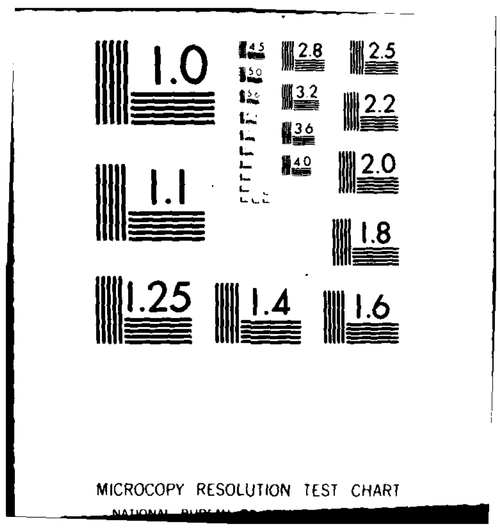
AD-A093 150 CALIFORNIA UNIV DAVIS INTERCOLLEGE DIV OF STATISTICS F/G 12/1
ESTIMATING A DISTRIBUTION FUNCTION WHEN NEW IS BETTER THAN USED--ETC(U)
SEP 80 F J SAMANIEGO, R A BOYLES AFOSR-77-3180
UNCLASSIFIED TR-22 AFOSR-TR-80-1159 NL

1 OF 1
AD A

END
DATE
FILMED
1-81
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
MICROCOPY RESOLUTION TEST CHART

3

LEVEL II



# TECHNICAL REPORT SERIES
# OF THE
# INTERCOLLEGE DIVISION OF STATISTICS

412106



University of California
Davis, California 95616

Estimating a Distribution Function
When New is Better Than Used

Francisco J. Samaniego and Russell A. Boyles

Technical Report No. 22

September 1980

AFOSR- 77-3180

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) UNCLASSIFIED

# REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

| | |
|---|---|
| 1. REPORT NUMBER (18) (19) AFOSR-TR-80-1159 | 2. GOVT ACCESSION NO. AD-A093 |
| 3. RECIPIENT'S CATALOG NUMBER 150 | |
| 4. TITLE (and Subtitle) (6) ESTIMATING A DISTRIBUTION FUNCTION WHEN NEW IS BETTER THAN USED, (9) | 5. TYPE OF REPORT & PERIOD COVERED Interim Technical Rept. |
| | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s) (10) Francisco J. Samaniego and Russell A. Boyles | 8. CONTRACT OR GRANT NUMBER(s) (15) AFOSR-77-3180 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS University of California Division of Statistics Davis, California 95616 (12) 30 (17) A5 | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F (16) 2304/A5 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332 | 12. REPORT DATE (11) September 1980 |
| | 13. NUMBER OF PAGES 27 |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (14) TR-22 | 15. SECURITY CLASS. (of this report) UNCLASSIFIED |
| | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

New better than used, nonparametric estimation, survival function, strong uniform consistency, type I censoring, convergence rates.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)


Let F be a distribution function on $(0,\infty)$, and let $S=1-F$ be its corresponding survival function. F is said to be New Better than Used (NBU) if $S(x)S(y) \quad S(x+y)$ for all x and y. Let $S_n(x)$ be the empirical survival function based on a random sample of size n from an NBU distribution function F. This paper is dedicated to the study of the estimator $\hat{S}_n(x)$ defined as $\sup[S_n(x+y)/S_n(y)[$, where the pupremum is taken over all y for which $S_n(y)$ $\theta$. It is shown that $\hat{S}_n$ is an NBU survival curve, and that it is relatively


DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 412 206

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) easy to compute. While $\hat{S}_n$ is not consistent in general, it is shown that $S_n$ is strongly uniformly consistent for S when the underlying distribution has compact support (for example, when sampling is subject tc type I censoring). Moreover, in such problem, the rate of convergence of $\hat{S}_n$ is shown to be optimal.

# Estimating a Distribution Function When New is Better Than Used

Francisco J. Samaniego and Russell A. Boyles
University of California, Davis

## Abstract

Let $F$ be a distribution function on $(0,\infty)$, and let $S = 1 - F$ be its corresponding survival function. $F$ is said to be New Better than Used (NBU) if $S(x)S(y) \geq S(x+y)$ for all $x$ and $y$. Let $S_n(x)$ be the empirical survival function based on a random sample of size $n$ from an NBU distribution function $F$. This paper is dedicated to the study of the estimator $\hat{S}_n(x)$ defined as $\sup\{S_n(x+y)/S_n(y)\}$, where the supremum is taken over all $y$ for which $S_n(y) > 0$. It is shown that $\hat{S}_n$ is an NBU survival curve, and that it is relatively easy to compute. While $\hat{S}_n$ is not consistent in general, it is shown that $\hat{S}_n$ is strongly uniformly consistent for $S$ when the underlying distribution has compact support (for example, when sampling is subject to type I censoring). Moreover, in such problems, the rate of convergence of $\hat{S}_n$ is shown to be optimal.

1. Introduction and Summary. A variety of nonparametric classes of life distributions have been introduced into the literature of statistical reliability theory over the last decade and a half. The appeal of such classes is that they directly model a qualitative characteristic of components or systems without additional (and often unnatural) constraints. The difficulty and the challenge that such nonparametric models pose is

that inference results tend to be more difficult to develop. This paper is dedicated to the estimation of a particular nonparametric class, the class of distributions with the "New Better than Used" (NBU) property.

