

AN OCCUPATION TIME THEOREM FOR A CLASS OF STOCHASTIC PROCESSES⁽¹⁾

BY
JOHN LAMPERTI

1. Introduction and statement of the theorem. Known general results on the problem of finding the limiting behavior of the occupation time of a set of states in a Markov process are of two types: when the set consists of one state or is small in a suitable sense, limiting distributions for the occupation time (properly normalized) are the Mittag-Leffler distributions, while in the case of a process consisting of sums of independent, identically-distributed random variables, limiting distributions of the fraction of the partial sums which are positive must be the generalized arc-sine laws. (Recent work on these problems is contained in [2; 1], and [3].) In this paper processes (not necessarily Markovian) will be considered having the property that the states are divided into two classes which communicate through the occurrence of a recurrent event (see [2]). More precisely, let $X_n(\omega)$ denote the state at time n of a stochastic process whose state space, except for one special state σ , is divided into two sets A and B . The assumptions are that occupation of state σ is a certain recurrent event (the initial condition is $X_0 = \sigma$), and that if $X_{n-1} \in A$ and $X_{n+1} \in B$ or vice versa, then $X_n = \sigma$. The main theorem describes the limiting behavior of the fraction of the time during which the process occupies set A , and as in the cases mentioned above the limiting distributions must be members of a definite class of distributions, depending on two parameters and of a new type.

Let $N_n(\omega)$ denote the occupation time up to time n of the set A , with the convention that occupation of the state σ is counted or not according to whether the last other state occupied was in A . 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 σ is n . Now the main theorem can be stated:

THEOREM 1. *Let $X_n(\omega)$ be the process described above. Then*

$$(1.1) \quad \lim_{n \rightarrow \infty} \Pr(N_n/n \leq t) = G(t)$$

exists if and only if

$$(1.2) \quad \lim_{n \rightarrow \infty} E(N_n/n) = \alpha \text{ exists,}$$

Received by the editors November 30, 1956.

⁽¹⁾ This work constitutes a portion of the author's California Institute of Technology doctoral thesis, written under the direction of Professor Frank Spitzer. Preparation of the paper was supported in part by the Office of Naval Research under contract Nonr 220(16). Another paper based on this thesis will appear in the Journal of Mathematics and Mechanics.

and also

$$(1.3) \quad \lim_{x \rightarrow 1^-} (1-x)F'(x)/(1-F(x)) = \delta \text{ exists,} \quad 0 \leq \delta \leq 1.$$

If both these conditions hold, $G(t) = G_{\alpha, \delta}(t)$ is the distribution on $[0, 1]$ which, provided α and $\delta \neq 0$ or 1 , has the density

$$(1.4) \quad G'_{\alpha, \delta}(t) = \frac{a \sin \pi \delta}{\pi} \frac{t^\delta(1-t)^{\delta-1} + t^{\delta-1}(1-t)^\delta}{a^2 t^{2\delta} + 2at^\delta(1-t)^\delta \cos \pi \delta + (1-t)^{2\delta}},$$

where $a = \frac{1-\alpha}{\alpha}$,

while in other cases

$$(1.5) \quad G_{0, \delta}(t) = 1 \quad \text{and} \quad G_{1, \delta} = 0 \quad \text{for} \quad 0 < t < 1,$$

$$G_{\alpha, 1}(t) = \begin{cases} 0 & \text{for } t < \alpha \\ 1 & \text{for } t \geq \alpha \end{cases} \quad \text{and} \quad G_{\alpha, 0}(t) = 1 - \alpha \quad \text{for} \quad 0 \leq t < 1.$$

The class of processes to which Theorem 1 is applicable is quite extensive, including all Markov chains in which one state separates the rest into two sets; in particular, it applies to all random walks. It is interesting to observe that this result is, like the theorem [3, Theorem 6.1] on sums of random variables, a generalization of the ordinary arc-sine law which holds for the coin-tossing or simple random walk process; when both α and δ are $1/2$, as they are for that case, (1.4) becomes the classical law. The intersection of the cases covered by the two theories is the class of recurrent processes consisting of sums of random variables which also are "continuous" at some point, and is therefore (except that the random variables added may have positive probability of being 0) just the coin-tossing process. It is also worth noting that the "ergodic" case, when the expected recurrence time for state σ is finite, is not an interesting one: in this case $\delta=1$ and the trivial distribution of (1.5) holds. In the other cases when the recurrence time is not finite, the convention used in the definition of N_n for the occurrences of state σ is a technical convenience only and does not affect the limiting distributions obtained.

