

RENEWAL THEORY FROM THE POINT OF VIEW OF THE THEORY OF PROBABILITY

BY
J. L. DOOB

Renewal theory is ordinarily reduced to the theory of certain types of integral equations. Since the basis for the integral equations is a simple probability process, however, it is to be expected that a treatment in terms of the theory of probability, which uses the modern developments of this theory, will shed new light on the subject. The purpose of this paper is to present such a treatment.

Renewal theory, in the simplest case, deals with the following situation. A population of individuals is at hand; when any individual dies he is immediately replaced by a newly born individual. The problem is to investigate the development of the population, particularly the age distribution. Since the individuals are supposed to live and die independently of each other it is sufficient for most purposes to consider a population consisting at any time of only a single individual. At birth, the future lifetime of the individual is a chance variable with distribution function $F(x)$ (the probability of death before age x). At any age x , the lifetime remaining then has distribution function

$$F_x(y) = \frac{F(x+y) - F(x)}{1 - F(x)}.$$

We shall always assume that the individual has a chance for survival, that is that $F(0+) < 1$. The course of the population can be described as follows: at $t=0$ the individual is of age x_0 , where x_0 is a chance variable with distribution function $\Phi(x)$; the individual lives a further time x_1 , dying at age x_0+x_1 ; his successor lives to age x_2 , the next to age x_3 , and so on. The chance variables x_1, x_2, \dots are mutually independent:

x_1 has distribution function $F_{x_0}(x)$;

x_j has distribution function $F(x)$, if $j > 1$.

Every problem of renewal theory must thus be expressible in terms of such a sequence of chance variables, and all the standard theorems relating to stationary sequences of independent chance variables must have significance for renewal theory.

Let $n(t)$ be the number of sums

$$x_1, x_1 + x_2, \dots$$

which are less than t , that is, the number of deaths which occur before time t .

For each t , $n(t)$ is a chance variable. Let $x(t)$ be the age of the population at time t :

$$\begin{aligned} x(t) &= t - [x_1 + \cdots + x_{n(t)}] & (\text{if } n(t) \geq 1) \\ &= x_0 + t & (n(t) = 0). \end{aligned}$$

For each t , $x(t)$ is a chance variable, and renewal theory becomes the study of this $x(t)$ process. The theoretical problems of renewal theory center about the asymptotic properties of $x(t)$ for large t . In particular it is important to ascertain under what conditions the $x(t)$ process is asymptotically stationary ($t \rightarrow \infty$) and asymptotically independent of the initial distribution Φ . The latter properties are usually formulated as follows. Define $U_x(t)$ as the expected number of deaths before time t , under the hypothesis that the initial age x_0 is x ,

$$U_x(t) = E\{x_0 = x; n(t)\}.$$

Then the above properties are interpreted to mean that $U'_0(t)$, the probability density of a death at time t (if $x_0 = 0$), converges to a limit when $t \rightarrow \infty$.

Taking up the problem from a more general point of view, Feller showed [2, p. 253]⁽¹⁾ that $\lim_{t \rightarrow \infty} U_0(t)/t$ exists in all cases. Täcklind [4; 5] has obtained accurate asymptotic expressions for $U_0(t)$ under various restrictions on the moments of order one and higher of the lifetime distribution $F(x)$. In the present paper it will be shown that Feller's result follows from the law of large numbers for mutually independent chance variables with a common distribution function, and that the theory of Markoff processes can be used to extend Täcklind's results using hypotheses of an entirely different type.

In the following, $\Phi(x)$ will always be the distribution function of the initial age x_0 , $F(x)$ will be the distribution function of lifetime, that is, the distribution function of x_2, x_3, \cdots , and $U(t) = E\{n(t)\}$ will be the expected number of deaths before time t . In particular if the initial distribution is concentrated at x , we write $U_x(t)$ for $U(t)$. It will be seen in Theorem 1 that $U(t)$ is always finite-valued. Finally we shall always write Δ as the mean lifetime,

$$\Delta = \int_0^\infty x dF(x);$$

Δ will usually, but not always, be supposed finite.

THEOREM 1. *Let x_1, x_2, \cdots be mutually independent chance variables. It is supposed that x_2, x_3, \cdots have a common distribution function and that*

$$(1.1) \quad \begin{aligned} x_1 &\geq 0, & x_2 &\geq 0, \\ 0 < E\{x_2\} &= \Delta \leq \infty. \end{aligned}$$

⁽¹⁾ Numbers in brackets refer to the bibliography at the end of the paper.

Let $n(t)$ be the number of sums x_1, x_2, \dots which are less than t . Then

$$(1.2) \quad \lim_{t \rightarrow \infty} \frac{n(t)}{t} = \frac{1}{\Delta} \quad (2).$$

Moreover $E\{n(t)\} < \infty$ and

$$(1.3) \quad \lim_{t \rightarrow \infty} \frac{E\{n(t)\}}{t} = \frac{1}{\Delta}.$$

In other words, in the notation of renewal theory, if $F(0+) < 1$, $U(t)$ is always finite and

$$(1.4) \quad \lim_{t \rightarrow \infty} \frac{U(t)}{t} = \frac{1}{\Delta}.$$

This theorem was first proved by Feller [2, p. 253] using Laplace transform methods. It is treated here only to show its close relationship to the strong law of large numbers.

