A "RENEWAL" LIMIT THEOREM FOR GENERAL STOCHASTIC PROCESSES[1]

BY V. E. BENEŠ

Bell Telephone Laboratories and Dartmouth College

1. Introduction. Let $\{x_t, -\infty < t < \infty\}$ be a stochastic process that takes values in a (possibly abstract) space X. Specifically, assume as given a space Ω of points ω, a Borel field $\mathfrak{F}$ of ω-sets, a probability measure P of $\mathfrak{F}$-sets, a Borel field $\mathfrak{A}$ of X-sets, and for each real t a function $x_t(\cdot)$ from Ω to X such that $\{x_t \varepsilon A\}$ is an $\mathfrak{F}$-set for $A \varepsilon \mathfrak{A}$. The existence and calculation of

$$\lim_{t\to\infty} \Pr\{x_t \varepsilon A\}, \tag{1}$$

under the weakest possible conditions, is a problem of considerable interest in probability theory.

Some processes for which this problem has received much attention are the Markov, semi-Markov, and regenerative processes. (The term 'semi-Markov' is Lévy's in [6], while the more inclusive 'regenerative' is Smith's, in [10] and [11].) These random processes share the very strong property of having *regeneration points*, junctures at which the previous history (except for the present state, possibly) of the process becomes irrelevant to its future development. In Smith's phrase, the past loses all prognostic significance at a regeneration point. In a sense, then, the process consists of segments that are mutually independent. (See Smith's construction, [11], p. 256.) Because of this independence, the renewal theorems of Feller [4], Blackwell [1], and Smith [9] have been directly useful in studying the limit (1) for these processes.

Our limit theorem is of "renewal" type in the sense that it depends on a relationship between x_t and a discrete parameter real process $\{S_k, k \text{ an integer}\}$ that is analogous to the classical (sequence of) renewal points. However, we do not assume that any *regeneration* occurs at these points. Our theorem differs from previous results in requiring no assumptions of independence, and resting only on mild stationarity properties, to the effect that certain kernels are difference kernels. The approach that we use to study (1) amounts to carefully distinguishing between the *independence* and the stationarity properties of x_t, and using only the latter. The approach is fruitful because (as will appear) our theory applies to many processes that do not have regeneration points.

The present paper was inspired in large part by the work of W. L. Smith (on renewal theory and regenerative processes); he will recognize in it generalizations and copies of many of his ideas. A detailed discussion of the relation of our work to Smith's is given in Section 4. Also, this paper is a general develop-

Received February 1, 1960; revised December 3, 1960.

[1] This paper was supported in part by a grant from the National Science Foundation to Dartmouth College.

ment of a method used by the author in a special problem (see [1a], especially Section XII).

2. Summary. Let $\{S_k, k \text{ an integer}\}$ be a discrete-parameter real stochastic process defined on the same measure space as x_t, with $S_{k+1} > S_k$ a.s. and

$$H(t) - H(s) = E\{\text{number of } S_k \text{ in } (s, t]\} < \infty, \qquad s < t.$$

According to Theorem 1 $\Pr\{x_t \varepsilon A\}$ can always be represented in the form

$$\int_{-\infty}^{t} Q(A, t, u)\, dH(u).$$

The class of events of the form $\{x_u \varepsilon A\}$, $-\infty < u < \infty$, is called *weakly stationary* with respect to $\{S_k\}$ if its *representative kernel* $Q(A, \cdot, \cdot)$ is a difference kernel, $Q(A, t, u) = Q(A, t - u)$. The class $\{x_u \varepsilon A\}$ is *weakly stationary* if it is so with respect to *some* sequence $\{S_k\}$ with the above properties. The process y_t is defined as the time from t to the next S_k after t, i.e.,

$$y_t = \min\{S_k - t \mid S_k > t\}.$$

The hypotheses of the main result (Theorem 3) are: (i) the classes (of events)

$$\{y_u < \infty\}, \qquad -\infty < u < \infty$$

$$\{x_u \varepsilon A\}, \qquad -\infty < u < \infty$$

are both weakly stationary with respect to $\{S_k\}$ with the respective kernels $Q(X, \cdot)$ and $Q(A, \cdot)$; (ii) the $\{S_k\}$ are "aperiodic" in the sense that the Fourier-Stieltjes transform of $Q(X, \cdot)$ vanishes at zero only; (iii) $H(\cdot + 1) - H(\cdot)$ is bounded. The conclusions of Theorem 3 are as follows: (i) if $Q(X, \cdot)$ is in L_1 and $Q(A, \cdot)$ belongs to the function-class K_H of [1], then

$$\Pr\{x_t \varepsilon A\} \to \frac{\int_0^\infty Q(A, u)\, du}{\int_0^\infty Q(X, u)\, du} \quad \text{as} \quad t \to \infty;$$

(ii) if $Q(X, \cdot)$ is not in L_1, and $Q(A, \cdot)$ belongs to N (defined below), then $\Pr\{x_t \varepsilon A\} \to 0$ as $t \to \infty$; (iii) if $Q(X, \cdot)$ is not in L_1, and $Q(X, \cdot) - Q(A, \cdot)$ is in N, then $\Pr\{x_t \varepsilon A\} \to 1$ as $t \to \infty$.

The function-class K_H (defined in [1] and later in this paper) depends on the measure induced by H. If f is continuous a.e., it belongs to K_α for any positive measure α that satisfies the hypothesis of Wiener's Theorem (Widder [12], p. 214). The function-class N is just Wiener's M with the requirement of continuity removed.

