*Some Limit Theorems for Stochastic Processes*¹

JOHN LAMPERTI

Communicated by M. ROSENBLATT

1. Introduction. The purpose of this paper is to study the asymptotic behavior, for large values of t , of several random variables associated with certain types of stochastic processes. The simplest type of process considered is a "recurrent event" in discrete time. (See [7] or [8] for the definition and terminology of recurrent events; other terms will be defined precisely in Section 2 of this paper.) The random variables to be studied are Y_n , the time elapsed (measured at time n) since the last occurrence of the event, Z_n , which is the time until the next occurrence, and T_n , the time between occurrences. It turns out (Theorem 3.2) that the limiting distribution of Y_n/n must be a "generalized arc sine" law; the condition for the existence of this limit distribution is also a familiar one.²

A recent paper of DYNKIN [6] contains results very closely related to those of this paper. He studies sums of positive, independent, identically distributed random variables X_m , and finds the condition for $\xi(t)/t$ to have a limiting distribution where

$$(1.1) \quad \xi(t) = \min \{S_m \mid S_m = X_1 + \cdots + X_m \geq t\} - t.$$

Suppose that X_m is the time between the $(m - 1)^{\text{th}}$ and the m^{th} occurrence of a certain recurrent event E , so that $S_m = X_1 + \cdots + X_m$ is the time of the m^{th} occurrence. These random variables X_m are positive, independent and identically distributed, and $\xi(n)$ defined by (1.1) with these X 's can be identified with Z_n for the event E . Since (as DYNKIN points out; also see Section 5 of this paper) it is easy to pass from the distribution of Z_n/n to that of Y_n/n and *vice versa*,

¹ This paper, like [9], is based upon the author's California Institute of Technology Ph.D. thesis written under the direction of Professor F. SPITZER, whose assistance and inspiration are gratefully acknowledged.

² The generalized arc sine laws are the limiting distributions of the relative number of positive partial sums in the addition of independent, identically distributed random variables; this problem has been studied by ANDERSEN and, more completely, by SPITZER [12]. The condition referred to arises in studying the time during which a recurrent event occurs (FELLER, [7]).

DYNKIN's theorem contains our Theorem 3.2. However, our proof is different from DYNKIN's, and lends itself to numerous generalizations and applications. It is these extensions of the basic Theorem 3.2 to cases not falling within the scope of DYNKIN's work which, it is hoped, justify the present paper.

An outline of the results to follow may clarify our intentions. Section 2 contains notation and the statements of two Tauberian theorems needed later. The basic theorem for recurrent events (contained in the results of [6] as noted above) is proved in section 3 by a method, different from DYNKIN's, which has been used in [9] and is similar to a technique of SPITZER's [12]. In Section 4 it is shown that with the aid of a combinatorial theorem due to ANDERSEN, a theorem on the number of positive partial sums of random variables (see Footnote 2) can be deduced from Theorem 3.2. Then, in Paragraph 5, we give the results showing how the existence of a limiting distribution for Y_n/n leads to a distribution for Z_n/n and for $T_n/n = Y_n/n + Z_n/n$. These results are very similar to ones obtained and used by DYNKIN, and so the details are omitted.

Next, Theorem 3.2 is generalized by relaxing the condition that the event considered be precisely a recurrent event. This is somewhat in the spirit of some recent work by DARLING & KAC [3] on occupation times in Markov processes, and is carried out in Section 6. As an application, sums of stable random variables are treated. Finally, in Section 8 we consider the case of continuous time. The method adopted employs the previous results (especially those of Paragraph 6) rather than carrying through a parallel analysis. No attempt has been made to apply the main theorem (Theorem 8.1) in its maximum generality; rather, an interesting special class of cases has been treated, which includes linear Brownian motion and the Cauchy process. Other cases to which these methods apply will perhaps occur to the reader.

2. Preliminaries.

Definition 2.1. Let $X_n(\omega)$ be a stochastic process in discrete time, and A a measurable subset of its state space. Denote

- (i) $Y_n = n - \max \{i \leq n \mid X_i \in A\}$, (time since the last visit to A)
- (ii) $Z_n = \min \{i \geq n \mid X_i \in A\} - n$, (time until the next visit)
- (iii) $T_n = Y_n + Z_n$, (time between visits)
- (iv) $N_n = \sum_{i=1}^n \phi_A(X_i)$ (occupation time of A)

(ϕ_A is the characteristic function of the set A).

In this paper the $\{X_n\}$ will always be a Markov process with stationary transition probabilities. This includes the case of a recurrent event, as may be seen from the following lemma (which is both well known and simple):

Lemma 2.1. If E is a recurrent event, there is a denumerable Markov chain $\{X_n\}$ with non-negative integer states such that $X_n = 0$ if and only if E occurs at

time n . (In fact, if Y_n is defined directly as the time from the last occurrence of E to n , then $\{Y_n\}$ form such a Markov chain.)

The Tauberian theorems which will be used in the sequel are stated below, since they are needed in perhaps not their most familiar forms.

