4. Acknowledgment. I am grateful to Dr. Robert E. Greenwood for his advice and guidance.

REFERENCES

[1] T. Austin, R. Fagen, T. Lehrer, and W. Penney, "The distribution of the number of locally maximal elements in a random sample," *Ann. Math. Stat.*, Vol. 28 (1957), pp. 786–90.

[2] Tomlinson Fort, *Finite Differences and Difference Equations*, Clarendon Press, Oxford, 1948.

[3] Clarence Hudson Richardson, *An Introduction to the Calculus of Finite Differences*, D. Van Nostrand, New York, 1954.

THE LIMITING JOINT DISTRIBUTION OF THE LARGEST AND SMALLEST SAMPLE SPACINGS[1]

By Lionel Weiss

Cornell University

1. Introduction and summary. $X_1, X_2, \cdots, X_n$ are independent chance variables, each with the same distribution. This common distribution assigns all the probability to the closed interval [0, 1], and has a density function $f(x)$ whose graph consists of any finite number of horizontal line segments. That is, there are H non-degenerate subintervals

$$I_1, I_2, \cdots, I_H, \qquad I_1 = [0, z_1), I_2 = [z_1, z_2), \cdots, I_H = [z_{H-1}, 1],$$

and for each x in I_j, $f(x) = a_j$. We assume that a_j is positive for all j. Let z_0 denote zero, and z_H denote unity. M will denote $\min_j a_j$, B will denote

$$\sum_{j: a_j = M} (z_j - z_{j-1}),$$

and S shall denote $\int_0^1 f^2(x)\,dx = \sum_{j=1}^H a_j^2(z_j - z_{j-1})$.

Let $Y_1 \leqq Y_2 \leqq \cdots \leqq Y_n$ denote the ordered values of $X_1, \cdots, X_n$, and define $W_1 = Y_1$, $W_2 = Y_2 - Y_1, \cdots, W_n = Y_n - Y_{n-1}$, $W_{n+1} = 1 - Y_n$, $U_n = \min(W_1, \cdots, W_{n+1})$, $V_n = \max(W_1, \cdots, W_{n+1})$. In [1] it is shown that if $f(x)$ is the uniform density function over [0, 1], then

$$\lim_{n\to\infty} P\left[U_n > \frac{u}{(n+1)^2}, \qquad V_n < \frac{\log(n+1) - \log v}{n+1}\right] = \exp\{-(u+v)\},$$

for any positive numbers u, v. It is easy to see that the convergence must be uniform over any bounded rectangle in the space of u and v. In this paper it is shown that if $f(x)$ is of the type described above, then

$$\lim_{n\to\infty} P\left[U_n > \frac{u}{(n+1)^2}, V_n < \frac{\log(n+1) + \log M - \log v}{M(n+1)}\right] = \exp\{-(Su + Bv)\},$$

Received August 27, 1958; revised December 10, 1958.

[1] Research under contract with the Office of Naval Research.

for any positive values u, v. This result can be used to study the asymptotic power of various tests of fit based on U_n and V_n which have been proposed (see [1], p. 253).

2. Derivation of the distribution. Let N_j denote the number of the values $X_1, \cdots, X_n$ which fall in the subinterval I_j, and let ${}_jY_1 \leqq {}_jY_2 \leqq \cdots \leqq {}_jY_{N_j}$ be the ordered values of the observations in I_j. Denote z_{j-1} by ${}_jY_0$, and z_j by ${}_jY_{N_j+1}$. Define ${}_jW_h$ as ${}_jY_h - {}_jY_{h-1}$ for $h = 1, \cdots, N_j + 1$. The joint conditional distribution of

$$\frac{{}_jW_1}{z_j - z_{j-1}}, \cdots, \frac{{}_jW_{N_j+1}}{z_j - z_{j-1}}$$

given N_j is the same as the joint distribution of $N_j + 1$ sample spacings created by N_j independent observations on a uniform distribution over [0, 1], and depends only on N_j. Denote min $({}_jW_1, \cdots, {}_jW_{N_j+1})$ by $U(N_j)$, and

$$\max ({}_jW_1, \cdots, {}_jW_{N_j+1})$$

by $V(N_j)$. Then we have from the theorem of ([1], p. 252)

$$\lim_{\substack{\min N_j \to \infty \\ j}} P\left[\frac{U(N_j)}{z_j - z_{j-1}} > \frac{u_j}{(N_j+1)^2}, \quad \frac{V(N_j)}{z_j - z_{j-1}} < \frac{\log (N_j + 1) - \log v_j}{N_j + 1};\right.$$
$$\left. \text{all } j \mid N_1, \cdots, N_H\right] = \exp\left\{-\sum_{j=1}^{H} (u_j + v_j)\right\},$$

and this approach is uniform over any bounded subset of $u_1, \cdots, u_H, v_1, \cdots, v_H$ space. Now we set

$$u_j = \left(\frac{N_j + 1}{n + 1}\right)^2 \frac{u}{z_j - z_{j-1}}$$

$$\log v_j = \log (N_j + 1) - \frac{N_j + 1}{M(n+1)(z_j - z_{j-1})} \log\left[\frac{M(n+1)}{v}\right].$$

Then the inequalities in the conditional probability above become

$$U(N_j) > \frac{u}{(n+1)^2}$$

$$V(N_j) < \frac{\log (n+1) + \log M - \log v}{M(n+1)}.$$

$(N_j + 1) / (n + 1)$ can be written as $a_j(z_j - z_{j-1}) + Z_j(n)$, where $Z_j(n)$ is a chance variable such that $n^{\frac{1}{2}-\delta}Z_j(n)$ converges to zero with probability one as n increases, for any positive δ. This means that $Z_j(n) \log (n + 1)$ converges to zero with probability one as n increases. Then it is easily verified that as n increases, with probability one u_j converges to $ua_j^2(z_j - z_{j-1})$ and v_j converges to $v(z_j - z_{j-1})$ if $a_j = M$ and to zero if $a_j > M$. We turn our conditional probability above into an unconditional probability by taking the expectation with respect to $N_1, \cdots, N_H$, and we find

$$\lim_{n\to\infty} P\left[\min_j U(N_j) > \frac{u}{(n+1)^2}, \quad \max_j V(N_j) < \frac{\log(n+1) + \log M - \log v}{M(n+1)}\right] = \exp\{-(Su + Bv)\},$$

for any positive values u, v.

