NONPARAMETRIC ESTIMATORS OF AVAILABILITY UNDER PROVISIONS OF SPARE AND REPAIR

by

P.K. Sen
Department of Biostatistics
University of North Carolina at Chapel Hill

and

M.C. Bhattacharjee
Indian Institute of Management
Calcutta, India

Institute of Statistics Mimeo Series No. 1461

May 1984

NONPARAMETRIC ESTIMATORS OF AVAILABILITY UNDER PROVISIONS OF SPARE AND REPAIR

M. C. Bhattacharjee and P. K. Sen

Indian Institute of Management, Calcutta, India

University of North Carolina, Chapel Hill, NC, USA

Nonparametric (as well as jackknifed) estimators are developed for the 'availability' of an equipment supported by a single spare and a repair facility, where down-time occurs whenever no spare/repaired unit is available at the point of failure of an operating unit. Asymptotic properties of the natural and the jackknifed estimators of availability are studied (without assuming the stochastic independence of life and repair times). Fixed sample size as well as sequential procedures are considered , and progressively censored schemes are also introduced in this context.

AMS Subject Classifications: 62G05, 62N05, 62L10.

Key Words and Phrases: *Availability; down time; jackknifing; progressive censoring schemes; renewal theorems; sequential confidence intervals; sequential tests; up time; weak convergence.*

1. INTRODUCTION

To introduce the basic model, we consider a *single-unit system* supported by a *repair facility* and a *single spare*. At the time when the operating unit fails, simultaneously, the spare takes over (instantaneously) as the new operating unit and the failed unit is sent for repair at the same instant which is a *regeneration point*. This system fails when the unit,currently operating , fails before the repair of the currently failed unit is completed. Otherwise, a failed unit on completion of repair assumes the role of a spare in cold stand-by attitude. Usually, it is assumed that the repair of a failed unit restores it to its new condition, and also that the life distributions of the original initial spare and the initial operating unit are the same, say F , defined on $R^+ = (0,\infty)$. If G be the distribution of the repair time of the failed unit (also defined on R^+), and if μ_F and μ_G denote respectively the means of the life time X and repair time Y having the distributions F and G, then the *limiting average availability* (i.e., the limiting expected proportion of *system up-times*) is defined as

$$A_{FG} = EO/\{EO + ED\}, \tag{1.1}$$

where EO is the mean time until system failure , measuring from the regeneration point, and ED is the mean *system down time*. If we denote by

$$\alpha_{FG} = P\{ X < Y \}, \tag{1.2}$$

then, we have

$$EO = \{1 - \alpha_{FG}\}^{-1}\mu_F \quad \text{and} \quad ED = \{1- \alpha_{FG}\}^{-1}E\{(Y-X)I(Y > X)\}, \tag{1.3}$$

where I(A) stands for the indicator function of the set A. The renewal theorem based results in (1.3) assume that for different i, the vectors (X_i,Y_i) are stochastically independent, though for each i, X_i and Y_i may not be independent. Assuming the operating periods (X_i) and repair periods (Y_i) to be mutually independent, a simplified version of (1.1) is given in Barlow and Proschan (1975, Ch.7), where other parametric models have also been briefly discussed. When both F and G are exponential d.f.'s, this reduces to

$$\mu_F(\mu_F+ \mu_G)(\mu_F^2 + \mu_F\mu_G + \mu_G^2)^{-1} = (1+\rho)/(1+\rho+\rho^2) , \tag{1.4}$$

where $\rho = \mu_G/\mu_F$. But, in general, A_{FG} in (1.1) depends on F and G in a more involved manner (and not just on their means). In practice, however, F and G are generally not known, and hence, for a desirable maintenance of the system, one may like to estimate A_{FG} in a nonparametric way (and also to provide some sharp bounds on A_{FG}) . Recently, Bhattacharjee and Kandar (1983) have considered some useful bounds for A_{FG} which are computable with only a limited knowledge about a few parameters of the unknown life and repair distributions.

One of the problems with such parametric procedures is their lack of *robustness* for departures from the assumed model (e.g., F actually Weibull against the assumed exponential form). Even for a small departure, there may be considerable loss of efficiency of the parametric procedure, while for major departures, they may even turn out to be inefficient or inconsistent. The main objective of the current study is to focus on some nonparametric developments both in the fixed sample size and sequential sampling schemes. Progressively censored schemes (PCS) based on the (regeneration) *cycle times* $\{T_i\}$ are also considered in this context.

Looking at (1.1), (1.2) and (1.3), we may gather that A_{FG} is a function of μ_F, α_{FG} and ED , each one of which being a regular functional of the d.f. F (and G). Thus, these estimable parameters (i.e., μ_F, α_{FG} and ED) can be estimated in a nonparametric fashion (under quite general regularity conditions) , wherein the validity and reliability of the estimates will be ensured for a broader class of F and G. Or , in other words, these nonparametric procedures will be robust and have broader scopes. However, in view of the fact that A_{FG} is not a linear function of these parameters, the resulting estimates may not be unbiased, and hence, some resampling schemes may be incorporated to reduce the effective bias, and *jackknifing techniques* are therefore adopted to achieve this goal.

In Section 2, without assuming the independence of the failure and repair times, a reformulation of A_{FG} in terms of some other functionals is provided, and this enables us to consider a very natural estimator. Fixed sample size estimation procedures are then introduced in Section 3. Section 4 is devoted to the development of sequential procedures pertaining to the confidence interval problem as well as the sequential testing problem. The concluding section deals briefly with PCS and some of the useful applications in the current context.