Theorem 2.1. (Corollary 2.2 of [9].) Let $P(x) = \sum_{n=1}^{\infty} p_n x^n$ where $p_n \geq 0$ and $\sum_{n=1}^{\infty} p_n = 1$. Then the following are equivalent conditions:

$$(A) \quad \lim_{x \rightarrow 1-} \frac{(1-x)P'(x)}{1-P(x)} = \alpha,$$

$$(B) \quad 1 - P(x) = (1-x)^{\alpha} L\left(\frac{1}{1-x}\right),$$

where $L(y)$ is slowly varying.³

Theorem 2.2. (A form of KARAMATA'S theorem; c.f. [13].) Suppose that $\sum_{n=0}^{\infty} c_n x^n$ is convergent for $|x| < 1$, and that for some positive integer k

$$\lim_{x \rightarrow 1-} (1-x)^k \sum_{n=0}^{\infty} c_n x^n = a \text{ exists.}$$

Suppose also that there exists a polynomial $Q(x)$ of degree at most $k-1$ such that $c_n \geq Q(n)$. Then

$$\lim_{n \rightarrow \infty} \frac{c_1 + c_2 + \cdots + c_n}{n^k} = \frac{a}{k!}.$$

(The stronger hypothesis $c_n \geq 0$ will not serve for our purposes.)

3. Y_n for Recurrent Events. Let E be a recurrent event; $E_n(E'_n)$ is the event that E occurs (does not occur) at time n . As in [7] we denote

$$f_n = \Pr(E'_{k+1}, E'_{k+2}, \dots, E'_{k+n-1}, E_{k+n} \mid E_k), \quad u_n = \Pr(E_{k+n} \mid E_k),$$

and put

$$F(x) = \sum_{n=1}^{\infty} f_n x^n, \quad U(x) = \sum_{n=0}^{\infty} u_n x^n.$$

It follows easily that

$$(3.1) \quad U(x) = U(x)F(x) + 1.$$

In case $\sum_n f_n = 1$ (E is "certain"), define also

$$t_n = \sum_{i=n+1}^{\infty} f_i; \quad T(x) = \sum_{n=0}^{\infty} t_n x^n = \frac{1-F(x)}{1-x}.$$

³ $L(y)$ is a slowly varying function if it is continuous and positive for large y , and $\lim_{y \rightarrow \infty} [L(cy)/L(y)] = 1$ for any $c > 0$. This definition is due to KARAMATA.

The first result following serves mainly to define the case of greatest interest.

Theorem 3.1. *If E is uncertain, then $Y_n/n \rightarrow 1$ with probability one, while if E is ergodic,*

$$(3.2) \quad \lim_{n \rightarrow \infty} \Pr(Y_n = k) = \frac{t_k}{M},$$

where M is the mean recurrence time for the event E .

Proof. The first statement follows from the fact that with probability one an uncertain event occurs only a finite number of times. Next, observe that the event $Y_n = k$ is equivalent to $E_{n-k}, E'_{n-k+1}, \dots, E'_n$. Hence

$$(3.3) \quad \Pr(Y_n = k) = u_{n-k} t_k.$$

But since E is ergodic, $u_n \rightarrow 1/M$ [8], which proves (3.2). Although it was not explicitly included in the theorem, the periodic case (with $M < \infty$) also offers no difficulty.

The interesting case, therefore, is the “null-recurrent” one, where $M = \infty$ but the event is still certain.

Theorem 3.2. *Suppose $\sum f_n = 1$ and $M = \sum n f_n = \infty$. In order that, for some distribution function $G(t)$,*

$$(3.4) \quad \lim_{n \rightarrow \infty} \Pr\left(\frac{Y_n}{n} \leq t\right) = G(t),$$

it is necessary and sufficient that the first moment converge; i.e.,

$$(3.5) \quad \lim_{n \rightarrow \infty} \mathbb{E}\left(\frac{Y_n}{n}\right) = \alpha \quad (0 \leq \alpha \leq 1).$$

If (3.5) holds, $G(t) = F_\alpha(t)$ where F_α is a distribution on $[0, 1]$ and, for $0 < t < 1$,

$$(3.6) \quad F_\alpha(t) = \frac{\sin \pi \alpha}{\pi} \int_0^t y^{\alpha-1} (1-y)^{-\alpha} dy \quad \text{if} \quad 0 < \alpha < 1,$$

$$F_0(t) = 1 \text{ and } F_1(t) = 0.$$

Finally, condition (3.5) is equivalent to

$$(3.7) \quad 1 - F(x) = (1-x)^{1-\alpha} L\left(\frac{1}{1-x}\right), \quad L \text{ slowly varying.}^4$$

Proof. The proof proceeds by operating with the double generating function of the quantities $\Pr(Y_n = k)$. First, by arguments of Abelian type we show that the generating functions of the moments of Y_n have the proper asymptotic behavior, and then KARAMATA's Tauberian theorem is used to show the actual convergence of the moments of Y_n/n .

⁴ Condition (3.7) is equivalent to FELLER's necessary and sufficient condition for the existence, under suitable normalization, of a non-degenerate limiting distribution for N_n [7].