It is shown in the course of proving Theorem 3 that the limit (1) depends only on H_+ defined to be 0 for negative argument, and $H(t) - H(0-)$ for $t \geqq 0$, and that H_+ must have the mathematical form of the classical renewal function (see Smith [9]) even though the $\{S_k\}$ do not necessarily form a renewal process. That is,

$$H_+ = F + F * G + F * G * G + \cdots \tag{2}$$

where F, G are distribution functions, and $*$ denotes Stieltjes convolution. The weak stationarity condition on y_t alone, without any independence assumptions, implies that the study of the asymptotic behavior of H reduces to a problem of renewal theory.

The preceding definitions and results require only minor modifications to cover the case in which x_t is defined for $0 \leqq t < \infty$ only, and the $\{S_k\}$ are positive and defined for $k \geqq 0$ only, with $S_0 = 0$.

The last section describes some ways of defining classes of events that are weakly stationary from arbitrary stochastic processes. The constructions show that in spite of (2), weakly stationary classes of events are not limited to stochastic processes of regenerative type.

It is important to understand that the validity of (2) is implied by, but does not imply, the fact that the $\{S_k\}$ form a renewal process. In other words, it is possible for (2) to hold without $\{S_k\}$ being a renewal process. From the representation (2) one can prove that there is some general renewal process (Smith's phrase) for which H_+ is a (modified) renewal function, but not that

$$\{S_k, k \geqq 1\}$$

is this process. In Section 6 we show how to construct an example in which (2) holds and the $\{S_k\}$ are not sums of independent random variables. It is a special feature of the notion of weak stationarity that it permits the use of the mathematical properties of the classical renewal function without incurring hypotheses of independence. That (2) holds is a relatively nontrivial (and perhaps unexpected) consequence of the fact that $Q(X, t, u)$ is a *difference* kernel, and it neither depends on nor implies any independence properties.

The point to grasp is that if $Q(X, t, u)$ and $Q(A, t, u)$ are difference kernels, then the event $\{x_t \varepsilon A\}$ behaves as though the $\{S_k\}$ formed a renewal process, irrespective of independences that may or may not obtain.

3. A Representation for $\Pr\{x_t \varepsilon A\}$. We suppose that $\{S_k, k = 0, \pm 1, \pm 2, \cdots\}$ is a discrete-parameter real stochastic process, defined on the same measure space as x_t, and such that

$$\infty > S_{k+1} > S_k, \qquad \text{a.s.,} \tag{3}$$

$$H(t) - H(s) = \sum_{k=-\infty}^{\infty} \Pr\{s < S_k \leqq t\} < \infty, \qquad s < t. \tag{4}$$

The $\{S_k\}$ are a generalization of renewal points obtained by discarding all hypotheses of independence. It is easily seen that

$$H(t) - H(s) = E\{\text{number of } S_k \text{ in } (s, t]\},$$

so that H is analogous to the renewal function of Smith [9], [10], [11].

THEOREM 1. *The probability of* $\{x_t \varepsilon A\}$ *can always be represented in the form*

$$\Pr\{x_t \varepsilon A\} = \int_{-\infty}^{t} Q(A, t, u)\, dH(u).$$

PROOF. Let $\mathfrak{F}_k$ be the smallest Borel field with respect to which S_k is measurable. From (3),

$$\begin{aligned}\Pr\{x_t \varepsilon A\} &= \sum_{k=-\infty}^{\infty} \Pr\{x_t \varepsilon A \;\&\; S_k \leqq t < S_{k+1}\} \\ &= \sum_{k=-\infty}^{\infty} \int_{\{S_k \leqq t\}} \Pr\{x_t \varepsilon A \;\&\; S_{k+1} > t \mid \mathfrak{F}_k\} P(d\omega).\end{aligned} \tag{5}$$

Choose particular versions of the conditional probabilities in (5), to be fixed henceforth. These versions are $\mathfrak{F}_k$-measurable ω-functions and so (Doob [3], p. 603, Theorem 1.5) have the form

$$\Pr\{x_t \varepsilon A \;\&\; S_{k+1} > t \mid \mathfrak{F}_k\} = F_k(A, t, S_k(\omega)), \tag{6}$$

where $F_k(A, t, \cdot)$ is a Baire function. To keep the interpretation[2] clear we write

$$F_k(A, t, u) = \Pr\{x_t \varepsilon A \;\&\; S_{k+1} > t \mid S_k = u\}. \tag{7}$$

An obvious transformation (Halmos [5], p. 207) then gives

$$\Pr\{x_t \varepsilon A\} = \sum_k \int_{-\infty}^{t} \Pr\{x_t \varepsilon A \;\&\; S_{k+1} > t \mid S_k = u\}\, d\Pr\{S_k \leqq u\}. \tag{8}$$

Each $\Pr\{S_k \leqq \cdot\}$ measure is absolutely continuous with respect to H, with Radon-Nikodým derivative

$$\varphi_k(u) = \frac{d\Pr\{S_k \leqq u\}}{dH(u)},$$

$$\sum_k \varphi_k(u) = 1 \tag{9}$$

$$0 \leqq \varphi_k(u) \leqq 1 \tag{10}$$

(9), (10) a.e. $[H]$.

