Chapter 2
Models for Perfect Repair

2.1 Introduction

According to the definition given by Ascher and Feingold [2], a repairable system is understood to be a system which, after failure, can be restored to a functioning condition by some maintenance action other than replacement of the entire system. Replacing the entire system may be an option, but it is not the only one. In this part of the book, we will assume that the description of the system state at any time is reduced to two levels: operative and failed. More detailed specifications of the state space are considered in Part III of the book.

Model deterioration (performance) of a repairable system can be tackled in several different ways. On the one hand, the interest may lie mainly in modeling the number of failures suffered by the system up to time t. If $N(t)$ is the number of failures of a repairable system occurring in the interval $(0, t]$, the most appropriate approach is to consider the counting process given by $\{N(t),\ t \geq 0\}$ as model deterioration. Attention is usually focused on the expected value, variance, and probability distribution of $N(t)$. The homogeneous Poisson process (*HPP*) is the counting process most frequently used throughout the extensive literature on the subject. *HPP* may also be characterized in terms of the random length of the times between two consecutive failures, exponentially distributed with the same parameter.

Data from repairable systems are usually given as times between failures $T_1, T_2, \ldots$. A common assumption made is that these failure times are independent and identically distributed, with distribution F. This assertion implies that after a failure, the system behavior is exactly the same as if it were new; thus, a perfect repair maintenance action is being carried out in the system environment. As explained in Kijima [26], in practice, the perfect repair assumption may be reasonable for systems with one structurally simple unit.

When F denotes a general family of distributions, the sequence $\{T_1, T_2, \ldots\}$ is referred to as a renewal process (*RP*). Therefore, the *HPP* may be seen as a

M. L. Gámiz et al., *Applied Nonparametric Statistics in Reliability*,
Springer Series in Reliability Engineering, DOI: 10.1007/978-0-85729-118-9_2,

particular case of *RP*. Of course, there is an evident duality between "time domain" and "counts domain", i.e., $\{N(t) \geq k\}$ if and only if $\{T_1 + T_2 + \cdots + T_k \leq t\}$. Putting this into words, there have been at least k renewals until time t if and only if the kth renewal has occurred before t.

Another general assumption made when using a counting process to model the time-evolution of a repairable system is that the time-to-repair is negligibly small compared to its time-to-failure. In many practical applications, where it is reasonable to expect that the system is not under repair for long in relation to its operating time, this assumption is fairly realistic. Otherwise, the system is not feasible.[1] So, in a case where it is assumed that the system is repaired and put into new operation immediately after the failure, the deterioration model will be given by an Ordinary Renewal Process (*ORP*).

On the other hand, there are some situations where one is interested in estimating other important measures such as availability (unavailability) of the system, which is the probability that the system is functioning at a given time. In this case, the modeling tool indicated is the Alternating Renewal Processes (*ARP*) where operative periods alternate with repair periods. Within the scope of an *ARP*, data collected consist of a sequence of alternating lifetimes and repair times, i.e. (T_1, R_1), (T_2, R_2), …, where T_1, T_2, … are the successive lifetimes of the system and these are independent and identically distributed (*i.i.d.*) with *CDF F*; and R_1, R_2, …, the corresponding repair times, which are i.i.d. with *CDF G*. It is also assumed that T_i and R_i are independent, for any $i = 1, 2, \ldots$. Every random length obtained as a lifetime plus a repair time is called a renewal cycle. Repairable system data are collected to estimate among other measures, quantities such as:

- The distribution of lifetimes (respectively, repair times);
- The expected number of renewals in an interval $(0, t]$, which is the renewal function;
- The probability that the system is operative at a given time t, which is the instantaneous availability;
- The proportion of time the system is in a functioning condition, which is the steady state or limiting availability.

Inference studies are carried out without assuming any particular functional form for distribution functions F and G. We therefore use a nonparametric approach. From a given data set, empirical estimators are constructed for the performance measures of a repairable system whose time evolution is modeled by an *ORP* or an *ARP* and we also obtain smooth estimators based on kernel functions. The implicit bandwidth parameter is derived by means of data-driven procedures, specifically bootstrap techniques, which prove very easy to implement and give very good results, as pointed out in the simulation examples included.

[1] Sometimes one may let "operating time" be the time parameter; or possibly "number of cycles" or "number of kilometers" (for cars). Then, repair times are 0 in these time axes.

2.2 Ordinary Renewal Process

In this section, we study probabilistic models for systems where after the occurrence of a random event (failure), at a random time, everything in the system starts over at the beginning. So, we consider systems under perfect repair maintenance policies, which means that the system operating state is restored to "as good as new" conditions after failure. This approach is appropriate for systems such as light bulbs or thermometers, where the occurrence of failures implies the substitution of the entire system, not being available partial repairs to recover the system function. Furthermore, if a reliability system is understood as a set of components or elementary parts, renewal processes are plausible models for the time behavior of the parts better than for the whole system, since after a failure occurs in a component, the replacement of the component is usually carried out, instead of repair.

Under perfect repair models, at time $t = 0$ a repairable system is put into operation and is functioning. At each failure time, the system is replaced by a new one of the same type. This process is repeated along time, and the replacement time is considered negligible. As a result, a sequence of lifetimes or random variables which are independent and identically distributed is obtained. Renewal processes have been extensively used by many researches interested in reliability (Barlow and Proschan [3] or Rausand and Hoyland [36] are classical references), the most simple case being the Homogenous Poisson Process (*HPP*), where the random time between successive renewals has an exponential distribution.

2.2.1 The Renewal Function

Renewal theory arises from the study of stochastic systems whose time evolution appear as successive life cycles. A life cycle is a time interval during which the system is functioning. At the start of every interval, the system is stochastically reinitiated. In this section, we introduce the main features that characterize an ordinary renewal process, paying special attention to the renewal function.

As stated above, an *ORP* may be represented by means of a sequence of random independent and identically distributed variables $\{T_k; k = 1, 2, \ldots\}$ (we will consider only non-negative variables) or equivalently by means of the counting process $\{N(t);\ t \geq 0\}$, where $N(t) = max\{k\colon T_1 + T_2 + \cdots + T_k \leq t\}$, that is, the number of renewals occurring in the interval $(0, t]$. Let F denote the *CDF* common to all T_k, and let us define $S_0 = 0$ and $S_k = T_1 + T_2 + \cdots + T_k$, for $k = 1, 2, \ldots,$ as the so-called *waiting times*, making it obvious that $P\{N(t) \geq k\} = P\{S_k \leq t\}$. This random quantity, S_k is obtained as the sum of k independent random variables, so that its *CDF*, which we call F_k, is given by the k-fold *Stieltjes-convolution* of F for $k \geq 1$, that is,

$$F^{(k)} = P\{S_k \le t\} = F^{(k-1)} * F(t) = \int_0^t F^{(k-1)}(t-u)dF(u).$$

The main objective in renewal theory is to derive the properties of $N(t)$, in particular its expected value, which is called the *renewal function*.

Definition 2.1 (*Renewal Function*) Let $F(t) = P\{T \le t\}$ be the *CDF* of the lifetime of a repairable system with perfect and instantaneous repair. Let $\{N(t), t \ge 0\}$ be the corresponding renewal counting process. The renewal function is defined as

$$M(t) = E[N(t)], \quad \text{for } t \ge 0.$$

It is easy to check the following equalities

$$M(t) = E[N(t)] = \sum_{k=1}^{\infty} P\{N(t) \ge k\} = \sum_{k=1}^{\infty} P\{S_k \le t\} = \sum_{k=1}^{\infty} F^{(k)}(t).$$

It can also be stated that $M(t) < \infty$ for all $0 \le t < \infty$. Furthermore, the expression above may be given via the following integral representation

$$M(t) = F(t) + \int_0^t M(t-u)dF(u) = F(t) + M(t) * F(t), \quad \text{for } t \ge 0, \tag{2.1}$$

which is a particular case of a wider class of equations called *renewal equations*,

$$W(t) = v(t) + \int_0^t W(t-u)dF(u) = v(t) + W(t) * F(t), \quad \text{for } t \ge 0,$$

where $v(t)$ and $F(t)$ are known, whereas $W(t)$ is an unknown function. In other words, $M(t)$ satisfies the renewal equation given by (2.1), and moreover, it is the unique solution that is bounded on finite intervals.

Closed form analytic expressions for $F^{(k)}$ are not generally available, special cases are Erlang and Normal distributions. Based on a central result of renewal theory, the ***key renewal theorem*** (there exists an extensive literature over the subject, see for instance [37]), simple asymptotic approximations can be obtained in the case where $E[T] = \mu < \infty$ and $Var[T] = \sigma^2 < \infty$,

$$\lim_{t\to\infty}\left[M(t) - \frac{t}{\mu}\right] = \frac{\sigma^2}{2\mu^2} - \frac{1}{2}.$$

This expression suggests the following asymptotic expression for the renewal function

$$M_\infty(t) = \frac{t}{\mu} + \frac{\sigma^2}{2\mu^2} - \frac{1}{2}.$$

With these considerations, the following estimator for the renewal function, valid for large values of t, may be defined,

$$\widehat{M}_\infty(t) = \frac{t}{\widehat{\mu}} + \frac{\widehat{\sigma}^2}{2\widehat{\mu}^2} - \frac{1}{2},$$

where $\widehat{\mu}$ and $\widehat{\sigma}$ are estimators of μ and σ respectively, based on data recorded up to time t. Nevertheless, in certain application areas such as reliability engineering, the interest is rather in the initial part of the life of a device, that is, for $t \in [0, 3\mu]$, see Frees [13] and Gertsbakh and Shpungin [17], where the estimation problem becomes more difficult. It seems natural to estimate the renewal function based on a sum of estimators of the convolutions of F, that is, we define

$$\widehat{M}(t) = \sum_{k=1}^{\kappa} \widehat{F}^{(k)}(t) \tag{2.2}$$

where the number of terms in the summation, the parameter κ, has to be determined. Various ways have been proposed by different authors in the literature on the subject.

2.2.2 *Nonparametric Estimation of the k-Fold Convolution of Distribution Functions*

The problem of dealing with the function M involves estimating k-fold convolution functions, which is not an easy task. Recently, a number of authors have tackled the problem, revealing the inherent difficulty in most cases. Below, we present some of these. Let $T_1, T_2, \ldots, T_n$ be non-negative independent random variables, with cumulative distribution function F. Let $F^{(k)}$ be the k-fold convolution function of F.

