

STATISTICAL ANALYSIS OF SOME RELIABILITY MODELS :  
PARAMETRICS, SEMI-PARAMETRICS AND NONPARAMETRICS

by

Pranab Kumar Sen

Department of Biostatistics, University of  
North Carolina at Chapel Hill, NC.

Institute of Statistics Mimeo Series No. 2108

November 1992

# Statistical Analysis of Some Reliability Models : Parametrics, Semi-parametrics and Nonparametrics

Pranab Kumar Sen

Departments of Biostatistics & Statistics, University of North Carolina,  
Chapel Hill, NC 27599-7400, USA

**Abstract.** The four basic reliability models considered are (i) Bundle strength of filaments, (ii) K-component system - in Series, (iii) K-component system - in Parallel, and (iv) Systems availability- under spare and repair. Along with the parametric formulations, some semi-parametric and nonparametric statistical solutions are considered. The role of resampling plans in this context is discussed critically.

AMS Subject Classification : 62N05, 62G99

Key words and phrases : Bundle strength of filaments; components in parallel, components in series; bootstrapping; jackknifing; system availability.

## 1. Introduction

Four prominent statistical problems arising in reliability theory are considered here; for such problems, appropriate models, ranging from the classical parametric to semi-parametric and nonparametric ones are presented side by side and their relative merits and demerits (with especial emphasis on robustness aspects) are discussed. In this context, the role of jackknifing and bootstrapping is also critically examined.

First, we consider the bundle strength of parallel filaments, as introduced by Daniels (1945). Consider a bundle of  $n$  parallel filaments. The individual filaments have breaking strengths which we denote by  $X_1, \dots, X_n$  respectively. It is easy to conceive that in general these are random variables which we can take independent and identically distributed (i.i.d.) with a continuous distribution function (d.f.)  $F$ , defined on the positive half of the real line  $R^+ = [0, \infty)$ . The most ideal case relates to the model where  $F$  is degenerate at the point  $x_0$  :  $0 < x_0 < \infty$ , so that the bundle strength of the  $n$  parallel filaments is equal to  $nx_0$ . However, the process in which the filaments are created may generally introduce some chance variation, and hence, such an ideal degenerate case is very seldom met in practice. On the contrary, if  $F$  is not degenerate at a point  $x_0$ , the bundle strength may not behave as the sum of the  $X_i$  ( $i \leq n$ ) nor as the maximum of these individual  $X_i$ . To see this, we arrange the  $X_i$  in an ascending order and denote the ordered random variables (r.v.) by  $X_{n:1} \leq \dots \leq X_{n:n}$ ; by virtue of the assumed continuity of  $F$ , ties among the  $X_i$  (and hence,  $X_{n:i}$ ) can be ignored with probability one. For any load  $x < X_{n:1}$ , all the filaments are in tact, so

that the strength of the bundle of  $n$  filaments is not smaller than  $nx$ . For  $x$  belonging to the half-open interval  $[X_{n:1}, X_{n:2})$ , one of the filaments is broken, so that the strength is not smaller than  $(n-1)x$ , and, in general, for  $x \in [X_{n:k}, X_{n:k+1})$ , it can not be smaller than  $(n-k)x$ , for  $k = 0, \dots, n$ , where  $X_{n:n+1} = +\infty$ . Since the jump-points for this segmented lines are the order statistics  $X_{n:k}$  themselves, Daniels (1945) had no problem in setting the Bundle Strength as

$$B_n = \max\{ (n-k+1)X_{n:k} : 1 \leq k \leq n \} . \quad (1.1)$$

Defined this way,  $B_n$  is neither a sum of the  $X_i$  nor an extreme value of these  $X_i$ . However, if we define some sample functions :

$$Y_{nk} = (n-k+1)X_{n:k} , \text{ for } k = 1, \dots, n, \quad (1.2)$$

then  $B_n$  is an extremum of these  $Y_{nk}$ ; note that the  $Y_{nk}$  are neither independent nor identically distributed r.v.'s. In Section 2, we shall study some statistical aspects of  $B_n$  blending parametrics to nonparametrics in a robust way.

Consider next a chain with  $k$  loops whose individual breaking strengths are denoted by  $X_1, \dots, X_k$  respectively. Note that the chain breaks when at least one of the loops is broken. Thus, the breaking strength of the chain is defined as

$$C_k = \min\{ X_i : 1 \leq i \leq k \} = X_{k:1} . \quad (1.3)$$

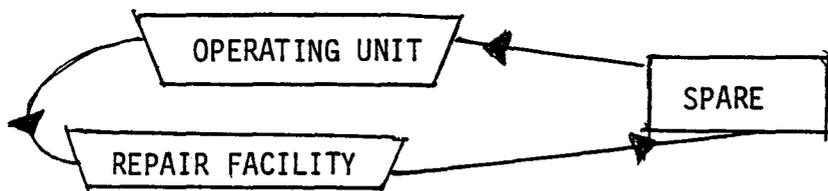
In this setup, we have the conventional sample minimum, and hence, one may use a variety of tools to study various properties of  $C_k$ . Section 3 deals with this in a broader perspective.

Suppose now that we have a system with  $n$  components of a system connected in parallel, so that the system functions if and only if at least one component functions. Thus, the life-time of the system is defined as

$$L_n = \max\{ X_i : 1 \leq i \leq n \} = X_{n:n} . \quad (1.4)$$

Unlike the case of  $C_k$ , here the distribution of  $L_n$  may not be that simple even for some simple parametric models. We relegate these details to Section 4.

As a final example, consider a one-unit system supported by a single spare and a repair facility. When the operating unit fails, it is instantaneously



replaced by the spare while the failed unit is sent to the repair shop. Upon repair, the unit is sent to the spare box. The system fails when an operating unit fails and no spare unit is there to replace it (which occurs when the

repairing time of a failed unit is larger than the life time of the operating unit about to fail). Let  $\pi$  be the probability of this event (i.e., repair time exceeds the life time). Then the limiting System Availability is defined by

$$A = E(\text{life time}) / \{ E(\text{life time}) + (1 - \pi)E(\text{Down time}) \} \quad (1.5)$$

[viz., Barlow and Proschan (1975)]. Although for some specific parametric models, simple and convenient estimators of  $A$  in (1.5) can be obtained, they may not be so robust to plausible departures from the assumed model. As such, in Section 5, we consider some nonparametric solutions.

For all the four problems sketched above, estimation of the asymptotic or finite sample variance of the estimators poses some serious analytical problems. In this context, the conventional resampling methods such as the Jackknifing and Bootstrapping are of immense help, and we shall discuss them in due course. We conclude this section with a remark that for simplicity of presentation, we will not include the relatively more complex case where the life times etc. are subject to some censoring schemes. Modifications for such schemes are fairly routine and will be briefly treated as and when appropriate.

## 2. Bundle strength

Daniels (1945) has an elegant (albeit lengthy) proof of the asymptotic normality of  $B_n$  defined in (1.1), and a much shorter proof is due to Sen, Bhattacharyya and Suh (1973). Before we proceed to discuss these results, we make some simplifications of notations which will enable us to present the parametric and semi-parametric models more appropriately. We denote the sample or empirical d.f.  $F_n$  by

$$F_n(x) = n^{-1} \sum_{i=1}^n I(X_i \leq x), \quad x \in R^+, \quad n \geq 1. \quad (2.1)$$

Then note that by (1.1) and (2.1),

$$\begin{aligned} Z_n &= n^{-1} B_n = \max\{ (1 - (i-1)/n) X_{n:i} : 1 \leq i \leq n \} \\ &= \sup\{ x[1 - F_n(x)] : x \in R^+ \}. \end{aligned} \quad (2.2)$$

This, by virtue of the Glivenko-Cantelli Lemma intuitively suggests that  $Z_n$ , the per unit bundle strength, estimates the parameter

$$\theta = \sup\{ x[1 - F(x)] : x \in R^+ \}. \quad (2.3)$$

As such, an approach to the asymptotic normality via  $Z_n$  may be much more convenient. Before we do that, let us examine the parameter  $\theta$  in (2.3) and in a parametric setup, try to obtain an optimal estimator of it.

In reliability models, researchers have a special soft corner for the simple exponential law, and hence, we start with this model first. The probability density function  $f(x)$  corresponding to the d.f.  $F(x)$  is taken as

$$f(x; \mu) = \mu \exp(-\mu x), I(x \in \mathbb{R}^+), \quad \mu > 0, \quad (2.4)$$

so that

$$EX = \int_0^\infty xf(x; \mu) dx = \mu^{-1} \text{ and } V(X) = \mu^{-2}. \quad (2.5)$$

Further, by (2.3) and (2.4),

$$(d/dx)\{x[1-F(x)]\} = [1-F(x)] - xf(x) = e^{-\mu x}\{1 - \mu x\}, \quad (2.6)$$

so that a maximum of  $x[1-F(x)]$  is attained at the point  $x_0 = \mu^{-1}$ , and we have

$$\theta = (\mu e)^{-1} = e^{-1} EX. \quad (2.7)$$

Since the minimum variance unbiased estimator (MVUE) of  $\mu^{-1}$  is the sample mean  $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$ , we are naturally led to consider the following estimator of  $\theta$ :

$$\hat{\theta}_n = e^{-1} \bar{X}_n. \quad (2.8)$$

It is easy to verify that

$$E(\hat{\theta}_n) = \theta \quad \text{and} \quad \text{Var}(\hat{\theta}_n) = e^{-2} \mu^{-2} n^{-1}. \quad (2.9)$$

We may note that  $n\bar{X}_n$  has a gamma density, so that the exact distribution of  $\hat{\theta}_n$  can be obtained from a suitable gamma one (involving the unknown scale parameter  $\mu$ ), while by the classical central limit theorem, we have from (2.8)-(2.9),

$$n^{1/2}(\hat{\theta}_n - \theta) \xrightarrow{D} \mathcal{N}(0, \theta^2), \text{ as } n \rightarrow \infty. \quad (2.10)$$