The first is true by definition; if the second fails on a set B of positive H-measure, then either $\Pr\{S_k \varepsilon B\} < 0$ or

$$\Pr\{S_k \varepsilon B\} = \int_B \varphi_k(u)\, dH(u) > H(B),$$

both of which conclusions are false. In (8) set $A = X$ and obtain

$$\begin{aligned}1 &= \sum_k \int_{-\infty}^{t} \Pr\{S_{k+1} > t \mid S_k = u\}\varphi_k(u)\, dH(u) \\ &= \int_{-\infty}^{t} \sum_k \Pr\{S_{k+1} > t \mid S_k = u\}\varphi_k(u)\, dH(u).\end{aligned} \tag{11}$$

---

[2] This useful abuse will be used repeatedly. If the r.v. y generates the field $\mathfrak{F}$, we write $\Pr\{\Lambda \mid y = u\}$ for the a.s. value of $\Pr\{\Lambda \mid \mathfrak{F}$ when $y(\omega) = u$.

The interchange of $\sum$ and $\int$ in the last step is effected by the monotone convergence theorem (Loève [7], p. 124). By the same argument

$$\Pr\{x_t \varepsilon A\} = \int_{-\infty}^{t} Q(A, t, u)\, dH(u),$$

where for each t, $Q(A, t, \cdot)$ is defined a.e. $[H]$ by

$$Q(A, t, u) = \sum_k \Pr\{x_t \varepsilon A \,\&\, S_{k+1} > t \mid S_k = u\}\varphi_k(u). \tag{12}$$

Theorem 1 justifies calling $Q(A, \cdot, \cdot)$ the *representative* kernel of the set of events $\{x_u \varepsilon A\}$, $-\infty < u < \infty$. This result stems entirely from the facts that x_t and the $\{S_k\}$ are defined on the same measure space, and that (3) and (4) hold. It is tempting to think of $Q(A, t, u)$ as a version of

$$\Pr\{x_t \varepsilon A \mid \text{last } S_k \text{ prior to } t \text{ was at } u\};$$

however, this interpretation is not correct, as will appear in the next section.

4. Weak Stationarity. Some additional relationship between x_t and the $\{S_k\}$ is required for study of the limit (1). If v_t is a stochastic process taking values in X, and $B \subset X$, the class of events of the form $\{v_u \varepsilon B$, $-\infty < u < \infty$, is called *weakly stationary* with respect to $\{S_n\}$ if its representative kernel depends only on the difference of its (last two) arguments, i.e., is a difference kernel. This usage resembles calling a second-order process stationary in the wide sense if its covariance is a difference kernel. It will be shown that the regenerative and equilibrium processes of Smith [10] are weakly stationary with respect to the renewal processes on which they are based.

In his paper [10] Smith has introduced the notion of an *equilibrium* process. Since this notion resembles that of weak stationarity, it is important to distinguish between the two notions, and to discuss the relations between them. Smith works only (with stochastic processes defined) on the interval $[0, \infty)$, while we work on $(-\infty, \infty)$; this is not an important difference. To facilitate comparison with Smith's work, the discussion ensuing in this section is for a process $\{x_t, t \geqq 0\}$ defined for $t \geqq 0$ only.

Stripped of inessentials and couched in our notation, Smith's equilibrium process can be described as follows. The sequence $\{S_k, k \geqq 0\}$ of random variables is assumed to form a *general renewal process*: that is, $S_0 = 0$, and the differences

$$S_{k+1} - S_k, \qquad k \geqq 0, \tag{13}$$

are mutually independent, nonnegative, and except possibly $S_1 - S_0$, identically distributed. The random processes $\{n(t), t \geqq 0\}$ and $\{S_{n(t)}, t \geqq 0\}$ are defined by the respective conditions

$$\left.\begin{aligned} n(t) &= k \\ S_{n(t)} &= S_k \end{aligned}\right\} \text{ if and only if } \quad S_k \leqq t < S_{k+1}, \qquad k \geqq 0.$$

Clearly, $n(t)$ is the number of S_k in $(0, t]$, and $S_{n(t)}$ is the epoch of the last S_k prior to t. Smith starts (essentially although not explicitly) with the representation, on $t \geqq 0$,

$$\Pr\{x_t \varepsilon A\} = \int_\Omega \Pr\{x_t \varepsilon A \mid \mathfrak{F}_t\} P(d\omega),$$

where $\mathfrak{F}_t$ is the Borel field generated by $S_{n(t)}$. By an obvious transformation of the space of integration, this can be written as

$$\begin{aligned}\Pr\{x_t \varepsilon A\} &= \int_{0-}^{t} \Pr\{x_t \varepsilon A \mid S_{n(t)} = u\}\, d_u \Pr\{S_{n(t)} \leqq u\} \\ &= \Pr\{x_t \varepsilon A \ \&\ n(t) = 0\} + \int_{0-}^{t} \Pr\{x_t \varepsilon A \mid S_{n(t)} \\ &\qquad = u, n(t) > 0\}\, du \Pr\{S_{n(t)} \leqq u \ \&\ n(t) > 0\}.\end{aligned} \tag{14}$$

Smith says that "an equilibrium process has been defined" if the kernel of the second integral of (14) has the form

$$\Pr\{x_t \varepsilon A \mid S_{n(t)} = u, \quad n(t) > 0\} = \varphi_A(t - u). \tag{15}$$

This kernel is correctly interpretable as

$$\Pr\{x_t \varepsilon A \mid \text{last } S_k \text{ prior to } t \text{ was at } u, n(t) > 0\}$$

and it is not the same kernel as would appear in our analogue of Theorem 1 for a process on $[0, \infty)$. Smith goes on to compute the measures

$$\Pr\{S_{n(t)} \leqq u, \quad n(t) > 0\}, \quad 0 < u \leqq t. \tag{16}$$

Using the fact that the $\{S_k\}$ form a general renewal process, he finds that these are given by

$$\int_{0-}^{u} [1 - G(t - v)]\, dH_+(v), \qquad 0 < u \leqq t,$$

where

$$H_+(v) = E\{\text{number of } S_k \text{ in } (0, v]\}, \tag{17}$$

$$\begin{aligned} H_+ &= K + K * G + K * G * G + \cdots \\ K &= \text{distr}\{S_1\} \\ G &= \text{distr}\{S_{k+1} - S_k\}, \qquad k > 1. \end{aligned} \tag{18}$$

