ON THE ASYMPTOTIC BEHAVIOR OF RECURRENT AND "ALMOST RECURRENT" EVENTS.

Thesis by

John Williams Lamperti

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1957

ABSTRACT

In connection with a recurrent event E, random variables of interest include Y_n, the time since the last occurrence of E, N_n, the total time during which E occurs, and T_n, the time between occurrences. In this thesis theorems are given concerning the limiting distributions of these and other quantities as $n \to \infty$. It is shown that the same distributions apply in certain cases where E is not a recurrent event; this suggests the concept of an "almost recurrent event" which is shown to have the same asymptotic behavior as an "associated" recurrent event. These results extend and correlate previous work of Feller, Dynkin, Spitzer, and Darling and Kac. Finally, an occupation time theorem is proved for processes whose states comprise two classes separated by the occurrence of a recurrent event.

1. Introduction.

In this thesis we study the asymptotic behavior, for large values of the time parameter, of some random variables associated with certain types of stochastic processes. The simplest type of process considered is a "recurrent event". Intuitively, this means some repetitive pattern or event associated with a process such that whenever the event (say E) occurs, the process "starts over" as far as all future probabilities are concerned. More precisely, we state

<u>Definition 1.1</u>: Let X_n be the state at time n of a stochastic process in discrete time with state space X. Denote by $\omega = \{X_i\}$ a path function of the process. Let $\Gamma_n(\omega) = \Gamma_n(X_1, \ldots, X_n)$ be a measurable function of ω depending only on the first n coordinates and assuming the value 1 or 0. Suppose that

(i) $\{\omega \mid \Gamma_m(\omega) = 1, \Gamma_{m+n}(\omega) = 1\} = \{\omega \mid \Gamma_m(\omega) = 1, \Gamma_n(T^m\omega) = 1\}$

where $T\omega = T\{X_1, X_2, \ldots\} = \{X_2, X_3, \ldots\}$, and also that

(ii) $\Pr(\Gamma_i(\omega) = \gamma_i,\ i = m+1,\ldots,m+n \mid \Gamma_m(\omega) = 1,\ \Gamma_j(\omega) = \gamma_j,\ j=1,\ldots,m-1)$

$$= \Pr(\Gamma_i(\omega) = \gamma_i,\ i = m+1,\ldots,m+n \mid \Gamma_m(\omega) = 1)$$

where $\gamma_i = 0$ or 1.

Then the event E which occurs at time n if and only if $\Gamma_n(\omega) = 1$ is a recurrent event. (This is similar to Feller's definition in (7) but somewhat more general.)

When concerned with such an event, it is natural to study the random variables which measure the time during which the event occurs, and the elapsed time (measured at a fixed value of n) since E last occurred, until it next occurs, and between occurrences. We shall investigate these quantities in greater generality than for recurrent events however, and so give a definition which includes all the cases to be considered.

<u>Definition 1.2</u>: Let $\sigma_n(\omega)$ be a stochastic process with only the two points 0 and 1 in its state-space. Then we denote:

(i) $N_n(\omega) = \sum_{i=1}^{n} \sigma_i(\omega)$ (the occupation time of state 1)

(ii) $Y_n = n - \max_{i=1,\ldots,n} (i \mid \sigma_i = 1)$ (time since last visit to 1)

(iii) $Z_n = \min_{i=n,n+1,\ldots} (i \mid \sigma_i = 1) - n$ (time until next visit)

(iv) $T_n = Y_n + Z_n$ (time between visits).

It is clear that this includes the case of a recurrent event, for we have only to choose

(1.1) $\sigma_n = 1$ if E occurs at time n, $\sigma_n = 0$ otherwise.

More generally, processes σ_n arise naturally from any stochastic process X_n by choosing a fixed subset A of the state space and letting

(1.2) $\sigma_n = \phi_A(X_n)$; i.e., $\sigma_n = 1$ if $X_n \varepsilon A$, $\sigma_n = 0$ if $X_n \not\varepsilon A$.

In the theorems we shall derive concerning the limiting behavior of the random variables of definition 1.2, σ_n does arise in this way, where the underlying process X_n is a Markov process with stationary transition probabilities. The basic results concern the convergence, under suitable conditions, of the distribution of Y_n/n to a limit law which is one of the "generalized arc-sine" laws found by Spitzer (11) in the study of sums of random variables. From this, theorems can be derived concerning the quantities Z_n and T_n. These results both extend the theory of recurrent events, and suggest the concept of an "almost recurrent event" possessing many of the properties of recurrent events and generalizing them.

Dynkin (6) has studied the random variables Y_n, Z_n and T_n for recurrent events in continuous time, and has obtained a necessary and sufficient condition for limiting distributions to exist. His point of view is to consider the sum of independent, identically distributed, positive random variables, and to identify the events (E occurs for the kth time at time t) and $(S_k = t)$. In his approach the quantity Z_n is basic. In this thesis we proceed quite differently, by using the double generating function of $Pr(Y_n = k)$. From this the limiting behavior of Y_n may be derived, using essentially a method employed by Spitzer in (11). The condition obtained for a limit distribution appears different from Dynkin's; the equivalence of the two yields a theorem on slowly varying functions (this is the subject of sec. 4) which is very useful in the sequel. This approach is adaptable to generalization to the case of (1.2) when the set A contains more than one state.

In sec. 2, we consider the case where E is a recurrent event. The main theorem gives necessary and sufficient conditions for Y_n/n to have a limiting distribution. It is also shown that Spitzer's theorem mentioned above can be obtained as a corollary with the aid of a combinatorial theorem due to E.S. Andersen (1). The next section is devoted to generalizing the results of sec. 2; here the random variable Y_n is defined by definition 1.2 together with (1.2), where A is a suitable "small" subset of the state space of a Markov process. In this case, the same limit laws are found to hold as for recurrent events.

In sec. 5, we discuss the limiting behavior of N_n, Z_n and T_n. N_n has been quite extensively studied by, for instance, Feller (7) in the case of recurrent events, and Darling and Kac (2) for a more general situation. Here we observe with the aid of the theorem of sec. 4 that if a suitable regularity condition is satisfied, in particular in the case of a recurrent event, then the condition previously found for Y_n/n to have a limiting distribution is also the condition required for the distribution of N_n/c_n to converge to a non-degenerate limit for suitable constants c_n. The limiting behavior of Z_n and T_n, on the other hand, may be quite easily derived from that of Y_n.

Also in sec. 5, we study another random variable beside the ones already discussed. We note that the event $Y_n/n > t$ is the same as the occupation time of the interval $(n(1-t), n]$ being zero. Thus it is natural to ask more detailed questions about this quantity.

<u>Definition 1.3</u>: If σ_n is the same process as before, denote

$$N_n(t) = \sum_{i=1+[n(1-t)]}^{n} \sigma_i \quad \text{(where } [x] \text{ means the integer part of } x\text{).}$$

We obtain (in the case of recurrent events) a limit theorem for $N_n(t)$ containing as special cases previous results for Y_n and N_n.

Next, in sec. 6, the concept of an "almost recurrent event" is formalized. The previous results are shown to apply and yield that these events exhibit much the same limiting behavior as do recurrent events. Finally, in the concluding sec. 7, we proceed in a different direction, and consider a large set of states not even "almost recurrent", but having a recurrent event for a boundary. An occupation time theorem is proved for this case, in which limiting distributions of a new type appear.

2. The Random Variable Y_n for Recurrent Events.

Suppose that X_n is the state at time n of a denumerable Markov chain in discrete time; then we may suppose that X_n takes the values 1, 2, 3, ... We shall also assume that the chain is irreducible (i.e., that passage from state i to state j is possible eventually for every i and j) and impose the initial condition $X_o = 1$. Under these conditions the event E that the process is in state 1 is a recurrent event. Conversely, given a recurrent event E (such that E occurs at $t = 0$) we can construct a Markov chain satisfying the above conditions by letting X_n be the time elapsed since the last occurrence of E prior to time n; this is the simplest case of theorem 6.1 to be proved later. Definition 1.2 applies either upon choosing the set A to be the single state 1 and using (1.2), or by (1.1), giving a precise definition for Y_n, the time since the last visit to state 1. The purpose of this section is to study the limiting behavior of Y_n for large n, and to derive as an application a theorem due to Spitzer concerning sums of random variables.

We shall now give the rather trivial results which hold when E is an uncertain or an ergodic event. (The state 1 is transient or has finite mean recurrence time and is aperiodic.) It is convenient to first introduce some notation we shall use throughout this section. As in Feller (8) we let

$$f_k = \Pr(X_k = 1, X_i \neq 1 \text{ for } i = 1,2, \ldots, k-1 | X_o=1), \quad F(x) = \sum_{n=1}^{\infty} f_n x^n,$$

$$u_k = \Pr(X_k = 1 | X_o = 1), \qquad U(x) = \sum_{n=o}^{\infty} u_n x^n ;$$

then it may easily be shown that:

(2.1) $\qquad U(x) = U(x) F(x) + 1.$

We further define

(2.2) $\quad t_n = f_{n+1} + f_{n+2} + \cdots ; \quad$ then $\quad T(x) = \sum_{n=o}^{\infty} t_n x^n = (1 - F(x))/(1 - x).$

<u>Theorem 2.1</u>: If E is uncertain, then $Y_n/n \to 1$ with probability one, while if E is ergodic, we have

(2.3) $\qquad \lim_{n\to\infty} \Pr(Y_n = k) = t_k/u,$

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

<u>Proof</u>: The first statement follows from the fact that with probability one an uncertain event occurs only a finite number of times. Next we observe that the event $Y_n = k$ is equivalent to E occurs at time $n-k$ and then occurs next at least $k+1$ time-units later. Hence

(2.4) $\qquad \Pr(Y_n = k) = u_{n-k}\, t_k.$

But since E is ergodic, $u_n \to 1/u$ by a theorem in, for instance, (8) which proves (2.3). Although it was not included in the theorem, evidently the periodic case (with finite expected recurrence time) also offers no difficulty.