Taking double generating functions in (3.3), we obtain

$$(3.8) \quad P(x, y) = \sum_{k,n} \Pr(Y_n = k) x^n y^k = U(x)T(xy).$$

Now observe (as does SPITZER in [12]) that

$$(3.9) \quad \begin{aligned} (1-x) \sum_{n=0}^{\infty} \mathbb{E}[\exp\{-\lambda(1-x)Y_n\}]x^n \\ = (1-x) \sum_n x^n \sum_k \frac{\mathbb{E}(Y_n^k)}{k!} [-\lambda(1-x)]^k = \sum_k \lambda^k f_k(x), \end{aligned}$$

where

$$f_k(x) = \frac{(1-x)^{k+1}}{k!} (-1)^k \sum_n \mathbb{E}(Y_n^k) x^n.$$

For $|x| < 1$, the quantity in (3.9) is an analytic function of x and of λ , if $\lambda \geq 0$. We shall see that the limit as $x \rightarrow 1-$ exists and is analytic at $\lambda = 0$, say $\sum b_n \lambda^n$, and so will conclude that $\lim_{x \rightarrow 1-} f_k(x) = b_k$.

Since, for all n , $0 \leq Y_n/n \leq 1$, (3.5) is obviously necessary for (3.4); assume (3.5) holds. But

$$P(x, y) = \sum_n \mathbb{E}(y^{Y_n}) x^n,$$

so that, using (3.8) and a form of Abel's theorem,

$$(3.10) \quad xU(x)T'(x) = \frac{\partial P}{\partial y} \Big|_{y=1} = \sum_n \mathbb{E}(Y_n) x^n \sim \frac{\alpha}{(1-x)^2}.$$

Expressed in terms of $F(x)$, (3.10) becomes

$$(1-x)F'(x) \sim (1-\alpha)[1-F(x)],$$

which by Theorem 2.1 is equivalent to (3.7). We shall see that (3.7) implies (3.4), and the equivalence of (3.7) and (3.5) will follow.

Applying (3.8) and (3.7) to (3.9), we obtain

$$\begin{aligned} \sum_k \lambda^k f_k(x) &= (1-x)P(x, e^{-\lambda(1-x)}) = \frac{1-x}{1-F(x)} \frac{1-F(xe^{-\lambda(1-x)})}{1-xe^{-\lambda(1-x)}} \\ &\sim (1-x)^\alpha L^{-1}\left(\frac{1}{1-x}\right) (1-xe^{-\lambda(1-x)})^{-\alpha} L\left(\frac{1}{1-xe^{-\lambda(1-x)}}\right). \end{aligned}$$

Taking account of the slowly varying property of L and noting that $1-xe^{-\lambda(1-x)} \sim (1-x)(1+\lambda)$ yields

$$(3.11) \quad \lim_{x \rightarrow 1-} \sum_k \lambda^k f_k(x) = (1+\lambda)^{-\alpha} = \sum_k \binom{-\alpha}{k} \lambda^k \quad \text{for } |\lambda| < 1.$$

⁵ Here and hereafter " $A(x) \sim B(x)$ " means $\lim_{x \rightarrow 1-} A(x)/B(x) = 1$.

Therefore

$$\lim_{x \rightarrow 1-} f_k(x) = \binom{-\alpha}{k},$$

or

$$(3.12) \quad (1-x) \sum_n \mathbb{E}(Y_n^k) x^n \sim (-1)^k k! \binom{-\alpha}{k} (1-x)^{-k}.$$

This completes the Abelian part of the argument.

In applying Theorem 2.2, we shall take the left member of (3.12) as $-\sum_n c_n x^n$. Then

$$\begin{aligned} -c_n &= \mathbb{E}(Y_n^k - Y_{n-1}^k) \\ &= \mathbb{E}[(Y_n - Y_{n-1})(Y_n^{k-1} + Y_n^{k-2} Y_{n-1} + \cdots + Y_{n-1}^{k-1})] \leq kn^{k-1} \end{aligned}$$

since $Y_{n-1}, Y_n \leq n$ and $Y_n - Y_{n-1} \leq 1$. Therefore $-kn^{k-1} = Q(n)$ will serve for the polynomial in the hypothesis of KARAMATA'S theorem (2.2), and we can conclude

$$(3.13) \quad \lim_{n \rightarrow \infty} \frac{-1}{n^k} \sum_{i=1}^n c_i = \lim_{n \rightarrow \infty} \mathbb{E} \left[\left(\frac{Y_n}{n} \right)^k \right] = (-1)^k \binom{-\alpha}{k}.$$

But it can be verified that the right hand side of (3.13) is the k^{th} moment of the law $F_\alpha(t)$ given by (3.6). Since Y_n/n always has values in a finite interval $[0, 1]$, the convergence of the moments implies the convergence of the distribution functions, and so the theorem is proved.

4. Application to the Number of Positive Partial Sums. Let $S_n = X_1 + \cdots + X_n$, where the X_i are independent, identically distributed random variables, and throughout this section let N_n denote the number of positive terms among $S_1, \dots, S_n$. We shall show that the limiting behavior of N_n can be deduced from Theorem 3.2 with the aid of the following combinatorial theorem by ANDERSEN [1]:

$$(4.1) \quad \Pr(N_n = k) = \Pr(N_k = k) \Pr(N_{n-k} = 0).$$

Lemma 4.1. *There exists a recurrent event E such that*

$$(4.2) \quad \Pr(N_n = k) = \Pr(Y_n = k).$$

Proof. A recurrent event is determined by specifying the sequence $\{u_n\}$, provided that the corresponding $\{f_n\}$ determined by (3.1) are non-negative and $\sum f_n \leq 1$. Let

$$(4.3) \quad u_n = \Pr(N_n = 0).$$

Then using (4.1) we have, for $|x| < 1$,

$$\begin{aligned} \sum_n \Pr(N_n = n)x^n \cdot \sum_n \Pr(N_n = 0)x^n &= \sum_n x^n \sum_{k=0}^n \Pr(N_k = k) \Pr(N_{n-k} = 0) \\ &= \sum_n x^n \sum_{k=0}^n \Pr(N_n = k) = \sum_n x^n = \frac{1}{1-x}. \end{aligned}$$

