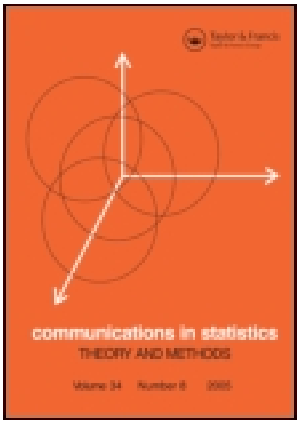

Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lsta20

Nonparametric Estimation of the Average Availability

N. Balakrishna [a] & Angel Mathew [a]
[a] Department of Statistics, Cochin University of Science and Technology, Cochin, India

Available online: 01 Apr 2009

Communications in Statistics—Theory and Methods, 38: 1207–1218, 2009

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610920802382627

Nonparametric Estimation of the Average Availability

N. BALAKRISHNA AND ANGEL MATHEW

Department of Statistics, Cochin University of Science and Technology, Cochin, India

The average availability of a repairable system is the expected proportion of time that the system is operating in the interval $[0, t]$. *The present article discusses the nonparametric estimation of the average availability when* (i) *the data on 'n' complete cycles of system operation are available,* (ii) *the data are subject to right censorship, and* (iii) *the process is observed upto a specified time 'T'. In each case, a nonparametric confidence interval for the average availability is also constructed. Simulations are conducted to assess the performance of the estimators.*

Keywords Censored data; Empirical distribution function; Nonparametric confidence interval; Product limit estimator; Renewal function.

Mathematics Subject Classification 62G05; 62N05.

1. Introduction

Consider a repairable system which is at any time either in operation or under repair after failure. Suppose that the system starts to operate at time $t = 0$. Let $\{X_n\}$ and $\{Y_n\}$ denote the sequences of operating and repair times, respectively. The first operating time and repair time constitute the first cycle of the system. Assuming that the sequences of operating and repair times constitute an alternating renewal process, a number of useful measures of the availability of such a system may be constructed. For example, we can determine the probability that the system is available at a given time (point availability) and the expected proportion of time that the system is operating in a given time interval (average availability). If we define

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

Received May 12, 2008; Accepted August 1, 2008
Address correspondence to N. Balakrishna, Department of Statistics, Cochin University of Science and Technology, Cochin 682 022, India; E-mail: nb@cusat.ac.in

then the point availability of the system is defined by $A(t) = P[\xi(t) = 1]$ and the average availability is defined as

$$A_{\text{avg}}(t) = \frac{1}{t}\int_0^t A(u)du = \frac{\bar{\alpha}(t)}{t}, \tag{1}$$

where $\bar{\alpha}(t)$ is the average time the system is operating within $[0, t]$; see Barlow and Proschan (1975). The properties of these measures are usually studied using the successive operating and repair times.

Let $\{X_n\}$ and $\{Y_n\}$ be independent sequences of independent and identically distributed (i.i.d.) non negative random variables with common distribution functions $F_X(\cdot)$ and $F_Y(\cdot)$, respectively. Assume that $F_X(\cdot)$ and $F_Y(\cdot)$ have positive mean μ_X and μ_Y and finite variance σ_X^2 and σ_Y^2, respectively. Define $Z_n = X_n + Y_n$. Let $F_Z(\cdot)$ be the marginal distribution function of the sequence $\{Z_n\}$ having mean $\mu_Z = \mu_X + \mu_Y$.

Let $S_n = \sum_{i=1}^n Z_i$ and define $N(t) = \text{Sup}\{n : S_n \le t\}$. Then $N(t)$ counts the number of cycles completed in the interval $[0, t]$ and $M(t) = E[N(t)]$ is the renewal function associated with the sequence $\{Z_n\}$. By definition, $M(t) = \sum_{k=1}^{\infty} F_Z^{(k)}(t)$, where $F_Z^{(k)}(t) = P[S_k \le t]$ is the k-fold convolution of $F_Z(t)$ and $F_Z(t) = F_X * F_Y(t)$, where $*$ denotes the convolution operator. Now the expression for the point availability $A(t)$ can be written as

$$A(t) = \overline{F}_X(t) + \overline{F}_X * M(t), \quad \text{where } \overline{F}_X(\cdot) = 1 - F_X(\cdot).$$

Since it is difficult to obtain closed form expressions for $A(t)$, except for few simple cases, in the literature more attention is being paid to the limiting measure $A = \lim_{t\to\infty} A(t) = \mu_X/(\mu_X + \mu_Y)$ called the limiting availability; see, for example, Mi (1995), Baxter and Li (1996), and Abraham and Balakrishna (2000). The nonparametric point and interval estimation of the point availability has been discussed by Baxter and Li (1994) and Li (1999) in the case of complete and censored observations, respectively. Ouhbi and Liminios (2003) constructed a nonparametric confidence interval for the point availability as a special case of semi-Markov process. But we have not come across any work on the estimation of the average availability. However, it is a valuable measure of performance of a repairable system as it captures availability behavior over a finite period of time. In this article, we consider the nonparametric estimation of the average availability of a system over the interval $[0, t]$.