Of much more interest is the case of a certain event (recurrent state) with infinite expected recurrence time. Concerning this case, the main theorem of this section will now be proved.

Theorem 2.2: We have, with $G(t)$ a distribution function,

$$\lim_{n\to\infty} \Pr(Y_n/n < t) = G(t), \tag{2.5}$$

if and only if

$$\lim_{n\to\infty} E(Y_n/n) = \alpha \text{ exists} \qquad (0 \leq \alpha \leq 1). \tag{2.6}$$

If (2.6) holds, $G(t) = F_\alpha(t)$, where

$F_\alpha(t) = 0$ if $t \leq 0$, $F_\alpha(t) = 1$ if $t \geq 1$, and for $0 < t < 1$,

$$F_\alpha(t) = \frac{\sin \pi\alpha}{\pi} \int_0^t y^{\alpha-1} (1-y)^{-\alpha}\, dy \text{ if } 0 \neq \alpha \neq 1, \text{ with } F_0(t) = 1 \text{ and } F_1(t) = 0. \tag{2.7}$$

Finally, condition (2.6) is equivalent to the following:

$$\lim_{x\to 1-} (1 - x)\, F'(x)/(1 - F(x)) = 1 - \alpha\,. \tag{2.8}$$

Proof: The proof proceeds by operating with the double generating function of the $\Pr(Y_n = k)$'s. First, by Abelian arguments we show that the generating functions of the moments of Y_n have the proper asymptotic behavior; this part of the proof is almost identical with Spitzer's proof of his theorem 6.1 of (11). (See theorem 2.3 below.) Then the Karamata Tauberian theorem is used to show the actual convergence of the moments of Y_n/n.

Taking double generating functions in (2.4) we get

$$(2.9) \qquad P(x,y) = \sum_{k,n} \Pr(Y_n = k)\, x^n\, y^k = U(x)\, T(xy).$$

Now following Spitzer's argument, we observe that

$$(2.10) \quad (1-x) \sum_{n=o}^{\infty} E(\exp(-\lambda(1-x)Y_n))\, x^n = (1-x) \sum_n x^n \left[\sum_k \frac{E(Y_n^k)}{k!}(-\lambda(1-x))^k\right]$$

$$= \sum_k \lambda^k f_k(x), \quad \text{where} \quad f_k(x) = \frac{(1-x)^{k+1}}{k!} (-1)^k \left[\sum_n E(Y_n^k) x^n\right].$$

For $|x| < 1$, this is an analytic function of x and λ. We shall show that the limit as $x \to 1-$ exists and is an analytic function of λ, say $\sum b_n \lambda^n$, and hence conclude that the limit of $f_k(x)$ is b_k.

It is convenient to define a new sequence of numbers by putting

$$a_1 + a_2 + \cdots + a_n = E(Y_n).$$

Let

$$A(x) = \sum a_n x^n \;; \quad \text{then} \quad A(x)/(1-x) = \sum_n E(Y_n)\, x^n .$$

But

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

so that

$$\left.\frac{\partial P(x,y)}{\partial y}\right|_{y=1} = \sum_n E(Y_n)\, x^n = \frac{A(x)}{1-x} .$$

From (2.9), (2.1) and (2.2),

$$\left.\frac{\partial P}{\partial y}\right|_{y=1} = x\; U(x)\; T'(x) = \frac{x\; T'(x)}{(1-x)\; T(x)} ;$$

hence

$$\frac{T'(x)}{T(x)} = \frac{A(x)}{x} .$$

Thus we obtain

$$(2.11) \qquad T(x) = \exp\Big(\int_0^x \frac{A(y)}{y}\, dy\Big) = \exp\Big(\sum_{k=1}^{\infty} \frac{a_k}{k}\, x^k\Big) .$$

We therefore can also write

$$(2.12) \qquad U(x) = \frac{1}{1-x} \exp\Big(- \sum_{k=1}^{\infty} \frac{a_k}{k}\, x^k\Big) .$$

Using these expressions, we have

$$(2.13) \qquad (1-x) \sum_n E(\exp(-\lambda(1-x)Y_n))\, x^n = (1-x)\; P(x, \exp(-\lambda(1-x))$$

$$= (1-x)\; U(x)\; T(x \exp(-\lambda(1-x)) = \exp\Big[-\sum_k \frac{a_k}{k}\, x^k(1-\exp(-\lambda k(1-x))\Big] .$$

Next a lemma, also from (11), is required:

<u>Lemma 2.1</u>: Suppose $\dfrac{a_1 + a_2 + \cdots + a_n}{n} \to \alpha \geq 0$, and $\lambda \geq 0$. Then

$$\lim_{x \to 1-} \sum_{k=1}^{\infty} \frac{a_k}{k}\, x^k(1 - \exp(-\lambda(1-x)k)) = \alpha \log(1+\lambda).$$

Now by definition of the a_n's, if (2.6) holds, the hypothesis of the lemma is satisfied. Therefore combining (2.10), (2.13) and the lemma yields

$$\lim_{x\to 1-} \sum \lambda^k f_k(x) = (1+\lambda)^{-\alpha} = \sum_n \binom{-\alpha}{n} \lambda^n \quad \text{for} \quad |\lambda| < 1. \tag{2.14}$$

This implies that

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

or

$$(1-x)^k \left[(1-x) \sum_n E(Y_n^k)\, x^n\right] \to (-1)^k\, k!\, \binom{-\alpha}{k}. \tag{2.15}$$

This completes the Abelian part of the argument.

Next, we shall apply to (2.15) the following form of Karamata's Tauberian theorem (see, for instance, (12)):

<u>Theorem</u>: Suppose that $\sum_i c_i x^i$ is convergent for $x < 1$, and that

$$\lim_{x\to 1-} (1-x)^k \sum_i c_i x^i = a \quad \text{exists for some positive integer } k.$$

Suppose also that for some polynomial P of degree $\leq k-1$, we have

$$c_i \geq P(i) .$$

Then

$$\frac{c_1 + c_2 + \cdots + c_n}{n^k} \to \frac{a}{k!} \qquad \text{as} \quad n \to \infty .$$

(Note that the less general form of the theorem in which it is required that $c_i \geq 0$ would not suffice, since $E(Y_n^k)$ is not necessarily increasing.) We apply the theorem by putting

$$-\sum_n c_n x^n = (1-x) \sum_n E(Y_n^k)\, x^n ,$$

so that

$$-c_n = E(Y_n^k - Y_{n-1}^k) = E[(Y_n - Y_{n-1})(Y_n^{k-1} + Y_n^{k-2}\, Y_{n-1} + \cdots + Y_{n-1}^{k-1})] \leq k\, n^{k-1}$$

since $Y_n \leq n$ and $Y_{n+1} - Y_n \leq 1$. Therefore $c_n \geq k\, n^{k-1} = P(n)$. Hence we can conclude that as $n \to \infty$,

$$(2.16)\quad -\frac{1}{n^k}\sum_{i=1}^{n} c_i \to (-1)^k \binom{-\alpha}{k}, \quad \text{or} \quad E[(Y_n/n)^k] \to (-1)^k \binom{-\alpha}{k} .$$

But $(-1)^k \binom{-\alpha}{k}$ is the kth moment of the "generalized arc sin" law $F_\alpha(t)$ of order α, as may be verified from the expression (2.7). Since for all n, $0 \leq Y_n/n \leq 1$, the convergence of the moments implies the convergence of the distribution functions of Y_n/n to $F_\alpha(t)$. Thus the sufficiency part of the theorem is proved.

The necessity of (2.6) being obvious, it only remains to prove the equivalence with (2.8). But we have seen that

$$\sum_n E(Y_n)\, x^n = \frac{x}{(1-x)}\frac{T'(x)}{T(x)}$$

and by an Abelian and a Tauberian argument we can conclude that (2.6) holds if and only if

$$(2.17)\qquad \lim_{x\to 1-} \frac{(1-x)\, T'(x)}{T(x)} = \alpha .$$

But by (2.2), this is the same as (2.8). This completes the proof.

Now we shall show that theorem 2.2 is actually a generalization of Spitzer's theorem. Let $S_n = X_1 + X_2 + \cdots + X_n$, where the X_i are independent and identically distributed random variables. Define

$$(2.18)\qquad N_n = \text{no. of } S_i > 0, \quad i = 1, 2, \ldots, n.$$

In (1), Andersen proved that

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

With the aid of this result, we shall prove

<u>Lemma 2.2</u>: There exists a recurrent event E such that

$$Pr(N_n = k) = Pr(Y_n = k)$$

where N_n is as in (2.18) and Y_n refers to the event E.

