & \text{for all } i = 0, \cdots, a-1 \\ i - \lfloor (4is + s^2)^{1/2} - s \rfloor - 1 & \text{for all } i = a, \cdots, b \\ \nu(b) & \text{for all } i > b, \end{cases}$$

$$\mu(i) = \begin{cases} \mu(a) & \text{for all } i = 0, \cdots, a-1 \\ i + \lfloor (4ir + r^2)^{1/2} + r \rfloor & \text{for all } i = a, \cdots, b \\ M & \text{for all } i > b. \end{cases}$$

Now

$$P\{A_M \leq (2s/M)^{1/2}\} = 2^{-M} \sum{}' f_i(M-i),$$

where the summation ranges over all i with $i + \tilde{\nu}_i \leq M \leq i + \hat{\mu}_i$, and (f_n) is the $\hat{\mu}$-Sheffer sequence for ∇ with roots in $\tilde{\nu}$ (see (3.11) and (3.12) for $\tilde{\nu}$ and $\hat{\mu}$). We give some percentage points for $M^{1/2}A_M$ in Table 4, where $\alpha = 1 - P(M^{1/2}A_M \leq z_\alpha)$. Discontinuities are marked by bars and $a = M - b$ is chosen close to θ as in Example 4. The asymptotic values are computed from Theorem 3*A*a) in Borokov/Sycheva (1968), as suggested in the work of Aaberge, Doksum and Fenstad (1977). Our asymptotic results differ slightly from their Table 3, because we use one more term in the asymptotic expansion.

APPENDIX

The purpose of this appendix is to summarize the algebraic results which we used in the preceding chapters. All proofs are straightforward verifications of the definitions. Hence, we give only some hints and leave the details to the reader. A more general approach including Eulerian polynomials can be found in Niederhausen (1980). The "Finite Operator Calculus" of Rota, Kahaner and Odlyzko (1973) is the fundament of the whole theory.

Let **P** be the algebra of polynomials over a field K with characteristic zero. In our rank test applications K always equals $\mathbb{Z}$, for the order tests choose $K = \mathbb{R}$. We will deal with **linear** operators $\mathbf{P} \to \mathbf{P}$ only, and omit the word "linear" in the sequel. For all $a \in K$ the **shift operator** is denoted by $E^a: p(x) \mapsto p(x+a)$. An operator Q on **P** is a delta operator, if

$$Q \text{ is shift invariant: } QE^a = E^aQ \quad \text{for all} \quad a \in K, \text{ and}$$

$$Qx \text{ is a non-zero constant.}$$

The derivative operator D is a delta operator if $K = \mathbb{R}$, and the following properties show how Q generalizes D:

(A.1) $$Qa = 0 \text{ for every constant } a$$

(A.2) $$\deg(Qp) = \deg(p) - 1 \text{ for each } p \in \mathbf{P} \text{ with } \deg(p) \geq 1;$$

see Rota *et al* (1973, page 687). Hence, the kernel of Q consists only of the constant polynomials. A sequence of polynomials $(s_n)_{n \in \mathbb{N}_0}$ is a **Sheffer sequence** for Q, if

(A.3) $$s_0 \text{ is a non-zero constant}$$

(A.4) $$Qs_n = s_{n-1} \quad \text{for all} \quad n \geq 1.$$

We make the convention $s_n = 0$ if $n < 0$. For instance, $(x^n/n!)$ is a Sheffer sequence for D.

Lemma A.1. *If (s_n) is a Sheffer sequence for Q then* $\deg(s_n) = n$.

Proof. (A.1)–(A.4) □

Lemma A.2. *If (s_n) and (t_n) are both Sheffer sequences for Q with the property*

$$s_n(\nu_n) = t_n(\nu_n)$$

for a given sequence (ν_n) in K, then the two sequences are equal.

Proof. Induction over n. Use $\ker(Q) =$ constant functions □

$$(s_n) \text{ has **roots** in } \nu: \mathbb{N}_0 \to K, \text{ say, if } s_n(\nu_n) = \delta_{0,n}$$

for all $n \in \mathbb{N}_0$. The Sheffer sequence for Q with roots in 0 is called the **basic sequence** for Q and always denoted by (q_n). Obviously,

(A.5) $$(x^n/n!) \text{ is the basic sequence for } D.$$

It is easy to verify that

(A.6) $$\left(\binom{x+n-1}{n}\right)_{n \in \mathbb{N}_0} \quad \text{is the basic sequence for } \nabla = I - E^{-1}.$$

More examples can be found in Rota *et al* (1973).