2.2.2.1 The Empirical Convolution Function

Frees [14] defines two alternative estimators of the renewal function. The first is based on the sum of the convolutions without replacing the empirical distribution function, and the second, called the *empirical renewal function*, is obtained as the renewal function of the empirical distribution corresponding to F. Let us introduce the estimators of the convolutions which Frees defines for constructing estimators of type $\widehat{M}(t)$ as given above.

If $\{i_1, i_2, \ldots, i_k\}$ is a subset of size k of $\{1, 2, \ldots, n\}$, then an estimator of $F^{(k)}(t)$ is

$$\widehat{F}_{C1}^{(k)}(t) = \frac{1}{\binom{n}{k}} \sum_{(n,\,k)} I(T_{i_1} + T_{i_2} + \cdots + T_{i_k} \leq t), \tag{2.3}$$

where $\sum_{(n,\ k)}$ denotes the sum over all $\binom{n}{k}$ different index combinations $\{i_1, i_2, \ldots, i_k\}$ of length k. The estimator (2.3) is a U-statistic and therefore it can be established that, for each $k \geq 1$, and for each $t \geq 0$, that

$$\widehat{F}_{C1}^{(k)}(t) \longrightarrow F^{(k)}(t), \text{almost surely}(a.s.).$$

Moreover, it is an unbiased minimum-variance estimator of $F^{(k)}(t)$. Based on the estimator in (2.3), Frees obtained the uniform consistency of $\widehat{M}_{C1}(t)$ $a.s.$ in compact intervals $[0,\ t]$, on the assumption that the number of terms in (2.2), $\kappa = n$ and that T has a positive mean and finite variance. The asymptotic normality of $\widehat{M}_{C1}(t)$ is also proven under some moment conditions.

The drawback of this estimator is the considerable number of computations needed to evaluate it, even though Frees introduced the design parameter $\kappa \leq n$. Schneider et al. [39] propose a new algorithm to compute the Frees estimator in order to reduce the computation time. They define a family of characteristic functions based on the sample that can be determined recursively, and then use Fourier transforms to recover the distributions $\widehat{F}_{C1}^{(k)}(t)$.

An alternative estimator of $M(t)$ is defined in the Concluding Remarks section in Frees [14]. In this case, $F^{(k)}(t)$ is estimated by means of the k-fold convolution of the empirical distribution function obtained from $T_1, T_2, \ldots, T_n$,

$$\widehat{F}_{C2}^{(1)}(t) = \frac{1}{n} \sum_{i=1}^{n} I(T_i \leq t) \tag{2.4}$$

for $k = 1$, which is the empirical distribution function, and

$$\widehat{F}_{C2}^{(k)}(t) = \int \widehat{F}_{C2}^{(k-1)}(t-u) d\widehat{F}_{C2}^{(1)}(u). \tag{2.5}$$

Although $\widehat{F}_{C2}^{(k)}(t)$ is a biased estimate of $F^{(k)}(t)$ (for $k \geq 2$), it is the nonparametric maximum likelihood estimator. It can also be expressed as

$$\widehat{F}_{C2}^{(k)}(t) = \frac{1}{n^k} \sum_{i_1, i_2, \cdots, i_k} I(T_{i_1} + T_{i_2} + \cdots + T_{i_k} \leq t),$$

$\widehat{F}_{C2}^{(k)}(t)$ is a V-statistic and is closely related to $\widehat{F}_{C1}^{(k)}(t)$, in fact, under some conditions for F it is possible to show that $\widehat{M}_{C2}(t)$, the estimator of $M(t)$ based on (2.4) and (2.5), is also consistent and has the same asymptotic distribution as $\widehat{M}_{C1}(t)$.

The computation of this estimator is also tackled by Schneider et al. [39]. They designate $\widehat{M}_{C2}(t)$ *the empirical renewal function*. Although this estimator is not easy to compute, these authors propose solving the following renewal equation

$$\widehat{M}_{C2}(t) = \widehat{F}_{C2}^{(1)}(t) + \int \widehat{M}_{C2}(t-u)d\widehat{F}_{C2}^{(1)}(u). \tag{2.6}$$

They propose an efficient method that consists of solving a discretized version of Eq. (2.6), given by

$$\widehat{M}_{C2}^{d}(r) = \widehat{F}_{C2,d}^{(1)}(r) + \sum_{j=1}^{r} \widehat{M}_{C2}^{d}(r-j)\left(\widehat{F}_{C2,d}^{(1)}(j) - \widehat{F}_{C2,d}^{(1)}(j-1)\right).$$

This method involves approximating the empirical distribution by a lattice distribution. The statistical properties of the estimator in (2.6), i.e. consistency and asymptotic normality, are discussed in Grübel and Pitts [19].

More recently, From and Li [15] construct, among other things, nonparametric confidence intervals for $F^{(k)}(t)$ based on the estimator $\widehat{F}_{C2}^{(k)}(t)$. First of all, they give a numerical procedure for approximating the k-fold convolution of F starting with the empirical distribution function. Next, they obtain the asymptotic distribution of $\sqrt{n}\left[\widehat{F}_{C2}^{(k)}(t) - F^{(k)}(t)\right]$ as a Normal law with mean 0 and derive an estimator of the variance. Finally, they give the expression of an approximate $100(1-\alpha)\%$ confidence interval for $F^{(k)}(t)$. However, as the authors admit, the computational burden is again very high.

2.2.2.2 The Histogram-Type Estimator

Markovich [29] investigates a histogram-type estimator of the renewal function similar to the first Frees estimator. This estimator is based on a new estimator of the k-fold convolution function, where, in contrast to the Frees estimators, only one combination of adjacent renewal times T_i is used.

To describe the estimator, let $[r]$ be the integer part of a real number r. Let $S_k = T_1 + T_2 + \cdots + T_k$, for $k = 1, 2, \ldots$, the waiting times, as defined previously. The estimation of $P\{S_k < t\}$, i.e. the k-fold convolution function of F, is obtained as an empirical distribution function based on an artificially constructed sample of the random variable S_k, from the initial data set of renewal times. For example, to estimate $P\{S_2 < t\} = F^{(2)}(t)$, proceed as follows. Construct the values $\tau_2^i = \sum_{q=2i-1}^{2i} T_q$, for $i = 1, 2, \ldots, n_2 = \left[\frac{n}{2}\right]$.

This procedure produces a sample of size n_2 of the random variable S_2. The associate empirical distribution function is then obtained as

$$\widehat{F}_{HT}^{(2)} = \frac{1}{n_2}\sum_{i=1}^{n_2} I\left(\tau_2^i \leq t\right).$$

In a similar way, continue with $k \geq 3$. Define the sequence

$$\tau_k^i = \sum_{q=1+k(i-1)}^{ki} T_q, \text{for } i = 1, 2, \ldots, n_k = \left[\frac{n}{k}\right].$$

and construct

$$\widehat{F}_{HT}^{(k)} = \frac{1}{n_k}\sum_{i=1}^{n_k} I\left(\tau_k^i \leq t\right). \tag{2.7}$$

Note that $n_k = 1$ for $k > n/2$; thus, the estimator above is defined for $k \leq n/2$.

The estimator of the renewal function based on (2.7) is therefore given by

$$\widehat{M}_{HT}(k, \kappa) = \sum_{k=1}^{\kappa} \widehat{F}_{HT}^{(k)}(t). \tag{2.8}$$

In the notation, the dependence of the number of terms in the summation is highlighted. The convergence properties of the estimator in (2.8) are investigated in Markovich [29]. The method is based on exploring the error term $\sum_{k=\kappa+1}^{\infty} F^{(k)}(t)$. To do this, some information about F, is required, such as the existence of a moment generating function. The number of terms κ in (2.8) can be determined by two alternative methods. One is to obtain κ, as a function of the sample size n, in order to provide the *a.s.* uniform convergence of the estimator to the true renewal function for small t. The results are established for both light- and heavy-tailed renewal time distributions. The other is to use a plot method to determine a desirable value for κ. The histogram-type estimator is plotted versus κ for fixed t. Then, based on the uniform convergence of the $\widehat{M}_{HT}(t, \kappa)$ to $M(t)$, κ is selected according to

$$\kappa^* = \arg\min\left\{\kappa : \widehat{M}_{HT}(t, \kappa) = \widehat{M}_{HT}(t, \kappa + 1); \kappa = 1, 2, \ldots, n - 1\right\}.$$

Compared to Frees estimate, the histogram-type method gives a more computationally tractable estimator. Moreover, Markovich [29] shows that although $\widehat{M}_{HT}(t, \kappa)$ has a greater bias than $\widehat{M}_{C1}(t)$, the mean squared error is smaller. Markovich and Krieger [30] present an alternative data-dependent selection of κ based on a bootstrap method similar to the one we will develop in a subsequent section.

2.2.2.3 Monte-Carlo Estimators

The next group of estimators of the k-fold convolution function of F is based on works by Brown et al. [5] and Gertsbakh and Shpungin [17], who use numerical Monte-Carlo methods to approximate the convolution functions of type $F^{(k)}(t) = P\{T_1 + T_2 + \cdots + T_k \leq t\}$, where T_i are random variables with known CDF, F, for $i = 1, 2, \ldots, k$.

The underlying idea is that the expected value of a random variable may be approximated by generating a large number of samples of the variable and computing the average value toward the samples. Once again, let $T_1, T_2, \ldots, T_n$ be a realization of a renewal process $N(t)$, so that we have a sequence of non-negative independent random variables with CDF F, unknown. Let $F^{(k)}$ be the k-fold convolution function of F. Starting with a random sample from distribution F, our objective is to adapt the k-fold convolutions approximated by the authors cited above. The function F_i's that they use to define their respective procedures are replaced by an empirical distribution based on the sample information. Or, equivalently, the role of each random variable T_i, with known distribution F, is developed by a random variable τ_i with distribution function $\widehat{F}$, for $i = 1, 2, \ldots, k$.