Let $F$ be the distribution of a nonnegative random variable $X$. Let $S(x) \equiv 1 - F(x)$ represent the corresponding survival function. $F$ is said to be NBU if the inequality

$$S(x)S(y) \geq S(x+y) \tag{1.1}$$

obtains for all nonnegative $x$ and $y$. From a modeling viewpoint, the inequality (1.1) simply formalizes the notion that the residual lifetime of a used item tends to be shorter than the lifetime of a new item. A good deal is known about the class of NBU distributions. Much of Chapter 6 in Barlow and Proschan (1975) is dedicated to their study. In particular, it is well known that the class contains all distributions whose failure rate is increasing on the average (IFRA) and the still smaller class of distributions with increasing failure rate (IFR). By far, the major portion of the literature on NBU distributions deals with their stochastic properties rather than with inference questions. An inference paper worth noting is that of Hollander and Proschan (1972) which studies the problem of testing exponentiality against NBU alternatives.

The problem we study in this paper is the estimation of the distribution function $F$ under the assumption that $F$ belongs to the class of NBU distributions. Without the NBU restriction, the empirical distribution function $F_n$ converges to $F$ in several senses and at the best possible rate. However, $F_n$ need not be NBU, and, in fact, it is not difficult to show that $P(F_n \text{ is NBU}) \to 0$ when sampling from some NBU distributions (for example, the exponential distribution). Our goal here is to construct a sequence

$\{\hat{F}_n\}$ of NBU distributions which achieve the same asymptotic optimality as the sequence $\{F_n\}$.

Consistent estimators of IFR distributions were obtained by Marshall and Proschan (1965) and, in a somewhat different context, by Crow and Shimi (1972). Consistent estimators of IFRA distributions are given in Barlow et al. (1972) as a special case of their treatment of star-ordered distributions. All of these solutions make use of the methodology of isotonic regression. Such an approach is not useful for estimating NBU distributions since the natural monotonicity present in these smaller classes (IFR, IFRA) is absent in the general NBU class. Hence, a substantially different approach is taken in the current study.

Let $X_1,\ldots,X_n$ be a sample of size $n$ from a distribution $F$, assumed to be NBU. We propose to study here the properties of the function

$$\hat{S}_n(x) = \sup_y \frac{S_n(x+y)}{S_n(y)} \tag{1.2}$$

as an estimator of the survival function $S(x)$, where $S_n$ is the empirical survival function, and where the supremum is taken over nonnegative $y$ for which $S_n(y) > 0$. The form of $\hat{S}_n$ is motivated from the following considerations: If $S_n$ were an NBU survival function, then the inequality

$$S_n(x) \geq \frac{S_n(x+y)}{S_n(y)} \tag{1.3}$$

would be satisfied for all appropriate $x$ and $y$. If inequality (1.3) is violated for some $x$ and $y$, then a new function is constructed in an attempt to rectify this violation. For each fixed $x$, the new function seeks to accommodate all violations of (1.3) by taking a supremum over all violations. It is not clear at first sight that what results is a legitimate survival function. Our initial hope was that several iterations of the transformation given in (1.2) would result in an NBU survival function that would serve as an estimator. We have been able to show, however, that the $\hat{S}_n(x)$ itself is an NBU survival function, that it is strongly *uniformly* consistent in a class of problems of substantial practical importance (the sampling from an NBU distribution under type I censoring), and that, in such problems, it converges to the underlying NBU survival function $S(x)$ at an optimal rate. The general properties of $\hat{S}_n(x)$ are given in Section 2, consistency is studied in Section 3, and rate results are given in Section 4. In the final section, we discuss some remaining open problems and make some concluding general remarks.

## 2. Properties of $\hat{S}_n(x)$.

The properties we wish to establish for $\hat{S}_n(x)$ follow from general properties of the transformation $T : S \to T \circ S$ defined by

$$T \circ S(x) = \sup_{\{y \geq 0 | S(y) > 0\}} \frac{S(x+y)}{S(y)} . \qquad (2.1)$$

For simplicity of notation, we will usually delete the index over which suprema are taken; unless otherwise stated, the index should be understood to be the set of nonnegative $y$ for which the denominator of the fractional expression is positive.

Theorem 2.1. Let $F$ be distribution function on $(0,\infty)$, and let $S(x) = 1 - F(x)$ be the corresponding survival function. Then, the function $T \circ S(x)$ defined by (2.1) has the following properties:

(a) $T \circ S(x)$ is well defined, and $0 \leq T \circ S(x) \leq 1 \quad \forall\, x$, with $T \circ S(0) = 1$.

(b) $T \circ S(x)$ is nonincreasing.

(c) $T \circ S(x)$ is right continuous.

(d) If $S(x) = 0$ for $x \geq M$, then $T \circ S(x) = 0$ for $x \geq M$.

(e) For all nonnegative $x$ and $y$, $[T \circ S(x)][T \circ S(y)] \geq T \circ S(x+y)$.

(f) $T \circ S(x) \geq S(x)$ for all $x$.