2. A REFORMULATION OF A_{FG}

Note that the ith cycle involves the life time X_i and repair time Y_i, and we do not necessarily assume that X_i and Y_i are mutually independent, although it is assumed that $\{(X_i,Y_i), i \geq 1\}$ form a sequence of independenent and identically distributed (i.i.d.) random vectors (r.v.) with a (bivariate) distribution function (d.f.) $H(x,y)$, defined on $[0,\infty)^2$. Thus, $F(x) = H(x,\infty)$ and $G(y) = H(\infty,y)$ are the two marginal d.f.'s. We also denote by $Z_i = X_i - Y_i$ and denote its d.f. by $P(z)$, $-\infty < z < \infty$. Then, note that parallel to (1.2),

$$\alpha_{FG} = P\{X \geq Y\} = P\{Z < 0\} = P(0) = \alpha_P , \qquad (2.1)$$

which may not be a sole functional of the marginal d.f.'s, but is simply expressible in terms of P(0). Thus, starting from the regeneration point, $E0$, the mean time upto the system failure is given by

$$E0 = \mu_F /(1-\alpha_P) = (1-\alpha_P)^{-1} \int_0^\infty x dF(x) . \qquad (2.2)$$

Note that if the system fails (i.e., encounters down time) in the mth cycle for the first time, for some $m \geq 1$, then $X_i \geq Y_i$ for every $i \leq m-1$ and $Y_m > X_m$. Therefore for the *stopping number* N , we have

$$P\{ N = m \} = [P(0)]^{m-1}[1-P(0)] , \text{ for } m = 1,2,\ldots\ldots \text{ ad inf.} \qquad (2.3)$$

Also, note that D, the system down time,is given by $Y_N - X_N$, where the stopping number N is defined earlier. Hence, we obtain that

$$\begin{aligned} ED &= \Sigma_{m\geq 1} E\{ (Y_m-X_m) I(N = m) \} \\ &= \Sigma_{m\geq 1} E\{ (Y_m-X_m) I(Y_i \leq X_i,\ i \leq m-1,\ Y_m > X_m)\} \\ &= \Sigma_{m\geq 1} [P(0)]^{m-1}[1-P(0)]\, E\{ (Y_m - X_m) \mid Y_m > X_m \} \\ &= [1-P(0)]\, \Sigma_{m\geq 1} [P(0)]^{m-1} \int_0^\infty z dP(z)/\{1-P(0)\}. \\ &= \{\int_0^\infty z dP(z)\}\, \Sigma_{m\geq 1} [P(0)]^{m-1} \\ &= \{1 - P(0)\}^{-1}\{\int_0^\infty z dP(z) \} \\ &= \{1 - P(0)\}^{-1} E\{(Y-X) I(Y > X)\} . \qquad (2.4) \end{aligned}$$

Hence, from (1.1), (2.2) and (2.4), we obtain that

$$A_{FG} = A_H = (EX)/(\ EX + E((Y-X)I(Y > X))\ )$$
$$= (EX)/(\ E(XI(X > Y)) + E(YI(Y > X))\ )$$
$$= (EX)/(E\{X \vee Y\}) \quad , \text{ where } a \vee b = \max(a,b). \qquad (2.5)$$

This gives a clear picture of the availability in terms of the life and repair times. In the special case where X and Y are independent with d.f. F and G, respectively, (2.5) reduces to

$$A_{FG} = \{\int_0^\infty (1-F(x))dx\ \} / \{\int_0 (1-F(x)G(x))dx\ \}. \qquad (2.6)$$

In the sequel, we shall find these expressions very useful.

3. ESTIMATION OF A_H : NON-SEQUENTIAL CASE.

Note that for a single copy of the system, we have the set of observations X_i, Y_i, $i=1,\dots,N$, where the stopping number N is defined by $\min\{k \geq 1: X_k < Y_k\}$. There is a down time $(Y_N - X_N)$, and following which, we have a regeneration point $(X_1+\dots+X_{N-1}+Y_N)$ where the repaired unit resumes the role of the operating one and the unit failing at time point $X_1+\dots+X_N$ (and remaining unattended for the down time period) is put to the repairing facility. Thus, from one regeneration point to the next one, we have a cycle of time length $T = X_1+\dots+X_{N-1} + Y_N = \Sigma_{i=1}^{N}(X_i \vee Y_i\ )$.

Now, for the ith cycle, we denote the associated r.v.'s by X_{ij} , Y_{ij} , $j=1,\dots,N_i$, N_i and T_i , so that $N_i = \min\{k \geq 1: X_{ik} < Y_{ik}\}$ and $T_i = \Sigma_{j=1}^{N_i}(X_{ij} \vee Y_{ij})$, for $i=1,2,\dots$. In a non-sequential setup, we are given these observable r.v.'s for $m(\geq 1)$ cycles, and based on these, we desire to provide a non-parametric estimate of A_H in (2.5). For latter use, we also define $O_i = \Sigma_{j=1}^{N_i} X_{ij}$ (the ith cycle on time), for $i=1,\dots,m$. Since the N_i are stopping times, we have

$$E(O_i) = EX_{i1}.EN_i = \mu_F\ \{\Sigma_{m=1}^{\infty} m[P(0)]^{m-1}[1-P(0)]\ \}$$
$$= \mu_F/[1-P(0)]\ . \qquad (3.1)$$

Similarly, by (2.4) and (3.1),

$$E(T_i) = E(O_i) + E(Y_{iN_i} - X_{iN_i}) = [1-P(0)]^{-1}\{\ \mu_F + E((Y-X)I(Y > X))\}. \qquad (3.2)$$

Thus, we may be tempted in estimating A_H by

$$\hat{A}_m = m^{-1} \Sigma_{i=1}^m (O_i/T_i) \ . \tag{3.3}$$

However, this estimator may have a basic undesirable property. Note that the O_i/T_i are i.i.d.r.v. (in fact, bounded between 0 and 1), so that $\hat{A}_m$ converges almost surely (a.s.) to $E(O_1/T_1)$. But, $E(O_1/T_1)$ need not be equal to A_H ; the amount of bias ,i.e., $E(O_1/T_1) - A_H$, depends on the unknown d.f. H. Even when H=FG, O_1/T_1 may not unbiasedly estimate A_H , so that even if m is large, $\hat{A}_m$ may not converge to A_H (but to $E(O_1/T_1)$ which may differ from A_H). For this reason, we consider the alternative estimator

$$A_m^* = \overline{O}_m/\overline{T}_m \quad \text{where } \overline{O}_m = m^{-1}\Sigma_{i=1}^m O_i \text{ and } \overline{T}_m = m^{-1}\Sigma_{i=1}^m T_i \ . \tag{3.4}$$