This yields his final representation:

$$\begin{aligned}\Pr\{x_t \varepsilon A\} &= \Pr\{x_t \varepsilon A \ \&\ n(t) = 0\} \\ &\quad + \int_{0-}^{t} \varphi_A(t - u)[1 - G(t - u)]\, dH_+(u).\end{aligned} \tag{19}$$

The development just given only used the assumption that $\{S_k\}$ is a general renewal process in the computation of (16). Without this assumption one can still write (16) as

$$\begin{aligned}\Pr\{S_{n(t)} \leqq u, n(t) > 0\} &= \sum_{k=1}^{\infty} \Pr\{S_k \leqq u \;\&\; S_{k+1} > t\} \\ &= \int_{0-}^{u} \sum_{k=1}^{\infty} \Pr\{S_{k+1} > t \mid S_k = v\}\psi_k(v)\, dH_+(v),\end{aligned}$$

where H_+ is given by (17), without (18),

$$\psi_k(v) = \frac{d \Pr\{S_k \leqq v\}}{dH_+(v)}, \qquad k \geqq 1,$$

and the $\psi_k(\cdot)$ may be taken to form a probability distribution in $k \geqq 1$ for each $v \geqq 0$. The general form of the representation (19), using neither of Smith's assumptions (13) and (15), is then

$$\begin{aligned}\Pr\{x_t \,\varepsilon\, A\} &= \Pr\{x_t \,\varepsilon\, A \;\&\; n(t) = 0\} \\ &+ \int_{0-}^{t} \Pr\{x_t \,\varepsilon\, A \mid S_{n(t)} = u\} \sum_{k=1}^{\infty} \Pr\{S_{k+1} > t \mid S_k = u\}\psi_k(u)\, dH_+(u).\end{aligned} \tag{20}$$

When $\{S_k\}$ is a general renewal process, we have

$$\begin{aligned}\Pr\{S_{k+1} > t \mid S_k = u\} &= \Pr\{S_{k+1} - S_k > t - u \mid S_k = \}, \\ &= 1 - G(t - u), \qquad u \leqq t,\end{aligned}$$

and the sum in (20) reduces to $1 - G(t - u)$.

These comments may now be made: Smith's definition of an equilibrium process explicitly assumes that the $\{S_k\}$ form a general renewal process; consequently, he can simplify the sum in (20), and replace it by $1 - G(t - u)$; thus, in the presence of the assumption that $\{S_k\}$ is a general renewal process, the condition (15) suffices to make the second term of (20) a convolution.

The analog of the representation (20) provided by our construction (Section 3) is

$$\begin{aligned}\Pr\{x_t \,\varepsilon\, A\} &= \Pr\{x_t \,\varepsilon\, A \;\&\; n(t) = 0\} \\ &+ \int_{0-}^{t} \sum_{k=1}^{\infty} \Pr\{x_t \,\varepsilon\, A \;\&\; S_{k+1} > t \mid S_k = u\}\psi_k(u)\, dH_+(u).\end{aligned} \tag{21}$$

The natural modification of. the definition of weak stationarity for a process x_t on $[0, \infty)$ is to require that the kernel

$$\sum_{k=1}^{\infty} \Pr\{x_t \,\varepsilon\, A \;\&\; S_{k+1} > t \mid S_k = u\}\psi_k(u)$$

depend on $(t - u)$ only. Although both (20) and (21) are valid representations of $\Pr\{x_t \,\varepsilon\, A\}$, it does not follow that their respective kernels agree in u a.e. $[H_+]$ for each t.

To further clarify the relationship between our representation of $\Pr\{x_t \,\varepsilon\, A\}$ and Smith's, we note that

$$\{S_k = u,\ S_{k+1} > t\} = \{n(t) = k,\ S_{n(t)} = u\}, \qquad k \geqq 1,$$

$$\Pr\{n(t) = k \mid \mathfrak{F}_k\} = \Pr\{S_{k+1} > t \mid \mathfrak{F}_k\} \quad \text{a.s.} \qquad \text{on } \{S_k \leqq t\}.$$

Then we can write our kernel as

$$\sum_{k=1}^{\infty} \Pr\{x_t \,\varepsilon\, A \ \&\ S_{k+1} > t \mid S_k = u\}\psi_k(u)$$

$$(22) \quad = \sum_{k=1}^{\infty} \Pr\{x_t \,\varepsilon\, A \mid S_k = u,\ S_{k+1} > t\}\Pr\{S_{k+1} > t \mid S_k = u\}\psi_k(u)$$

$$= \sum_{k=1}^{\infty} \Pr\{x_t \,\varepsilon\, A \mid n(t) = k,\ S_{n(t)} = u\}\Pr\{S_{k+1} > t \mid S_k = u\}\psi_k(u).$$

Smith's kernel, on the other hand, is

$$(23) \qquad \Pr\{x_t \,\varepsilon\, A \mid S_{n(t)} = u\} \sum_{k=1}^{\infty} \Pr\{S_{k+1} > t \mid S_k = u\}\psi_k(u),$$

and the first factor of (23) can be written as

$$\Pr\{x_t \,\varepsilon\, A \mid S_{n(t)} = u\} = \sum_{k=1}^{\infty} \Pr\{x_t \,\varepsilon\, A \mid n(t) = k,\ S_{n(t)} = u\} \cdot \Pr\{n(t) = k \mid S_{n(t)} = u\}.$$

