Chapter 2
Introduction to Renewal Theory

2.1 Introduction

Let $\{X_i\}_{i=1}^{\infty}$ be a series of independent and identically distributed nonnegative random variables. Assume they are continuous. In particular, there exists some density function $f_X(x)$, $x \geq 0$, such that $F_X(x) \equiv \mathrm{P}(X_i \leq x) = \int_{t=0}^{x} f_X(t)\,dt$, $i \geq 1$. Imagine X_i representing the life span of a lightbulb. Specifically, there are infinitely many lightbulbs in stock. At time $t = 0$, the first among them is placed. It burns out after a (random) time of X_1. Then it is replaced by a fresh lightbulb that itself is replaced after an additional (random) time of X_2, etc. Note that whenever a new lightbulb is placed all statistically starts afresh. Let $S_n = \Sigma_{i=1}^{n} X_i$, $n \geq 1$, and set $S_0 = 0$. Of course, $S_{n+1} = S_n + X_{n+1}$, $n \geq 0$.

For $t \geq 0$, let

$$N(t) \equiv \sup\{n \geq 0 | S_n \leq t\}.$$

In words, $N(t)$ is the number of lightbulbs which were burnt out during the time interval $[0, t]$. Note that it is possible that $N(t)$ equals zero. The set of random variables $N(t)$, $t \geq 0$, is called a *renewal process*. The following equality between events clearly holds:

$$\{N(t) = n\} = \{S_n \leq t, S_{n+1} > t\},\ t \geq 0, n \geq 0.$$

We are interested in the following three processes. The first is $A(t) = t - S_{N(t)}$, called the *age process*, and the second is $R(t) = S_{N(t)+1} - t$, called the *residual process*. Indeed, $A(t)$ is the length of time since the last replacement prior to time t, and $R(t)$ is the length of time until the next replacement. For the third process, let $L(t) = A(t) + R(t)$, called the *length process*. Note that $L(t) = S_{N(t)+1} - S_{N(t)} = X_{N(t)+1}$. Also, $L(t)$ is the *total* life span of the lightbulb that is functioning at time t. We are interested in the limit distributions of these three sequences of random variables when t goes to infinity.

M. Haviv, *Queues: A Course in Queueing Theory*, International Series in Operations Research & Management Science 191, DOI 10.1007/978-1-4614-6765-6_2,

2.2 Main Renewal Results

2.2.1 The Length Bias Distribution and the Inspection Paradox

Let L be the *length bias* random variable associated with X. Specifically, L has the same support as X but a different density function. Yet, its density function stems from that of X:

$$f_L(\ell) = \frac{\ell f_X(\ell)}{\mathrm{E}(X)},\ \ell \geq 0. \tag{2.1}$$

The definition of the density $f_L(\ell)$ suits cases in which the sampling is favorably biased towards observations with large values, as for example when sampling for the life span is done randomly sometime during the individual's lifetime (and not at birth or at death). Here, the density is not only proportional to the original likelihood but also to the value itself, i.e., the life span. Hence, $f_L(\ell)$ is proportional to the product between ℓ and $f_X(\ell)$. Finally, one has to divide this product by $\mathrm{E}(X)$ in order to get a density function; i.e., the integral between zero and infinity is then equal to one. We claim without a proof that when t goes to infinity, $L(t)$ as defined above, follows the length bias distribution. In other words,

$$\lim_{t\to\infty} \mathrm{P}(L(t) \leq x) = \int_{t=0}^{x} f_L(\ell)\,dt = \int_{\ell=0}^{x} \frac{\ell f_X(\ell)}{\mathrm{E}(X)}\,d\ell,\ \ x \geq 0.$$

The interested reader is referred, e.g. to the text [41], pp. 117–118, for a formal argument. An alternative justification is given below in Sect. 2.3. See Example 3 there. The intuition is clear: a lightbulb whose life span is twice as long as that of another is twice as likely to be sampled when one inspects the process at a random time (and not at the time of replacement).

It is easy to see from (2.1) that $\mathrm{E}(L) = \mathrm{E}(X^2)/\mathrm{E}(X)$ or, in general, that

$$\mathrm{E}(L^n) = \frac{\mathrm{E}(X^{n+1})}{\mathrm{E}(X)},\ \ n = 0, 1, 2, \ldots \tag{2.2}$$

Example 1 (The exponential distribution). It is possible to see that if X is exponentially distributed with parameter λ, then L is with

$$f_L(\ell) = \lambda(\lambda\ell)e^{-\lambda\ell},\ \ \ell \geq 0$$

which is an Erlang distribution with parameters 2 and λ.