Note that by the Khintchine strong law of large numbers, as $m \to \infty$,

$$\begin{aligned} \overline{O}_m &\to E(O_1) = \mu_F/(1-P(0)) \text{ a.s.,} \\ \overline{T}_m &\to E(T_1) = (1-P(0))^{-1}\{ \mu_F + E[(Y-X)I(Y > X)]\} \quad \text{a.s.,} \end{aligned} \tag{3.5}$$

so that

$$A_m^* \to A_H \text{ a.s., as } m \to \infty \ . \tag{3.6}$$

Moreover, A_m^* is a bounded valued random variable $(0 \le A_m^* \le 1)$, so that (3.6) also ensures that $E(A_m^*)$ converges to A_H as $m \to \infty$. Thus, A_m^* is asymptotically unbiased for A_H. We denote the dispersion matrix of (O_1,T_1) by

$$\underset{\sim}{\Sigma} = \begin{pmatrix} \sigma_{OO} & \sigma_{OT} \\ \sigma_{TO} & \sigma_{TT} \end{pmatrix} . \tag{3.7}$$

Then, by the classical central limit theorem (on the (O_i,T_i)), as $m \to \infty$,

$$m^{\frac{1}{2}} (\overline{O}_m - EO_1 , \overline{T}_m - ET_1) \longrightarrow \mathcal{N}_2(\underset{\sim}{0}, \underset{\sim}{\Sigma}) \ . \tag{3.8}$$

Therefore, it follows by some standard steps that as $m \to \infty$,

$$m^{\frac{1}{2}}(A_m^* - A_H) \longrightarrow \mathcal{N}(0, \{E(T_1)\}^{-2}\{\sigma_{OO} - 2A_H\sigma_{OT} + A_H^2\sigma_{TT}\}). \tag{3.9}$$

Natural estimates of the parameters appearing on the right hand side of (3.9) may be used to test for a suitable hypothesis on A_H or to attach a confidence interval on A_H . In the absence of the knowledge of these paramters, the sample size m may not be fixed in a manner that the test has a specified power against a specified alternative or the width of the confidence interval is bounded by some prefixed positive number (2d). For either of these problems, we may need a sequential procedure, and we shall consider the same in the next section.

4. SEQUENTIAL METHODS

We have already noticed in Section 3 that A_m^* is generally not unbiased for A_H. Moreover, we need to estimate the parameters in the asymptotic distribution in (3.9). For this purpose, we employ *jackknifing* , by which we are able to reduce the bias and to estimate the asymptotic variance as well.

For every m($\geq$ 2), we define the $\bar{O}_m$, $\bar{T}_m$ and A_m^* , as in (3.4). Also, let $\bar{O}_{m-1}^{(i)}$ and $\bar{T}_{m-1}^{(i)}$ be defined as in (3.4) , but based on a sample of size m-1 obtained from the original sample (of size m) by omitting (O_i, T_i), for i=1,...,m. For every i(=1,....,m), we define

$$A_{m-1}^{*(i)} = \bar{O}_{m-1}^{(i)}/\bar{T}_{m-1}^{(i)} \quad ; \; A_{m,i}^* = mA_m^* - (m-1)A_{m-1}^{*(i)} \; . \tag{4.1}$$

Then, the jackknifed estimator of A_H is defined by

$$A_m^{**} = m^{-1}\Sigma_{i=1}^m A_{m,i}^* = A_m^* + [(m-1)/m]\Sigma_{i=1}^m (A_m^* - A_{m-1}^{*(i)}) \; . \tag{4.2}$$

Side by side, we introduce the jackknifed variance estimator

$$\begin{aligned} s_m^2 &= (m-1)^{-1} \Sigma_{i=1}^m (A_{m,i}^* - A_m^{**})^2 \\ &= (m-1)[\Sigma_{i=1}^m \{ A_{m-1}^{*(i)} - m^{-1}\Sigma_{j=1}^m A_{m-1}^{*(j)} \}^2]. \end{aligned} \tag{4.3}$$

Note that in the current situation, we have a function

$$g(a,b) = a/b, \text{where } 0 < a < b < \infty , \tag{4.4}$$

and moreover, in (3.4) or (4.1), the estimators $\bar{O}_m$ and $\bar{T}_m$ (or $\bar{O}_{m-1}^{(i)}$ and $\bar{T}_m^{(i)}$) are sample means (and hence, U-statistics of degree 1). Note further that

$$\partial g/\partial a = a/b , \quad \partial g/\partial b = -a/b^2 , \quad \partial^2 g/\partial a^2 = 0, \; \partial^2 g/\partial a\partial b = -b^{-2} \text{ and } \partial^2 g/\partial b^2 = 2a/b^3. \tag{4.5}$$

Therefore, we may formally write

$$g(a,b) = g(\alpha,\beta) + (a-\alpha)/\beta - \alpha(b-\beta)/\beta^2 - (a-\alpha)(b-\beta)/b^{o2} + (b-\beta)^2 a^o/b^{o3} , \tag{4.6}$$

where a^o lies between a and α and b^o lies between b and β . The second and third terms on the right hand sude of (4.6) are the so called linear terms, while the last two terms are the quadratic ones.