If in fact the kernels (22) and (23) agree, then for all $t > 0$ a.e. $[H_+]$ on $[0, t]$ we have

$$\Pr\{x_t \,\varepsilon\, A \mid S_{n(t)} = u\} = \sum_{k=1}^{\infty} \Pr\{x_t \,\varepsilon\, A \mid n(t) = k,\ S_{n(t)} = u\}p_k(t, u)$$

where

$$p_k(t, u) = \frac{\Pr\{S_{k+1} > t \mid S_k = u\}\psi_k(u)}{\sum_{m=1}^{\infty} \Pr\{S_{m+1} > t \mid S_m = u\}\psi_k(u)}, \qquad k \geqq 1.$$

This suggests but does not prove that in some cases it may be true that for $k \geqq 1$,

$$(24) \qquad \Pr\{n(t) = k \mid S_{n(t)} = u\} = p_k(t, u) \quad \text{a.e.} \quad [H_+] \text{ on } [0, t].$$

Formula (24) is *consistent*, in that

$$\sum_{k=1}^{\infty} \Pr\{n(t) = k \mid S_{n(t)} = u\} = 1, \qquad u > 0.$$

Conversely though, if (24) holds, then the kernels agree. In the case where $\{S_k\}$ is a general renewal process, (24) is equivalent to

$$(25) \qquad \Pr\{n(t) = k \mid S_{n(t)} = u\} = \psi_k(u) \quad \text{a.e.} \quad [H_+] \text{ on } [0, t].$$

The following weaker result will suffice.

LEMMA 1. *If* $\{S_k, k \geqq 0\}$ *is a general renewal process, then the kernels* (22) *and* (23) *agree* a.e. *with respect to* H_+.

PROOF. If Y is a Borel subset of $[0, t]$, then

$$\{n(t) = k \;\&\; S_{n(t)} \in Y\} = \{n(t) = k \;\&\; S_k \in Y\}.$$

Now

$$\begin{aligned}\Pr\{n(t) = k \;\&\; S_{n(t)} \in Y\} &= \int_{\{S_{n(t)} \in Y\}} \Pr\{n(t) = k \mid \mathfrak{F}\} P(d\omega) \\ &= \int_Y \Pr\{n(t) = k \mid S_{n(t)} = u\}\, d \Pr\{S_{n(t)} \leqq u\} \\ &= \int_Y \Pr\{n(t) = k \mid S_{n(t)} = u\}[1 - G(t-u)]\, dH_+(u).\end{aligned}$$

However

$$\begin{aligned}\Pr\{n(t) = k \;\&\; S_k \in Y\} &= \int_{\{S_k \in Y\}} \Pr\{n(t) = k \mid \mathfrak{F}_k\} P(d\omega) \\ &= \int_{\{S_k \in Y\}} \Pr\{S_{k+1} > t \mid \mathfrak{F}_k\} P(d\omega) \\ &= \int_Y \Pr\{S_{k+1} > t \mid S_k = u\}\, d \Pr\{S_k \leqq u\} \\ &= \int_Y [1 - G(t-u)] \psi_k(u)\, dH_+(u).\end{aligned}$$

Since $Y \subset [0, t]$ is arbitrary, the result is proven. With all these preliminary observations we can prove Theorem 2.

THEOREM 2. *An equilibrium process* (*in the sense of Smith* [10]) *is* (*a class of*) *weakly stationary* (*events*) *on* $[0, \infty)$ *with respect to the renewal process on which it is based.*

PROOF. Lemma 1 implies that the representative kernels (22) and (23) agree. Hence Smith's condition (15) suffices to establish that (22) is a difference kernel.

The definition of a weakly stationary class of events in no way involves the statistical dependence or independence of the random variables $\{S_k - S_{k-1}\}$, and none can be deduced therefrom. The fact that H_+ has the same form as it would if the $\{S_k\}$ were sums of independent identically distributed variates is a consequence of the assumption

$$Q(X, t, u) = Q(X, t + y, u + y), \qquad \text{all } y.$$

This assumption we interpret as a *stationarity* condition because it is a statement of *invariance under translation*, and we see no hope of interpreting it as an independence condition.

We conclude that a weakly stationary class of events is a more general notion than what Smith calls an equilibrium process, because the requirement that $\{S_k\}$ be a (general) renewal process is an explicit part of the latter, and is absent in the former. The precise extent of this greater generality must be a topic for future study, but some hints of it can be obtained from examples such as are given in Section 6.

Having discussed the relation between our work and Smith's for x_t defined on $[0, t)$, we revert to our original assumptions for the proof of Theorem 3: $\{S_k, k = 0, \pm 1, \pm 2, \cdots\}$ and $\{x_t, -\infty < t < +\infty\}$. It is convenient to record next some preliminary results.