(g) If $S$ is NBU, then $T \circ S = S$.

<u>Proof</u>.

(a) By hypothesis, $S(0) = 1$. Thus the set $\{y \geq 0 | S(y) > 0\}$ is nonempty, and $T \circ S(x)$ is well defined $\forall\, x$. Moreover, $T \circ S(x)$ is the supremum of a set of fractions, each of which is nonnegative and less than or equal to one. Thus,

$$0 \leq T \circ S(x) \leq 1$$

Also,

$$1 \geq T \circ S(0) = \sup_y \frac{S(y)}{S(y)} \geq \frac{S(0)}{S(0)} = 1$$

(b) Suppose $u < v$. Then for all $y$,

$$S(u+y) \geq S(v+y).$$

This implies that

$$T \circ S(u) = \sup_y \frac{S(u+y)}{S(y)} \geq \sup_y \frac{S(v+y)}{S(y)} = T \circ S(v).$$

(c) Let $\epsilon > 0$ be arbitrary, and let $x$ be fixed but arbitrary. By the definition of $T \circ S(x)$, we can find $y^*$ such that

$$\frac{S(x+y^*)}{S(y^*)} > T^\circ S(x) - \epsilon .$$

Then, for any $\delta > 0$, we have by part (b)

$$\begin{aligned} T \circ S(x) &\geq T \circ S(x^+) \\ &\geq T \circ S(x+\delta) \\ &\geq \frac{S(x+\delta+y^*)}{S(y^*)} , \end{aligned}$$

where $T^\circ S(x^+) = \lim\limits_{z \to x^+} T^\circ S(z)$. Since $S$ is right continuous, we have

$$\frac{S(x+\delta+y^*)}{S(y^*)} \to \frac{S(x+y^*)}{S(y^*)}$$

as $\delta \to 0$. Thus

$$\begin{aligned} T^\circ S(x) &\geq T \circ S(x^+) \\ &\geq \frac{S(x+y^*)}{S(y^*)} \\ &\geq T^\circ S(x) - \epsilon . \end{aligned}$$

Since $\epsilon$ is arbitrary, we have $T \circ S(x^+) = T \circ S(x)$.

(d) Suppose $S(x) = 0$. Then by (a), $S(x+y) = 0 \ \forall \ y \geq 0$, and thus

$$\begin{aligned} T \circ S(x) &= \sup_y \frac{S(x+y)}{S(y)} \\ &= 0 \end{aligned}$$

since the fraction $\frac{S(x)}{S(y)}$ is well defined for $y$ sufficiently small.

(e) $[T \circ S(x)][T \circ S(y)] = \sup_z \frac{S(x+z)}{S(z)} \sup_z \frac{S(y+z)}{S(z)}$

$$\geq \sup_z \frac{S(x+y+z)}{S(y+z)} \sup_z \frac{S(y+z)}{S(z)}$$

$$\geq \sup_z \frac{S(x+y+z)}{S(y+z)} \cdot \frac{S(y+z)}{S(z)}$$

$$= T \circ S(x+y)$$

(f) $S(0) = 1$ by hypothesis. Thus,

$$S(x) = \frac{S(x)}{S(0)}$$

$$\leq \sup_y \frac{S(x+y)}{S(y)}$$

$$\leq T \circ S(x).$$

(g) If $S$ is NBU, then for each fixed $x$ and all appropriate $y$,

$$S(x) \geq \frac{S(x+y)}{S(y)} .$$

Hence

$$S(x) \geq \sup_y \frac{S(x+y)}{S(y)}$$

$$= T \circ S(x).$$

This inequality, together with part (f) proves the claim. This completes the proof of the theorem. ■

Remarks. (1) While Theorem 2.1 does not deal with distributions $F$ having a point mass at $0$, such distributions cause no essential difficulty. If a distribution is degenerate at zero, then $T \circ S(x)$ is undefined; we can deal with this situation by simply defining $T \circ S(x) \equiv 0$ when $S \equiv 0$ on $[0,\infty)$. With this understanding, $T \circ S$ is well defined for any distribution on $[0,\infty)$, and all properties in Theorem 2.1 apply to the extension of $T \circ S$, with the exception of the equation $T \circ S(0) = 1$. If $F$ has a point mass at $0$, then the empirical c.d.f. $F_n$ is degenerate at zero with positive probability, and the above extension of the transformation $T$ is necessary. However, when the true distribution function $F$ is NBU, we must have $S(0) = 0$ or $1$, since $S(0)^2 \geq S(0)$. Thus, $F$ will either be degenerate at zero or give no mass to zero, and these properties are inherited by $F_n$.
(2) In general, $T \circ S$ need not be a survival function, that is, it need not be true that $T \circ S(x) \to 0$ as $x \to \infty$. For example, if $F$ is the Pareto distribution with survival function

$$S(x) = \frac{1}{x} \qquad \text{for } x \geq 1 ,$$

it is easy to see that $T \circ S(x) \equiv 1$. However, part (d) of Theorem 2.1 shows this cannot happen when $F$ has compact support.

