ON THE SELECTION OF SUBSEQUENCES

OYSTEIN ORE

The problem we shall discuss in the following is concerned with the possibility of extracting from a given sequence a subsequence whose terms obey some relatively simple mathematical law. As a prototype of such questions one may take the prime representing function determined by W. H. Mills [3] which has the property of defining within the sequence of *all* prime numbers a subsequence whose terms have a mathematical expression of simple type.

We shall establish a general theorem of this kind which makes it possible to formulate selection laws for a great variety of sequences when only quite weak conditions are prescribed in advance.

Among the various applications of this result let us mention especially its bearing upon the so-called frequency definition of probability. Here, as one may recall, the natural symbol for the repetition of an event is taken to be a sequence of signs s and f, denoting success and failure. Such a sequence can equally well be described by a series of increasing integers

$$(1) \qquad\qquad A_i = \{a_i\}$$

where each a_i denotes the location of the ith success.

The first requirement on these sequences, according to von Mises' theory, is that the relative frequency of successes shall have a limit

$$(2) \qquad\qquad \lim_{n\to\infty} \frac{\nu(n)}{n} = p$$

where $\nu(n)$ denotes the number of successes in the first n trials, that is, the number of terms in (1) which are $\leq n$.

The second requirement is the so-called *irregularity condition*: It shall be impossible to extract a subsequence of the given sequence such that its limit frequency is different from (2). It is immediately clear that since there exist subsequences corresponding only to successes, that is, having a frequency 1, it is necessary in the definition of irregularity to introduce some restriction as to the method of choice for the terms in those subsequences to which the irregularity condition shall refer.

Our Theorem 5 gives a contribution to this discussion and adds to the difficulties which would have to be surmounted in phrasing the

permitted mathematical operations, should one wish to maintain the frequency view, since it is shown that for any sequence satisfying (2) with $p>0$ it is possible to formulate in simple mathematical terms a process of selection which produces a subsequence of frequency 1.

The main principle of the following considerations is an extension of the idea introduced by Mills.

1. We shall assume for the present only that (1) is a sequence of real, positive numbers tending to infinity. Furthermore, there shall be associated with it some function $f(x)$ having the following properties:

(a) $f(x)$ is continuous and monotone increasing for $x>M$, where M is some fixed positive number.

(b) for each $a_i>M$ one can find two quantities $\delta_i \geqq 0$, $\epsilon_i>0$ such that within the interval

$$(f(a_i - \delta_i), f(a_i + \epsilon_i))$$

is always contained another term a_j in (1), with $a_j>a_i$.

It is clear that when a function of this kind exists, it is always possible to extract an increasing subsequence

$$(3) \qquad\qquad B = \{b_i\}, \qquad\qquad b_i > M,$$

from (1) such that

$$(4) \qquad\qquad f(b_i - \delta_i) < b_{i+1} < f(b_i + \epsilon_i)$$

where δ_i and ϵ_i now shall indicate the quantities in (b) which correspond to the term b_i in (1).

We shall now make the condition (b) more strict and assume instead that our selection process has the following *nesting property*:

(b*) It shall be possible to select the terms in the subsequence (3) in such a manner that

$$(5) \quad f(b_i - \delta_i) < b_{i+1} - \delta_{i+1} \leqq b_{i+1} < b_{i+1} + \epsilon_{i+1} < f(b_i + \epsilon_i).$$

Any function satisfying the conditions (a) and (b*) we shall call a *selection function* for the sequence (1). We may observe that in most applications the selection function $f(x)$ is *universally defined* so that for *every* $x>M$ there exists some term $a_i>x$ of (1) in the interval

$$(f(x - \delta_x), f(x + \epsilon_x)).$$

2. The inverse function of the selection function $f(x)$ shall be denoted by $f_{-1}(x)$ and the ith iterations of $f(x)$ and $f_{-1}(x)$ respectively by $f_i(x)$ and $f_{-i}(x)$. It is evident from (a) that they are all con-

tinuous and monotone increasing functions for sufficiently large values of x.

