ORDER STATISTICS AND STATISTICS OF STRUCTURE (*d*)

By Herbert T. David

Iowa State University

0. Summary and introduction. This note discusses the asymptotic independence of the "(d)-structured" order statistics $F(X^{(i)})$ and some of the more usual statistics of structure (d). (For a discussion of structure (d) and related concepts, see references [2], [3] and [4].) Asymptotic independence holds, in particular, in the case of the Kolmogoroff-Smirnoff statistic, and thus provides approximate significance levels for the simultaneous test of the hypothesis that the population c.d.f. has specified form and the hypothesis that the sample contains no outlying observations.

If the Kolmogoroff-Smirnoff statistic is used in conjunction with the largest order statistic $X^{(n)}$, the acceptance region of the resulting test can be characterized as follows: For acceptance (of the hypothesis that the population c.d.f. has specified shape and that the sample contains no outliers), the n "risers" of the sample c.d.f. must fall within a "three-sided" region. This region is the intersection of the usual Kolmogoroff-Smirnoff region with the half-plane to the left of the critical value for $X^{(n)}$. The example of the last section deals with this case.

Section 1 contains the theory on which the subsequent development is based. It involves the multivariate probability integral transformation T_G discussed in [10], which has the property that $T_G(X)$, X distributed according to G, is uniform over the unit cube, so that $T_H^{-1}T_G(X)$ has distribution H. The transformation enters the argument in essentially this way: If Y is a vector distributed according to H, f and g are two functions of Y, and $H(t)$ is the conditional distribution of Y, given $g(Y) \leqq t$, then $f(Y)$ and $g(Y)$ are independent if and only, for essentially all t, $f(Y)$ and $f(T_H^{-1}T_{H(t)}(Y))$ have the same distribution conditionally on $g(Y) \leqq t$. In the present application, Y is a random sample from the uniform distribution, g is an order statistic of this sample, and f is the function of Y corresponding to a statistic of structure (d) such as the Kolmogoroff-Smirnoff statistic or Sherman's statistic [12].

It may appear that the above method recommends itself primarily on grounds of novelty, rather than suitability. This is borne out by the fact that the relatively weak requirement that the distributions of $f(Y)$ and $f(T_H^{-1}T_{H(t)}, Y))$ agree asymptotically is verified below by showing that in fact $f(Y)$ and $f(T_H^{-1}T_{H(t)}(Y))$ themselves agree asymptotically, i.e., that their difference converges in probability to zero. This last is made to follow in turn from the very rapid covergence of $|Y - T_H^{-1}T_{H(t)}(Y)|$ to zero.

One has the option of having H or g perform the ordering of the sample Y. By this is meant that H can be the uniform distribution over the unit cube, and g

Received 5 October 1962; revised 20 April 1964.

the function "kth largest coordinate of Y". Or H can be the joint distribution of the order statistics of Y and g the function "kth coordinate of Y".

The second of these two approaches yields a form for $T_H^{-1}T_{H(t)}$ especially suitable (due to the common multiplier $C_n(x, t, l)$ appearing in Equations (1)) for establishing the convergence to zero of $f(Y) - f(T_H^{-1}T_{H(t)}(Y))$ for a fairly large class of statistics f. These include the statistics of Kolmogoroff-Smirnoff type, whose essential combining operation is "sup", and also certain other statistics with less tractable combining operations, such as Sherman's statistic, where the combining operation is addition. Section 2 is devoted to this second approach, applied to Sherman's statistic.

The first approach, though seemingly unnatural for statistics with combining operations more demanding than "sup", has the advantage of yielding expressions as easy to manipulate for bivariate (e.g., two order statistics) as for univariate g. Section 3 is devoted to this first approach, applied to statistics of Kolmogoroff-Smirnoff type, with g bivariate.

The last section contains an application, together with a computation indicating that independence sets in quite rapidly.

1. Theory.

DEFINITION 1. $Y_1, Y_2, \cdots, Y_n, \cdots \equiv (X_{11}), (X_{21}, X_{22}), \cdots, (X_{n1}, X_{n2}, \cdots, X_{nn}), \cdots, X_{ij}$ real, is a triangular sequence of random variables, with Y_n distributed according to the n-dimensional probability measure $H(n)$ over the Borel sets of Euclidean n-space E_n.

DEFINITION 2. f_n and g_n are two scalar-valued measurable functions of Y_n.

ASSUMPTION 1. The c.d.f.'s $\lambda_g^{(n)}$ of $g_n(Y_n)$ converge weakly to a c.d.f. λ_g, and similarly for $\lambda_f^{(n)}$, $f_n(Y_n)$ and λ_f.

