For $n = 2$ and $a_1 = a_2 = b_1 = 1$, $b_2 = -1$ we obtain from (22)

$$f_s(t) = \exp [-(\sigma_1^2 + \sigma_2^2)t^2/2] \qquad s = 1, 2.$$

This shows that $\sigma_1^2 = \sigma_2^2$ and establishes Bernstein's theorem.

REFERENCES

[1] S. BERNSTEIN, "Sur une proprieté caractéristique de la loi de Gauss, *Trans. Leningrad Polytechnic Institute* (1941).

[2] H. CRAMÉR, "Über Eine Eigenschaft Der Normalen Verteilungsfunktion," *Math. Z.* Vol. 41 (1936), pp. 405–414.

[3] M. FRÉCHET, "Généralisation de la loi de probabilité de Laplace, *Ann. Inst. H. Poincaré*, Vol. 13 (1951), pp. 1–29.

[4] M. KAC, "A characterization of the normal distribution," *Amer. J. Math.*, Vol. 61 (1939), pp. 726–728.

[5] YU. V. LINNIK, "Remarks concerning the classical derivation of Maxwell's law," *Doklady Akad. Nauk SSSR*, Vol. 85 (1952), pp. 1251–1254.

[6] J. MARCINKIEWICZ, "Sur une proprieté de la loi de Gauss," *Math. Z.* Vol. 44 (1938), pp. 612–618.

ADDENDUM

The authors are indebted to Professor G. Darmois for calling their attention to his note in the *C. R. Acad. Sci. Paris*, Vol. 232 (1951), pp. 1999–2000 in which he proved the theorem for $n = 2$ without assuming the existence of moments. He later extended this to the case of arbitrary n. His paper will be published in the *Bulletin of the International Statistical Institute*. The method of proof used by Professor Darmois is different from the one presented in this paper. The authors learned that these results were also obtained by methods similar to Darmois' by B. V. Gnedenko (*Izvestiya Akad. Nauk. SSSR, Ser. Mat.*, Vol. 12 (1948), pp. 97–100) for the case $n = 2$ and by V. P. Skitovich (*Doklady Akad. Nauk. SSSR* (N.S.) Vol. 89 (1953), pp. 217–219) for any n.

While reading the proofs of this paper the authors learned that the theorem was also discussed by M. Loève in the appendix to P. Lévy's "Processus stochastiques", Gauthier-Villars, Paris, 1948, pp. 337–338.

ON OPTIMAL SYSTEMS[1]

BY DAVID BLACKWELL

Howard University

1. Summary. For any sequence $x_1, x_2, \cdots$ of chance variables satisfying $|x_n| \leqq 1$ and $E(x_n | x_1, \cdots, x_{n-1}) \leqq -u(\max |x_n| | x_1, \cdots, x_{n-1})$, where u is a fixed constant, $0 < u < 1$, and for any positive number t,

$$\Pr \{ \sup_n (x_1 + \cdots + x_n) \geqq t\} \leqq \left(\frac{1-u}{1+u}\right)^t.$$

Received 7/14/53.

[1] This paper was written under contract Nonr-1197(00) with the Office of Naval Research.

Equality holds for integral t when $x_1, x_2, \cdots$ are independent with

$$\Pr\{x_n = 1\} = (1-u)/2, \qquad \Pr\{x_n = -1\} = (1+u)/2.$$

This has a simple interpretation in terms of gambling systems, and yields a new proof of Lévy's extension of the strong law of large numbers to dependent variables [2], with an improved estimate for the rate of convergence.

2. The theorem and its interpretation. We consider a gambling house which will play any game named by the customer, provided that (1) the customer's maximum gain or loss does not exceed one unit and (2) the customer's expectation does not exceed $-ug$, where g is his maximum gain or loss. A customer with unlimited credit wishes to devise a system of play which will maximize his probability of eventually becoming at least t units ahead, where t is a fixed positive number. Thus a system is a sequence $x_1, x_2, \cdots$ of chance variables satisfying

$$|x_n| \leqq 1 \tag{1}$$

$$E(x_n \mid x_1, \cdots, x_{n-1}) \leqq -u(\max |x_n| \mid x_1, \cdots, x_{n-1}). \tag{2}$$

