The Annals of Probability
1975, Vol. 3, No. 1, 166–168

THE EXTENT OF DISCRETE DISTRIBUTIONS

BY R. M. MEYER

University of Rochester

The *extent* $\mathscr{E}$ of a probability distribution is defined to be the range of its probability measure, thus, it represents the set of possible probabilities of events. The structure of $\mathscr{E}$ is examined for discrete distributions with (countably) infinite support. An application is considered.

1. Introduction. For any random variable X with cdf F it is clear that as B varies over the class $\mathscr{B}$ of Borel sets, $P(B) = \int_B dF$ varies within $[0, 1]$. The range $\mathscr{E}$ of P is termed the *extent* of the corresponding distribution. Although $\mathscr{E} \equiv [0, 1]$ if F is continuous, the nature of $\mathscr{E}$ is not so clear in other cases. $\mathscr{E}$ is the set of possible probabilities of events. Here we investigate $\mathscr{E}$ in case the distribution is discrete with infinite support.

2. Preliminary results. We begin with the basic

DEFINITION. Let $\{a_n\}_0^\infty$ be a sequence of real numbers such that $\sum |a_n| < \infty$, and let $S = \{s\}$ be the family of all subsets of $\{0, 1, 2, \cdots\}$. The extent $\mathscr{E}$ of $\{a_n\}$ is the set of real numbers $\{\sum_{n \in s} a_n;\ s \in S\}$.

For purposes of investigating the extent of discrete distributions with countably infinite support, there is no loss of generality in assuming that the support is $\{0, 1, 2, \cdots\}$ and each term is positive. For such distributions $\{a_n\}_0^\infty$, $\mathscr{E}$ is a subset of $[0, 1]$, symmetric about $\frac{1}{2}$ and including 0 and 1. The (geometric) distribution $a_n = (\frac{1}{2})^{n+1}$ has *full* extent $\mathscr{E} \equiv [0, 1]$ whereas the distribution $\{\frac{5}{12}, \frac{4}{12}, a_2, a_3, \cdots\}$ has not since any ξ satisfying $\frac{3}{12} < \xi < \frac{4}{12}$ or $\frac{8}{12} < \xi < \frac{9}{12}$ fails to belong to $\mathscr{E}$.

By using a Cantor Diagonalization argument on any subsequence $\{a_{n_i}\}$ of $\{a_n\}$ satisfying: $0 < a_{n_0} < \frac{1}{2}$, $a_{n_i} < (a_{n_{i-1}})^{i+1}$, $n_{i-1} < n_i$ $(i = 1, 2, \cdots)$ it can be proved that the extent of $\{a_{n_i}\}$ hence of $\{a_n\}$ is uncountable.

The set $\mathscr{E}$, which is dense in itself, is closed and hence perfect. For, let $\{\delta_n\}$ be a sequence in $\mathscr{E}$ converging to (some) ξ. ($\delta_n = \sum_0 \delta_{in} a_i$; $\delta_{in} = 0, 1$; $n = 1, 2, \cdots$.) First let $\{\delta_n^{(0)}\}$ be any subsequence of $\{\delta_n\}$ chosen so that $\delta_{0n}^{(0)} \equiv \delta_0^* \equiv \liminf_m \delta_{0m}$; thus $\delta_n^{(0)} = \delta_0^* a_0 + \sum_1 \delta_{in}^{(0)} a_i$ $(n = 1, 2, \cdots)$. Next let $\{\delta_n^{(1)}\}$ be any subsequence of $\{\delta_n^{(0)}\}$ chosen so that $\delta_{1n}^{(1)} \equiv \delta_1^* = \lim_n \inf \delta_{1n}^{(0)}$; thus $\delta_n^{(1)} = \delta_0^* a_0 + \delta_1^* a_1 + \sum_2 \delta_{in}^{(1)} a_i$ $(n = 1, 2, \cdots)$ and so on. Since the common limit of all of these (sub) sequences is ξ, it follows that $\xi = \sum_0 \delta_i^* a_i$ and so $\xi \in \mathscr{E}$.

Upon examining relationships among the extents $\mathscr{E}_m$ of $\{a_n\}_m^\infty$ $(m = 0, 1, 2, \cdots)$ we obtain the following:

THEOREM 1. *For a distribution $\{a_n\}$ define $T_k = a_k + a_{k+1} + \cdots$. If for all n*

Received March 23, 1973; revised March 29, 1974.

AMS 1970 *subject classifications.* Primary 60; Secondary 02, 20.

Key words and phrases. Extent, discrete distribution, events, probability inequalities, gap.

sufficiently large $a_n > T_{n+1}$ *and, as* $n \to \infty$, $2^n T_n \to 0$ *then the measure of the extent* $\mathscr{E}$ *of* $\{a_n\}$ *is zero.*

Proof consists of showing that the conditions guarantee that the measure of the set of points in [0, 1] not belonging to $\mathscr{E}$ is one.

From Theorem 1 follows that the Geometric distribution $a_n = p(1-p)^n$ has extent of measure zero if $\frac{1}{2} < p < 1$ (in fact, for $p = \frac{2}{3}$, $\mathscr{E}$ is the Cantor Ternary set) and the Poisson distribution $a_n = e^{-\lambda}\lambda^n/n!$ has extent of measure zero for any $\lambda > 0$.