- *The Crude Monte-Carlo estimator*, $\widehat{F}^{(k)}_{CMC}$. The first estimator is easy to implement and is obtained according to the following steps:
  - Simulate N random samples of size k, from the distribution $\widehat{F}$, i.e. the empirical distribution function. Let $t_1^j, t_2^j, \ldots, t_k^j$ be a realization of the jth sample, for $j = 1, 2, \ldots, N$;
  - Define $\varphi^{(j)}(t) = I\left(t_1^j + t_2^j + \ldots + t_k^j \leq t\right)$, for $j = 1, 2, \ldots, N$;
  - Define $\widehat{F}^{(k)}_{CMC}(t) = \frac{1}{N}\sum_{j=1}^{N} \varphi^{(j)}(t)$.
- *The Brown estimator*, $\widehat{F}^{(k)}_B$. Next, the approximation given by Brown et al. [5] is adapted to the present case. It is obtained according to the following.
  - Define the random variable

$$Z_k(t) = \begin{cases} \widehat{F}\left(t - \sum_{i=1}^{k-1} \tau_i\right), & \tau_1 + \tau_2 + \cdots + \tau_{k-1} \leq t \\ 0, & \text{otherwise} \end{cases} \tag{2.9}$$

    where τ_j are considered as independent random variables with distribution function $\widehat{F}$.
  - It is easy to prove that $E[Z_k(t)] = \widehat{F}^{(k)}_{C2}(t)$, the estimator of the k-fold convolution function given in (2.5),where expectation is with respect to $\widehat{F}$.
  - Generate N independent random variables as $Z_k(t)$ defined by (2.9). Denote the jth replication by Z_k^j and approximate the value of the k-fold convolution by

$$\widehat{F}^{(k)}_B(t) = \frac{1}{N}\sum_{j=1}^{N} Z_k^j(t).$$

2.2.2.4 Kernel Estimator and Bandwidth Parameter Estimation

Let us now propose a family of smooth estimators for the successive k-fold convolution functions generated from a distribution function F, based on kernel-type estimators. First, given T_1, T_2, …, T_n, a random sample of *i.i.d.* with *CDF* F, define an estimator of F by means of

$$\widehat{F}(t,h) = \frac{1}{n}\sum_{i=1}^{n} W\left(\frac{t-T_i}{h}\right),$$

where $W(x) = \int_{-\infty}^{x} w(u)\ \mathrm{d}u$, [32] with w a kernel function in the context of nonparametric estimation, usually taken to be a non-negative, symmetric function that integrates to one, and h is a bandwidth parameter that controls the amount of smoothness (also called *smoothing parameter*). For our own particular convenience (see [16]), we will consider

$$\widehat{F}_S(t,h) = \frac{1}{n}\sum_{i=1}^{n} \Phi\left(\frac{t-T_i}{h}\right), \tag{2.10}$$

where $\Phi(u)$ is the Gaussian kernel, that is $\Phi\left(\frac{t-T_i}{h}\right)$ represents, for each $i = 1$, 2, …, n, the distribution function of a Normal law with mean T_i and standard deviation h. In order to estimate the k-fold convolution function of F, we can consider the following estimator

$$\widehat{F}_{S1}^{(k)}(t,h) = \frac{1}{n^k}\sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \Phi\left(\frac{t-T_{i_1}}{h}\right) * \overset{(k)}{\cdots} * \Phi\left(\frac{t-T_{i_k}}{h}\right). \tag{2.11}$$

The convolution of the kernel functions in the expression (2.11) may be seen as the distribution function of the sum of k independent Normal random variables with standard deviation h, and means $T_{i_1}, T_{i_2}, \ldots, T_{i_k}$, respectively. With the properties of the Normal family, this is the distribution function of a Normal variable with mean $T_{i_1} + T_{i_2} + \cdots + T_{i_k}$ and standard deviation $\sqrt{k}h$; therefore, it can be noted that

$$\widehat{F}_{S1}^{(k)}(t,h) = \frac{1}{n^k}\sum_{i_1,i_2,\ldots,i_k=1}^{n} \Phi\left(\frac{t-(T_{i_1}+T_{i_2}+\cdots+T_{i_k})}{\sqrt{k}h}\right). \tag{2.12}$$

On the other hand, for large k, the function in (2.12) is tractable only with difficulty from a computational point of view and so we consider a more feasible expression given by

$$\widehat{F}_{S2}^{(k)}(t,h) = \frac{1}{\binom{n}{k}}\sum_{(n,k)} \Phi\left(\frac{t-(T_{i_1}+T_{i_2}+\cdots+T_{i_k})}{\sqrt{k}h}\right), \tag{2.13}$$

where $\sum_{(n,k)}$ denotes the sum over all $\binom{n}{k}$ distinct subsets of size k, $\{i_1, i_2, \ldots, i_k\}$ of $\{1, 2, \ldots, n\}$.

The kernel smoothing of the convolution functions requires the choice of a bandwidth parameter. The general criterion for choosing a value for the smoothing parameter, h, is to minimize some measure of the error of the kernel estimator. One of the most popular measures of such error is the *Mean Integrated Squared Error* (*MISE*), defined by

$$MISE(k,h) = E\left[\int \left[F^{(k)}(t) - \widehat{F}_{S2}^{(k)}(t,h)\right]^2 \mathrm{d}t\right]. \tag{2.14}$$

So, in principle, expression (2.14) suggests that the choice of h seems to depend on k. However, looking at the different k-fold convolution estimators, the parameter h has been inherited in $\widehat{F}_{S2}^{(k)}(t,h)$ from the smooth estimator in (2.10). So the definitive factor for selecting the bandwidth is to asymptotically minimize the following:

$$MISE(h) = E\left[\int \left[F(t) - \widehat{F}_S(t,h)\right]^2 \mathrm{d}t\right], \tag{2.15}$$

which reduces the problem to selecting the bandwidth for a smooth distribution function estimation.

We follow the guidelines given in Hansen [21]. A manageable expression for the asymptotic *MISE* (i.e. *AMISE*) may be obtained using Gaussian kernels. If $h\sqrt{n} \to \infty$ as $n \to \infty$

$$AMISE = \frac{V}{n} - \frac{h}{n\sqrt{\pi}} + \frac{h^4 R_1}{4} + O\left(h^4\right), \tag{2.16}$$

which is the result obtained by Jones [23] for the particular case of Gaussian kernels.

The first term to appear in (2.16) does not depend on $h, V = \int_0^\infty F(u)(1 - F(u))du$. Further, $R_1 = \int_0^\infty \left(d^2F(u)\right)^2 du$ is a measure of the roughness of F, where d^2 denotes the second derivative operator. This expression can be generalized to

$$R_m = \int_0^\infty \left(d^{m+1}F(u)\right)^2 \mathrm{d}u,$$

d^{m+1} being the operator indicating the derivative of order $m + 1$, for $m \geq 1$. When using Gaussian kernels, the *AMISE* is minimized for the value of h given by

$$h_0 = \left(\frac{1}{n\sqrt{\pi}R_1} \right)^{\frac{1}{3}}. \tag{2.17}$$

Obviously, the h_0 in expression (2.17) is not known since it depends on the value of R_1 which, in turn, depends on the second derivative of F. Therefore, a plug-in method is used to replace R_1 in (2.17) with a consistent estimate [21].

If F is a Normal distribution with standard deviation $\sigma, R_1 = \left(\sigma^3 4\sqrt{\pi}\right)^{-1}$, see Hansen [21], and thus $\widehat{h}_{0,r} = \widehat{\sigma}\left(4n^{-1}\right)^{1/3}$. This particular estimate of h is called the *reference bandwidth* in Hansen [21]. It will be used later.

According to the plug-in rule developed by Hansen [21], (see Eq. (2.7) therein), it is possible to define the estimator of R_m, for $m \geq 1$, obtained by Jones and Sheather [24], as

$$\widehat{R}_m(b) = (-1)^m \frac{1}{n^2} \sum_{i,j=1}^{n} d^{2m} \phi_b\left(T_i - T_j\right), \tag{2.18}$$

where $\phi_b\left(T_i - T_j\right)$ is the *pdf* of a Normal variable with mean T_j and standard deviation b, that is, it is a Gaussian kernel with bandwidth given by b. Jones and Sheather [24] show that the optimal b, the one that minimizes the corresponding *AMISE*, depends on R_{m+1} by means of

$$b_m(R_{m+1}) = \left(\frac{2^{m+\frac{1}{2}}\Gamma\left(m + \frac{1}{2}\right)}{\pi n R_{m+1}} \right)^{\frac{1}{2m+3}}.$$

This equation indicates that the b_1 needed to estimate a value of R_1, required for the estimation of h_0 in (2.17), depends on R_2, which must also be estimated. For estimating R_2, a new bandwidth b_2 will be involved that will depend on R_3, and so on. In other words, it could be expressed as

$$\widehat{R}_1 = \widehat{R}_1(R_2) = \widehat{R}_1\left(\widehat{R}_2(R_3)\right) = \cdots = \widehat{R}_1\left(\widehat{R}_2\left(\widehat{R}_3\left(\ldots \widehat{R}_{m-1}(R_m)\right)\right)\right).$$

This relationship suggests the sequential plug-in rule proposed by Hansen [21], which we detail below,

- Fix $N \geq 1$ and take $\widehat{R}_{N+1} = R_{N+1}\left(\widehat{h}_{o,r}\right)$, by means of (2.18), with $\widehat{h}_{o,r}$ the reference bandwidth;
- Obtain recursively, $\widehat{R}_{m-1} = \widehat{R}_{m-1}\left(\widehat{R}_m\right)$, for $m = 2, \ldots, N$;
- Finally, the estimated bandwidth $\widehat{h}_{o,N}$, will result from substituting $\widehat{R}_1 = \widehat{R}_1\left(\widehat{R}_{N+1}\right)$, obtained in the previous step, in Eq. (2.17).

2.3 Alternating Renewal Process

Let us now consider that the renewal procedure is not an instantaneous event, in such a way that the time to repair or replacement cannot be considered negligible. In other words, we now think of a renewal cycle as a two-phase phenomenon, whose duration is determined by two random variables, say, failure time plus renewal time.

2.3.1 Introduction and Some Applications of the ARP in Reliability

A single unit that evolves in time is considered. Only two states are observed for the system: operative and failed. Let $\varphi(t)$ be, by the value zero versus one, the state of the system at time t; thus,

$$\varphi(t) = \begin{cases} 0, & \text{if the system is operative at time } t \\ 1, & \text{otherwise} \end{cases}$$

Let T be the failure time and R the repair time, respectively. It is assumed that the starting state of the system is operative. Many electrical devices respond to this kind of functioning, for example light bulbs simply function or do not function. T and R are completely unknown in the sense that we do not assume any functional form for their distribution functions. In addition, we suppose once again that perfect repairs are carried out on the system, that is, once the system has failed and a repair has been completed, its behavior is exactly the same as if it were new. Under these conditions, $\{\varphi(t), t \geq 0\}$ is an *Alternating Renewal Process* (*ARP*).

Let F (f) and G (g) be the cumulative distribution (density) function corresponding to the failure time T and repair time R, respectively, both of which are supposed to be absolutely continuous. We do not assume any parametric distribution family for T and R.