§2 is devoted to a theorem on slowly varying functions which is needed as a lemma for the proof of Theorem 1; the theorem is a close relative of a well-known one [4, Chapter V]. It is felt that this theorem is of independent interest, however, as it relates two conditions which arise frequently in Tauberian arguments in probability. The proof given here is due to Professor H. F. Bohnenblust, and is both conceptually simpler and more general than the author's original proof. With the aid of this theorem, and upon reference to [2], it appears that condition (1.3) is precisely the necessary and sufficient condition that the occupation time of state σ itself, suitably normalized, shall possess a limiting distribution, for (1.3) can be rewritten as

$$(1.6) \quad 1 - F(x) = (1 - x)^\delta L\left(\frac{1}{1 - x}\right), \quad \text{where } L(y) \text{ is slowly varying,}$$

which is equivalent to condition (5.6) of [2].

2. A theorem on slowly varying functions.

DEFINITION (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

$$(2.1) \quad \lim_{y \rightarrow \infty} \frac{L(cy)}{L(y)} = 1 \quad \text{for all } c > 0,$$

THEOREM 2. For $A > x > 0$, let $f(x)$ be a positive function. If $f'(x)$ exists for small enough x and satisfies

$$(2.2) \quad \lim_{x \rightarrow 0^+} \frac{xf'(x)}{f(x)} = \alpha,$$

then

$$(2.3) \quad f(x) = x^\alpha L(1/x)$$

where $L(y)$ is a slowly varying function. Conversely, if (2.3) holds for some number α , and if $f'(x)$ exists and is monotone in a neighborhood of 0, then (2.2) holds.

Proof. Assume (2.2), and define $L(y)$ by putting $f(x) = x^\alpha L(1/x)$. Then $\log f(x) = \alpha \log x + \log L(1/x) = \alpha \log x + g(x)$, say. We must show that $L(y)$ satisfies (2.1); this is equivalent to

$$(2.4) \quad \lim_{x \rightarrow 0} (g(cx) - g(x)) = 0 \quad \text{for all } c > 0.$$

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 (2.2),

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

To prove the converse, suppose $f'(x)$ is nondecreasing and that (2.3) holds. (Exactly the same argument works also for f' nonincreasing.) 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 nondecreasing, we can write

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

or

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

Now by (2.3) and (2.1) we obtain upon letting $x \rightarrow 0+$,

$$\limsup_{x \rightarrow 0} \frac{xf'(x)}{f(x)} \leq (c^\alpha - 1) \frac{1}{c - 1} \leq c^{\alpha-1} \liminf_{x \rightarrow 0} \frac{xf'(x)}{f(x)}.$$

But this holds for all $c > 1$; letting $c \rightarrow 1$ yields precisely (2.2).

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, $f''(x) = -\int_0^\infty t^2 e^{-tx} dG(x) < 0$ so that $f'(x)$ is a decreasing function of x . Thus:

COROLLARY 2.1. *Let $\phi(x) = \int_0^\infty e^{-tx} dG(t)$, where $G(t)$ is monotone and is of total variation one. Then the following are equivalent:*

- (A)
$$\lim_{x \rightarrow 0+} \frac{x\phi'(x)}{1 - \phi(x)} = -\alpha,$$
- (B)
$$1 - \phi(x) = x^\alpha L\left(\frac{1}{x}\right)$$

where $L(y)$ is a slowly varying function.