Suppose first that $\Delta = \infty$, and that the theorem has already been proved for finite Δ . Define $x_j^{(N)}(t)$ to be x_j for $x_j \leq N$ and 0 otherwise. Then if $n^{(N)}(t)$ is defined as the number of sums of truncated variables less than t , $n(t) \leq n^{(N)}(t)$. Hence, since the theorem is supposed true for the truncated variables,

$$(1.5) \quad \limsup_{t \rightarrow \infty} \frac{n(t)}{t} \leq \lim_{t \rightarrow \infty} \frac{n^{(N)}(t)}{t} = \frac{1}{E\{x_2^{(N)}\}},$$

with probability 1, and

$$(1.6) \quad \limsup_{t \rightarrow \infty} \frac{E\{n(t)\}}{t} \leq \lim_{t \rightarrow \infty} \frac{E\{n^{(N)}(t)\}}{t} = \frac{1}{E\{x_2^{(N)}\}}.$$

When $N \rightarrow \infty$ the denominator $E\{x_2\}$ in (1.5) and (1.6) becomes infinite, giving the desired result. In the following, therefore, it will be sufficient to suppose that Δ is finite.

According to the strong law of large numbers, and using the fact that $\lim_{n \rightarrow \infty} x_1/(n-1) = 0$ with probability 1,

$$(1.7) \quad \lim_{n \rightarrow \infty} \frac{x_2 + \dots + x_n}{n-1} = \lim_{n \rightarrow \infty} \frac{x_1 + \dots + x_n}{n-1} = \lim_{n \rightarrow \infty} \frac{x_1 + \dots + x_n}{n} = \Delta$$

with probability 1. Hence if $0 < \epsilon < \Delta$,

$$(1.8) \quad (\Delta - \epsilon)n \leq x_1 + \dots + x_n \leq (\Delta + \epsilon)n$$

for sufficiently large n , with probability 1. In particular if $n = n(t)$ the left side of (1.8) implies

(2) Here and in the following we interpret $1/\Delta$ as 0 if $\Delta = \infty$.

$$n(t) \leq \frac{t}{\Delta - \epsilon}$$

and if $n = n(t) + 1$, the right side implies

$$n(t) + 1 \geq \frac{t}{\Delta + \epsilon}.$$

Hence for large t

$$(1.9) \quad \frac{1}{\Delta + \epsilon} - \frac{1}{t} \leq \frac{n(t)}{t} \leq \frac{1}{\Delta - \epsilon},$$

with probability 1. This implies (1.2), and (1.3) will follow when term by term integration is justified.

Before going further with the general case we note the results for the following particular case.

$$(1.10) \quad \begin{aligned} P\{x_1 = 0\} &= 1, \\ P\{x_j = 1\} &= p > 0, \\ P\{x_j = 0\} &= 1 - p. \end{aligned} \quad j > 1,$$

In this case if t is an integer, $n(t+1) - n(t)$ is the number of times the sums take on the value t , so that for all t , if $[t]$ is the largest integer less than t ,

$$(1.11) \quad \frac{E\{n(t)\}}{t} = \frac{1}{t} \sum_{j < t} E\{n(j+1) - n(j)\} = \frac{[t] + 1}{tp},$$

which implies (1.3) for this special case. Moreover

$$(1.12) \quad E\left\{\left[\frac{n(t)}{t}\right]^2\right\} = \frac{[t] + 1}{t^2} \cdot \frac{1 - p}{p^2} + \left(\frac{[t] + 1}{tp}\right)^2.$$

Now returning to the general case, since $P\{x_2 = 0\} < 1$, there is a positive number λ such that $P\{x_j \geq \lambda\} > 0$. Replacing x_j by x_j/λ if necessary, which does not alter the significance of the theorem, we can suppose that $P\{x_j \geq 1\} = p > 0$. Now define \bar{x}_j by

$$(1.13) \quad \begin{aligned} \bar{x}_1 &\equiv 0, \\ \bar{x}_j &= 0 \quad \text{if } x_j < 1, \\ \bar{x}_j &= 1 \quad \text{if } x_j \geq 1. \end{aligned} \quad j \geq 2,$$

Then if $\bar{n}(t)$ is the number of times the sums $\bar{x}_1, \bar{x}_1 + \bar{x}_2, \dots$ are less than t , it is clear that $n(t) \leq \bar{n}(t)$. We have, using (1.12),

$$(1.14) \quad \limsup_{t \rightarrow \infty} E\left\{\left[\frac{n(t)}{t}\right]^2\right\} \leq \limsup_{t \rightarrow \infty} E\left\{\left[\frac{\bar{n}(t)}{t}\right]^2\right\} = \frac{1}{p^2}.$$

The boundedness of these second moments at $t = \infty$ implies that (1.2) can be integrated term by term, giving (1.3).

This theorem is a first step in the investigation of the asymptotic stationarity of the $x(t)$ process; there is an average long run frequency of births and deaths, independent of the initial conditions, with no restrictive hypotheses whatever (except that $F(0+) < 1$ so that there is a chance for temporary survival).

The explicit evaluation of $U(t)$ is furnished by the following corollary. We use the convolution notation

$$F \star G = \int_0^x F(x-y) dG(y),$$

for distribution functions and

$$f \star g = \int_0^x f(x-y) g(y) dy$$

for density functions. In general we shall use small letters for derivatives, so that $f = F'$, and so on.