Let us now proceed to study the robustness of  $Z_n$  and  $\hat{\theta}_n$  when the d.f. may depart from an exponential one. We let  $\bar{F}(x) = 1 - F(x)$  and define the failure rate or hazard rate by

$$h_F(x) = f(x)/\bar{F}(x) = -(d/dx) \log \bar{F}(x), \quad x \in \mathbb{R}^+. \quad (2.11)$$

Note that by (2.6), we have the implicit functional equation

$$x_0 h_F(x_0) = 1, \quad (2.12)$$

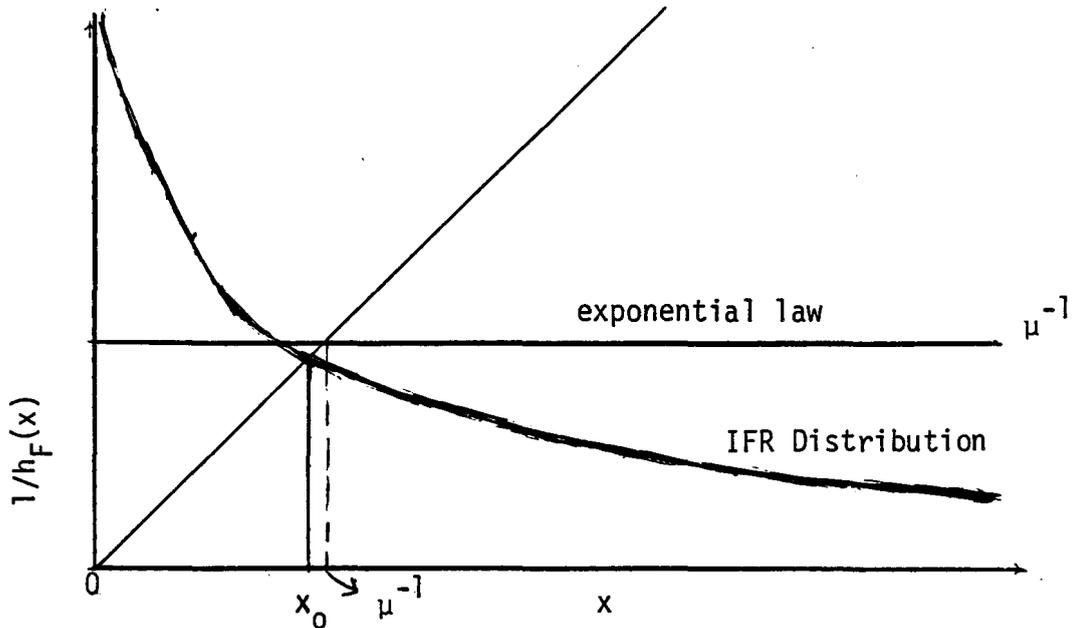
and this may provide a clear picture for a large class of d.f.'s. Consider first the increasing failure rate (IFR) distributions (on  $\mathbb{R}^+$ ) which are characterized by the monotone nondecreasing nature of the failure rate  $h_F(x)$ ,  $x \in \mathbb{R}^+$ . If we let

$$H_F(x) = \int_0^x h_F(x) dx, \quad x \in \mathbb{R}^+ \quad (2.13)$$

be the integrated hazard function, then we have

$$\bar{F}(x) = \exp\{-H_F(x)\}, \text{ where } H_F'(x) = h_F(x) \text{ is nonnegative.} \quad (2.14)$$

Since  $\bar{F}(0) = 1$ ,  $\bar{F}(\infty) = 0$  and  $\bar{F}(x)$  is monotone nonincreasing, we have  $H_F(0) = 0$ ,  $H_F(\infty) = +\infty$  and  $H_F(x)$  monotone nondecreasing in  $x \in \mathbb{R}^+$ . Moreover  $1/h_F(x)$  is monotone nonincreasing in  $x$ . Thus, plotting  $1/h_F(x)$  against  $x$  (over  $\mathbb{R}^+$ ), we obtain the following figure, where the upper asymptote may either be 0 or a finite positive constant depending on whether  $h_F(x)$  converges to  $\infty$  or to a finite upper limit. Thus, here a solution  $x_0$  to (2.12) exists, and is unique. Note that the exponential law relates to a boundary of IFR laws (as  $h_F(x) = \text{constant}$ ).



Note that whenever  $h_F(x)$  is nondecreasing,  $xh_F(x)$  is convex ( assuming that  $h_F$  is continuous), so that if we consider an exponential law with  $\mu = 1/x_0$ , then for both the d.f.  $F$  and the constructed exponential model, the point at which the supremum is attained is the same ( equal to  $x_0$  ). However, under the exponential model, the expected value of  $X$  is  $x_0$  while the expected value of  $X$  under  $F$  depends on  $F$  through other parameters as well.

To illustrate this point further we consider the so called proportional hazards model (PHM) where some ordering of the bundle strength can be studied in a convenient manner. Suppose now that  $F$  belongs to a class  $\mathcal{F}$ , such that there exists a (fixed)  $F_0 \in \mathcal{F}$ , for which

$$h_F(x) = c(F, F_0)h_0(x), \quad h_0(x) = h_{F_0}(x), \quad \text{for every } x \in \mathbb{R}^+, \quad (2.15)$$

for every  $F \in \mathcal{F}$ , where  $c(F, F_0)$  is a positive constant, depending on  $F$ . We denote the integrated hazard function as in (2.14). Then, we have from (2.14) and (2.15),

$$H_F(x) = c(F, F_0)H_0(x), \quad H_0(x) = H_{F_0}(x), \quad \text{for every } x \in \mathbb{R}^+. \quad (2.16)$$

We define  $\theta$  as in (2.3), and express it as  $\theta(F)$ ,  $F \in \mathcal{F}$ ; also, let  $\theta_0 = \theta(F_0)$ . We denote the solution in (2.12) by  $x_0(F)$ , so that by (2.12) and (2.15),

$$x_0(F)h_F(x_0(F)) = 1 = x_0(F_0)h_0(x_0(F_0)), \quad (2.17)$$

where the left hand side of (2.17) is expressible as  $c(F, F_0)x_0(F)h_0(x_0(F))$ . If  $xh_0(x)$  is nondecreasing in  $x$  ( insured by the IFR nature of the hazard function), then from the above, we obtain that

$$x_0(F) \text{ is } \begin{matrix} \geq \\ \leq \end{matrix} x_0(F_0) \text{ according as } c(F, F_0) \text{ is } \begin{matrix} \leq \\ \geq \end{matrix} 1. \quad (2.18)$$

Actually, for (2.18) to hold, it is not necessary to assume IFR ; nondecreasing nature of  $xh_F(x)$  suffices.

Note that  $\bar{F}(x)$  is decreasing in  $x$  and by virtue of the assumed IFR property of  $F$ ,  $[h_F(x)]^{-1}$  is also nonincreasing in  $x \in R^+$ . Therefore,  $\bar{F}(x)/h_F(x)$  is nonincreasing in  $x \in R^+$ , so that by (2.14), we conclude that

$$F \text{ IFR} \Rightarrow [h_F(x)]^{-1} \exp(-H_F(x)) \text{ is nonincreasing in } x \in R^+. \quad (2.19)$$

Combining (2.15), (2.18) and (2.19), we obtain that for all  $F \in \mathcal{F}$ ,

$$\begin{aligned} [\exp(-H_F(x_0(F)))]/h_F(x_0(F)) &= [\exp(-c(F, F_0)H_0(x_0(F)))]/[c(F, F_0)h_0(x_0(F))] \\ &\stackrel{\geq}{\leq} [\exp(-H_0(x_0(F_0)))]/h_0(x_0(F_0)) \text{ according as } c(F, F_0) \text{ is } \stackrel{\leq}{\geq} 1. \end{aligned} \quad (2.20)$$

Finally, by (2.3), (2.17) and (2.20), we conclude that

$$\begin{aligned} \theta(F) &= x_0(F) [\exp(-H_F(x_0(F)))] = [\exp(-H_F(x_0(F)))]/h_F(x_0(F)) \\ &= \frac{[\exp(-c(F, F_0)H_0(x_0(F)))]/[c(F, F_0)h_0(x_0(F))]}{[\exp(-H_0(x_0(F_0)))/h_0(x_0(F_0))]} \theta(F_0) \\ &\stackrel{\geq}{\leq} \theta(F_0) \text{ according as } c(F, F_0) \text{ is } \stackrel{\leq}{\geq} 1. \end{aligned} \quad (2.21)$$

Therefore, the PHM ordering of the d.f.'s  $F$  (within the class  $\mathcal{F}$ ) renders an ordering of the parameters  $\theta(F)$ ,  $F \in \mathcal{F}$ , albeit in an opposite direction. In the particular case of exponential models [viz., (2.4)],  $h_F(x) = \mu(F) = 1/E_F(X)$ , so that recognizing that  $c(F, F_0) = [\theta(F)/\theta(F_0)]$ , we obtain from (2.21) that

$$\theta(F) = \theta(F_0) / \{\mu(F)/\mu(F_0)\}, \text{ for all exponential } F. \quad (2.22)$$

This simple scalar relationship does not hold for non-exponential d.f.'s. Also, in general, the PHM assumption may not be tenable (even within the class of IFR distributions). To stress these points, we consider the following (Weibull) model:

We consider the survival function

$$F(x) = \exp(-\mu x^\gamma) I(x \in R^+), \quad \mu > 0, \gamma > 0. \quad (2.23)$$

Here  $\gamma$  is referred to as the shape parameter. Note that here

$$h_F(x) = \mu\gamma\{x^{\gamma-1}\} \text{ and it is IFR/DFR according as } \gamma \text{ is } \stackrel{\geq}{\leq} 1. \quad (2.24)$$

We denote the left hand side of (2.24) by  $h(x; \mu, \gamma)$ , and note that

$$\frac{h(x; \mu_1, \gamma_1)}{h(x; \mu_2, \gamma_2)} = (\mu_1\gamma_1/\mu_2\gamma_2) \{x^{\gamma_1 - \gamma_2}\}. \quad (2.25)$$

