RENEWAL THEORY

STEVEN P. LALLEY
UNIVERSITY OF CHICAGO

1. Renewal processes

A *renewal process* is the increasing sequence of random nonnegative numbers $S_0, S_1, S_2, \dots$ gotten by adding i.i.d. *positive* random variables $X_0, X_1, \dots$, that is,

$$S_n = S_0 + \sum_{i=1}^{n} X_i \tag{1}$$

When $S_0 = 0$ the renewal process is an *ordinary* renewal process; when S_0 is a nonnegative random variable the renewal process is a *delayed* renewal process. In either case, the individual terms S_n of this sequence are called *renewals*, or sometimes *occurrences*. With each renewal process is associated a *renewal counting process* $N(t)$ that tracks the total number of renewals (not including the initial occurrence) to date: the random variable $N(t)$ is defined by

$$N(t) = \max\{n : S_n \le t\} = \tau(t) - 1 \quad \text{where} \tag{2}$$

$$\tau(a) = \min\{n \ge 1 : S_n > a\}. \tag{3}$$

Two cases arise in applications, the *arithmetic* case, in which the inter-occurrence times X_i are integer-valued, and the *non-arithmetic* case, in which the distribution of X_i is not supported by any arithmetic progression $h\mathbb{Z}$. The arithmetic case is of particular importance in the theory of discrete-time Markov chains, because the sequence of times at which the Markov chain returns to a particular state x is an arithmetic renewal process, as we will show. Since the theories in the arithmetic and non-arithmetic cases follow mostly parallel tracks, we shall limit our discussion to the arithmetic case.

2. The Feller-Erdös-Pollard Renewal Theorem

Assume that $\{S_n\}_{n\ge 0}$ is an ordinary, arithmetic renewal process with inter-occurrence times $X_i = S_i - S_{i-1}$ and inter-occurrence time distribution $f(k) = f_k = P\{X_i = k\}$. Define the *renewal measure*

$$u(k) = u_k = P\{S_n = k \text{ for some } n \ge 0\} = \sum_{n=0}^{\infty} P\{S_n = k\}. \tag{4}$$

Proposition 1. *The renewal measure u satisfies the* renewal equation

$$u_m = \delta_{0,m} + \sum_{k=1}^{m} f_k u_{m-k} \tag{5}$$

where $\delta_{0,m}$ is the Kronecker delta function (1 if $m = 0$ and 0 otherwise).

Proof. Exercise. (Condition on the first step X_1 of the random walk.) □

The cornerstone of renewal theory is the *Feller-Erdös-Pollard* theorem, which describes the asymptotic behavior of hitting probabilities in a renewal process.

Theorem 1. *(Feller-Erdös-Pollard) Let $(S_n)_{n\geq 0}$ be an ordinary arithmetic renewal process whose inter-occurrence time distribution $f_k = P\{X_i = k\}$ has finite mean $0 < \mu < \infty$ and is not supported by any proper additive subgroup of the integers (i.e., there is no $m \geq 2$ such that $P\{X_i \in m\mathbb{Z}\} = 1$). Let u be the associated renewal measure. Then*

$$\lim_{k\to\infty} u(k) = \lim_{k\to\infty} P\{S_n = k \text{ for some } n \geq 0\} = 1/\mu. \tag{6}$$

Corollary 1. *If $\{S_n\}_{n\geq 0}$ is a* delayed *renewal process whose inter-occurrence time distribution $f_k = P\{X_1 = k\}$ satisfies the hypotheses of the Feller-Erdös-Pollard theorem, then*

$$\lim_{k\to\infty} P\{S_n = k \text{ for some } n \geq 0\} = 1/\mu. \tag{7}$$

Proof. Condition on the initial delay:

$$P\{S_n = k \text{ for some } n \geq 0\} = \sum_{m=0}^{\infty} P\{S_0 = m\}P\{S'_n = k - m \text{ for some } n \geq 0\}$$

where $S'_n = S_n - S_0 = \sum_{i=1}^{n} X_i$. The Feller-Erdös-Pollard theorem implies that for each m the hitting probability $P\{S'_n = k - m \text{ for some } n \geq 0\}$ converges to $1/\mu$ as $k \to \infty$, and so the dominated convergence theorem (applied to the infinite sum above) implies (7). □

The Feller-Erdös-Pollard theorem is, in effect, equivalent to Kolmogorov's limit theorem for aperiodic, irreducible, positive recurrent Markov chains. We will give two proofs, first, a proof based on Kolmogorov's theorem, and second, a direct proof based on a coupling argument.