COROLLARY 1.1. *If $F(0+) < 1$, the function $U(t)$ is always finite-valued and is expressed in terms of the initial age distribution function $\Phi(x)$ and the lifetime distribution function $F(x)$ by means of the equation*

$$(1.15) \quad U(t) = \Phi_1 + \Phi_1 \star F + \Phi_1 \star F \star F + \dots$$

where

$$(1.16) \quad \Phi_1(t) = \int_0^\infty F_\lambda(t) d\Phi(\lambda) = \int_0^\infty \frac{F(\lambda+t) - F(\lambda)}{1 - F(\lambda)} d\Phi(\lambda).$$

In particular

$$(1.17) \quad \begin{aligned} U_x &= F_x + F_x \star F + F_x \star F \star F + \dots, \\ U_0 &= F + F \star F + F \star F \star F + \dots. \end{aligned}$$

In these equations the infinite series converge for all values of the arguments.

If $U_0(t)$ is absolutely continuous, $F(t)$ is also absolutely continuous, and conversely if $F(t)$ is absolutely continuous, $U_x(t)$ is absolutely continuous in t for all x and (1.15) and (1.17) can be differentiated term by term,

$$(1.18) \quad \begin{aligned} u(t) &= \phi_1 + \phi_1 \star f + \phi_1 \star f \star f + \dots, \\ u_x(t) &= f_x + f_x \star f + f_x \star f \star f + \dots, \quad \phi_1(t) = \int_0^\infty \frac{f(\lambda+t)}{1 - F(\lambda)} d\Phi(\lambda), \\ u_0(t) &= f + f \star f + f \star f \star f + \dots. \end{aligned}$$

To prove the corollary, we note first that the n th term of (1.15) is the

distribution function of the sum $x_1 + \cdots + x_n$. Hence

$$\begin{aligned}
 U(t) &= E\{n(t)\} = \sum_{v=1}^{\infty} P\{n(t) \geq v\} \\
 (1.19) \quad &= \sum_{v=1}^{\infty} P\{x_1 + \cdots + x_n < t\} \\
 &= \Phi_1 + \Phi_1 \star F + \Phi_1 \star F \star F + \cdots,
 \end{aligned}$$

where all the series converge, since $U(t)$ exists. This proves (1.15), and thereby the particular case (1.17). The fact that the series converge is also easily deduced by elementary methods.

If $U_0(t)$ is absolutely continuous, that is, if its increase on any set of Lebesgue measure zero is zero, (1.17) shows that the same is true of $F(t)$, whose increase on any set is at most that of $U_0(t)$. Conversely if $F(t)$ is absolutely continuous, its convolution with any other distribution function is also absolutely continuous, and according to Fubini's series theorem $U_0(t)$ is absolutely continuous, the series can be differentiated term by term, and the equations (1.18) hold. The equations hold almost everywhere, or everywhere if the left sides are defined as the sums on the right. This completes the proof.

The following further corollary will be useful below.

COROLLARY 1.2. *There is a number K depending on c but neither on t nor on the initial distribution Φ , such that*

$$(1.20) \quad E\{[n(t+c) - n(t)]^2\} \leq K \quad (c > 0).$$

It has already been proved that $E\{n(t)^2\} < \infty$, and this implies the finiteness of the left side of (1.20). Moreover if $c > 0$,

$$\begin{aligned}
 E\{[n(t+c) - n(t)]^2\} &= E\{E\{x(t) = x; [n(t+c) - n(t)]^2\}\} \\
 &= E\{E\{x_0 = x; n(c)^2\}\} \\
 (1.21) \quad &\leq E\{E\{x_0 = 0; [n(c) + 1]^2\}\} \\
 &= E\{x_0 = 0; [n(c) + 1]^2\},
 \end{aligned}$$

and K can be taken as this last conditional expectation.

THEOREM 2. *Let $\{x(t)\}$ be the variables of any renewal process with $0 < \Delta < \infty$. Then*

$$(2.1) \quad \lim_{t \rightarrow \infty} \frac{x(t)}{t} = 0$$

with probability 1, and if in addition the initial age has a finite expectation, $E\{x(0)\} < \infty$, it follows that

$$(2.2) \quad \lim_{t \rightarrow \infty} \frac{E\{x(t)\}}{t} = 0.$$

In fact for large t , if $s_r = x_1 + \cdots + x_r$, $x(t) = t - s_{n(t)}$ so that

$$(2.3) \quad \frac{x(t)}{t} = 1 - \frac{s_{n(t)}}{n(t)} \cdot \frac{n(t)}{t}.$$

Now according to the strong law of large numbers, as applied to the chance variables x_2, x_3, \cdots which are independent and have a common distribution function,

$$(2.4) \quad \Delta = \lim_{n \rightarrow \infty} \frac{x_2 + \cdots + x_n}{n-1} = \lim_{n \rightarrow \infty} \frac{x_1 + \cdots + x_n}{n-1} = \lim_{t \rightarrow \infty} \frac{s_{n(t)}}{n(t)}$$

with probability 1. Combining this with (1.2) and (2.3) gives the desired result. Since

$$x(t) \leq t + x_0$$

where x_0 is the initial age,

$$\frac{x(t)}{t} \leq 1 + \frac{x_0}{t} \leq 1 + x_0 \quad (t \geq 1).$$

Hence if $E\{x(0)\} < \infty$, $x(t)/t$ is dominated by an integrable function independent of t , and (2.1) can be integrated to the limit to get (2.2).