From the definition of the average availability stated in (1), it follows that $A_{\text{avg}}(t)$ is not a probability, but represents the expected proportion of "uptime" over the interval $[0, t]$ of system operation. At any time 't', we have $M(t)\mu_Z \le t < (M(t) + 1)\mu_Z$. Assuming that the system is operating at time $t = 0$, $\bar{\alpha}(t)$, the average up time in the interval $[0, t]$ can be written as

$$\bar{\alpha}(t) = \begin{cases} t - M(t)\mu_Y & \text{if } M(t)\mu_Z \le t < M(t)\mu_Z + \mu_X \\ (M(t) + 1)\mu_X & \text{if } M(t)\mu_Z + \mu_X \le t < (M(t) + 1)\mu_Z \end{cases}.$$

That is,

$$\bar{\alpha}(t) = \lambda(t)\{(M(t) + 1)\mu_X\} + (1 - \lambda(t))\{t - M(t)\mu_Y\}, \tag{2}$$

where $\lambda(t) = I\{M(t)\mu_Z + \mu_X \leq t\}$ and $I(B)$ denotes the indicator function of an event B.

Thus,

$$A_{\text{avg}}(t) = \frac{1}{t}\left[\lambda(t)\{(M(t)+1)\mu_X\} + (1-\lambda(t))\{t - M(t)\mu_Y\}\right]. \tag{3}$$

If the system is under repair at time $t = 0$, then the expression for average up time takes the form

$$\bar{\alpha}^*(t) = \eta(t)\{t - (M(t)+1)\mu_Y\} + (1-\eta(t))M(t)\mu_X,$$

where $\eta(t) = I\{M(t)\mu_Z + \mu_Y \leq t\}$ and hence the expression for the average availability will be

$$A^*_{\text{avg}}(t) = \frac{1}{t}\left[\eta(t)\{t - (M(t)+1)\mu_Y\} + (1-\eta(t))M(t)\mu_X\right]. \tag{4}$$

As $t \to \infty$, the estimators of $A_{\text{avg}}(t)$ and $A^*_{\text{avg}}(t)$ have similar asymptotic properties and their proofs are almost identical. Hence, in this article, we present the asymptotic properties of $A_{\text{avg}}(t)$ defined by (3).

In Sec. 2, we discuss the nonparametric estimation of $A_{\text{avg}}(t)$ based on complete observations. Section 3 discusses the estimation in the case of censored observations and in Sec. 4, we consider the estimation in the case of continuous observation over a fixed period. Section 5 presents some numerical illustrations.

2. Estimation of Average Availability in the Case of Complete Observations

Suppose that observations on the failure times $X_1, X_2, \ldots, X_n$ and the repair times $Y_1, Y_2, \ldots, Y_n$ are available. Let $\widehat{F}_X(t)$ and $\widehat{F}_Y(t)$ denote the empirical distribution function of the random variables X and Y, respectively. By definition,

$$\widehat{F}_X(t) = \frac{1}{n}\sum_{i=1}^{n} I\{X_i \leq t\} \quad \text{and} \quad \widehat{F}_Y(t) = \frac{1}{n}\sum_{i=1}^{n} I\{Y_i \leq t\}.$$

Then a natural nonparametric estimator of μ_X and μ_Y are given by

$$\hat{\mu}_X = \int_0^\infty x d\widehat{F}_X(x) = \overline{X} \quad \text{and} \quad \hat{\mu}_Y = \int_0^\infty x d\widehat{F}_Y(x) = \overline{Y}, \quad \text{respectively.}$$

Nonparametric estimation of the renewal function has been discussed by many authors; see, for example, Frees (1986), Grubel and Pitts (1993), and Harel et al. (1995). For fixed t, Baxter and Li (1994) proposed a method for constructing nonparametric confidence intervals for the renewal function which is easier to compute than that of Frees (1986). Thus, an estimator for $M(t)$ is given by

$$\widehat{M}_n(t) = \sum_{k=1}^{\infty} \widehat{F}_Z^{(k)}(t), \quad \text{where } \widehat{F}_Z(t) = \widehat{F}_X * \widehat{F}_Y(t). \tag{5}$$

We propose an estimator for the average availability as

$$\widehat{A}_{\text{avg}}(t) = \frac{\hat{\bar{\alpha}}_n(t)}{t}, \tag{6}$$

where $\hat{\bar{\alpha}}_n(t) = \hat{\lambda}_n(t)\{(\widehat{M}_n(t)+1)\hat{\mu}_X\} + (1-\hat{\lambda}_n(t))\{t - \widehat{M}_n(t)\hat{\mu}_Y\}$, with $\hat{\lambda}_n(t) = I\{\widehat{M}_n(t)\hat{\mu}_Z + \hat{\mu}_X \leq t\}$ and $\hat{\mu}_Z = \hat{\mu}_X + \hat{\mu}_Y$.

We prove the strong consistency of the proposed estimator in the following theorem.

Theorem 2.1. *As* $n \to \infty$, $\widehat{A}_{\text{avg}}(t) \to A_{\text{avg}}(t)$ *almost surely* (*a.s.*).

Proof. Baxter and Li (1994) studied asymptotic properties of the estimator $\widehat{M}_n(t)$ defined by (5) and shown that $\widehat{M}_n(t) \to M(t)$ (a.s.) as $n \to \infty$. By the strong law of large numbers, we have $\hat{\mu}_X \to \mu_X$, $\hat{\mu}_Y \to \mu_Y$, and $\hat{\mu}_Z \to \mu_Z$ (a.s.) as $n \to \infty$. Using the fact that $\widehat{M}_n(t)\hat{\mu}_Z + \hat{\mu}_X \to M(t)\mu_Z + \mu_X$ (a.s.), we can conclude that $\hat{\lambda}_n(t) \to \lambda(t)$ (a.s.) as $n \to \infty$. Thus, $\hat{\bar{\alpha}}_n(t) \to \bar{\alpha}(t)$ (a.s.) and hence $\widehat{A}_{\text{avg}}(t) \to A_{\text{avg}}(t)$ (a.s.) as $n \to \infty$.