We shall now introduce two new sequences

$$U = \{u_i\}, \qquad V = \{v_i\}$$

whose terms are defined as follows:

(6) $$u_i = f_{-i}(b_i - \delta_i), \qquad v_i = f_{-i}(b_i + \epsilon_i).$$

The first sequence U is increasing since we obtain stepwise from (5)

$$u_{i+1} = f_{-i-1}(b_{i+1} - \delta_{i+1}) = f_{-i}(f_{-1}(b_{i+1} - \delta_{i+1})) > f_{-i}(b_i - \delta_i) = u_i.$$

Similarly one establishes from (5) that V is a decreasing sequence since

$$v_{i+1} = f_{-i-1}(b_{i+1} + \epsilon_{i+1}) = f_{-i}(f_{-1}(b_{i+1} + \epsilon_{i+1})) < f_{-i}(b_i + \epsilon_i) = v_i.$$

Finally the definitions (6) show that

$$u_i < v_i,$$

so that U is bounded from above.

We conclude that the limit

$$\lim_{i \to \infty} u_i = K$$

exists where K is some positive number satisfying the conditions

$$u_i < K < v_i$$

for every i, hence also

$$f_i(u_i) < f_i(K) < f_i(v_i).$$

This according to (6) leads to the result:

THEOREM 1. *If $f(x)$ is a selection function for the sequence* (1), *then there exists a subsequence $B = \{b_i\}$ and a positive constant K such that*

(7) $$b_i - \delta_i < f_i(K) < b_i + \epsilon_i.$$

From this basic result we shall draw several consequences in the following. Let us say that $f(x)$ is a *convergent selection function* if it satisfies the conditions

$$\lim_{i \to \infty} \delta_i = \lim_{i \to \infty} \epsilon_i = 0.$$

We can then state:

THEOREM 2. *If the sequence* (1) *has a convergent selection function*

$f(x)$, *it is possible to determine a positive constant K such that the values $f_i(K)$ converge to a subsequence $B = \{b_i\}$ of* (1), *in the sense that*

$$\lim_{i \to \infty} (b_i - f_i(K)) = 0.$$

There is, as one sees, no particular limitation in the preceding consideration by taking the interval in (b) such that always $\delta_i = 0$. The condition (5) in (b*) then assumes the "one-sided form":

(b**) $f(b_i) < b_{i+1} < b_{i+1} + \epsilon_{i+1} < f(b_i + \epsilon_i).$

Correspondingly, when a selection function fulfills this condition, the main result in (7) becomes

(8) $b_i < f_i(K) < b_i + \epsilon_i$

and in the case of a convergent selection function the approach to the subsequence in Theorem 2 is always from above.

3. From now on we shall suppose that our infinite sequence (1) consists of positive integers. We obtain as an immediate consequence of (8):

THEOREM 3. *If* (1) *is a sequence of integers and $f(x)$ a selection function satisfying* (b**) *with $\epsilon_i < 1$ for every i, then one can find a positive constant K such that*

(9) $b_i = [f_i(K)]$ $(i = 1, 2, \cdots)$

is a subsequence of (1).

In the formula (9) for the subsequence the brackets as usual denote "the greatest integer contained in."

The general result in Theorem 3 may be illustrated on the prime representing function given by Mills [3]. One can show that for the sequence of primes the function

(10) $f(x) = x^3$

is a selection function satisfying the conditions of this theorem and therefore there exists a constant K such that

$$p_i = [K^{3^i}]$$

for every i.

A slight extension of Mills' reasoning gives the somewhat better result that in (10) we actually have a convergent selection function so that K^{3^i} converges to primes for increasing i. To verify this we use the result given by A. E. Ingham [1] that if p_n and p_{n+1} are consecu-

tive primes, then

$$p_{n+1} < p_n + k \cdot p_n^{5/8}$$

for some suitable positive constant k. Now let N be an arbitrary positive integer and p_n the greatest prime below N^3. Then one obtains

$$p_n < N^3 < p_{n+1} < p_n + k p_n^{5/8} < N^3 + k N^{15/8}$$

or

$$N^3 < p_{n+1} < (N + \epsilon_N)^3 - 1 < (N + \epsilon_N)^3$$

so that (10) is a universally defined selection function for the sequence of primes, satisfying (b**). Since one can take

$$\epsilon_N = \frac{k}{3} N^{-1/8},$$

the selection function is convergent as we asserted.