Immediately from the shift-invariance follows: If (s_n) is a Sheffer sequence for Q with roots in ν, then $(E^a s_n)$ is a Sheffer sequence for Q with roots in $\nu - a$.

Deeper than all the other results in this appendix is the following

LEMMA A.3. *If $\nu(n) = an + b$ $(a, b \in K)$, then*

$$s_n(x) = (x - an - b)(x - b)^{-1} q_n(x - b)$$

defines the Sheffer sequence for Q with roots in ν. (For $n = 0$ we have to define $\frac{0}{0} = 1$.)

PROOF. See Niederhausen (1978, equation (16.2); 1980, equation (2.4) □

Now we come to a representation theorem for Sheffer sequences with roots in

$$\nu(i) = \begin{cases} \varphi(i) & \text{for all} \quad 0 \le i \le L \\ ci + d & \text{for all} \quad i > L, \end{cases}$$

where $L \in \mathbb{N}_0$; $c, d \in K$ and $\varphi: \mathbb{N}_0 \to R$ arbitrary.

THEOREM A.1. *If (s_n) is the Sheffer sequence for Q with roots in ν as above, then*

$$\text{(A.7)} \qquad \begin{aligned} s_n(x) &= \textstyle\sum_{i=0}^{L} s_i(ci + d)(x - cn - d) \\ &\quad \cdot (x - ci - d)^{-1} q_{n-i}(x - ci - d) \qquad \text{for all} \quad n \in \mathbb{N}_0. \end{aligned}$$

PROOF. Check recurrence and side conditions, using Lemma A.3 □

COROLLARY A.1. (**Binomial Theorem**). *If (s_n) is the Sheffer- and (q_n) the basic sequence for Q, then*

$$s_n(x + y) = \textstyle\sum_{i=0}^{n} s_i(y) q_{n-i}(x) \qquad \textit{for all} \quad n \in \mathbb{N}_0.$$

PROOF. Choose $c = 0$, $d = y$ and $L = \infty$ in (A.7) □

Avoiding alternating summation in (A.7), it may be sometimes preferable to use the "outside method" (a term, introduced by Hodges(1957)):

$$\text{(A.8)} \qquad s_n(x) = r_n(x) - \textstyle\sum_{i=L+1}^{n} r_i(ci + d)(x - cn - d)(x - ci - d)^{-1} q_{n-i}(x - ci - d)$$

where (r_n) is the Sheffer sequence for Q with roots in φ (follows by summation over *all* $i = 0, \cdots, n$ in (A.7)).

Repeated use of (A.7) yields a representation of the Sheffer sequence (s_n) for Q with roots defined in the piecewise affine function

$$\nu(i) = ia_j + b_j \qquad \text{for all} \quad L_j < i \le L_{j+1},$$

where $-1 = L_0 < L_1 < \cdots$, each L_j integer, and $a_j, b_j \in K$ for all $j \in \mathbb{N}_0$. Then for all $L_j < n \le L_{j+1}$

$$\text{(A.9)} \qquad s_n(x) = \textstyle\sum_{k_j=0}^{L_j} \cdots \sum_{k_1=0}^{L_1} p_j(x) p_{j-1}(\nu_j(k_j)) \cdots p_0(\nu_1(k_1)),$$

if

$$p_i(x) = \frac{x - \nu_i(k_{i+1})}{x - \nu_i(k_i)} q_{k_{i+1}-k_i}(x - \nu_i(k_i)),$$

where $k_0 = 0$ and $k_{j+1} = n$. Because of its importance we explicitly write down the special case of (A.9) where

$$\nu(i) = \begin{cases} ia + b & \text{for all} \quad i = 0, \cdots, L \\ ic + d & \text{for all} \quad i > L. \end{cases}$$

Then

$$\text{(A.10)} \qquad s_n(x) = \textstyle\sum_{i=0}^{L} \dfrac{i(c - a) + d - b}{ic + d - b} q_i(ci + d - b) \dfrac{x - cn - d}{x - ci - d} q_{n-i}(x - ic - d).$$

For $n \leq L$, the right hand side equals $(x - an - b)(x - b)^{-1}q_n(x - b)$ by Lemma A.3.