First Proof. Let $(R_k)_{k\geq 0}$ be the *residual age* process, defined as follows:

$$R_k = S_{\tau(k)} - k \quad \text{where} \quad \tau(k) := \min\{n \geq 0 : S_n \geq k\}.$$

It is easily verified that the residual age process is a Markov chain (exercise!) with transition probabilities

$$\begin{aligned} p(0,k) &= f_{k+1} \quad && \text{for } k = 0,1,2,\cdots, \\ p(k,k-1) &= 1 && \text{for } k \geq 1. \end{aligned}$$

Occurrences of the renewal process coincide with visits to state 0 by the residual lifetime process, so for any integer $m \geq 0$, the event $\{S_n = m \text{ for some } n\}$ is the same as the event $\{X_m = 0\}$. Thus,

$$\lim_{k\to\infty} P\{S_n = k \text{ for some } n \geq 0\} = \lim_{k\to\infty} P^0\{X_k = 0\}$$

provided the latter limit exists. By Kolmogorov's limit theorem, this limit will exist and equal the steady-state probability of state 0 for the residual lifetime chain if this chain can be shown to be aperiodic, irreducible, and positive recurrent.

Irreducibility is fairly obvious, once the appropriate state space is specified. The issue is this: if the inter-occurrence time distribution f_k has finite support, and $K = \max\{k : f_k > 0\}$, then the residual lifetime will never be larger than $K - 1$, and so in particular integer states $k \geq K$ will be inaccessible from state 0. Thus, in this case, the appropriate state space is $[K-1] = \{0,1,2,\cdots,K-1\}$. On the other hand, if the inter-occurrence time distribution f_k has infinite support, then the appropriate state space is the set $\mathbb{Z}_+$ of all nonnegative integers. In either case, all states are clearly accessible from 0, and 0 is accessible from any state k, so the Markov chain is irreducible.

To show that the residual lifetime chain is aperiodic it suffices, since the chain is irreducible, to prove that state 0 has period 1. Now the set of times $n \geq 1$ for which $p_n(0,0)$ is the additive semigroup generated by the set of positive integers k for which $f_k > 0$ (why?). The hypothesis of the theorem ensures that the GCD of this set is 1, so state 0 has period 1.

Finally, to show that the chain is positive recurrent, it suffices to show that state 0 has finite expected return time. But the time until first return to 0 is just $S_1 = X_1$, the first inter-occurrence time, and we have assumed that this has finite expectation μ. Therefore, the steady-state probability of state 0 is

$$\pi(0) = 1/\mu.$$

□

The second proof of the Feller-Erdös-Pollard theorem will rely on a *coupling* argument to show that the probabilities $u(k)$ and $u(k-1)$ are nearly equal for large k:

Proposition 2. *Under the hypotheses of the Feller-Erdös-Pollard theorem,*

$$\begin{aligned} &\lim_{k\to\infty} u(k) - u(k-1) = 0 \quad \textit{and so} \\ &\lim_{k\to\infty} u(k) - u(k-j) = 0 \quad \textit{for every } j \geq 1. \end{aligned} \tag{8}$$

Proof. Assume first that the distribution $\{f_k\}_{k\geq 1}$ of the inter-occurrence times X_i is not supported by any *coset* of a proper subgroup of the integers (i.e., there do not exist integers k and $m \geq 2$ such that $P\{X_i \in k + m\mathbb{Z}\} = 1$). Let $\{X_i\}_{i\geq 1}$ and $\{X_i'\}_{i\geq 1}$ be two independent sequences of identically distributed random variables, all with distribution $\{f_k\}_{k\geq 1}$. Since the inter-occurrence time distribution has finite mean, the differences $Y_i := X_i - X_i'$ have mean zero. Furthermore, since $\{f_k\}_{k\geq 1}$ of the inter-occurrence times X_i is not supported by any coset of a proper subgroup of the integers, the distribution of the differences Y_i is not supported by any proper subgroup of the integers, and so by the recurrence theorem for one-dimensional random walks the sequence

$$S_n^Y = \sum_{i=1}^{n} Y_i$$

will visit every integer infinitely often, with probability one. Let T be the smallest n such that $S_n^Y = 1$; then T is a stopping time for the sequence $\{(X_n, X_n')\}_{n\geq 1}$ (that is, for the minimal filtration generated by this sequence). Define

$$\begin{aligned} S_n &= \sum_{i=1}^{n} X_i \quad \text{and} \\ S_n' &= 1 + \sum_{i=1}^{n\wedge T} X_i' + \sum_{i=1+n\wedge T}^{n} X_i. \end{aligned}$$

By construction, $S_n = S_n'$ for all $n \geq T$, Moreover, since T is a stopping time, the process $\{S_n' - 1\}_{n\geq 0}$ has the same joint distribution as does $\{S_n\}_{n\geq 0}$. Therefore, the limit relation (8) follows by the coupling principle: in particular, for large k there will be probability close to 1 that the two walks have joined before either reaches k, and so $u(k) - u(k-1) \approx 0$.