LEMMA 1. *Let t be a continuity point of λ_g such that $\lambda_g(t) > 0$. For large n, $H(n)$ induces a uniquely specified conditional probability measure $H(n, t)$ of Y_n, given the condition $C(n, t): g_n(Y_n) \leqq t$; the* c.d.f. *corresponding to $H(n, t)$ is*

$$\Pr\{Y_n \leqq y; g_n(Y_n) \leqq t\}/\lambda_g^{(n)}(t).$$

PROOF. By Assumption 1, $\lambda_g^{(n)}(t) \to \lambda_g(t) > 0$, so that $\lambda_g^{(n)}(t) > 0$ for large n.

ASSUMPTION 2. For every n, there exists a measurable transformation $T_{H(n,t)}$ of E_n into the unit cube Q_n, such that the measure induced over the Borel sets of Q_n by the probability measure $H(n, t)$ over the Borel sets of E_n is Lebesgue measure $L(n)$.

ASSUMPTION 3. For every n, there exists a measurable transformation $T_{H(n)}^{-1}$ of Q_n into E_n, such that the measure induced over the Borel sets of E_n by the probability measure $L(n)$ over Q_n is the probability measure $H(n)$.

DEFINITION 3. Two sequences of random variables $\{f_n\}$ and $\{g_n\}$ are said to be asymptotically independent if there exist two c.d.f.'s, λ_f and λ_g, such that the c.d.f. sequences for $\{f_n\}$, $\{g_n\}$ and $\{f_n, g_n\}$ converge weakly, respectively, to λ_f, λ_g and $\lambda_f \cdot \lambda_g$.

THEOREM 1. *Let t be a continuity point of λ_g such that $\lambda_g(t) > 0$, and let $Z_{n,t}$ be distributed according to $H(n, t)$. For the asymptotic independence of $f_n(Y_n)$ and*

$g_n(Y_n)$ *it is necessary and sufficient that, for all* t, *the* c.d.f.'s *of* $f_n(Z_{n,t})$ *and the* c.d.f.'s *of* $f_n(T^{-1}_{H(n)}T_{H(n,t)}(Z_{n,t}))$ *converge weakly to the same* c.d.f.

Proof of sufficiency. (a) Consider any continuity point t' of λ_g at which λ_g is zero. Then, for such a t' and any continuity point u of λ_f, we have, using Assumption 1, that $0 \leqq \Pr\{f_n(Y_n) \leqq u, g_n(Y_n) \leqq t'\} \leqq \lambda_g^{(n)}(t') \to \lambda_g(t') = 0$, so that $\Pr\{f_n(Y_n) \leqq u; g_n(Y_n) \leqq t'\} \to 0 = \lambda_f(u)\cdot\lambda_g(t')$, without appeal to the condition of the theorem.

(b) Consider any continuity point t of λ_g at which λ_g exceeds zero, and any continuity point u of λ_f. Using in order Lemma 1, the condition of the theorem, Assumptions 2 and 3, and Assumption 1, we then have

$$\begin{aligned}\lim_{n\to\infty} &\Pr\{f_n(Y_n) \leqq u; g_n(Y_n) \leqq t\}/\lambda_g^{(n)}(t) \\ &= \lim_{n\to\infty} \Pr\{f_n(Y_n) \leqq u \mid C(n, t)\} = \lim_{n\to\infty} \Pr\{f_n(Z_{n,t}) \leqq u\} \\ &= \lim_{n\to\infty} \Pr\{f_n(T^{-1}_{H(n)}T_{H(n,t)}(Z_{n,t})) \leqq u\} \\ &= \lim_{n\to\infty} \Pr\{f_n(Y_n) \leqq u\} = \lambda_f(u).\end{aligned}$$

And this yields

$$\begin{aligned}\lim_{n\to\infty} \Pr\{f_n(Y_n) \leqq u; g_n(Y_n) \leqq t\} &= \lambda_f(u)\cdot\lim_{n\to\infty}\lambda_g^{(n)}(t) \\ &= \lambda_f(u)\cdot\lambda_g(t).\end{aligned}$$

Proof of necessity. Part (b) of the sufficiency argument is easily inverted: for any continuity point t of λ_g with $\lambda_g(t) > 0$ and any continuity point u of λ_f, asymptotic independence implies that $\lim_{n\to\infty} \Pr\{f_n(Y_n) \leqq u \mid C(n, t)\} = \lambda_f(u)$, i.e., that $\lim_{n\to\infty} \Pr\{f_n(Z_{n,t}) \leqq u\} = \lambda_f(u)$. On the other hand, $\Pr\{f_n(T^{-1}_{H(n)}T_{H(n,t)}(Z_{n,t})) \leqq u\} \to \lambda_f(u)$ by Assumptions 1, 2 and 3.

Corollary 1. *Let* t *be a continuity point of* λ_g *such that* $\lambda_g(t) > 0$, *and let* $Z_{n,t}$ *be distributed according to* $H(n, t)$. *It is sufficient for the asymptotic independence of* $f_n(Y_n)$ *and* $g_n(Y_n)$ *that, for all* t, $\operatorname{plim} |f_n(Z_{n,t}) - f_n(T^{-1}_{H(n)}T_{H(n,t)}(Z_{n,t}))| = 0$.

