Selected Techniques of Statistical Analysis, chap. 1, Statistical Research Group, Columbia University, New York, McGraw-Hill Book Co., 1947.
[3] N. L. JOHNSON AND B. L. WELCH, "Applications of the noncentral t-distribution," *Biometrika*, Vol. 31 (1939–1940), pp. 362–389.
[4] ERICH LEHMANN, *Theory of Testing Hypotheses*, Mimeographed notes recorded by Colin Blyth, ASUC Bookstore, University of California, Berkeley, 1948–1949.
[5] S. RUSHTON, "On a sequential t-test," *Biometrika*, Vol. 37 (1950), pp. 326–333.
[6] ABRAHAM WALD, *Sequential Analysis*, John Wiley and Sons, Inc., New York, 1947.

AN EXTENSION OF THE BOREL-CANTELLI LEMMA

BY STANLEY W. NASH

University of British Columbia

1. Introduction. Consider a probability space $(\Omega, \mathfrak{F}, P)$ and a sequence of events $\{A_n\}$, $A_n \varepsilon \mathfrak{F}$, $n = 1, 2, \cdots$. The upper limiting set of the sequence is defined to be

$$\lim_{n\to\infty} \sup_{k\geq n} A_k = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

It is the event that infinitely many of the A_n occur. The purpose of this paper is to find necessary and sufficient conditions for $P(\limsup A_n) = 1$.

The general problem of finding the probability of an infinite number of a sequence of events occurring was considered by Borel [1], [2] and Cantelli [3]. In what follows we shall use the following notations. Let $\alpha_n = I(A_n)$, the indicator of the event A_n (or characteristic function of the set A_n), that is

$$\alpha_n = \begin{cases} 1 \text{ when } A_n \text{ occurs} \\ 0 \text{ when } A_n \text{ fails to occur.} \end{cases}$$

Let $P(A_n \mid \alpha_1\alpha_2 \cdots \alpha_{n-1})$ denote the conditional probability of the event A_n, given the outcomes of the previous $n - 1$ trials. When $n = 1$, the expression is taken to represent the unconditional probability $P(A_1)$. The 1912 Borel criterion stated:

If $0 < p_n' \leqq P(A_n \mid \alpha_1\alpha_2 \cdots \alpha_{n-1}) \leqq p_n'' < 1$ for every n, whatever be α_1, $\alpha_2, \cdots, \alpha_{n-1}$, then $\sum_{j=1}^{\infty} p_j'' < \infty$ implies that $P(\limsup A_n) = 0$, and $\sum_{j=1}^{\infty} p_j' = \infty$ implies that $P(\limsup A_n) = 1$.

Cantelli proved that $\sum_{j=1}^{\infty} P(A_j) < \infty$ always implies that $P(\limsup A_n) = 0$.

Paul Lévy [4] clarified the general problem by proving the following theorem. The subset K (or K') of the sample space Ω for which

$$\sum_{j=1}^{\infty} P(A_j|\alpha_1\alpha_2 \cdots \alpha_{j-1}) < \infty \quad (\text{or } = \infty)$$

and the subset H (or H') of Ω for which $\limsup A_n$ fails to occur (or occurs) differ at most by a set of probability 0. In other words $P(KH') = P(K'H) = 0$ and $P(KH) + P(K'H') = 1$. The hypothesis of the theorem proved in the next section ensures that KH is a null set, so that $P(K'H') = 1$. However, the proof given is direct and independent of Lévy's result.

Loève [5] found necessary and sufficient conditions for $P(\lim\sup A_n) = 0$. Let $p_{nk} = P(A_k \mid \alpha_n = \alpha_{n+1} = \cdots = \alpha_{k-1} = 0)$ for $k > n$, and let $p_{nn} = P(A_n)$. The criterion states:

If $\lim_{n\to\infty} \sum_{k=n}^{\infty} p_{nk} = 0$, then, and only then $P(\lim\sup A_n) = 0$.

Chung and Erdös [6] mentioned the sufficiency of the following criterion for $P(\lim\sup A_n) = 1$.[1]

If $\sum_{k=n}^{\infty} p_{nk} = \infty$ for every n, then, and only then, $P(\lim\sup A_n) = 1$.

2. A necessary and sufficient condition for $P(\lim\sup A_n) = 1$.

CRITERION. If $\sum_{j=1}^{\infty} P(A_j \mid \alpha_1\alpha_2 \cdots \alpha_{j-1}) = \infty$ for every sequence

$$\alpha_1\alpha_2 \cdots \alpha_n \cdots$$

of outcomes of trials for which only finitely many $\alpha_n \equiv I(A_n) = 1$ and no $P(\alpha_1\alpha_2 \cdots \alpha_n) = 0$, then, and only then, $P(\lim\sup A_n) = 1$.

PROOF. The class H of sequences $\lambda = \alpha_1\alpha_2 \cdots \alpha_n \cdots$, for which only finitely many $\alpha_n \equiv I(A_n) = 1$ is denumerable, for its members can be put into one-to-one correspondence with rational numbers between 0 and 1, say with those whose binary expansions have the corresponding sequences of zeros and ones. It follows that, if $P(\lambda) = 0$ for every $\lambda \in H$, then and only then, $\sum_{\lambda\in H} P(\lambda) = 0$, that is $P(\text{Only finitely many } A_n) = 0$, and consequently $P(\text{Infinitely many } A_n) = 1$.