Let us mention in this connection that, in general, when an everywhere defined selection function is known, it is usually not difficult to produce from it a family of others (compare Kuipers [2], Niven [4], Wright [5]).

4. We shall now return to our probability sequences and deduce the following result:

THEOREM 4. *Let (1) be a probability sequence, that is, a sequence of increasing natural numbers, satisfying the density condition*

(11) $$\lim_{n \to \infty} \frac{v(n)}{n} = p > 0$$

where $v(n)$ is the number of terms $\leq n$. Then any exponential function

(12) $$f(x) = a^x, \qquad\qquad a > 1,$$

is a selection function satisfying the condition (b**) *and one may take*

$$\epsilon_n = \epsilon < 1$$

arbitrarily fixed.

PROOF. Let

$$0 < \epsilon < 1, \qquad a > 1$$

be otherwise arbitrary real numbers. The frequency condition (11) implies that for any two constants k_1 and k_2 such that

$$0 < k_1 < p < k_2$$

one can always select some $N = N(k_1, k_2)$ so large that

(13)
$$k_1 n < \nu(n) < k_2 n$$

for $n > N$. Since the quotient k_2/k_1 can be taken arbitrarily close to unity, we shall assume that these quantities satisfy the condition

$$\frac{k_2}{k_1} < a^\epsilon$$

or

(14)
$$k_1 a^\epsilon - k_2 > 0.$$

To prove that $f(x)$ is a selection function satisfying (b**) it suffices to establish that there always exists an element of the sequence (1) contained in the interval

(15)
$$(a^n, a^{n+\epsilon} - 1).$$

If, however, this were not the case, a contradiction is at hand. When the interval (15) contains no element of (1), one must have

$$\nu(a^n) = \nu(a^{n+\epsilon} - 1).$$

Using the inequalities (13) one concludes that

$$k_2 a^n > k_1(a^{n+\epsilon} - 1)$$

or

$$a^n(k_1 a^\epsilon - k_2) < k_1.$$

But the condition (14) shows that for sufficiently large n this is impossible.

When Theorem 4 is combined with Theorem 3, we have obtained:

THEOREM 5. *Let (1) be a probability sequence with a density $p > 0$. Then there exists a positive constant K such that for sufficiently large n the sequence $\{b_n\}$ with*

$$b_1 = [2^K], \qquad b_2 = [2^{2^K}], \qquad b_3 = [2^{2^{2^K}}], \cdots$$

is a subsequence of (1).

In Theorem 5 we have preferred to take $a = 2$; any other constant $a > 1$ could have been used.

Theorem 5 yields the result which we wished to establish, namely, that for any probability sequence it is possible to define, by relatively

simple mathematical means, a process of selection which produces a subsequence corresponding exclusively to successes from a certain point on.

It is clear that a similar argument may be used for any sequence for which there exists a limiting function $\phi(x)$ for the density such that

$$k_1\phi(n) < \nu(n) < k_2\phi(n).$$

BIBLIOGRAPHY

1. A. E. Ingham, *On the difference between consecutive primes*, Quart. J. Math. Oxford Ser. vol. 8 (1937) pp. 255–266.

2. L. Kuipers, *Prime-representing functions*, Neder. Akad. Wetensch. vol. 53 (1950) pp. 309–310.

3. W. H. Mills, *A prime representing function*, Bull. Amer. Math. Soc. vol. 53 (1947) p. 604.

4. I. Niven, *Functions which represent prime numbers*, Proceedings of the American Mathematical Society vol. 2 (1951) pp. 753–755.

5. E. M. Wright, *A prime representing function*, Amer. Math. Monthly vol. 58 (1951) pp. 616–618.

YALE UNIVERSITY