By virtue of (4.6), for enery $i: 1 \le i \le m$ and $m > 1$, we have

$$A^{*(i)}_{m-1} = A^*_m + (\bar{O}^{(i)}_{m-1} - \bar{O}_m)/\bar{T}_m - \bar{O}_m(\bar{T}^{(i)}_{m-1} - \bar{T}_m)/\bar{T}^2_m - (\bar{O}^{(i)}_{m-1} - \bar{O}_m)(\bar{T}^{(i)}_{m-1} - \bar{T}_m)/\tilde{T}^2_{m,i}$$
$$+ (\bar{T}^{(i)}_{m-1} - \bar{T}_m)^2 \tilde{O}_{m,i}/\tilde{T}^3_{m,i}, \text{ where } \tilde{O}_{m,i} \in (\bar{O}^{(i)}_{m-1}, \bar{O}_m) \text{ and } \tilde{T}_{m,i} \in (\bar{T}^{(i)}_{m-1}, \bar{T}_m). \quad (4.7)$$

Therefore, we have

$$A^*_{m,i} = A^*_m + (O_i - \bar{O}_m)/\bar{T}_m - (T_i - \bar{T}_m)/\bar{T}^2_m + (m-1)(\bar{O}^{(i)}_{m-1} - \bar{O}_m)(\bar{T}^{(i)}_{m-1} - \bar{T}_m)/\tilde{T}^2_{m,i}$$
$$- (m-1)(\bar{T}^{(i)}_{m-1} - \bar{T}_m)^2 \tilde{O}_{m,i}/\tilde{T}^3_{m,i}, \quad i=1,\dots,m. \quad (4.8)$$

By (4.2) and (4.8), we obtain that

$$A^{**}_m = A^*_m + \frac{m-1}{m}\Sigma^m_{i=1}(\bar{O}^{(i)}_{m-1} - \bar{O}_m)(\bar{T}^{(i)}_{m-1} - \bar{T}_m)/\tilde{T}^2_{m,i} - \frac{m-1}{m}\Sigma^m_{i=1}(\bar{T}^{(i)}_{m-1} - \bar{T}_m)^2 \tilde{O}_{m,i}/\tilde{T}^3_{m,i}. \quad (4.9)$$

Since $\bar{O}_m$ and $\bar{T}_m$ are both U-statistics (of degree 1), we may now appeal to the proof of Theorem 3.1 of Sen (1977) and conclude that as $m \to \infty$,

$$\max\{ (\bar{O}^{(i)}_{m-1} - \bar{O}_m)^2 : 1 \le i \le m \} = O(m^{-1}) \text{ a.s.}, \quad (4.10)$$

$$\max\{ (\bar{T}^{(i)}_{m-1} - \bar{T}_m)^2 : 1 \le i \le m \} = O(m^{-1}) \text{ a.s.}, \quad (4.11)$$

$$(m-1)\Sigma^m_{i=1}(\bar{O}^{(i)}_{m-1} - \bar{O}_m)^2 \to \sigma_{OO} \text{ a.s.}, \quad (4.12)$$

$$(m-1)\Sigma^m_{i=1}(\bar{T}^{(i)}_{m-1} - \bar{T}_m)^2 \to \sigma_{TT} \text{ a.s}, \quad (4.13)$$

and (3.5) holds. As such, $\max\{\tilde{T}^{-2}_{m,i} : 1 \le i \le m\}$ and $\max\{\tilde{O}_{m,i}/\tilde{T}^3_{m,i} : 1 \le i \le m\}$ can both be bounded a.s., by some positive, finite numbers, as $m \to \infty$. Therefore, from (4.9) through (4.13), we conclude that

$$|A^{**}_m - A^*_m| = O(m^{-1}) \text{ a.s., as } m \to \infty. \quad (4.14)$$

Moreover, by (4.8), (4.14) and the a.s. orders in (4.10) through (4.13), we conclude that as $m \to \infty$,

$$\max\{|(A^*_{m,i} - A^{**}_m) - (O_i - \bar{O}_m)/\bar{T}_m + \bar{O}_m(T_i - \bar{T}_m)/\bar{T}^2_m| : 1 \le i \le m\} = O(m^{-\frac{1}{2}}) \text{ a.s.}, \quad (4.15)$$

so that by (4.3) and (4.15), we conclude that as $m \to \infty$,

$$|s^2_m - s^{o2}_m| = O(m^{-1}) \text{ a.s.}, \quad (4.16)$$

where

$$s^{o2}_m = \bar{T}^{-2}_m (m-1)^{-1}\Sigma^m_{i=1}(O_i - \bar{O}_m)^2 - 2\bar{O}_m\bar{T}^{-3}_m (m-1)^{-1}\Sigma^m_{i=1}(O_i - \bar{O}_m)(T_i - \bar{T}_m)$$
$$+ \bar{O}^2_m\bar{T}^{-4}_m (m-1)^{-1}\Sigma^m_{i=1}(T_i - \bar{T}_m)^2. \quad (4.17)$$

By the strong convergence of U-statistics, we conclude that as $m \to \infty$,

$$(m-1)^{-1}\Sigma_{i=1}^{m}(O_i-\bar{O}_m)^2 \to \sigma_{OO} \text{ a.s.}, \ (m-1)^{-1}\Sigma_{i=1}^{m}(O_i-\bar{O}_m)(T_i-\bar{T}_m) \to \sigma_{OT} \text{ a.s.},$$
$$(m-1)^{-1}\Sigma_{i=1}^{m}(T_i - \bar{T}_m)^2 \to \sigma_{TT} \text{ a.s.}, \tag{4.18}$$

and for these, the existence of the second order moments suffices. Thus, by (3.5), (4.17) and (4.18), we conclude that whenever the d.f. H has finite second order moments, as $m \to \infty$,

$$s_m^{o2} \to \sigma_A^2 \ (= \{E(T_1)\}^{-2}\{\sigma_{OO} - 2A_H\sigma_{OT} + A_H^2\sigma_{TT}\}) \text{ a.s.} \tag{4.19}$$

Therefore, from (4.16) and (4.19), we obtain that

$$s_m^2 \to \sigma_A^2 \quad \text{a.s., as } m \to \infty . \tag{4.20}$$

Having obtained the a.s. convergence results in (3.6), (4.14) and (4.20), we are in a position to present the sequential estimation and testing procedures.