Stating this in terms of generating functions yields

COROLLARY 2.2. *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)
$$\lim_{x \rightarrow 1-} \frac{(1-x)P'(x)}{1 - P(x)} = \alpha,$$
- (B)
$$1 - P(x) = (1-x)^\alpha L\left(\frac{1}{1-x}\right)$$

where $L(y)$ is slowly varying.

3. Proof of Theorem 1. The proof proceeds by operating with the double generating function of the quantities $p_{k,n} = \Pr(N_n = k)$, and is analogous in its overall plan to Spitzer's proof of his Theorem 6.1 in [3]. First by Abelian arguments we show that the generating functions of the moments of N_n have the proper asymptotic behavior, and then Karamata's Tauberian theorem is used to show the convergence of the moments. This yields the Stieltjes

transform of the limiting distribution of N_n , which can be inverted explicitly yielding (1.4) and (1.5).

Let $p = \Pr (X_{n+1} \in A | X_n = \sigma)$. It is convenient to assume for the time being that state σ can not repeat itself, so that

$$q = 1 - p = \Pr (X_{n+1} \in B | X_n = \sigma).$$

We now define

$$(3.1) \quad \begin{aligned} f_n^{(1(2))} &= \Pr (X_n = \sigma, X_i \neq \sigma, 0 < i < n | X_0 = \sigma, X_1 \in A(B)), \\ F_j(x) &= \sum_{n=1}^{\infty} f_n^{(j)} x^n, \end{aligned} \quad j = 1, 2,$$

and also denote the “tails” by

$$(3.2) \quad t_n^{(j)} = f_{n+1}^{(j)} + f_{n+2}^{(j)} + \dots \quad \text{and} \quad T_j(x) = \sum_{n=0}^{\infty} t_n^{(j)} x^n = \frac{1 - F_j(x)}{1 - x}, \quad j = 1, 2.$$

Now by the fact that state σ is a recurrent event, the following difference equation holds:

$$(3.3) \quad p_{k,n} = p \sum_m f_m^{(1)} p_{k-m,n-m} + q \sum_m f_m^{(2)} p_{k,n-m} + p l_n^{(1)} \delta_{k,n} + q l_n^{(2)} \delta_{k,0}.$$

Taking double generating functions in (3.3) yields

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

so that upon rearrangement

$$(3.4) \quad P(x, y) = \frac{pT_1(xy) + qT_2(x)}{p(1 - xy)T_1(xy) + q(1 - x)T_2(x)}.$$

Now 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

$$(3.5) \quad \left. \frac{\partial P(x, y)}{\partial y} \right|_{y=1} = \sum_n E(N_n)x^n \sim \frac{\alpha}{(1-x)^2} \quad \text{where} \quad \alpha = \lim_{n \rightarrow \infty} E(N_n/n),$$

and also that

$$(3.6) \quad \left. \frac{\partial^2 P(x, y)}{\partial y^2} \right|_{y=1} = \sum_n E(N_n^2 - N_n)x^n \sim \frac{2\beta}{(1-x)^3} \quad \text{where} \quad \beta = \lim_{n \rightarrow \infty} E \left[\left(\frac{N_n}{n} \right)^2 \right].$$

(Above and throughout, $f(x) \sim g(x)$ means that $\lim_{x \rightarrow 1-} f(x)/g(x) = 1$.) To apply these conditions, we can find the left hand sides of (3.5) and (3.6)

directly from (3.4). A straightforward computation yields

$$(3.7) \quad \lim_{x \rightarrow 1-} \frac{pT_1(x)}{pT_1(x) + qT_2(x)} = \alpha,$$

$$(3.8) \quad \lim_{x \rightarrow 1-} \frac{p(1-x)T_1'(x)}{pT_1(x) + qT_2(x)} = \gamma, \quad \text{where } \beta = \gamma(1-\alpha) + \alpha^2.$$

Assume now that α and $\delta \neq 0$ or 1. (These cases, as will be checked later, lead to degenerate distributions.) Then by means of (3.7), (3.8) may be re-expressed as

$$(3.9) \quad \lim_{x \rightarrow 1-} \frac{(1-x)T_1'(x)}{T_1(x)} = \frac{\gamma}{\alpha} = 1 - \delta \quad \text{say, where } \delta < 1.$$

We shall show that (3.7) and (3.9) are sufficient as well as necessary for (1.1) to hold with a nondegenerate distribution function $G(t)$.