If the distribution $\{f_k\}_{k\geq 1}$ is supported by $\ell + m\mathbb{Z}$ for some $m \geq 2$ and $1 \leq k < m$ then the differences Y_i take their values in the subgroup $m\mathbb{Z}$, so the recurrence theorem no longer implies that the random walk S_n^Y will visit every integer. However, S_n^Y will visit every point of $m\mathbb{Z}$, so the coupling argument above (with the obvious modifications) shows that

$$\begin{aligned} &\lim_{k\to\infty} u(k) - u(k-m) = 0 \quad \Longrightarrow \\ &\lim_{k\to\infty} u(k) - u(k-jm) = 0 \quad \forall\, j \in \mathbb{N}. \end{aligned} \tag{9}$$

Recall that u satisfies the renewal equation $u(k) = Eu(k - K_1)$ for all integers $k \geq 1$ (compare equation (5)), so if the distribution $\{f_k\}_{k\geq 1}$ is supported by $\ell + m\mathbb{Z}$ for some $m \leq 2$ and $1 \leq \ell \leq m-1$ then (9) implies that

$$\begin{aligned} &\lim_{k\to\infty} u(k) - u(k-jm-\ell) = 0 \quad \forall\, j \in \mathbb{N}, \quad \text{which further implies} \\ &\lim_{k\to\infty} u(k) - u(k-jm-2\ell) = 0 \quad \forall\, j \in \mathbb{N}, \\ &\lim_{k\to\infty} u(k) - u(k-jm-3\ell) = 0 \quad \forall\, j \in \mathbb{N}, \\ &\text{etc.} \end{aligned}$$

The sequence $\ell, 2\ell, 3\ell, \dots$ must exhaust the integers mod m, because otherwise the distribution $\{f_k\}_{k\geq 1}$ would be supported by a proper subgroup of $\mathbb{Z}$, contrary to our standing assumptions. The proposition now follows. □

Second Proof of the Feller-Erdös-Pollard theorem. Proposition 2 implies that for large k the renewal measure $u(k)$ differs by only a negligible amount from $u(k-j)$ for any j. To deduce that $u(k)$ converges to $1/\mu$ we use the renewal equations $u_m = \delta_0(m) + \sum_{k=1}^{m} f_k u_{m-k}$. Summing over all m from 0 to n gives

$$\sum_{k=0}^{n} u_{n-k} = 1 + \sum_{k=0}^{n} u_{n-k} \sum_{j=1}^{k} f_j,$$

which can be re-written as

$$\sum_{k=0}^{n} u_{n-k}(1 - F_k) = 1 \tag{10}$$

where $F_k = \sum_{j=1}^{k} f_j$.

The Feller-Erdös-Pollard theorem follows easily from equation (10) and Proposition 2. Since $\sum_{k=1}^{\infty}(1 - F_k) = \mu$, the sequence $\{(1 - F_k)/\mu\}_{k\geq 1}$ is a probability distribution on the positive integers. For any $\varepsilon > 0$ there exists $k(\varepsilon) < \infty$ such that this probability distribution puts at least $1-\varepsilon$ of its mass on the interval $[1, k(\varepsilon)]$. By Proposition 2, for sufficiently large m, say $m \geq m(\varepsilon)$, the function $u(m-j)$ will not differ by more than ε from $u(m)$, and in any case $u(j) \leq 1$ for all j. Consequently, for $m \geq m(\varepsilon)$,

$$|u_m \mu - 1| \leq \varepsilon + \varepsilon\mu.$$

Since $\varepsilon > 0$ is arbitrary, the theorem follows. □

3. The Choquet-Deny Theorem[1]

There are a number of other interesting proofs of the Feller-Erdös-Pollard theorem, most of which proceed by proving the intermediate Proposition 2 in some other way. In this section I will discuss one of these alternative approaches, which brings to light another important property of random walks on the integer lattices, the so-called *Liouville property*.

Recall that a *harmonic function* for a denumerable-state Markov chain with transition probabilities $p(\cdot,\cdot)$ is a real-valued function h on the state space $\mathscr{X}$ that satisfies the *mean-value property*

$$h(x) = \sum_{y \in \mathscr{X}} p(x,y)h(y) \quad \text{for all } x \in \mathscr{X}. \tag{11}$$

(The function h must also be integrable against each of the measures $p(x,\cdot)$ in order that the sum on the right side of (11) be well-defined.) A *random walk* on the $d-$dimensional integer lattice $\mathbb{Z}^d$ is a Markov chain with state space $\mathbb{Z}^d$ whose transition probabilities are spatially homogeneous, i.e., such that there is a probability distribution $\{p_x\}_{x \in \mathbb{Z}^d}$ such that

$$p(x,y) = p_{y-x} \quad \forall\, x,y \in \mathbb{Z}^d.$$

More generally, a *(right) random walk* on a (multiplicative) finite or denumerable *group* Γ is is a Markov chain with state space Γ whose transition probabilities satisfy

$$p(x,y) = p_{x^{-1}y}$$

for some probability distribution $\{p_z\}_{z \in \Gamma}$. (In general, groups need not be abelian, so the order of multiplication matters.) It is of considerable importance to know if a group supports random walks with nontrivial (that is, non-constant) bounded harmonic functions. Groups for which random walks have no nontrivial bounded harmonic functions are said to have the *Liouville property*.[2] The Choquet-Deny theorem states that $\mathbb{Z}^d$ has the Liouville property.