<u>Proof</u>: A recurrent event is determined by the sequence $\{f_n\}$, provided that $f_n \geq 0$ and $\sum f_n \leq 1$. We shall see that these conditions are satisfied if we define u_n to be $Pr(N_n = 0)$. By (2.19), we have

$$\sum_n Pr(N_n = n)\ x^n \cdot \sum_n Pr(N_n = 0) x^n = \sum_n x^n \sum_{k=o}^{n} Pr(N_k = k) \cdot Pr(N_{n-k} = 0)$$

$$= \sum_n x^n \sum_{k=o}^{n} Pr(N_n = k) = \sum_n x^n = \frac{1}{1-x} .$$

Therefore

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

which we must take as $T(x)$; thus we must choose $t_n = Pr(N_n = n)$, yielding $f_n = Pr(N_{n-1} = n-1) - Pr(N_n = n)$ which has the required properties, so $\{f_n\}$ does define a recurrent event E. But by (2.19),

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

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

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

Hence by theorem 2.2 and lemma 2.2 we can deduce

<u>Theorem 2.3</u>: (Spitzer (11)) If $\frac{a_1 + a_2 + \cdots + a_n}{n} \to \alpha$, then

$$\lim_{n\to\infty} \Pr(N_n/n < t) = F_\alpha(t).$$

If $\frac{a_1 + \cdots + a_n}{n}$ does not tend to a limit, then neither does $\Pr(N_n/n < t)$.

<u>Example</u>. Let S_n be the "coin-tossing" process; that is, the random variables X_i are plus or minus one with probabilities 1/2. It is easy to see that $\frac{a_1 + a_2 + \cdots + a_n}{n} \to \frac{1}{2}$ and therefore by theorem 2.3 the fraction of the time that $S_n > 0$ has the limiting distribution $F_{1/2}(t)$; this was the first case of the theorem to have been discovered (Lévy (10)). But theorem 2.2 tells us more, for return to equilibrium $(S_n = 0)$ is a recurrent event, and it is easy to compute $F(x) = 1 - (1-x^2)^{1/2}$ which satisfies (2.8) with $\alpha = 1/2$, so that in the limit, the time since the last equalization also has the arc-sine law distribution $F_{1/2}$.

3. Y_n for More General Processes.

In this section we shall prove analogues of theorems 2.1 and 2.2 for a more general situation than that of sec. 2. We shall have as our state space X, an abstract set upon which a Borel field F of subsets is defined. We also suppose given a stochastic transition function $p(\xi, E)$ which is to define the probability of a transition in unit time from ξ into the set E; such a function must define, for each $\xi \in X$, a probability measure on the sets $E \in F$, while for each E it must be a measurable function of ξ. From this function, and the Markov property, the n-step transition probabilities $p_n(\xi, E)$ may be easily computed. It is well known that under these conditions, and given the initial condition $X_o = \xi_o$, there actually exists a Markov process $X_n(\omega)$ having the transition probabilities $p_n(\xi, E)$. (For discussion, see for instance (5), chapter V, no. 5, where will also be found other results to which we will refer.)

We will obtain results in the case in which the process is "recurrent"; more precisely, we assume the following condition holds:

Condition (C). A countably additive sigma-finite measure $m(E)$ is defined upon the field F, with the property that $m(E) > 0$ implies

$$\Pr(X_n \varepsilon E \text{ infinitely often} | X_o = \xi_o) = 1, \text{ for all } \xi_o \varepsilon X.$$

This condition, in the case of a denumerable Markov chain (which is included in the discussion of this section), is equivalent to the statement that the chain is irreducible and the states are all recurrent.

In (9), Harris has proved the following fundamental result:

Theorem: Let X_n be a Markov process in discrete time (as above) which satisfies (C). Then there exists a non-negative sigma-finite measure Q defined on the sets of F, unique up to a positive factor, such that $m(E) > 0$ implies $Q(E) > 0$, and satisfying

$$(3.1) \qquad Q(E) = \int_X p(\xi, E)\, Q(d\xi).$$

The measure Q is called the invariant measure of the process. If $Q(X)$ is finite, it may be normalized so that $Q(X) = 1$, and then by (3.1) defines a stationary probability distribution for the process; if $Q(X) = \infty$, no such stationary distribution exists. In case the Markov chain has a denumerable number of states, and is (by (C)) irreducible and recurrent, the cases $Q(X) = 1$ or ∞ correspond to finite or infinite mean recurrence times for the states. In the denumerable case, a stronger form of the above theorem is known (Derman (3) based on a result of Doeblin (4)), which states that the measure (now expressed by attaching the discrete mass q_i to the ith state) is given by the expression

$$(3.2) \qquad \frac{q_i}{q_j} = \lim_{n\to\infty} \frac{\sum_{l=1}^{n} p_{ci}^{(l)}}{\sum_{l=1}^{n} p_{dj}^{(l)}} \qquad \text{for any } c \text{ and } d,$$

where $p_{ab}^{(n)}$ is the transition probability of going from state a into state b in n-steps. An analogue of (3.2) for the non-denumerable case has been proved by Harris under conditions more restrictive than (C).

We shall now turn to the behavior of the random variable Y_n for large values of n. We choose a fixed set $A \varepsilon F$, and let Y_n be defined by (1.2) and definition 1.2. As in sec. 2, we denote

$$(3.3) \qquad \begin{aligned} f_n(\xi, E) &= \Pr(X_n \varepsilon E,\ X_i \not\varepsilon A \text{ for } i = 1,2, \ldots, n-1 \mid X_o = \xi) \text{ and} \\ t_n(\xi) &= f_{n+1}(\xi, A) + f_{n+2}(\xi, A) + \cdots . \end{aligned}$$

We also need to introduce the generating functions

$$(3.4) \qquad \begin{aligned} U_x(\xi, E) &= \sum_{n=o}^{\infty} p_n(\xi, E)\, x^n, \quad F_x(\xi, E) = \sum_{n=1}^{\infty} f_n(\xi, E)\, x^n, \text{ and} \\ T_x(\xi) &= \sum_{n=o}^{\infty} t_n(\xi)\, x^n = (1 - F_x(\xi, A))/(1-x) . \end{aligned}$$

Next, we shall prove a generalization of theorem 2.1, in the case where $Q(X) = 1$. We impose the following regularity condition:

Condition (D). There is a finite measure π on the sets of F (with $\pi(X) > 0$) such that for some positive integer v and positive number ε, $p^{(v)}(\xi, E) \leq 1 - \varepsilon$, if $\pi(E) \leq \varepsilon$.

Theorem 3.1: Suppose the Markov process X_n satisfies (C) and (D), has $Q(X) = 1$, and contains no cyclically moving subclasses of states. Then, for any initial condition $X_o = \xi_o$, we have

$$(3.5) \qquad \lim_{n\to\infty} \Pr(Y_n = k) = \int_A t_k(\xi)\, Q(d\xi).$$

Proof: Again the event $Y_n = k$ means that $X_{n-k} \varepsilon A$ and the next k values of $X_i \not\varepsilon A$. Therefore,

$$(3.6) \qquad \Pr(Y_n = k) = \int_A p_{n-k}(\xi_o, d\xi)\, t_k(\xi).$$

But under the hypotheses we have imposed, it follows from theorem 5.7 of (5), chapter V, that the limit of (3.6) is (3.5).

We note that in the case of theorem 2.1, the limiting distribution (2.3) depends very strongly on the particular form of the recurrence time distribution $\{f_n\}$, and that again the distribution (3.5) is determined in a complicated way by both the transition probabilities of the process and the "shape" of the set A. The case of infinite expected recurrence time, treated by theorem 2.2 (and generalized by theorem 3.2 below) is more pleasing, in that an easily characterized one-parameter class of limit laws emerges, to which the limit distribution of Y_n (if one exists) must belong.

In order to be able to prove the desired results, we shall need a regularity condition which insures that the set A is "sufficiently coherent". From here on we assume that condition (C) is satisfied, $X_o = \xi_o$, $Q(X) = \infty$, and that the measure is normalized so that $Q(A) = 1$; Y_n is defined with respect to A.

Condition (U). If E is a set belonging to F such that $Q(E) < \infty$, we can put

$$(3.7) \qquad U_x(\xi, E) = Q(E)\,[h(x) + h_1(x, \xi, E)]$$

where $h(x) \to \infty$ as $x \to 1-$, and where

$$(3.8) \qquad \lim_{x \to 1-} \sup_{E \subseteq A} |h_1(x, \xi_o, E)/h(x)| = 0.$$

This says that the error, relative to the "size" of the set E, in approximating $U_x(\xi_o, E)$ by $Q(E)\,h(x)$ becomes relatively small uniformly with respect to E for all sets $E \subseteq A$ for the fixed starting point ξ_o. (U) is not the weakest condition under which the proof of theorem 3.2 could be carried out, but to assume it avoids some analytical complication in the proof. It is not hard to see, however, that (3.8) could be weakened to uniform boundedness over all sets $E \subseteq A$ plus (3.8) holding when A has been replaced by a set A_n belonging to a denumerable family such that each $A_n \subseteq A$ and $Q(A - A_n) \to 0$.