Now (3.9) is the same as

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

so that by Corollary 2.2 we have

$$(3.10) \quad 1 - F_1(x) = (1-x)^\delta L\left(\frac{1}{1-x}\right),$$

where L is slowly varying. Using (3.7) it also follows that

$$(3.11) \quad 1 - F_2(x) = \frac{p}{q} \frac{1-\alpha}{\alpha} c(x)(1-x)^\delta L\left(\frac{1}{1-x}\right),$$

where $c(x) \rightarrow 1$ as $x \rightarrow 1$. From these two equations, using Corollary 2.2 and the fact that $F(x) = pF_1(x) + qF_2(x)$, the necessity of (1.3) may be seen.

With this much preparation, we can proceed to the description of the generating functions of the moments of N_n . Observe (as in [3]) that

$$\begin{aligned} (1-x) \sum_{n=0}^{\infty} E(\exp(-\lambda(1-x)N_n))x^n \\ &= (1-x) \sum_k x^n \left[\sum_n \frac{E(N_n^k)}{k!} (-\lambda(1-x))^k \right] \\ (3.12) \quad &= \sum_k \lambda^k f_k(x), \quad \text{where } f_k(x) = \frac{(1-x)^{k+1}}{k!} (-1)^k \left[\sum_n E(N_n^k) x^n \right]. \end{aligned}$$

For $|x| < 1$, this is an analytic function of x and λ . We shall show that the limit as $x \rightarrow 1-$ exists and is an analytic function of λ , say $g(\lambda) = \sum_n c_n \lambda^n$,

and hence conclude that the limit of $f_k(x)$ is c_k . To achieve this, we combine (3.10) and (3.11) with (3.4) to obtain

$$(3.13) \quad \begin{aligned} (1-x) \sum_n E(\exp(-\lambda(1-x)N_n))x^n &= (1-x)P(x, \exp[-\lambda(1-x)]) \\ &= \frac{(1-x)e^{-(1-x)\lambda} L\left(\frac{1}{1-xe^{-(1-x)\lambda}}\right) + \frac{1-\alpha}{\alpha} c(x)(1-x)^{\delta-1} L\left(\frac{1}{1-x}\right)}{(1-xe^{-(1-x)\lambda})^\delta L\left(\frac{1}{1-xe^{-(1-x)\lambda}}\right) + \frac{1-\alpha}{\alpha} c(x)(1-x)^\delta L\left(\frac{1}{1-x}\right)}. \end{aligned}$$

Using the slowly varying property of $L(u)$ and the fact that $1-xe^{-\lambda(1-x)} \sim (1-x)(1+\lambda)$, taking limits in (3.13) yields

$$(3.14) \quad g(\lambda) = \frac{(1+\lambda)^{\delta-1} + \frac{1-\alpha}{\alpha}}{(1+\lambda)^\delta + \frac{1-\alpha}{\alpha}}$$

which is analytic at $\lambda=0$, justifying the expression as $\sum_n c_n \lambda^n$. Therefore $\lim_{x \rightarrow 1^-} f_k(x) = c_k$, or

$$(3.15) \quad (1-x) \sum_n E(N_n^k)x^n \sim (-1)^k c_k k!(1-x)^{-k}.$$

Karamata's Tauberian theorem (see, for instance, [4, Chapter V]) may be applied to this set of asymptotic expressions since N_n (and therefore $E(N_n^k)$) are nondecreasing with n , and the conclusion is

$$(3.16) \quad \lim_{n \rightarrow \infty} E[(N_n/n)^k] \text{ exists, } = (-1)^k c_k.$$

But each N_n/n is a random variable taking only values between 0 and 1, and it has just been shown that each moment converges to a limit as $n \rightarrow \infty$. Therefore the $(-1)^k c_k$'s must form the moment sequence for a distribution to which the distribution of N_n/n converges.