Theorem 2. *(Choquet-Deny) An irreducible random walk on $\mathbb{Z}^d$, or more generally on any abelian group, admits no non-constant bounded harmonic functions.*

Proof. (Doob-Snell-Williamson; Szekely-Zeng) Recall that a Markov chain is *irreducible* if any state is accessible from any other, that is, if for any two states x,y there is a finite path from x to y such that each step x_i, x_{i+1} of the path has $p(x_i,x_{i+1}) > 0$. Consequently, to prove that a bounded harmonic function h is constant it suffices to show that $h(x) = h(y)$ for any two states x,y with $p(x,y) > 0$.

If h is a bounded harmonic function for the Markov chain $\{X_n\}_{n \geq 0}$ then the sequence $h(X_n)$ is a bounded martingale under P^x, for any initial state x. Hence, the martingale convergence theorem implies that $H_\infty := \lim_{n\to\infty} h(X_n)$ exists almost surely. Since the martingale is bounded, the convergence also holds in L^1 and so for any initial state x and any $n \geq 1$

$$h(X_n) = E^x(H_\infty \mid \mathscr{F}_n) \tag{12}$$

where $\mathscr{F}_n$ is the $\sigma-$algebra generated by the random variables $\{X_j\}_{j \leq n}$.

[1]This section, which presents yet another approach to the Feller-Erdös-Pollard theorem, can be skipped

[2]The terminology is used because Liouville proved the first important theorem having to do with the existence of bounded harmonic functions. See any elementary text on complex variables for more.

If $\{X_n\}_{n\geq 0}$ is a random walk on an *abelian* group Γ then the increments $\xi_n = X_n - X_{n-1}$ are independent and identically distributed. In this case, the limit random variable H_∞ must be measurable with respect to the exchangeable $\sigma-$algebra, because any permutation of the first n increments ξ_i of the random walk will not change the value of X_{n+k} for any $k \geq 0$. (Observe that this uses, in an essential way, the hypothesis that the group is abelian.) But the Hewitt-Savage 0-1 Law implies that any exchangeable random variable must be constant a.s., and so $H_\infty = C$ for some $C \in \mathbb{R}$. It now follows by equation (12) that $h(X_n) = C$ almost surely for every $n \geq 1$, and in particular for $n = 1$. Thus, if $p(x, y) > 0$ then $h(y) = h(x) = C$. □

Alternative proof of Proposition 2. Extend the sequence $\{u_k\}_{k\geq 0}$ to the integers $k \in \mathbb{Z}$ by setting $u(k) = 0$ if $k \leq -1$. Since the values u_k are uniformly bounded (they are probabilities, so they must be in $[0,1]$), the Bolzano-Weierstrass theorem implies that any sequence $m_n \to$ has a subsequence m_r such that

$$\lim_{r\to\infty} u_{k+m_r} = v_k = v(k)$$

exists for every $k \in \mathbb{Z}$. Since $m_r \to \infty$, the renewal equation (5) implies that any such limit must satisfy

$$v(k) = Ev(k - X_1),$$

that is, v is a harmonic function for the random walk with steps $-X_i$. By the Choquet-Deny theorem, v is constant. This clearly implies Proposition 2. □

Exercise 1. Give another proof of the Choquet-Deny theorem based on a coupling argument. HINT: There are two important issues you must deal with: (i) the Choquet-Deny theorem does not require that the increments of the random walk have finite first moment, so you must be careful about using the recurrence theorem; and (ii) the Choquet-Deny theorem holds in all dimensions $d \geq 1$. To deal with (i), try a truncation in your coupling scheme. For (ii) try coupling each of the d component random walks one at a time.

4. THE RENEWAL EQUATION AND THE KEY RENEWAL THEOREM

4.1. **The Renewal Equation.** The usefulness of the Feller-Erdös-Pollard theorem derives partly from its connection with another basic theorem called the Key Renewal Theorem (see below) which describes the asymptotic behavior of solutions to the *Renewal Equation.* The Renewal Equation is a convolution equation relating bounded sequences $\{z(m)\}_{m\geq 0}$ and $\{b(m)\}_{m\geq 0}$ of real numbers:

Renewal Equation, First Form:

$$z(m) = b(m) + \sum_{k=1}^{m} f(k) z(m-k). \tag{13}$$

Here $f(k) = f_k$ is the interoccurrence time distribution for the renewal process. There is an equivalent way of writing the Renewal Equation that is more suggestive of how it actually arises in practice. Set $z(m) = b(m) = 0$ for $m < 0$; then the upper limit $k = m - 1$ in the sum in the Renewal Equation may be changed to $m = \infty$ without affecting its value. The Renewal Equation may now be written as follows, with X_1 representing the first interoccurrence time:

Renewal Equation, Second Form:

$$z(m) = b(m) + Ez(m - X_1). \tag{14}$$

As you will come to appreciate, renewal equations crop up all over the place. In many circumstances, the sequence $z(m)$ is some scalar function of time whose behavior is of some interest; the renewal equation is gotten by conditioning on the value of the first interoccurrence time. In carrying out this conditioning, it is crucial to realize that the sequence $S_1^*, S_2^*, \dots$ defined by

$$S_n^* = S_n - X_1 = \sum_{j=2}^{n} X_j$$

is itself a renewal process, independent of X_1, and with the same interoccurrence time distribution $f(x)$.