Note that for  $\gamma_1 \geq \gamma_2$ , (2.25) is a nondecreasing function of  $x \in R^+$ . If  $\gamma_1 > \gamma_2$ , at  $x = 0$ , (2.25) is equal to 0, while as  $x \rightarrow \infty$ , it also tends to  $\infty$ . Thus, the two hazard functions  $h(x; \mu_1, \gamma_1)$  and  $h(x; \mu_2, \gamma_2)$  are not proportional to each other, and hence, the PHM model may not apply here. The exponential law in (2.4) corresponds to the case of  $\gamma = 1$ . We now compare (2.4) and (2.23) with a view to studying the robustness properties of the parametric estimator in (2.8). For the sake of simplicity we let  $\mu_1 = \mu_2 = \mu = 1$ . Then we have on dropping  $\mu$ ,

$$x_0(1) = 1/1 = 1 \text{ and } x_0(\gamma) = 1/h(x_0(\gamma), \gamma) = \gamma^{-1} \{x_0(\gamma)\}^{1-\gamma}$$

Thus,

$$x_0(\gamma) = \gamma^{-1/\gamma} \tag{2.26}$$

This leads to

$$\begin{aligned} \theta(\gamma) &= x_0(\gamma) [\exp(-\{x_0(\gamma)\}^\gamma)] = \gamma^{-1/\gamma} \exp(-\gamma^{-1}) \\ &= \theta(1) \{ \exp(1 - \gamma^{-1}) \} / \gamma^{1/\gamma}, \text{ for } \gamma \in (0, \infty). \end{aligned} \tag{2.27}$$

It is easy to verify that the right hand side of (2.27) attains a unique minimum at  $\gamma = 1$ , so that we conclude that

$$\theta(\gamma) \geq \theta(1), \text{ where the equality sign holds for } \gamma = 1. \tag{2.28}$$

The picture becomes a bit different for an arbitrary  $\mu$ . Denoting by  $\theta(\mu, \gamma)$  the parameter  $\theta(F)$  when  $F$  is Weibull with parameters  $\mu, \gamma$ , we have on parallel lines,

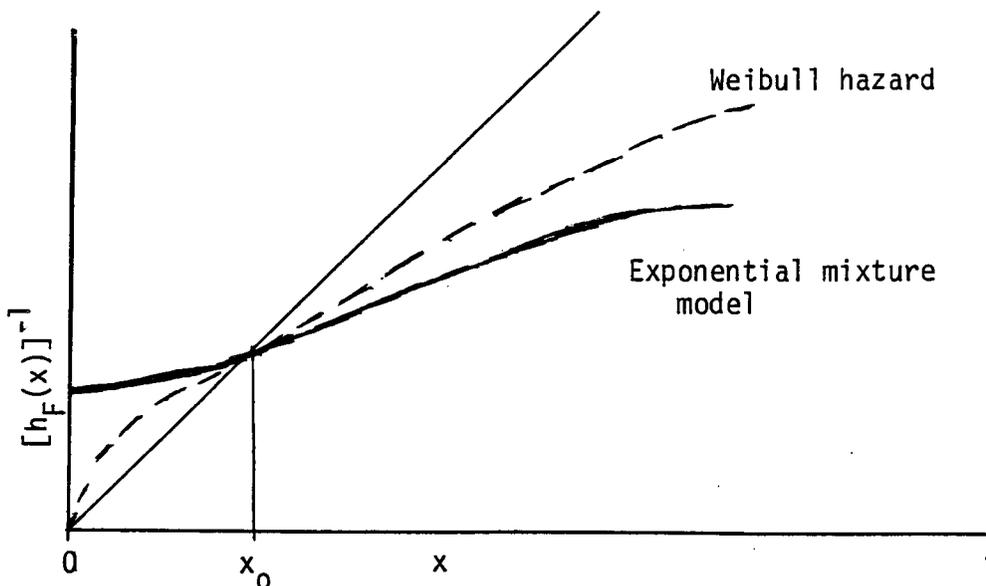
$$\log \theta(\mu, \gamma) = -\gamma^{-1} \log(\mu\gamma) - \gamma^{-1}, \tag{2.29}$$

so that

$$(\partial/\partial\gamma) \log \theta(\mu, \gamma) = \gamma^{-2} \log(\mu\gamma) \leq 0 \text{ according as } \gamma \text{ is } \leq \mu^{-1}. \tag{2.30}$$

Thus,  $\theta(\mu, \gamma)$  attains a minimum value ( $= \exp(-\mu)$ ) at  $\gamma = 1/\mu$ , its upper asymptote is equal to 1 (as  $\gamma \rightarrow \infty$ ) and at the lower end, as  $\gamma \rightarrow 0$ , it goes to  $\infty$ .

In view of (2.24), we proceed now to consider the class of decreasing failure rate (DFR) distributions, for which  $h_F(x)$  is nonincreasing in  $x \in \mathbb{R}^+$ . In such a case, we have a regular behavior whenever  $h_F(0)$  is  $> 0$  and  $h_F(x)/x$  lies below the mark 1 for all  $x \geq x_0$ , for some  $x_0 < \infty$ , or  $[h_F(x)]^{-1}$  lies above the line  $y=x$  for



for  $x \leq x_0$  albeit it may be equal to 0 at  $x = 0$ . The Weibull law belongs to the latter case when  $\gamma < 1$ . As a nice example of the former regular case, we present the following mixture model for which the density function is expressed as

$$f(x) = \mu [ \pi \exp(-\mu x) + (1-\pi)k \exp(-k\mu x) ] I(x \in \mathbb{R}^+), \tag{2.31}$$

where  $k \neq 1$ ,  $\pi \in (0,1)$  and  $\mu > 0$ . For  $\pi = 1$  or  $k = 1$ , (2.31) reduces (2.4).

Note that

$$[h_F(x)]^{-1} = \mu^{-1} [\pi e^{-\mu x} + (1-\pi) e^{-k\mu x}] [\pi e^{-\mu x} + k(1-\pi) e^{-k\mu x}]^{-1}, \quad (2.32)$$

and it fails to be equal to  $\mu^{-1}$  for all  $x \in R^+$ . Moreover, the first order partial derivative of (2.32) with respect to  $x$  is nonnegative for all  $x \in R^+$ . Finally,  $[h_F(x)]^{-1}$  assumes the value  $\mu^{-1}(\pi + k(1-\pi))^{-1} > 0$  at  $x = 0$  and at  $x = \infty$ , it has the asymptote  $\mu^{-1}$  or  $(k\mu)^{-1}$  depending on whether  $k$  is greater than or less than 1. Therefore, the picture conforms to first DFR case treated earlier.

In view of the discussion made above, it is quite clear that the estimator of  $\theta$  based either on a parametric or semi-parametric model may have serious bias if the model based assumptions are not all tenable, and in this respect the nonparametric estimator  $Z_n$  in (2.2) is highly robust. To study its properties, we assume that there is a unique  $x_0 (= x_0(F)) : 0 < x_0 < \infty$ , such that

$$\theta = \sup\{x[1-F(x)] ; x \in R^+\} = x_0[1-F(x_0)], \quad (2.33)$$

and moreover,  $F$  has a continuous density  $f$  in a neighborhood of  $x_0$  where  $f(x_0) > 0$ .

We define an empirical process  $W_n = \{W_n(t) : t \in [0,1]\}$  by letting

$$W_n(t) = n^{1/2} \{F_n(F^{-1}(t)) - t\}, \quad 0 \leq t \leq 1. \quad (2.34)$$

Then, it is well known that

$$W_n \text{ converges weakly to a Brownian Bridge } W^0, \text{ as } n \rightarrow \infty. \quad (2.35)$$

This ensures that on any compact interval  $[0,C]$ ,  $C < \infty$ ,

$$n^{1/2} \{x[1-F_n(x)] - x[1-F(x)] : x \in [0,C]\} \text{ weakly converges to a Gaussian random function.} \quad (2.36)$$

This weak convergence has been tacitly incorporated by Sen, Bhattacharyya and Suh (1973) in providing a vastly simplified proof of the following:

$$n^{1/2} (Z_n - \theta(F)) \overset{D}{\rightarrow} N(0, \gamma^2), \quad (2.37)$$

where

$$\gamma^2 = x_0^2 F(x_0)[1 - F(x_0)] = \theta^2(F) \{F(x_0)/\bar{F}(x_0)\}. \quad (2.38)$$

The asymptotic variance in (2.38) is a function of the underlying d.f.  $F$  (which is not properly known), and hence, to make full statistical use of (2.38), we need to estimate this in a nonparametric way. We also note that by definition,

$$Z_n = \sup\{x[1-F_n(x)] : x \in R^+\} \geq x_0[1-F_n(x_0)], \quad (2.39)$$

so that

$$E_F(Z_n) \geq x_0 E_F[1-F_n(x_0)] = x_0[1-F(x_0)] = \theta(F), \quad (2.40)$$

and hence,  $Z_n$  has a nonnegative bias. Further, note that [viz., Sen(1981,Ch.8)]

$$\{[x[1-F_n(x)] ; x \in R^+], n \geq 1\} \text{ is a reverse martingale process} \quad (2.41)$$

(as  $E[F_n | C_{n+1}] = F_{n+1}$  a.e.,  $n \geq 1$ ) and  $\sup\{\dots\}$  is a convex function. Thus,

$$\begin{aligned} E_F(Z_n | C_{n+1}) &\geq \sup\{E[x[1-F_n(x)] : x \in R^+] | C_{n+1}\} \\ &= \sup\{x[1-F_{n+1}(x)] : x \in R^+\} = Z_{n+1} \text{ a.e., } n \geq 1, \end{aligned} \quad (2.42)$$

so that  $\{Z_n, n \geq 1\}$  is a nonnegative reverse sub-martingale. Also, note that

$$\begin{aligned} Z_n &= \max\{(n-i+1)X_{n:i}/n : 1 \leq i \leq n\} \leq n^{-1} \max\{(\sum_{j=i}^n X_{n:j}) : 1 \leq i \leq n\} \\ &\leq n^{-1} \sum_{i=1}^n X_{n:i} = n^{-1} \sum_{i=1}^n X_i = \bar{X}_n, \text{ say.} \end{aligned} \quad (2.43)$$