A renewal cycle duration is given by $T + R$. Let $H = F * G$, the *CDF* of $T + R$, where * denotes Stieltjes convolution product. The renewal function is now obtained as $M(t) = \sum_{k=1}^{\infty} H^{(k)}(t)$, where $H^{(k)}(t) = \left(H * \overset{(k)}{\cdots} * H\right)(t)$ is the k-fold convolution. In this case, $H^{(k)}(t)$=P{"k renewal cycles are completed in (0, t]"}=$P\{(T_1 + R_1) + (T_2 + R_2) + \cdots + (T_k + R_k) \leq t\}$.

Alternating renewal processes have proved their usefulness as stochastic models in many reliability applications. In fact, they have been widely used as models for diverse phenomena in the engineering field. A typical example is the analysis of a machine which periodically fails, undergoes a technical service, which consists of replacement or perfect repair, and is put to work again. This time, non-negligible repair or replacement times are taken into account. An important application is described by Dickey [10], which includes an example that occurs frequently in nuclear safety systems, where a component is continuously monitored with attention to pressure conditions. When the failure is detected, the component is repaired. This situation may be analyzed by using an alternating renewal process.

Another illustrative example where this type of stochastic process appears particularly suitable for modeling is in air-conditioning loads on electrical power systems, as provided by Mortensen [31].

Di Crescenso [9] gives a generalization of the telegrapher's random process, a stochastic process that describes a motion on the real line characterized by two alternating velocities with opposite directions, where the random times separating consecutive reversals of direction perform an *ARP*. The telegrapher's random process has wide applications in diverse areas such as physics, for describing fluorescence intermittency, for example, or in finance, for describing stock prices.

Chen and Yuan [7] calculate performance measures, such as expected value and variance of the transient throughput and the probability that measures the delivery in time for a balanced serial production with no interstage buffers. The work is based on two fundamental assumptions: that each machine alternates between normal and failed, and that up times and down times are i.i.d.; therefore, an *ARP* is considered.

Bernardara et al. [4], present a new model of rain in time. The alternation of meteorological states (namely, wet and dry) is represented by a strict *ARP* with a Generalized Pareto law of wet and dry periods.

Vanderperre and Makhanov [40] introduce a robot safety device system consisting of a robot with internal safety device. The goal is to obtain the availability measures of the system. The system is characterized by the following safety shut-down rule: "Any repair of the failed safety device requires a shut-down of the operative robot". On the other hand, the safety unit must not operate if the robot is under repair. The system is attended to by two different repair men, and any repair is supposed to be perfect and general. The safety device has a constant failure rate and a general repair time. Both the lifetime and the repair time of the robot are general.

The goal in this section is to obtain a nonparametric estimator for the performance measures of a repairable system modeled by a general *ARP*. In particular, we are interested in estimating the point availability and the long-run availability.

2.3.2 Availability Measures of a Repairable System

Availability is probably the most usual measure for the effectiveness of a repairable system. It was defined by Barlow and Proschan [3] as "the probability that the system is operating at a specified time t", which means that the system has not failed in the interval $(0, t]$ or it has been restored after failure so that it is operational at time t. This measure does not tell us how many times the system has failed before t, the availability of a system just quantifies the chance of finding the system operative when it is required. So, availability measures concern both reliability and maintainability properties of the system and increase with improving either time to failure or maintenance conditions.

Different types of availability measures can be established according to underlying criterions, such as time interval considered and the relevant types of maintenance policies. Next, we present different coefficients of availability for single or one-unit system.

- ***Instantaneous or Point Availability***, $A(t)$
  Instantaneous availability is the probability that the system will be operational at a given time, t, that is

$$A(t) = P[\varphi(t) = 0].$$

  When renewals or repairs are not being carried out in the system, the point availability reduces to the reliability function, $A(t) = P\{T > t\}$.
  In case of repairable systems, availability incorporates maintainability information, and therefore, the operative state of a system at an arbitrary time t is guaranteed if either the system has not failed until t or it has successively failed and been repaired and it is functioning properly since the last repair which occurred at time u, $0 < u < t$. As a consequence, it is easy to see that

$$\begin{aligned}
A(t) &= P\{T_1 > t\} + \sum_{k=1}^{\infty} \int_0^t dP\left\{\sum_{i=1}^{k}(T_i + R_i) \le u\right\} P\{T_{k+1} > t - u\} \\
&= 1 - F(t) + \int_0^t \sum_{k=1}^{\infty} \mathrm{d}H^{(k)}(u)(1 - F(t-u)) \\
&= 1 - F(t) + \int_0^t (1 - F(t-u))dM(u) \\
&= 1 - F(t) + M(t) * (1 - F(t)).
\end{aligned}$$

  We will return to this expression in Sect. 2.3.5.
- ***Average Availability***, $A_{av}(t)$
  This measure gives the proportion of time that the system is available for use. It is calculated as the average value of the point availability function over a period $(0, t]$,

$$A_{av}(t) = \frac{1}{t}\int_0^t A(u)\mathrm{d}u,$$

  which may be interpreted as the average proportion of working time of the system over the first t time units in which the system is operative.

- ***Steady State Availability***, A

  The steady state or limiting availability is the most commonly used availability measure. It gives the long-run performance of a repairable system and is defined as the limit of the instantaneous availability function as time approaches infinity, that is

$$A = \lim_{t \to \infty} A(t).$$

  As a consequence of the key renewal theorem, an important and useful expression for A can be derived. Classical renewal theory (see [37]) establishes that since $1 - F$ is a bounded function, and, as reasoned previously, it is verified that $A(t) = 1 - F(t) + M(t)*(1 - F(t))$, then, point availability is the unique solution of the equation $A(t) = 1 - F(t) + H(t) *A(t)$ that is bounded on finite intervals. So, by the key renewal theorem, it is deduced, under mild conditions over H, that

$$\lim_{t \to \infty} A(t) = \frac{1}{E[T+R]} \int_0^{\infty} (1 - F(t))\mathrm{d}t = \frac{E[T]}{E[T] + E[R]}.$$

  In other words, it is derived the expression so celebrated in reliability literature that states that

$$A = \frac{MTTF}{MTTF + MTTR},$$

  where $MTTF$ ($MTTR$) denotes *mean time to failure (repair)*.

  In practical applications, it is acceptable that point availability approaches its limiting value after a time period. Thus, it can be thought that after a reasonable period of time the system availability is almost invariant with time. However, in many practical cases, the interest is not in a so long period of time in which a steady situation may have been reached. Consider, for instance, that the useful life of any electrical device, from a user viewpoint, could be much shorter than the time the system availability reaches such a steady value.

Other definitions for the availability of a repairable system could be introduced if we distinguish between different types of maintenance strategies, more explicitly, if we consider only corrective downtime (*inherent availability*) or if shutdowns are scheduled for preventive maintenance (*achieved availability*). The most complex case is when all experienced sources of downtime are considered, such as administrative downtime, logistic downtime, preventive and corrective maintenance downtime. The ratio of the system uptime and total time is then defined as the *operational availability*, and it is the more realistic availability measure in the sense that it is the one that the customer actually experiences. For more details, see for example Kumar et al. [27].

2.3.3 Nonparametric Estimation of Steady-State Availability

The problem of estimating the availability measures of a repairable system has been extensively discussed in the recent literature. Many authors have dealt with this topic in various situations, e.g. Ananda [1] constructs confidence intervals and performs hypotheses testing for the long-run availability of a parallel system with multiple components that have exponential failure and repair times.

Phan-Gia and Turkkan [33] consider a gamma alternating renewal system and obtain several results with regard to the availability function. Although they do not carry out an estimation study of point availability, they do obtain interesting results on the variable representing the random proportion of time that the system is on during a renewal period.

Claasen et al. [8] consider a two unit standby system where random variables involving time duration, i.e. lifetime, repair time, and warm up time for the repair facility, are considered as exponential laws. They obtain an estimator of the steady-state availability under such conditions.

Finally, Hwan Cha et al. [22] and Ke & Chu [25] conduct some procedures for obtaining confidence intervals for the steady-state availability of a repairable system.

The long-run performance of a repairable system is assessed in terms of steady-state availability, which was defined in the previous section as

$$A = \lim_{t\to\infty} P\{\varphi(t) = 0\},$$

the probability that the system is functioning at a large time t. In *ARP*, it is well known that

$$A = \frac{MTTF}{MTTF + MTTR},$$

where, as defined previously, $MTTF = E[T]$ and $MTTR = E[R]$.

The aim of this section is to conduct inferences on A based on distribution-free estimators of failure and repair time. Let us consider a system that is activated and functioning at time $t = 0$, and replaced by a new one whenever it fails. We observe such a system in a fixed time interval $[0,\tau]$ and let (T_1, R_1), (T_2, R_2), …, (T_n, R_n) be the registered sample, where $T_1, T_2, \ldots, T_n$ are the observed lifetimes of the system, which are *i.i.d.* with *CDF* F, $E[T_i] = \mu_T > 0$ and $Var(T_i) = \sigma_T^2$. Likewise, $R_1, R_2, \ldots, R_n$ are the observed repair times, which are *i.i.d.* with *CDF* G, $E[R_i] = \mu_R > 0$ and $Var(R_i) = \sigma_R^2$.

The natural estimator of A is given by

$$\widehat{A} = \frac{\overline{T}}{\overline{T} + \overline{R}},$$

where $\overline{T}$ and $\overline{R}$ represent the sample means of the T's and R's, respectively.