The simplest circumstances under which (U) holds are given by:

Lemma 3.1: If the X_n process is a denumerable Markov chain, and A is a finite set of states, then (U) is satisfied.

Proof: Denote by $\{q_i\}$ the invariant measures of the states of the chain, and by $U_{ij}(x)$ the generating function of the $p_{ij}^{(n)}$'s. Then we may choose $U_{11}(x)/q_1$ as the $h(x)$ of (3.7). It follows that

$$h_1(x,i,j) = U_{ij}(x)/q_j - U_{11}(x)/q_1 .$$

Thus,

$$h_1(x,i,j)/h(x) = (q_1/q_j)\, U_{ij}(x)/U_{11}(x) - 1,$$

which approaches 0 as $x \to 1-$ by (3.2) plus an Abelian theorem. The uniformity with respect to subsets of A required by (U) is automatic since A consists of a finite number of states.

Analogously to our choice of $h(x)$ in the above lemma, we next state

Lemma 3.2: If (U) holds, $h(x)$ may be chosen as $U_x(\xi_0,A)$.

Proof: If (U) holds, (3.7) and (3.8) imply that

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

With this much preparation we shall now prove the main result of this section. First another notation is convenient:

(3.9) $$U(x) = U_x(\xi_0,A) = h(x), \quad \text{and} \quad F(x) = \int_A F_x(\xi,A)\, Q(d\xi).$$

Theorem 3.2: Let the Markov process X_n be as described above, with (U) holding for the set A with respect to which Y_n is defined. Then the statement of theorem 2.2 given on page 7 continues to be true.

Proof. It is easy to see that if we can show that (2.15) continues to be valid in the present case, then the Tauberian part of the proof of theorem 2.2 will also still hold without modification. We shall again use (2.10), and so the problem is to evaluate

$$(3.10)\qquad \lim_{x\to 1-} (1-x) \sum_{n=o}^{\infty} E(\exp(-\lambda(1-x)\,Y_n))\,x^n.$$

Introducing the further notation

$$(3.11)\qquad T(x) = \int_A T_x(\xi)\,Q(d\xi) = (1-F(x))/(1-x),$$

we proceed to evaluate (3.10) with the aid of two lemmas.

Lemma 3.3: As in sec. 2, let $P(x,y)$ be the double generating function $\sum_{n=o}^{\infty} E(y^{Y_n})\,x^n$. Then as $x \to 1-$,

$$(3.12)\qquad P(x,y) = U(x)\,T(xy)\,(1+o(1)),\ \text{and}\ \left.\frac{\partial P(x,y)}{\partial y}\right|_{y=1} = xU(x)T'(x)(1+o(1)).$$

Proof. Taking double generating functions in (3.6), we obtain

$$(3.13)\qquad P(x,y) = \sum_{n,k} \Pr(Y_n = k)\,x^n y^k = \int_A U_x(\xi_o, d\xi)\,T_{xy}(\xi)\,.$$

But by (3.7) and (3.9) we can write

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

so that

$$\left|\frac{P(x,y)}{U(x)\ T(xy)} - 1\right| = \frac{\left|\int_A T_{xy}(\xi)\ d_\xi[Q(\xi)\ h_1(x,\xi_o,\xi)]\right|}{T(x,y)\ U(x)}$$

$$\leq \sup_{E\subseteq A} \frac{|h_1(x,\xi_o,E)|}{U(x)} \cdot \frac{\int_A T_{xy}(\xi)\ Q(d\xi)}{T(x,y)}$$

$$= \sup_{E\subseteq A} \frac{|h_1(x,\xi_o,E)|}{U(x)},$$

and thus by (3.8) we obtain that

$$\lim_{x\to 1-} P(x,y)/U(x)\ T(xy) = 1.$$

The second statement of the lemma may be proved in exactly the same way from (3.13) and (U).

<u>Lemma 3.4</u>: $\lim_{x\to 1-} U(x)\ (1 - F(x)) = 1.$

<u>Proof</u>. Setting $y = 1$ in the first part of (3.12), we obtain

$$1/(1-x) = P(x,1) = U(x)\ T(x)\ (1 + o(1)) = U(x)\ (1 - F(x))(1 + o(1))/(1-x)$$

which proves the statement.

With the aid of these results, the procedure of sec. 2 may be imitated rather closely. We again define $\{a_n\}$ by means of the relation

$$a_1 + a_2 + \cdots + a_n = E(Y_n).$$

By lemma 3.3 we obtain in analogy to (2.11)

$$T(x) = \exp\left(\sum_{k=1}^{\infty} \frac{a_k}{k} x^k\right) \cdot (1 + o(1)), \tag{3.14}$$

while lemma 3.4 and (3.11) yield with (3.14)

$$(3.15) \qquad U(x) = \frac{1}{1-x} \exp\left(- \sum_{k=1}^{\infty} \frac{a_k}{k} x^k\right) \cdot (1 + o(1)) .$$

Combining these expressions with lemma 3.3, and using lemma 2.1, we find that (3.10) is equal to $(1+\lambda)^{-\alpha}$, provided that $\frac{a_1 + \cdots a_n}{n} \to \alpha$ (i.e., that $E(Y_n/n) \to \alpha$). Thus the proof goes through just as before. Finally, again with the aid of lemma 3.3, it may be verified that

$$E(Y_n/n) \to \alpha \quad \text{if and only if} \quad (1-x)\, F'(x)/(1 - F(x)) \to 1 - \alpha.$$

This completes the proof of the theorem.

4. A Theorem on Slowly Varying Functions.

In a recent paper (6), Dynkin also obtains a necessary and sufficient condition for (2.5) to hold; however, his condition, which is obtained by a quite different method from that of sec. 2, does not resemble (2.8). Comparison of the two suggested the theorem below; the present proof is due to H.F. Bohnenblust, and is both conceptually simpler and considerably more general than the original one.

Definition 4.1: (Karamata) A real function of a real variable, say $L(y)$, is slowly varying if it is continuous, positive for large enough y, and satisfies

$$(4.1) \qquad \lim_{y \to +\infty} \frac{L(cy)}{L(y)} = 1 \qquad \text{for all } c > 0.$$

<u>Theorem 4.1</u>: For $x > 0$, let $f(x)$ be a positive function. If $f'(x)$ exists for small enough x and satisfies

$$(4.2) \qquad \lim_{x\to 0} \frac{x\, f'(x)}{f(x)} = \alpha , \qquad \text{then we have}$$

$$(4.3) \qquad f(x) = x^{\alpha}\, L(\tfrac{1}{x}) \quad \text{where } L(y) \text{ is a slowly varying function.}$$

Conversely, if (4.3) holds for some number α, and if $f'(x)$ exists and is monotone in a neighborhood of 0, then (4.2) holds.

<u>Proof</u>: Assume (4.2), and define $L(y)$ by putting $f(x) = x^{\alpha} L(\frac{1}{x})$. Then $\log f(x) = \alpha \log x + \log L(\frac{1}{x}) = \alpha \log x + g(x)$, say. We must show that $L(y)$ satisfies (4.1); this is equivalent to

$$(4.4) \qquad \lim_{x\to 0} (g(cx) - g(x)) = 0 \quad \text{for all } c > 0.$$

But

$$g(cx) - g(x) = \log f(cx) - \log f(x) - \alpha \log c = \int_x^{cx} \frac{f'(t)}{f(t)}\, dt - \alpha \log c.$$

Hence by (4.2),

$$\lim_{x\to 0} (g(cx) - g(x)) = \lim_{x\to 0} \int_x^{cx} \frac{\alpha}{t}\, dt - \alpha \log c = 0.$$

To prove the converse, suppose $f'(x)$ is non-decreasing and that (4.3) holds. (Exactly the same argument works also for f' non-increasing.) By the mean value theorem, we have for $x < x_1$,

$$f'(\eta) = \frac{f(x_1) - f(x)}{x_1 - x} \quad \text{for some } x < \eta < x_1 .$$

Choosing $x_1 = cx$ where $c > 1$, and since $f'(x)$ is non-decreasing, we can write

$$f'(x) \leq f'(\eta) = \frac{f(cx) - f(x)}{x(c-1)} \leq f'(cx) \;, \quad \text{or}$$

$$\frac{x\, f'(x)}{f(x)} \leq \left(\frac{f(cx)}{f(x)} - 1\right) \frac{1}{c-1} \leq cx\, \frac{f'(cx)}{f(cx)} \cdot \frac{1}{c}\, \frac{f(cx)}{f(x)}$$

Now by (4.3) and (4.1) we obtain upon letting $x \to 0$,

$$\overline{\lim_{x\to 0}}\, \frac{x\, f'(x)}{f(x)} \leq (c^{\alpha} - 1)\, \frac{1}{c-1} \leq c^{\alpha-1} \underline{\lim_{x\to 0}}\, \frac{x\, f'(x)}{f(x)} \;.$$

But this holds for all $c > 1$; letting $c \to 1$ we get precisely (4.2).