Now, let $S_n$ be the emprirical survival function. Parts (a)-(d) of Theorem 2.1 prove that $\hat{S}_n(x) = T \circ S_n(x)$ is a survival function. Moreover, $\hat{S}_n(x)$ is NBU. Since NBU distributions are precisely the fixed points of the transformation $T$, $\hat{S}_n(x)$ is the empirical survival function whenever $S_n(x)$ is itself NBU.

We close this section with some remarks on the computation of $\hat{S}_n(x)$. For a given $S$, the computation of $T \circ S(x)$ can involve uncountably many maximization problems, each of which can be difficult analytically. It is therefore obvious that for some survival curves, the computation of $T \circ S(x)$ is intractable. Fortunately, such is not the case for the estimator $\hat{S}_n(x)$. The following result reduces the computation of $\hat{S}_n(x)$ to a simple programming problem.

Theorem 2.2. Let $x_1 \leq x_2 \leq \ldots \leq x_n$ be the order statistics from a random sample of size $n$ from the distribution $F$ on $(0,\infty)$. Let $S_n(x)$ be the empirical survival function, and let $\hat{S}_n(x)$ be as in (1.2). Then $\hat{S}_n(x)$ has the following characteristics:

(a) For any $x \geq 0$,

$$S_n(x) = \sup_i \frac{S_n(x+x_i)}{S_n(x_i)}$$

(b) $\hat{S}_n(x)$ is a step function, and if $\hat{S}_n(x)$ has a jump at $x$, then $x = x_r - x_s$ for some $r$ and $s$, where $0 \leq s < r \leq n$, with $x_0$ defined as zero.

Proof.

(a) Since $S_n$ is a step function, the ratio

$$\frac{S_n(x+y)}{S_n(y)}$$

will change values only when $y = x_i$ for some $i$ or when $x+y = x_i$ for some $i$. Now suppose the latter equation obtains for some $y$ that is not an order statistic. Then, for some positive $\epsilon$, we have

$$\frac{S_n(x+y-\epsilon)}{S_n(y-\epsilon)} > \frac{S_n(x+y)}{S_n(y)}$$

since, for $\epsilon$ sufficiently small,

$$S_n(y-\epsilon) = S_n(y)$$

while

$$S_n(x+y-\epsilon) > S_n(x+y).$$

This shows that for the purpose of defining $\hat{S}_n(x)$, we need only consider ratios of the form

$$\frac{S_n(x+x_i)}{S_n(x_i)}.$$

(b) We show that unless $x = x_r - x_s$ for some r,s such that $0 \leq s < r \leq n$, there exists a positive number $\epsilon$ for which $\hat{S}_n(x) = \hat{S}_n(x+\epsilon)$. Suppose $x \in (0,\infty)$ is such that $x \neq x_r - x_s$ for $0 \leq s < r \leq n$. From part (a), we know that

$$\hat{S}_n(x) = \sup_i \frac{S_n(x+x_i)}{S_n(x_i)}$$

Since no $x+x_i$ is equal to an order statistic, it is possible to find $\epsilon > 0$ such that

$$S_n(x+\epsilon+x_i) = S_n(x+x_i)$$

But then

$$\frac{S_n(x+\epsilon+x_i)}{S_n(x_i)} = \frac{S_n(x+x_i)}{S_n(x_i)} \qquad \forall\ i$$

so that

$$\hat{S}_n(x+\epsilon) = \hat{S}_n(x).$$

Thus, $x$ is not a jump point of $\hat{S}_n(x)$. This argument shows, in fact, that $\hat{S}_n$ is a step function whose jump points may only occur at numbers in the finite set $\{x \in (0,\infty) | x = x_r - x_s,\ 0 \le s < r \le n\}$. ∎

The result above shows that the computation of $\hat{S}_n$ involves checking at most $\binom{n+1}{2}$ points for possible jumps. At each potential jump point $x$, the comparison of the numbers

$$\frac{S_n(x+x_i)}{S_n(x_i)}$$

reveals the value of $\hat{S}_n$. Writing a program for computing $\hat{S}_n$ is thus entirely elementary. The computation is in fact even simpler than indicated, since a fair number of potential jump points may be eliminated from consideration by arguments exemplified in the following result.

<u>Theorem 2.3</u>. Let $x_0 = 0$, and let $x_1 \le x_2 \le \ldots \le x_n$ be the order statistics from a random sample of size $n$ from a distribution $F$ on $(0,\infty)$. Then

$$\hat{S}_n(x) = 1 \qquad \forall\ x < \max_i(x_i - x_{i-1}).$$

<u>Proof</u>: If $x < x_j - x_{j-1}$ for some $j$, then

$$S_n(x_{j-1}) = S_n(x + x_{j-1})$$

so that

$$\hat{S}_n(x) \ge \frac{S_n(x+x_{j-1})}{S_n(x_{j-1})} = 1\ .$$ ∎

Besides being relevant to computation, the result above is useful in our discussion of the consistency of $\hat{S}_n$.