We have just seen (at least when α and $\delta \neq 0$ or 1) that $(-1)^k c_k$ is the k 'th moment of $G(t)$, so that

$$g(x) = \sum_k c_k x^k = \sum_k (-1)^k \int_0^1 t^k dG(t)x^k = \int_0^1 \sum_k (-tx)^k dG(t) = \int_0^1 \frac{dG(t)}{1+tx},$$

which on putting $u=1/x$ becomes

$$(3.17) \quad \int_0^1 \frac{dG(t)}{t+u} = (1/u)g(1/u) = \frac{(u+1)^{\delta-1} + \frac{1-\alpha}{\alpha} u^{\delta-1}}{(u+1)^\delta + \frac{1-\alpha}{\alpha} u^\delta}.$$

If the integral in (3.17) is extended from 0 to ∞ instead of 0 to 1, the expression becomes the Stieltjes transform of $G(t)$, which may be inverted (see [4, Chapter VIII]) yielding the density (1.4) for $0 \leq x \leq 1$ and 0 elsewhere⁽²⁾.

Most of the hard work is now done. We have shown the necessity (for (1.1)) of (1.2) and (1.3), and the sufficiency of (1.2) and (3.9). However, (1.2) implies (3.7), and this, together with (1.3), Corollary 2.2, and the relation $F = pF_1 + qF_2$ imply (3.9), so that the conditions of the theorem are also sufficient. Next we observe that the assumption (used in writing (3.3)) that state σ does not repeat itself is superfluous. For, if $\delta \neq 1$, (1.3) implies that the expected value of the recurrence time for state σ is infinite, and therefore the fraction of the time, up to time n , during which the process is in state σ approaches 0 with probability one. Therefore the altered process obtained by deleting all repetitions of the state σ 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. (Similarly, when the mean recurrence time for σ is infinite, the convention for state σ used in the definition of N_n could be dropped and N_n simply defined as the occupation time of set A .)

Finally, we must examine the previously excluded extreme cases. If $\alpha = 1$ and (1.3) holds for some δ , the argument leading to (3.14) is valid, and then (3.16) states that all moments converge to 1, so that $N_n/n \rightarrow 1$ in probability. If $\alpha = 0$, by the same method we can see that the relative occupation time of set B approaches 1, so that in this case $N_n/n \rightarrow 0$ in probability. If now $\alpha \neq 0$ or 1 and $\delta = 1$, as when the recurrence time of state σ is finite, (3.14) again holds and implies that the moments of N_n/n converge to the successive powers of α , so that the limiting distribution has mass 1 at α . Finally, if $\alpha \neq 0$ or 1 and $\delta = 0$, all the limiting moments are α , so that $G(t)$ in this case has mass α at $t = 1$, and mass $1 - \alpha$ at $t = 0$. This proves (1.5) and completes the proof of the theorem.

REFERENCES

1. D. A. Darling and M. Kac, *On occupation times for Markov processes*, Trans. Amer. Math. Soc. vol. 84 (1957) pp. 444-458.
2. W. Feller, *Fluctuation theory of recurrent events*, Trans. Amer. Math. Soc. vol. 67 (1949) pp. 98-119.
3. F. Spitzer, *A combinatorial lemma and its application to probability theory*, Trans. Amer. Math. Soc. vol. 82 (1956) pp. 323-340.
4. D. V. Widder, *The Laplace transform*, Princeton, 1941.

CALIFORNIA INSTITUTE OF TECHNOLOGY,
PASADENA, CALIF.

⁽²⁾ The author is indebted to Dr. Spitzer for pointing out the possibility of the explicit formula (1.4).