Therefore

$$\sum \Pr(N_n = n)x^n = \frac{1}{(1-x)U(x)}$$

must be identified with $T(x)$, so that $t_n = \Pr(N_n = n)$ and

$$f_n = \Pr(N_{n-1} = n-1) - \Pr(N_n = n)$$

which does have the required properties. For the event E defined by (4.3),

$$\Pr(Y_n = k) = u_{n-k}t_k = \Pr(N_{n-k} = 0) \Pr(N_k = k) = \Pr(N_n = k).$$

If a_n is defined to be $\Pr(S_n > 0)$,

$$\frac{a_1 + a_2 + \cdots + a_n}{n} = \mathbb{E}\left(\frac{N_n}{n}\right) = \mathbb{E}\left(\frac{Y_n}{n}\right).$$

Then from the lemma above and Theorem 3.2 we deduce

Theorem 4.1. (SPITZER [12].) *If $(a_1 + \cdots + a_n)/n \rightarrow \alpha$, then*

$$\lim_{n \rightarrow \infty} \Pr\left(\frac{N_n}{n} \leq x\right) = F_\alpha(x).$$

If $(a_1 + \cdots + a_n)/n$ does not tend to a limit, then neither does $\Pr(N_n/n \leq x)$.

5. The Random Variables Z_n and T_n . The results concerning Z_n and T_n are all based on the following identities:

$$(5.1) \quad \Pr(Z_n > m, Y_n > k) = \Pr(Y_{n+m} > m+k) = \Pr(Z_{n-k} > m+k).$$

The truth of (5.1) is obvious from Definition 2.1. DYNKIN [6] uses identities and obtains results very similar to the ones to be given here, and therefore, except for Lemma 5.1 as a sample, no proofs will be included.

Lemma 5.1. *If Y_n/n has a continuous limiting distribution $G(t)$, then*

$$(5.2) \quad \lim_{n \rightarrow \infty} \Pr\left(\frac{Z_n}{n} > u, \frac{Y_n}{n} > v\right) = 1 - G\left(\frac{u+v}{1+u}\right), \quad u \geq 0, \quad 1 \geq v \geq 0.$$

Proof. Choose two sequences $\{m_n\}$ and $\{k_n\}$ of non-negative integers with $k_n \leq n$ and such that

$$(5.3) \quad \lim_{n \rightarrow \infty} \frac{m_n}{n} = u, \quad \lim_{n \rightarrow \infty} \frac{k_n}{n} = v \leq 1.$$

Using the hypothesis of the lemma, (5.3), and (5.1) gives

$$\begin{aligned} 1 - G\left(\frac{u+v}{1+u}\right) &= \lim_{n \rightarrow \infty} \Pr\left(\frac{Y_{n+m_n}}{n+m_n} > \frac{u+v}{1+u}\right) = \lim_{n \rightarrow \infty} \Pr(Y_{n+m_n} > m_n + k_n) \\ &= \lim_{n \rightarrow \infty} \Pr(Z_n > m_n, Y_n > k_n) = \lim_{n \rightarrow \infty} \Pr\left(\frac{Z_n}{n} > u, \frac{Y_n}{n} > v\right). \end{aligned}$$

This lemma states, in other words, that Z_n/n and Y_n/n have a joint limit distribution if Y_n/n has a continuous one. The same is true of Z_n/n :

Lemma 5.2. *If Z_n/n has a continuous limiting distribution, then Z_n/n and Y_n/n have a joint limiting distribution.*

With the aid of these two lemmas, the theorem below follows almost at once:

Theorem 5.1. *The following are equivalent statements:*

$$(5.4) \quad \lim_{n \rightarrow \infty} \Pr\left(\frac{Y_n}{n} \leq t\right) = F_\alpha(t),$$

$$(5.5) \quad \lim_{n \rightarrow \infty} \Pr\left(\frac{Z_n}{n} \leq t\right) = F_\alpha\left(\frac{t}{1+t}\right) = \frac{\sin \pi \alpha}{\pi} \int_0^t \frac{\tau^{\alpha-1}}{1+\tau} d\tau.$$

It is easy also to deduce the limiting behavior of T_n/n :

Theorem 5.2. *If the equivalent conditions (5.4) and (5.5) hold, then*

$$(5.6) \quad \lim_{n \rightarrow \infty} \Pr\left(\frac{T_n}{n} \leq t\right) = \frac{\sin \pi \alpha}{\pi} \begin{cases} \int_0^t \tau^{\alpha-2} [1 - (1-\tau)^{1-\alpha}] d\tau, & 0 \leq t \leq 1 \\ \int_0^t \tau^{\alpha-2} d\tau, & t \geq 1 \end{cases}.$$

Remark. Nowhere has essential use been made of the discreteness of time; in fact, DYNKIN uses very similar arguments in a continuous time situation. Provided Y_t or Z_t (and hence both), defined in the manner analogous to Definition 2.1, are random variables, the results of this section apply to them. Theorem 5.1 especially will be needed in Section 8 below.