Example 2 (The Erlang distribution). The result in the previous example can be generalized as follows. If X follows an Erlang distribution with parameters n and λ, then L follows an Erlang distribution with parameters $n+1$ and λ. Indeed, if $f_X(x) = \lambda(\lambda x)^{n-1}e^{-\lambda x}/(n-1)!$ with mean n/λ, then

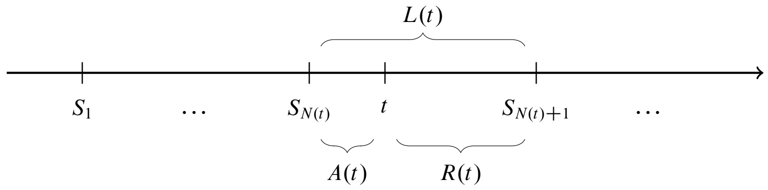

Fig. 2.1 The renewal process

$$f_L(\ell) = \ell \frac{\lambda(\lambda\ell)^{n-1}e^{-\lambda\ell}}{(n-1)!} / \frac{n}{\lambda}$$

$$= \lambda(\lambda\ell)^n \frac{e^{-\lambda\ell}}{n!}, \ \ell \geq 0$$

which is the density function of an Erlang random variable with parameters $n+1$ and λ, and mean $\frac{n+1}{\lambda}$.

Example 3 (The hyper-exponential distribution).

$$f_L(\ell) = \frac{\ell}{\sum_n \frac{p_n}{\lambda_n}} \sum_n p_n \lambda_n e^{-\lambda_n \ell}, \ \ell \geq 0.$$

Example 4 (Mixture of Erlang distributions). Let $\overline{d}$ be the mean number of stages. Then, $\overline{d}/\lambda$ is the mean value of the random variable. Hence, by (1.13),

$$f_L(\ell) = \ell \sum_n p_n \frac{\lambda(\lambda\ell)^{n-1}e^{-\lambda\ell}}{(n-1)!} / (\overline{d}/\lambda).$$

Remark. We next give an explanation for the fact that $E(L) = E(X^2)/E(X)$. Suppose x_i, $i \geq 1$, is a realization of the random variables X_i, $i \geq 1$. Then, the sample average of $L(t)$ along the time interval $[0, T]$ is defined by

$$\overline{L}(T) = \frac{1}{T} \int_{t=0}^{T} L(t)\, dt.$$

Recall that $L(t)$ is the length of the renewal interval covering point t, $t \geq 0$, i.e., $L(t) = x_{N(t)+1}$. Figure 2.1 depicts the case where $x_1 = 1$, $x_2 = 0.5$, $x_3 = 0.7$ and $x_4 = 1.3$.

Clearly,

$$\overline{L}(T) = \frac{1}{T} \left[\sum_{j=1}^{N(T)} x_j^2 + x_{(N(T)+1)}(T - \sum_{j=1}^{N(T)} x_j) \right]$$

or

$$\overline{L}(T) = \frac{N(T)}{T} \frac{1}{N(T)} \left[\sum_{j=1}^{N(T)} x_j^2 + x_{N(T)+1}(T - \sum_{j=1}^{N(T)} x_j) \right].$$

When $T \to \infty$, $N(T)/T$ goes with probability one (see, e.g. [41], p. 133) to $1/\mathrm{E}(X)$ as the latter is the renewal rate. Moreover, as its second term goes to zero,

$$\frac{1}{N(T)} \left[\sum_{j=1}^{N(T)} x_j^2 + x_{N(T)+1}(T - \sum_{j=1}^{N(T)} x_j) \right]$$

goes to $\mathrm{E}(X^2)$ with probability one. In summary,

$$\lim_{T \to \infty} \overline{L}(T) = \frac{\mathrm{E}(X^2)}{\mathrm{E}(X)}$$

as required.

The inspection paradox. The fact that $\mathrm{E}(L) = \mathrm{E}(X^2)/\mathrm{E}(X)$ leads immediately to the conclusion that $\mathrm{E}(L) \geq$E(X) with an equality if and only if X is deterministic. This inequality is known as the *inspection paradox.* Suppose for example that the renewal process under consideration is that of lightbulbs that are replaced one by another as soon as one is burnt out. One who inspects the current functioning lightbulb assesses the distribution of the lifespan of this light bulb (age plus residual) by that as the distribution of L, in particular its mean equals $\mathrm{E}(X^2)/\mathrm{E}(X)$ which is greater than or equal to $\mathrm{E}(X)$. Thus, an inspected light bulb is on average better than an average light bulb! This seams as a paradox. Yet, as we have seen throughout this section, among those lightbulbs which are inspected there is a bias towards the long lightbulbs due to the fact that long ones are more likely to be sampled (even if under the distribution of X they are equally likely).