Example 1. *(Total Lifetime Distribution)*Let $L(m) = A(m) + R(m)$ be the total lifetime of the component in use at time m. (Here $A(m)$ and $R(m)$ are the age and residual lifetime random variables, respectively.) Fix $r \geq 1$, and set $z(m) = P\{L(m) = r\}$. Then z satisfies the Renewal Equation (14) with

$$\begin{aligned} b(m) &= f(r) \quad \text{for } m = 0, 1, 2, \dots, r-1 \\ &= 0 \quad \text{for } m \geq r \end{aligned} \tag{15}$$

EXERCISE: Derive this.

4.2. **Solution of the Renewal Equation.** Consider the Renewal Equation in its second form $z(m) = b(m) + Ez(m - X_1)$ where by convention $z(m) = 0$ for all negative values of m. Since the function $z(\cdot)$ appears on the right side as well as on the left, it is possible to resubstitute on the right. This leads to a sequence of equivalent equations:

$$\begin{aligned} z(m) &= b(m) + Ez(m - X_1) \\ &= b(m) + Eb(m - X_1) + Ez(m - X_1 - X_2) \\ &= b(m) + Eb(m - X_1) + Eb(m - X_1 - X_2) + Ez(m - X_1 - X_2 - X_3) \end{aligned}$$

and so on. After m iterations, there is no further change (because $S_{m+1} > m$ and $z(l) = 0$ for all negative integers l), and the right side no longer involves z. Thus, it is possible to solve for z in terms of the sequences b and p:

$$\begin{aligned} z(m) &= \sum_{n=0}^{\infty} Eb(m - S_n) \\ &= \sum_{n=0}^{\infty} \sum_{x=0}^{\infty} b(m-x) P\{S_n = x\} \\ &= \sum_{x=0}^{\infty} \sum_{n=0}^{\infty} b(m-x) P\{S_n = x\} \\ &= \sum_{x=0}^{\infty} b(m-x) u(x). \end{aligned}$$

Note that only finitely many terms in the series are nonzero, so the interchange of summations is justified. Thus, the solution to the Renewal Equation is the convolution of the sequence $b(m)$ with the renewal measure:

$$\boxed{z(m) = \sum_{x=0}^{\infty} b(m-x)u(x)} \tag{16}$$

4.3. **The Key Renewal Theorem.** The formula (16) and the Feller-Erdös-Pollard theorem now combine to give the asymptotic behavior (as $m \to \infty$) of the solution z.

Theorem 3. *(Key Renewal Theorem) Let $z(m)$ be the solution to the Renewal Equation (14). If the sequence $b(m)$ is absolutely summable, then*

$$\lim_{m\to\infty} z(m) = \mu^{-1} \sum_{k=0}^{\infty} b(k). \tag{17}$$

Proof. The formula (16) may be rewritten as

$$z(m) = \sum_{k=0}^{\infty} b(k)u(m-k) \tag{18}$$

For each fixed k, the sequence $u(m-k) \to \mu^{-1}$ as $m \to \infty$, by the Feller-Erdös-Pollard theorem. Thus, as $m \to \infty$, the kth term of the series (18) converges to $b(k)/\mu$. Moreover, because $u(m-k) \le 1$, the kth term is bounded in absolute value by $|b(k)|$. By hypothesis, this sequence is summable, so the Dominated Convergence Theorem implies that the series converges as $m \to \infty$ to the right side of (17).

Example 2. *Residual Lifetime.* For each fixed $r \ge 1$, the sequence $z(m) = P\{R(m) = r\}$ satisfies the renewal equation

$$z(m) = P\{X_1 = m+r\} + \sum_{k=1}^{m} z(m-k)P\{X_1 = k\} = f_{m+r} + \sum_{k=1}^{m} z(m-k)f_k. \tag{19}$$

This reduces to (14) with $b(m) = f(m+r)$. The sequence $b(m)$ is summable, because $\mu = EX_1 < \infty$ (why?). Therefore, the Key Renewal Theorem implies that for each $r = 1,2,3,\dots,$

$$\lim_{m\to\infty} P\{R(m) = r\} = \mu^{-1} \sum_{k=0}^{\infty} f(k+r) = \mu^{-1} P\{X_1 \ge r\}. \tag{20}$$

This could also be deduced from the convergence theorem for Markov chains, using the fact that the sequence R_m is an irreducible, positive recurrent Markov chain with stationary distribution (**??**).