In order to prove the weak convergence of $\widehat{A}_{\text{avg}}(t)$, let us define $\Delta\mu_X = \hat{\mu}_X - \mu_X$, $\Delta\mu_Y = \hat{\mu}_Y - \mu_Y$, $\Delta\mu_Z = \hat{\mu}_Z - \mu_Z$, $\Delta M(t) = \widehat{M}_n(t) - M(t)$, and $\Delta\lambda(t) = \hat{\lambda}_n(t) - \lambda(t)$.

Introducing the notations $K_1(t) = \lambda(t)[M(t)+1]/t$, $K_2(t) = [\lambda(t)\mu_X - (1-\lambda(t))\mu_Y]/t$, $K_3(t) = M(t)[1-\lambda(t)]/t$, $J_X(t) = F_X * M * M(t)$, $J_Y(t) = F_Y * M * M(t)$, and writing $\Delta M(t)$ in the form of Eq. (2.2) of Harel et al. (1995) we can write

$$\sqrt{n}[\widehat{A}_{\text{avg}}(t) - A_{\text{avg}}(t)] = \sqrt{n}(I_1 + I_2 + I_3),$$

where $I_1 = K_1(t)\Delta\mu_X + K_2(t)J_Y * \Delta F_X(t)$, $I_2 = K_2(t)J_X * \Delta F_Y(t) - K_3(t)\Delta\mu_Y$, and I_3 contains terms involving $\Delta\lambda(t)$ and terms of the form $\Delta A \Delta B$ or $\Delta A * \Delta B$.

By writing $\sqrt{n}\Delta A * \Delta B = \sqrt{n}\Delta A * \widehat{B}_n - \sqrt{n}\Delta A * B$, it is easy to see that the two terms on the right-hand side converge almost surely to the same limit by Lemma 2.1 of Baxter and Li (1994). Since $\sqrt{n}\Delta\lambda(t) \to 0$ in probability as $n \to \infty$, it is straight forward to verify that $\sqrt{n}I_3 \to 0$ in probability as $n \to \infty$.

Now proceeding along the lines of Baxter and Li (1994), we can show that $\sqrt{n}I_1 \xrightarrow{L} N(0, \sigma_1^2(t))$ and $\sqrt{n}I_2 \xrightarrow{L} N(0, \sigma_2^2(t))$ as $n \to \infty$, where

$$\begin{aligned}\sigma_1^2(t) = {} & K_1^2(t)\sigma_X^2 + K_2^2(t)[J_Y^2 * F_X(t) - [J_Y * F_X(t)]^2] \\ & + 2K_1(t)K_2(t)[J_Y * V_X(t) - \mu_X J_Y * F_X(t)]\end{aligned} \tag{7}$$

and

$$\begin{aligned}\sigma_2^2(t) = {} & K_2^2(t)[J_X^2 * F_Y(t) - [J_X * F_Y(t)]^2] + K_3^2(t)\sigma_Y^2 \\ & - 2K_2(t)K_3(t)[J_X * V_Y(t) - \mu_Y J_X * F_Y(t)],\end{aligned} \tag{8}$$

with $V_X(t) = \int_0^t x dF_X(x)$ and $V_Y(t) = \int_0^t x dF_Y(x)$.

Since ΔF_X and ΔF_Y are independent, I_1 and I_2 are also independent. This leads to the following theorem.

Theorem 2.2. *As* $n \to \infty$, $\sqrt{n}[\widehat{A}_{\text{avg}}(t) - A_{\text{avg}}(t)] \xrightarrow{L} N(0, \sigma^2(t))$, *where* $\xrightarrow{L}$ *denotes convergence in distribution and*

$$\sigma^2(t) = \sigma_1^2(t) + \sigma_2^2(t), \tag{9}$$

with $\sigma_1^2(t)$ *and* $\sigma_2^2(t)$ *are given by* (7) *and* (8), *respectively.*

Let $S_X^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$ and $S_Y^2 = \frac{1}{n-1}\sum_{i=1}^n (Y_i - \overline{Y})^2$ be estimators of σ_X^2 and σ_Y^2, respectively. Then an estimator $\hat{\sigma}^2(t)$ of $\sigma^2(t)$ can be obtained on replacing $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, F_X(\cdot), F_Y(\cdot)$, and $M(\cdot)$ by $\overline{X}, \overline{Y}, S_X^2, S_Y^2, \widehat{F}_X(\cdot), \widehat{F}_Y(\cdot)$, and $\widehat{M}_n(\cdot)$, respectively in (9). Using Lemma 2.1 of Baxter and Li (1994), it can be shown that $\hat{\sigma}^2(t) \to \sigma^2(t)$ almost surely as $n \to \infty$. Thus, given a significance level $\alpha \in (0, 1)$, an approximate large sample $100(1-\alpha)\%$ confidence interval for $A_{\text{avg}}(t)$ is

$$\widehat{A}_{\text{avg}}(t) - z_{\alpha/2}\frac{\hat{\sigma}(t)}{\sqrt{n}} \le A_{\text{avg}}(t) \le \widehat{A}_{\text{avg}}(t) + z_{\alpha/2}\frac{\hat{\sigma}(t)}{\sqrt{n}},$$

where $z_{\alpha/2}$ denotes the upper $\alpha/2$ quantile of the standard normal distribution.