In the case (as in the original form of the theorem) that $f(x) = 1 - \int_0^\infty e^{-tx}\, dG(t)$ where $G(t)$ is the distribution function of a positive random variable, it follows that $f''(x) = - \int_0^\infty t^2\, e^{-tx}\, dG(x) < 0$ so that $f'(x)$ is a decreasing function of x. Thus we have:

<u>Corollary 4.1</u>: Let $\phi(x) = \int_0^\infty e^{-tx}\, dG(t)$. Then the following are equivalent:

A) $\lim_{x\to 0} \dfrac{x\, \phi'(x)}{1 - \phi(x)} = - \alpha$

B) $1 - \phi(x) = x^{\alpha}\, L(\frac{1}{x})$ where $L(y)$ is a slowly varying function.

Stating this in terms of generating functions, we have:

<u>Corollary 4.2</u>: 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:

A) $\lim\limits_{x \to 1-} \dfrac{(1-x)\, P'(x)}{1-P(x)} = \alpha$

B) $1 - P(x) = (1-x)^{\alpha}\, L(\frac{1}{1-x})$ where $L(y)$ is slowly varying.

Applying corollary 4.2 to the situation of theorem 2.2 yields

Corollary 4.3: Condition (2.8), which is necessary and sufficient for (2.5), is itself equivalent to:

$$1 - F(x) = (1-x)^{1-\alpha}\, L(\tfrac{1}{1-x}) \quad \text{where } 0 \leq \alpha \leq 1 \text{ and } L(y) \text{ is}$$

a slowly varying function.

5. Limit Theorems for N_n, Z_n, T_n and $N_n(t)$.

In the case of a recurrent event with infinite expected recurrence time, the condition for the existence of

$$(5.1) \qquad V(t) = \lim_{n \to \infty} \Pr(N_n < c_n\, t),$$

where $V(t)$ is a non-degenerate distribution function and the c_n are appropriate norming constants, has been known for some time (see (7) theorem 6) to be

$$(5.2) \qquad 1 - G(t) \sim t^{\alpha-1}\, L(t) \qquad \text{as } t \to \infty,$$

where $G(t)$ is the distribution function of the recurrence times, $0 < \alpha \leq 1$, and $L(t)$ is a slowly varying function. If (5.2) holds, $V(t)$ turns out to be the Mittag-Leffler distribution $M_{1-\alpha}(t)$ of index $1-\alpha$, given by

$$M_\beta(t) = \frac{1}{\pi\beta} \int_0^t \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j!} \sin(\pi\beta j)\, \Gamma(\beta j + 1)\, \tau^{j-1}\, d\tau \,. \tag{5.3}$$

Letting, as in sec. 2, $F(x)$ denote the generating function of the recurrence time distribution, we find (using Karamata's and Abel's theorems, respectively) that (5.2) holds if and only if

$$1 - F(x) = (1-x)^{1-\alpha}\, L\left(\frac{1}{1-x}\right) , \tag{5.4}$$

with $L(y)$ slowly varying. But by corollary 4.3 and theorem 2.2, (5.4) is the condition under which the limit distribution $F_\alpha(t)$ holds for Y_n/n .

Recent work of Darling and Kac (2) has shown that essentially the above situation for N_n holds under more general conditions than for recurrent events. We consider again the Markov process X_n of sec. 3, and define the random variables with respect to a set A of invariant measure one. Sufficient for our purpose is the following:

<u>Condition (U')</u>. (U) holds for all $\xi_0 \in A$, and the convergence in (3.8) is uniform over A with respect to ξ_0.

We summarize the relationship of N_n and Y_n in

<u>Theorem 5.1</u>: Suppose the Markov process X_n and the set A satisfy (U'). If

$$h(x) = (1-x)^{\alpha-1}\, L\left(\frac{1}{1-x}\right) , \tag{5.5}$$

where L is slowly varying and $0 \le \alpha \le 1$, then Y_n/n has the limiting distribution $F_\alpha(t)$ (given by (2.7)) while, provided $\alpha \neq 0$, N_n suitably normalized is distributed in the limit by $M_{1-\alpha}(t)$. If (5.5) does not hold, then $Pr(Y_n/n < t)$ does not have a limit, and (as is also the case if $\alpha = 0$) there is no normalization for which N_n has a non-degenerate limit distribution.

Proof: The statements about N_n follow from (2); condition U' is somewhat stronger than necessary, but allows a unified statement of results. The statements about Y_n are seen to be a paraphrase of theorem 3.2 when it is observed, with the aid of lemmas 3.2 and 3.4 and corollary 4.3, that (5.5) is equivalent to (2.8).

The study of Z_n and T_n will be based upon the identity

$$(5.6) \qquad Pr(Z_n > m,\ Y_n > k) = Pr(Y_{n+m} > m+k).$$

The truth of (5.6) is obvious from definition 1.2. This identity is analogous to ones used by Dynkin in (6) for a similar purpose; as previously commented, in his approach Z_n rather than Y_n is the known quantity.

Lemma 5.1: Under the conditions of theorem 3.2, Z_n and Y_n have a joint limiting distribution if and only if Y_n/n has a limiting distribution. If (2.8) or equivalently (5.5) does hold, then

$$(5.7) \qquad \lim_{n\to\infty} Pr(Z_n/m > u,\ Y_n/n > v) = 1 - F_\alpha\left(\frac{u+v}{1+u}\right) \quad \text{for} \quad u \geq 0.$$

Proof: Suppose that (5.5) holds, and choose two sequences $\{m_n\}$ and $\{k_n\}$ of non-negative integers with $k_n \leq n$ and satisfying

$$(5.8) \qquad \lim_{n\to\infty} m_n/n = u, \qquad \lim_{n\to\infty} k_n/n = v \leq 1.$$

Now since $F_\alpha(t)$ is (for $\alpha \neq 0$ or 1) a continuous distribution function, we have from theorem 3.2, (5.6) and (5.8)

$$1 - F_\alpha\left(\frac{u+v}{1+u}\right) = \lim_{n\to\infty} \Pr(Y_n/n > (u+v)/(1+u)) = \lim_{n\to\infty} \Pr(Y_{n+m_n} > m_n + k_n)$$

$$= \lim_{n\to\infty} \Pr(Z_n > m_n, Y_n > k_n) = \lim_{n\to\infty} \Pr(Z_n/n > u, Y_n/n > v).$$

The converse is clear, as is the situation for the trivial cases when $\alpha = 1$ or 0.

The lemma immediately yields as an application

<u>Theorem 5.2</u>: Under the conditions of theorem 3.2, if (5.5) holds then

$$(5.9) \qquad \lim_{n\to\infty} \Pr(Z_n/n \leq t) = F_\alpha\left(\frac{t}{1+t}\right) = \frac{\sin \pi\alpha}{\pi} \int_0^t \frac{\tau^{\alpha-1}}{1+\tau}\, d\tau .$$

We can also use Lemma 5.1 to obtain a result on T_n.

<u>Theorem 5.3</u>: Under the conditions of theorem 3.2, if (5.5) holds then

$$(5.10) \qquad \lim_{n\to\infty} \Pr(T_n/n \leq t) = \frac{\sin \pi\alpha}{\pi} \int_0^t \tau^{\alpha-2}(1 - (1-\tau)^{1-\alpha})\, d\tau \quad \text{for } 0\leq t\leq 1,$$

$$= \frac{\sin \pi\alpha}{\pi} \int_0^t \tau^{\alpha-2}\, d\tau \qquad \text{for} \quad t > 1.$$

<u>Proof</u>: From lemma 5.1 we can compute the joint limiting density of $(Z_n/n, Y_n/n)$, obtaining

$$(5.11) \qquad p_\alpha(u,v) = (1-\alpha)\, \frac{\sin \pi\alpha}{\pi}\, (1-v)^{-\alpha}\, (u+v)^{\alpha-2}$$

$$\text{for} \quad 0 \leq u,\ 0 \leq v \leq 1.$$

But since $T_n = Z_n + Y_n$ we have for the density of T_n/n

$$t_\alpha(t) = \int_0^{\min(t,1)} p_\alpha(t-v, v)\, dv;$$

upon substituting (5.11) we obtain (5.10).

Finally we shall prove some results about $N_n(t)$ (definition 1.3) which provide an extension of theorem 2.2. For simplicity we confine the discussion to the situation of sec. 2, where the set A is a single state in a denumerable Markov chain. Now the event $Y_n/n < t$ is the same as the event $N_n(t) > 0$. Since we have already found in theorem 2.2 conditions for the probability of this event to have a limit as $n \to \infty$, it is natural to ask if any further information can be obtained about the limiting behavior of $N_n(t)$. First we shall show that $N_n(t)$ itself does not remain of moderate size, but either is zero or grows arbitrarily large.