2.2.2 The Age and the Residual Distributions

When an individual is sampled during his lifetime, it makes sense to define the following two random variables. Specifically, denote two nonnegative random variables by A and R, called *age* and *residual*, respectively. They are the limit random variables of $A(t)$ and $R(t)$ as defined in Sect. 2.1. The former reflects the age of the sampled individual, while the latter reflects how much life is still ahead of him. Of course, their sum gives the total life span. Assuming that for a given life span, all points of time during one's life are equally likely to be sampled. Hence, conditioning on $L = x$, we assume that the age follows a uniform distribution whose support is the $[0, x]$ interval. Formally,

$$f_{A|L=x}(a) = \begin{cases} \frac{1}{x} & a \le x \\ 0 & a > x. \end{cases} \tag{2.3}$$

Since $R = L - A$, we conclude that $R|L = x$ follows the same distribution as $A|L = x$. Moreover, the marginal, i.e., unconditional, distributions of both A and R are identical. Indeed, note the symmetry here between A and R: if one reverses the orientation of time, then age and residual life swap their meanings. An alternative way to define A and R given L, is to say that $(A, R) = (UL, (1-U)L)$ when U is a continuous zero-one uniformly distributed random variable that is independent of L.

Remark. Our point of departure here was that $A|L$ was distributed uniformly in $[0, L]$. However, it is possible to prove formally that

$$\lim_{t\to\infty} \mathrm{P}(A(t) \le aL(t) = \ell) = \frac{a}{l}, \quad 0 \le a \le \ell$$

from which the joint distribution of A and L (and hence the conditional distribution of $A|L$) follow. See Exercise 9 for the approach suggested below in Sect. 2.3.

Next we find the marginal density function of A (and hence of R):

$$f_A(a) = \int_{x=0}^{\infty} f_{A|L=x}(a) f_L(x)\, dx = \int_{x=a}^{\infty} \frac{x f_X(x)}{\mathrm{E}(X)} \frac{1}{x}\, dx = \frac{\overline{F}_X(a)}{\mathrm{E}(X)}\,, a \ge 0\,. \tag{2.4}$$

Note that $f_A(a)$ is monotone decreasing with a, $a \ge 0$. The intuition behind that is simple: if $a \le b$, then whoever's current age is b has been at age a at some time in the past. The converse is not always true: If one's age has been a it is not necessarily true that he/she will reach age b.

Equation (2.4), coupled with (1.3), leads to the fact that $\mathrm{E}(A) = \mathrm{E}(X^2)/2\mathrm{E}(X)$. In general, from (1.3) we can learn that

$$\mathrm{E}(R^n) = \mathrm{E}(A^n) = \int_{x=0}^{\infty} x^n \frac{\overline{F}_X(x)}{\mathrm{E}(X)}\, dx = \frac{\mathrm{E}(X^{n+1})}{(n+1)\mathrm{E}(X)}, \; n \ge 0. \tag{2.5}$$

As $L = A + R$ and as $\mathrm{E}(A) = \mathrm{E}(R)$, we conclude that $\mathrm{E}(A) = \mathrm{E}(L)/2$ or $\mathrm{E}(A) = \mathrm{E}(R) = \mathrm{E}(L)/2$, whereby we get an alternative proof that $\mathrm{E}(A) = \mathrm{E}(R) = \mathrm{E}(X^2)/2\mathrm{E}(X)$. This result should be compared with *Solomon's Wisdom.* Finally, since the conditional distribution of A and R given L coincide, the same is true of their marginal distributions. In summary, A and R are identically distributed. Of course, in general, they are not independent.

Remark. The counterpart of Fig. 2.2 for the sample paths of the age and residual processes is given below in Figs. 2.3 and 2.4, respectively. Comparing the three figures, it is clear that $\mathrm{E}(A) = \mathrm{E}(R) = \mathrm{E}(L)/2$.