6. Extension to More General Processes. In this section X will denote an abstract set with a Borel field F of subsets, and $p(\xi, E)$ is supposed to be a Markov transition function when $\xi \in X$ and $E \in F$ (see [5] v, Section 5). That is, for each E in F , $p(\xi, E)$ is measurable in ξ , while, for each ξ , $p(\xi, E)$ is a probability measure on F . Given the initial condition $X_0 = \xi_0$, there is a Markov process $\{X_n\}$ satisfying $p(\xi, E) = \Pr(X_{n+1} \in E \mid X_n = \xi)$; the n -step transition probabilities will be denoted by $p_n(\xi, E)$. If A is a set belonging to F , then Y_n , Z_n , etc. are defined by Definition 2.1. Under suitable regularity assumptions, Theorem 3.2 will be extended to this more general situation. It will be assumed that A is visited infinitely often with probability one, if $X(0) = \xi_0 \in A$.

We will use the following notation, chosen to resemble that of Section 3 as closely as possible:

$$(6.1) \quad f_n(\xi, E) = \Pr (X_{m+n} \in E, X_{m+i} \notin A \text{ for } 1 \leq i < n \mid X_m = \xi),$$

$$t_n(\xi) = f_{n+1}(\xi, A) + f_{n+2}(\xi, A) + \cdots; \quad t_0(\xi) = 1, \text{ if } \xi \in A.$$

The generating functions are also needed:

$$(6.2) \quad U_x(\xi, E) = \sum_{n=0}^{\infty} p_n(\xi, E)x^n, \quad F_x(\xi, E) = \sum_{n=1}^{\infty} f_n(\xi, E)x^n,$$

$$T_x(\xi) = \sum_{n=0}^{\infty} t_n(\xi)x^n = \frac{1 - F_x(\xi, A)}{1 - x}, \quad \text{if } \xi \in A.$$

In order to prove Theorem 6.1 below, a rather strong assumption will be made concerning U_x and the set A .

Condition (U). *There is a measure $Q(E)$ defined for $E \in \mathcal{F}$ and $E \subset A$ such that $Q(A) = 1$, and for such sets E ,*

$$(6.3) \quad U_x(\xi, E) = Q(E)[h(x) + h_1(x, \xi, E)],$$

where $h(x) \rightarrow \infty$ as $x \rightarrow 1-$, and where, for E measurable,

$$(6.4) \quad \lim_{x \rightarrow 1-} \sup_{E \subset A} \left| \frac{h_1(x, \xi_0, E)}{h(x)} \right| = 0.$$

This condition says that the error, relative to the Q measure of the set E , in approximating $U_x(\xi_0, E)$ by $Q(E)h(x)$ becomes small uniformly with respect to E as $x \rightarrow 1-$. (U) is probably not the weakest possible condition, but makes for an easier proof of the theorem.

The following simple results will serve to illustrate the meaning of condition (U), and also will be useful later.

Lemma 6.1. *If $\{X_n\}$ is a denumerable, irreducible and recurrent Markov chain, and A is a finite set of states, then (U) holds.*

Proof. Let $\{q_i\}$ be the invariant measure of the chain [4], normalized so that $\sum_{i \in A} q_i = 1$; $U_{ii}(x)$ will denote the generating function of the transition probabilities $p_{ii}^{(n)}$. Let $h(x) = (1/q_1)U_{11}(x)$, so that

$$h_1(x, i, j) = \frac{1}{q_i} U_{ii}(x) - \frac{1}{q_1} U_{11}(x).$$

Now it is known [4] that

$$\frac{q_i}{q_j} = \lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n p_{ai}^{(k)}}{\sum_{k=1}^n p_{bj}^{(k)}} \quad \text{for any } a \text{ and } b.$$

With the aid of this result and Abel's theorem, it is easy to see that $h_1(x, i, j)/h(x) \rightarrow 0$ as $x \rightarrow 1-$. The uniformity required for (U) is automatic, since A is a finite set.

Lemma 6.2. *If (U) holds, we may choose $h(x) = U_x(\xi_0, A)$.*

Proof. (6.3) and (6.4) imply that

$$\lim_{x \rightarrow 1-} \frac{U_x(\xi_0, A)}{h(x)} = 1.$$

Some more notation is introduced, and then the theorem: Let

$$(6.5) \quad U(x) = U_x(\xi_0, A) = h(x), \quad F(x) = \int_A F_x(\xi, A) Q(d\xi),$$

and

$$T(x) = \int_A T_x(\xi) Q(d\xi) = \frac{1 - F(x)}{1 - x}.$$

Theorem 6.1. *Let the Markov process $\{X_n\}$ be as above, with (U) holding for the set A with respect to which Y_n is defined. Then in order that*

$$(6.6) \quad \lim \Pr \left(\frac{Y_n}{n} \leq x \right) = G(x)$$

exist, it is necessary and sufficient that

$$(6.7) \quad \lim E \left(\frac{Y_n}{n} \right) = \alpha$$

exist, which is in turn equivalent to

$$(6.8) \quad h(x) = (1 - x)^{\alpha-1} L \left(\frac{1}{1 - x} \right), \quad L \text{ slowly varying.}$$

If (6.7) holds, $G(x)$ must be equal to $F_\alpha(x)$, given in (3.6).