Proof. By Assumptions 1, 2 and 3 the c.d.f.'s of $f_n(T^{-1}_{H(n)}T_{H(n,t)}(Z_{n,t})) \equiv A_n$ converge weakly to λ_f for all t. But, by the assumption of the corollary, $f_n(Z_{n,t}) = A_n + \epsilon_n$, where $\operatorname{plim} \epsilon_n = 0$, so that the c.d.f.'s of $f_n(Z_{n,t})$ also converge weakly to λ_f, and Theorem 1 applies.

Note that the arguments in this section remain essentially unchanged if f_n or g_n are vectors rather than scalars. In Section 3, the random variables g_n are two-dimensional.

2. The asymptotic independence of Sherman's statistic and extreme (d)-structured order statistics. Consider the following specialization of the quantities introduced in Section 1. Y_n : the set of order statistics $(U_1^{(n)}, \cdots U_n^{(n)})$ for a random sample from the uniform distribution on [0, 1]; $H(n)$: the uniform distribution over region $0 \leqq x_1 \leqq \cdots \leqq x_n \leqq 1$; f_n : the function Φ_n of the uniform order statistics corresponding to Sherman's statistic, suitably normed to insure convergence to a normal distribution $N(0, c)$; λ_f : the c.d.f. for $N(0, c)$;

$g_n(x_1, \cdots, x_n) = n(x_{n-l} - 1)$; λ_g : the c.d.f. for a gamma distribution on $[0, -\infty)$ (see [6], p. 371); $C(n, -t)$: $U^{(n)}_{n-l} \leqq 1 - (t/n)$; $H(n, t)$: the uniform distribution over the region $0 \leqq x_1 \leqq \cdots \leqq x_n \leqq 1$; $x_{n-l} \leqq 1 - (t/n)$.

Given an n-variate distribution G, define T_G as in [10]. A direct computation then shows that $T_{H(n,t)}$ satisfies Assumption 2, that $T_{H(n)}$ possesses an inverse $T^{-1}_{H(n)}$, and that $T^{-1}_{H(n)}$ satisfies Assumption 3. It then remains, only, to calculate $T^{-1}_{H(n)}T_{H(n,t)}$ and to verify the condition of Corollary 1.

To begin with, the transformation $(y_1, \cdots, y_n) = T_{H(n)}(x_1, \cdots, x_n)$ is given by: $y_{n-m} = \Pr\{U^{(n)}_{n-m} \leqq x_{n-m} \mid U^{(n)}_{n-m+1} = x_{n-m+1}, \cdots, U_n^{(n)} = x_n\}$:

$$y_n = (x_n)^n$$

$$y_{n-1} = (x_{n-1}/x_n)^{n-1}$$

$$\vdots$$

$$y_1 = (x_1/x_2)$$

so that $(x_1, \cdots, x_n) = T^{-1}_{H(n)}(y_1, \cdots, y_n)$ is given by:

$$x_n = (y_n)^{1/n}$$

$$x_{n-1} = (y_n)^{1/n}(y_{n-1})^{1/n-1}$$

$$\vdots$$

$$x_1 = (y_n)^{1/n}(y_{n-1})^{1/n-1} \cdots (y_2)^{\frac{1}{2}}(y_1).$$

As for $T_H(n,t)$, the subset of Q_n assigned probability one by $H(n, t)$ is the union $\bigcup_{b=0}^{l} S_b$ of the sets S_b : $\{(x_1, \cdots, x_n): 0 \leqq x_1 \leqq \cdots \leqq x_{n-b} \leqq 1 - t/n \leqq x_{n-b+1} \leqq \cdots \leqq x_n \leqq 1\}$, and, for $x \varepsilon S_b$, $y = T_{H(n,t)}(x)$ is given by: $y_{n-m} = \Pr\{U^{(n)}_{n-m} \leqq x_{n-m} \mid U^{(n)}_{n-m+1} = x_{n-m+1}, \cdots, U_n^{(n)} = x_n; C(n, -t)\}$:

$$\begin{aligned}
y_n &= [\textstyle\sum_{j=0}^{l} \binom{n}{j}(x_n - (1 - t/n))^j(1 - t/n)^{n-j}] \\
&\quad \cdot [\textstyle\sum_{j=0}^{l} \binom{n}{j}(t/n)^j(1 - t/n)^{n-j}]^{-1} \equiv A_n(x, t, l) \\
y_{n-1} &= [\textstyle\sum_{j=0}^{l-1} \binom{n-1}{j}(x_{n-1} - (1 - t/n))^j(1 - t/n)^{n-1-j}] \\
&\quad \cdot [\textstyle\sum_{j=0}^{l-1} \binom{n-1}{j}(x_n - (1 - t/n))^j(1 - t/n)^{n-1-j}]^{-1} \equiv A_{n-1}(x, t, l) \\
&\vdots \\
y_{n-b+1} &= [\textstyle\sum_{j=0}^{l-b+1} \binom{n-b+1}{j}(x_{n-b+1} - (1 - t/n))^j(1 - t/n)^{n-b+1-j}] \\
&\quad \cdot [\textstyle\sum_{j=0}^{l-b+1})\binom{n-b+1}{j}(x_{n-b+2} - (1 - t/n))^j(1 - t/n)^{n-b+1-j}]^{-1} \equiv A_{n-b+1}(x, t, l) \\
y_{n-b} &= (x_{n-b})^{n-b}[\textstyle\sum_{j=0}^{l-b} \binom{n-b}{j}(x_{n-b+1} - (1 - t/n))^j(1 - t/n)^{n-b-j}]^{-1} \\
&\equiv (x_{n-b})^{n-b}/B_{n-b}(x, t, l) \\
y_{n-b-1} &= (x_{n-b-1}/x_{n-b})^{n-b-1} \\
&\vdots
\end{aligned}$$

$$y_2 = (x_2/x_3)^2$$
$$y_1 = x_1/x_2 .$$

Hence, compositing $T_{H(n)}^{-1}$ and $T_{H(n,t)}$, one obtains, for $(y_1, \cdots, y_n) \equiv T_{H(n)}^{-1} T_{H(n,t)}(x_1, \cdots, x_n)$ and x in S_b,

$$\begin{aligned}
y_n &= (A_n(x, t, l))^{1/n} \\
y_{n-1} &= (A_n(x, t, l))^{1/n}(A_{n-1}(x, t, l))^{1/n-1} \\
&\vdots \\
y_{n-b+1} &= (A_n(x, t, l))^{1/n} \cdots (A_{n-b+1}(x, t, l))^{1/n-b+1} \\
y_{n-b} &= (A_n(x, t, l))^{1/n} \cdots (A_{n-b+1}(x, t, l))^{1/n-b+1}[x_{n-b}/(B_{n-b}(x, t, l)^{1/n-b})] \\
&\equiv (x_{n-b})(C_n(x, t, l)) \\
y_{n-b-1} &= (x_{n-b-1})(C_n(x, t, l)) \\
&\vdots \\
y_1 &= (x_1)(C_n(x, t, l)).
\end{aligned} \tag{1}$$

The next step is the evaluation of $|\Phi_n(T_{H(n)}^{-1} T_{H(n,t)}(x)) - \Phi_n(x)|$. To this end, recall that, for Sherman's statistic, Φ_n is given by

$$\Phi_n(x_1, \cdots, x_n) = n^{\frac{1}{2}}(\textstyle\sum_{i=0}^{n} |x_{i+1} - x_i - (n+1)^{-1}| - 2e^{-1}),$$

so that, defining y as in (1),

$$\begin{aligned}
|\Phi_n(y_1, &\cdots, y_n) - \Phi_n(x_1, \cdots, x_n)| \\
&= |n^{\frac{1}{2}}(|y_1 - (n+1)^{-1}| + |y_2 - y_1 - (n+1)^{-1}| + \cdots \\
&\quad + |y_n - y_{n-1} - (n+1)^{-1}| + |1 - y_n - (n+1)^{-1}|) \\
&\quad - n^{\frac{1}{2}}(|x_1 - (n+1)^{-1}| + |x_2 - x_1 - (n+1)^{-1}| + \cdots \\
&\quad + |x_n - x_{n-1} - (n+1)^{-1}| + |1 - x_n - (n+1)^{-1}|)| \\
&\leqq n^{\frac{1}{2}}(|y_1 - x_1| + |(y_2 - y_1) - (x_2 - x_1)| + \cdots) \\
&= n^{\frac{1}{2}}(|y_1 - x_1| + |(y_2 - x_2) - (y_1 - x_1)| + \cdots \\
&\quad + |(y_n - x_n) - (y_{n-1} - x_{n-1})| + |y_n - x_n|) \\
&\equiv n^{\frac{1}{2}}(d_0 + d_1 + \cdots + d_{n-1} + d_n).
\end{aligned}$$

It is now convenient, for x in S_b, to partition this last summation at $n - b$, so that we write

$$\begin{aligned}
|\Phi_n(y) - \Phi_n(x)| &\leqq n^{\frac{1}{2}}(\textstyle\sum_{i=0}^{n-b-1} d_i + \sum_{i=n-b}^{n} d_i) \\
&\equiv n^{\frac{1}{2}}(K_{1,n} + K_{2,n}).
\end{aligned} \tag{2}$$

Consulting (1), we find that

$$K_{1,n} \leqq |C_n(x, t, l) - 1| \, |x_1 + \textstyle\sum_{i=2}^{n-b} (x_i - x_{i-1})| \leqq |C_n(x, t, l) - 1|. \tag{3}$$