3. Additional results. For any distribution of the type being considered and any $\xi \in [0, 1]$ define a bounded, non-decreasing sequence $\{P_n\}_0^\infty$ as follows:

$$\begin{aligned} P_0 &= a_0, \quad \text{if} \quad a_0 \leqq \xi, \qquad & P_n &= P_{n-1} + a_n, \quad \text{if} \quad P_{n-1} + a_n \leqq \xi \\ &= 0, \quad \text{if} \quad a_0 > \xi, & &= P_{n-1}, \quad \text{if} \quad P_{n-1} + a_n > \xi \qquad (n > 0). \end{aligned}$$

Now $P_n \uparrow P \leqq \xi$ and clearly if $P = \xi$ then $\xi \in \mathscr{E}$. Furthermore, if $P_{n-1} = P_n$ for some n then $\xi - P_{n-1} < a_n$ so that if $P_{n-1} = P_n$ for infinitely many n, $P_n \to \xi$ and $\xi \in \mathscr{E}$. Thus we have

THEOREM 2. *For a distribution* $\{a_n\}$ *define* $T_k = a_k + a_{k+1} + \cdots$. *If* $a_n < T_{n+1}$ $(n = 0, 1, \cdots)$ *then* $\mathscr{E} \equiv [0, 1]$; *if* $a_n < T_{n+1}$ $(n \geqq N)$ *then* $\mathscr{E}$ *has positive measure.*

For the first part, the hypotheses guarantee that for any $\xi \in (0, 1)$, $P_{n-1} = P_n$ for infinitely many n, hence $\xi \in \mathscr{E}$ and $\mathscr{E}$ is full. The second part follows by considering $\mathscr{E}_N$, the extent of $\{a_n\}_N^\infty$. Thus, for example, the Geometric distribution has full extent for $0 < p < \frac{1}{2}$.

Alternate conditions determining the measure of $\mathscr{E}$ are based upon the asymptotic behavior of the familiar ratios a_{n+1}/a_n. Thus we have

THEOREM 3. *For the distribution* $\{a_n\}$ *define* $\underline{A} = \liminf_n a_{n+1}/a_n$ *and* $\bar{A} = \limsup_n a_{n+1}/a_n$. *If* $\bar{A} < \frac{1}{2}$ *then* $\mathscr{E}$ *has measure zero; if* $\underline{A} > \frac{1}{2}$ *then* $\mathscr{E}$ *has positive measure (with no conclusion if* $\underline{A} \leqq \frac{1}{2} \leqq \bar{A}$).

The above theorem follows since the conditions $\bar{A} < \frac{1}{2}$ and $\underline{A} > \frac{1}{2}$ imply the conditions (respectively) of Theorem 1 and Theorem 2.

From Theorem 3 follows that

(a) log-convex distributions $(a_n^2 \geqq a_{n-1}a_{n+1}, n = 1, 2, \cdots)$ for which $a_{n+1}/a_n \downarrow A$ always exists have full extent if $A > \frac{1}{2}$ and measure zero extent if $A < \frac{1}{2}$.

(b) log-concave distributions $(a_n^2 \leqq a_{n-1}a_{n+1}, n = 1, 2, \cdots)$ for which $a_{n+1}/a_n \uparrow A$ always exists have extent of measure zero if $A < \frac{1}{2}$ and positive measure if $A > \frac{1}{2}$. The Logarithmic series distribution $a_n = \alpha\theta^{n+1}/n + 1$ $(0 < \theta < 1)$ is log-concave, for example, with $A = \theta$.

4. Continuation. Let S be a subset of [0, 1]. An open interval $I \subset [0, 1]\backslash S$ is called a *gap interval* of S. The supremum g of the lengths of all gap intervals of S is called the *gap* of S. Clearly if $S_1 \subset S_2$ then $g_1 \geqq g_2$ where g_i is the gap of S_i $(i = 1, 2)$. If $S = \mathscr{E}$, the extent of a distribution $\{a_n\}$, then g is called the gap of the distribution (or its extent). Clearly, $g = 0$ if $\mathscr{E}$ is full whereas the Geometric distribution, for example, has gap $g = 2p - 1$ for $\frac{1}{2} < p < 1$.

THEOREM 4. *For a distribution* $\{a_n\}$ *with extent* $\mathscr{E}$ *the gap of the distribution satisfies*: $g \leqq \max_n a_n$.

The proof follows since $S = \{T_0, T_1, \cdots\} \subset \mathscr{E}$ and the gap of S is $\max_n a_n$.

For a distribution such as the Poisson (λ), g is small even for $\lambda \geqq 1$ ($g \to 0$ as $\lambda \to \infty$); however for $\lambda \leqq 1$ g is surprisingly large. Specifically, we have the following values for (λ, g): (.20, .62), (.40, .34), (.60, .19), (.80, .15), (1.00, .10). Thus, for example, with $\lambda = .20$ there is an interval of length .62 in which the probability of no event can fall.

Possible improvement of probability inequalities could result using the knowledge of extent and gap. Suppose A is an event (subset of 0, 1, 2, $\cdots$) such that $P(A) \leqq \alpha$ is known. If α lies in a gap interval $I = (\alpha_1, \alpha_2)$ of the distribution we can automatically improve the inequality to $P(A) \leqq \alpha_1$. Similarly for reverse inequalities. Maximum improvement is g, the gap of $\mathscr{E}$.

5. Concluding remarks. Further properties related to extent have been considered and will be presented later. Thanks to M. C. Trivedi who determined gaps for the Poisson distribution.

DEPARTMENT OF STATISTICS
THE UNIVERSITY OF ROCHESTER
ROCHESTER, NEW YORK 14627