4.1. Bounded-width confidence interval for A_H. The underlying d.f. H ,and hence, A_H being unknown, we intend to provide a confidence interval for A_H, such that the confidence coefficient is $1 - \alpha$, for some preassigned α $(0 < \alpha < 1)$, and the width of this interval is bounded from above by 2d, for some preassigned $d(>0)$. For every $m > 1$, and $d > 0$, let

$$I_m(d) = \{ t : (A_m^{**} - d)\vee 0 \le t \le (A_m^{**} + d)\wedge 1\}. \tag{4.21}$$

Also, let τ_α be thr upper $100\alpha\%$ point of the standard normal d.f., and let m_o $(=m_o(d))$ be an initial sample size , usually greater than 2. Then, we may consider a *stopping number M (= M(d))*, defined by

$$M(d) = \min\{ m \ge m_o : s_m^2 \le md^2/\tau_{\alpha/2}^2 \} ; \tag{4.22}$$

if no such m exists, we let $M(d) = +\infty$. Whenever $M < \infty$, the confidence interval for A_H is taken as $I_M(d)$, defined as in (4.21), for M=m. Since M(d) is a stopping number, we can verify that for every (fixed) $d > 0$, $M(d) < \infty$, with probability 1. Now, by definition in (4.21), $I_M(d)$ has a width $\le 2d$, so that we need to verify that $I_M(d)$ has the coverage probability $1-\alpha$. Towards this, we proceed as in Chow and Robbins(1965) and work out the properties of this procedure in the asymptotic case where d is made to converge to 0.

We may note that by virtue of the law of iterated logarithm,

$$\overline{\lim} \{(m/\log\log m)^{\frac{1}{2}} \mid \overline{O}_m - EO_1 \mid\} = \{2\,\sigma_{OO}\}^{\frac{1}{2}} \text{ a.s.,} \tag{4.23}$$

$$\overline{\lim} \{(m/\log\log m)^{\frac{1}{2}} \mid \overline{T}_m - ET_1 \mid\} = \{2\,\sigma_{TT}\}^{\frac{1}{2}} \text{ a.s.,} \tag{4.24}$$

so that by (3.4), (4.6), (4.14) and (4.23)-(4.24), we conclude that as $m \to \infty$,

$$\begin{aligned} A_m^{**} - A_H &= (\overline{O}_m - EO_1)/ET_1 - (EO_1)(\overline{T}_m - ET_1)/(ET_1)^2 + O(m^{-1}\log\log m) \text{ a.s.} \\ &= m^{-1}\Sigma_{i=1}^m (ET_1)^{-2}\{(O_i - EO_i)ET_1 - (T_i - ET_i)EO_1\} + O(m^{-1}\log\log m) \text{ a.s.} \\ &= m^{-1}\Sigma_{i=1}^m W_i + O(m^{-1}\log\log m) \text{ a.s.} \end{aligned} \tag{4.25}$$

where the W_i are i.i.d.r.v. with mean 0 and a finite positive variance σ_A^2, defined by (4.19). This basic representation provides the basic tools for our subsequent analysis. For every m, we introduce a stochastic process $Q_m = \{Q_m(t), t \in [0,1]\}$ by letting $Q_m(t) = m^{-\frac{1}{2}}[mt](A_{[mt]}^{**} - A_H)/\sigma_A$, $t \in [0,1]$. Also, let $Q = \{Q(t), t \in [0,1]\}$ be a Wiener process on the unit interval [0,1]. Then, by an appeal to weak invariance principles for sums of i.i.d.r.v.'s [see for example, Sen(1981, Ch.2)], we conclude that under the assumed regularity conditions,

$$Q_m \underset{D}{\to} Q \text{, in the } J_1\text{-topology on } D[0,1] \text{, as } m \to \infty, \tag{4.26}$$

and further, the compactness of $\{Q_m\}$ remains in tact with respect to the uniform topology as well. A direct consequence of this weak invariance principle is that for every $\varepsilon > 0$ and $\eta > 0$, there exist a $\delta : 0 < \delta < 1$ and an m_o, such that

$$P\{\sup_{0 \le s < t \le s+\delta \le 1} | Q_n(t) - Q_n(s) | > \varepsilon\} < \eta, \quad \forall\, m \ge m_o, \tag{4.27}$$

so that if $\{T_m\}$ is any sequence of positive r.v., such that as $m \to \infty$,

$$T_m \to t, \text{ in probability, for some fixed } t \in [0,1], \tag{4.28}$$

then, by (4.26) and (4.27),

$$Q_m(T_m) \text{ is asymptotically normal with 0 mean and variance } t. \tag{4.29}$$

Now, for every $d(>0)$, we define

$$m(d) = \min\{ m \ge m_o : \sigma_A^2 \le md^2/\tau_{\alpha/2}^2 \}. \tag{4.30}$$

Then, by virtue of (3.9) and (4.14), along with the definition of σ_A^2 in (4.19),

$$\lim_{d \downarrow 0} P\{ A_H \in I_{m(d)}(d) \} = 1 - \alpha. \tag{4.30}$$

On the other hand, by (4.20), (4.22) and (4.30), as $d \downarrow 0$,

$$M(d)/m(d) \to 1 \text{ a.s.} \tag{4.31}$$

Note that if we define $T_{m(d)} = M(d)/m(d)$, $d > 0$, then, by (4.31), $T_{m(d)} \xrightarrow{p} 1$, as $d \downarrow 0$, so that by (4.28) and (4.29) (with $t = 1$), we obtain that as $d \downarrow 0$,

$$\{M(d)\}^{\frac{1}{2}} \left(A^{**}_{M(d)} - A_H \right)/\sigma_A \sim \mathcal{N}(0,1), \qquad (4.32)$$

and combining (4.32) with (4.20), we have by the Slutzky theorem, as $d \downarrow 0$,

$$\{M(d)\}^{\frac{1}{2}} \left(A^{**}_{M(d)} - A_H \right)/s_{M(d)} \sim \mathcal{N}(0,1). \qquad (4.33)$$