3. Estimation of Average Availability in the Case of Censored Observations

Suppose that observations on the failure and repair time are subject to right censorship. In practice, a censored failure time occurs when the system is removed before failure for some preventive maintenance and a censored repair time occurs when the repair work is terminated before the repair is completed due to some technical reason; for example, see Baxter and Li (1996) and Li (1999). Let $X_1, X_2, \ldots, X_n$ $(Y_1, Y_2, \ldots, Y_n)$ denote the failure (repair) times and $C_1, C_2, \ldots, C_n$ $(D_1, D_2, \ldots, D_n)$ denote the random censoring times associated with the failure (repair) times having distribution functions $F_X(F_Y)$ and $G_C(G_D)$, respectively. Suppose that the four random sequences $\{X_i\}, \{Y_i\}, \{C_i\}$, and $\{D_i\}$ are mutually independent and continuous. Under the censoring model, instead of observing X_i, we observe the pair $(L_i, \delta_i), i = 1, 2, \ldots, n$, where $L_i = \min(X_i, C_i)$ and $\delta_i = I(X_i \le C_i)$. Let $H_X(t) = 1 - (1 - F_X(t))(1 - G_C(t))$ be the distribution function of L_i and $\tau_X = \inf\{x : H_X(x) = 1\} \le \infty$ be the least upper bound for the support of H_X. With right-censored data, the most commonly used nonparametric estimator of F_X is the product limit estimator (PLE) (Kaplan and Meier, 1958)

$$\widehat{F}_{X,c}(t) = 1 - \prod_{i=1}^{n}\left[1 - \frac{\delta_{(i)}}{n-i+1}\right]^{I(L_{(i)} \le t)} \quad \text{for } t \le L_{(n)}, \quad \text{and 1 for } t > L_{(n)}$$

where $L_{(1)} \le L_{(2)} \le \cdots \le L_{(n)}$ are the order statistics of $L_1, L_2, \ldots, L_n$ and $\delta_{(i)}$ denotes the concomitant associated with $L_{(i)}$. Similarly, we can construct the product limit estimator $\widehat{F}_{Y,c}$ of F_Y. Let $H_Y(t) = 1 - (1 - F_Y(t))(1 - G_D(t))$ and $\tau_Y = \inf\{x : H_Y(x) = 1\}$.

Then a natural nonparametric estimator of $\mu_X(\mu_Y)$ is

$$\hat{\mu}_{X,c} = \int_0^\infty \hat{\overline{F}}_{X,c}(t)dt \left(\hat{\mu}_{Y,c} = \int_0^\infty \hat{\overline{F}}_{Y,c}(t)dt\right), \quad \text{where } \overline{F}_X = 1 - F_X \ (\overline{F}_Y = 1 - F_Y).$$

Let $\widehat{M}_{c,n}(t)$ be an estimator of the renewal function $M(t)$ obtained by replacing F_X and F_Y with $\widehat{F}_{X,c}$ and $\widehat{F}_{Y,c}$, respectively. Then $\widehat{M}_{c,n}(t) = \sum_{k=1}^{\infty} \widehat{F}_{Z,c}^{(k)}(t)$, where $\widehat{F}_{Z,c}(t) = \widehat{F}_{X,c} * \widehat{F}_{Y,c}(t)$.

In this case, a nonparametric estimator of $A_{\text{avg}}(t)$ is given by

$$\widehat{A}_{\text{avg},c}(t) = \frac{\hat{\bar{\alpha}}_{c,n}(t)}{t}, \tag{10}$$

where $\hat{\bar{\alpha}}_{c,n}(t) = \hat{\lambda}_c(t)\{(\widehat{M}_{c,n}(t)+1)\hat{\mu}_{X,c}\} + (1-\hat{\lambda}_c(t))\{t - \widehat{M}_{c,n}(t)\hat{\mu}_{Y,c}\}$, with $\hat{\lambda}_{c,n}(t) = I\{\widehat{M}_{c,n}(t)\hat{\mu}_{Z,c} + \hat{\mu}_{X,c} \le t\}$ and $\hat{\mu}_{Z,c} = \hat{\mu}_{X,c} + \hat{\mu}_{Y,c}$.

Before going to study the asymptotic properties of the estimator $\widehat{A}_{\text{avg},c}(t)$, we shall define

$$\mu_{X,c} = \int_0^{\tau_X} \overline{F}_X(t)dt, \quad \mu_{Y,c} = \int_0^{\tau_Y} \overline{F}_Y(t)dt, \quad \lambda_c(t) = I\{M(t)\mu_{Z,c} + \mu_{X,c} \le t\},$$
$$\mu_{Z,c} = \mu_{X,c} + \mu_{Y,c}, \quad \text{and} \quad \bar{\alpha}_c(t) = \lambda_c(t)\{(M(t)+1)\mu_{X,c}\} + (1-\lambda_c(t))\{t - M(t)\mu_{Y,c}\}.$$

Theorem 3.1. *As $n \to \infty$, $\widehat{A}_{\text{avg},c}(t) \to A_{\text{avg},c}(t)$ almost surely for $t < \tau$, where $\tau = \min(\tau_X, \tau_Y)$ and $A_{\text{avg},c}(t) = \bar{\alpha}_c(t)/t$.*