The fundamental $x(t)$ process is a Markoff process, since if $x(t_0)$ is known the conditional probabilities for $x(t)$ when $t > t_0$ are certainly unaffected by knowledge of $x(s)$ for $s < t_0$. The transition probabilities are stationary, that is $P\{x(s); x(t) < \lambda\}$ depends only on λ and $t-s$, and are easily written explicitly:

$$\begin{aligned} P\{x(0) = x; x(t) \geq \lambda\} &= 0 & (t + x < \lambda) \\ &= \frac{1 - F(x+t)}{1 - F(x)} & (t < \lambda < t+x) \\ &= \frac{1 - F(x+t)}{1 - F(x)} + \int_0^{t-\lambda} [1 - F(t-s)] dU_x(s) & (\lambda < t). \end{aligned} \quad (3.1)$$

The first two evaluations are obvious. The first term in the third evaluation represents the probability of survival in the complete time interval $(0, t)$. The second term is the probability of attaining age λ or greater after one or more deaths. The integral is clearly formally correct, with s interpreted as the time of the last death before t . It can be justified by pursuing the probability approach further to make the intuitive argument rigorous, or as follows. From (1.17) with $x_1 = x$

$$\begin{aligned}
 & \int_0^{t-\lambda} [1 - F(t-s)] dU_x(s) \\
 (3.2) \quad &= \sum_{r=1}^{\infty} \int_0^{t-\lambda} [1 - F(t-s)] dP\{x_1 + \cdots + x_r < s\} \\
 &= \sum_{r=1}^{\infty} \int_0^{t-\lambda} P\{x_{r+1} \geq t-s\} dP\{x_1 + \cdots + x_r < s\} \\
 &= \sum_{r=1}^{\infty} P\{x_1 + \cdots + x_r < t-\lambda, x_1 + \cdots + x_{r+1} \geq t\},
 \end{aligned}$$

and this last expression is the probability that $x(t) \geq \lambda$ after passing through one or more zeroes. More generally, if the distribution of $x(0)$ is given by Φ , the $x(t)$ probabilities are evaluated by substituting (3.1) in

$$(3.3) \quad P\{x(t) \geq \lambda\} = \int_0^{\infty} P\{x(0) = x; x(t) \geq \lambda\} d\Phi(x).$$

The $x(t)$ process will now be investigated in more detail, from a purely qualitative point of view. Suppose that this process is asymptotically stationary, so that the distribution function of $x(t)$, the age at time t , converges to a limiting distribution function when $t \rightarrow \infty$. Then it is to be expected that if this distribution is chosen as the initial age distribution the resulting process will actually be stationary, that is, the distribution of age will be independent of t . The investigation of all possible stationary processes corresponding to the given transition probabilities is therefore in order. The following theorem, which is well known in various forms and degrees of generality, answers this question completely.

THEOREM 4. *A renewal process with $F(0+) < 1$ is stationary if and only if the average lifetime is finite,*

$$(4.1) \quad 0 < \Delta = \int_0^{\infty} \lambda dF(\lambda) < \infty$$

and if the initial distribution $\Phi(x)$ for $x(0)$ is absolutely continuous, with density Φ' given by

$$(4.2) \quad \Phi'(x) = \frac{1 - F(x)}{\Delta}.$$

If $\Delta < \infty$, (4.2) defines a function Φ' which is non-negative and has integral 1. Suppose that $x(0)$ has density of distribution given by (4.2) and that $U(t) = E\{n(t)\}$ as usual. Then (1.16) yields $\Phi_1 = \Phi$, and (1.15) yields $U(t) = t/\Delta$. Finally $P\{x(t) \geq \lambda\}$ can be evaluated using (3.1) and (3.3); putting in $U(t) = t/\Delta$ yields

$$(4.3) \quad P\{x(t) \geq \lambda\} = 1 - \Phi(\lambda) = P\{x(0) \geq \lambda\}.$$

This shows that the $x(t)$ process with the initial distribution density (4.2) is stationary. Conversely suppose that the initial distribution Φ is chosen to make the $x(t)$ process stationary. Then $U(s+t) - U(s)$ is the expected number of zeros in the interval $[s, t)$ which is the same as the number expected in the interval $[0, t)$,

$$(4.4) \quad U(s+t) = U(s) + U(t).$$

Since $U(t)$ is obviously monotone, the only solution to (4.5) is $U(t) = ct$ for some constant $c \geq 0$; $c=0$ is impossible, since it implies that $n(t)$ is never positive. The constant c must be $1/\Delta$ by Theorem 1; hence $\Delta < \infty$. If $t > \lambda$, we have, using (3.3),

$$(4.5) \quad \begin{aligned} 1 - \Phi(\lambda) &= P\{x(0) \geq \lambda\} \\ &= \int_0^\infty \frac{1 - F(x+t)}{1 - F(x)} d\Phi(x) + \frac{1}{\Delta} \int_0^{t-\lambda} [1 - F(t-s)] ds \\ &= \int_0^\infty \frac{1 - F(x+t)}{1 - F(x)} d\Phi(x) + \frac{1}{\Delta} \int_\lambda^t [1 - F(s)] ds. \end{aligned}$$

When $t \rightarrow \infty$ the first integrand converges boundedly to 0, so that

$$(4.6) \quad 1 - \Phi(\lambda) = \frac{1}{\Delta} \int_\lambda^\infty [1 - F(s)] ds$$

as was to be proved.