Now we assume that K is completely ordered. Let μ: $\mathcal{N}_0 \to K$ be a non-decreasing function and $(t_{n,i})_{n,i\in\mathcal{N}_0}$ be a double sequence in **P** with the properties

$$\text{(A.11)} \quad \begin{aligned} t_{n,i}(\mu_i) &= t_{n,i+1}(\mu_i) \quad &&\text{for all} \quad 0 \leq i \leq r(n) = \min\{m \in \mathcal{N}_0 \mid \mu(m) = \mu(n)\}, \\ t_{n,i} &= 0 &&\text{for all} \quad i > r(n). \end{aligned}$$

Define an associated sequence (f_n) to $(t_{n,i})$ by

$$\text{(A.12)} \qquad f_n(x) = t_{n,i}(x) \qquad \text{for all} \quad \mu_{i-1} < x \leq \mu_i \ (\mu_{-1} = -\infty).$$

We call (f_n) a **μ-Sheffer sequence**, if $(t_{m+n,r(m)})_{n\in\mathcal{N}_0}$ is a Sheffer sequence for all $m \in \mathcal{N}_0$. From Corollary A.1 we get a first representation of $f_n(x)$

$$\text{(A.13)} \qquad f_n(x) = \textstyle\sum_{k=i}^{n} f_k(y)q_{n-k}(x - y) \quad \text{if} \quad x, y \in [\mu_{i-1}, \mu_i].$$

If (f_n) has roots in ν, i.e.

$$\text{(A.14)} \qquad f_n(\nu_n) = \delta_{0,n} \qquad \text{for all} \quad n \in \mathcal{N}_0,$$

then any value $f_n(z)$ can be computed from (A.13) by stepping through all the intervals $[\mu_j, \mu_{j+1}]$ until z is enclosed. We give only a brief description of this trivial algorithm:

ALGORITHM A.1. Assume $f_{r(j)}(\mu_j), \cdots, f_i(\mu_j)$ are already computed such that $j \leq n$ and $\nu_{i+1} > \mu_j$

(a) If $\nu_{i+1} < \mu_{j+1}$ then define $x = \nu_{i+1}$, $y = \mu_j$, and compute $f_{r(j)}(\nu_{i+1}), \cdots, f_i(\nu_{i+1})$ from (A.13). Of course, $f_{i+1}(\nu_{i+1}) = 0$. Therefore, the i-index increased by one, and it increases again if ν_{i+2} lies also in the same interval (define $x = \nu_{i+2}$ and $y = \nu_{i+1}$). Finally a k is reached such that $\mu(j) < \nu_k < \mu_{j+1} < \nu_{k+1}$ (the case $\nu_k = \mu_{j+1}$ is left to the reader). Then choose $x = \mu_{j+1}$, $y = \nu_k$, and compute $f_{r(j)}(\mu_{j+1}), \cdots, f_k(\mu_{j+1})$ from (A.13). Now we are in the same situation as in the beginning.

(b) If $\nu_{i+1} > \mu_{j+1} > \mu_j$ then define $x = \mu_{j+1}$, $y = \mu_j$, and compute $f_{r(j)}(\mu_{j+1}), \cdots, f_k(\mu_{j+1})$ from (A.13). Again, we are in the same situation as in the beginning.

(c) If $\mu_j = \mu_{j+1}$ increase j by one.

In special cases this algorithm can be simplified, as in Example 4 of Section 4.

A one-dimensional recursion can be obtained from

THEOREM A.2: *Let (f_n) be a μ-Sheffer sequence for Q (with basic sequence (q_n)). If (f_n) is associated to $(t_{n,i})$ then*

$$\text{(A.15)} \qquad t_{n,i}(x) = \textstyle\sum_{k=i}^{n} f_k(\mu_k)q_{n-k}(x - \mu_k) \qquad \textit{for all} \quad n \in \mathcal{N}_0 \textit{ and } i = 0, \cdots, n.$$

PROOF. Verify side conditions (A.11).

See Niederhausen (1980, Theorem 4.1) for a general version of this theorem. The announced one-dimensional recursion follows, when we write (A.15) as

$$\text{(A.16)} \qquad f_n(x) = \textstyle\sum' f_k(\mu_k)q_{n-k}(x - \mu_k) \qquad \text{for all} \quad n \in \mathcal{N}_0,$$

where the summation runs over all k such that $\mu_k > x$. Thus, a system of equations for the unknown $f_k(\mu_k)$ is obtained, if only one value $f_n(\nu_n)$ with $\nu_n \leq \mu_n$ is known for each n. By Cramer's rule, $f_n(\mu_n)$ can be expressed as a determinant.