Proof. Li (1999) discussed the nonparametric estimation of the renewal function with right-censored data and proved that $\widehat{M}_{c,n}(t) \to M(t)$ almost surely as $n \to \infty$. Asymptotic properties of the mean survival time for right-censored data have been discussed by Susarla and Van Ryzin (1980) and Stute and Wang (1994). Based on their results, it is easy to see that $\hat{\mu}_{X,c} \to \mu_{X,c}$ (a.s.) as $n \to \infty$, where $\mu_{X,c}$ may not be equal to μ_X since the data (L_i, δ_i), $i = 1, 2, \ldots, n$ provide no information about F_X beyond τ_X. Similarly, $\hat{\mu}_{Y,c} \to \mu_{Y,c}$ (a.s.) as $n \to \infty$. Then for $t < \tau$, it can be shown that $\widehat{M}_{c,n}(t)\hat{\mu}_{Z,c} + \hat{\mu}_{X,c} \to M(t)\mu_{Z,c} + \mu_{X,c}$ (a.s.) and hence $\hat{\lambda}_{c,n}(t) \to \lambda_c(t)$ (a.s.). Thus, $\hat{\bar{\alpha}}_{c,n}(t) \to \bar{\alpha}_c(t)$ (a.s.) leads to the conclusion that $\widehat{A}_{\text{avg},c}(t) \to A_{\text{avg},c}(t)$ almost surely as $n \to \infty$.

Remark 3.1. If F_X, F_Y, G_C, and G_D have unbounded support, then $\tau_X = \tau_Y = \infty$ and hence $A_{\text{avg},c}(t) = A_{\text{avg}}(t)$. Also, if the least upper bound for the support of F_X and F_Y are less than or equal to τ_X and τ_Y, respectively, even if they have bounded support, $A_{\text{avg},c}(t) = A_{\text{avg}}(t)$, as $\mu_{X,c} = \mu_X$ ($\mu_{Y,c} = \mu_Y$).

In order to establish the weak convergence of $\widehat{A}_{\text{avg},c}(t)$, introduce the notations

$$U_X(t) = \int_0^t \frac{dF_X(x)}{\overline{F}_X(x)\overline{H}_X(x)}, \quad U_Y(t) = \int_0^t \frac{dF_Y(x)}{\overline{F}_Y(x)\overline{H}_Y(x)}, \quad Q_X(t) = \int_t^{\tau_X} \overline{F}_X(x)dx,$$
$$Q_Y(t) = \int_t^{\tau_Y} \overline{F}_Y(x)dx, \quad K_{1,c}(t) = \lambda_c(t)[M(t)+1]/t,$$
$$K_{2,c}(t) = [\lambda_c(t)\mu_{X,c} - (1-\lambda_c(t))\mu_{Y,c}]/t$$

and $K_{3,c}(t) = M(t)[1-\lambda_c(t)]/t$. Now by proceeding in the lines of the proof of Theorem 2.2 and using Lemma 3 of Li (1999), we can prove the following theorem.

Theorem 3.2. *As* $n \to \infty$, $\sqrt{n}[\widehat{A}_{\text{avg},c}(t) - A_{\text{avg},c}(t)] \xrightarrow{L} N(0, \sigma_c^2(t))$, *with*

$$\sigma_c^2(t) = \sigma_{1,c}^2(t) + \sigma_{2,c}^2(t), \tag{11}$$

where

$$\sigma_{1,c}^2(t) = \int_0^{\tau_X} \left[K_{2,c}(t)J_Y(t-x)\overline{F}_X(x) + K_{2,c}(t)R_X(t-x) - K_{1,c}(t)Q_X(x)\right]^2 dU_X(x)$$

and

$$\sigma_{2,c}^2(t) = \int_0^{\tau_Y} \left[K_{2,c}(t)J_X(t-x)\overline{F}_Y(x) + K_{2,c}(t)R_Y(t-x) + K_{3,c}(t)Q_Y(x)\right]^2 dU_Y(x)$$

with

$$R_X(t-x) = \int_x^t J_Y(t-y)d\overline{F}_X(y) \quad \textit{and} \quad R_Y(t-x) = \int_x^t J_X(t-z)d\overline{F}_Y(z).$$

In order to obtain a consistent estimator of $U_X(t)$ and $U_Y(t)$, we use the arguments given in Baxter and Li (1996). They propose a consistent estimator $\widehat{U}_X(t)$ of $U_X(t)$ defined by

$$\widehat{U}_X(t) = n\int_0^t \frac{dN_{1X}(x)}{N_{2X}(x)[N_{2X}(x) - I_X(x)]},$$

where $N_{1X}(t) = \#\{k : Z_k = X_k \le t\}$, $N_{2X}(t) = \#\{k : Z_k > t\}$ with $\#\{B\}$ denotes the cardinality of the set B and $I_X(t) = 1$ if there is a k such that $Z_k = X_k = t$, $I_X(t) = 0$ otherwise. Similarly, an estimator $\widehat{U}_Y(t)$ of $U_Y(t)$ can be constructed. On replacing $\mu_{X,c}$, $\mu_{Y,c}$, $F_X(\cdot)$, $F_Y(\cdot)$, $U_X(\cdot)$, $U_Y(\cdot)$ by their corresponding consistent estimators in (11), a consistent estimator $\hat{\sigma}_c^2(t)$ of $\sigma_c^2(t)$ is obtained. Thus, given a significance level $\alpha \in (0, 1)$, an approximate large sample $100(1-\alpha)\%$ confidence interval for $A_{\text{avg},c}(t)$ is

$$\widehat{A}_{\text{avg},c}(t) - z_{\alpha/2}\frac{\hat{\sigma}_c(t)}{\sqrt{n}} \le A_{\text{avg},c}(t) \le \widehat{A}_{\text{avg},c}(t) + z_{\alpha/2}\frac{\hat{\sigma}_c(t)}{\sqrt{n}}.$$