Example 3. *Total Lifetime.* Recall (Example 3 above) that the sequence $z(m) = P\{L(m) = r\}$ satisfies the Renewal Equation (14) with $b(m)$ defined by (15). Only finitely many terms of the sequence $b(m)$ are nonzero, and so the summability hypothesis of the Key Renewal Theorem is satisfied. Since $\sum_{k\ge 0} b(m) = rf(r)$, it follows from (17) that

Corollary 2.

$$\lim_{m\to\infty} P\{L(m) = r\} = rf(r)/\mu. \tag{21}$$

Example 4. *Fibonacci numbers.* Discrete convolution equations arise in many parts of probability and applied mathematics, but often with a "kernel" that isn't a proper probability distribution. It is important to realize that such equations can be converted to (standard) renewal equations by the device known as *exponential tilting*. Here is a simple example.

Consider the *Fibonacci sequence* 1,1,2,3,5,8,.... This is the sequence a_n defined by the recursion

$$a_{n+2} = a_{n+1} + a_n \tag{22}$$

and the initial conditions $a_0 = a_1 = 1$. To convert the recursion to a renewal equation, multiply a_n by a geometric sequence:

$$z_n = \theta^{-n} a_n$$

for some value $\theta > 0$. Also, set $z_n = 0$ for $n \leq -1$. Then (22) is equivalent to the equation

$$z_n = z_{n-1}\theta^1 + z_{n-2}\theta^2 + \delta_{0,n} \tag{23}$$

where $\delta_{i,j}$ is the Kronecker delta function. (The delta function makes up for the fact that the original Fibonacci recursion (22) does not by itself specify a_0 and a_1.) Equation (23) is a renewal equation for any value of $\theta > 0$ such that $\theta^1 + \theta_2 = 1$, because then we can set $f_1 = \theta^1$, $f_2 = \theta^2$, and $f_m = 0$ for $m \geq 3$. Is there such a value of θ? Yes: it's the *golden ratio*

$$\theta = \frac{-1+\sqrt{5}}{2}.$$

(The other root is negative, so we can't use it to produce a renewal equation from (22).) The Key Renewal Theorem implies that

$$\lim_{n\to\infty} z_n = \lim_{n\to\infty} \theta^{-n} a_n = 1/(\theta + 2\theta^2).$$

Thus, the Fibonacci sequence grows at an exponential rate, and the rate is the inverse of the golden ratio.

5. Generating Functions and Renewal Theory

5.1. Generating Functions.

The Renewal Theorem tells us how the age and residual lifetime distributions behave at large times, and similarly the Key Renewal Theorem tells us how the solution of a renewal equation behaves at infinity. In certain cases, *exact* calculations can be done. These are usually done using *generating functions*.

Definition 1. The *generating function* of a sequence $\{a_n\}_{n\geq 0}$ of real (or complex) numbers is the function $A(z)$ defined by the power series

$$A(z) = \sum_{n=0}^{\infty} a_n z^n. \tag{24}$$

Observe that for an arbitrary sequence $\{a_n\}$ the series (24) need not converge for all complex values of the argumetn z. In fact, for some sequences the series (24) diverges for *every* z except $z = 0$: this is the case, for instance, if $a_n = n^n$. But for many sequences of interest, there will exist a positive number R such that, for all complex numbers z such that $|z| < R$, the series (24) converges absolutely. In such cases, the generating function $A(z)$ is said to have positive *radius of convergence*. The generating functions in all of the problems considered in these notes will have positive radius of convergence. Notice that if the entries of the sequence a_n are probabilities, that is, if $0 \le a_n \le 1$ for all n, then the series (24) converges absolutely for all z such that $|z| < 1$.

If the generating function $A(z)$ has positive radius of convergence then, at least in principal, all information about the sequence $\{a_n\}$ is encapsulated in the generating function $A(z)$. In particular, each coefficient a_n can be recovered from the function $A(z)$, since $n!a_n$ is the nth derivative of $A(z)$ at $z = 0$. Other information may also be recovered from the generating function: for example, if the sequence $\{a_n\}$ is a discrete probability density, then its mean may be obtained by evaluating $A'(z)$ at $z = 1$, and all of the higher moments may be recovered from the higher derivatives of $A(z)$ at $z = 1$.