Consider first those sequences $\lambda \in H$ for which $P(\alpha_1\alpha_2 \cdots \alpha_n \text{ as in } \lambda) = 0$ for some finite n. For such sequences $P(\lambda) \leqq P(\alpha_1\alpha_2 \cdots \alpha_n \text{ as in } \lambda) = 0$, since the event λ is a subset of the event that the outcomes of the first n trials are the same as they are in λ, an infinite sequence of trials. Restrict further consideration then to those sequences $\lambda \in H$ for which $P(\alpha_1\alpha_2 \cdots \alpha_n) > 0$ for every n. For such sequences all conditional probabilities $P(\alpha_n \text{ as in } \lambda \mid \alpha_1\alpha_2 \cdots \alpha_{n-1} \text{ as in } \lambda)$ are defined and positive. Accordingly

$$P(\lambda) = \prod_{j=1}^{\infty} P(\alpha_j \text{ as in } \lambda \mid \alpha_1\alpha_2 \cdots \alpha_{j-1} \text{ as in } \lambda)$$
$$= \prod_{j=1}^{\infty} \{1 - P(\alpha_j \textit{ not} \text{ as in } \lambda \mid \alpha_1\alpha_2 \cdots \alpha_{j-1} \text{ as in } \lambda)\}.$$

The infinite product for $P(\lambda)$ is zero if, and only if

$$\sum_{j=1}^{\infty} P(\alpha_j \textit{ not} \text{ as in } \lambda \mid \alpha_1\alpha_2 \cdots \alpha_{j-1} \text{ as in } \lambda) = \infty.$$

For any $\lambda \in H$ all but finitely many $\alpha_n = 0$. Thus the series

$$\sum_{j=1}^{\infty} P(\alpha_j \textit{ not} \text{ as in } \lambda \mid \alpha_1\alpha_2 \cdots \alpha_{j-1} \text{ as in } \lambda)$$

Received 5/19/53.

[1] In a communication to the referee of this paper, Chung and Erdös point out that the statement given in their paper is wrong. There the condition "for every n" is omitted. Also, the proof should not have been attributed to Borel.

and $\sum_{j=1}^{\infty} P(\alpha_j = 1 \mid \alpha_1\alpha_2 \cdots \alpha_{j-1} \text{ as in } \lambda)$ differ in only finitely many terms, hence converge or diverge together. Therefore, $P(\lambda) = 0$ if, and only if,

$$\sum_{j=1}^{\infty} P(A_j \mid \alpha_1\alpha_2 \cdots \alpha_{j-1} \text{ as in } \lambda) = \infty.$$

But $\sum_{\lambda \varepsilon H} P(\lambda) = P(\text{Only finitely many } A_n) = 0$ in this case, and so $P(\limsup A_n) = P(\text{Infinitely many } A_n) = 1$. Q.E.D.

3. An application. Borel's criterion for $P(\limsup A_n) = 1$ is easily seen to be a special case of the criterion of Section 2. To show that the generalization achieved is not trivial consider the following example. Two urns each contain a red and a black ball at the beginning of the experiment. A ball is drawn at random from the first urn, its color noted, and the ball is returned to the urn. This is repeated until a black ball is drawn. Each time a red ball is drawn from the first urn, the number of balls in the second urn is doubled by putting in as many red balls as there were balls of either color in the urn before. Once a black ball has been drawn from the first urn, all further draws are at random from the second urn with replacement after each draw. No further change is made in the composition of the contents of the second urn. Let A_n designate the drawing of a black ball in the nth trial. Consider the sequences of trials for which the $(k - 1)$st trial is the first time a black ball is drawn. Then $\alpha_{k-1} = 1$ but $\alpha_h = 0$ for $h < k - 1$. The second urn will contain 2^k balls at the kth trial and thereafter, $2^k - 1$ red and 1 black. Thus

$$P(A_n \mid \alpha_1\alpha_2 \cdots \alpha_{n-1}) = \begin{cases} \frac{1}{2} \text{ for } n < k \\ 2^{-k} \text{ for } n \geqq k. \end{cases} \qquad (k > 1)$$

Then $p_n' = \inf_k P(A_n \mid \alpha_1\alpha_2 \cdots \alpha_{n-1}) = 2^{-n} > 0$. But $\sum_{j=1}^{\infty} p_j' = \sum_{j=1}^{\infty} 2^{-j} = 1$ converges, so the hypothesis of Borel's criterion does not hold and its conclusion can not be inferred. But $\sum_{j=1}^{\infty} P(A_j \mid \alpha_1\alpha_2 \cdots \alpha_{j-1})$ diverges for every possible sequence λ. In particular it diverges for every $\lambda \varepsilon H$, where only finitely many α_n are ones. Thus, by the criterion of Section 2, $P(\text{Infinitely many } A_n) = P(\text{Black drawn infinitely often}) = 1$.

REFERENCES

[1] ÉMILE BOREL, "Les probabilités dénombrables et leurs applications arithmétiques," *Rend. Circ. Mat. Palermo*, Vol. 27 (1909), pp. 247–271.

[2] ÉMILE BOREL, "Sur un problème de probabilités relatif aux fractions continues," *Math. Ann.*, Vol. 72 (1912), pp. 578–587.

[3] FRANCESCO PAOLO CANTELLI, "Sulla probabilità come limite della frequenza," *Rend. Accad. Lincei*, Series 5, Vol. 24 (1917), Semester 1, pp. 39–45.

[4] PAUL LÉVY, *Théorie de l'Addition des Variables Aléatoires*, Gauthier-Villars, Paris, 1937, pp. 249–250.

[5] MICHEL LOÈVE, "On almost sure convergence," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1951, pp. 279–303, (see pp. 282–283).

[6] KAI-LAI CHUNG AND PAUL ERDÖS, "On the application of the Borel-Cantelli lemma," *Trans. Amer. Math. Soc.*, Vol. 72 (1952), pp. 179–186.