The theoretical basis of the present paper is now complete. The renewal process is a Markoff process with known transition probabilities, stationary if the initial distribution is chosen suitably. Before proceeding to study the process, however, the connection with integral equations will be given. It is clear from Corollary 1.1 that $U(t)$ satisfies the equation

$$(5.1) \quad U(t) = \Phi_1(t) + \int_0^t U(t-s) dF(s).$$

In particular

$$(5.2) \quad U_x(t) = F_x(t) + \int_0^t U_x(t-s) dF(s)$$

and ($x=0$)

$$(5.3) \quad U_0(t) = F(t) + \int_0^t U_0(t-s) dF(s).$$

Here Φ_1 given by (1.16) is the distribution function of x_1 , the time to the first

death. If $F(t)$ is absolutely continuous, we have seen that $U_x(t)$ is also, and (5.2) and (5.3) become

$$(5.2') \quad u_x(t) = \frac{f(x+t)}{1-F(x)} + \int_0^t u_x(t-s)f(s)ds,$$

$$(5.3') \quad u_0(t) = f(t) + \int_0^t u_0(t-s)f(s)ds.$$

There is a similar density version of (5.1). These equations hold for almost all t , or for all t , if the left sides of (5.2') and (5.3') are properly defined. If Δ is finite and if the distribution of $x(0)$ is absolutely continuous, with density $1-F(\lambda)\Delta$, $U(t)$ reduces to t/Δ , and $u(t)=1/\Delta$. Conversely if $U(t)=\text{const.} \cdot t$, the constant must be $1/\Delta$ by Theorem 1. If this is substituted in (5.1), Φ_1 can be evaluated; the latter function, just as in the stationary case, is absolutely continuous, with derivative $[1-F(\lambda)]/\Delta$. But Φ_1 and F determine the $x(t)$ process completely; hence the process must be stationary and $\Phi=\Phi_1$. This argument shows that $U(t)$ is proportional to t only if the process is stationary.

From the present point of view, the fact that $U(t)$ is always finite is an existence theorem for the solution of an integral equation. We state the existence theorem only for (5.3) and (5.3'), to give the principle involved.

THEOREM 6. *Let $F(s)$ be any distribution function with $F(s)=0$ for $s<0$, and $F(0+)<1$. Then there is a monotone nondecreasing function $U_0(t)$ which satisfies (5.3). Moreover if $\tilde{U}_0(t)$ is any function defined and measurable for $t\geq 0$, and bounded in every finite interval, which satisfies (5.3), $\tilde{U}_0(t)\equiv U_0(t)$.*

Let $f(t)$ be any non-negative Lebesgue measurable function, defined for $t\geq 0$, and integrable. Then there is a non-negative Lebesgue measurable function $u_0(t)$, integrable over every finite interval, which satisfies (5.3'). Moreover if $\tilde{u}_0(t)$ is any function defined and Lebesgue measurable for $t\geq 0$, and integrable over every finite interval which satisfies (5.3'), $\tilde{u}_0(t)\equiv u_0(t)$.

The existence proofs have already been given. The uniqueness proofs are carried through as follows. If U_0 satisfies the stated conditions, then clearly

$$(6.1) \quad \begin{aligned} \tilde{U}_0(t) &= F + \tilde{U}_0 \star F = F + F \star F + \tilde{U}_0 \star F \star F \\ &= F + {}_2F + \cdots + {}_nF + \tilde{U}_0 \star {}_nF \end{aligned}$$

for every n , where ${}_nF$ is the n th iterated convolution of F . The series of convolutions is convergent to U_0 . Hence

$$(6.2) \quad \tilde{U}_0 = U_0 + \lim_{n \rightarrow \infty} \tilde{U}_0 \star {}_nF.$$

Now since the series of convolutions converges, ${}_nF \rightarrow 0$,

$$(6.3) \quad (\tilde{U}_0 \star_n F)(t) = \int_0^t \tilde{U}_0(t-s) d_n F(s) \leq \text{L.U.B.}_{s \leq t} |\tilde{U}_0(s)| \cdot {}_n F(t) \rightarrow 0,$$

and thus (6.2) becomes $\tilde{U}_0 = U_0$.

If $u_0(t)$ satisfies the stated conditions, $U_0(t)$ defined by

$$(6.4) \quad U_0(t) = \int_0^t u_0(s) ds$$

satisfies the conditions of the first part of the theorem. Hence $\tilde{U}_0(t) \equiv U_0(t)$ so that $\tilde{u}_0(t) = u_0(t)$ for almost all t . The right side of (5.3') is thus the same for \tilde{u}_0 as for u_0 . Since both satisfy the equation, the two functions must be identical.

It was shown above that the $x(0)$ distribution can be chosen to make the renewal process stationary, so that $P\{x(t+h) < \lambda\}$ is independent of h ; from this it follows more generally that

$$P\{x(t_j + h) < \lambda_j, j = 1, \dots, n\} \quad (t_1 < \dots < t_n, t_1 + h \geq 0)$$

is independent of h . The definition of the process can then be extended to all negative values of t , by setting

$$P\{x(s_j) < \lambda_j, j = 1, \dots, n\} = P\{x(s_j + h) < \lambda_j, j = 1, \dots, n\}$$

whenever $s_1 < \dots < s_n$, choosing h so large that $s_1 + h \geq 0$. This stationary process will be called the stationary renewal process below. It is determined completely by the lifetime distribution function $F(\lambda)$, which must be supposed to have a finite first moment. It is easily shown that (for the stationary renewal process) given any s , the lifetime of the individual living at $t=s$ has distribution function

$$(7.1) \quad \frac{1}{\Delta} \int_0^\lambda x dF(x);$$

more abstractly stated, if ξ_1 is the first zero of $x(t)$ to the left of s and ξ_2 the first to the right, then $\xi_2 - \xi_1$ has distribution function (7.1). One way to define the stationary renewal process is the following. Let $\dots, x_{-2}, x_{-1}, \bar{x}, x_1, \dots$ be mutually independent chance variables, with the common distribution function $F(\lambda)$ except that \bar{x} has the distribution function (7.1). Let intervals of lengths $\dots, x_{-2}, x_{-1}, \bar{x}, x_1, \dots$ be placed end to end on a line, and choose the origin of coordinates in the interval of length x (choosing it at random, equidistributed, in this interval). Then $x(t)$ is the distance from t to the nearest end point preceding t .