As for $K_{2,n}$, bounding $|(y_i - x_i) - (y_{i-1} - x_{i-1})|$ by $|y_i - x_i| + |y_{i-1} - x_{i-1}|$ and recalling the expression for y_{n-b} in (1), one obtains

$$K_{2,n} \leqq x_{n-b} \, |C_n(x, t, l) - 1| + 2 \textstyle\sum_{m=0}^{b-1} |y_{n-m} - x_{n-m}|, \tag{4}$$

and (2), (3) and (4) imply, for x in S_b, that

$$|\Phi_n(y) - \Phi_n(x)| \leqq 2n^{\frac{1}{2}}(|C_n(x, t, l) - 1| + \textstyle\sum_{m=0}^{b-1} |y_{n-m} - x_{n-m}|). \tag{5}$$

Now, for x in S_b, the factors $A_{n-m}(x, t, l)$ in (1) satisfy

$$(1 - t/n)^{n-m} \leqq A_{n-m} \leqq 1, \tag{6}$$

so that, in view of (1), for x in S_b and $0 \leqq m \leqq b - 1$, $(1 - t/n)^{m+1} \leqq y_{n-m} \leqq 1$. On the other hand, for such x and m, one also has that $1 - t/n \leqq x_{n-m} \leqq 1$, so that, for x in S_b,

$$\textstyle\sum_{m=0}^{b-1} |y_{n-m} - x_{n-m}| \leqq \textstyle\sum_{m=0}^{b-1} (1 - (1 - t/n)^{m+1}). \tag{7}$$

Finally, using (6) for each of the b factors A_{n-m} of $C_n(x, t, l)$ and noticing that $1 \geqq B_{n-b} \geqq (1 - t/n)^{n-b}$, one finds that

$$(1 - t/n)^b - 1 \leqq C_n(x, t, l) - 1 \leqq (1 - t/n)^{-1} - 1, \tag{8}$$

and (5), (7) and (8) imply that, for x in the region $\bigcup_{b=0}^{l} S_b$ assigned probability 1 by $H(n, t)$,

$$|\Phi_n(T_{H(n)}^{-1} T_{H(n,t)}(x)) - \Phi_n(x)| \equiv |\Phi_n(y) - \Phi_n(x)| \leqq K(t)/n^{\frac{1}{2}}.$$

This implies the condition of Corollary 1, and hence the desired asymptotic independence.

It may be of interest to note, in view of (7) and (8), that the above argument, and hence asymptotic independence, remains valid when l is allowed to grow slowly with n.

3. The asymptotic independence of the Kolmogoroff-Smirnoff statistic and two extreme (d)-structured order statistics. (The argument in this section applies as well, essentially without change, to the statistics proposed in [1], [7], and [13].) The quantities of Section 1 now are specialized as follows: Y_n : a random sample $(U_1, \cdots, U_n)$ from the uniform distribution on [0, 1]; $H(n)$: the uniform distribution over the unit n-cube Q_n ; f_n : the function Φ_n, symmetric over Q_n, corresponding to the Kolmogoroff-Smirnoff statistic, suitably normed to insure convergence to the Kolmogoroff-Smirnoff limit distribution; λ_f : the c.d.f. for this distribution; $g_n(x) = (g_{n,1}(x), g_{n,2}(x))$, where $g_{n,1}(x) = -(n)((k + 1)$st smallest coordinate of $x)$, and $g_{n,2}(x) = (n)((l + 1)$st largest coordinate of $x - 1)$, with $n - l > k + 1$; λ_g : the bivariate c.d.f. for two mutually independent random variables, one with a gamma distribution on $[0, +\infty)$, the other with a gamma distribution of $[0, -\infty)$; the condition $C(n, -s, -t): g_{n,1}(U_1, \cdots, U_n) \leqq -s$, $g_{n,2}(U_1, \cdots, U_n) \leqq -t$, with $1 - t/n > s/n$; $H(n, t)$: the uniform distribution over the sub-region of Q_n for which the $(k + 1)$st smallest coordinate is no smaller than s/n, and the $(l + 1)$st largest coordinate is no greater than $1 - t/n$; this last distribution will be denoted by $J(n, k, l, s/n, t/n)$.

In this application, a transformation $T_{H(n)}^{-1}$ satisfying Assumption 3 is the identity, and $T_{H(n,t)}$, computed below in accordance wth [10], is found to satisfy Assumption 2. Hence, as in Section 2, all that remains after computing $T_{H(n,t)}$ is verifying the condition of Corollary 1. $T_{H(n,t)}$ is computed as follows:

DEFINE.

$A(i, j, n - (i + j), \delta, \epsilon)$: the sub-region of Q_n for which i coordinates lie between 0 and δ, j coordinates lie between $1 - \epsilon$ and 1, and $n - (i + j)$ coordinates lie between δ and $1 - \epsilon$.

$$B(n, k, l, \delta, \epsilon): \bigcup_{R(i,j)} A(i, j, n - (i + j), \delta, \epsilon),$$

where $R(i, j) = [(i, j): i = 0, \cdots, k; j = 0, \cdots, l; i + j \leqq n]$.