4. Estimation of Average Availability in the Case of Continuous Observation Over a Fixed Period

Suppose that the process is observed continuously over a fixed period $[0, T]$. Let $N_X(T)$ and $N_Y(T)$ denote the number of completed failures and repairs up to time T. Then the empirical estimators for the distribution functions $F_X(t)$ and $F_Y(t)$ can be defined as:

$$\widehat{F}_{X,T}(t) = \frac{1}{N_X(T)}\sum_{i=1}^{N_X(T)} I\{X_i \le t\} \quad \text{and} \quad \widehat{F}_{Y,T}(t) = \frac{1}{N_Y(T)}\sum_{i=1}^{N_Y(T)} I\{Y_i \le t\}.$$

In this case, natural nonparametric estimators for μ_X and μ_Y are given by

$$\hat{\mu}_X = \int_0^\infty x d\widehat{F}_{X,T}(x) = \frac{1}{N_X(T)}\sum_{i=1}^{N_X(T)} X_i = \overline{X}_{N_X(T)} \quad \text{and}$$

$$\hat{\mu}_Y = \int_0^\infty x d\widehat{F}_{Y,T}(x) = \frac{1}{N_Y(T)} \sum_{i=1}^{N_Y(T)} Y_i = \overline{Y}_{N_Y(T)}, \quad \text{respectively.}$$

An estimator of the renewal function $M(t)$ in this case is given by

$$\widehat{M}_T(t) = \sum_{k=1}^{\infty} \widehat{F}^{(k)}_{Z,T}(t), \quad \text{where } \widehat{F}_{Z,T}(t) = \widehat{F}_{X,T} * \widehat{F}_{Y,T}(t).$$

As a nonparametric estimator of $A_{\text{avg}}(t)$ we consider

$$\widehat{A}_{\text{avg},T}(t) = \frac{\hat{\hat{\alpha}}_T(t)}{t}, \tag{12}$$

where $\hat{\hat{\alpha}}_T(t) = \hat{\lambda}_T(t)\{(\widehat{M}_T(t) + 1)\hat{\mu}_{X,T}\} + (1 - \hat{\lambda}_T(t))\{t - \widehat{M}_T(t)\hat{\mu}_{Y,T}\}$ with $\hat{\lambda}_T(t) = I\{\widehat{M}_T(t)\hat{\mu}_{Z,T} + \hat{\mu}_{X,T} \le t\}$ and $\hat{\mu}_{Z,T} = \hat{\mu}_{X,T} + \hat{\mu}_{Y,T}$.

The strong consistency of the proposed estimator is stated in the following theorem whose proof follows parallel to that of Theorem 2.1 once we note that $N_X(T)$ and $N_Y(T)$ tends to infinity as $T \to \infty$.

Theorem 4.1. *As $T \to \infty$, $\widehat{A}_{\text{avg},T}(t) \to A_{\text{avg}}(t)$ almost surely.*

In order to study the weak convergence of $\widehat{A}_{\text{avg},T}(t)$ by introducing the notation $\Delta_T A = \widehat{A}_T - A$ and then proceeding as in Theorem 2.2, we can write

$$\sqrt{T}[\widehat{A}_{\text{avg},T}(t) - A_{\text{avg}}(t)] = \sqrt{T}(I_{1,T} + I_{2,T} + I_{3,T}),$$

where

$$I_{1,T} = K_1(t)\Delta_T\mu_X + K_2(t)J_Y * \Delta_T F_X(t), \quad I_{2,T} = K_2(t)J_X * \Delta_T F_Y(t) - K_3(t)\Delta_T\mu_Y.$$

Further, $I_{3,T}$ is obtained by replacing Δ by Δ_T in I_3.

Following the arguments in Theorem 2.2 and using Lemma 3.1 stated in Ouhbi and Liminios (2003), it follows that $\sqrt{T}I_{3,T} \to 0$ in probability as $T \to \infty$.

Writing $\sqrt{T}I_{1,T} = \sqrt{\frac{T}{N_X(T)}}\sqrt{N_X(T)}I_{1,T}$ and using the fact that $N_X(T)/T \to 1/\mu_X$ as $T \to \infty$, we can show that $\sqrt{T}I_{1,T}$ follows a normal distribution with mean 0 and variance $\sigma^2_{1,T}(t)$, where

$$\begin{aligned}\sigma^2_{1,T}(t) = {} & K_1^2(t)\mu_X\sigma_X^2 + K_2^2(t)\mu_X[J_Y^2 * F_X(t) - [J_Y * F_X(t)]^2] \\ & + 2K_1(t)K_2(t)\mu_X[J_Y * V_X(t) - \mu_X J_Y * F_X(t)].\end{aligned} \tag{13}$$

On the similar lines,

$$\sqrt{T}I_{2,T} \xrightarrow{L} N(0, \sigma^2_{2,T}(t)),$$

where

$$\begin{aligned}\sigma^2_{2,T}(t) = {} & K_2^2(t)\mu_Y[J_X^2 * F_Y(t) - [J_X * F_Y(t)]^2] + K_3^2(t)\mu_Y\sigma_Y^2 \\ & - 2K_2(t)K_3(t)\mu_Y[J_X * V_Y(t) - \mu_Y J_X * F_Y(t)].\end{aligned} \tag{14}$$

This leads to the following theorem.