Let us now derive the asymptotic distribution of $\widehat{A}$, in order to obtain an asymptotic confidence interval for A. To do so, we consider the following function

$$f(x,y) = \frac{x}{x+y},$$

and the Taylor series to first order around the point (a, b), which is given as

$$f(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b),$$

where the subscripts denote the respective partial derivative. Let us consider the last expression for values $x = \overline{T}$; $y = \overline{R}$; $a = \mu_T$ and $b = \mu_R$. As we know,

$$\sqrt{n}\left(\overline{T} - \mu_T\right) \xrightarrow{d} N\left(0, \sigma_T^2\right)$$

and

$$\sqrt{n}\left(\overline{R} - \mu_R\right) \xrightarrow{d} N\left(0, \sigma_R^2\right),$$

where $\xrightarrow{d}$ denotes convergence in distribution. Thus, we can write

$$\widehat{A} = \frac{\mu_T}{\mu_T + \mu_R} + \frac{\mu_R}{\left(\mu_T + \mu_R\right)^2}\left(\overline{T} - \mu_T\right) + \frac{\mu_T}{\left(\mu_T + \mu_R\right)^2}\left(\overline{R} - \mu_R\right),$$

or equivalently,

$$\sqrt{n}\left(\widehat{A} - A\right) = \sqrt{n}\frac{\mu_R}{\left(\mu_T + \mu_R\right)^2}\left(\overline{T} - \mu_T\right) + \sqrt{n}\frac{\mu_T}{\left(\mu_T + \mu_R\right)^2}\left(\overline{R} - \mu_R\right).$$

Evaluating the above limits in this expression and using the Delta method, we can obtain

$$\sqrt{n}\left(\widehat{A} - A\right) \xrightarrow{d} N\left(0, \sigma_A^2\right),$$

where

$$\sigma_A^2 = \frac{\mu_R^2\sigma_T^2 + \mu_T^2\sigma_R^2}{\left(\mu_T + \mu_R\right)^4}.$$

Let $\widehat{\sigma}_T^2 = \frac{1}{n}\sum_{i=1}^n \left(T_i - \overline{T}\right)^2$ and $\widehat{\sigma}_R^2 = \frac{1}{n}\sum_{i=1}^n \left(R_i - \overline{R}\right)^2$, these being the estimators of σ_T^2 and σ_R^2, respectively. Then we can estimate the variance of the limiting availability by means of

$$\widehat{\sigma}_A^2 = \frac{\overline{R}^2\widehat{\sigma}_T^2 + \overline{T}^2\widehat{\sigma}_R^2}{\left(\overline{T} + \overline{R}\right)^4}.$$

It can be seen that $\widehat{\sigma}_A^2 \xrightarrow{a.s.} \sigma_A^2$, as $n \to \infty$. For a given confidence level $1 - \alpha$, an approximate large sample $100(1 - \alpha)\%$ confidence interval for A can be given by

$$\left(\widehat{A} - z_{\frac{\alpha}{2}} \frac{\widehat{\sigma}_A^2}{\sqrt{n}}, \widehat{A} + z_{\frac{\alpha}{2}} \frac{\widehat{\sigma}_A^2}{\sqrt{n}}\right),$$

with $z_{\frac{\alpha}{2}}$ being the quantile of order $100(1 - \frac{\alpha}{2})\%$ of $N(0,1)$.

2.3.4 Smooth Estimation in the ARP

2.3.4.1 Kernel Estimation

Suppose that the system has been observed up to the nth cycle of the alternating renewal process. During the observation period, we have recorded $0 = S_0 < S_1 < S_2 < \cdots < S_r$, the sequence representing the successive arrival times. For the sake of simplicity in our exposition, we assume that the initial state of the system is operative, and also that the last event recorded is a repair of the system. Under these assumptions, we have no loss of generality, and there exists n such that $r = 2n$.

Let us define the following alternative and independent sequences:

$$T_j = S_{2j-1} - S_{2j-2}, j = 1, 2, \ldots, n,$$

that is, the failure times sequence, and

$$R_j = S_{2j} - S_{2j-1}, j = 1, 2, \ldots, n,$$

the repair times sequence. In other words, we have on the one hand, T_1, T_2, …, T_n the successive lifetimes of the system, which are *i.i.d.* with *CDF F*. On the other hand, R_1, R_2, …, R_n, the corresponding repair times, are *i.i.d.* with *CDF G*. We also assume that $\{T_i, R_i\}$ are independent, and therefore we have an *ARP*.

Since the distribution functions F and G are considered to be absolutely continuous, let f and g denote the corresponding density functions. It is possible to give nonparametric estimators of F and G, respectively, based on kernel estimator functions. That is, define

$$\widehat{F}(t, h_1) = \frac{1}{n} \sum_{i=1}^{n} W_1\left(\frac{t - T_i}{h_1}\right), \tag{2.19}$$

and

$$\widehat{G}(t, h_2) = \frac{1}{n} \sum_{i=1}^{n} W_2\left(\frac{t - R_i}{h_2}\right), \tag{2.20}$$

where $W_j(x) = \int_{-\infty}^{x} w_j(u)du$, with w_j a kernel function in the context of non-parametric estimation [32], and h_j a bandwidth parameter or smoothing parameter that we need to determine, for $j = 1, 2$.

Remark Given that T and R are non-negative random variables, exponential kernel functions could be used (see Guillamón et al. [20]). That is, $w_1(u) = (1/\bar{t})e^{-u/\bar{t}} I_{[0,+\infty)}(u)$ and $w_2(u) = (1/\bar{r})e^{-u/\bar{r}} I_{[0,+\infty)}(u)$, with $\bar{t}$ and $\bar{r}$ being the observed mean values of failure and repair times, respectively, and $I_{[0,+\infty)}(u) = 1$, for $u \in [0, +\infty)$ and 0, otherwise. One important advantage of using this kind of kernels, which are asymmetric, is that the bias that arises when estimating near the origin is considerably reduced. Nevertheless, we use mainly the Gaussian kernel in our simulation studies.

With respect to the smoothing parameter, we suggest the use of two different values, h_1 and h_2, in the definition of the respective estimators for F and G, since in general, failure and repair times are expected to have different ranges.

2.3.4.2 A Bootstrap Method for Choosing the Bandwidth

One of the most important aspects of kernel estimation is the choice of smoothing parameter. Many different proposals have been made to address this dilemma (see Sect. 2.2.2.4). There exists a vast literature on the use of bootstrap methods for selecting the bandwidth.

Bootstrap resampling techniques were introduced by Efron [11]. One of the earliest references on the subject is the work by Cao [6], who introduced a smooth bootstrap for choosing the bandwidth in kernel density estimation. The method, which exhibited a reasonably reliable behavior, has been subsequently extended to confront the problem of bandwidth selection in other contexts, hazard rate estimation, for instance, and under different sampling schemes, in particular, in the presence of censoring in González-Manteiga et al. [18].

These techniques have already been developed in many reliability applications, producing very good results, see for example, Phillips [34, 35], Marcorin and Abackerli [28] and Gámiz and Román [16]. In all cases, bootstrap techniques reveal significant improvements in estimation compared to traditional techniques. In this section, we present a method based on bootstrap resampling to select the smoothing parameters involved in the kernel estimators of the distribution functions in an *ARP*.

The bandwidth parameter, $\mathbf{h} = (h_1, h_2)'$, can be selected as the minimizer of the mean integrated squared error. In this context of the *ARP*, we define the following *MISE*

$$MISE(\mathbf{h}) = MISE(h_1, F) + MISE(h_2, G), \tag{2.21}$$

where the *MISE*s on the right-hand side are the usual mean integrated squared errors associated with both kernel distribution estimators, this term having been defined previously in Eq. (2.15) as

$$MISE(h_1, F) = E\left[\int \left[\widehat{F}(t, h_1) - F(t)\right]^2 \mathrm{d}t\right].$$

This expression is obviously unknown, since we do not have the distribution for which the expectation of (2.21) is calculated. Therefore, we describe below a procedure to approximate the *MISE* in (2.21). This is based on a bootstrap method that consists of imitating the random procedure from which the original sample is drawn, and so we replace the role of the true distribution functions F and G by estimators of the type given by (2.19) and (2.20). In other words, we use the smoothed bootstrap method, where the bootstrap sample is obtained from the estimated values of the distributions F and G. Next, we describe an algorithm for realizing *ARP* trajectories that imitate the original sample. This algorithm is based on the embedded Markov chain. It can be explained as follows:

Let $\mathbf{h}$ be bandwidth parameters.

Algorithm *Smoothed Bootstrap for ARP*

Step 1. Put $m = 0$, $s_0 = 0$;
Step 2. Generate random variable $T^\bullet \sim \widehat{F}(\cdot; h_1)$ and set $t = T^\bullet(\omega)$;
Step 3. Put $m = m + 1$ and $s_m = s_{m\text{-}1} + t$. If $m \geq 2n$ then end;
Step 4. Generate random variable $R^\bullet \sim \widehat{G}(\cdot; h_2)$ and set $r = R^\bullet(\omega)$;
Step 5. Put $m = m + 1$ and $s_m = s_{m\text{-}1} + r$. If $m \geq 2n$ then end, otherwise continue to Step 2.

Once the bootstrap sample is drawn, consider the bootstrap version of the estimators in (2.19) and (2.20), $F^\bullet(t, h_1)$ and $G^\bullet(t, h_2)$, for which we have replaced the original sample by the bootstrap sample in the expressions (2.19) and (2.20). Now, define the bootstrap estimate of the mean integrated squared error by

$$MISE^\bullet(\mathbf{h}) = MISE^\bullet(h_1, F) + MISE^\bullet(h_2, G),$$

that is,

$$MISE^\bullet(\mathbf{h}) = E_\bullet\left[\int \left[\widehat{F}^\bullet(t, h_1) - \widehat{F}(t, h_1)\right]^2 \mathrm{d}t\right] + E_\bullet\left[\int \left[\widehat{G}^\bullet(t, h_2) - \widehat{G}(t, h_2)\right]^2 \mathrm{d}t\right]$$

The minimizer of the above function is the bootstrap bandwidth selector. Although $MISE^\bullet(\mathbf{h})$ can be written in terms of the original sample and, therefore, from a theoretical viewpoint no resampling is needed, an explicit expression is quite hard to obtain (see Cao [6]). Therefore, in practice, Monte-Carlo methods are proposed to calculate the values of $MISE^\bullet(\mathbf{h})$.

The resampling procedure consists in drawing B bootstrap samples in the following way: for each bootstrap sample, b, a replication of $\widehat{F}^\bullet(t, h_1)$ and $\widehat{G}^\bullet(t, h_2)$

are obtained, i.e. $\widehat{F}_b^{\bullet}(t, h_1)$ and $\widehat{G}_b^{\bullet}(t, h_2)$, so that the bootstrap estimation of the standard error may be obtained as the sample mean of the bootstrap samples

$$\begin{aligned}&\widehat{MISE}^{\bullet}(\mathbf{h})\\&=\frac{1}{B}\sum_{b=1}^{B}\left[\int\left(\left[\widehat{F}_b^{\bullet}(t,h_1)-\widehat{F}(t,h_1)\right]^2+\left[\widehat{G}_b^{\bullet}(t,h_2)-\widehat{G}(t,h_2)\right]^2\right)\mathrm{d}t\right].\end{aligned}\quad (2.22)$$

The integral in (2.22) is approximated by numerical methods if necessary. Finally, we choose the vector of bandwidths, $\mathbf{h}_{boot}$ that minimizes the expression (2.22).