A fundamental problem of renewal theory is, as already remarked, the determination of conditions under which a renewal process with an arbitrary initial age distribution is asymptotically ($t \rightarrow \infty$) identical with the corre-

sponding stationary renewal process. This problem amounts to that of the determination of conditions under which the transition probability functions of the stationary renewal process, $P\{x(0)=x; x(t)\in A\}$, become the absolute probabilities $P\{x(t)\in A\}$ for large t . In order to solve this problem, the stationary renewal process will be examined in considerable detail. Let T_s be the transformation of the space of sample functions taking $x(t)$ into $x(t+s)$. Then T_s is a transformation which preserves probabilities. If α is any chance variable connected with the stationary renewal process, it becomes, under T_s , a new chance variable $T_s\alpha$. For example the chance variable $x(0)$ becomes $T_s x(0)=x(s)$. The condition that the age distribution in a renewal process is independent of the initial age distribution (asymptotically for large t) is, roughly speaking, that if α is a chance variable, α and $T_s\alpha$ are asymptotically independent for large s .

In the preceding paper in this volume [1] it was shown how the analysis of the properties of T_s served to solve precisely the problem considered here. Following the notation of that paper we denote by X the class of real numbers, by \mathfrak{F}_x the field of Borel sets of real numbers, and by Φ measure the measure of Borel sets determined by the distribution with density $(1-F(\lambda))/\Delta$. We shall use ρ to denote the maximum life,

$$\rho = \text{L.U.B. } x. \\ F(x) < 1$$

The sets of Φ measure 0 in the interval $(0, \rho)$ are then simply the sets of Lebesgue measure 0 in that interval.

In [1] the parameter s was restricted to be integral. The renewal process parameter is continuous-valued, and the proper adaptations will be made below. Knowledge of the general theory of continuous parameter processes will be assumed. In particular measure theorems like Fubini's theorem will be applied with no explanatory commentary.

In [1] it was made clear that the concepts of metric transitivity and angle variables play a fundamental role in studies of this type. It can be shown that under fairly general conditions continuous parameter Markoff processes do not have angle variables, but it will be unnecessary to apply such theorems, since a complete analysis can be made in the renewal theory case. This is done in the following theorem.

THEOREM 8. *The stationary renewal process is metrically transitive. It has angle variables if and only if $F(x)$ increases only in jumps, and if the jumps are all at integral multiples of some number ξ (³).*

Suppose that A is an invariant x set, that is, A is a Borel set with the property that the two sets of sample functions

(³) According to a letter from Kakutani, he also has proved this theorem, before the war, in a somewhat different form.

$$\{x(0) \in A\}, \quad \{x(t) \in A\}$$

differ by at most a set of probability 0. We show that the process is metrically transitive by showing that then A has Φ measure 0 or 1. By hypothesis if $\alpha(x)$ is defined as 1 on A and 0 otherwise, $\alpha[x(t)]$ is an invariant function which takes on only the values 0 and 1. Hence, according to Fubini's theorem, neglecting a set of sample functions of total probability 0, for each sample function $x(t)$, $\alpha[x(t)] = 1$ for almost all t or $\alpha[x(t)] = 0$ for almost all t . In other words $x(t) \in A$ for almost all t or $x(t) \in X - A$ for almost all t . Now the sample functions oscillate between 0 and numbers which come arbitrarily close to the maximum lifetime ρ . Hence A must contain almost all points of the interval $(0, \rho)$ or almost none. In the first case A will have Φ measure 1, in the second case Φ measure 0. The process is therefore metrically transitive.

Suppose that there is an angle variable, that is, that there is a number $\lambda \neq 0$ and a function $\alpha[x(0)]$ satisfying

$$(8.1) \quad P\{\alpha[x(t)] = e^{i\lambda t} \alpha[x(0)]\} = 1, \quad P\{\alpha[x(t)] = 0\} < 1,$$

for each t . There are then values of a (which can be taken arbitrarily small) such that

$$(8.2) \quad \alpha(a + t) = e^{i\lambda t} \alpha(a)$$

for almost all values of t in the interval $(0, \rho)$. If μ is any point with the property that $F(\mu + \epsilon) - F(\mu - \epsilon) > 0$ for all $\epsilon > 0$, $x(t)$ must sometimes jump from values near μ to 0. Hence $e^{i\lambda t}$ must have period μ , $\mu = 2k\pi/\lambda$, for some integer k ; that is, the points of increase μ must be integral multiples of $2\pi/\lambda$. Conversely suppose that F increases only in jumps at integral multiples of some number ξ ; we can write $\xi = 2\pi/\lambda$. Then

$$\alpha[x(t)] = e^{i\lambda x(t)}$$

satisfies (8.1). This completes the proof of the theorem.