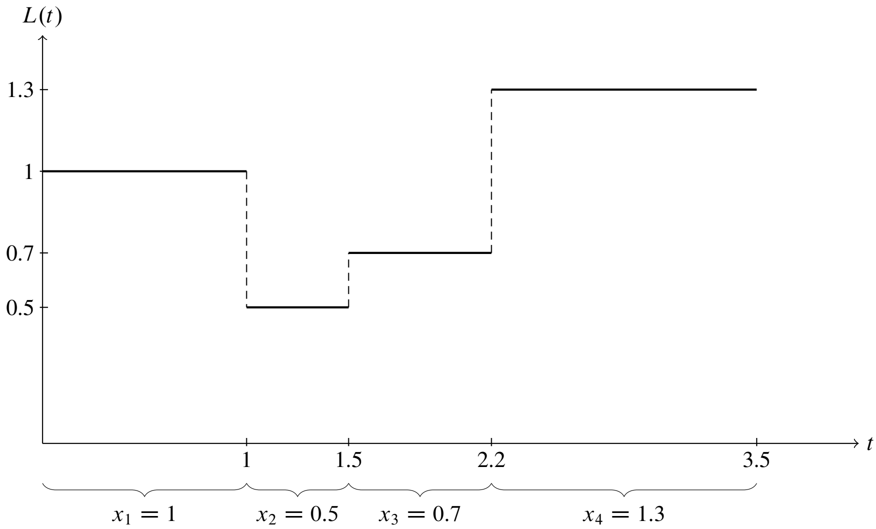

Fig. 2.2 The length process

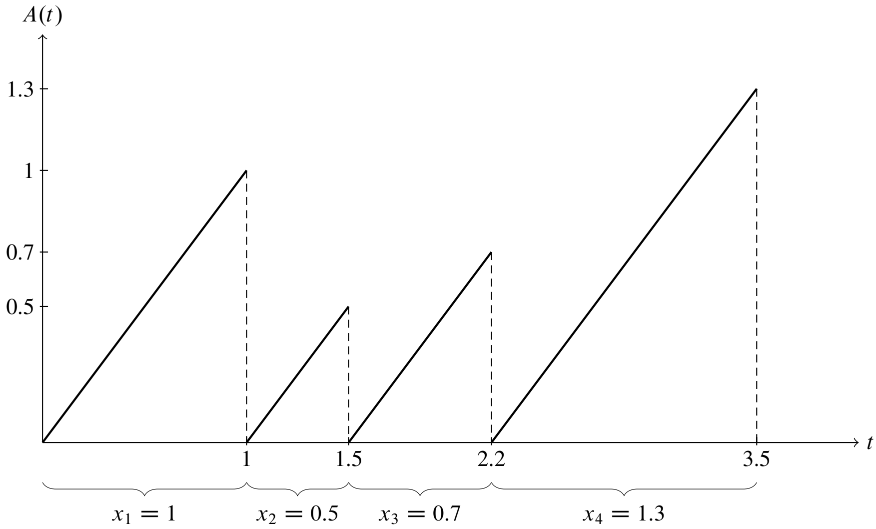

Fig. 2.3 The age process

Remark. Trivially, $\mathrm{E}(A) = \mathrm{E}(R) \leq \mathrm{E}(L)$. Yet, all the following three options are possible: $\mathrm{E}(A) < \mathrm{E}(X)$, $\mathrm{E}(A) = \mathrm{E}(X)$ and $\mathrm{E}(A) > \mathrm{E}(X)$. The third option might look counterintuitive but it is possible. This phenomenon can be explained by the inspection paradox as it possible that $\mathrm{E}(L) \geq 2\mathrm{E}(X)$. See Exercise 5 for an example.

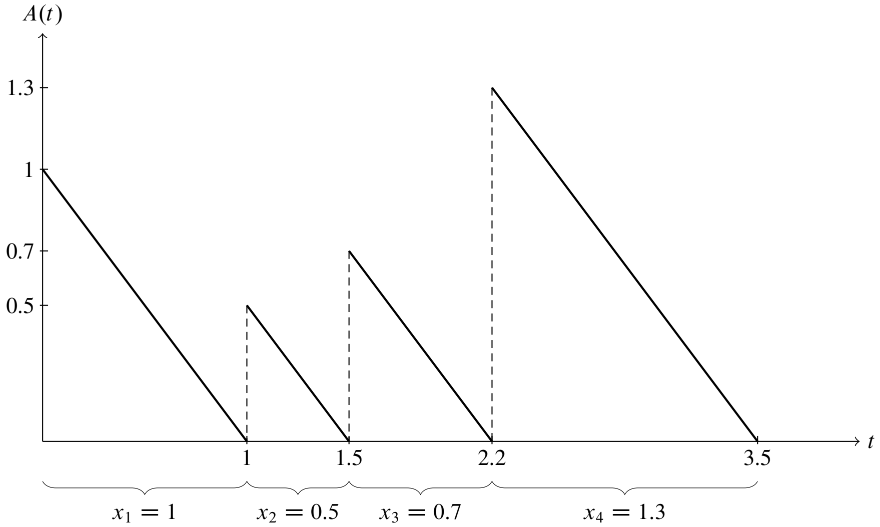

Fig. 2.4 The residual process

*Example 1 (*cont.). Note that in the case where X follows an exponential distribution with parameter λ

$$f_R(a) = f_A(a) = \frac{e^{-\lambda a}}{1/\lambda} = \lambda e^{-\lambda a}\ , a \geq 0.$$

In other words, the age (as well as the residual) follows an exponential distribution with the same parameter of λ. It is also possible to show that A (or R) and X follow the same distribution, only if X is exponential. In fact R (or A) and X having the same distribution can be looked at as an alternative definition of the memoryless property.