LEMMA 2. *For almost all* u $[H]$,

$$U(t - u) - Q(X, t, u)$$

is a distribution function in t, $U(\cdot)$ *being the unit step at zero, continuous from the right.*

PROOF. According to Doob [3], p. 29, Theorem 9.4, there is a conditional distribution of S_{k+1} in the wide sense, relative to $\mathfrak{F}_k$, for each k. That is, if Y is a generic Borel set on the line, there is a function p_k of Y and ω such that p_n is a probability measure of Y when ω is fixed, for fixed Y, $p_k(Y, \cdot)$ is measurable relative to $\mathfrak{F}_k$, and

$$\Pr\{S_{k+1} \varepsilon Y \mid \mathfrak{F}_k\} = p_k(Y, \omega), \qquad \text{w.p. 1.}$$

Choose $Y = (t, \infty)$; then almost surely

$$F_k(X, t, S_k(\omega)) = p_k((t, \infty), \omega),$$

and $U(t - S_k(\omega)) - F_k(X, t, S_k(\omega))$ is a distribution function in t for almost all ω. Hence $U(t - u) - F_k(X, t, u)$ is a distribution function in t for almost all u with respect to $\Pr\{S_k \leqq \cdot\}$-measure. The lemma now follows from formulas (9), (10), and (12).

LEMMA 3. *If* $Q(X, t, u) = Q(X, t - u)$, *then* $U(\cdot) - Q(X, \cdot)$ *is a distribution function. If* $Q(A, t, u) = Q(A, t - u)$, *then* $Q(A, \cdot)$ *is Borel measurable.*

PROOF. The first part is obvious if we pick a u for which $U(t - u) - Q(X, t, u)$ is a distribution function in t. For the second part, we note that

$$Q(A, t, u) = \sum_k F_k(A, 0, u - t)\varphi_k(u - t),$$

and each φ_k, $F_k(A, 0, \cdot)$ is Borel measurable. We shall denote the distribution function $U(\cdot) - Q(X, \cdot)$ by $G(\cdot)$; clearly, G vanishes for negative argument.

If α is a regular Borel measure on the line, a Borel set B is called *strongly regular* with respect to α if for every $\epsilon > 0$ there exist a compact $C \subset B$ and an open $U \supset B$ such that for all t sufficiently large[3]

$$\alpha(t - U) - \alpha(t - C) < \epsilon.$$

K_α is the class of Borel measurable functions f such that given $k > 0$ and $\epsilon > 0$

[3] $t - U$ is the set $\{t - x \mid x \varepsilon U\}$.

exist f^+ and f^- defined on $(-k, k)$, with the properties

(i) $$f^- \leqq f \leqq f^+$$

(ii) $$\int_{-k}^{k} [f^+ - f^-]\, du < \epsilon$$

(iii) f^+ and f^- are of the form[4] $\sum_{j=1}^{n} b_j \chi_{B_j}$ where the B_j are Borel sets and strongly regular with respect to α.

It is shown in [1] that if α is positive and satisfies the hypothesis of Wiener's theorem (Widder [12], p. 214), then $g \in K_\alpha$ implies

$$\int_{-\infty}^{\infty} g(t-u)\, d\alpha(u) \to A \int_{-\infty}^{\infty} g(u)\, du, \qquad A \geqq 0,$$

provided that

$$\sum_n \sup_{n \leqq x < (n+1)} |g(x)| < \infty.$$[5]

5. The Principal Result.

THEOREM 3. *If*

(i) *y_t is the time from t to the next S_k after t, i.e.,* $y_t = \min\{S_k - t \mid S_k > t\}$,

(ii) *the two classes of events*

$$\{y_u < \infty\}, \qquad -\infty < u < \infty \qquad (\text{or } \{x_u \in X\},\ -\infty < u < \infty)$$
$$\{x_u \in A\}, \qquad -\infty < u < \infty,$$

are both weakly stationary with respect to $\{S_k\}$, *with the (respective) representative kernels* $Q(X, \cdot)$ *and* $Q(A, \cdot)$,

(iii) *The* $\{S_k\}$ *are "aperiodic" in the sense that* $U(\cdot) - Q(X, \cdot)$ *is not a lattice distribution, or that the Fourier-Stieltjes transform of* $Q(X, \cdot)$ *vanishes at zero only,*

(iv) $H(\cdot + 1) - H(\cdot)$ *is bounded,*

then, as $t \to \infty$

$$\Pr\{x_t \in A\} \to \frac{\int_0^\infty Q(A, u)\, du}{\int_0^\infty Q(X, u)\, du},$$

provided that $Q(A, \cdot) \in K_H$ *and*

(26) $$\int_0^\infty Q(X, u)\, du < \infty.$$

If (26) *is not true, the restriction that* $Q(A, \cdot)$ *belong to* K_H *is unnecessary, and*

$$\Pr\{x_t \in A\} \to 0 \quad \text{as} \quad t \to \infty$$

---

[4] χ_B is the characteristic function of B.

[5] The class of g satisfying this condition is designated by N in [1].

provided that $Q(A, \cdot) \varepsilon N$. Finally, if $Q(X, \cdot)$ is not in L_1, and if

$$k(\cdot) = Q(X, \cdot) - Q(A, \cdot)$$

belongs to N, then

$$\Pr\{x_t \varepsilon A\} \to 1 \quad \text{as} \quad t \to \infty.$$

PROOF. From (11), it is easy to see that

$$\begin{aligned} 1 = \Pr\{y_t < \infty\} &= \int_{-\infty}^{t} \sum_k \Pr\{S_{k+1} > t \mid S_k = u\}\, \varphi_k(u)\, dH(u) \\ &= \int_{-\infty}^{t} Q(X, t-u)\, dH(u). \end{aligned} \tag{27}$$

Assume first that (26) holds; since $Q(X, \cdot)$ is bounded, non-increasing in $(0, \infty)$, and zero for negative argument, it follows that both $Q(X, \cdot)$ and $Q(A, \cdot)$ belong to N, i.e.,

$$\sum_{n=-\infty} \sup_{n<x<(n+1)} |Q(B, x)| < \infty, \qquad B = X \quad \text{or} \quad A.$$