Theorem 4.2. *As* $T \to \infty$, $\sqrt{T}[\widehat{A}_{\text{avg},T}(t) - A_{\text{avg}}(t)] \xrightarrow{L} N(0, \sigma_T^2(t))$, *where*

$$\sigma_T^2(t) = \sigma_{1,T}^2(t) + \sigma_{2,T}^2(t), \tag{15}$$

with $\sigma_{1,T}^2(t)$ *and* $\sigma_{2,T}^2(t)$ *are given in* (13) *and* (14), *respectively.*

This result can be used to construct $100(1-\alpha)\%$ asymptotic confidence interval for $A_{\text{avg}}(t)$ as before.

5. Numerical Studies

In this section, we present a simulation study in order to assess the performance of the proposed estimator in the case of (i) complete observations, (ii) censored observations, and (iii) continuous observation over a fixed period. We use the algorithm proposed by Schneider et al. (1990) for computing the renewal function. Let $0 = t_0 < t_1 < \cdots < t_m = t$ be an equally spaced partition of $[0, t]$, where the choice of m depends on t and on the data. An algorithm for computing the estimates and the confidence interval for $A_{\text{avg}}(t)$ can be summarized as follows.

1. Compute $\widehat{F}_X, \widehat{F}_Y, \hat{\mu}_X, \hat{\mu}_Y$ and the standard deviations $\hat{\sigma}_X$ and $\hat{\sigma}_Y$.
2. Find $\widehat{F}_Z(t_i) = \sum_{j=1}^{m} \widehat{F}_X(t_i - t_j)[\widehat{F}_Y(t_j) - \widehat{F}_Y(t_{j-1})]$ for $i = 1, 2, \ldots, m$.
3. Evaluate $\widehat{M}(t)$ using the recursive relationship

$$\widehat{M}(t_i) = \widehat{F}_Z(t_i) + \sum_{j=1}^{i} \widehat{M}(t_i - t_j)[\widehat{F}_Z(t_j) - \widehat{F}_Z(t_{j-1})], \quad \text{for } i = 1, 2, \ldots, m.$$

   and compute $\widehat{A}_{\text{avg}}(t)$.
4. Compute $\widehat{J}_X(t_i), \widehat{J}_Y(t_i), \widehat{V}_X(t_i)$, and $\widehat{V}_Y(t_i)$ then $\widehat{J}_X * \widehat{F}_Y(t_i), \widehat{J}_X^2 * \widehat{F}_Y(t_i), \widehat{J}_X * \widehat{V}_Y(t_i)$, $\widehat{J}_Y * \widehat{F}_X(t_i), \widehat{J}_Y^2 * \widehat{F}_X(t_i)$, and $\widehat{J}_Y * \widehat{V}_X(t_i)$ recursively for $i = 1, 2, \ldots, m$.
5. Substitute the values obtained in the above steps to evaluate $\hat{\sigma}^2(t)$.

The same algorithm can be used to compute the confidence interval for $A_{\text{avg},c}(t)$ and $\widehat{A}_{\text{avg},T}(t)$ defined in (10) and (12), respectively, after appropriate modifications.

Consider first the case of complete observations. Suppose that the distribution of the failure times is gamma with shape parameter 3 and scale parameter 2 and the repair times also follow a gamma distribution with shape parameter 1 and scale parameter 2. Three time points $t = 2.5$, $t = 5$, and $t = 7.5$ are considered for the simulation. The exact values of $A_{\text{avg}}(t)$ at these points are obtained using *Mathematica*. In Table 1, 'n' denotes the number of observations of operating and repair times, $\bar{\widehat{A}}_{\text{avg}}(t)$ denotes the average of $\widehat{A}_{\text{avg}}(t)$ over 100 repetitions at 't', $\bar{\hat{\sigma}}(t)$ denotes the sample mean of the estimated standard error of the estimate and $A_{\text{avg},L}(t)$ and $A_{\text{avg},U}(t)$ denote the 95% lower and upper confidence limits for $A_{\text{avg}}(t)$, respectively. The values given in parenthesis represent the standard error of the corresponding estimator.

In order to check the performance of the estimator under censoring we suppose that F_X is a gamma distribution with shape parameter 3 and scale parameter 2, and that F_Y is a gamma distribution with shape parameter 2 and scale parameter 1. Further assume that censoring distributions are exponential with $G_C(t) = 1 - e^{-0.05t}$

Table 1
Simulation results for average availability in the case of complete observations

t	$A_{avg}(t)$	n	$\bar{\widehat{A}}_{avg}(t)$	$\bar{\hat{\sigma}}(t)$	$A_{avg,L}(t)$	$A_{avg,U}(t)$
		25	0.96171 (0.0324)	0.00537 (0.0048)	0.95118	0.97223
2.5	0.95852	75	0.95809 (0.0242)	0.00346 (0.0021)	0.95131	0.96486
		150	0.95933 (0.0131)	0.00236 (0.0008)	0.95470	0.96396
		25	0.88572 (0.0358)	0.01532 (0.0056)	0.85570	0.91574
5	0.88641	75	0.88895 (0.0209)	0.00895 (0.0019)	0.87141	0.90650
		150	0.88462 (0.0151)	0.00674 (0.0010)	0.87141	0.89783
		25	0.84120 (0.0324)	0.02231 (0.0059)	0.79746	0.88493
7.5	0.84232	75	0.84136 (0.0199)	0.01297 (0.0021)	0.81595	0.86677
		150	0.84190 (0.0138)	0.00913 (0.0010)	0.82400	0.85980