Therefore the Lebesgue Dominated Convergence Theorem insures that whenever  $EX < \infty$ ,  $Z_n$  is uniformly integrable. All these results in turn imply that

$$E(Z_n - \theta) = o(1/n^{1/2}), \text{ as } n \rightarrow \infty, \quad (2.44)$$

although this bias is not typically  $O(n^{-1})$ . This feature calls for some cautions in the unrestricted use of the delta-method of variance estimation. Basically, if  $\hat{x}_n$  (one of the  $X_{n:i}$ ) is the sample point at which  $Z_n$  is attained, we may set an estimator of  $\gamma^2$  as

$$\hat{\gamma}_n^2 = \hat{x}_n^2 F_n(\hat{x}_n)[1-F_n(\hat{x}_n)] \text{ and } Z_n = \hat{x}_n[1-F_n(\hat{x}_n)]. \quad (2.45)$$

But, because of the nonlinear nature of the function, it is likely to be biased (albeit of order  $o(n^{-1/3})$ ), depending on the order of convergence of  $\hat{x}_n$  to  $x_0$ . This bias picture has been thoroughly studied in Sen (1992), and hence, we omit the details. Basically, both jackknifing and bootstrapping work out well in this context. Although jackknifing makes some adjustment for bias (reduction), it is not as much effective as in the regular case where the bias is  $O(n^{-1})$ . The bootstrapping works out well, but is likely to have comparatively larger bias. Some modification of jackknifing works out even better.

### 3. Multi-component system ; in series

Consider a chain having  $k$  ( $\geq 1$ ) loops. The chain breaks down when at least one of the loops is broken. Thus, the breaking stress of the chain depends on the number of loops it has as well as their individual strengths. If we denote these individual failure times by  $X_1, \dots, X_k$  respectively, then we may set

$$Y = \min\{X_1, \dots, X_k\} \quad (3.1)$$

as the failure time for the chain. Often, it is plausible to make an assumption that the  $X_i$  are i.i.d.r.v. with a d.f.  $F$ , defined on  $R^+$ . In the parametric setup, the most common model rests on the assumption that  $F$  is the simple exponential model given by (2.4). Let us denote the d.f. of  $Y$  by  $G$  and let  $\bar{G}$  be the corresponding survival function. Then we may write

$$\bar{G}(y) = [\bar{F}(y)]^k, \text{ for every } y \in R^+. \quad (3.2)$$

We consider the functional (mean life time)

$$\theta(G) = \int_{R^+} \bar{G}(x) dx \quad \text{and} \quad \theta(F) = \int_{R^+} \bar{F}(x) dx, \quad (3.3)$$

and we want to draw statistical conclusions on such parameters. In the most simple parametric case of exponential law given by (2.4), we have

$$\theta(G) = (k\mu)^{-1} = k^{-1}\theta(F), \quad (3.4)$$

so that when an optimal estimator of  $\theta(F)$  is available, it can be incorporated to

deriving an optimal estimator of  $\theta(G)$ . Actually, for this simple exponential model,

$$\bar{F}(x) = \exp(-\mu x)I(x \in R^+) \text{ and } \bar{G}(x) = \exp(-k\mu x)I(x \in R^+), \quad (3.5)$$

so that for the estimation of  $\theta(G)$  or the d.f.  $G$ , it suffices to estimate  $\mu$  in an optimal manner. If there are  $N$  independent observations  $X_1, \dots, X_N$  on the individual life times ( $X$ ) and if we assume that  $X$  has the simple exponential law in (2.4), then an optimal estimator of  $\theta(F)$  is  $\bar{X}_N = N^{-1}(X_1 + \dots + X_N)$ , and hence, by (3.4), an optimal estimator of  $\theta(G)$  is

$$\hat{\theta}(G) = k^{-1} \bar{X}_N. \quad (3.6)$$

This simple prescription may not work out well for d.f.'s other than the exponential ones. As an illustration, we consider the mixture model:

$$\bar{F}(x) = \pi \exp(-\mu x) + (1-\pi) \exp(-c\mu x), \quad x \in R^+, \quad 0 \leq \pi \leq 1; \quad c \neq 1. \quad (3.7)$$

Then, we have

$$\bar{G}(x) = \exp(-k\mu x) [ \pi + (1-\pi) \exp(-(c-1)\mu x) ]^k, \quad x \in R^+, \quad (3.8)$$

The second factor on the right hand side of (3.8) is equal to 1 at  $x = 0$ , while it is greater ( or less ) than 1 for  $x > 0$  depending on whether  $c$  is less ( or greater ) than 1. Thus,  $\bar{G}(x)$  is dominated from above ( or below ) by the simple exponential law with parameter  $(k\mu)$  depending on whether  $c$  is  $<$  or  $>$  1. As a result,  $\theta(G)$  is  $>$  or  $<$   $(k\mu)^{-1}$  depending on  $c$  being  $<$  or  $>$  1. This explains the lack of robustness of the parametric estimator in (3.6) against possible departure from the assumed exponentiality. In this context of robustness study, it is not uncommon to assume only local departures from the assumed model. This, in the context of the mixture model, can be set by letting

$$\pi = 1 - \epsilon \quad \text{and} \quad c = 1 \pm \eta, \quad \epsilon, \eta \text{ both small (positive)}. \quad (3.9)$$

Then, from (3.8) and (3.9), we have

$$\begin{aligned} \bar{G}(x) &= \exp(-k\mu x) [ 1 - \epsilon(1 - \exp(\mp \eta\mu x)) ]^k \\ &= \exp(-k\mu x) [ 1 - (\pm \eta\mu x - \frac{1}{2}(\eta\mu x)^2 + \dots) ]^k \end{aligned} \quad (3.10)$$

so that the alteration to the simple exponential model will be dominated by the factor  $k\epsilon\eta$ . Thus, if  $k$  is not too small, even a small error contamination may cause a perceptible change in the value of  $\theta(G)$  or the d.f.  $G$  itself.

As a second example, consider the Weibull law in (2.23). Here, we have  $\bar{G}(x) = [\bar{F}(x)]^k$ , although the simple relationship between  $\theta(G)$  and  $\theta(F)$  in (3.4) may not hold when  $\gamma \neq 1$ ; this proportionality factor ( $k$ ) depends then on both  $k$  and the unknown shape parameter  $\gamma$ .

We may recall that by (3.2), on denoting by  $h_F(x)$  and  $h_G(x)$  the hazard functions corresponding to the d.f.  $F$  and  $G$  respectively, we have

$$h_G(x) = -(d/dx) \log \bar{G}(x) = -k(d/dx) \log \bar{F}(x) = k h_F(x), \quad x \in R^+, \quad (3.11)$$

so that for the corresponding integrated hazard functions, we have

$$H_G(x) = \int_0^x h_G(y) dy = k \int_0^x h_F(y) dy = k.H_F(x), \quad x \in R^+. \quad (3.12)$$

As a result, we have

$$\theta(G) = \int_0^\infty \exp(-H_G(x)) dx = \int_0^\infty \exp(-kH_F(x)) dx, \quad (3.13)$$

so that it suffices to estimate the integrated hazard rate  $H_F(x)$ . Now,  $H_F(x)$ ,  $x \in R^+$ , can be estimated in a semi-parametric or nonparametric manner. People who love to incorporate smoothing techniques would be delighted at this prospect. But, generally, such smooth estimators may eliminate the bumps to a greater extent at the cost of a possibly slower rate of convergence. In any case, in a completely nonparametric setup, we have a much simpler solution which we prescribe below.

Define a kernel of degree  $k$  by letting

$$\phi(x_1, \dots, x_k) = \min\{x_1, \dots, x_k\}. \quad (3.14)$$

Then, note that

$$E_F \phi(X_1, \dots, X_k) = \int_0^\infty \bar{G}(x) dx = \theta(G). \quad (3.15)$$

Consider then the Hoeffding (1948) U-statistic corresponding to the kernel in (3.14) when  $N$  is greater than or equal to  $k$ . This is given by

$$U_N = \binom{N}{k}^{-1} \sum_{\{1 \leq i_1 < \dots < i_k \leq N\}} \phi(X_{i_1}, \dots, X_{i_k}). \quad (3.16)$$