3. <u>The Question of Consistency</u>. In spite of its appealing stochastic and computational properties, the transformation $\hat{S}_n(x) = T \circ S_n(x)$ does not provide the complete answer to the problem of estimating an NBU distri' - tion. Indeed, $\hat{S}_n$ is not in general a consistent estimator of $S$. Suppose, for example, that $S(x) = e^{-\lambda x}$, the survival function of the exponential distribution. The exponential is of course a boundary case in the family of NBU distributions, being the unique continuous distribution for which $S(x)S(y) = S(x+y)$. One can demonstrate the inconsistency of $\hat{S}_n(x)$ when sampling from the exponential distribution as follows. Fix $x$. Since $\{\hat{S}_n(x)\}$ are bounded random variables, it suffices to show that, for all $n$,

$$E\hat{S}_n(x) \geq M_x > e^{-\lambda x} .$$

Now it is well known that the spacing between the two largest order statistics from an exponential sample has the same distribution as the original observables; that is,

$$X_n - X_{n-1} \sim \text{Exp}(\lambda).$$

Now define

$$R_n(x) = \begin{cases} 1 & \text{if } x < X_n - X_{n-1} \\ S_n(x) & \text{if } x \geq X_n - X_{n-1} \end{cases} .$$

By Theorem 2.3 and part (f) of Theorem 2.1, we have

$$\hat{S}_n(x) \geq R_n(x).$$

Thus,

$$\begin{aligned} E\,\hat{S}_n(x) &\geq E\,R_n(x) \\ &\geq 1 \cdot P(X_n - X_{n-1} > x) \\ &\qquad + E\,S_n(x) \cdot P(X_n - X_{n-1} \leq x) \\ &= p + (1-p)e^{-\lambda x} \\ &> e^{-\lambda x}, \end{aligned}$$

where $p = P(X_n - X_{n-1} > x) > 0$.

It is easy to pinpoint the reason for the inconsistency demonstrated above. It is clear that the tail behavior of $S_n$ plays a crucial role in the asymptotic behavior of $\hat{S}_n$. This leads us to investigate the consistency of $\hat{S}_n$ in problems in which the tail behavior of $S_n$ is controlled. We establish in the following two theorems the strong uniform consistency of $\hat{S}_n$ when sampling from an NBU distribution with compact support.

<u>Theorem 3.1</u>. Let $F$ be an NBU distribution on $(0,\infty)$, and let $S = 1 - F$. Suppose the constant $T = \inf\{M | F(M) = 1\}$ is well defined and finite. Then for all $x$,

$$\hat{S}_n(x) \to S(x)$$

with probability one.

Proof. By the Glivenko-Cantelli theorem, there exists a null set $\mathcal{N}$ such that $\forall \omega \in \Omega/\mathcal{N}$,

$$\sup_z |F_n(z) - F(z)| \to 0.$$

Our notation suppresses the fact that the sequences we will consider depend on $\omega$. Since $\hat{S}_n(0) = 1$ and $\hat{S}_n(T) = 0$ for all $n$, we need only establish convergence of $\hat{S}_n(x)$ for $x \in (0,T)$. Let such an $x$ be fixed. For any $\omega \in \Omega/\mathcal{N}$, we have

$$\sup_y \frac{S_n(x+y)}{S_n(y)} \geq S_n(x)$$

which implies that

$$\underline{\lim}_{n\to\infty} \hat{S}_n(x) \geq \lim_{n\to\infty} S_n(x)$$

$$= S(x).$$

In order to obtain an upper bound for $\overline{\lim}_{n\to\infty} \hat{S}_n(x)$, we note that $\forall\, y$ for which $S(y) > 0$,

$$\lim_{n\to\infty} \frac{S_n(x+y)}{S_n(y)} = \frac{S(x+y)}{S(y)}$$

$$\leq S(x)$$

since $F$ is NBU.

We wish to show that for any $\epsilon > 0$,

$$\overline{\lim}_{n\to\infty} \sup_y \frac{S_n(x+y)}{S_n(y)} \leq S(x) + \epsilon. \tag{3.1}$$

Suppose (3.1) fails for some fixed $\epsilon_0$. Then there exists a sequence $\{y_{n_i}\}$ such that

$$\frac{S_{n_i}(x+y_{n_i})}{S_{n_i}(y_{n_i})} > S(x) + \epsilon_0 \qquad \forall\ i. \tag{3.2}$$

The sequence $\{y_{n_i}\}$ is bounded; indeed $y_{n_i} \le T - x \ \ \forall\ i$. We will assume, without loss of generality, that

$$y_{n_i} \to y^*,$$

for otherwise, we could extract a convergent subsequence satisfying (3.2). We now consider two cases: either $\{y_{n_i}\}$ contains a subsequence converging to $y^*$ from above, or it contains a subsequence converging to $y^*$ from below. Again, without loss of generality, we will assume first that $y_{n_i}$ increases to $y^*$ and then that $y_{n_i}$ decreases to $y^*$. We will derive a contradiction to (3.2) in either case.

Case 1: If $y_{n_i} \to y^*$ from below, then

$$\frac{S_{n_i}(x+y_{n_i})}{S_{n_i}(y_{n_i})} \to \frac{S(x+y^*-)}{S(y^*-)},$$

where $S(a-) = \lim_{z \to a^-} S(z)$. But

$$\frac{S(x+y^*-)}{S(y^*-)} \le S(x),$$

for otherwise, $\exists\ \epsilon > 0$ such that

$$\frac{S(x+y^* - \epsilon)}{S(y^* - \epsilon)} > S(x)$$

contradicting the fact that $F$ is NBU. Thus, for sufficiently large $n_i$,

$$\frac{S_{n_i}(x+y_{n_i})}{S_{n_i}(y_{n_i})} < S(x) + \epsilon_0,$$

contradicting (3.2).