By (4.22) and (4.33), we conclude that

$$\lim_{d\downarrow 0} P\{ A_H \in I_{M(d)}(d) \} = 1 - \alpha . \qquad (4.34)$$

In the literature, (4.34) is termed the *asymptotic consistency* of the sequential procedure. In a sense, (4.31) reflects the *asymptotic efficiency* of the sequential procedure ; however, one usually defines this in terms of the following:

$$\lim_{d\downarrow 0} EM(d)/m(d) = 1 , \qquad (4.35)$$

which would fit with the Chow-Robbins(1965) definition. Verification of (4.35) naturally requires the convergence results on the moments of M(d), and, this in trun, requires the study of the moment convergence properties of s_m^2 . Note that unlike the case in Chow and Robbins(1965), we do not have here a reversed (sub-) martingale property of $\{s_m^2 , m \geq m_o\}$, so that the computation of the first order moment of $\sup_m s_m^2$, needed for the purpose, requires more elaborate analysis and more stringent regularity conditions. We note in this context that by (4.1), for every $i : 1 \leq i \leq m$, $m > 1$, $0 \leq A^{*(i)}_{m-1} \leq 1$, so that by (4.3)

$$s_m^2 \leq (m-1) \Sigma_{i=1}^m (A^{*(i)}_{m-1})^2 \leq m(m-1) \text{ , with probability 1.} \qquad (4.36)$$

Moreover, using the inequality that $\Sigma_{i=1}^m (O_i - \bar{O}_m)^2 \leq \Sigma_{i=1}^m O_i^2 \leq (\Sigma_{i=1}^m O_i)^2$ along with the same for the T_i , we obtain from (4.17) that

$$s_m^{o2} \leq 4m^2/(m-1) \text{ , with probability 1 } (\forall\, m \geq 2). \qquad (4.37)$$

Therefore, from (4.36) and (4.37), we obtain that

$$|s_m^2 - s_m^{o2}| \leq 5m^2 \text{ , with probability 1, for every } m \geq 2. \qquad (4.38)$$

Now, we assume that for some $r > 8$, $E|O_i|^r < \infty$ and $E|T_i|^r < \infty$. Then, in (4.10)-(4.11), in the right hand side, we may replace 'a.s.' by 'with a probability greater than $1- O(m^{-r/2+1})$, while in (4.12)-(4.13), we have with a probability greater than $1-O(m^{-r/2})$. As such, using (4.8)-(4.9), (4.38), and the modified steps in (4.10)-(4.13), we obtain that

$$E|s_m^2 - s_m^{o2}| = O(m^{-r/2+3}), \text{ for every } m \geq 2, \tag{4.39}$$

and this ensures that

$$E\{\sup_{m\geq 2}|s_m^2 - s_m^{o2}|\} \leq \Sigma_{m\geq 2} E|s_m^2 - s_m^{o2}| < \infty \quad , \text{ for } r > 8. \tag{4.40}$$

Moreover, we may rewrite

$$\sup_m s_m^2 \leq \sigma_A^2 + \sup_m |s_m^{o2} - \sigma_A^2| + \sup_m |s_m^2 - s_m^{o2}| , \tag{4.41}$$

so that, for our purpose, it suffices to show that for some $q \geq 1$,

$$E\{\sup_m |s_m^{o2} - \sigma_A^2|^q\} \leq \Sigma_{m\geq m_o} E|s_m^{o2} - \sigma_A^2|^q \text{ converges.} \tag{4.42}$$

By virtue of (4.37), we may now virtually repeat the proof of a theorem due to Cramér(1946, pp.353-356) and conclude that for $q = r/2$, $r > 8$,

$$E|s_m^{o2} - \sigma_A^2|^q = O(m^{-q/2}) + O(m^{-q/2-1/2}) , \tag{4.43}$$

so that (4.42) holds by noting that $q/2 = r/4 > 2$.

Note that by (4.22), for every $d > 0$,

$$s^2_{M(d)} \leq (d^2/\tau^2_{\alpha/2})M(d) \text{ and } s^2_{M(d)-1} > (d^2/\tau^2_{\alpha/2})(M(d)-1) . \tag{4.44}$$

Combinning (4.44) with (4.20), (4.30) and (4.39)-(4.42), we conclude that (4.35) holds. However, this result is not really needed from the practical application point of view, and (4.31) requiring only the finiteness of the second moments would suffice. We may notice that in view of (4.16), (4.17), (4.18), (4.44) and the (joint) asymptotic normality of Hoeffding's (1948) U-statistics, weak invariance principles are applicable for the $\{m(d)\}^{\frac{1}{2}}[s^2_{m(d)} - \sigma_A^2]$, as $d \downarrow 0$, so that we have under the finiteness of the fourth order moments,

$$\{m(d)\}^{-\frac{1}{2}}(M(d) - m(d)) \sim \mathcal{N}(0, \nu^2) , \tag{4.45}$$

for some finite ν, and this provides the *asymptotic normality of the stopping time*. The asymptotic normality result in (4.45) requires less stringent regularity conditions than the asymptotic efficiency result in (4.35). (4.45) also casts light on the variability of M(d) around m(d), the optimal sample size if σ_A were known.

4.2. Sequential tests for Availability. Consider the hypothesis testing problem :

$$H_0' : A_H = \theta_o \quad \text{vs.} \quad H_1' : A_H = \theta_1 = \theta_o + \Delta , \qquad (4.46)$$

where θ_o and Δ are specified, and we like the test to have the prescribed strength (α,β), where we restrict ourselves to positive α ,β for which $\alpha < \frac{1}{2}$ and $\beta < \frac{1}{2}$. Since the underlying d.f. H is not known, no fixed sample size test may meet the desired goal, and hence, we take recourse to a sequential procedure. In this context, (4.20) and (4.25) play the vital role.