As in Section 2, if we denote the order statistics by  $X_{N;1} \leq \dots \leq X_{N:N}$ , then some simple manipulations lead us to

$$U_N = \binom{N}{k}^{-1} \sum_{i=1}^N \binom{N-i}{k-1} X_{N;i}, \quad (3.17)$$

so that  $U_N$  is an L-estimator of  $\theta(G)$  and is unbiased too. From the general expression for the variance of U-statistics given by Hoeffding (1948), we have

$$\text{Var}(U_N) = E_F [U_N - \theta(G)]^2 = \binom{N}{k}^{-1} \sum_{c=1}^k \binom{k}{c} \binom{N-k}{k-c} \zeta_c, \quad (3.18)$$

where

$$\zeta_c = \text{Cov}\{\phi(X_1, \dots, X_k), \phi(X_{k-c+1}, \dots, X_{k+c})\}, c = 0, 1, \dots, k, \quad (3.19)$$

and  $\zeta_0 = 0$ . Note that all these  $\zeta_c$  are themselves estimable parameters of degree less than or equal to  $2k$ . Moreover,  $U_N$  is an optimal estimator of  $\theta(G)$  in a non-parametric sense [discussed in detail in Sen (1981, Ch.3)]. Also, we note that if  $N$  is not too small, we may write (3.18) as

$$\text{Var}(U_N) = k^2 N^{-1} \zeta_1 + O(N^{-2}), \quad (3.20)$$

so that invoking the Hoeffding(1948) asymptotic normality of U-statistics, we obtain that for large  $N$ ,

$$N^{1/2} (U_N - \theta(G)) / k \zeta_1^{1/2} \xrightarrow{D} N(0, 1), \quad (3.21)$$

so that to draw statistical conclusions on the parameter  $\theta(G)$ , all we need is to provide a consistent estimator of  $\zeta_1$ , and this can be done in a variety of ways. We may also note in this context that by virtue of the reverse martingale property of U-statistics,  $U_N$  converges almost surely (a.s.) to  $\theta(G)$  as  $N \rightarrow \infty$ . While it is

possible to estimate unbiasedly each of the  $\zeta_c$ ,  $c \leq k$ , when  $N \geq 2k$ , so that an unbiased estimator of the variance in (3.18) may be obtained in terms of a set of U-statistics, the task becomes quite laborious as  $N$  increases. For this reason, we do not recommend this unbiased estimation of variance of  $U_N$ . We may proceed as in Sen (1960) and construct a variance estimator of  $U_N$  (which turns out to be the same as the jackknife variance estimator), and using latter results of Sen (1977) we can establish the a.s. convergence of such an estimator. Because of the fact that  $U_N$  in (3.16) is an L-estimator and is a U-statistic too, it satisfies the reverse martingale property mentioned before, and hence, we may obtain a more simplified version of this variance estimator by using Theorem 7.4.2 of Sen (1981). Looking back at (3.17), we may introduce the notations :

$$c_{N,i} = k \binom{N-1}{k-1}^{-1} \binom{N-i}{k-1}, \text{ for } i = 1, \dots, N, \text{ and } N \geq k. \quad (3.22)$$

Then, we have

$$U_N = N^{-1} \sum_{i=1}^N c_{N,i} X_{N:i} = \int_0^{\infty} J_N(F_N(x)) x dF_N(x), \quad (3.23)$$

where

$$J_N(u) = c_{N,[Nu]}, \text{ for } 0 < u < 1. \quad (3.24)$$

We let

$$D_{Ni} = X_{N:i+1} - X_{N:i}, \text{ } i=1, \dots, N-1, \quad (3.25)$$

(these are termed the spacings), and let

$$V_N = N(N-1) \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} c_{N-1,i} c_{N-1,j} [(i \wedge j)N - ij] D_{N,i} D_{N,j} / N^2. \quad (3.26)$$

It may then be noted that  $V_N$  is a variant of the usual jackknife variance estimator and in this specific problem,

$$V_N \rightarrow k^2 \zeta_1^2 \text{ a.s., as } N \rightarrow \infty. \quad (3.27)$$

By the Slutsky Theorem, (3.21) and (3.27), we have

$$N^{1/2} (U_N - \theta(G)) / V_N^{1/2} \xrightarrow{\mathcal{D}} N(0, 1), \text{ as } N \rightarrow \infty, \quad (3.28)$$

so that we may use (3.28) conveniently in setting a confidence interval for  $\theta(G)$  or for testing a suitable null hypothesis on  $\theta(G)$ .

We may recall that the d.f.  $G$  is related to the d.f.  $F$  by (3.2), and in a non-parametric setup, as nothing is known precisely about  $F$ , it is also of interest to estimate  $G$  in a nonparametric manner. Towards this, we define

$$G_{U,N}(x) = \binom{N}{k}^{-1} \sum_{\{1 \leq i_1 < \dots < i_k \leq N\}} I\{\phi(X_{i_1}, \dots, X_{i_k}) \leq x\}, \text{ } x \in R^+, \quad (3.29)$$

Then,  $G_{U,N} = \{G_{U,N}(x), x \in R^+\}$  is a U-process, and  $G_{U,N}$  unbiasedly estimates  $G$ . Convergence properties of such U-processes have been studied by a host of research workers [we may refer to Sen (1988b) where other references are cited]. Basically, we have the following:

(i) Glivenko-Cantelli type Convergence: As  $N \rightarrow \infty$ ,

$$\|G_{U,N} - G\| = \sup\{|G_{U,N}(x) - G(x)| : x \in R^+\} \rightarrow 0 \text{ a.s.} \quad (3.30)$$

(ii) Weak Invariance Principle: As  $N$  increases,

$$N^{1/2}(G_{U,N} - G) \text{ converges weakly to a Gaussian function.} \quad (3.31)$$

(iii) Hoeffding-Decomposition/Projection: There exists a function  $\psi_x(\cdot; F)$ , such that  $E_F \psi_x(X; F) = 0$ ,  $E_F \psi_x^2(X; F) < \infty$ , and

$$G_{U,N}(x) - G(x) = kN^{-1} \sum_{i=1}^N \psi_x(X_i, F) + R_N(x; F), \quad (3.32)$$

where

$$\|R_N(F)\| = \sup\{|R_N(x; F)| : x \in R^+\} = o_p(N^{-1/2}). \quad (3.33)$$

In (3.33),  $o_p(\cdot)$  may be replaced by 2nd mean convergence also. Further,  $\psi_x(\cdot; F)$  can easily be computed from the kernel in (3.29) by taking the first order projection. Actually, the weak invariance principle is a direct consequence of (3.32) and the smoothness properties of the conditional expectation  $\psi_x(\cdot; F)$  (as a function of  $x \in R^+$ , given  $F$ ). Moreover, the decomposition in (3.32) (also known as the first order asymptotic representation) provides the justification for the adaptability of the usual resampling methods for this process too; we may refer to the last section of Sen (1988b) where these are discussed in detail.

#### 4. Multi-component system - in parallel

A parallel structure functions if and only if at least one of the components functions. Hence, the life time ( $Y$ ) of the system is characterized by

$$Y = \max\{X_1, \dots, X_k\}, \quad (4.1)$$

where the  $X_i$  are the life times of the individual components. We assume that the  $X_i$  are i.i.d.r.v.'s with a d.f  $F$ , defined on  $R^+$ , and we denote the d.f. of  $Y$  by  $H$ . Then, we have by (4.1),

$$H(y) = [F(y)]^k \text{ or } \bar{H}(y) = 1 - [1 - \bar{F}(y)]^k. \quad (4.2)$$

By the usual binomial expansion, we have from (4.2),

$$\bar{H}(y) = \sum_{j=1}^k (-1)^{j-1} \binom{k}{j} [\bar{F}(y)]^j. \quad (4.3)$$

We define  $G$  and  $\theta(G)$  as in (3.2) and (3.3), and to be more specific, we attach the subscript  $k$  to  $G$ . Then, we have from (4.3),

$$\bar{H}(y) = \sum_{i=1}^k (-1)^{i-1} \binom{k}{i} \bar{G}_i(y), \quad (4.4)$$

so that

$$\theta(H) = \int_{R^+} \bar{H}(y) dy = \sum_{j=1}^k (-1)^{j-1} \binom{k}{j} \theta(G_j). \quad (4.5)$$

Thus, we may proceed as in Section 3 for the estimation of the individual  $\theta(G_j)$  and incorporate the linearity in (4.5) to provide an estimator of  $\theta(H)$ . In a similar manner, in (3.14), we replace the kernel  $\phi(X_1, \dots, X_k)$  by

$$\phi(x_1, \dots, x_k) = \max\{x_1, \dots, x_k\}, \quad (4.6)$$

and note that

$$E_F \phi(X_1, \dots, X_k) = \int_{R^+} \bar{H}(y) dy = \theta(H). \quad (4.7)$$

This leads us to consider the U-statistic

$$U_N = \binom{N}{k}^{-1} \sum_{\{1 \leq i_1 < \dots < i_k \leq N\}} \phi(X_{i_1}, \dots, X_{i_k}) \quad (4.8)$$

which is an unbiased, symmetric and optimal nonparametric estimator of  $\theta(H)$ . Also proceeding as in (3.17), we obtain that (4.8) simplifies to

$$U_N = \binom{N}{k}^{-1} \sum_{i=1}^N \binom{i-1}{k-1} X_{N:i}, \quad (4.9)$$

so that  $U_N$  is again an L-estimator with smooth weights. As such, we may adopt (3.18) through (3.21) with a definition of the  $\zeta_c$ ,  $c \leq k$  as in (3.19) but for the kernel in (4.6). In (3.22) if we replace the  $\binom{N-i}{k-1}$  by  $\binom{i-1}{k-1}$  and define the  $J_N(u)$  accordingly as in (3.24), then (3.26) as modified also applies to the variance estimation of  $U_N$ , so that (3.27)-(3.28) also hold for  $U_N$  in (4.8). In a similar manner, in (3.19), we replace the kernel  $I\{\phi(X_1, \dots, X_k) \leq x\}$  by  $I\{\max(X_1, \dots, X_k) \leq x\}$ , and denote the resulting U-process by  $H_{U,N}(x)$ ,  $x \in R^+$ . Then (3.30) through (3.33) hold with  $G$  replaced by  $H$ . As such, the prospects for adaptation of resampling plans (i.e., jackknifing and bootstrapping) are excellent, and much of the details discussed in Sen (1988b) remains pertinent here.