*Example 2 (*cont.). In the case where X follows an Erlang distribution with parameters n and λ,

$$\overline{F}_X(x) = \sum_{k=0}^{n-1} e^{-\lambda x} \frac{(\lambda x)^k}{k!} \text{ and } \mathrm{E}(X) = n/\lambda,$$

(see (1.11)) we get that

$$f_R(a) = f_A(a) = \frac{1}{n} \sum_{k=1}^{n} \lambda e^{-\lambda a} \frac{(\lambda a)^{k-1}}{(k-1)!}\ .$$

This distribution can be seen as a mixture between n random variables, each of which is Erlang with the same scale parameter of λ but with a different number of stages, where the number of stages is uniformly distributed between $1, 2, \ldots, n$. This leads to the following observation. When one inspects a component whose longevity follows an Erlang distribution with parameters n and λ, i.e., the sum of n random stages (which are independent and exponentially distributed, and share the same expected length of λ^{-1}), then the index of the current stage (and the number of stages to be completed inclusive of the current one) are uniformly distributed between 1 and n. In other words, given stage d, the age follows an Erlang distribution with parameters d and λ while the residual follows an Erlang distribution with parameters $n - d + 1$.[1] Since all stages are equally likely, we conclude that

$$\mathrm{E}(R) = \mathrm{E}(A) = \frac{n+1}{2\lambda}.$$

The following lemma contains two facts. The first is trivial while for the second we supply a short proof.

Lemma 2.1.

$$f_{L|A=a}(l) = \frac{f_X(l)}{\overline{F}_X(a)}, \;\; l \geq a \geq 0.$$

Of course, the corresponding value when $l < a$ is zero. Also,

$$f_{(A,R)}(a, r) = \frac{f_X(a+r)}{\mathrm{E}(X)}, \; r, a \geq 0. \tag{2.6}$$

Proof of Equation (2.6).

$$f_{(A,R)}(a, r) = f_A(a) f_{R|A=a}(r) = \frac{\overline{F}_X(a)}{\mathrm{E}(X)} \frac{f_X(a+r)}{\overline{F}_X(a)} = \frac{f_X(a+r)}{\mathrm{E}(X)}.$$

□

Remark. Note that the joint density in Eq. (2.6) is a function of a and r only through their sum. This is not a surprise given that $L = A + R$ and that both $A|L$ and $R|L$ are uniformly distributed in $[0, L]$.

2.2.3 The Memoryless Property (Versions 4 and 5)

As can be seen, the joint density function (2.6) does not in general equal to $f_A(a) f_R(r)$, and hence A and R are not necessarily independent. These properties

[1] The current stage is counted both in terms of age and residual lifetime.

do obtain, however, when X follows an exponential distribution, as the reader can easily check. We next show that this is the only case resulting in such an independence. Thus, the independence of A and R is equivalent to memorylessness. Put differently, the mutual independence of the age and the residual is in fact equivalent to the memoryless property. To show this, we need the following lemma:

Lemma 2.2.

$$P(A \geq a, R \geq r) = P(A \geq a + r) = P(R \geq a + r). \tag{2.7}$$

Proof. The right equality is trivial as A and R are identically distributed. The left equality follows from Eq. (2.6):

$$\begin{aligned} P(A \geq a, R \geq r) &= \int_{a' \geq a} \int_{r' \geq r} f_{(A,R)}(a', r')\, da'\, dr' \\ &= \frac{1}{E(X)} \int_{a' \geq a} \int_{r' \geq r} f_X(a' + r')\, da'\, dr' \\ &= \int_{a' \geq a} \frac{\overline{F}_X(a' + r)}{E(X)}\, da' = \int_{a' \geq a} f_A(a' + r)\, da' = P(A \geq a + r), \end{aligned}$$

as required. □

Remark. Note from (2.7) that $P(A \geq a, R \geq r)$ is a function of a and a only through $a + r$.

Equation (2.7) immediately leads to the following:

$$P(A \geq a | R \geq r) = \frac{P(A \geq a, R \geq r)}{P(R \geq r)} = \frac{P(R \geq a + r)}{P(R \geq r)} . \tag{2.8}$$

Theorem 2.1. *The random variables A and R are independent if and only if X follows an exponential distribution.*

Proof. From (2.8) and the fact that A are R are identically distributed, we learn that A and R are independent if and only if

$$\frac{P(R \geq a + r)}{P(R \geq r)} = P(R \geq a) \ .$$

This is equivalent to R possessing the memoryless property (see (1.7)). Hence,

$$f_R(r) = \lambda e^{-\lambda r}, r \geq 0, \tag{2.9}$$

for some $\lambda > 0$. Then, by (2.4), $\overline{F}_X(x) = E(X)\lambda e^{-\lambda x}, x \geq 0$. This implies that X has an exponential tail which is possible if and only if X is exponentially distributed. This concludes the proof. □

Remark. To conclude: $f_{A,R}(a,r) = f_A(a) f_R(r)$, a, $r \geq 0$, namely the age and the residual being independent, is our fourth version of the memoryless property. It is possible to see from (2.9) that this is equivalent for the residual (and hence the age) to follow an exponential distribution. Hence, this will be our fifth version for memoryless.