Proof. By exactly the argument of Section 3, (3.12) will imply the truth of (6.6) and establish that $G(x) = F_\alpha(x)$. Hence if it can be shown that

$$(6.9) \quad \lim_{x \rightarrow 1-} (1 - x) \sum_{n=0}^{\infty} E [e^{-\lambda(1-x)Y_n}] x^n = (1 + \lambda)^{-\alpha},$$

then the argument used in proving Theorem 3.2 will apply without change. (6.9) will be demonstrated with the aid of the following lemma and its corollary.

Lemma 6.3. *Let $P(x, y) = \sum_{n,k} \Pr(Y_n = k) x^n y^k$. Then*

$$(6.10) \quad P(x, y) \sim U(x)T(xy), \quad \left. \frac{\partial P}{\partial y} \right|_{y=1} \sim U(x)T'(x).$$

Proof. It is clear that

$$(6.11) \quad \Pr(Y_n = k) = \int_A p_{n-k}(\xi_0, d\xi) t_k(\xi).$$

Taking double generating functions yields

$$(6.12) \quad P(x, y) = \int_A U_x(\xi_0, d\xi) T_{xy}(\xi).$$

Using (6.3) and the definitions of $U(x)$ and $T(x)$,

$$\frac{P(x, y)}{U(x)T(xy)} = \frac{\int_A U(x)T_{xy}(\xi) Q(d\xi) + \int_A T_{xy}(\xi) Q(d\xi)h_1(x, \xi_0, d\xi)}{U(x)T(xy)}$$

so that

$$\begin{aligned} \left| \frac{P(x, y)}{U(x)T(xy)} - 1 \right| &= \frac{\int_A T_{xy}(\xi) Q(d\xi)h_1(x, \xi_0, d\xi)}{T(xy)U(x)} \\ &\leq \sup_{E \subset A} \frac{|h_1(x, \xi_0, E)|}{U(x)} \cdot \frac{\int_A T_{xy}(\xi) Q(d\xi)}{T(xy)} = \sup_{E \subset A} \frac{|h_1(x, \xi_0, E)|}{h(x)}, \end{aligned}$$

and so the first part of the lemma follows from (6.4). The second statement can be proved in the same way, by differentiating (6.12).

Corollary.

$$U(x) \sim \frac{1}{1 - F(x)}.$$

Proof. Setting $y = 1$ in the first part of (6.10) gives

$$\frac{1}{1 - x} = P(x, 1) \sim U(x)T(x) = U(x) \frac{1 - F(x)}{1 - x}.$$

Now it is easy to finish the proof of Theorem 6.1. From Lemma 6.3 and its corollary we conclude as in Section 3 that (6.7) (necessary for (6.6)) implies (6.8), and then the truth of (6.9) can be shown again by the same argument used in the proof of Theorem 3.2. Since the argument from (6.9) to (6.6) need not be changed at all, the proof is complete.

Remark. The proofs of Theorems 5.1 and 5.2 concerning limit distributions for Z_n/n and T_n/n used only the existence of a limiting distribution for Y_n/n , and so whenever Theorem 6.1 applies these results are also available.

7. Application to Sums of Stable Random Variables.⁶ Condition (U) is not too satisfactory; intuitively it seems to be stronger than necessary and it

⁶ [11], Section 22 contains a general discussion of the stable laws.

is certainly not very easy to apply. In this and the next section we shall treat important special classes of processes which do fall within the scope of Theorem 6.1. These results should be looked upon as examples only; no attempt has been made to exhaust the generality of the method.

Let X be a random variable such that

$$(7.1) \quad \log \{E(e^{i\lambda X})\} = -b |\lambda|^\beta \left\{ 1 + ic \frac{\lambda}{|\lambda|} \tan \frac{\pi}{2} \beta \right\}, \quad 1 < \beta \leq 2,$$

or else $\log \{E(e^{i\lambda X})\} = -b |\lambda|$ ($\beta = 1$). The stochastic process we shall consider is

$$(7.2) \quad S_n = X_1 + X_2 + \cdots + X_n$$

where the X_i are independent, identically distributed random variables satisfying (7.1). Let A be a bounded, measurable, non-null subset of the real line; we shall assume without loss of generality that $\mu(A) = 1$. The random variables Y_n , etc., will be defined with respect to this set A and the process S_n .

Theorem 7.1. *For the situation described above,*

$$(7.3) \quad \lim_{n \rightarrow \infty} \Pr \left(\frac{Y_n}{n} \leq x \right) = F_{1/\beta}(x).$$

Proof. First of all, it follows from (7.1) and the results of [2] that S_n lies infinitely often in any open interval, with probability one. (This would not be the case were $\beta < 1$ in (7.1).) Also from (7.1)

$$\int_{-\infty}^{\infty} |E(e^{i\lambda X})| d\lambda < \infty,$$

so that X has a continuous probability density $p(x)$ ([11], Section 12). These two facts together insure that A is visited infinitely often with probability one.