It may be remarked that for the particular case of  $F$  being exponential,  $G$  in Section 3 also turns out to be exponential. Even in this simple parametric case, for  $k \geq 2$ ,  $H$  is not an exponential d.f., and hence, some of the simplicities in Section 3 are not applicable here. We may, however, note that

$$\begin{aligned} h_H(y) &= h(y)/\bar{H}(y) \\ &= \{ \sum_{j=1}^k (-1)^{j-1} \binom{k}{j} [\bar{F}(y)]^{j-1} j f(y) \} / \bar{H}(y) \\ &= k f(y) \{ \sum_{j=1}^k (-1)^{j-1} \binom{k-1}{j-1} [\bar{F}(y)]^{j-1} \} / \bar{H}(y) \\ &= k f(y) \bar{H}_{k-1}(y) / \bar{H}_k(y), \end{aligned} \quad (4.10)$$

where  $\bar{H} = \bar{H}_k$  and  $\bar{H}_{k-1}$  relates to the survival function of  $\max(X_1, \dots, X_k)$ . Thus, we obtain the chain relation for the hazard functions :

$$h_H(x) = k h_F(x) \{ \bar{F}(x) \bar{H}_{k-1}(x) / \bar{H}_k(x) \}. \quad (4.11)$$

Note that

$$\begin{aligned} \bar{H}_{k-1}(x) &= P\{ \max(X_1, \dots, X_{k-1}) > x \} \\ &\leq P\{ \max(X_1, \dots, X_k) > x \} = \bar{H}_k(x), \quad x \in R^+, \end{aligned} \quad (4.12)$$

while  $\bar{F}(x)$  is  $\leq 1$ , for all  $x$ . Hence, from (4.11)- (4.12), we obtain that

$$h_H(x) \leq k \cdot h_F(x) = h_G(x), \quad x \in R^+. \quad (4.13)$$

Both (4.10) and (4.11) provide useful hints in providing smooth and efficient estimators of the hazard rate  $h_H(x)$ . For example  $H_{U,N}(x)$ , as described before (4.10), and denote by  $H_{U,N}^{(k)}$  to make its dependence on  $k$  explicit, is a consistent and optimal nonparametric estimator of  $H_k(x)$ , for every  $k \geq 1$ . This, in turn, provide a smooth nonparametric estimator of  $\bar{H}_{k-1}(x)/\bar{H}_k(x)$ . Hence, the kernel/nearest neighbor method of smooth estimation of the density  $f(y)$  when worked in conjunction with the smooth estimator of  $\bar{H}_{k-1}(x)/\bar{H}_k(x)$  provides a smooth estimator of  $h_H(x)$ . Note that although the raw estimates  $H_{U,N}$  are step functions, they may be very effectively smoothed out by using any smoothing formula. Similarly, if an appropriate estimator of  $h_F(x)$  is available, one may use (4.11) along with some estimators of  $\bar{F}(x)$ ,  $\bar{H}_{k-1}(x)$  and  $\bar{H}_k(x)$  to estimate  $h_H(x)$ . In this context, semi-parametric models may be adopted to estimate the hazard rate  $h_F(\cdot)$  and use (4.11) to produce a parallel estimator of  $h_H(\cdot)$ .

### 5. System with a spare and repair facility

In a single-unit system supported by a single spare and a repair facility, when the operating unit fails, it is instantaneously replaced by the spare while it is being despatched to the repair facility. Upon repair, it is sent to the spare box. The system fails when an operating unit fails and no spare is there to replace it. This occurs when the repairing time ( $Y$ ) is more than the life time ( $X$ ) of the operating unit. Let  $X$  have a distribution function  $F$  (survival function  $\bar{F}$ ) defined on  $R^+$ , and let  $G$  (and  $\bar{G}$ ) be the d.f. (and survival function) of the r.v.  $Y$ . Both  $F$  and  $G$  (and hence,  $\bar{F}$  and  $\bar{G}$ ) are defined on  $R^+$ . Since, to start with, we have a spare whose life time has the d.f.  $F$ , if we denote by  $N$  the number of operating units failure culminating in a system failure, then  $N$  is a positive integer valued r.v. with

$$P\{N = k + 1\} = \alpha^{k-1}(1 - \alpha), \text{ for } k = 1, 2, 3, \dots \quad (5.1)$$

and

$$\alpha = P\{X > Y\} = \int_{R^+} G(x)dF(x). \quad (5.2)$$

Let us denote the mean of the d.f.  $F$  and  $G$  by  $\theta(F)$  and  $\theta(G)$  respectively. Then the mean time until the first system failure is given by

$$E(T_1) = [(2 - \alpha)/(1 - \alpha)]\theta(F), \quad (5.3)$$

while the mean time, measuring from a regeneration point, is

$$E(T) = E(T_1) - \theta(F) = \theta(F)/(1 - \alpha). \quad (5.4)$$

Note that the system is null recurrent when  $\alpha = 1$ , but for  $\alpha < 1$ , both (5.3) and (5.4) are finite positive quantities. In such a case, there is a possibility of the system being down, and the mean ED of system downtimes is

$$ED = \int_0^\infty \int_0^\infty \{\bar{G}(x+t)/\bar{G}(x)\} dF(x)dt. \quad (5.5)$$

Usually, it is assumed that (i) the repair of a failed unit restores it to its new condition, (ii) the original unit and the spare both have the same d.f.  $F$ , and (iii) the life time  $X$  and repair time  $Y$  are mutually independent [ viz., Barlow and Proschan (1975, pp.202-205 )]. The limiting average availability (i.e., the limiting expected proportion of system up times ) is defined by

$$A_{FG} = ET / \{ ET + ED \} . \quad (5.6)$$

Recall that  $\alpha$ , defined by (5.2), is a functional of  $F$  and  $G$  and so is  $ED$ . Thus,  $A_{FG}$  is a functional of  $(F,G)$ . As regards the definition of  $A_{FG}$ , the stochastic independence of  $X$  and  $Y$  ( as assumed in (5.1)-(5.2)) is not that crucial, and Sen and Bhattacharjee (1986) were able to relax this assumption to a greater extent. For simplicity of presentation, here, we take  $X$  and  $Y$  to be mutually independent. By virtue of (5.4), (5.5) and (5.6), we have

$$A_{FG} = \theta(F) / \{ \theta(F) + (1 - \alpha)ED \} . \quad (5.7)$$

Since  $ED$  is a nonlinear functional of  $(F,G)$ ,  $A_{FG}$  as defined in (5.7) is also so.

Let us first consider the simplest parametric model where both  $F$  and  $G$  are simple exponential d.f.'s. We let

$$F(x) = \exp(-x/\theta_1), \text{ and } \bar{G}(x) = \exp(-x/\theta_2), \quad x \in R^+, \quad (5.8)$$

so that

$$\theta_1 = \theta(F), \quad \theta_2 = \theta(G) \text{ and } \alpha = \theta_1 / (\theta_1 + \theta_2). \quad (5.9)$$

In this special case,

$$\bar{G}(x+t)/\bar{G}(x) = \exp(-t/\theta_2) = \bar{G}(t), \text{ for every } x, t \in R^+, \quad (5.10)$$

so that

$$ED = \int_0^\infty \int_0^\infty \bar{G}(t) dF(x) dt = \int_0^\infty \bar{G}(t) dt = \theta_2 . \quad (5.11)$$

As a result, we have

$$A_{FG} = [ \theta(F) \{ \theta(F) + \theta(G) \} ] / [ \theta^2(F) + \theta(F)\theta(G) + \theta^2(G) ] . \quad (5.12)$$

If we define the ratio of the means by

$$\rho = \rho(F,G) = \theta(G)/\theta(F), \quad (5.13)$$

then (5.12) reduces to

$$A_{FG} = (1 + \rho) / (1 + \rho + \rho^2) . \quad (5.14)$$

Thus, for the simple exponential model,  $A_{FG}$  is a monotone (nonincreasing) function of  $\rho$ ; it is equal to 1 when  $\rho = 0$  and it converges to 0 as  $\rho \rightarrow \infty$ . In this case, it suffices to estimate  $\rho$ , and hence, standard parametric theory is readily available. Let us next consider some semi-parametric models and examine the robustness aspects of this simple parametric estimator.

As a natural extension of the exponential model, we consider the proportional hazard model (PHM) where we set

$$\bar{G}(x) = [\bar{F}(x)]^c, \text{ for some } c > 0. \quad (5.15)$$

Then, we have

$$\begin{aligned} \alpha &= \int_0^\infty G(x)dF(x) = 1 - \int_0^\infty \bar{G}(x)dF(x) \\ &= 1 - \int_0^\infty [\bar{F}(x)]^c dF(x) = 1 - \int_0^1 (1-u)^c du \\ &= 1 - (1+c)^{-1} = c/(1+c). \end{aligned} \quad (5.16)$$

On the other hand,

$$\theta(G) = \int_0^\infty \bar{G}(x)dx = \int_0^\infty [\bar{F}(x)]^c dx = \theta(F, c) \quad (5.17)$$

depends on  $c$  as well as  $F$  (in an involved way), so that the ratio  $\theta(G)/\theta(F)$  may not be a sole function of  $c$ . Similarly,

$$ED = \int_0^\infty \int_0^\infty \{\bar{F}(x+t)/\bar{F}(x)\}^c dF(x)dt = \Delta(F, c) \quad (5.18)$$

may depend on  $c$  as well as  $F$  (in an involved way). Thus, inspite of the simplicity of the expression for  $\alpha$  in (5.16), in general, for a PHM,  $A_{FG}$  may not have a very simple form (comparable to (5.14)).