A crucially important property of generating functions is the *multiplication law*: The generating function of the convolution of two sequences is the product of their generating functions. This is the basis of most uses of generating functions in random walk theory, and all of the examples considered below. For *probability* generating functions, this fact is a consequence of the multiplication law for expectations of independent random variables: If X and Y are independent, nonnegative-integer valued random variables, then

$$Ez^{X+Y} = Ez^X Ez^Y. \tag{25}$$

5.2. **The Renewal Equation.** Let $\{f_k\}_{k\ge1}$ be a probability distribution on the positive integers with finite mean $\mu = \sum kf_k$, and let $X_0, X_1, \dots$ be a sequence of independent, identically distributed random variables with common discrete distribution $\{f_k\}$. Define the *renewal sequence* associated to the distribution $\{f_k\}$ to be the sequence

$$\begin{aligned} u_m &= P\{S_n = m \text{ for some } n \ge 0\} \\ &= \sum_{n=0}^{\infty} P\{S_n = m\} \end{aligned} \tag{26}$$

where $S_n = X_1 + X_2 + \cdots + X_n$ (and $S_0 = 0$). note that $u_0 = 1$ The *renewal equation* is obtained by conditioning on the first step $S_1 = X_1$:

$$u_m = \sum_{k=1}^{m} f_k u_{m-k} + \delta_0(m) \tag{27}$$

where δ_0 is the Kronecker delta (1 at 0; and 0 elsewhere).

The renewal equation is a particularly simple kind of recursive relation: the right side is just the *convolution* of the sequences f_k and u_m. The appearance of a convolution should always suggest the use of some kind of generating function or transform (Fourier or Laplace), because these always convert convolution to multiplication. Let's try it: define generating functions

$$U(z) = \sum_{m=0}^{\infty} u_m z^m \qquad \text{and} \tag{28}$$

$$F(z) = \sum_{k=1}^{\infty} f_k z^k. \tag{29}$$

Observe that if you multiply the renewal equation (27) by z^m and sum over m then the left side becomes $U(z)$, so

$$
\begin{aligned}
U(z) &= 1 + \sum_{m=0}^{\infty}\sum_{k=1}^{m} f_k u_{m-k} z^m \\
&= 1 + \sum_{k=1}^{\infty}\sum_{m=k}^{\infty} f_k z^k u_{m-k} z^{m-k} \\
&= 1 + F(z)U(z).
\end{aligned}
\tag{30}
$$

Thus, we have a simple *functional equation* relating the generating functions $U(z)$ and $F(z)$. It may be solved for $U(z)$ in terms of $F(z)$:

$$
\boxed{U(z) = \frac{1}{1-F(z)}}
\tag{31}
$$

5.3. **Partial Fraction Decompositions.** Formula (31) tells us how the generating function of the renewal sequence is related to the probability generating function of the steps X_j. Extracting useful information from this relation is, in general, a difficult analytical task. However, in the special case where the probability distribution $\{f_k\}$ has finite support, the method of *partial fraction decomposition* provides an effective method for recovering the terms u_m of the renewal sequence. Observe that when the probability distribution $\{f_k\}$ has finite support, its generating function $F(z)$ is a *polynomial*, and so in this case the generating function $U(z)$ is a *rational function*.[3]

The strategy behind the method of partial fraction decomposition rests on the fact that a *simple pole* may be expanded as a geometric series: in particular, for $|z| < 1$,

$$
(1-z)^{-1} = \sum_{n=0}^{\infty} z^n.
\tag{32}
$$

Differentiating with respect to z repeatedly gives a formula for a pole of order $k+1$:

$$
(1-z)^{-k-1} = \sum_{n=k}^{\infty} \binom{n}{k} z^{n-k}.
\tag{33}
$$

Suppose now that we could write the generating function $U(z)$ as a sum of poles $C/(1-(z/\zeta))^{k+1}$ (such a sum is called a *partial fraction decomposition*). Then each of the poles could be expanded in a series of type (32) or (33), and so the coefficients of $U(z)$ could be obtained by adding the corresponding coefficients in the series expansions for the poles.

Example: Consider the probability distribution $f_1 = f_2 = 1/2$. The generating function F is given by $F(z) = (z+z^2)/2$. The problem is to obtain a partial fraction decomposition for $(1-F(z))^{-1}$. To do this, observe that at every pole $z = \zeta$ the function $1-F(z)$ must take the value 0. Thus, we look for potential poles at the zeros of the polynomial $1-F(z)$. In the case under consideration,

[3] A *rational function* is the ratio of two polynomials.

the polynomial is quadratic, with roots $\zeta_1 = 1$ and $\zeta_2 = -2$. Since each of these is a simple root both poles should be simple; thus, we should try

$$\frac{1}{1-(z+z^2)/2} = \frac{C_1}{1-z} + \frac{C_2}{1+(z/2)}.$$

The values of C_1 and C_2 can be gotten either by adding the fractions and seeing what works or by differentiating both sides and seeing what happens at each of the two poles. The upshot is that $C_1 = 2/3$ and $C_2 = 1/3$. Thus,

$$U(z) = \frac{1}{1-F(z)} = \frac{2/3}{1-z} + \frac{1/3}{1+(z/2)}. \tag{34}$$

We can now read off the renewal sequence u_m by expanding the two poles in geometric series:

$$u_m = \frac{2}{3} + \frac{1}{3}(-2)^m. \tag{35}$$

There are several things worth noting. First, the renewal sequence u_m has limit 2/3. This equals $1/\mu$, where $\mu = 3/2$ is the mean of the distribution $\{f_k\}$. We should be reassured by this, because it is what the Feller-Erdös-Pollard Renewal Theorem predicts the limit should be. Second, the remainder term $(1/3)(-2)^{-m}$ decays exponentially in m. This is always the case for distributions $\{f_k\}$ with finite support. It is not always the case for arbitrary distributions $\{f_k\}$, however.