In order to perform the resampling procedure, it is necessary to start with a pilot bandwidth $\mathbf{h}^0 = \left(h_1^0, h_2^0\right)'$ as the initial value and the bootstrap bandwidth parameters are achieved by means of the following iterative method [20]:

$$\begin{aligned}&\widehat{MISE}^{\bullet}(\mathbf{h}^j)\\&=\frac{1}{B}\sum_{b=1}^{B}\left[\int\left(\left[\widehat{F}_b^{\bullet}(t,h_1^j)-\widehat{F}\left(t,h_1^{j-1}\right)\right]^2+\left[\widehat{G}_b^{\bullet}(t,h_2^j)-\widehat{G}\left(t,h_2^{j-1}\right)\right]^2\right)\mathrm{d}t\right],\end{aligned}$$

where $\widehat{F}_b^{\bullet}\left(t, h_1^j\right)$ and $\widehat{G}_b^{\bullet}\left(t, h_2^j\right)$ are the estimators of the *ARP* distributions based on the bootstrap sample with $\mathbf{h}^j$ as bandwidth vector value; and, $\widehat{F}\left(t, h_1^{j-1}\right)$ and $\widehat{G}\left(t, h_2^{j-1}\right)$ are the estimators based on the original sample and with parameter $\mathbf{h}^{j\text{-}1}$, the one that minimizes the $\widehat{MISE}^{\bullet}$ at the previous iteration.

Thus, starting with a pilot bandwidth $\mathbf{h}^0$, the idea is to find the value of $\mathbf{h}^1$ that minimizes $\widehat{MISE}^{\bullet}(\mathbf{h}^j)$, for $j = 1$; then make $j = j + 1$ and repeat the procedure until an appropriate convergence criterion is achieved, that could be, for example that two consecutive $\mathbf{h}^{j-1}$ and $\mathbf{h}^j$ are close enough.

To illustrate this, we apply the method to an alternating renewal process where the failure (repair) times have been simulated from a particular parametric distribution family.

Example Weibull Lifetime and Lognormal Repair Time

First, we consider a system that evolves in time, passing through successive up and down states. The lengths of the up periods are considered to be random variables with Weibull distribution with scale parameter 20 and shape parameter 2. On the other hand, it is known that the Lognormal distribution is a suitable model for the repair times in many cases. We consider a two-parameter Lognormal distribution for the repair time, where the mean log time is chosen as 2 and the standard deviation of the log time is 1.5. Under these conditions, we have simulated 100 renewal cycles, each consisting of an *up* period plus a *down* period. We have applied the smoothed method obtaining the curves given in Fig. 2.1. The bootstrap approximation for the

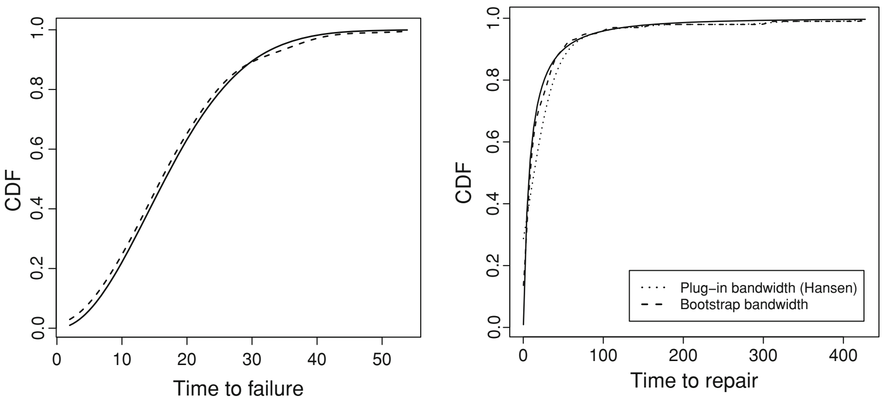

Fig. 2.1 Nonparametric estimation an *ARP*: Theoretical curve (*solid line*) and bootstrap estimate (*dashed line*)

vector of bandwidths is given by $\mathbf{h}_{boot} = (3.159268, 4.543223)'$. These values have been obtained after performing the iterative procedure of the previous section. To do this, we needed a pilot vector of bandwidths to initiate the procedure. We have considered as initial values of the bandwidth parameters the plug-in values suggested in Hansen [21] and given by $\widehat{h}_{0,r} = \widehat{\sigma}(4n^{-1})^{1/3}$, where n is the sample size and $\widehat{\sigma}$ is the sampling standard deviation. In our example, we have obtained as initial vector of bandwidths $\mathbf{h}_{ini} = (3.314656,\ 18.96642)'$. We stopped when the difference between the values of the $\mathbf{h}'s$ estimated in two consecutives iterations is below 10^{-2}. This convergence was attained after 8 iterations. The results are presented in Fig. 2.1. The solid line in each graph represents the theoretical cumulative distribution function of the Weibull and Lognormal distribution. We have not included the estimator based on the initial value in the left panel, since the difference with the curve obtained with the bootstrap approximation of the bandwidth is not appreciated in the graph. The estimators show a lower precision near the origin. This problem can be solved by means of local estimation procedures or the use of suitable asymmetric kernels; however, this issue will not be addressed here.

2.3.5 Smooth Estimation of the Availability Function

Let $H = F * G$, the *CDF* of a renewal cycle duration, where $*$ denotes Stieltjes convolution product, and $M(t) = \sum_{k=1}^{\infty} H^{(k)}(t)$, where $H^{(k)}(t) = \big(H * \overset{(k)}{\cdots} * H\big)(t)$ is the k-fold convolution.

As shown in classical renewal theory (see Sect. 2.3.2), the availability function satisfies the following renewal type equation,

$$A(t) = 1 - F(t) + \int_0^t (1 - F(t-u))dM(u) = 1 - F(t) + M(t) * (1 - F(t)) \tag{2.23}$$

which is generally not easy to evaluate.

The estimation of the availability is now based on kernel estimators of the functions F and G defined in expressions (2.19) and (2.20), respectively, hence

$$\widehat{A}(t, \mathbf{h}) = 1 - \widehat{F}(t, h_1) + \widehat{M}(t, \mathbf{h}) * \left(1 - \widehat{F}(t, h_1)\right) \tag{2.24}$$

where $\widehat{M}(t, \mathbf{h}) = \sum_{k=1}^{\infty} \widehat{H}^{(k)}(t, \mathbf{h})$. Expression (2.24) depends on $\mathbf{h} = (h_1, h_2)'$, the bootstrap bandwidth parameters for estimating F and G, respectively, obtained previously.

The aim is to give the value of the above expression for the availability. For this purpose, note that the most tedious problem to arise in expression (2.23) is that of deriving the renewal function $M(t)$, and therefore we proceed as follows. First, we approximate the value of $M(t)$ by plug-in into the expression of M (given above), the functions F and G by their respective exponential kernel estimator as explained in the previous section, that is (2.19) and (2.20), respectively.

So, for estimating the k-fold convolution function $H^{(k)}(t)$ we have the following expression

$$\begin{aligned}\widehat{H}_n^{(k)}(t, \mathbf{h}) = \frac{1}{n^{2k}} \sum_{i_1, \ldots, i_k; j_1, \ldots, j_k}^{n} & W_1\left(\frac{t - T_{i_k}}{h_1}\right) * \cdots * W_1\left(\frac{t - T_{i_k}}{h_1}\right) \\ & * W_2\left(\frac{t - R_{j_1}}{h_2}\right) * \cdots * W_2\left(\frac{t - R_{j_k}}{h_2}\right)\end{aligned} \tag{2.25}$$

for $k = 1, 2, \ldots$; where $\{T_{i_m}\}$ and $\{R_{j_m}\}$ are the observed failure and repair times; $\mathbf{h}_{boot} = (h_1, h_2)'$ is the bootstrap bandwidth vector and, W_1 and W_2 represent the *CDF* of an exponential random variable with scale parameters $\bar{t}$ and $\bar{r}$, respectively, i.e. the observed sample means. For each $i, j = 1, 2, \ldots, n$, the convolution into the sum can be viewed as the *CDF* of the sum of two independent random variables, one with exponential distribution with scale parameter $h_1\bar{t}$ and location T_i, and the other with scale $h_2\bar{r}$ and location R_j. So, these convolutions can be obtained analytically (we have used the package *distr* [38] of R for that purpose).

Given that we have chosen exponential kernel functions, the convolutions that appear in expression (2.25) can be simplified as follows

$$\begin{aligned}\widehat{H}_n^{(k)}(t, \mathbf{h}) = \frac{1}{n^{2k}} \sum_{i_1, \ldots, i_k; j_1, \ldots, j_k}^{n} & \mathbf{W}_1\left(\frac{t - (T_{i_1} + \cdots + T_{i_k})}{h_1}\right) \\ & * \mathbf{W}_2\left(\frac{t - \left(R_{j_1} + \cdots + R_{j_k}\right)}{h_2}\right),\end{aligned}$$

Table 2.1 Kernel estimation of the renewal function

t_0	1	3	5	6	7	10	15
$\widehat{M}(t_0)$	1.0042 (1.2869)	1.5492 (2.2393)	2.4998 (3.0631)	2.9183 (3.4552)	3.3512 (3.8391)	4.57678 (4.9550)	6.8928 (6.6132)

where $\mathbf{W}_1$ and $\mathbf{W}_2$ are respectively, the distribution functions of the corresponding Gamma family.

To conclude, the number of terms, k_0, considered in $M(t)$ can be determined by the normal approximation of the number of renewals in $(0, t]$, that is, if $Y_i = T_i + R_i$ is the length of a renewal cycle $k_0 = \min\{k : P[Y_1 + Y_2 + \cdots + Y_k \leq t] \leq \varepsilon\}$, with ε fixed small enough.

Example Kernel Estimation of the Renewal Function

To study the empirical performance, we carried out the following simulation study. Let $Y_1, Y_2, \ldots, Y_n$ be the length of $n = 10$ renewal cycles, each obtained by means of a life period simulated from a *CDF* Weibull with scale $\beta = 1$ and shape $\alpha = 2$, plus a down period simulated from a *CDF* Weibull with scale $\beta = 1$ and shape $\alpha = 0.5$. Next, we applied the bootstrap procedure as indicated above to estimate the distributions associated with the *ARP*.

A computational procedure that gives the value of the renewal function $M(t_0)$ for a fixed t_0 may be implemented by means of some functions working in R. First, the method approximates the number of significant terms in $M(t_0)$, i.e. k_0 (defined above); and then it uses the convolution function of gamma distributions, which is implemented in the package *distr* of R, to obtain the probability distribution functions that appear in the successive $\widehat{H}^{(k)}$. Some of the results are displayed in Table 2.1.