We consider now some special class of d.f.'s for which we may have some useful bounds on  $A_{FG}$ . First consider the NBU/NWU class. A d.f.  $F$  is NBU (new better than used) if  $\bar{F}(x+y) \leq \bar{F}(x)\bar{F}(y)$ , for all  $x, y \geq 0$ , while it is NWU (new worse than used) if  $\bar{F}(x+y) \geq \bar{F}(x)\bar{F}(y)$ , for all  $x, y \geq 0$ . Note that the common boundary of the NBU and NWU families is the exponential d.f. [where by (5.10) the strict equality sign holds everywhere]. We may also note that

$$IFR \Rightarrow NBU \quad \text{and} \quad DFR \Rightarrow NWU. \quad (5.19)$$

Thus, for  $F$  NBU, we have under the PHM in (5.15) [ by (5.18)]

$$\begin{aligned} ED &\leq \int_0^\infty \int_0^\infty \{\bar{F}(t)\}^c dF(x)dt = \int_0^\infty \{\bar{F}(t)\}^c dt \\ &= \int_0^\infty \bar{G}(t) dt = \theta(G). \end{aligned} \quad (5.20)$$

Similarly,

$$ED \geq \theta(G) \text{ whenever (5.15) holds and } F \text{ is NWU.} \quad (5.21)$$

Therefore, from (5.7), (5.20) and (5.21), we obtain that under the PHM in (5.15),

$$\begin{aligned} A_{FG} \text{ is } \begin{cases} \geq & \theta(F)/\{\theta(F) + (1-\alpha)\theta(G)\} = (1+c)\theta(F)/\{(1+c)\theta(F) + \theta(G)\}, \\ \leq & \end{cases} \\ \text{according as } F \text{ is NBU/Exponential/NWU.} \end{aligned} \quad (5.22)$$

Moreover, we may note that

$$\begin{aligned} \theta(G)/\theta(F) &= \{\int_0^\infty \{\bar{F}(x)\}^c dx\} / \{\int_0^\infty \bar{F}(x)dx\} \\ &= \{\int_0^\infty \{\bar{F}(x)\}^{c-1} \cdot \bar{F}(x)dx\} / \{\int_0^\infty \bar{F}(x)dx\} \\ &\begin{cases} \leq & 1 \text{ according as } c \text{ is } \geq 1. \\ \geq & \end{cases} \end{aligned} \quad (5.23)$$

Substitution of (5.23) in (5.22) may lead to some crude bounds for  $A_{FG}$ . For example, for  $F$  NBU and for  $c > 1$  (in (5.15)), we have  $A_{FG} \geq (1+c)/(2+c)$ , where for large  $c$  the bound becomes a better approximation. However, this still depends on the nature of  $F$ . For  $F$  exponential,  $c=1$ , this is exact.

It is clear from the above discussion that even one assumes a PHM for (F,G), the functional  $A_{FG}$  may depend on the subclass of NBU/NWU etc to which F may belong. Thus, an incorrect assumption of exponentiality of F and G may lead to serious bias for the estimator of  $A_{FG}$  based on (5.14) with  $\rho$  estimated by the ratio of the sample means of the repair times and operating times. It seems therefore quite natural to incorporate nonparametric methods in the estimation of  $A_{FG}$  and emphasize on their robustness. Proceeding as in Sen and Bhattacharjee (1986), we may write

$$ED = (1 - \alpha)^{-1} E[ (Y - X) I(Y > X) ], \quad (5.24)$$

so that by (5.4), (5.6) and (5.24), we have

$$\begin{aligned} A_{FG} &= \theta(F) / \{ \theta(F) + E[(Y-X)I(Y > X)] \} \\ &= \{ EX \} / \{ EX + E[(Y-X)I(Y > X)] \} \\ &= \{ EX \} / \{ E[XI(X > Y)] + E[YI(Y > X)] \} \\ &= \{ EX \} / \{ E[X \vee Y] \}, \end{aligned} \quad (5.25)$$

where  $(a \vee b) = \max(a, b)$ . This definition even permits one to replace the assumed independence of X and Y by that of independence of the stochastic vectors  $(X_i, Y_i)$  for different  $i (\geq 1)$ , yet allowing possible stochastic dependence between  $X_i$  and  $Y_i$ . Note that if we assume that X and Y are independent then we have

$$\begin{aligned} E[X \vee Y] &= \int_0^\infty P\{[X \vee Y] > t\} dt \\ &= \int_0^\infty \{1 - F(t)G(t)\} dt \\ &= \theta^*(F, G), \text{ say.} \end{aligned} \quad (5.26)$$

Thus,

$$A_{FG} = \theta(F) / \theta^*(F, G), \quad (5.27)$$

and hence, one may use the classical U-statistics corresponding to the two estimable parameters in the numerator and denominator of (5.27), and denoting them by  $U_{n1}$  and  $U_{n2}$ , respectively, one may consider the estimator

$$U_{n,FG} = U_{n1} / U_{n2}. \quad (5.28)$$

$U_{n,FG}$  may not be a strictly unbiased estimator of  $A_{FG}$ , but it is bounded from above by 1 (and from below by 0), and hence, it can be shown that

$$E[U_{n,FG}] = A_{FG} + n^{-1} \beta_{FG} + O(n^{-2}), \quad (5.29)$$

where

$$\beta_{FG} = A_{FG}^3 [ \theta(F) ]^{-2} \{ n \text{Var}(U_{n2}) - n \text{Cov}(U_{n1}, U_{n2}) \}. \quad (5.30)$$

This familiar form [in (5.29)] clearly suggests that by the classical jackknifing method, this leading term in the bias can be eliminated, and the resulting jackknife version of  $U_{n,FG}$  will have bias  $O(n^{-2})$ , but the same mean square error

(upto the order  $n^{-2}$ ). This procedure works out well even if the  $(X_i, Y_i)$  are not mutually independent [ viz., Sen (1986)]. We may also use the definition (5.7) and express  $A_{FG}$  as a function of  $\theta(F)$ ,  $\alpha$  and ED. If we denote by  $F_n$  and  $G_n$  the sample (empirical) d.f. of the  $X_i$  and  $Y_i$  respectively, then we may use the von Mises' functionals

$$\theta(F_n) = \int_0^\infty x dF_n(x) = n^{-1} \sum_{i=1}^n X_i = \bar{X}_n, \text{ say,} \quad (5.31)$$

$$\hat{\alpha}_n = \int_0^\infty G_n(x) dF_n(x) \quad (2\text{-sample Wilcoxon statistic}) \quad (5.32)$$

and

$$\hat{\xi}_n = \int_0^\infty \int_0^\infty \{ \bar{G}_n(x+t) / \bar{G}_n(x) \} dF_n(x) dt. \quad (5.33)$$

We denote the order statistics corresponding to the life time variable  $X$  by  $X_{n:1}, \dots, X_{n:n}$  and for the repair time variable  $Y$  by  $Y_{n:1}, \dots, Y_{n:n}$  respectively. Then, by partial integration in (5.33), we have

$$\hat{\xi}_n = n^{-1} \sum_{i=1}^n n_i^{-1} \sum_{j=1}^n (Y_{n:j} - X_{n:i}) I(Y_{n:j} > X_{n:i}), \quad (5.34)$$

where

$$n_i = \sum_{j=1}^n I(Y_{n:j} > X_{n:i}), \text{ for } i = 1, \dots, n. \quad (5.35)$$

As such, if we use the crude estimator

$$\hat{A}_{FG, n} = \bar{X}_n / \{ \bar{X}_n + (1 - \hat{\alpha}_n) \hat{\xi}_n \}, \quad (5.36)$$

we may incorporate jackknifing to eliminate the leading bias term and obtain the asymptotic normality of the jackknifed estimator to draw conclusions on the parameter  $A_{FG}$  in a nonparametric way. In all these cases, bootstrapping may also be used (instead of jackknifing). But, as we have here a nonlinear functional, jackknifing may have some advantage regarding the elimination of the leading bias term.

Basically, by (5.27),  $A_{FG}$  is the ratio of two estimable parameters, and hence, the usual ratio-type estimators may be employed to estimate it conveniently, and then jackknifing can take care of the reduction of the bias term due this non-linearity. Let us formulate a related functional for which an optimal unbiased nonparametric estimator can readily be obtained, and examine the closeness of the two parameters. Since  $X$  and  $Y$  are nonnegative r.v.'s, in (5.25), we replace the ratio of the two expectations by the expectation of the ratio and define

$$A_{FG}^0 = E[ X / \{ X \vee Y \} ]. \quad (5.37)$$

We rewrite this as

$$A_{FG}^0 = E[ I(X > Y) ] + E[ (X/Y) I(Y > X) ] = \alpha + E[ \frac{X}{Y} I(Y > X) ]. \quad (5.38)$$

By partial integration, we have

$$A_{FG}^0 = 1 - \int_0^\infty x \int_x^\infty y^{-2} \bar{G}(y) dy dF(x). \quad (5.39)$$

Recall that  $X^{1/2}I(Y > X) = Y^{1/2}I(Y > X) \cdot (X/Y)^{1/2}I(Y > X)$ , so that by the Cauchy-Schwarz inequality, we have

$$(E[X^{1/2}I(Y > X)])^2 \leq [E(YI(Y > X))][E((X/Y)I(Y > X))], \quad (5.40)$$

and hence,

$$\begin{aligned} E[(X/Y)I(Y > X)] &\geq (E[X^{1/2}I(Y > X)])^2 / E[YI(Y > X)] \\ &= E[XI(Y > X)] / E[YI(Y > X)] - \text{Var}[X^{1/2}I(Y > X)] / E[YI(Y > X)]. \end{aligned} \quad (5.41)$$

Also, note that

$$E[XI(Y > X)] = \int_0^\infty x\bar{G}(x)dF(x); \quad (5.42)$$

$$E[YI(Y > X)] = \int_0^\infty yF(y)dG(y) = \theta(G) - \int_0^\infty yF(y)dG(y). \quad (5.43)$$

We may use (5.38), (5.41), (5.42) and (5.43) to provide a lower bound for  $A_{FG}^0$ . A somewhat more simple (albeit crude) lower bound can be obtained by letting  $Z = X \vee Y$  and noting that by the Cauchy-Schwarz inequality

$$\begin{aligned} A_{FG}^0 &= E(X/Z) \geq [E(X^{1/2})]^2 / E(Z) = E(X) / E(Z) - [\text{Var}(X^{1/2})] / E(Z) \\ &= A_{FG} - A_{FG} [\text{Var}(X^{1/2})] / E(X) \\ &= A_{FG} [1 - (\text{coefficient of variation of } X^{1/2})(E[X^{1/2}])^2 / EX]. \end{aligned} \quad (5.44)$$