It remains to verify (U) and (6.8), and then the conclusion will follow from Theorem (6.1). S_n has a continuous density $p_n(x)$; from (7.1) if $\beta > 1$,

$$\begin{aligned} \int_{-\infty}^{\infty} e^{i\lambda x} p_n(x) dx &= \exp \left\{ -bn |\lambda|^\beta \left(1 + ic \frac{\lambda}{|\lambda|} \tan \frac{\pi}{2} \beta \right) \right\} \\ &= \int_{-\infty}^{\infty} e^{-\lambda n^{1/\beta} x} p(x) dx = \int_{-\infty}^{\infty} e^{i\lambda u} n^{-1/\beta} p(n^{-1/\beta} u) du, \end{aligned}$$

so that

$$(7.4) \quad p_n(x) = n^{-1/\beta} p(n^{-1/\beta} x);$$

this result also holds when $\beta = 1$. Therefore

$$(7.5) \quad U_x(y, E) = \int_E \sum_n x^n n^{-1/\beta} p(tn^{-1/\beta} - y) dt.$$

Now since $p(x)$ is continuous and A is bounded,

$$(7.6) \quad \lim_{n \rightarrow \infty} p(n^{-1/\beta} t) = p(0)$$

uniformly in t for $t \in A$. Therefore if $\beta > 1$,

$$(7.7) \quad U_x(0, A) \sim h(x) = p(0) \sum_n n^{-1/\beta} x^n \sim k(1-x)^{-1+1/\beta},$$

while if $\beta = 1$,

$$(7.8) \quad h(x) = p(0) \sum_n \frac{x^n}{n} \sim k \log \frac{1}{1-x}$$

where in both cases $k \neq 0$ is a constant, and so (6.8) holds. Furthermore, taking the measure Q of (U) as Lebesgue measure gives

$$(7.9) \quad h_1(x, 0, E) = \frac{U_x(0, E)}{Q(E)} - h(x) = \frac{1}{Q(E)} \int_E \sum_n x^n n^{-1/\beta} [p(tn^{-1/\beta}) - p(0)] dt.$$

But $|p(tn^{-1/\beta}) - p(0)|$ is bounded for all n and all $t \in A$, and by (7.6) is uniformly small if n is large; also $h(x)$ is unbounded as $x \rightarrow 1-$. Applying these facts to (7.9) it is easy to see that $|h_1(x, 0, E)| = o[h(x)]$ uniformly for any measurable set $E \subset A$, which completes the proof that (U) holds, and so also that of Theorem 7.1.

8. Extension to Continuous Time. Suppose that $\{X_t\}$ is a Markov process⁷ with real time-parameter, and that A is a measurable subset of its state space. Formally, we denote, as in Definition 2.1,

$$(8.1) \quad \begin{aligned} \text{(i)} \quad & Y_t = t - \sup \{ \tau \leq t \mid X_\tau \in A \} \\ \text{(ii)} \quad & Z_t = \inf \{ \tau \geq t \mid X_\tau \in A \} - t \\ \text{(iii)} \quad & T_t = Y_t + Z_t. \end{aligned}$$

A stochastic process is "measurable" if $X_t(\omega)$ is jointly measurable in the two variables t and ω ; for such a process, the quantities defined by (8.1) are random variables (see [5], II, Section 2).

To extend the previous results to this case, we shall consider also the following:

$$(8.2) \quad \begin{aligned} \text{(i)} \quad & Y'_t = t - \max \{ n \leq t \mid X_n \in A \} \\ \text{(ii)} \quad & Z'_t = \min \{ n \geq t \mid X_n \in A \} - t, \quad n = 0, 1, 2, \dots \end{aligned}$$

Under suitable conditions Theorem 6.1 can be applied to Y'_t and Z'_t . If, in addition, condition (R) below holds, then Y'_t and Y_t will differ by $O(1)$ with high probability, so that Y'_t/t and Y_t/t will have the same limiting distribution.

Condition (R). If $0 \leq t \leq 1$ and $\xi_0 \in A$,

$$(8.3) \quad \lim_{n \rightarrow \infty} \Pr(X_{t+i} \in A \text{ for some } i = 0, 1, 2, \dots, n \mid X_0 = \xi_0) = 1$$

uniformly in t and ξ_0 .

⁷ The Markov processes considered are always assumed to have stationary transition probabilities.

Theorem 8.1. *Let $\{X_t\}$ be a measurable Markov process satisfying condition (R), and suppose that $\{X_n\}$ ($n = 0, 1, 2, \dots$) satisfies condition (U). Then (6.8) is necessary and sufficient for the existence of a limiting distribution for Y_t/t . If (6.8) holds, then*

$$(8.4) \quad \lim_{t \rightarrow \infty} \Pr \left(\frac{Y_t}{t} \leq x \right) = F_\alpha(x).$$

Proof. Given $\epsilon > 0$, by (R) there is an N independent of t and ξ such that the probability in (8.3) exceeds $1 - \epsilon$. Hence

$$\Pr(Z_t \leq Z'_t \leq Z_t + N) \geq 1 - \epsilon.$$

Therefore either Z_t/t and Z'_t/t have the same limiting distribution, or both fail to have one at all. If $t = n + \theta$, $0 < \theta < 1$, then $Z'_t = Z'_{n+1} + (1 - \theta)$, and it is easy to see that Z'_t/t and Z'_n/n also behave in the same way. But since (U) holds, by Theorems 6.1 and 5.1 (6.8) is a necessary and sufficient condition that Z'_n/n have the limiting distribution (5.3). Hence (6.8) is the condition that Z_t/t satisfy (5.5), and, appealing this time to the continuous version of Theorem 5.1, (6.8) must also be necessary and sufficient for (8.4).

We shall now give two specific applications. The first will be stated without proof that (R) holds; this verification is simple but slightly tedious. For condition (U) we refer to Lemma 6.1.