<u>Theorem 5.4</u>: We have, provided the left hand side exists,

$$(5.12) \qquad \lim_{n\to\infty} \Pr(N_n(t) > 0) = \lim_{n\to\infty} \Pr(N_n(t) > k) \quad \text{for any} \quad k \geq 0.$$

<u>Proof</u>: Since we are dealing with a recurrent event, S_n, the time at which the event occurs for the nth time, is the sum of n positive independent random variables (the waiting times) with generating function $F(x)$.
We can write

$$(5.13) \qquad \Pr(N_n(t) > k) = \sum_{r=0}^{[nt]} \Pr(Z_{[n(1-t)]+1} = r)\, \Pr(S_k \leq nt - r).$$

But since the limit on the left in (5.12) exists, by theorem 2.2 condition (2.8) must hold, and so from theorem 5.2 the limiting distribution $F_\alpha(\frac{t}{1+t})$ holds for Z_n/n. Also, since k is fixed

$$\lim_{n\to\infty} \Pr(S_k \leq nt - r) = 1 \qquad \text{for } r/n \leq t-\varepsilon \ , \ \varepsilon > 0.$$

Therefor on taking the limit in (5.13) as $n \to \infty$ we get letting $\frac{r}{n(1-t)} \to \tau$,

$$\lim_{n\to\infty} \Pr(N_n(t) > k) = \int_0^{t/1-t} 1 \, dF_\alpha(\tfrac{\tau}{1+\tau}) = F_\alpha(t)$$

which proves the theorem.

It is possible to obtain more information about $N_n(t)$; we shall now show in conclusion of this section that the existence of the limit in (5.12) is sufficient for $N_n(t)$, when suitably normalized, to have a limit distribution which is non-degenerate provided $\alpha \neq 0$ or 1.

Theorem 5.5 : Suppose that (2.8), and hence (5.5), holds for the event E. Then for a suitable sequence $\{c_n\}$ of positive numbers, and $u \geq 0$,

$$(5.14) \qquad \lim_{n\to\infty} \Pr(N_n(t) > c_n u) = \int_0^{t/1-t} G_{1-\alpha}(u^{1/\alpha-1}[t-\tau(1-t)]) \, dF_\alpha(\tfrac{\tau}{1+\tau}),$$

where $G_{1-\alpha}$ is the distribution of the positive stable law of index $1-\alpha$ ((7), sec. 5).

Proof. Our procedure is based upon (5.13). As in the proof of theorem 5.4, the right hand side will, if (2.8) holds, converge to an integral over the range 0 to $t/(1-t)$ with respect to $F_\alpha(\frac{\tau}{1+\tau})$. The problem is to choose the c_n so that

$$\Pr(S_{[c_n u]} \le nt - r) \to \text{function of } t, u, \tau .$$

The possibility of doing this is based on a lemma proved in (7).

<u>Lemma</u>. If (5.2) holds (equivalent as we have seen to (2.8)) then

$$\lim_{n\to\infty} \Pr(S_n \le v b_n) = G_{1-\alpha}(v) , \tag{5.15}$$

where the b_n are a sequence satisfying

$$\lim_{n\to\infty} n(1 - G(b_n)) = 1. \tag{5.16}$$

We next see that it is possible to choose $f(u)$ and c_n so that for each A,

$$\lim_{n\to\infty} \Pr(S_{[c_n u]} \le A\ n) = \lim_{n\to\infty} \Pr(S_{[c_n]} \le Anf(u)) = G_{1-\alpha}(Af(u)). \tag{5.17}$$

For, by the lemma, the second term equals the third provided

$$\lim_{n\to\infty} [c_n]\ (1 - G(n)) = 1,$$

while in order for the first to equal the third we need

$$\lim_{n\to\infty} [c_n u](1 - G(n/f(u)) = 1.$$

Combining these two expressions, with the aid of (5.2) and the definition of a slowly varying function, we find that the choice

$$f(u) = u^{1/\alpha-1} ; \qquad c_n = n^{1-\alpha}/L(n) \tag{5.18}$$

is sufficient for (5.17) to hold.

In the passage to the limit we are letting $\frac{r}{n(1-t)} \to \tau$; therefore we substitute $A = t - \tau(1-t)$ in (5.17) and are now able to complete the limiting process. Since (5.17) holds we obtain

$$\lim_{n\to\infty} \Pr(N_n(t) \geq c_n u) = \int_0^{t/1-t} G_{1-\alpha}(f(u)\ (t-\tau(1-t))\ dF_\alpha(\tfrac{\tau}{1+\tau}) ,$$

which upon making the choice (5.18) is the same as (5.14) and proves the theorem. Note that in order to obtain the complete limit distribution of $N_n(t)$ it is also necessary to attach mass $1 - F_\alpha(t)$ to the point zero.

6. Almost Recurrent Events.

In this section we shall formalize the concept of an "almost recurrent event". To do so we consider processes $X_n(\omega)$ which are not necessarily Markovian; the basic idea is that the proof of theorems 2.1 and 2.2 and their generalizations in sec. 3 do not really depend on the Markov property for all the states of the process but only on the behavior of the process at the state 1 or set A. If this behavior has certain properties we can construct a new process which is Markovian and in which the set A has been "embedded".

We will start by giving another definition of recurrent event which is easily seen to be equivalent to definition 1.1 although apparently stated in a more special form.

Definition 6.1: Let $X_n(\omega)$ be a not necessarily Markovian stochastic process with stationary transition probabilities and state-space X. Let e be a state of X and let the event E be equivalent to the occupation of state e by the process. Assume

$$(6.1)\quad \Pr(\phi_e(X_{m+i}) = \gamma_{m+i},\ i = 1, \dots , n \mid X_m = e)$$

$$= \Pr(\phi_e(X_{m+i}) = \gamma_{m+i},\ i = 1,\dots,n \mid X_m = e,\ X_j = \xi_j \in X,\ j = 1,\dots,m-1)$$

where $\gamma_i = 0$ or 1 and ϕ_e is the characteristic function of $\{e\}$. Then E is a recurrent event.

Note that (6.1) essentially says that the process X_n is Markovian at the particular state e although it need not be elsewhere. If Y_n is defined with respect to the state e, then Y_n is a denumerable Markov chain with occupation of the state $Y_n = 0$ equivalent to the occurrence of E.

We will now turn to the more general case and give analogues of these facts. Again let $X_n(\omega)$ be a process with stationary transition probabilities, and let A be a fixed subset of the state space X; we define E to be the event that $X_n \in A$. We shall formulate a definition with the aid of several postulates on the process.

<u>P.1</u>) If $\xi \in A$, $\xi_i \in X$, and E_i is either a measurable subset of A or the complement of A, then

$$\Pr(X_{m+i} \in E_i,\ i = 1, \dots , n \mid X_m = \xi)$$

$$= \Pr(X_{m+i} \in E_i,\ i = 1, \dots , n \mid X_m = \xi,\ X_j = \xi_j,\ j = 1, \dots , m-1).$$

This is the analogue of (6.1) and requires the process to be "Markovian in the set A".

It will be convenient to speak of the "embedded process" $X'_n(\omega)$ which is obtained by observing the X_n process only at those times when E occurs; thus $X'_n(\omega) = X_N(\omega)$, where N is the time of the nth occurrence of E. The states of the embedded process are the points of A. Of course if A is not visited infinitely often the X'_n process can be considered defined for all n only by adding an extra state in which the process is absorbed after its last visit to A. If this is done the embedded process will, by P.1, always be a Markov process. The cases of interest here, however, will be those in which A is visited infinitely often with probability 1, and in which the embedded process itself has suitable recurrence properties; we insure this by

P.2) The embedded process X'_n satisfies condition (C) of sec. 3.

We now introduce again the notations

$$Pr(X_n \varepsilon E \subseteq X | X_o = \xi) = p_n(\xi, E); \quad \sum_{n=o}^{\infty} p_n(\xi, E)\, x^n = U_x(\xi, E).$$

Now by P.2 and the theorem of Harris quoted in sec. 3, there exists a measure $Q(E)$ defined on subsets E of A which is the invariant measure for the embedded process. Making use of this measure we further postulate an analogue of condition (U):

P.3) $Q(A) = 1$, and for $\xi, \xi_o \varepsilon A$, $E \subseteq A$ we have

$$U_x(\xi, E) = Q(E)\,(h(x) + h_1(x, \xi, E) \qquad \text{where as } x \to 1- ,$$

$$h(x) \to \infty , \quad \text{and} \quad \sup_{E \subseteq A} \left| \frac{h_1(x, \xi_o, E)}{h(x)} \right| \to 0 .$$

We may similarly give an equivalent to (U'):

P.3') P.3 holds for all $\xi_o \varepsilon A$, and the convergence is uniform in ξ_o over A.

Now we can state the following:

Definition 6.2: If an event E satisfies P.1, P.2 and P.3 or P.3', then E is an "almost recurrent event".

The main motivation for this definition is that almost recurrent events have many of the same limit theorems holding as do recurrent events. We shall show this with the aid of the theorems of the preceding sections, plus the following result which makes it possible to apply the theorems we have previously obtained.

Theorem 6.1: Let the process $X_n(\omega)$ and the set A satisfy P.1. Then there exists a (non-unique) Markov process M_n whose state space contains a subset A' in one to one correspondence $\xi' \varepsilon A' \longleftrightarrow \xi \varepsilon A$ with A such that the transition probabilities $q_n(\xi', E')$ of M_n, for $\xi' \varepsilon A'$ and $E' \subseteq A'$, satisfy

$$(6.2) \qquad q_n(\xi', E') = p_n(\xi, E) \quad \text{for all } n.$$

Furthermore, if P.2 holds for X_n, then (C) holds for M_n.