<u>Case 2</u>: $y_{n_i} \to y^*$ from above. In this case

$$\frac{S_n(x+y_{n_i})}{S_n(y_{n_i})} \to \frac{S(x+y^*)}{S(y^*)} \leq S(x)$$

by the right continuity of $S$ and the fact that $F$ is NBU. As in the preceding case, this contradicts (3.2), proving (3.1).

In conclusion, we have shown that for every $\epsilon > 0$,

$$S(x) \leq \underline{\lim}_{n \to \infty} \hat{S}_n(x) \leq \overline{\lim}_{n \to \infty} \hat{S}_n(x) \leq S(x) + \epsilon$$

with probability one. Thus,

$$\hat{S}_n(x) \xrightarrow{a.s.} S(x) \qquad \text{for all } x \in (0,\infty). \qquad \blacksquare$$

If $S$ is continuous, the uniform convergence of $\hat{S}_n$ follows immediately from Theorem 3.1. However, the reliability applications of interest to us will typically involve survival functions with at least one discontinuity. We therefore establish the following general result.

<u>Theorem 3.2</u>. Under the hypotheses of Theorem 3.1,

$$\sup_x |\hat{S}_n(x) - S(x)| \to 0$$

with probability one.

<u>Proof</u>. Let $\omega \in \Omega/\eta$, the set of strong uniform convergence of $S_n$. In order to apply a lemma in Chung (1968, page 124) we need

$$\hat{S}_n(x) \to S(x) \tag{3.3}$$

on a set of $x$ dense in $(0,\infty)$, which we take to be the set of rationals Q, and

$$\hat{S}_n(x^-) \to S(x^-) \tag{3.4}$$

for every jump point $x$ of $S$.

We have (3.3) $\forall\, x \in Q$ by Theorem 3.1. We now establish (3.4) for a fixed jump point $x$. Let $\epsilon > 0$, and choose $\delta > 0$ such that $x - \delta \in Q$ and

$$S(x - \delta) \le S(x^-) + \epsilon. \tag{3.5}$$

Now since

$$\hat{S}_n(x^-) \le \hat{S}_n(x - \delta)$$

we have, by (3.3) and (3.5), that

$$\overline{\lim_{n\to\infty}}\, S_n(x^-) \le S(x-) + \epsilon. \tag{3.6}$$

On the other hand,

$$\hat{S}_n(x-) \ge S_n(x-)$$

so that

$$\varliminf_{n\to\infty} \hat{S}_n(x-) \geq S(x-). \tag{3.7}$$

Together, (3.6) and (3.7) imply that

$$\hat{S}_n(x-) \to S(x-).$$

By the aforementioned lemma in Chung (1968),

$$\sup_x |\hat{S}_n(x) - S(x)| \xrightarrow{a.s.} 0. \qquad \blacksquare$$

We now examine the domain of applicability of the consistency results established here. It is clear upon a moment's reflection that all real reliability experiments have a finite time horizon. Thus, random sampling or sampling subject to type II censoring from a life distribution with unbounded support are unrealistic (though perhaps still useful) models. It should certainly not be considered onerous, therefore, to impose the restriction of boundedness to the support of $F$, first, and most importantly, because this restriction corresponds to reality, and secondly, because the upper bound of the support of $F$ is unspecified and may be as large as a particular application requires.

Let $G$ be an NBU distribution on $(0,\infty)$. Suppose a sample $X_1,\ldots,X_n$ is taken from $G$, subject to type I censoring, that is, subject to truncation at time $T$. Thus, the variables we observe are $Y_1,\ldots,Y_n$, where

$$Y_i = \min(X_i,T)$$

where $X_1,\ldots,X_n \overset{iid}{\sim} G$. Let $F$ be the distribution of the Y's, and let $S = 1 - F$. It is easy to show that $F$ is NBU. Moreover, $F$ has compact support, so that Theorem 3.2 implies the strong uniform consistency of $\hat{S}_n$ as an estimator of $S$. We note that in most problems of interest, $F$ has a discontinuity at $T$, and thus the generality of Theorem 3.2 is necessary.

4. Convergence Rates for $\hat{S}_n$. It is possible to study rates of convergence of nonparametric estimators of a distribution function in a variety of contexts. We will study in this section the rate at which $\hat{S}_n$ converges to $S$ in mean square as well as the rate of almost sure pointwise convergence of $\hat{S}_n$ to $S$. It is well known that the empirical survival curve $S_n$ converges to $S$ in the two senses above at the best possible rates. Specifically, $S_n$ is mean square consistent with rate $\mathcal{O}(n^{-1})$ and almost surely pointwise consistent with rate $\mathcal{O}(n^{-\frac{1}{2}}(\log\log n)^{\frac{1}{2}})$. We establish below a fundamental inequality from which optimal convergence rates for $\hat{S}_n$ follow. We also obtain a rate result for almost sure uniform convergence of $\hat{S}_n$ to $S$ under an additional regularity condition.