THEOREM 9. *For any renewal process, if $F(0+) < 1$ and if*

$$(9.1) \quad \int_0^\infty \lambda^2 dF(\lambda) < \infty$$

then

$$(9.2) \quad \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t x(s) ds = \frac{1}{2\Delta} \int_0^\infty \lambda^2 dF(\lambda)$$

with probability 1.

This theorem states that there is always an average age, under the stated conditions. If the distribution of $x(0)$ is chosen to make the process stationary, (9.2) becomes simply a statement of the strong law of large numbers for sta-

tionary processes. In the stationary case (9.2) is true with $x(s)$ replaced by T_α , where α is any chance variable with a finite expectation, and the limit on the right is replaced by $E\{\alpha\}$. In the general case, the theorem is proved as follows. Using the notation $n(t)$ as before, and with $s_n = x_1 + \cdots + x_n$,

$$\begin{aligned}
 \frac{1}{t} \int_0^t x(s) ds &= \frac{1}{t} \int_0^{x_1} (x_0 + s) ds + \frac{1}{t} \int_{x_1}^{x_1+x_2} (s - x_1) ds + \cdots \\
 (9.3) \quad &+ \frac{1}{t} \int_{s_{n(t)}}^t [s - s_{n(t)}] ds \\
 &= \frac{x_0 x_1}{t} + \frac{1}{2t} \sum_{j=1}^{n(t)} x_j^2 + \frac{[t - s_{n(t)}]^2}{2t}, \quad n(t) \geq 1.
 \end{aligned}$$

When $t \rightarrow \infty$, the first term goes to 0, and if we apply the strong law of large numbers and Theorem 1, the second term goes to the right side of (9.2). The third term is dominated by $x_{n(t)+1}^2/2t$. Since $n(t)/t \rightarrow 1/\Delta$ with probability 1, (9.2) will be proved if it is shown that $x_n^2/n \rightarrow 0$. This is a trivial consequence of the strong law of large numbers applied to x_2^2, x_3^2, \dots , and is also easily verified directly.

The asymptotic properties of the conditional probabilities of the form $P\{x(0)=x; x(t) \in A\}$ (where A is any Borel set) and of the conditional expectations $U_x(t)$ will now be investigated by applying the general theory of Markoff processes to the renewal process. The problem is to find conditions under which

$$(10.1) \quad \lim_{t \rightarrow \infty} P\{x(0) = x; x(t) \in A\} = \frac{1}{\Delta} \int_A [1 - F(\lambda)] d\lambda,$$

and under which a corresponding limit relation for $U_x(t)$ will be true. The natural one from the point of view of the theory of Markoff processes is

$$(10.2) \quad \lim_{t \rightarrow \infty} [U_x(t+h) - U_x(t)] = \frac{h}{\Delta}.$$

It is clear that these limit relations do not hold if $F(x)$ has the special form described in Theorem 8. Even if (10.1) and (10.2) are true, however, they cannot be true uniformly in x without strong restrictions on $F(t)$. It is easily seen that uniformity implies that $F(t)$ has finite moments of all orders and even a finite generating function. Hence the conditions of Doebelin and of Kryloff and Bogoliouboff on the transition probabilities of Markoff processes (which imply uniformity)⁽⁴⁾ are rather restrictive in renewal theory. The fact that if $F(x)$ has a finite first moment the renewal process can be made stationary means that the weaker hypotheses of [1] are usable. In the lan-

(⁴) Cf. the last section of [1] for a general discussion of these conditions.

guage of [1], the stationary renewal process is metrically transitive, has no angle variables (excluding the degenerate case described in Theorem 8), has regular transition probabilities, and in addition has the important property that for any x, y there is a possible transition from x to y . This fact makes it possible to replace "almost everywhere" results by "everywhere" results.

The following theorem shows that (10.1) is always true if ordinary limits are replaced by Cesàro limits. The same reasoning would hold for (10.2), but (10.2) taken Cesàro reduces to (1.4) which has already been proved. More generally however, the argument would prove that $E\{x(0) = x; T_t\alpha\} \rightarrow E\{\alpha\}$ (Cesàro) for a large class of chance variables α .

THEOREM 11. *If $0 < \Delta < \infty$, and if A is any Borel set,*

$$(11.1) \quad \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t P\{x(0) = x; x(s) \in A\} ds = \frac{1}{\Delta} \int_A [1 - F(\lambda)] d\lambda,$$

for all values of x in the range of life ($x < \rho$).

This theorem is easily proved directly, but the proof will be omitted because the theorem is a special case of Theorem 1 of [1]. (The change from a discrete to a continuous parameter requires no change in the principle of the proof.)

Since the stationary renewal process is metrically transitive, and has no angle variables if $F(x)$ is not of the special type described in Theorem 8, there is necessarily a type of weak convergence of the transition probability $P\{x(0) = x; x(t) \in A\}$ toward the desired limit $(1/\Delta) \int_A [1 - F(x)] dx$. In fact if $\Delta < \infty$ and if $F(x)$ is not of the type described in Theorem 8, it is easily proved if we use a theorem of von Neumann and Koopman [3, pp. 36-37] that if $\psi(x)$ is a probability density in the interval $(0, \rho)$, then

$$(12.1) \quad \lim_{t \rightarrow \infty, t \in I} \int P\{x(0) = x; x(t) \in A\} \psi(x) dx = \frac{1}{\Delta} \int_A [1 - F(x)] dx.$$

Here I is a certain sum of intervals of relative measure 1^(*) which is independent of ψ, A . In other words, initial absolutely continuous distributions become the stationary distribution in the long run, at least if t is restricted to a set of relative measure 1. A similar statement can be made covering the convergence of $U_x(t+h) - U_x(t)$ to h/Δ as $t \rightarrow \infty$. It is more interesting however to find conditions under which there is ordinary convergence. Such conditions are given in the following theorem.