A particular system is obtained by letting $x_1, x_2, \cdots$ be independent, with $\Pr\{x_n = 1\} = (1-u)/2$ and $\Pr\{x_n = -1\} = (1+u)/2$. For this system, it is known ([1], p. 290) that

$$\Pr\{\max_n (x_1 + \cdots + x_n) \geqq t\} = \left(\frac{1-u}{1+u}\right)^t \tag{3}$$

for any positive integer t. Our theorem is that this is the best system in the sense of maximizing the probability of eventually attaining t, that is, we shall prove the

THEOREM. *For any system $x_1, x_2, \cdots$ satisfying* (1) *and* (2), *and any positive number t,*

$$\Pr\{(x_1 + \cdots + x_n \geqq t \text{ for some } n\} \leqq \left(\frac{1-u}{1+u}\right)^t.$$

PROOF. For any real number t and any system S, let

$$\phi(N, S, t) = \Pr\{\max_{0 \leqq k \leqq n} (x_1 + \cdots + x_k) \geqq t\}, \qquad \phi(N, t) = \sup_S \phi(N, S, t).$$

In particular $\phi(0, S, t) = 1$ for $t \leqq 0$, $= 0$ for $t > 0$. We shall show that

$$\phi(N+1, t) = \sup_{x \varepsilon X} E\phi(n, t-x), \tag{4}$$

where X consists of all chance variables x satisfying $|x| \leqq 1$ and $Ex \leqq -u \max |x|$. Actually (4) is intuitively clear; it asserts that, to maximize the probability of reaching t during $N + 1$ trials one must, for each value of x_1, use that system in the remaining N trials which maximizes the probability of attaining the new required sum $t - x_1$ and one must choose x_1 so that the average probability of attainment in the remaining trails is maximized.

To avoid tedious measurability difficulties, we remark that we may restrict attention to those systems for which (a) each x_n assumes a finite set of values, all rational, (b) $x_n = 0$ for sufficiently large n, and (c) the probability of any particular sequence $x_1, x_2, \cdots$ is rational. Denote by $\mathcal{S}$ the countable set of systems satisfying (a), (b), and (c).

Now any system $S \varepsilon \mathcal{S}$ is described by specifying the initial variable x_1 and for each value v of x_1, the system $S(v) \varepsilon \mathcal{S}$ to be used thereafter when $x_1 = v$.

We have, for $S \varepsilon \mathcal{S}$, $t > 0$,

$$\phi(N + 1, S, t) = E\phi(N, S(x_1), t - x_1), \tag{5}$$

so that

$$\phi(N + 1, S, t) \leqq E\phi(N, t - x_1) \leqq \sup_{x \varepsilon X} E\phi(N, t - x). \tag{6}$$

Taking the sup over $S \varepsilon \mathcal{S}$ in (6) yields

$$\phi(N + 1, t) \leqq \sup_{x \varepsilon X} E\phi(N, t - x). \tag{7}$$

On the other hand, (5) yields

$$E\phi(N, S(x_1), t - x_1) \leqq \phi(N + 1, t). \tag{8}$$

For a fixed initial variable x_1, allowing $S(x_1)$ to range over all $S \varepsilon \mathcal{S}$, independently for the different values of x_1, yields

$$E\phi(N, t - x_1) \leqq \phi(N + 1, t). \tag{9}$$

Since any $x \varepsilon X$ is an admissible initial variable, from (9) we obtain

$$\sup_{x \varepsilon X} E\phi(N, t - x) \leqq \phi(N + 1, t). \tag{10}$$

Inequalities (5) and (10) yield (4) for $t > 0$; for $t \leqq 0$, (4) is obvious, since $\phi(N + 1, t) = 1$ and $E\phi(N, t - x) = 1$ for $x \equiv 0$.

To continue the proof of the theorem, we consider the transformation U, taking Borel-measurable functions of t into Borel-measurable functions of t, defined by

$$Uf(t) = \sup_{x \varepsilon X} Ef(t - x). \tag{11}$$

Equation (4) asserts that $U\phi(N, t) = \phi(N + 1, t)$. We verify, for $g(t) = [(1 - u)/(1 + u)]^t$, that $Ug = g$. To see this, fix t_0 and d, with $0 < d \leqq 1$, and let $h(t)$ be the linear function of t with $h(t_0 - d) = g(t_0 - d)$ and $h(t_0 + d) = g(t_0 + d)$. Then for any $x \varepsilon X$ with $\sup | x | = d$

$$Eg(t_0 - x) \leqq Eh(t_0 - x) = h(t_0 - E(x)) \leqq h(t_0 + ud),$$

with equality if and only if x assumes only the values $+d$, $-d$ and $E(x) = -ud$. Now

$$h(t) = g(t_0 - d) + \frac{g(t_0 + d) - g(t_0 - d)}{2d} (t - t_0 + d),$$

so that

$$r(d) = h(t_0 + ud) = g(t_0)[\tfrac{1}{2}(1 + u)g(d) + \tfrac{1}{2}(1 - u)g(-d)].$$