Table 2
Simulation results for average availability in the case of censored observations

t	$A_{avg}(t)$	n	$\bar{\widehat{A}}_{avg,c}(t)$	$\bar{\hat{\sigma}}_c(t)$	$X\%$	$Y\%$	$A_{avg,L}(t)$	$A_{avg,U}(t)$
		25	0.96102 (0.0233)	0.00584 (0.0044)	25.16	16.28	0.94956	0.97247
2.5	0.95852	75	0.96077 (0.0136)	0.00351 (0.0014)	24.25	17.57	0.95389	0.96765
		150	0.96090 (0.0085)	0.00243 (0.0006)	24.57	17.80	0.95613	0.96567
		25	0.89173 (0.0338)	0.01792 (0.0189)	25.68	17.88	0.85661	0.92686
5	0.88641	75	0.88795 (0.0190)	0.00979 (0.0026)	24.35	17.85	0.86877	0.90713
		150	0.88728 (0.0130)	0.00709 (0.0013)	24.71	17.61	0.87339	0.90117
		25	0.84946 (0.0379)	0.02408 (0.0221)	25.28	17.48	0.80227	0.89665
7.5	0.84232	75	0.84543 (0.0193)	0.01325 (0.0023)	24.88	17.56	0.81946	0.87140
		150	0.84212 (0.0157)	0.00982 (0.0015)	24.33	17.29	0.82289	0.86136

Table 3
Simulation results for average availability in the case of continuous observations

t	$A_{\text{avg}}(t)$	T	$\widehat{\bar{A}}_{\text{avg},T}(t)$	$\bar{\hat{\sigma}}_T(t)$	$\overline{N}(T)$	$A_{\text{avg},L}(t)$	$A_{\text{avg},U}(t)$
		250	0.94723 (0.0308)	0.00343 (0.0022)	31.05	0.94380	0.95066
2.5	0.95852	500	0.96432 (0.0214)	0.00160 (0.0009)	61.93	0.96271	0.96592
		1000	0.95839 (0.0199)	0.00133 (0.0006)	125.42	0.95707	0.95972
		250	0.88814 (0.0134)	0.00627 (0.0015)	31.36	0.88187	0.89440
5	0.88641	500	0.88191 (0.0111)	0.00558 (0.0007)	63.61	0.87633	0.88749
		1000	0.88599 (0.0131)	0.00401 (0.0005)	126.37	0.88198	0.89000
		250	0.84661 (0.0393)	0.00856 (0.0019)	32.32	0.83805	0.85518
7.5	0.84232	500	0.84413 (0.0217)	0.00632 (0.0011)	60.43	0.83781	0.85046
		1000	0.84164 (0.0114)	0.00506 (0.0007)	127.69	0.83658	0.84670

and $G_D(t) = 1 - e^{-0.1t}$. The results of the simulation study are presented in Table 2. Here, $X\%$ and $Y\%$ denote the average censoring rate associated with the failure time and the repair time, respectively.

Table 3 presents the result of the simulation study in the case of continuous observation over a fixed period $[0, T]$ using the same distributions for generating the failure and repair times as in the case of complete observations. Here, $\overline{N}(T)$ denotes the average number of cycles completed upto time 'T'.

From Tables 1–3, it can be seen that even for moderate sample sizes, the standard deviation of the estimate is small and the width of the confidence interval is reasonably narrow.

6. Concluding Remarks

We have discussed the nonparametric estimation of the average availability when the operating and repair times of a system are mutually independent sequences of i.i.d. random variables. The proposed estimators of the average availability are proved to be consistent and asymptotically normal when (i) the data are complete, (ii) the data are subject to right censorship, and (iii) the data are observed over a fixed period. The simulation study shows that the proposed estimators perform well even for reasonable sample sizes.

References

Abraham, B., Balakrishna, N. (2000). Estimation of limiting availability for a stationary bivariate process. *J. Appl. Prob.* 37:696–704.

Barlow, R. E., Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing*. New York: Holt, Rinehart and Winston.

Baxter, L. A., Li, L. (1994). Nonparametric confidence intervals for the renewal function and the point availability. *Scand. J. Statist.* 21:277–287.

Baxter, L. A., Li, L. (1996). Nonparametric estimation of the limiting availability. *Lifetime Data Anal.* 2:391–402.

Frees, E. W. (1986). Nonparametric renewal function estimation. *Ann. Statist.* 14:1366–1378.

Grubel, R., Pitts, S. M. (1993). Nonparametric estimation in renewal theory I: the empirical renewal function. *Ann. Statist.* 21:1431–1451.

Harel, M., O'Cinneide, C. A., Schneider, H. (1995). Asymptotics of the sample renewal function. *J. Math. Anal. Appl.* 189:240–255.

Kaplan, E. L., Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* 53:457–481.

Li, L. (1999). Estimating the point availability with right censored data. *Naval. Res. Logistics* 46:119–127.

Mi, J. (1995). Limiting behaviour of some measures of system availability. *J. Appl. Probab.* 32:482–493.

Ouhbi, B., Liminios, N. (2003). Nonparametric reliability estimation of semi-Markov processes. *J. Statist. Plann. Infer.* 109:155–165.

Schneider, H., Lin, B., O'Cinneide, C. A. (1990). Comparison of nonparametric estimators for the renewal function. *Appl. Statist.* 39:55–61.

Stute, W., Wang, J. L. (1994). The strong law under random censorship. *Ann. Statist.* 21:1591–1607.

Susarla, V., Van Ryzin, J. (1980). Large sample theory for an estimator of the mean survival time from censored samples. *Ann. Statist.* 8:1002–1016.