THEOREM 12. *Suppose that $0 < \Delta < \infty$ and that for some $v \geq 1$ the v th iterated convolution of F with itself is not a singular function. Then for every $x < \rho$*

(*) A Lebesgue measurable set is said to have relative measure 1 if the measure of its intersection with the interval $(-t, t)$ divided by $2t$ approaches 1, when $t \rightarrow \infty$.

$$(12.2) \quad \lim_{t \rightarrow \infty} P\{x(0) = x; x(t) \in A\} = \frac{1}{\Delta} \int_A [1 - F(x)] dx$$

for every Borel set A , and

$$(12.3) \quad \lim_{t \rightarrow \infty} [U_x(t+h) - U_x(t)] = h/\Delta$$

for every h .

If F has the form described in Theorem 8, it and all its iterated convolutions are singular functions. Hence this case is excluded by the conditions of the theorem. The stationary renewal process under consideration is metrically transitive and has no age variables.

According to Theorem 4 of [1] (the change from a discrete to a continuous parameter requires no change in the principle of the proof) if α is a chance variable, restricted as described below,

$$(12.4) \quad \lim_{t \rightarrow \infty} E\{x(0) = x; T, \alpha\} = E\{\alpha\}$$

if only the singular component of the monotone function $P\{x(0)=x; x(t)<\lambda\}$ with respect to the Φ measure with density $[1-F(\lambda)]/\Delta$, or what amounts to the same thing, with respect to Lebesgue measure m_A , goes to zero when $t \rightarrow \infty$. The condition is, analytically,

$$(12.5) \quad \lim_{t \rightarrow \infty} \max_{mN=0} P\{x(0) = x; x(t) \in N\} = 0$$

(where N is a Borel set). Now the transition probability distribution $P\{x(0)=x; x(t) \in A\}$ is, according to (3.1), a distribution in which the point $x=t$ has positive probability $[1-F(x+t)]/[1-F(x)]$, which goes to 0 when $t \rightarrow \infty$, and otherwise, neglecting this point, a set A has conditional probability

$$(12.6) \quad - \int_A [1 - F(\lambda)] d_\lambda U_x(t - \lambda), \quad \lambda < t.$$

The singular part of this measure is

$$(12.7) \quad \max_{mN=0} \int_N [1 - F(t - \lambda)] d_\lambda U_x(\lambda) = \int_0^t [1 - F(t - \lambda)] d\bar{U}_x(\lambda),$$

where $\bar{U}_x(\lambda)$ is the singular component of $U_x(\lambda)$. Let $\bar{F}(x)$ be the singular component of $F(x)$. Since a convolution in which one term is absolutely continuous is itself absolutely continuous, (1.17) implies

$$(12.8) \quad \bar{U}_x(\lambda) \leq F_x + F_x \star \bar{F} + F_x \star \bar{F} \star \bar{F} + \dots$$

If the v th iterated convolution is not singular, the series (12.7) is dominated uniformly in λ by the series

$$1 + \cdots + 1 + a + \cdots + a + a^2 + \cdots + a^2 + \cdots = \nu/(1-a),$$

where there are ν 1's and ν a 's, and so on, and $1-a$ is the value at ∞ of the nonsingular convolution. Hence $\bar{U}_x(\infty) < \infty$, and the second integral in (12.7) goes to 0 when $t \rightarrow \infty$, since the integrand converges boundedly to 0. Thus the basic condition of Theorem 4 of [1] is satisfied. Hence (12.4) is true for certain chance variables α . In particular, according to this theorem, α can be the characteristic function of an $x(0)$ set,

$$\begin{aligned} \alpha &= 1, & \text{if } x(0) \in A, \\ \alpha &= 0, & \text{if } x(0) \in X - A. \end{aligned}$$

With this evaluation of α , (12.4) becomes (12.2). Finally according to Theorem 4 of [1], we can set $\alpha = n(h) - n(0)$ if

$$(12.9) \quad \text{L.U.B.}_x E\{x(0) = x; |n(t+h) - n(t)|\} < \infty.$$

This inequality is a consequence of (1.20)⁽⁶⁾.

BIBLIOGRAPHY

1. J. L. Doob, *Asymptotic properties of Markoff transition probabilities*, Trans. Amer. Math. Soc. vol. 63 (1948) pp. 293-321.
2. W. Feller, *On the integral equation of renewal theory*, Ann. Math. Statist. vol. 12 (1941) pp. 243-267.
3. E. Hopf, *Ergodentheorie*, Ergebnisse der Mathematik, vol. 5, no. 2.
4. S. Täcklind, *Elementare Behandlung vom Erneuerungsproblem*, Skandinavisk Aktuarietidskrift (1944) pp. 1-15.
5. ———, *Fourieranalytische Behandlung vom Erneuerungsproblem*, Skandinavisk Aktuarietidskrift (1945) pp. 68-105.

UNIVERSITY OF ILLINOIS,
URBANA, ILL.

⁽⁶⁾ In a paper to appear in Duke Math J., Blackwell proves that (12.3) is true more generally whenever $F(\lambda)$ is not of the form described in Theorem 8.