Since $r(0) = r(1) = g(t_0)$ and r is convex in d, $r(d) \leqq g(t_0)$ for all d, $Eg(t_0 - x) \leqq g(t_0)$ for all $x \varepsilon X$, with equality only for $x \equiv 0$ and $x = \pm 1$ with probabilities $\frac{1}{2}(1 \pm u)$, and $Ug = g$.

To complete the proof of the theorem, we note that $f_1 \leqq f_2$ for all t implies $Uf_1 \leqq Uf_2$ for all t. Since $\phi(0, t) = 1$ for $t \leqq 0$ and 0 for $t > 0$, $\phi(0, t) \leqq g(t)$ for all t. If $\phi(N, t) \leqq g(t)$ for all t, applying U yields

$$U\phi(N, t) = \phi(N + 1, t) \leqq Ug(t) = g(t)$$

for all t so that, by induction, $\phi(N, t) \leqq g(t)$ for all t, N. Consequently $\lim_{N\to\infty} \phi(N, t) \leqq g(t)$. But for any system S,

$$\begin{aligned}\Pr\{x_1 + \cdots + x_n \geqq t \text{ for some } n\} &= \lim_{n\to\infty} \phi(N, S, t) \leqq \lim_{N\to\infty} \phi(N, t) \leqq g(t) \\ &= [(1 - u)/(1 + u)]^t,\end{aligned}$$

and the proof is complete

Corollary. *If* $x_1, x_2, \cdots$ *satisfy* $|x_n| \leqq 1$ *and* $E(x_n \mid x_1, \cdots, x_{n-1}) = 0$, *then* $(x_1 + \cdots + x_n)/n \to 0$ *with probability* 1; *in fact*

$$\Pr\left\{\left|\frac{x_1 + \cdots + x_n}{n}\right| \geqq \epsilon \text{ for some } n \geqq N\right\} \leqq 2\left(\frac{1}{1+\epsilon}\right)^{\epsilon N/(2+\epsilon)} \tag{12}$$

Proof. $\Pr\left\{\frac{(x_1 + \cdots + x_n)}{n} \geqq \epsilon \text{ for some } n \geqq N\right\}$

$$\leqq \Pr\{(x_1 - \epsilon/2) + \cdots + (x_n - \epsilon/2) \geqq \epsilon N/2 \text{ for some } n\} \leqq \left(\frac{1}{1+\epsilon}\right)^{\epsilon N/(2+\epsilon)}$$

where the last inequality is obtained by applying the theorem to the sequence $y_n = (x_n - \epsilon/2)/(1 + \epsilon/2)$, with $t = \epsilon N/(2 + \epsilon)$. The same inequality holds for $\Pr\{(x_1 + \cdots + x_n)/n \leqq -\epsilon \text{ for some } n \geqq N\}$, and the corollary follows, The part of the corollary on convergence with probability 1 is due to Lévy ([2], p. 252). However, his method of proof does not yield a geometric rate of convergence in the sense specified by (12).

Added in proof. T. E. Harris has kindly called my attention to a result of S. Bernstein (see J. V. Uspensky, *Introduction to Mathematical Probability*, McGraw-Hill Book Company, Inc., New York and London, 1937, pp. 204–205, problems 12–15), which yields a geometric rate for independent variables under conditions weaker than uniform boundedness. Moreover, for the case of independent x_n with $|x_n| \leqq 1$, Bernstein's rate is slightly better than that given here, having an expansion $r = 1 - (\epsilon^2/2), (\epsilon^3/6) + \cdots$ as compared with $r = 1 - (\epsilon^2/2) + (\epsilon^3/2) + \cdots$ for the rate given here.

REFERENCES

[1] W. Feller, *Probability Theory and Its Applications*, John Wiley and Sons, Inc., 1950.
[2] P. Lévy, *Theorie de l'addition des variables aléatoires*, Gauthier-Villars, Paris, 1937.