The bootstrap approximation of the bandwidth involved in these calculations is given by $\mathbf{h} = (0.0505, 0.1304)'$. The numbers in parentheses express the values of $M(t_0)$ provided by using the functions that approximate the convolution of absolutely continuous distributions, performed in the package *distr* of R. These values are used here for reference.

In any case, from a computational point of view, expression (2.24) is extremely awkward to evaluate, so in Sect. 2.3.7, we propose a slightly more feasible procedure.

2.3.6 Consistency of the Bootstrap Approximation of the Availability Function

The Laplace transform may be used to prove the consistency of the estimate defined in (2.24), which is established in terms of the *unavailability function*, that is, $U(t) = 1 - A(t)$.

Taking the Laplace transform on both sides of Eq. (2.23), with the properties of the Laplace transform, we obtain

$$\widehat{U}^*(s,\mathbf{h}) = \frac{1}{s} - \widehat{A}^*(s,\mathbf{h}) = \frac{1}{s} - \frac{1-\widehat{f}^*(s,h_1)}{s\left[1-\widehat{f}^*(s,h_1)\widehat{g}^*(s,h_2)\right]}, \tag{2.26}$$

where $\widehat{f}^*(s,h_1)$ and $\widehat{g}^*(s,h_2)$ represent respectively the Laplace transform of the kernel densities, which are

$$\widehat{f}^*(s,h_1) = w^*(sh_1)\frac{1}{n}\sum_{i=1}^{n} e^{-sT_i}$$

and

$$\widehat{g}^*(s,h_2) = w^*(sh_2)\frac{1}{n}\sum_{i=1}^{n} e^{-sR_i},$$

given that we have used the same kernel function, W, in the estimation of F and G.

In the above expressions, $w^*\left(sh_j\right)$ is the Laplace transform of the function $w = dW$, say the derivative of W, evaluated in sh_j, $j = 1, 2$. That is, $w^*(s) = \int_0^\infty e^{-st}w(t)dt$. We assume that the kernel function w is such that its Laplace transform exists for $s > 0$ (which is valid in the case of Gaussian or exponential kernels, for example).

Considering expression (2.26), we find that

$$\widehat{U}^*(s,\mathbf{h}) = \frac{1}{s} - \frac{1-\widehat{w}^*(sh_1)\frac{1}{n}\sum_{i=1}^{n} e^{-sT_i}}{s\left[1-\left(\widehat{w}^*(sh_1)\frac{1}{n}\sum_{i=1}^{n} e^{-sT_i}\right)\left(\widehat{w}^*(sh_2)\frac{1}{n}\sum_{i=1}^{n} e^{-sR_i}\right)\right]}. \tag{2.27}$$

If the Laplace transform f^* does exist, we find that $f^*(s) = E[e^{-sT}]$. Since $\{T_i\}$ are *i.i.d.* with density function f, by the strong law of large numbers, for any s for which the above expectation exists,

$$\frac{1}{n}\sum_{i=1}^{n} e^{-sT_i} \xrightarrow{a.s.} f^*(s) \ \text{ as } n \to \infty,$$

and the same argument is valid for establishing that

$$\frac{1}{n}\sum_{i=1}^{n} e^{-sR_i} \xrightarrow{a.s.} g^*(s) \ \text{ as } n \to \infty.$$

Moreover, for fixed s, since $h_j \to 0$ as $n \to \infty$, with the properties of the Laplace transform, we obtain $w^*\left(sh_j\right) \to 1$, for $j = 1, 2$. In conclusion, from (2.27), we obtain

$$\widehat{U}^*(s, \mathbf{h}) \longrightarrow \frac{1}{s} - \frac{1 - f^*(s)}{s[1 - f^*(s)g^*(s)]}.$$

The right-hand side of the last expression corresponds to the Laplace transform of the unavailability function $U(t)$.

The unavailability function may be considered as a defective measure, so that we can apply the *extended continuity theorem of the Laplace transform for measures* (see [12]), according to which, if H_n is a measure with the Laplace transform φ_n, for $n = 1, 2, \ldots$, and $\varphi_n(s) \rightarrow \varphi(s)$ for $s > 0$, then φ is the Laplace transform of a measure H and $H_n \rightarrow H$, and this convergence is for each bounded interval of continuity of H.

In conclusion, we deduce the uniformly strong consistency of the estimator of availability given by (2.23), that is, for all $t \in R_+$,

$$\sup_{t \in [0,\tau]} \left| \widehat{A}(t, \mathbf{h}) - A(t) \right| \xrightarrow{a.s.} 0, \text{as } n \to \infty.$$

The Laplace transforms described above, all correspond to defective distributions since, as can easily be checked, for any $n = 1, 2, \ldots, \lim_{s \to 0} \widehat{U}^*(s, \mathbf{h}) \neq 1$. This result may be established by means of the following limit

$$\lim_{s \to 0} s\widehat{A}^*(s, \mathbf{h}) = \frac{h_1 + \bar{t}_n}{(h_1 + \bar{t}_n) + (h_2 + \bar{r}_n)}, \tag{2.28}$$

which is independent of the kernel function w. Here, $\bar{t}_n$ and $\bar{r}_n$ represent the mean sample values.

By the properties of the Laplace transform, we find that $\lim_{s \to 0} s\widehat{A}^*(s, \mathbf{h}) = \lim_{t \to \infty} \widehat{A}(t, \mathbf{h})$. Taking this property together with (2.28), we obtain via our kernel estimator of the availability, an expression for asymptotic availability which is congruent with the known result that establishes that

$$\lim_{t \to \infty} A(t) = \frac{MTTF}{MTTF + MTTR},$$

where *MTTF* denotes the mean time to failure and *MTTR*, the mean time to repair.

Example: Exploring Exponential Kernel Functions

The estimation procedure above may also be carried out by considering the sample information given by the observed times between failures, i.e. $Y_i = T_i + R_i$, for $i = 1, 2, \ldots, n$. Let $H(t)$ denote the theoretical cumulative distribution function; in this case, kernel estimation of the availability function would have the following Laplace transform

$$\widehat{A}^*(s, \mathbf{h}) = \frac{1 - w^*(s, h_1)\frac{1}{n}\sum_{i=1}^{n} e^{-sT_i}}{s\left[1 - w^*(s, h_2)\frac{1}{n}\sum_{i=1}^{n} e^{-sY_i}\right]}.$$

If we consider an exponential kernel function, the last expression is of the form

$$\widehat{A}^*(s,\mathbf{h}) = \frac{1-\left(\frac{1}{sh_1+1}\right)\frac{1}{n}\sum_{i=1}^{n} e^{-sT_i}}{s\left[1-\left(\frac{1}{sh_2+1}\right)\frac{1}{n}\sum_{i=1}^{n} e^{-sY_i}\right]}.$$

We can approximate the above exponential functions by their Taylor expansions, obtaining, for s near 0, that

$$\widehat{A}^*(s,\mathbf{h}) = \frac{1-\left(\frac{1-s\bar{t}_n}{sh_1+1}\right)}{s\left[1-\left(\frac{1-s\bar{y}_n}{sh_2+1}\right)\right]},$$

where $\bar{t}_n$ is the mean failure time and $\bar{y}_n$ is the mean duration of a renewal cycle, that is $\bar{y}_n = \bar{t}_n + \bar{r}_n$. Easy computations lead to

$$\widehat{A}^*(s,\mathbf{h}) = \frac{h_1+\bar{t}_n}{h_2+\bar{y}_n}\left(\frac{sh_2+1}{s(sh_1+1)}\right),$$

and, by inverting this expression, we obtain an estimator of the availability by the following

$$\widehat{A}(t,\mathbf{h}) = \frac{h_1+\bar{t}_n}{h_2+\bar{y}_n}\left[1+\left(\frac{h_2}{h_1}-1\right)e^{-\frac{t}{h_1}}\right],$$

which is valid for large values of t, given the equivalence between the values of the Laplace transform of a function, in this case A^*, near the origin, and the asymptotic values of the function, say A.

2.3.7 *Bootstrap Estimate of the k-Fold Convolution of a Distribution Function*

In Sect. 2.2.2.4, a kernel estimator for the k-fold convolution function was defined. There, a plug-in bandwidth selector was suggested, based on the asymptotic form of the mean integrated squared error (*MISE*).

In this section, we use a procedure based on bootstrap techniques similar to those in Sect. 2.3.4, in order to find the value of h that minimizes the *MISE*. As in Sect. 2.2.2.4, the particular form of $\widehat{F}_{S2}^{(k)}(t,h)$ for all $k \geq 2$ reduces the problem of finding a bandwidth for any k to the first step, that is for $k = 1$. So, the optimization problem is stated as finding the value of h that minimizes

$$MISE(h) = E\left[\int\left[F(t)-\widehat{F}_{S2}(t,h)\right]^2 \mathrm{d}t\right].$$

Using a similar rationale to that given in Sect. 2.3.4.2, the bootstrap bandwidth parameter is achieved by means of the following iterative method

$$\widehat{MISE}(h^j) = \frac{1}{B}\sum_{b=1}^{B}\left\{\int\left[\widehat{F}_b^{\bullet}\left(t, h^j\right) - \widehat{F_{S2}}\left(t, h^{j-1}\right)\right]^2 \mathrm{d}t\right\}.$$

We carry out a smoothed bootstrap to obtain bootstrap samples in the observation interval, which is determined, as above, by the occurrence of the nth renewal, the size of the original sample.

Algorithm *Smoothed Bootstrap*

Step 1. Put $m = 0$ and $s_0 = 0$;
Step 2. Generate random variable $T^{\bullet} \sim \widehat{F}(\cdot; h)$ and set $t = T^{\bullet}(\omega)$;
Step 3. Put $m = m + 1$ and $s_m = s_{m\text{-}1} + t$. If $m \geq n$ then end, otherwise continue to Step 2.