For n, k and l positive integers,

$$I(n, k, l, \delta, \epsilon) = \sum_{R(i,j)} n!/[i!j!(n - i - j)!]\delta^i\epsilon^j(1 - \epsilon - \delta)^{n-i-j},$$

where $R(i, j)$ is again as defined above.

For k or l a negative integer, $I(n, k, l, \delta, \epsilon) \equiv 0$; $I(0, 0, 0, \delta, \epsilon) \equiv 1$.

$$I_1 \equiv I(n - m, \min (n - m, k - a), \min (n - m, l - b), s/n, t/n)$$

$$I_2 \equiv I(n - m - 1, \min (n - m - 1, k - a - 1), \min (n - m - 1, l - b), s/n, t/n)$$

$$I_3 \equiv I(n - m - 1, \min (n - m - 1, k - a), \min (n - m - 1, l - b), s/n, t/n)$$

$$I_4 \equiv I(n - m - 1, \min (n - m - 1, k - a), \min (n - m - 1, l - b - 1), s/n, t/n)$$

We note that

$$(s/n)(I_2) + (1 - s/n - t/n)(I_3) + (t/n)(I_4) = I_1 . \tag{9}$$

Now suppose that $(X_1, \cdots, X_n)$ is distributed according to $J(n, k, l, s/n, t/n)$; then the conditional distribution of $(X_{m+1}, \cdots, X_n)$, given that $(X_1, \cdots, X_m) = (x_1, \cdots, x_m)$, is the distribution $J(n - m, \min (n - m, k - a), \min (n - m, l - b), s/n, t/n)$ that has density $1/I_1$ over the sub-region $B(n - m, \min (n - m, k - a), \min (n - m, l - b), s/n, t/n)$ of Q_{n-m}, where a and b are respectively, the number of x_i's less than s/n and greater than $1 - t/n$. In accordance with [10], the component y_{m+1} of $y = T_{H(n,t)}(x)$ now is simply the ensuing marginal c.d.f. of X_{m+1} :

For $0 \leqq x_{m+1} \leqq s/n$, $y_{m+1} = (x_{m+1})(I_2/I_1)$.

For $s/n \leqq x_{m+1} \leqq 1 - t/n$, $y_{m+1} = (s/n)(I_2/I_1) + (x_{m+1} - s/n)(I_3/I_1)$.

For $1 - t/n \leqq x_{m+1} \leqq 1$, $y_{m+1} = (s/n)(I_2/I_1) + (1 - s/n - t/n)(I_3/I_1) + (x_{m+1} - (1 - t/n))(I_4/I_1)$.

Using (9), it now follows that

For $0 \leqq x_{m+1} \leqq s/n$, $|y_{m+1} - x_{m+1}| = |(x_{m+1})(I_2 - I_1)/I_1| \leqq 2s/nI_1$.

For $s/n \leqq x_{m+1} \leqq 1 - t/n$, $|y_{m+1} - x_{m+1}| = |(s/n)(I_2 - I_3)/I_1 + (x_{m+1}) \cdot (I_3 - I_1)/I_1| = |(s/n)(I_2 - I_3)/I_1 + (x_{m+1})((s/n)(I_3 - I_2) + (t/n)(I_3 - I_4))/I_1| = |(s/n)(I_3 - I_2)(x_{m+1} - 1) + (t/n)(I_3 - I_4)|/I_1 \leqq 2s/nI_1 + 2t/nI_1$.

For $1 - t/n \leqq x_{m+1} \leqq 1$, $|y_{m+1} - x_{m+1}| = |(s/n)(I_2/I_1) + (1 - s/n - t/n) \cdot (I_3/I_1) + (t/n)(I_4/I_1) + (x_{m+1} - 1)(I_4/I_1) - x_{m+1}| = |1 + (x_{m+1} - 1) \cdot (I_4/I_1) - x_{m+1}| = |(x_{m+1} - 1)(I_4 - I_1)/I_1| \leqq 2t/nI_1$.