We define two positive numbers (A,B), such that $\beta/(1-\alpha) \le B < 1 < A < \alpha/(1-\beta) < \infty$, and let $a = \log A$ and $b = \log B$, so that $b < 0 < a$. Typically, for small values of Δ , the lower (and upper) bounds for B (and A) are taken to be equal. Further, for every Δ , we define a positive integer $m_o(\Delta)$, such that as $\Delta \to 0$, $m_o(\Delta)$ goes to ∞ but $\Delta^2 m_o(\Delta)$ converges to 0 . Further, for every $m > 1$, we define the jackknifed estimator A_m^{**} as in (4.2) and the variance estimator s_m^2 as in (4.3). Then, starting with the initial sample size $m_o(\Delta)$, we continue drawing observations one by one so long as

$$bs_m^2 < m\Delta[A_m^{**} - (\theta_o+\theta_1)/2] < as_m^2 , \qquad m \ge m_o(\Delta) ; \qquad (4.47)$$

if, for the first time , (4.47) is vitiated for $m = M$ $(=M(\Delta))$, then we stop at that time and accept H_0 or H_1 according as $m\Delta[A_m^{**} - (\theta_o+\theta_1)/2]$ is $\le bs_m^2$ or $\ge as_m^2$. Thus, $M(\Delta)$ is the *stopping variable*. Note that in the current context, at the stopping number M, we have the data relating to M completed cycles of total time length $T_1 + \ldots + T_M$.

Note that by virtue of (3.9), (4.14) and (4.20), for every fixed θ_o and Δ ,

$$P\{ M(\Delta) > m \mid A_H \} \to 0 \text{ as } m \to \infty , \qquad (4.48)$$

so that the proposed test terminates with probability 1; the proof of (4.48) runs parallel to that in Section 6 of Sen(1977), and hence, is omitted.

To study the OC and ASN functions of the proposed test, as in Sen(1981, Ch.9), we consider an asymptotic setup wherein we allow Δ to converge to 0 (comparable to $d \downarrow 0$ in the earlier problem in Section 4.1), and, we set

$$A_H = \theta_o + \phi\Delta, \text{ where } \phi \in \Phi = \{\phi : |\phi| \leq K < \infty\} \quad . \qquad (4.49)$$

Also, we denote by $L_H(\phi,\Delta)$ the OC (i.e., probability of accepting H_0 when actually $A_H = \theta_o + \phi\Delta$) of the test based on (4.47). Note that as $\Delta \to 0$, $L_H(\phi,\Delta)$, for every fixed $\phi \in \Phi$, converges to a limit $P(\phi)$, and this is given by :

$$\lim_{\Delta\downarrow 0} L_H(\phi,\Delta) = P(\phi) = \begin{cases} (A^{1-2\phi} - 1)/(A^{1-2\phi} - B^{1-2\phi}) , & \phi \neq \frac{1}{2} , \\ a/(a-b) , & \phi = \frac{1}{2} . \end{cases} \qquad (4.50)$$

Note that $P(0) = 1-\alpha$ and $P(1) = \beta$, so that the limiting strength of the test is (α,β). By virtue of (4.20), (4.25) and the Skorokhod-Strassen embedding of Wiener process for sums of i.i.d.r.v.'s, the proof of (4.50) follows precisely on the same line as in the case of Theorem 6.1 of Sen(1977), and hence, is omitted.

Under the additional regularity conditions needed to ensure the existence of the momemts of s_m^2 in (4.41)-(4.43), we may again adapt the proof of Theorem 6.2 of Sen(1977) and arrive at the following :

Under the asymptotic setup in (4.49), for every (fixed) $\phi \in \Phi$,

$$\lim_{\Delta\downarrow 0} \{\Delta^2 E[M(\Delta) | \theta = \theta_o + \phi\Delta]\} = \psi(\phi,\sigma_A) \qquad (4.51)$$

where

$$\psi(\phi,\sigma_A) = \begin{cases} \{bP(\phi) + a[1-P(\phi)]\}\sigma_A^2/(\phi-\frac{1}{2}) , & \phi \neq \frac{1}{2} \\ -\sigma_A^2 ab, & \phi = \frac{1}{2} \end{cases} \qquad (4.52)$$

and σ_A^2 and $P(\phi)$ are defined by (4.19) and (4.50), respectively.

The asymptotic results in (4.50)-(4.52) provide good approximations in practice when Δ is small.

5. PROGRESSIVE CENSORING SCHEMES

We may note that for the statistical inference procedures described in Section 4, one needs the set of observations (O_i,T_i), $i \geq 1$. In an operating scheme, each observation (vector) (O_i,T_i) requires the running of the system for the total cycle time T_i , and these are nonnegative random variables. Thus, given a predetermined duration of the study, one would have a random number of cycles completed, and based on this random number of observations, one would be required to draw inference on A_H . Alternatively, one would allow the system to be operative until we have a prespecified number (m) of cycles completed, and then to base the statistical analysis on (O_i,T_i), i=1,...,m; in such a case, the duration of the study is equal to $T_1+\dots+T_m$ and is therefore stochastic in nature. In the usual fashion, we may term the two schemes as *Truncated (Type I Censored)* and *Censored (Type II Censored) Schemes*. In either way, when the joint d.f. of the (O_i,T_i) is not (at least roughly) known, we may have some drawback. Either we may have very few observations to base a test or formulate an estimate with reliable precisions, or we may have to wait too long to obtain the desired number of observations to initiate the statistical inference procedures. It is not unnatural to have interim analysis, so that if at any early stage, there is any strong indication of sub-standard functionning of the system, one may curtail the study and make corrective adjustments before putting it back to operation. In this context, continuous monitoring of the system is often advocated, and under such a monitoring scheme, *progressively censored schemes* can naturally be adopted with some advantage.