The first part of Theorem 2 now follows, in view of (iv) and (14), by an extension of Wiener's general Tauberian theorem, Theorem 10 of [1]. When (26) is not true, the arguments become simpler. We set

$$H_+(t) = \begin{cases} H(t) - H(0-) & t \geqq 0 \\ 0 & t < 0 \end{cases}$$

$$H_-(t) = \begin{cases} 0 & t \geqq 0 \\ H(0-) - H(t) & t < 0 \end{cases}$$

$$\xi(t) = \begin{cases} \int_{-\infty}^{0-} Q(X, t-u)\, dH(u) & t \geqq 0 \\ 0 & t < 0 \end{cases}$$

$$F(t) = U(t) - \xi(t),$$

$$G(t) = U(t) - Q(X, t).$$

The equation (27) may be written as two equations, each valid for all t:

$$U(t) = \xi(t) + \int_{0-}^{t} Q(X, t-u)\, dH_+(u) \tag{28}$$

$$U(-t) = \int_{-\infty}^{t} Q(X, t-u)\, dH_-(u) - \xi(t). \tag{29}$$

Let $\epsilon > 0$ be given, and choose $t_0 > 0$ and T so that

$$\int_{-\infty}^{-T} Q(X, t_0 - u)\, dH(u) < \epsilon.$$

Since $Q(X, \cdot)$ is monotone nonincreasing in $(0, \infty)$, we have

$$\int_{-\infty}^{-T} Q(X, t-u)\, dH(u) < \epsilon, \qquad t > t_0.$$

However

$$\int_{-T}^{0-} Q(X, t-u)\, dH(u) \leqq Q(X, t) \int_{-T}^{0-} dH \to 0 \quad \text{as} \quad t \to \infty.$$

Hence, since ξ is nonincreasing in $(0, \infty)$, we have

$$\xi(t) \to 0 \quad \text{as} \quad t \to \infty, \tag{30}$$

and so F is the distribution function of a nonnegative variate.

It is obvious that $Q(X, t) \geqq Q(A, t)$ for all t. Hence, writing

$$\Pr\{x_t \, \varepsilon \, A\} = \int_{-\infty}^{0-} Q(A, t-u)\, dH_-(u) + \int_{0-}^{t} Q(A, t-u)\, dH_+(u),$$

it follows from (30) that the first term on the right goes to zero with increasing t. The limit of $\Pr\{x_t \, \varepsilon \, A\}$, if it exists, depends on H only through H_+.

From (iv) we have $H_+(t) = O(e^{\gamma t})$ as $t \to \infty$, for any $\gamma > 0$, and so the Laplace-Stieltjes transform

$$H_+^*(s) = \int_{0-}^{\infty} e^{-st}\, dH_+(t)$$

converges in $\mathrm{Re}(s) > 0$. It is clear that the respective transforms $F*$ and $G*$ converge in the same region. From (28) we find that

$$H_+^*(s) = \frac{F^+(s)}{1 - G*(s)},$$

so that H_+ turns out to be a slight modification of the renewal function of Smith [9], with the properties

$$H_+ = F*L$$

$$L = G + G*G + G*G*G + \cdots$$

$$L(x+y) - L(x) \leqq 1 + L(y), \qquad x, y > 0,$$

so that also

$$H_+(x+y) - H_+(x) \leqq 1 + L(y), \qquad x, y > 0.$$

None of the preceding remarks depends on the failure of condition (26).

Now if (26) is false, then $\int_0^\infty x\, dG(x) = \infty$, and we proceed as follows: let n be a large integer; for t much larger than n, since $0 \leqq Q(A, t-u) \leqq 1$ and $Q(A, \cdot) \, \varepsilon \, N$, we have

$$\int_{0-}^{t} Q(A, t-u)\, dH_+(u) \leqq H_+(t) - H_+(t-n)$$
$$+ [1 + L(1)] \sum_{k=n}^{\infty} \sup_{k \leqq x < k+1} |Q(A, x)|.$$

By Blackwell's renewal theorem [2] the first term on the right tends to zero as $t \to \infty$, because the mean renewal lifetime is infinite. The second term can be made arbitrarily small by a large enough[6] choice of n. Hence if (26) fails and $Q(A, \cdot) \in N$,

$$\Pr\{x_t \in A\} \to 0 \quad \text{as} \quad t \to \infty.$$

Finally, if $k(\cdot) = Q(X, \cdot) - Q(A, \cdot)$ is in N, we may use Theorem 1 of [1] again to conclude that

$$1 - \Pr\{x_t \in A\} = \int_{-\infty}^{t} k(t-u)\, dH(u) \to 0 \quad \text{as} \quad t \to \infty.$$

With minor changes, the preceding derivations all apply if x_t is defined for $0 \leqq t < \infty$ only, and the $\{S_k\}$ are positive and defined for $k = 0, 1, 2, \cdots$ only, with $S_0 = 0$. Then H can be taken to coincide with H_+, ξ is identically zero, and (29) is dropped.

6. Construction of examples. As has been pointed out, processes possessing regeneration points give rise to weakly stationary classes of events. However, it is of interest whether there are classes of events that are weakly stationary but are not based in any, or in any obvious, way upon regeneration points. The construction[7] to be given shows that any process can be used for defining a weakly stationary class of events $\{y_u < \infty\}$, $-\infty < u < \infty$.