Throughout this section, we assume that $F$ is an NBU distribution function on $(0,\infty)$ with $S = 1 - F$, and that $T = \inf\{M \mid F(M) = 1\}$ is well defined and finite. Let $\hat{S}_n$ be as in (1.2).

Lemma 4.1. Let $x \in (0,T)$, and let $\alpha \in (0,1)$. If $S_n(T-x) \geq \alpha S(T-x)$, then

$$|\hat{S}_n(x) - S(x)| \leq D_n\left[\frac{3}{\alpha S(T-x)} + 1\right],$$

where

$$D_n = \sup_{0 \leq z < \infty} |S_n(z) - S(z)|.$$

Proof. Since $\hat{S}_n(x) \geq S_n(x)$, we have

$$|\hat{S}_n(x) - S_n(x)| = \hat{S}_n(x) - S_n(x)$$

$$= \sup_y \left[ \frac{S_n(x+y)}{S_n(y)} - S_n(x) \right]$$

$$= \sup_y \left[ \frac{S_n(x+y) - S_n(x)S_n(y)}{S_n(y)} \right]$$

$$\leq \frac{1}{S_n(T-x)} \sup_y (S_n(x+y) - S_n(x)S_n(y)).$$

The inequality follows from the fact that $S_n$ is nonincreasing, and $S_n(T-x) > 0$ by hypothesis. Since $S_n(T-x) \geq \alpha S(T-x)$, which is positive by the definition of $T$, we have

$$|\hat{S}_n(x) - S_n(x)| \leq \frac{1}{\alpha S(T-x)} \cdot \sup_y \{S_n(x+y) - S(x+y) + S(x+y) - S(x)S(y) + S(x)S(y) - S_n(x)S(y) + S_n(x)S(y) - S_n(x)S_n(y)\}$$

$$\leq \frac{1}{\alpha S(T-x)} \sup_y |[S_n(x+y) - S(x+y)] + [S(x) - S_n(x)]S(y) + [S(y) - S_n(y)]S_n(x)|$$

since $S(x+y) - S(x)S(y) \leq 0 \quad \forall\, x, y$,

$$\leq \frac{1}{\alpha S(T-x)} (3\, D_n)$$

by the triangle inequality. Finally, we have

$$|\hat{S}_n(x) - S(x)| \leq |\hat{S}_n(x) - S_n(x)| + |S_n(x) - S(x)|$$

$$\leq \frac{3\, D_n}{\alpha S(T-x)} + D_n$$

$$= D_n \left[ \frac{3}{\alpha S(T-x)} + 1 \right]. \qquad \blacksquare$$

Our rate of convergence results for $\hat{S}_n$ will be obtained by applying the well-known exponential inequality due to Dvoretzki, Kiefer and Wolfowitz (1956), which we state below:

Lemma 4.2. Let $X_1,\ldots,X_n,\ldots \overset{iid}{\sim} F$, and let $F_n$ represent the empirical cdf based on $\{X_1,\ldots,X_n\}$. There exists a universal constant $c^*$ such that

$$P(\sqrt{n}\,D_n > \lambda) \le c^* e^{-2\lambda^2},$$

where

$$D_n = \sup_{-\infty < x < \infty} |F_n(x) - F(x)|.$$

We first study the rate of mean square consistency of $\hat{S}_n$. The following two results supply the answer.

Lemma 4.3. Let $x \in (0,T)$ and let $\alpha \in (0,1)$. Then

$$E|\hat{S}_n(x) - S(x)|^2 \le E\,D_n^2\left[\frac{3}{\alpha S(T-x)} + 1\right]^2 + \frac{(1 - S(T-x))}{n(1-\alpha)^2 S(T-x)}.$$

Proof. Let A denote the event $\{S_n(T-x) \ge \alpha S(T-x)\}$. Then

$$\begin{aligned} P(E^c) &= P(S_n(T-x) < \alpha S(T-x)) \\ &= P(S_n(T-x) - S(T-x) < -(1-\alpha)S(T-x)) \\ &\le P([S_n(T-x) - S(T-x)]^2 > (1-\alpha)^2(S(T-x))^2) \\ &\le \frac{S(T-x)(1-S(T-x))}{n(1-\alpha)^2 S(T-x)^2} \end{aligned}$$

by the Markov inequality

$$= \frac{1 - S(T-x)}{n(1-\alpha)^2 S(T-x)} .$$

Thus, by Lemma 4.1,

$$E|\hat{S}_n(x) - S(x)|^2 = \int_E |\hat{S}_n(x) - S(x)|^2 \, dP + \int_{E^c} |\hat{S}_n(x) - S(x)| \, dP$$

$$\leq \int_E \left(\frac{3}{\alpha S(T-x)} + 1\right)^2 D_n^{\,2} \, dP + P(E^c)$$

$$\leq E D_n^{\,2} \left[\frac{3}{\alpha S(T-x)} + 1\right]^2 + \frac{1 + S(T-x)}{n(1-\alpha)^2 S(T-x)} . \quad \blacksquare$$

We may now establish

<u>Theorem 4.1</u>. $E|\hat{S}_n(x) - S(x)|^2 = \mathcal{O}(n^{-1})$.