The final step in our derivation is to show that with probability approaching one as n increases, $U_n = \min_j U(N_j)$ and $V_n = \max_j V(N_j)$. If

$$U_n \neq \min_j U(N_j),$$

it means that for some j, $U(N_j) = {}_jW_1$ or ${}_jW_{N_j+1}$. But from the symmetry of the joint distribution of ${}_jW_1, \cdots, {}_jW_{N_j+1}$, it is easily seen that

$$\lim_{n\to\infty} P[U(N_j) = {}_jW_1 \text{ or } {}_jW_{N_j+1}] = 0.$$

Then $\lim_{n\to\infty} P[U_n = \min_j U(N_j)] = 1$. If $V_n \neq \max_j V(N_j)$, it implies that the sample spacing of maximum length contains one of the values $z_0, \cdots, z_H$. Denote by T_j the length of the sample spacing which contains z_j. Simple calculations show that the limiting distribution of $(n+1)T_j$ has a bounded density function. But $V_n \geqq \max_j V(N_j)$, and the limiting probability given in the preceding paragraph shows that $(n+1)\max_j V(N_j)$ grows without bound as n increases. Thus with probability approaching one as n increases,

$$V_n > \max_j T_j.$$

3. The large-sample power of certain tests of fit based on U_n and V_n. The statistics V_n/U_n and $V_n - U_n$ have been proposed to test the hypothesis that $f(x)$ is the uniform density function. We define U_n' as $(n+1)^2 U_n$, and V_n' as $M(n+1)V_n - \log(n+1)$. Above we showed that U_n', V_n' have a nondegenerate joint limiting distribution as n increases. We have

$$\frac{M}{(n+1)\log(n+1)} \frac{V_n}{U_n} = \frac{1}{\log(n+1)} \frac{V_n'}{U_n'} + \frac{1}{U_n'}$$

$$M(n+1)(V_n - U_n) - \log(n+1) = V_n' - \frac{MU_n'}{(n+1)}.$$

These relations imply that asymptotically, the test based on V_n/U_n is equivalent to the test based on U_n alone, and the test based on $V_n - U_n$ is equivalent to the test based on V_n alone. Small values of U_n are critical, large values of V_n are critical. Using the results above, we find that the asymptotic critical value for U_n if the desired level of significance is α is $-\log(1-\alpha)/(n+1)^2$, and the asymptotic power of the test based on U_n is $1 - (1-\alpha)^S$. Thus the test based on U_n is not consistent against any alternative of the type we are considering. However, it is not difficult to show that the test based on V_n is consistent against any such alternative.

REFERENCE

[1] D. A. DARLING, "On a class of problems related to the random division of an interval," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 239–253.

GENERALIZED D_n^+ STATISTICS[1]

By A. P. DEMPSTER

Harvard University

1. Introduction. The purpose here is to present simplified derivation methods which can be applied to generalizations of some distributions derived by Birnbaum and Tingey [1] and Birnbaum and Pyke [2]. In the case of [1] the generalization is explicitly written down as equation (5)′. Other authors have noticed this generalization; it appears implicitly in equation (31) of Chapman [3] and is given explicitly by Pyke [4]. However the derivation given in the following section differs from the methods of other authors and gives a probabilistic meaning to each term in the summation formula (5)′. In the case of [2] explicit formulas are given for a special case of our generalization different from that considered by Birnbaum and Pyke.

Consider a sample of n from the uniform distribution on (0, 1). Denote the sample c.d.f. by $F_n(x)$. The relevant part of the curve $y = F_n(x)$ is entirely contained by the closed unit square $0 \leqq x \leqq 1$ and $0 \leqq y \leqq 1$, and within this square the population c.d.f. is represented by the line $y = x$. For $0 \leqq \delta < 1$ and $0 < \epsilon < 1$ the line joining $(0, \delta)$ and $(1 - \epsilon, 1)$ will be referred to as barrier (δ, ϵ). A set of such barriers moving away from $y = x$ may be conceived of, and we are concerned with a set of probabilistic questions about which barriers are crossed and where by the curve $y = F_n(x)$ as it passes from (1, 1) to (0, 0)

2. The basic derivation. Denote by $f_j (0 \leqq j \leqq n - 1)$ the probability that $y = F_n(x)$ crosses the barrier (δ, ϵ) at level $y = (n - j) / n$ not having crossed it at any level $y = (n - i) / n$ for $i < j$. Denote the abscissa of the intersection of the barrier (δ, ϵ) and level $y = (n - j) / n$ by m_j. Then it is easily checked that

$$m_j = \frac{1 - \epsilon}{1 - \delta}\left(1 - \delta - \frac{j}{n}\right). \tag{1}$$

Finally, let us use $b(r, s, p)$ for the binomial probability $\binom{s}{r} p^r (1 - p)^{s-r}$

An expression for f_j may be derived as follows. Given that $y = F_n(x)$ passes

Received May 27, 1958; revised November 24, 1958.

[1] Work done in part at Princeton University while the author was supported by the National Research Council of Canada, and in part at Bell Telephone Laboratories while the author was a Member of Technical Staff.