Later on we will need the Laplace transforms of A, R, and L in terms of the Laplace transform of the random variable X. They are related as follows:

Lemma 2.3. *Let $F_X^*(s)$ be the LST of the nonnegative and continuous random variable X. Then,*

$$F_R^*(s) = F_A^*(s) = \frac{1 - F_X^*(s)}{\mathrm{E}(X)s} \tag{2.10}$$

is the LST of the age (and the residual) distribution. Also,

$$F_L^*(s) = -\frac{d\, F_X^*(s)}{d\, s} \frac{1}{\mathrm{E}(X)} \tag{2.11}$$

is the LST of the length bias distribution.

Proof.

$$F_A^*(s) = \frac{\int_{x=0}^{\infty} \overline{F}_X(x) e^{-sx}\, dx}{\mathrm{E}(X)}$$

which by integration by parts, equals

$$= -\overline{F}_X(x) e^{-sx} \frac{1}{s\mathrm{E}(X)} |_{x=0}^{\infty} - \int_{x=0}^{\infty} \frac{1}{s\mathrm{E}(X)} f_X(x) e^{-sx}\, dx = \frac{1}{s\mathrm{E}(X)} - \frac{F_X^*(s)}{s\mathrm{E}(X)}$$

as required. The proof of (2.11). □

The final limit result we would like to mention (without a proof) concerns the limit probability of a renewal during the next instant of time.

Theorem 2.2.

$$\lim_{t \to \infty} \mathrm{P}(N(t + \Delta t) - N(t) = 1) = \frac{1}{\mathrm{E}(X)} \Delta t + o(\Delta t).$$

As $\mathrm{E}(X)$ is the expected time between renewals, $1/\mathrm{E}(X)$ is the rate of renewals, i.e., the expected number of renewals per unit of time. What Theorem 2.2 states is stronger than an *average* rate result. It states that this rate, when looked at as a probability of renewal, holds, at the limit, for any instant of time. Alternatively, by (1.4),

$$\int_{x=0}^{\infty} h_X(x) f_A(x)\, dx = \int_{x=0}^{\infty} \frac{f_X(x)}{\overline{F}_X(x)} \frac{\overline{F}_X(x)}{\mathrm{E}(X)}\, dx = \frac{1}{\mathrm{E}(X)}.$$

In words, the average hazard rate with respect to the age distribution equals $1/\mathrm{E}(X)$; i.e., the average failure rate, which is in fact the average renewal rate, is the reciprocal of the mean time between renewals. Finally, note that in the case of exponential distribution, the theorem holds for any time and not only at the limit.

2.3 An Alternative Approach

Let $Z_i = X_i + Y_i$, $i \geq 1$, be a series of independent and identically distributed nonnegative and continuous random variables. The same can be said of the two series X_i and Y_i but, and this is worth noting, for a given i, X_i and Y_i are not necessarily independent. Consider now a renewal process in which any Z_i is followed by Z_{i+1}, $i \geq 1$. Moreover, each period of the Z_i, commences with X_i and then is followed by Y_i. We say that the process is in an "on" mode if currently an X_i is running and it is said to be "off" if currently a Y_i is running. Quite naturally, a realization of a Z_i is called a *cycle*.

It is clear that the process moves consecutively from an "on" mode to an "off" mode, then to an "on" mode again, etc.[2] We are interested in the limit probability that the process in "on". It is claimed here without a proof that it equals

$$\mathrm{P}(\text{“on”}) = \frac{\mathrm{E}(X)}{\mathrm{E}(X) + \mathrm{E}(Y)}. \tag{2.12}$$

The result is quite intuitive but it is somewhat surprising that it holds also when X and Y are not independent.

Example 1. Suppose a machine works for a time whose length X follows a uniform distribution in the unit interval. When the machine breaks, it undergoes repair which lasts $Y = X^2$. Clearly, $\mathrm{E}(X) = 1/2$ and $\mathrm{E}(Y) = 1/3$. Thus, the long-term probability that the machine is operational is

$$\frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{3}} = \frac{3}{5}\,.$$

Example 2. Our first application of the above result is in deriving the limit distribution of the age process. Fix an age of a and say that the process is "on" as long as its age is less than or equal to a and it is "off" otherwise. Note that in this example the "on" and the "off" periods are clearly not independent. Also, an "on" period can be followed with a zero length "off" period (if $x \leq a$). The expected cycle length is of course $\mathrm{E}(X)$. The expected time of "on" is $\mathrm{E}(\min\{a, X\})$. The latter value equals

[2] This does not rule out the possibility the length of one of these phases will equal zero.

$$E(\min\{a, X\}) = \int_{t=0}^{a} \overline{F}_X(t)\, dt,$$[3]

Hence, by (2.12),

$$F_A(a) = \frac{\int_{t=0}^{a} \overline{F}_X(t)\, dt}{E(X)}.$$

Taking derivative with respect to a we conclude that

$$f_A(a) = \frac{\overline{F}_X(a)}{E(X)},$$

which coincides with (2.4).