Proof: The states of M_n may be taken to be the pairs (ξ, n) where $\xi \varepsilon A$ and n is a non-negative integer. The event that $M_n = (\xi, m)$ is taken to mean that $X_{n-m} = \xi \varepsilon A$, and that $X_i \not\varepsilon A$ for $i = n-m+1, \ldots, n$. For the X_n process we again define

$$f_n(\xi, E) = \Pr(X_n \varepsilon E \subseteq X,\ X_i \not\varepsilon A \text{ for } i=1,2,\ldots,n-1 \mid X_o = \xi).$$

We choose for the M_n process the one-step transition probabilities

$$q_1[(\xi, 0), (E, 0)] = f_1(\xi, E), \qquad q_1[(\xi, 0), (E, 1)] = \phi_E(\xi)(1 - f_1(\xi, A)),$$

$$(6.3) \ q_1[(\xi, n), (E, 0)] = \frac{f_{n+1}(\xi, E)}{\sum_{i=n+1}^{\infty} f_i(\xi, A)},$$

$$(6.3)\quad q_1[(\xi,n),(E,n+1)] = \phi_E(\xi)\,\frac{\sum\limits_{i=n+2}^{\infty} f_i(\xi,A)}{\sum\limits_{i=n+1}^{\infty} f_i(\xi,A)} \;;$$

all other transitions have probability 0. Since M_n is assumed to be a Markov process, the other transition probabilities are determined by these. Let $A' = \{(\xi,0)\}$. Then it is easy to verify that (6.3) do indeed from a Markov transition operator, and that under the natural mapping of A' onto A (6.2) holds. This construction is the natural generalization of the way in which, for A a single state (E a recurrent event), Y_n provides an "equivalent" Markov process.

To verify that (C) holds for the process M_n, note that P.2 provides a measure $m(E)$ on subsets of A such that sets of positive measure are visited infinitely often. For a given set $E \subseteq A$, let

$$S_E = \{i \mid \Pr(M_n \in (E,i) \text{ inf. often}) = 1\} \;;$$

this set is well defined regardless of the initial state of the process by P.2. Now we define

$$m'(E,n) = m(E)\,\phi_{S_E}(n).$$

This provides a measure on the states of the M_n process satisfying the requirements of (C). Theorem 6.1 is now proved.

By (6.2) the probability distribution of any of the random variables of definition 1.2 or 1.3 is the same for either the process X_n or the process M_n, if the respective initial conditions are $X_o = \xi_o \in A$ and $M_o = \xi'_o = (\xi_o,0) \in A'$. Furthermore, P.3 or P.3' holds for X_n means that the M_n process satisfies condition (U) or (U') respectively for the set A'. Thus the results of sec. 3 and sec. 5 on limiting distributions can be applied in the somewhat more general setting of definition 6.2, so that the "almost recurrent" event E justifies its name by exhibiting much of the same limiting behavior as the "associated" recurrent event whose recurrence times have the generating function

$$F(x) = \int_A \sum_{n=1}^{\infty} f_n(\xi, A)\, x^n\, Q(d\xi).$$

7. Occupation Time for "Continuous" Processes.

The problem of occupation time for a "small" set of states in a Markov process has been discussed in sec. 5, and we have seen that the limiting distributions obtained are the Mittag-Leffler laws of (5.3). In sec. 2 we have proved Spitzer's theorem which asserts that the occupation time for the positive half-line in the case of sums of independent identically-distributed random variables approaches a generalized arc-sine law distribution. Few other general results are known. In this section, we shall treat a fairly extensive class of processes and obtain conditions for the occupation time of half-lines to have a limiting distribution.

Let $X_n(\omega)$ be a real-valued Markov process satisfying condition (C) of sec. 3, with the initial condition $X_0 = 0$. The fundamental assumption is that the path functions are continuous at 0, in the sense that

(7.1) $\quad X_{n-1}$ and X_{n+1} have different signs $\Longrightarrow$ $X_n = 0$.

Let N_n be the occupation time of the positive half-line up to time n, with the convention that if $X_n = 0$, it is counted or not counted as contributing to N_n according to whether the last preceding non-zero state assumed by the process was positive or not. (The interesting case will be when the expected recurrence time of state 0 is infinite, and in this case the time spent in 0 is a vanishingly small fraction of the total, so that the convention will not alter the limiting distribution of N_n/n.) We denote by $F(x) = \sum_{n=1}^{\infty} f_n x^n$ the generating function of the probabilities f_n that the recurrence time of state zero is n. Now the main theorem can be stated:

<u>Theorem 7.1</u>: Let X_n be as above with (7.1) holding. Then

$$\lim_{n\to\infty} \Pr(N_n/n < t) = V(t), \text{ a distribution function,} \tag{7.2}$$

if and only if

$$\lim_{n\to\infty} E(N_n/n) = \alpha \text{ exists, and also} \tag{7.3}$$

$$\lim_{x\to 1-} (1-x)\, F'(x)/(1-F(x)) = \delta \text{ exists, } 0 \le \delta \le 1. \tag{7.4}$$

If these conditions hold, $V(t) = F_{\alpha,\delta}(t)$ is the distribution with the moments $(-1)^k c_k$, where

$$\sum_{k=0}^{\infty} c_k z^k = \frac{(1+z)^{\delta-1} + \frac{1-\alpha}{\alpha}}{(1+z)^{\delta} + \frac{1-\alpha}{\alpha}} \quad . \tag{7.5}$$

<u>Proof:</u> The proof proceeds by finding the double generating function of N_n and operating with it in a similar manner to that used in the proof of theorem 2.2. Let $p = Pr(X_{n+1} > 0 | X_n = 0)$; it is convenient for the time being to assume that state 0 can not repeat itself, so that

$$q = 1 - p = Pr(X_{n+1} < 0 | X_n = 0).$$

We now define

$$f_n^{(1(2))} = Pr(X_n = 0,\ X_i \neq 0,\ 0 < i < n | X_o = 0,\ X_1 > (<)\ 0), \tag{7.6}$$

and

$$F_j(x) = \sum_{n=1}^{\infty} f_n^{(j)}\, x^n, \qquad j = 1,2,$$

and as we have done before denote

$$t_n^{(j)} = f_{n+1}^{(j)} + f_{n+2}^{(j)} + \cdots \quad \text{and} \quad T_j(x) = \sum_{n=o}^{\infty} t_n^{(j)}\, x^n = \frac{1 - F_j(x)}{1 - x}, \tag{7.7}$$

$$j = 1,2.$$

Then by the Markov property the following difference equation holds on letting $p_{k,n} = Pr(N_n = k)$:

$$p_{k,n} = p \sum_{1} f_1^{(1)}\, p_{k-1,n-1} + q \sum_{1} f_1^{(2)}\, p_{k,n-1} + pt_n^{(1)}\, \delta_{kn} + qt_n^{(2)}\, \delta_{ko}. \tag{7.8}$$

Taking double generating functions in (7.8) yields

$$P(x,y) = \sum_{k,n} p_{k,n}\, x^n\, y^k = pF_1(xy)\, P(x,y) + qF_2(x)\, P(x,y) + pT_1(xy) + qT_2(x)$$

so that upon rearrangement

$$P(x,y) = \frac{pT_1(xy) + qT_2(x)}{p(1-xy)\, T_1(xy) + q(1-x)\, T_2(x)} \quad . \tag{7.9}$$

If N_n/n is to have a limiting distribution, it is necessary for the first two moments to converge. Since $P(x,y) = \sum_n E(y^{N_n})\, x^n$, this implies that

$$(7.10)\quad \left.\frac{\partial P(x,y)}{\partial y}\right|_{y=1} = \sum_n E(N_n)\, x^n \sim \frac{\alpha}{(1-x)^2} \quad \text{where} \quad \alpha = \lim_{n\to\infty} E(N_n/n),$$

and also that

$$(7.11)\quad \left.\frac{\partial^2 P(x,y)}{\partial y^2}\right|_{y=1} = \sum_n E(N_n^2 - N_n)\, x^n \sim \frac{2\beta}{(1-x)^3} \quad \text{where } \beta = \lim_{n\to\infty} E\left(\left(\frac{N_n}{n}\right)^2\right).$$

(Above, and throughout this section, $A(x) \sim B(x)$ means that $\lim_{x\to 1-} A(x)/B(x) = 1$.) To see what this means for the generating functions, we can find the left hand side of (7.10) and (7.11) directly from (7.9). A straightforward computation yields

$$(7.12)\quad \lim_{x\to 1-} \frac{pT_1(x)}{pT_1(x) + qT_2(x)} = \alpha \text{ , and}$$

$$(7.13)\quad \lim_{x\to 1-} \frac{p(1-x)\, T_1'(x)}{pT_1(x) + qT_2(x)} = \gamma \text{ , where } \quad \beta = \gamma(1-\alpha) + \alpha^2 .$$

Now a degenerate distribution results if $\alpha = 0$ or 1, or, if this is not the case, if $\gamma = 0$. Therefore assume that $0 < \alpha < 1$, and that $0 < \gamma$. Then by means of (7.12), (7.13) may be reexpressed as

$$(7.14)\quad \lim_{x\to 1-} \frac{(1-x)\, T_1'(x)}{T_1(x)} = \frac{\gamma}{\alpha} = 1 - \delta \text{ ,} \qquad \delta < 1.$$

We shall show that (7.12) and (7.14) are sufficient as well as necessary for (7.2) to hold for a non-degenerate distribution $V(t)$.