Problem 1. Recall that the *Fibonacci sequence* $1, 1, 2, 3, 5, 8, \dots$ is the sequence a_n such that $a_1 = a_2 = 1$ and such that

$$a_{m+2} = a_m + a_{m+1}.$$

(A) Find a functional equation for the generating function of the Fibonacci sequence. (B) Use the method of partial fractions to deduce an *exact* formula for the terms of the Fibonacci sequence in terms of the golden ratio.

5.4. **Step Distributions with Finite Support.** Assume now that the step distribution $\{f_k\}$ has finite support, is nontrivial (that is, does not assign probability 1 to a single point) and is nonlattice (that is, it does not give probability 1 to a proper arithmetic progression). Then the generating function $F(z) = \sum f_k z^k$ is a polynomial of degree at least two. By the Fundamental Theorem of Algebra, $1 - F(z)$ may be written as a product of linear factors:

$$1 - F(z) = C \prod_{j=1}^{K} (1 - z/\zeta_j) \tag{36}$$

The coefficients ζ_j in this expansion are the (possibly complex) roots of the polynomial equation $F(z) = 1$. Since the coefficients f_k of $F(z)$ are real, the roots of $F(z) = 1$ come in conjugate pairs; thus, it is only necessary to find the roots in the upper half plane (that is, those with nonnegative imaginary part). In practice, it is usually necessary to solve for these roots numerically. The following proposition states that none of the roots is *inside* the unit circle in the complex plane.

Lemma 1. *If the step distribution $\{f_k\}$ is nontrivial, nonlattice, and has finite support, then the polynomial $1-F(z)$ has a simple root at $\zeta_1 = 1$, and all other roots ζ_j satisfy the inequality $|\zeta_j| > 1$.*

Proof. It is clear that $\zeta_1 = 1$ is a root, since $F(z)$ is a probability generating function. To see that $\zeta_1 = 1$ is a *simple* root (that is, occurs only once in the product (36)), note that if it were a multiple root then it would have to be a root of the derivative $F'(z)$ (since the factor $(1-z)$ would occur at least twice in the product (36)). If this were the case, then $F'(1) = 0$ would be the mean of the probability distribution $\{f_k\}$. But since this distribution has support $\{1,2,3\cdots\}$, its mean is at least 1.

In order that ζ be a root of $1-F(z)$ it must be the case that $F(\zeta) = 1$. Since $F(z)$ is a probability generating function, this can only happen if $|\zeta| \geq 1$. Thus, to complete the proof we must show that there are no roots of modulus one other than $\zeta = 1$. Suppose, then, that $\zeta = e^{i\theta}$ is such that $F(\zeta) = 1$, equivalently,

$$\sum f_k e^{i\theta k} = 1.$$

Then for every k such that $f_k > 0$ it must be that $e^{i\theta k} = 1$. This implies that θ is an integer multiple of $2\pi/k$, and that this is true for every k such that $f_k > 0$. Since the distribution $\{f_k\}$ is nonlattice, the greatest common divisor of the integers k such that $f_k > 0$ is 1. Hence, θ is an integer multiple of 2π, and so $\zeta = 1$. □

Corollary 3. *If the step distribution $\{f_k\}$ is nontrivial, nonlattice, and has finite support, then*

$$\frac{1}{1-F(z)} = \frac{1}{\mu(1-z)} + \sum_{r=1}^{R} \frac{C_r}{(1-z/\zeta_r)^{k_r}} \tag{37}$$

where μ is the mean of the distribution $\{f_k\}$ amd the poles ζ_r are all of modulus strictly greater than 1.

Proof. The only thing that remains to be proved is that the simple pole at 1 has residue $1/\mu$. To see this, multiply both sides of equation (37) by $1-z$:

$$\frac{1-z}{1-F(z)} = C + (1-z)\sum_{r} \frac{C_r}{(1-z/\zeta_r)^{k_r}}.$$

Now take the limit of both sides as $z \to 1-$: the limit of the right side is clearly C, and the limit of the left side is $1/\mu$, because μ is the derivative of $F(z)$ at $z = 1$. Hence, $C = 1$. □

Corollary 4. *Assume that the hypotheses of Corollary 3 hold, and assume also that the poles ζ_r are all simple (equivalently, the roots of the equation $F(z) = 1$ are all simple). Then the renewal measure has the explicit representation*

$$u(m) = \frac{1}{\mu} + \sum_{r=1}^{R} C_r \zeta_r^{-m}. \tag{38}$$