Example 3. Next we derive the limit distribution of the length of the entire renewal period when sampled at an arbitrary point of time (which from the above we know follows the length bias distribution). Now we say that the process is "on" when the length of the renewal interval is less than or equal to l (and it is "off" when it is greater than l). Note that the product between the lengths of the "on" and "off" periods is always zero. As before, $E(X)$ is the expected length of the cycle. Also, the expected "on" time equals $\int_{x=0}^{l} x f_X(x)\, dx$. Note that when $x \geq l$, the "on" period equals zero. Hence, by (2.12),

$$F_L(l) = \frac{\int_{x=0}^{l} x f_X(x)\, dx}{E(X)}, \ l \geq 0.$$

Taking the derivative with respect to l, we get that

$$f_L(l) = \frac{l f_X(l)}{E(X)}, \ l \geq 0,$$

which coincides with (2.1).

2.4 A Note on the Discrete Version

Suppose the life span of a component is measured only by whole numbers, say days. Then, $P(X = i) = p_i$, $i \geq 1$, for some nonnegative numbers that sum up to one. Denote by q_i, $i \geq 1$, the probability that the life span is at least i, i.e.,

$$q_i = P(X \geq i) = \Sigma_{j=i}^{\infty} p_j, i \geq 1.$$

[3]See (1.3) for the case where $n = 0$ and for the random variable $\min\{a, X\}$. Note that the tail function for all values from a and above equal zero.

It is possible to see that

$$E(X) = \Sigma_{i=1}^{\infty} q_i \tag{2.13}$$

(see Exercise 1 for a proof), which is the counterpart of (1.3) when $n = 1$ for integral random variables. Also, $h_i = p_i/q_i = P(X = i|X \geq i)$ is the *hazard* at i, $i \geq 1$ and it plays the same role as the hazard defined in (1.4) for continuous random variables.

Next, the length bias distribution in the case of discrete random variables is defined via

$$P(L = \ell) = \frac{\ell p_\ell}{E(X)}, \ \ell \geq 1\,. \tag{2.14}$$

Hence,

$$E(L) = \frac{E(X^2)}{E(X)} \tag{2.15}$$

as in the continuous version. Also, the age distribution is defined via

$$P(A = a|L = \ell) = \begin{cases} \frac{1}{\ell} & 1 \leq a \leq \ell \\ 0 & a > \ell \end{cases}$$

Hence,

$$P(A = a) = \sum_{\ell=a}^{\infty} P(L = \ell)P(A = a|L = \ell) = \sum_{\ell=a}^{\infty} \frac{\ell p_\ell}{E(X)} \frac{1}{\ell} = \frac{q_a}{E(X)}, \quad a \geq 1. \tag{2.16}$$

The residual here has the same meaning as in the continuous case but some care is needed due to the integrality requirement. Specifically, as we like R and A to be identically distributed, we need to define R as $L + 1 - A$ and hence both age and residual are *inclusive* of the current day. Clearly then,

$$P(A = a, R = r) = \frac{P(L = a + r - 1)}{a + r - 1} = \frac{p_{a+r-1}}{E(X)}, \quad a, r \geq 1\,. \tag{2.17}$$

Due to the double counting of the current day, $E(A) + E(R) - 1 = E(L)$. Since $E(A) = E(R)$, we conclude by (2.15) that

$$E(A) = E(R) = \frac{1}{2}\left[\frac{E(X^2)}{E(X)} + 1\right]. \tag{2.18}$$

Example 4 (Geometric random variables). Suppose X follows a geometric distribution with parameter p, i.e., $P(X = i) = p(1 - p)^{i-1}$, $i \geq 1$. Then, $q_i = (1 - p)^{i-1}$, $i \geq 1$, $E(X) = 1/p$ and $E(X^2) = (2 - p)/p^2$. Hence, $P(L = \ell) = \ell p^2(1 - p)^{\ell-1}$, $\ell \geq 1$, and $E(L) = (2 - p)/p$. Note that L follows a negative binomial distribution with parameters 2 and p, counting the number of trials until (exclusive) the second success. Also, $\mathrm{P}(A = a) = p(1-p)^{i-1}$, $i \geq 1$, which is the memoryless version of discrete random variables. The same of course is the distribution of R. Also, by (2.17), for any $a, r \geq 1$,

$$\mathrm{P}(A = a, R = r) = p(1-p)^{a+r-2}/(1/p) = p(1-p)^{a-1}p(1-p)^{r-1}$$
$$= \mathrm{P}(A = a)\mathrm{P}(R = r).$$

In other words, A and R are independent. Again, this is the memoryless phenomenon. Finally, by (2.18)

$$\mathrm{E}(A) = \mathrm{E}(R) = \frac{1}{2}(\frac{2-p}{p} + 1) = \frac{1}{p},$$

as expected.