We have shown that

$$Q(X, t, u) = \sum_{K} \Pr\{S_{k+1} > t \mid S_k = u\}\varphi_k(u),$$

where the $\varphi_k(u)$ may be taken to form a discrete probability distribution in n for each u. In order that $Q(X, \cdot, \cdot)$ be a difference kernel, it is sufficient (though not necessary), that

$$\Pr\{S_{k+1} > t \mid S_k = u\}$$

depend on $(t - u)$ only. Let z_k, $k = 0, 1, 2, \cdots$ be an arbitrary stochastic process which will represent the relevance of the past to $S_{k+1} - S_k = x_{k+1}$. The distance x_{k+1} between S_k and S_{k+1} will be chosen in a manner depending on z_h in such a way that knowledge of S_k is *irrelevant* to x_{k+1}, i.e.,

$$\Pr\{x_{k+1} \leqq w \mid \mathfrak{F}_k\}$$

is a.e. a constant depending on w. For simplicity allow z_k to take only the values 0 and 1.

Let G be any distribution function with $G(0) = 0$. To avoid complications, let all the mass of G be on a countable set. If z_k and S_k are known, choose $x_{k+1} =$

[6] The referee suggested this simple argument. The author's original version depended on Theorem 6 of Pitt [8] by being a direct adaption of case (ii), Theorem 1 of Smith [9], followed by an application of Theorem 1 of [1].

[7] This construction is based on ideas discussed by the author with J. L. Snell and A. J. Fabens.

$S_{k+1} - S_k$ independently with distribution

$$G_i(k+1, u, \cdot) \quad \text{if} \quad z_k = i, \qquad S_k = u, \qquad i = 0 \text{ or } 1, \tag{31}$$

where G_i is defined as follows: set

$$q_k(u) = \inf \{x \mid G(x) > \Pr \{z_k = 1 \mid S_k = u\}\}$$

$$G_0(k+1, u, w) = \begin{cases} 0, & w < q_k(u) \\ \dfrac{G(w) - \Pr\{z_k = 1 \mid S_k = u\}}{1 - \Pr\{z_k = 1 \mid S_k = u\}}, & w \geqq q_k(u), \end{cases}$$

$$G_1(k+1, u, w) = \begin{cases} \dfrac{G(w)}{\Pr\{z_k = 1 \mid S_k = u\}}, & w < q_k(u) \\ 1, & w \geqq q_k(u). \end{cases}$$

A verbal description of the above construction is as follows: If $S_k = u$ set

$$q_k(u) = \{\text{the least } x \text{ for which } G(x) > \Pr\{z_k = 1 \mid S_k = u\}\}.$$

Then, if $z_k = 0$ we sample x_{n+1} from the distribution $G(\cdot)$, but conditional on being greater than $q_k(u)$; if $z_k = 1$, x_{n+1} is sampled from $G(\cdot)$, but conditional on being less than $q_k(u)$. It is evident that

$$\Pr\{x_{k+1} \leqq w \mid S_k = u\} = G(w)$$

independently of k and u, and hence

$$Q(X, t, u) = G(t - u), \tag{32}$$

so that setting $S_0 = 0$ for definiteness, the events $\{y_u < \infty\}$, $-\infty < u < \infty$, are weakly stationary.

The given construction works for any process z_k. However, it is possible to make the 'relevance of the past,' the process z_k, depend explicitly on past S_k's or x_k's. For instance, we can define a two-dimensional process $\{z_k, x_k\}$, as follows: let $p(\cdot)$ be any function with range in [0, 1]; choose x_1 arbitrarily, and set

$$\Pr\{z_k = 1\} = p(x_k)$$

$$\Pr\{z_k = 0\} = 1 - p(x_k)$$

independently of everything in the past except x_k. If z_k and $S_k = \sum_{j=1}^{k} x_j$ are known, choose x_{k+1} independently of everything except z_k and S_k with the conditional distribution (31). Formula (32) is still true, but now x_{k+1} depends explicitly on x_k via z_k.

REFERENCES

[1a] V. E. Beneš, "General stochastic processes in traffic systems with one server," *Bell System Tech. J.*, Vol. 39 (1960), pp. 127–160.

[1] V. E. Beneš, "Extensions of Wiener's Tauberian Theorem for positive measures," *J. Math. Anal. and Appl.*, Vol. 2 (1961), pp. 1–20.

[2] D. Blackwell, "A renewal theorem," *Duke Math. J.*, Vol. 15 (1948), pp. 145–150.

[3] J. L. Doob, *Stochastic Processes*, John Wiley and Sons, New York, 1950.
[4] W. Feller, "On the integral equation of renewal theory," *Ann. Math. Stat.*, Vol. 12 (1941), pp. 243–267.
[5] P. R. Halmos, *Measure Theory*, Van Nostrand, Princeton, 1950.
[6] P. Lévy, "Processus semi-markoviens," *Proc. Internat. Congr. Math.*, Vol. 3 (1954), pp. 416–426.
[7] M. Loève, *Probability Theory*, 1st. ed., Van Nostrand, New York, 1955.
[8] H. R. Pitt, "General Tauberian Theorems II," *J. London Math. Soc.*, Vol. 15 (1940), pp. 97–112.
[9] W. L. Smith, "Asymptotic renewal theorems," *Proc. Roy. Soc. Edinburgh*, Ser. A, Vol. 64 (1954), pp. 9–48.
[10] W. L. Smith, "Regenerative stochastic processes," *Proc. Roy. Soc. London*, Ser. A, Vol. 232 (1955), No. 1188 (11 October 1955), pp. 6–31.
[11] W. L. Smith, "Renewal theory and its ramifications," *J. Roy. Stat. Soc.*, Ser. B, Vol. 20 (1958), pp. 243–302.
[12] D. V. Widder, *The Laplace Transform*, Princeton, 1941.