It remains to notice that, for n large (specifically, $n \geqq 2(s + t)$), I_1 is bounded away from zero uniformly in m: For $n \geq 2(s + t)$ and $n - m \leqq 2(s + t)$, one has that $I_1 \geqq [1 - ((s + t)/n)]^{n-m} \geqq (\frac{1}{2})^{n-m} \geqq (\frac{1}{2})^{2(s+t)}$, while, for $n - m \geqq 2(s + t)$, $I_1 \geqq [1 - ((s + t)/n)]^{n-m} \geqq \{1 - [(s + t)/(n - m)]\}^{n-m} \geqq \{1 - [(s + t)/2(s + t)]\}^{2(s+t)} = (\frac{1}{2})^{2(s+t)}$. Hence, for n large and x in the subregion $B(n, k, l, 2/n, t/n)$ of Q_n assigned probability 1 by $J(n, k, l, s/n, t/n)$ (indeed, for x in Q_n), the bounds computed for $|y_{m+1} - x_{m+1}|$ in fact imply that

$$(10) \quad \max_{0 \leqq m \leqq n-1} |y_{m+1} - x_{m+1}| \leqq 2(s + t)2^{2(s+t)}/n \equiv K(s, t)/n.$$

The bound (10) now is used to verify the condition of Corollary 1. The function Φ_n is given by

$$(11) \quad \Phi_n(x) = n^{\frac{1}{2}}\{\max_m [\max (x_m^{(n)} - (m - 1)/n, (m/n) - x_m^{(n)})]\},$$

where $x_m^{(n)}$ is the mth smallest coordinate of x, and the bound (10) implies that, for x in $B(n, k, l, s/n, t/n)$, $\max_{1 \leqq m \leqq n} |y_m^{(n)} - x_m^{(n)}| \leqq K(s, t)/n$, hence that

$$(12) \quad \begin{aligned} |\Phi_n(T_{H(n)}^{-1}T_{H(n,t)}(x)) - \Phi_n(x)| &= |\Phi_n(T_{H(n,t)}(x)) - \Phi_n(x)| \\ &\equiv |\Phi_n(y) - \Phi_n(x)| \leqq K(s, t)/n^{\frac{1}{2}}, \end{aligned}$$

and this implies the condition of Corollary 1, and hence asymptotic independence.

4. Bounds for departure from independence, and an example. Consider the quantities defined in Section 1, abbreviating $T_{H(n)}^{-1}T_{H(n,t)}$ to $T_{n,t}$. Suppose that, as was true in Sections 2 and 3, the condition of Corollary 1 is satisfied in the strong sense that, for x in a region assigned probability 1 by $H(n, t)$, $|f_n(T_{n,t}(x)) - f_n(x)| \leqq K(t)\psi(n)$, where $\psi(n)$ is of order less than 0. Then

$$\begin{aligned} \Pr \{f_n(Y_n) \leqq u \mid g_n(Y_n) \leqq t\} \\ &= \Pr \{f_n(Z_{n,t}) \leqq u\} = \Pr \{f_n(Z_{n,t}) \leqq u - K \cdot \psi\} \\ &\quad + \Pr \{u - K \cdot \psi < f_n(Z_{n,t}) \leqq u\} \leqq \Pr \{f_n(T_{n,t}(Z_{n,t})) \leqq u\} \\ &\quad + \Pr \{u - 2K \cdot \psi < f_n(T_{n,t}(Z_{n,t})) \leqq u + K \cdot \psi\} \end{aligned}$$

$$= \lambda_f^{(n)}(u) + [\lambda_f^{(n)}(u + K\cdot\psi) - \lambda_f^{(n)}(u - 2K\cdot\psi)]$$
$$\equiv \lambda_f^{(n)}(u) + A(n, u, K\cdot\psi).$$

Again,

$$\Pr\{f_n(Y_n) \leqq u \mid g_n(Y_n) \leqq t\}$$
$$= \Pr\{f_n(Z_{n,t}) \leqq u + K\cdot\psi\} - \Pr\{u < f_n(Z_{n,t}) \leqq u + K\cdot\psi\}$$
$$\geqq \Pr\{f_n(T_{n,t}(Z_{n,t})) \leqq u\} - \Pr\{u - K\cdot\psi < f_n(T_{n,t}(Z_{n,t})) \leqq u + 2K\cdot\psi\}$$
$$= \lambda_f^{(n)}(u) - [\lambda_f^{(n)}(u + 2K\cdot\psi) - \lambda_f^{(n)}(u - K\cdot\psi)] \equiv \lambda_f^{(n)}(u) - B(n, u, K\cdot\psi).$$

Then

$$(13) \qquad -\lambda_g^{(n)}(t)B(n, u, K\cdot\psi) \leqq \Pr\{f_n(Y_n) \leqq u; g_n(Y_n) \leqq t\}$$
$$-\lambda_g^{(n)}(t)\lambda_f^{(n)}(u) \leqq \lambda_g^{(n)}(t)A(n, u, K\cdot\psi),$$

and (13) furnishes bounds for the departure of $f_n(Y_n)$ and $g_n(Y_n)$ from independence.

The bounds (13) are now applied in connection with an example involving a specialization of the material in Section 3. In this example, 38 observations $V_1, \cdots, V_{38}$ were hypothesized to constitute a random sample from a population distributed according to a certain c.d.f. $F = F_0$. Alternatives feared were (1) $F \neq F_0$, and (2) the presence of one or more high outliers. It was therefore decided to perform simultaneously (1) a Kolmogoroff-Smirnoff test at the $1 - (.95)^{\frac{1}{2}}$ level, and (2) a $1 - (.95)^{\frac{1}{2}}$ level test based on $\max(V_i)$.