Now (7.14) is the same as

$$\lim_{x\to 1-} \frac{(1-x)\, F_1'(x)}{1 - F_1(x)} = \delta ,$$

so that by corollary 4.2 we have

(7.15) $\quad 1 - F_1(x) = (1-x)^{\delta}\ L(1/(1-x))$, where L is slowly varying.

Using (7.12), it also follows that

(7.16) $\quad 1 - F_2(x) = \frac{p}{q}\frac{1-\alpha}{\alpha} c(x)\ (1-x)^{\delta}\ L(1/(1-x))$, where $c(x)\to 1$ as $x\to 1$.

From these two equations, using corollary 4.2 and the fact that $F(x) = pF_1(x) + qF_2(x)$, the necessity of (7.4) may be seen.

After this preparation, the method of sec. 2 can be applied. To do so, it is necessary to evaluate

$$\lim_{x\to 1-} (1-x) \sum_n E(\exp(-\lambda(1-x)N_n))\, x^n = \lim_{x\to 1-} (1-x)\ P(x,\exp(-\lambda(1-x)))$$

$$= g(\lambda), \text{ say.}$$

But upon substitution of (7.15) and (7.16) into (7.9) there results

(7.17) $\quad (1-x)\ P(x,e^{-\lambda(1-x)}) =$

$$\frac{(1-xe^{-(1-x)\lambda})^{\delta-1}\ L(\frac{1}{1-xe^{-(1-x)\lambda}}) + \frac{1-\alpha}{\alpha}\ c(x)\ (1-x)^{\delta-1}\ L(\frac{1}{1-x})}{(1-xe^{-(1-x)\lambda})^{\delta}\ L(\frac{1}{1-xe^{-(1-x)\lambda}}) + \frac{1-\alpha}{\alpha}\ c(x)\ (1-x)^{\delta}\ L(\frac{1}{1-x})} .$$

Using the slowly varying property of $L(u)$ and the fact that $1 - xe^{-(1-x)\lambda} \sim (1-x)\ (1+\lambda)$, taking limits in (7.17) yields

$$(7.18)\qquad g(\lambda) = \frac{(1+\lambda)^{\delta-1} + \frac{1-\alpha}{\alpha}}{(1+\lambda)^{\delta} + \frac{1-\alpha}{\alpha}} ,$$

which is analytic at $\lambda = 0$, and so may be expressed there by

$$(7.19)\qquad g(\lambda) = \sum_{k=0}^{\infty} c_k \lambda^k .$$

From (2.10) applied to the present case, it follows that

$$c_k = \lim_{x\to 1-} f_k(x) = \lim_{x\to 1-} \frac{(1-x)^{k+1}\ (-1)^k}{k!} \sum_n E(N_n^k)\, x^n, \text{ or}$$

$$(7.20)\qquad (1-x) \sum_n E(N_n^k)\, x^n \sim (-1)^k c_k\, k!/(1-x)^k .$$

Karamata's Tauberian theorem (in its customary form) applies since N_n is a non-decreasing random variable, and yields

$$(7.21)\qquad \lim_{n\to\infty} E((N_n/n)^k) = (-1)^k c_k .$$

Each N_n/n is a random variable taking only values between 0 and 1, and we have just shown that each moment converges. Therefore the $|c_k|$'s must form the moment sequence for a distribution to which the distribution of N_n/n converges.

Little remains to complete the proof of the theorem. We have shown the necessity (for (7.2)) of (7.3) and (7.4), and the sufficiency of (7.3) and (7.14). However, (7.3) implies (7.12), and (7.12) together with (7.4), corollary 4.2, and the relation $F = pF_1 + qF_2$ imply (7.14), so that the conditions of the theorem are also sufficient. Next, we observe that the assumption (which was used in writing (7.8)) that state 0 does not repeat itself is superfluous. For, provided $\delta \neq 1$, (7.4) implies that the expected value of the recurrence time for state 0 is infinite, and therefore the fraction of the time, up to time n, during which the process is in state 0 approaches 0 with probability one. Therefore the altered process obtained by deleting all repetitions of the state 0 has the same limiting behavior for N_n/n as does the original process, and the result we have derived may be applied to it. (For the same reason, if $\delta \neq 1$ the convention used for the state 0 in the definition of N_n could be dropped and N_n simply defined as the occupation time of the positive half-line.)

Finally, we must examine the previously excluded cases which lead to trivial distributions. If $\alpha = 1$ and (7.4) holds for some δ, the argument leading to (7.18) is valid, and (7.21) states that all moments converge to 1 so that $N_n/n \to 1$ in distribution. If $\alpha = 0$, by the same method we can see that the relative occupation time of the negative half-line approaches 1 and therefore $N_n/n \xrightarrow{d} 0$. The remaining case is when $\alpha \neq 0, 1$ and $\delta = 1$; this occurs in particular when the recurrence time is finite. Here (7.18) still holds, and the moments of N_n/n turn out to be the powers of α, so that for this case, $N_n/n \xrightarrow{d} \alpha$. (If $\alpha \neq 0,1$ and $\delta = 0$, all the moments are α, so that the limiting distribution $F_{\alpha,0}(t)$ has mass α at $t = 1$, and mass $1-\alpha$ at $t = 0$.) The proof of theorem 7.1 is now complete.

It is possible to obtain $F_{\alpha,\delta}(t)$ itself in certain cases. Let $\alpha = 1/2$ (as in the case of a process symmetric about 0). If also $\delta = 1/2$, as for coin tossing, we have

$$g(\lambda) = \frac{1}{\sqrt{1+\lambda}} = \sum_k \binom{-1/2}{k} \lambda^k .$$

But $(-1)^k \binom{-1/2}{k}$ are the moments of the classical arc-sine law. More generally, if $F_\beta(t)$ are the generalized arc-sine laws (2.7),

$$(7.22) \qquad F_{1/2,2^{-n}}(t) = \sum_{i=1}^{2^n-1} (-1)^{i+1} F_{i/2^n}(t).$$

This fact follows from (7.18) upon repeatedly rationalizing the denominator until all radicals have been removed, and then recognizing the quantities $(-1)^k \binom{-\beta}{k}$ as the moments of $F_\beta(t)$; we use the fact that the moment problem on a finite interval has a unique solution.

It is interesting to observe that this result is, like Spitzer's theorem 2.3 on sums of random variables, a generalization of the ordinary arc-sine law which holds for coin-tossing; the intersection of the two theories is the class of recurrent processes consisting of sums of independent, identically distributed random variables (spatial homogeneity) which are also continuous at some point, and is therefore essentially the case of coin-tossing. It also seems noteworthy that, provided the expected value of N_n/n converges, the condition (7.4) for a limiting distribution is again the condition which implies that for the state 0, Y_n/n and N_n/c_n (for suitable constants c_n) have limit distributions, as we have seen in sec. 2 and sec. 5. In conclusion, we observe that theorem 7.1 holds as well for non-real valued processes which need not even be Markov, provided that there is a state, say "0", occupation of which by the process is a recurrent event, and which divides the state-space of the process into two disjoint parts such that the process can pass from one part to the other only by occupying state 0.

REFERENCES

[1] E. Sparre Andersen, "On sums of symmetrically dependent random variables", Skandinavisk Aktuarietidskrift, 123-138, (1953).

[2] D.A. Darling and M. Kac, "On occupation times for Markov processes" to appear in Trans. Amer. Math. Soc.

[3] C. Derman, "A solution to a set of fundamental equations in Markov chains", Proc. Amer. Math. Soc. 5, 332-334 (1954).

[4] W. Doeblin, "Sur Deux Problemes de M. Kolmogoroff", Bull. Soc. Math. France, 66, 210-220 (1938).

[5] J.L. Doob, "Stochastic processes", New York, Wiley, 1953.

[6] E.B. Dynkin, "Nekotorye Predelnye Teoremy Dlya summ Nezavisimyh Slucainy Veyicin S Beskouecnymi Matematiceskimi Ozidaniyami", Izvestiia Akademiia Nauk. S.S.S.R. 19, 247-266 (1955).

[7] W. Feller, "Fluctuation theory of recurrent events", Trans. Amer. Math. Soc. 67, 98-119 (1949).

[8] W. Feller, "Probability theory and its applications", New York, Wiley, 1950.

[9] T.E. Harris, "The existence of stationary measures for certain Markov processes", To appear in Proc. of the Third Berkeley Symposium.

[10] M.P. Lévy, "Sur certains processus stochastiques homogènes", Compositio Math. 7, 283-339 (1939).

[11] F. Spitzer, "A combinatorial lemma and its applications to probability theory", to appear in Trans. Amer. Math. Soc.

[12] D.V. Widder, "The Laplace Transform", Princeton, 1946.