Since  $(E[X^{1/2}])^2 / EX$  is less than or equal to 1, whenever the coefficient of variation of  $X^{1/2}$  is small, the second term on the right hand side of (5.44) is small, and the lower bound is close to  $A_{FG}$  itself. This is typically the case when  $EX$ , the expected life time of the operating unit is not small (so that  $\alpha$  is close to 1) while the repair time distribution  $G$  dominates  $F$  (i.e.,  $G(x) \geq F(x)$  for all  $x$ ). Next note that for every  $z \geq x \geq 0$ ,

$$\begin{aligned} P\{X > x, Z > z\} &= P\{X > x, X > z, Y \leq X\} + P\{X > x, Y > z, Y > X\} \\ &= \int_z^\infty G(x)dF(x) + \bar{F}(x)\bar{G}(z) - \int_z^\infty \bar{F}(y)dG(y) \\ &= \int_z^\infty G(x)dF(x) + \bar{F}(x)\bar{G}(z) + \bar{F}(z)G(z) - \int_z^\infty G(y)dF(y) \\ &= \bar{F}(x) - \bar{F}(x)G(z) + \bar{F}(z)G(z) \\ &= \bar{F}(x)[1 - F(z)G(z)] + \bar{F}(x)[F(z)G(z) - G(z)] + \bar{F}(z)G(z) \\ &= \bar{F}(x)[1 - F(z)G(z)] + G(z)\bar{F}(z)F(x) \\ &\geq \bar{F}(x)[1 - F(z)G(z)] = P\{X > x, Z > z\}. \end{aligned} \quad (5.45)$$

Thus,  $X$  and  $Z$  are positively associated (and nonnegative) r.v.'s, so that  $X$  and  $Z^{-1}$  are negatively associated. Therefore,

$$A_{FG}^0 = E(XZ^{-1}) \leq E(X)E(Z^{-1}) = A_{FG} \cdot [E(Z)E(Z^{-1})], \quad (5.46)$$

This simple upper bound works out well when  $E(Z)E(Z^{-1})$  is close to 1. Since, by (5.25) and (5.39), 1 is an upper bound for  $A_{FG}^0$ , we may use (5.39) otherwise.

Actually, for  $\alpha$  close to 1, we may consider some simple asymptotics too. Let us denote by  $W = (Z-X)/Z$ , so that we have

$$A_{FG}^0 = 1 - EW \quad \text{and} \quad A_{FG} = 1 - E(ZW)/E(Z). \quad (5.47)$$

We rewrite

$$\begin{aligned} A_{FG} &= 1 - EW - E[(Z-EZ)W]/E(Z) \\ &= A_{FG}^0 - \text{Cov}(Z, W) / E(Z). \end{aligned} \quad (5.48)$$

Next, we note that  $W = 0$  whenever  $Z = X$  (which happens with probability  $\alpha$ ), while, for  $Y > X$ ,  $W = (Y-X)/Y = 1 - X/Y$  is  $\uparrow$  in  $Y$  (for a given  $X$ ). Thus, conditionally,  $X = x$ ,  $W$  and  $Z$  are concordant. Hence, we may write

$$\begin{aligned} \text{Cov}(Z, W) &= E[\text{Cov}(Z, W | X)] + \text{Cov}[E(Z|X), E(W|X)] \\ &\geq \text{Cov}[E(Z|X), E(W|X)], \end{aligned} \quad (5.49)$$

so that we have from (5.48) and (5.49) that

$$A_{FG}^0 \geq A_{FG} + [\text{Cov}(E(Z|X), E(W|X))]/E(Z). \quad (5.50)$$

However,

$$E(Z | X=x) = xG(x) + \int_x^\infty yF(y)dG(y) \quad \text{is } \uparrow \text{ in } x, \quad (5.51)$$

$$E(W | X=x) = \int_x^\infty (1 - x/y)g(y)dy \quad \text{is } \downarrow \text{ in } x, \quad (5.52)$$

so that

$$E(Z | X) \quad \text{and} \quad E(W | X) \quad \text{are discordant,} \quad (5.53)$$

and hence, the covariance term in (5.50) is non-positive. Thus, we are not in a position to claim that  $A_{FG}^0$  is bounded from below or above by  $A_{FG}$ . On the other hand, if  $EZ$  is large and  $(Z-X)$  is 0 with probability  $\alpha$  close to one, the covariance term in (5.48) will be small compared to  $EZ$ , and hence  $A_{FG}$  and  $A_{FG}^0$  will be close to each other. Recall that

$$|\text{Cov}(Z, W)/EZ| \leq \{(\text{Coff. var. of } Z)(\text{var. of } W)\}^{1/2}, \quad (5.54)$$

so that whenever the variance of  $W$  is small or the coefficient of variation of  $Z$  is small, the second term on the right hand side of (5.48) will be small.

Given this affinity of  $A_{FG}$  and  $A_{FG}^0$ , we are in a position to advocate the use of  $A_{FG}^0$  for which we have a simple U-statistic estimator :

$$U_n^0 = n^{-1} \sum_{i=1}^n (X_i / (X_i \vee Y_i)), \quad (5.55)$$

in the matched sample case, and

$$U_n^* = n^{-2} \sum_{i=1}^n \sum_{j=1}^n (X_i / (X_i \vee Y_j)), \quad (5.56)$$

in the independent samples case. In either case, the mean, variance etc, of the statistic can be computed as in before and its asymptotic normality follows from standard results on U-statistics. Further, jackknifing and bootstrapping can be used as in before for the variance estimation purpose. Since these U-statistics are unbiased estimators of  $A_{FG}^0$ , there is no need to reduce the bias by jackknifing.

Throughout this section and earlier, we have considered the uncensored case where the life time  $X$  is an observable r.v. Often, in practice, censoring may arise due to a variety of causes. In reliability setting, for example, age or block replacement policies may induce censoring. In the literature, various types of censoring have been discussed in various contexts. In the current context, Type I censoring and random censoring appear to be most appropriate. In a Type I censoring, an operating unit if not failed before a prefixed time, say,  $T$ , is replaced at that time by the spare if the spare is available. Thus, the effective life time is given by  $X^* = \min(X, T)$ . If we denote the d.f. of  $X^*$  by  $F^*$ , defined on  $R^+$ , then we may virtually repeat whatever we have discussed with  $F$  replaced by  $F^*$ . In this context, often,  $T$  is itself allowed to be a positive r.v. with a d.f.  $P(t)$ , and it is assumed that  $T$  and  $X$  are mutually independent. That leads to

$$1 - F^*(y) = \bar{F}^*(y) = [1-F(y)][1-P(y)] = \bar{F}(y)\bar{P}(y), \quad y \in R^+. \quad (5.57)$$

In the literature, this is referred to as random censoring. As regards the estimation of availability  $A_{FG}$  or  $A_{FG}^0$ , one may then equivalently estimate  $A_{F^*G}$  or  $A_{F^*G}^0$ . Since the impact of censoring is automatically taken into account in the definition of  $A_{F^*G}$  or  $A_{F^*G}^0$ , there is no need to go for any further complications than to replacing the  $X_i$  by  $X_i^*$  in the estimators considered earlier. Incidentally, by comparing  $A_{FG}$  and  $A_{F^*G}$  ( or  $A_{FG}^0$  and  $A_{F^*G}^0$  ), we may notice that  $F^* \leq F$ , so that

$$A_{FG} \geq A_{F^*G} \quad \text{and} \quad A_{FG}^0 \geq A_{F^*G}^0, \quad (5.58)$$

and the difference  $A_{FG} - A_{F^*G}$  (or  $A_{FG}^0 - A_{F^*G}^0$ ) indicates the loss due to censoring.

Type I and random censoring may also be adopted in the context of the estimation problems treated in Sections 3 and 4. There is however an important consideration. If, for example, there are  $k$  units, either in series or parallel, if a censoring takes place, it may simultaneously affect the entire system. Thus, one needs to introduce a model which should be capable of explaining this phenomenon in a more natural way. This is indeed possible, but a bit more complex than the simple structure explained here. We intend to communicate this in a future work.

The mechanism of censoring is much more complex with respect to the bundle strength of filaments treated in Section 2. One may need to bring-in some appropriate 'stress-strain' model to explain how a censoring may arise in such a context. Although this is very possible, a more elaborate analysis may be required to justify the basic assumptions appropriate in practical applications, and hence, we shall not deal with them here.

## REFERENCES

- Barlow, R.E. and Proschan, F. (1975). Statistical Theory of Reliability and Life Testing: Probability Models. Holt, Rinehart and Winston, New York.
- Cox, D.R. (1972). Regression models and life tables. J. Roy. Statist. Soc. Ser.B 34, 187-220.
- Daniels, H.A. (1945). The statistical theory of the strength of bundles of threads. Proc. Roy. Soc., Ser.A, 183, 405-435.
- Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. Ann. Math. Statist. 19, 293-325.
- Sen, P.K. (1960) . On some convergence properties of U-statistics. Calcutta Statist. Assoc. Bull. 10, 1-18.
- Sen, P.K.(1973). On fixed-size confidence bands for the bundle strenght of filaments. Ann. Statist. 1, 526-537.
- Sen, P.K.(1977). Some invariance principles relating to jackknifing and their role in sequential analysis. Ann. Statist. 5, 315-329.
- Sen, P.K. (1981). Sequential Nonparametrics: Invariance Princples and Statistical Inference. John Wiley and Sons, New York.
- Sen, P.K. (1986). Nonparametric estimators of Availability under provisions of spare and repair, II. In Reliability and Quality Control (ed:A.P.Basu), North Holland, Amsterdam, pp. 297-308.
- Sen, P.K.(1988a). Functional jackknifing : Rationality and general asymptotics. Ann. Statist. 16, 450-469.
- Sen, P.K.(1988b). Functional approaches in resampling plans: A review of some recent developments. Sankhya, Ser.A 50, 394-435.
- Sen, P.K. (1992). Resampling methods for the extrema of certain sample functions. In Proc. Calcutta Triennial Internat. Conf. Statist. (eds:S.K.Basu & B.K.Sinha), in press.
- Sen, P.K. and Bhattacharjee, M.C. (1986). Nonparametric estimators of availability under provisions of spare and repair, I. In Reliability and Quality Control (ed: A.P. Basu), North Holland , Amsterdam, pp. 281-296.
- Sen, P.K., Bhattacharyya, B.B. and Suh, M.W. (1973). Limiting behavior of the extrema of certain sample functions. Ann. Statist. 1, 297-311.