COROLLARY A.2: *If (f_n) is a μ-Sheffer- and (q_n) the basic sequence for Q, then*

$$f_n(\mu_n) = \det(\alpha_{i,j})_{i,j=1,\cdots,n+1},$$

where

$$\alpha_{i,j} = \begin{cases} q_{i-j}((\nu_{i-1} - \mu_{j-1})_-) & \textit{for all } j = 1, \cdots, n \\ f_{i-1}(\nu_{i-1}) & \textit{if } j = n+1, \end{cases}$$

for any $\nu \leq \mu$.

If, in addition, (f_n) *has roots in* ν, *then*

$$f_n(\mu_n) = (-1)^n \det(q_{i+1-j}((\nu_i - \mu_{j-1})_-))_{i,j=1,\cdots,n}. \tag{A.17}$$

In some applications (see Example 4) μ is of the special form $\mu(i) = ic + d$ with $c, d \in K$. If there exists a $\Delta > 0$ such that $f_i(\mu_i - \Delta)$ is known for all $i \in \mathbb{N}_0$, (A.16) becomes a generating function identity; as follows. Define, for $z \geq 0$,

$$\varphi_z(t) = \textstyle\sum_{i\geq 0} f_i(\mu_i - z)t^i \text{ and } \gamma_z(t) = \sum_{i=0}^{\lfloor z/c \rfloor} q_i(ic - z)t^i. \tag{A.18}$$

Now we get from (A.16) with $x = \mu_n - z$ that

$$\varphi_z(t) = \varphi_0(t)\gamma_z(t).$$

Choose $z = \Delta$, then $\varphi_0(t) = \gamma_\Delta(t)^{-1}\varphi_\Delta(t)$, hence,

$$\varphi_z(t) = \frac{\gamma_z(t)\varphi_\Delta(t)}{\gamma_\Delta(t)}. \tag{A.19}$$

REFERENCES

AABERGE, R., DOKSUM, K. A., and FENSTAD, G. (1977). Plots and tests for symmetry. *Biometrika* **64** 473–487.

ALTER, R., GOVINDARAJULU, Z., and GRAGG, L. E. (1975). The exact distribution of the one-sample Kolmogorov statistic. *Trabajos Estadíst. Investigación Oper.* **26** 407–431.

BIRNBAUM, Z. W., and LIENTZ, B. P. (1969). Exact distributions for some Rényi-type statistics. *Zastos Mat.* **10** 179–192.

BOROKOV, A. A., and SYCHEVA, N. M. (1968). On asymptotically optimal non-parametric criteria. *Theory Probability Appl.* **13** 359–393.

BUTLER, C. C. (1969). A test for symmetry using the sample distribution function. *Ann. Math. Statist.* **40** 2209–2210.

CANNER, P. L. (1975). A simulation study of one- and two-sample Kolmogorov-Smirnov statistics with a particular weight function. *J. Amer. Statist. Assoc.* **70** 209–211.

CHANG, LI-CHIEN (1955). On the ratio of the empirical distribution function to the theoretical distribution function. *Acta Math. Sinica* **5** 347–368 (*IMS & AMS Sel. Transl. Math. Statist. Prob.* **4** (1963)17–38).

CHATTERJEE, S. K., and SEN, P. K. (1973). On Kolmogorov-Smirnov-type tests for symmetry. *Ann. Inst. Statist. Math.* **25** 287–300.

DANIELS, H. E. (1945). The statistical theory of the strength of bundles of threads, I. *Proc. Roy. Soc. London Ser. A* **183** 405–435.

DOKSUM, K. A., and SIEVERS, G. L. (1976). Plotting with confidence: Graphical comparisons of two populations. *Biometrika* **63** 421–434.

DURBIN, J. (1973). Distribution theory for tests based on the sample distribution function. *Regional conference series in appl. math.* SIAM, **9**.

EICKER, F. (1970). On the probability that a sample distribution function lies below a line segment. *Ann. Math. Statist.* **41** 2075–2092.

EPANECHNIKOV, V. A. (1968). The significance level and power of the two-sided Kolmogorov test in the case of small sample sizes. *Theory Probability Appl.* **13** 686–690.

FRAY, R. D., and ROSELLE, D. P. (1971). Weighted lattice paths. *Pacific J. Math.* **37** 85–96.