Theorem 8.2. *Let A consist of a finite number of non-instantaneous states in a denumerable, recurrent, irreducible Markov chain. Then*

$$(8.5) \quad \sum_{n=0}^{\infty} \Pr(X_n \in A) x^n \sim (1 - x)^{\alpha-1} L\left(\frac{1}{1-x}\right)$$

is a necessary and sufficient condition that Y_t/t have a limiting distribution, which must be F_α if (8.5) holds. (Condition (8.5) could be replaced by an equivalent Laplace-transform condition which might look more natural for a continuous parameter chain.)

The final theorem is an analogue of Theorem 7.1. Let $\{X_t\}$ be a separable process with stationary independent increments⁸ such that (7.1) holds for $X = X_{t+1} - X_t$. Suppose $X_0 = 0$ and that A is a bounded measurable real set, $\mu(A) = 1$.

Theorem 8.3. *For such a process $\{X_t\}$,*

$$\lim_{t \rightarrow \infty} \Pr \left(\frac{Y_t}{t} \leq x \right) = F_{1/\beta}(x).$$

Proof. Since Theorem 7.1 has already been proved, it is only necessary to verify condition (R). We saw with the aid of [2] that the process S_n of (7.2)

⁸ [5], VIII and [10] give background on this type of process.

visits A infinitely often; this can be strengthened to give, for any $\epsilon > 0$, the existence of an N such that, uniformly for ξ in a finite interval $[-M, M]$,

$$(8.6) \quad \Pr(X_n + \xi \in A \text{ for some } n = 0, 1, 2, \dots, N) > (1 - \epsilon)^{\frac{1}{2}}.$$

In fact, because of the existence of continuous densities, the probability on the left side of (8.6) is a continuous function of ξ , and so the uniformity in (8.6) follows from the compactness of $[-M, M]$.

Now, just as in demonstrating (7.4), it is easy to show that X_t has continuous density $p_t(x) = t^{-1/\beta} p_1(t^{-1/\beta} x)$. Therefore if $0 < t \leq 1$,

$$\Pr(-B \leq X_t \leq B) = t^{-1/\beta} \int_{-B}^B p_1(t^{-1/\beta} x) dx \geq \Pr(-B \leq X_1 \leq B).$$

Suppose $A \subset [-C, C]$. Then for any $0 \leq t \leq 1$ and $\xi_0 \in A$,

$$(8.7) \quad \Pr(-(B+C) \leq X_t \leq B+C \mid X_0 = \xi_0) \geq \Pr(-B \leq X_1 \leq B \mid X_0 = 0).$$

Choose B so that the right hand side of 8.7 is not less than $(1 - \epsilon)^{\frac{1}{2}}$, and then let $M = B + C$ in (8.6); the conclusion is that with probability in excess of $1 - \epsilon$, $X_{t+i} \in A$ for some $i \leq N$ for any $X_0 = \xi_0 \in A$ and $0 \leq t \leq 1$, which is the same as (R).

Corollary 8.1. *If X_t is a Cauchy process, $Y_t/t \rightarrow 1$ and $Z_t/t \rightarrow \infty$ in probability.*

Corollary 8.2. *If X_t is linear Brownian motion, Y_t/t has the classical arc sine law for a limiting distribution.*

Remark. It is not difficult to prove Corollary 8.2 directly. This result and the corresponding statements for Z_t and T_t were proved by LÉVY [10] in the case where A is the single state zero, recurrent because of the continuity of the path functions. In this case, the distributions hold exactly for each finite t instead of appearing as limiting distributions for t large.

REFERENCES

- [1] E. S. ANDERSEN, On sums of symmetrically dependent random variables, *Skandinavisk Aktuarietidskrift*, 1953, pp. 123-138.
- [2] K. L. CHUNG & W. H. J. FUCHS, On the distribution of values of sums of random variables, *Mem. Amer. Math. Soc.* **6** (1951).
- [3] D. A. DARLING & M. KAC, On occupation times for Markov processes, *Trans. Amer. Math. Soc.* **84** (1957) pp. 444-458.
- [4] C. DERMAN, A solution to a set of fundamental equations in Markov chains, *Proc. Amer. Math. Soc.* **5** (1954) pp. 332-334.
- [5] J. L. DOOB, *Stochastic processes*, New York, 1953.
- [6] E. B. DYNKIN, Limit theorems for sums of independent random quantities, *Izvestiya Akad. Nauk S.S.S.R.* **19** (1955, in Russian) pp. 247-266.
- [7] W. FELLER, Fluctuation theory of recurrent events, *Trans. Amer. Math. Soc.* **67** (1949) pp. 98-119.
- [8] ———, *An introduction to probability theory and its applications*, New York, 1950.

- [9] J. LAMPERTI, An occupation time theorem for a class of stochastic processes, to appear in *Trans. Amer. Math. Soc.*
- [10] P. LÉVY, *Processus Stochastiques et Mouvement Brownien*, Paris, 1948.
- [11] M. LOÈVE, *Probability Theory*, New York, 1955.
- [12] F. SPITZER, A combinatorial lemma and its applications to probability theory, *Trans. Amer. Math. Soc.* **82** (1956) pp. 323–339.
- [13] D. V. WIDDER, *The Laplace Transform*, Princeton, 1946.

Stanford University
Stanford, California