The random variables $F_0(V_i)$ now play the role of the random variables U_i of Section 3. If $\Phi_{38}(F_0(V))$ is defined as in (11), reference to formula (5) of [8], with $n = 38$, $\alpha = (1 - (.95)^{\frac{1}{2}})/2 = .01266$, and $A(\alpha) = .17$, shows that the critical $100(1 - (.95)^{\frac{1}{2}})$ % value for $\Phi_{38}(F_0(V))$ is $((38)^{\frac{1}{2}})(.2347)$. Also, the critical $100(1 - (.95)^{\frac{1}{2}})$ % value for $F_0(\max(V_i)) = \max(F_0(V_i))$ is $((.95)^{\frac{1}{2}})^{1/38} = .99932$. As indicated in Section 0, simultaneously performing the test based on $\Phi_{38}(F_0(V))$ and the test based on $\max(F_0(V_i))$ amounts to verifying whether the 38 "risers" of the sample c.d.f. lie within a region formed by the intersection of the two-sided Kolmogoroff-Smirnoff acceptance band with boundaries $F_0(v) \pm .2347$, and the half-plane to the left of $v = F_0^{-1}(.99932)$, which, for the particular population c.d.f. in question here, equals 70. We now use (13) to bound the actual level of this nominally 5 % joint test.

Consulting (10) and (12) of Section 3, we find, since $s = 0$ in this application, that $K(s, t) = (2t)(2^t)$; also that $\psi(n) = n^{-\frac{1}{2}}$. However, the further simplification $l = 0$ of the present example enables us to replace $(2t)(2^t)$ by t; this because $T_{H(n,t)}$ now has the very tractable form $T_{H(n,t)}(x) = x(1 - t/n)^{-1}$, enabling us to replace the right-hand side of (10) by t/n.

It remains to identify the various quantities appearing in (13): $n = 38$; $u = (38)^{\frac{1}{2}}(.2347) = 1.447$; $\psi = 1/(38)^{\frac{1}{2}}$; $K = t = 38(1 - .99932)$; $K\cdot\psi = .0042$; $B(n, u, K\cdot\psi) = \lambda_f^{(38)}(1.447 + .0084) - \lambda_f^{(38)}(1.447 - .0042) \doteq$

.002, and $A(n, u, K \cdot \psi) = \lambda_f^{(38)}(1.447 + .0042) - \lambda_f^{(38)}(1.447 - .0084) \doteq .002$, where $\lambda_f^{(38)}$ may be computed with the help of [8], or, to an approximation adequate for the present computation, from tabulations of the asymptotic distribution λ_f ; $\lambda_g^{(38)}(t) = (.95)^{\frac{1}{2}}$. The actual level of the joint test is thus no less than 4.8 %, nor greater than 5.2 %.

REFERENCES

[1] ANDERSON, T. W. and DARLING, D. A. (1952). Asymptotic theory of certain 'goodness of fit' criteria based on stochastic processes. *Ann. Math. Statist.* **23** 193–212.

[2] BELL, C. B. (1960). On the structure of distribution-free statistics. *Ann. Math. Statist.* **31** 703–709.

[3] BIRNBAUM, Z. W. (1953). Distribution-free tests of fit for continuous distribution functions. *Ann. Math. Statist.* **24** 1–8.

[4] BIRNBAUM, Z. W. and RUBIN, H. (1954). On distribution-free statistics. *Ann. Math. Statist.* **25** 593–598.

[5] CHAPMAN, DOUGLAS G. (1958). A comparative study of several one-sided goodness-of-fit tests. *Ann. Math. Statist.* **29** 655–674.

[6] CRAMÉR, HARALD (1946). *Mathematical Methods of Statistics*. Princeton Univ. Press.

[7] DEMPSTER, A. P. (1959). Generalized Dn^+ statistics. *Ann. Math. Statist.* **30** 593–597.

[8] MILLER, LESLIE H. (1956). Table of percentage points of Kolmogorov statistics. *J. Amer. Statist. Assoc.* **51** 111–121.

[9] PYKE, RONALD (1959). The supremum and infimum of the Poisson process. *Ann. Math. Statist.* **30** 568–576.

[10] ROSENBLATT, MURRAY (1952). Remarks on a multivariate transformation. *Ann. Math. Statist.* **23** 470–472.

[11] SCHEFFÉ, HENRY (1947). A useful convergence theorem for probability distributions. *Ann. Math. Statist.* **18** 434–438.

[12] SHERMAN, B. (1950). A random variable related to the spacing of sample values. *Ann. Math. Statist.* **21** 339–361.

[13] WALD, A. and WOLFOWITZ, J. (1939). Confidence limits for continuous distribution functions. *Ann. Math. Statist.* **10** 105–118.