GNEDENKO, B. V., and KOROLJUK, V. S. (1951). On the maximum discrepancy between two empirical distributions. *Dokl. Adad. Nauk SSSR* **80** 525–528 (*IMS & AMS Sel. Transl. Math. Statist. Prob.* **1** (1961) 13–16).

HODGES, J. L. (1957). The significance probability of the Smirnov two-sample test. *Ark. Mat.* **3** 469–486.

ISHII, G. (1959). On the exact probabilities of Rényi's tests. *Ann. Inst. Statist. Math.* **11** 17–24.

KEMPERMAN, J. H. B. (1961). The passage problem for a stationary Markov chain. *Statist. research monographs* **1**. Univ. Chicago Press.

KOROLJUK, V. S. (1955). On the discrepancy of empiric distributions for the case of two independent samples. *Izv. Akad. Nauk SSSR Ser. Mat.* **19** 81–96 (*IMS&AMS Sel. Transl. Math. Statist. Prob.* **4** (1963) 105–122).

KREWERAS, G. (1965). Sur une classe de problèmes des dénombrement lies au treillis dés partitions des entiers. *Cahiers Bureau Univ. Recherche Opér.* Inst. Statist. Univ. Paris.

KUIPER, N. H. (1960). Tests concerning random points on a circle. *Nederl. Akad. Wetensch. Proc. Ser. A.* **63** 38–47.

LAUWERIER, H. A. (1963). The asymptotic expansion of the statistical distribution of N. V. Smirnov. *Z. Wahrsch. Verw. Gebiete* **2** 61–68.

MASSEY, F. J. (1950). A note on the estimation of a distribution function by confidence limits. *Ann. Math. Statist.* **21** 116–119.

MASSEY, F. J. (1951). The distribution of the maximum deviation between two sample cumulative step functions. *Ann. Math. Statist.* **22** 125–128.

MOHANTY, S. G. (1971). A short proof of Steck's result on two-sample Smirnov statistics. *Ann. Math. Statist.* **42,** 413–414.

NIEDERHAUSEN, H. (1978). A general Ballot-problem and its applications for the Kolmogorov-Smirnov two sample test (in German). *Thesis; Techn. Univ. of Graz, Austria.*

NIEDERHAUSEN, H. (1980). Linear recurrences under side conditions. *European J. Combin.* **1** 353–368.

NÓE, M. (1972). The calculation of distributions of two-sided Kolmogorov-Smirnov type statistics. *Ann. Math. Statist.* **43** 58–64.

PITMAN, E. J. G. (1972). Simple proofs of Steck's determinantal expressions for probabilities in the Kolmogorov and Smirnov tests. *Bull. Austral. Math. Soc.* **7** 227–232.

RÉNYI, A. (1968). On a group of problems in the theory of ordered samples. *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.* **18** 23–30. (*IMS & AMS Sel. Transl. Math. Statist. Prob.* **13** (1973) 289–298.)

ROBBINS, H. (1954). A one-sided confidence interval for an unknown distribution function. *Ann. Math. Statist.* **25** 409.

ROTA, G.-C., KAHANER, D., and ODLYZKO, A. (1973). On the foundations of a combinatorial theory: VII. Finite operator calculus. *J. Math. Anal. Appl.* **42** 684–760.

SMIRNOV, N. V. (1944). Approximate laws of distribution or random variables from empirical data. (in Russian) *Uspehi Mat. Nauk* **10** 179–206.

SMIRNOV, N. V. (1947). Sur un critère de symétrie de la loi de distribution d' une variable aléatorie. *Akad. Nauk SSSR, C.R. (Doklady)* **56** 11–14.

STECK, G. P. (1969). The Smirnov two sample tests as rank tests. *Ann. Math. Statist.* **40** 1449–1466.

STECK, G. P. (1971). Rectangle probabilities for uniform order statistics and the probability that the empirical distribution function lies between two distribution functions. *Ann. Math. Statist.* **42** 1–11.

STECK, G. P. (1974). A new formula for $P(R_i \leq b_i, 1 \leq i \leq m \mid m, n, F = G^k)$. *Ann. Probab.* **2** 155–160.

STEPHENS, M. A. (1965). The goodness-of-fit statistics V_N: distribution and significance points. *Biometrika* **52** 309–321.

TAKÁCS, L. (1967). *Combinatorial methods in the theory of stochastic processes.* Wiley, New York.

DEPARTMENT OF STATISTICS
UNIVERSITY OF TORONTO
TORONTO, CANADA M55 1A1