To be more general, we consider $n(\geq 1)$ independent copies of the system, and, for the ith copy, we denote the random variables by (O_{ij},T_{ij}), $j \geq 1$, for i=1,...,n. Also, for every $t(>0)$, we define the non-negative integer-valued random variables $\{ m_i(t), i=1,\dots,n \}$ by letting $T_{i0} = 0$, $i \geq 1$, and

$$m_i(t) = \max\{ k \geq 0 : T_{i0}+\dots+T_{ik} \leq t \}, \quad t > 0 , \ i \geq 1. \tag{5.1}$$

Then, if all the n copies of the system start operating at time 0, we have at time t, $N(t) = m_1(t) + \dots + m_n(t)$ observations (vectors) (O_{ij},T_{ij}), j=1,..., $m_i(t)$, i=1,...,n , and we desire to draw inference on A_H (as in Sections 3 and 4) in a repeated scheme , where we allow t to vary over an interval [0,T], for some positive and finite T. This situation is quite comparable to the sequential schemes in Section 4, excepting that the $N(t)$ are non-negative integer valued random variables, and hence, some additional adjustments are necessary to make the theory applicable.

In Section 4.1, we have studied sequential confidence intervals for A_H, and, in this context, the stopping number M(d), d > 0 plays the vital role. To adapt this sequential procedure in the setup of progressive censoring, we define for every d > 0,

$$\omega_d = \min\{ \ t > 0 : N(t) \leq s^2_{N(t)} d^2/\tau^2_{\alpha/2} \} . \tag{5.2}$$

Thus, the stopping time is given by ω_d, and, based on the number $N(\omega_d)$ of observations, we construct the confidence interval for A_H as in (4.21),with $m = N(\omega_d)$. At this stage, we may make use of the elemetary renewal theorem and conclude that

$$N(t)/(nt) \quad \to \quad \{ET_1\}^{-1} \quad \text{a.s., as t increases,} \tag{5.3}$$

so that by (4.20), (4.30), (5.2) and (5.3), we obtain that as $d \downarrow 0$,

$$\begin{aligned} n\,\omega_d &\sim m(d)E(T_1) \\ &\sim d^{-2}\,\tau^2_{\alpha/2}\,\sigma^2_A\{ET_1\} \quad \text{a.s.,} \end{aligned} \tag{5.4}$$

and,

$$\omega_d \sim n^{-1}d^{-2}\,\tau^2_{\alpha/2}\,\sigma^2_A\,(ET_1)\ \text{a.s.} \tag{5.5}$$

Note that in the above derivation, we have tacitly assumed that either n is fixed or n is allowed to be large, but nd^2 is small, so that (5.3) remains adaptable. The stopping time ω_d thus depends on n as well. Having obtained these results, we are in a position to verify (4.31) and (4.34) without any difficulty. To verify (4.35), with M(d) replaced by $N(\omega_d)$, we make use of the related *elementary renewal theorem* [viz., Ross(1970,pp.40-41)], and conclude that

$$\lim_{t\to\infty}\{EN(t)/(nt)\} = \{E(T_1)\}^{-1}. \tag{5.6}$$

As such, the results in Section 4.1 can readily be adapted to show that (4.35) holds in the current context too. Further, given (5.3), (5.5) and the asymptotic normality result in (4.45), we have the parallel result for the stopping time ω_d. The inversion theorem via (5.2) yields the desired result.

For the sequential testing problem in Section 4.2, we may replace in (4.47), m by N(t), $t \geq t_o$ for some positive t_o. The stopping number M(Λ), defined after

(4.47), has to be replaced by the stopping time $\omega^*(\Delta)$, defined by

$$\omega^*(\Delta) = \min\{\ t \geq t_o\ :\ N(t)\Delta[A^{**}_{N(t)} - (\theta_o+\theta_1)/2]/s^2_{N(t)} \notin (b,a)\}, \qquad (5.7)$$

where a and b are defined as in (4.47). With this minor change, we may prove (4.48) by replacing $M(\Delta)$ and m by $\omega^*(\Delta)$ and t, respectively. Under (4.49), the limiting OC function in (4.50) remains valid, while, for the limiting ASN function in (4.51), if we replace $M(\Delta)$ by $\omega^*(\Delta)$, then in (4.52), we need to replace σ_A^2 by $\{ET_1\}^{-1}\ \sigma_A^2$; this will then be the average stopping time in the limiting case. Control chart type sampling inspection plans (for A_H) under progressive censoring may also be adopted under the usual setup.

6. ACKNOWLEDGEMENTS.

Work of the second author was partially supported by the Office of the Naval Research, Contract N00014-83-K-0387

REFERENCES

Barlow, R.E. (1962). Repairman Problems. In *Studies in Applied Probability and Management Sciences (Ch.2)*, ed. Arrow, Karlin and Scarf., Stanford Univ.Press.

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

Bhattacharjee, M.C. and Kandar, R. (1983). Simple bounds on Availability in a model with unknown life and repair distributions. *J. Statist. Plan. Infer. 8*, 129-142.

Chow, Y.S. and Robbins, R. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *Ann. Math. Statist. 36*, 457-462.

Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton Univ. Press, N.J.

Gnedenko. B.V., Belyayev, Yu.K. and Solovyev, A.D. (1969). *Mathematical Methods of Reliability Theory*. Academic Press; New York.

Hoeffding, W. (1948). On a class of statistics with asymptotically normal distribution. *Ann. Math. Statist. 19*, 293-325.

Ross, S.M. (1970). *Applied Probability Models with Optimization Applications*. Holden-Day ; San Francisco.

Sen, P. K. (1977). Some invariance principles relating to jackknifing and their role in sequential analysis. *Ann. Statist.* 5, 316-329.

Sen, P. K. (1981). *Sequential Nonparametrics: Invariance Principles and Statistical Inference*. John Wiley; New York.