More involved is the case where we have a random sum of independent and identically distributed random variables. Specifically, let $Y = \Sigma_{i=1}^{N} X_i$, where $\{X_i\}_{i=1}^{\infty}$ are independent and identically distributed random variables and N is an independent discrete random variable. Let L_X, L_Y, and L_N be the length bias distribution of X, Y, and N, respectively. Note that the special case where N is constant and X_i, $1 \leq i \leq N$, follows exponential distribution is dealt with in Example 2 in Sect. 2.2.1 since Y now follows an Erlang distribution.

Theorem 2.3. *L_Y is distributed as $L_X + \Sigma_{i=1}^{L_N - 1} X_i$, where the summation here is between independent random variables.*

Proof. See Exercise 4.

A possible example is the case where Y is a mixture of Erlang random variables: the X_i's, $i \geq 1$, are exponentially distributed and N has some discrete distribution. Then, by Example 2 of Sect. 2.2.1, L_X has an Erlang distribution whose first parameter equals two. The second summand is also a mixture of Erlang random variables.

2.5 Exercises

1. Prove formula (2.13).
2. Derive the density function of the age in the case where the original random variable follows hyper-exponential distribution.
3. Derive the density function of the age in the case where the original random variable is a mixture of Erlang random variables.
4. Prove Theorem 2.3 and state it in your own words.

5. Recall that if X follows an Erlang distribution with parameters n and λ, then X is in fact a sum of n independent and exponentially distributed random variables with parameter λ, called stages. Let $p_i(t)$ be the probability that the stage at time t is i, given that $X \geq t$, $1 \leq i \leq n$.

   (a) Find $p_i(t)$, $1 \leq i \leq n$.
   (b) Show that $h_X(t) = \lambda p_n(t)$. State this result in your own words.

6. Show that if X is with a DHR distribution then $\mathrm{E}(A) \geq \mathrm{E}(X)$. Also, show that $\mathrm{E}(A) \geq \mathrm{E}(X)$ if and only if the coefficient of variation of X is greater than or equal to 1.
7. Show that the following families of continuous distributions are closed in the sense that if X's distribution belong to them, the same is the case with L's. Specifically,

   (a) If $X \sim \Gamma(\alpha, \beta)$ then $L \sim \Gamma(\alpha + 1, \beta)$.
   (b) If $X \sim \mathrm{beta}(\alpha, \beta)$ then $L \sim \mathrm{beta}(\alpha + 1, \beta)$

8. Show that the following families of discrete distributions are closed in the sense that if X's distribution belong to them, the same is the case with $L - 1$'s. Specifically,

   (a) If $X \sim \mathrm{Bin}(n, p)$ then $L - 1 \sim \mathrm{Bin}(n - 1, p)$
   (b) If $X \sim \mathrm{NB}(r, p)$, then $L - 1 \sim \mathrm{NB}(r + 1, p)$
   (c) If $X \sim \mathrm{Pois}(\lambda)$, then $L - 1 \sim \mathrm{Pois}(\lambda)$. Moreover, show that if X and $L - 1$ follow the same distribution then X follows a Poisson distribution.

9. Define a renewal process as being 'on' when its age is larger than or equal to a and when its residual is larger than or equal to r.[4]

   (a) Show that the expected time in which the process is 'on' during one renewal period equals
   $$\mathrm{E}(\max\{X - a - r, 0\}).$$
   Express this expected value in terms of $f_X(x)$ and/or $F_X(x)$.
   (b) Deduce the limit joint distribution of A and R. In particular, show that
   $$f_{A,R}(a, r) = \frac{f_X(a + r)}{\mathrm{E}(X)}.$$

10. Define a renewal process as being 'on' when its age is smaller than or equal to a and when its residual is smaller than or equal to r.[5]

   (a) Show that the expected time in which the process is 'on' during one renewal period equals

[4]This exercise is due to Yoav Kerner.

[5]This exercise is due to Binyamin Oz.

$$\mathrm{E}(\min\{a, X\} + \min\{r, X\} - \min\{a + r, X\}).$$

Express this expected value in terms of $f_X(x)$ and/or $F_X(x)$.

(b) Deduce the limit joint distribution of A and R. In particular, show that

$$f_{A,R}(a, r) = \frac{f_X(a + r)}{\mathrm{E}(X)}.$$

http://www.springer.com/978-1-4614-6764-9

Queues
A Course in Queueing Theory
Haviv, M.
2013, XIV, 221 p., Hardcover
ISBN: 978-1-4614-6764-9