Example: Kernel Estimation of the k-fold Convolution of a Distribution Function

To illustrate, we now consider the following simulation study. Let $T_1, T_2, \ldots, T_n$ be the lengths of $n = 10$ simulated renewal cycles, i.i.d. with *CDF* Weibull with scale parameter $\beta = 1$ and shape parameter $\alpha = 3$. We construct the estimator $\widehat{F}^k_{S2}(t, h)$, as defined in Eq. (2.13), for $k = 2, 3, 4, 5, 6, 7$, and obtain the corresponding values for t in [0, 10]. The approximation of bandwidth is obtained by bootstrap techniques in the first step, that is for $k = 1$, which gives the value $h_{boot} = 0.2786$. The results are displayed in the figures below. We compared our results to those obtained with the estimator proposed by Frees [14], that is, M_{C1} which is defined in Sect. 2.2.2.1 based on Eq. (2.3), and found that the kernel estimator gives greater accuracy. We made use once again of the *distr* package provided by the *R* programming system in order to approximate, for $k = 2, 3, 4, 5, 6, 7$, the convolution functions, unfeasible in theory, for the Weibull distribution F with scale parameter $\beta = 1$ and shape parameter $\alpha = 3$. We use the functions performed under R as reference values to contrast the accuracy of the

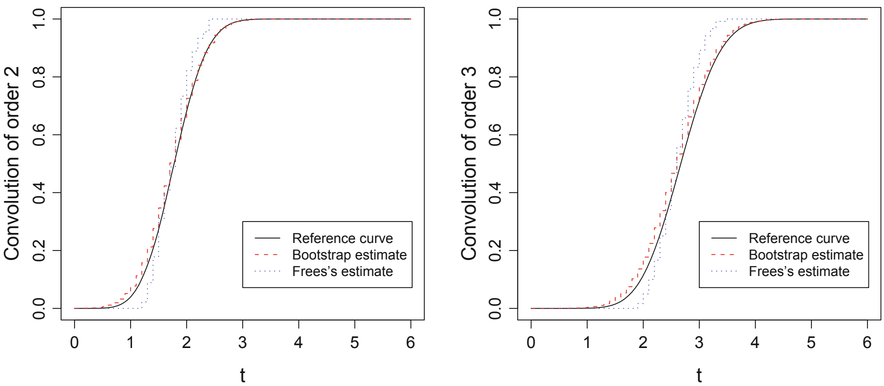

Fig. 2.2 Nonparametric estimation of the k-fold convolution, for $k = 2, 3$

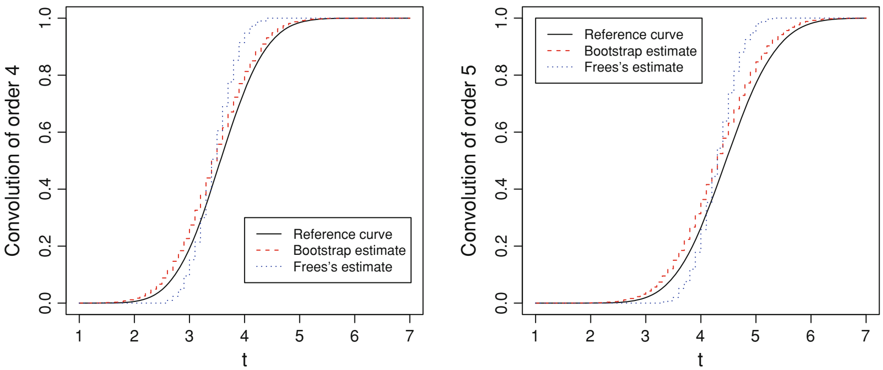

Fig. 2.3 Nonparametric estimation of the k-fold convolution, for $k = 4, 5$

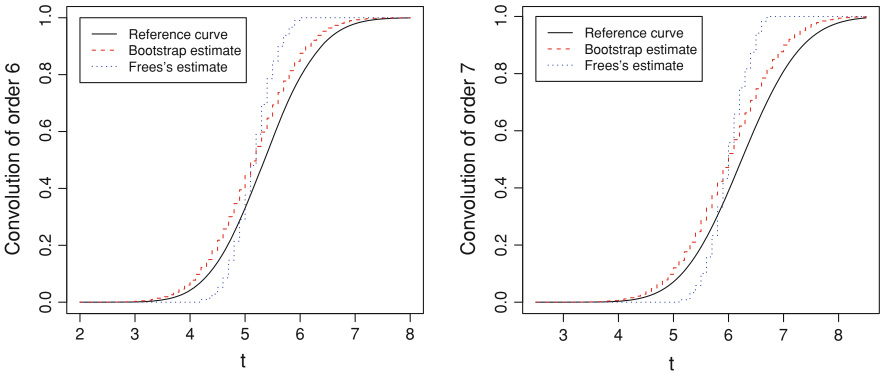

Fig. 2.4 Nonparametric estimation of the k-fold convolution, for $k = 6, 7$

two estimators. The bootstrap curve captures the "theoretical" behavior better than Frees's curve, as can be appreciated from Figs. 2.2, 2.3, 2.4.

Acknowledgments Section 2.3 of this chapter is an extension into book-length form of the article *Nonparametric estimation of the availability in a general repairable system*, originally published in *Reliability Engineering and System Safety* **93** (8), 1188–11962 (2008).The authors express their full acknowledgement of the original publication of the paper in the journal cited above, edited by Elsevier.

References

1. Ananda MMA (1999) Estimation and testing of availability of a parallel system with exponential failure and repair times. J Stat Plan Inf 77:237–246
2. Ascher H, Feingold H (1984) Repairable systems reliability: modelling, inference, misconceptions and their causes. Marcel Dekker Inc, New York

3. Barlow RE, Proschan F (1975) Statistical theory of reliability and life testing. Probability Models. Holt, Rinehart and Winston, New York
4. Bernardara P, De Michele C, Rosso R (2007) A simple model of rain in time: an alternating renewal process of wet and dry states with a fractional (non-Gaussian) rain intensity. Atmos Res 84:291–301
5. Brown M, Solomon H, Stevens MA (1981) Monte-Carlo simulation of the renewal function. J Appl Probab 13:426–434
6. Cao R (1993) Bootstrapping the mean integrated squared error. J Multivar Anal 45:137–160
7. Chen C-T, Yuan J (2003) Throughput analysis using alternating renewal process in balanced serial production lines with no interstage buffers. J Chin Insst Ind Eng 20(5):522–532
8. Claasen SJ, Joubert JW, Yadavalli VSS (2004) Interval estimation of the availability of a two unit standby system with non instantaneous switch over and 'dead time'. Pak J Stat 20(1):115–122
9. Di Crescenso A (2001) On random motions with velocities alternating at Erlang-distributed random times. Adv Appl Probab 33(3):690–701
10. Dickey JM (1991) The renewal function for an alternating renewal process, which has a Weibull failure distribution and a constant repair time. Reliab Eng Syst Saf 31:321–343
11. Efron B (1979) Bootstrap methods: another look at the jackknife. Ann Stat 7:1–26
12. Feller W (1971) An introduction to probability theory and its applications, vol II, 2nd edn. Willey, New York
13. Frees EW (1986a) Warranty analysis and renewal function estimation. Naval Res Logist Quart 33:361–372
14. Frees EW (1986b) Nonparametric renewal function estimation. Ann Stat 14(4):1366–1378
15. From SG, Li L (2003) Nonparametric estimation of some quantities of interest from renewal theory. Naval Res Logist 50(6):638–649
16. Gámiz ML, Román Y (2008) Non-parametric estimation of the availability in a general repairable system. Reliab Eng Syst Saf 93:1188–1196
17. Gerstbakh I, Shpungin Y (2004) Renewal function and interval availability: a numerical Monte Carlo study. Commun Stat Theory Methods 33(3):639–649
18. González-Manteiga W, Cao R, Marron JS (1996) Bootstrap selection of the smoothing parameter in nonparametric hazard rate estimation. J Am Stat Assoc 91(435):1130–1140
19. Grübel R, Pitts SM (1993) Nonparametric estimation in renewal theory I: the empirical renewal function. Ann Stat 21(3):1431–1451
20. Guillamón A, Navarro J, Ruiz JM (1999) A note on kernel estimators for positive valued random variables. Sankhya: Indian J Stat 61 Ser A (Part 2):276–281
21. Hansen BE (2004) Bandwidth selection for nonparametric distribution estimation. Discussion paper, University of Wisconsin, Madison
22. Hwan Cha J, Sangyeol L, Jongwoo J (2006) Sequential confidence interval estimation for system availability. Qual Reliab Eng Int 22:165–176
23. Jones MC (1990) The performance of kernel density functions in kernel distribution function estimation. Stat Probab Lett 9:129–132
24. Jones MC, Sheather SJ (1991) Using non-stochastic terms to advantage in kernel-based estimation of integrated squared density derivatives. Stat Probab Lett 11:511–514
25. Ke JC, Chu YK (2007) Nonparametric analysis on system availability: confidence bound and power function. J Math Stat 3(4):181–187
26. Kijima M (1989) Some results for repairable systems with general repair. J Appl Probab 26(1):89–102
27. Kumar UD, Crocker J, Chitra T, Saranga H (2006) Reliability and six sigma. Springer, New York
28. Marcorin AJ, Abackerli AJ (2006) Field failure data: an alternative proposal for reliability estimation. Qual Reliab Eng Int 22:851–862
29. Markovich NM (2004) Nonparametric renewal function estimation and smoothing by empirical data. Preprint ETH, Zuerich

30. Markovich NM, Krieger UR (2006) Nonparametric estimation of the renewal function by empirical data. Stoch Models 22:175–199
31. Mortensen RE (1990) Alternating renewal process models for electric power system loads. IEEE Trans Autom Control 35(11):1245–1249
32. Mugdadi AR, Ghebregiorgis GS (2005) The Kernel distribution estimator of functions of random variables. Nonparametr Stat 17(7):807–818
33. Pham Gia T, Turkkan N (1999) System availability in a gamma alternating renewal process. Naval Res Logist 46:822–844
34. Phillips MJ (2000) Bootstrap confidence regions for the expected ROCOF of a repairable system. IEEE Trans Reliab 49(2):204–208
35. Phillips MJ (2001) Estimation of the expected ROCOF of a repairable system with bootstrap confidence region. Qual Reliab Eng Int 17:159–162
36. Rausand M, Hoyland A (2004) System reliability theory: models, statistical methods and applications, 2nd edn. Wiley, New York
37. Ross SM (1992) Applied probability models with optimization applications. Courier Dover Publications, New York
38. Ruckdeschel P, Kohl M, Stabla T, Camphausen F (2006) S4 classes for distributions. R News 6(2):2–6. http://www.uni-bayreuth.de/departments/math/org/mathe7/DISTR/distr.pdf
39. Schneider H, Lin B-S, O'Cinneide C (1990) Comparison of nonparametric estimation for the renewal function. Appl Stat 39(1):55–61
40. Vanderperre EJ, Makhanov SS (2008) Point availability of a robot with internal safety device. Contemp Eng Sci 1(1):15–25

http://www.springer.com/978-0-85729-117-2

Applied Nonparametric Statistics in Reliability
Gámiz, M.L.; Kulasekera, K.B.; Limnios, N.; Lindqvist, B.H.
2011, XIII, 230 p., Hardcover
ISBN: 978-0-85729-117-2