<u>Proof</u>. By Lemma 4.3, we need only show that $E(D_n^{\,2}) = \mathcal{O}(n^{-1})$. But

$$n E(D_n^{\,2}) = E((\sqrt{n}\, D_n)^2)$$

$$= \int_0^\infty P((\sqrt{n}\, D_n)^2 > \lambda) d\lambda$$

$$= \int_0^\infty P(\sqrt{n}\, D_n > \sqrt{\lambda}) d\lambda$$

$$\leq \int_0^\infty c^* e^{-2\lambda} d\lambda$$

$$= \frac{c^*}{2} < \infty . \quad \blacksquare$$

The rate of almost sure pointwise convergence is contained in the following result.

Theorem 4.2. $\hat{S}_n(x) - S(x) = \mathcal{O}(n^{-\frac{1}{2}}(\log\log n)^{\frac{1}{2}})$ almost surely.

Proof. For any fixed $\alpha \in (0,1)$, the inequality of Lemma 4.1 holds with probability one for all $n$ sufficiently large. ■

We may establish rates of uniform convergence under the additional hypothesis that $S(T-) > 0$.

Corollary 4.1. If $S(T-) > 0$, then

$$\sup_{0 \leq x \leq T} E|\hat{S}_n(x) - S(x)|^2 = \mathcal{O}(n^{-1})$$

and

$$\sup_{0 \leq x \leq n} |\hat{S}_n(x) - S(x)| = \mathcal{O}(n^{-\frac{1}{2}}(\log\log n)^{\frac{1}{2}}) \quad \text{a.s.}$$

Proof: If $S(T-) > 0$, we may obtain universal bounds in Lemma 4.1 and 4.3 by replacing $S(T-x)$ by $S(T-)$. ■

5. Discussion. We have studied in this paper the properties of an ad hoc estimator $\hat{S}_n(x)$ of the survival function of an NBU distribution. We have demonstrated the strong uniform consistency of $\hat{S}_n$ when the underlying distribution has compact support, and have verified that $\hat{S}_n$ tends to $S$ at the best possible rate in the sense of mean square convergence and almost sure pointwise convergence. Thus, this paper presents an asymptotically optimal solution to the problem of estimating an NBU distribution with compact support. An application of particular interest is that of sampling from an NBU distribution subject to type I censoring.

The problem of estimating an NBU distribution with unbounded support remains open. This problem is of theoretical rather than practical interest, since real reliability experiments have a finite horizon, even though this fact is ignored by most models. It is nevertheless of interest to construct a consistent NBU sequence in the general problem, and to compare the behavior of such a sequence to that of $\hat{S}_n$ in problems with compact support. These challenges are left to a future investigation.

The estimator $\hat{S}_n$ has the properties that $\hat{S}_n$ is NBU and $\hat{S}_n(x) \geq S_n(x)$ $\forall$ x. We have therefore found the name "upper NBU-izer" appropriate in referring to $\hat{S}_n$. The motivation for the definition of $\hat{S}_n$ suggests other possible NBU-izing transformations. An immediate candidate is the transformation

$$\tilde{S}_n(x) = \inf_{0 \leq y \leq x} S_n(y) S_n(x-y), \tag{5.1}$$

an estimator which is also motivated by the inequality (1.1). It can be shown that $\tilde{S}_n$ is strongly uniformly consistent for $S$, the survival function of an <u>arbitrary</u> underlying NBU distribution! Emphoria over this result is mitigated by the fact that $\tilde{S}_n$ lacks a crucial property, namely, $\tilde{S}_n$ need not be NBU. There is, however, a "lower NBU-izer" which, while being more complex, appears promising. We can show that the transformation

$$\tilde{T} \circ S(x) = \inf_{\underset{\sim}{u} \in C_x} \Pi\, S(u_i)$$

where

$$C_x = \{\underset{\sim}{u} \in (0,x]^k \mid \sum_{i=1}^{k} u_i \leq x; k=1,2,\ldots\},$$

transforms an arbitrary survival function into an NBU survival function. Moreover,

$$\tilde{T} \circ S(x) \leq S(x) \qquad \forall\, x,$$

so that $\hat{S}_n(x) = \tilde{T} \circ S_n(x)$ is indeed a lower NBU-izer. We will report in detail on the properties of $\tilde{S}_n(x)$ in a forthcoming paper.

6. Acknowledgments. We have benefitted from helpful conversations with Julius Blum and Thomas Ferguson.

## REFERENCES

Barlow, R.E., Bartholomew, D.J., Bremner, J.M. and Brunk, H.D. (1972). Statistical Inference under Order Restrictions. John Wiley and Sons, New York.

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

Chung, K.L. (1968). A Course in Probability Theory. Harcourt, Brace and World, Inc., New York.

Crow, L.H. and Shimi, I.N. (1972). Maximum likelihood estimation of life-distributions from renewal testing. Annals of Mathematical Statistics, **43**, 1827-1838.

Dvoretzki, A., Kiefer, J. and Wolfowitz, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. Annals of Mathematical Statistics, **27**, 642-669.

Hollander, M. and Proschan, F. (1972). Testing whether new is better than used. Annals of Mathematical Statistics, **43**, 1136-1146.

Marshall, A. and Proschan, F. (1965). Maximum likelihood estimation for distributions with monotone failure rate. Annals of Mathematical Statistics, **36**, 69-77.