The Annals of Statistics
1986, Vol. 14, No. 2, 607–618

TIME SEQUENTIAL ESTIMATION OF THE EXPONENTIAL MEAN UNDER RANDOM WITHDRAWALS[1]

By Joseph C. Gardiner, V. Susarla and John Van Ryzin

Michigan State University, State University of New York-Binghamton and Columbia University

In the context of lifetesting, an asymptotically risk-efficient procedure for the estimation of the exponential mean lifetime is considered when the survival times of the units are subject to random censorship. The loss function is the sum of squared error due to estimation, cost of recruitment of the units, and cost of total time on test. Asymptotic properties of the sequential estimator and stopping time are described as the per unit cost of total time on test decreases to zero.

1. Introduction. In several statistical experiments pertaining to reliability, life tests, and other longitudinal investigations a sample of units on test are under continual surveillance until one or the other specified terminal response is recorded for each unit. Such experiments may entail a considerable expenditure in costs and time particularly if the per unit cost of recruitment of subjects into the study and of follow-up time are high. It is then desirable to curtail observation at an intermediate state, prior to the last response being recorded, and base analyses on the current accumulated statistical evidence should it seem warranted for the study under consideration.

In this article we address the problem of estimation of the mean exponential lifetime θ from a sample of subjects whose survival times are deterred from complete observation due to random withdrawals or censorship. For each unit, the censoring variable Y is assumed independent of the survival time X, but is otherwise unknown, and the investigator only observes the datum (Z, δ) where $Z = \min(X, Y)$ and $\delta = 1$ or 0 according as $Z = X$ or Y. Let $Z_{(1)}, \dots, Z_{(n)}$ denote the order statistic corresponding to the random sample $Z_1, \dots, Z_n$ and $\delta_{[j]} = 1$ if $Z_{(j)}$ is uncensored and $= 0$ otherwise. At the kth response, $1 \le k \le n$, we have at our disposal the data $\{(Z_{(i)}, \delta_{[i]}): 1 \le i \le k\}$ on the basis of which an appropriate estimator $\hat{\theta}_{n,k}$ of θ can be constructed. The *loss* incurred up to this stage is measured by

$$L_{n,k} = a(\hat{\theta}_{n,k} - \theta)^2 + bn + cV_{n,k}, \tag{1.1}$$

where the *total time on test* (TTT) or follow-up time expended is

$$V_{n,k} = \sum_{i=1}^{k} Z_{(i)} + (n-k)Z_{(k)}. \tag{1.2}$$

The weights a, b, and c are given positive constants; b may be interpreted as the

Received June 1984; revised February 1985.

[1] Research supported in part by the National Institute of General Medical Sciences under Grant 1R01-GM28405 and in part by the Office of Naval Research under Contract N00014-79-C0522.

AMS 1980 *subject classifications*. Primary 62L12, 62L15; secondary 62G05.

Key words and phrases. Asymptotic normality, martingale, order statistics.

per unit *cost of recruitment* of units into the study and c the per unit *cost of follow-up time*. The risk in estimating θ by $\hat{\theta}_{n,k}$ is thus

$$R_{n,k} = aE(\hat{\theta}_{n,k} - \theta)^2 + bn + cEV_{n,k}, \tag{1.3}$$

and quite naturally, we seek to minimize the risk with appropriate choice of k, for a given sample size n and specified constants a, b, and c. With the underlying censoring distribution left unspecified and the variables $\{(Z_{(i)}, \delta_{[i]}: 1 \le i \le k\}$ being neither independent nor identically distributed, we do not have a tractable expression for the risk (1.3). However, when n is large, (1.3) can be reduced to a more transparent form from which the optimal choice of k can be readily obtained. Since θ and the censoring distribution are both unknown it turns out that no unique choice of k minimizes the risk universally and hence we explore an alternative sequential procedure.

In this paper we describe a time-sequential scheme for the estimation of θ which under some natural assumptions are *asymptotically risk efficient* as the per unit cost $c \to 0$. Furthermore, we derive pertinent asymptotic properties of the sequential estimator and stopping number of the proposed scheme. Section 2 states the main results of the paper together with the notation and assumptions to be used in the sequel. The proofs are given in Section 3 followed by several auxiliary lemmata whose detailed proofs are available in Gardiner, Susarla, and Van Ryzin (1984), which also contains other relevant references.

A time-sequential procedure analogous to that described in this article has been considered by Sen (1980) when censorship is absent. In this case the risk (1.3) takes on a particularly simple form for each n, k and further (1.2) is expressible as the sum of independent exponential (mean θ) variates. In fact we may take $\hat{\theta}_{n,k} = k^{-1}V_{n,k}$ and thus (1.3) reduces to $k^{-1}a\theta^2 + bn + kc\theta$. Sen's treatment is remarkably elegant and exploits fully these special circumstances. In the present article, our time-sequential procedure is based on the data $\{\{(Z_{(i)}, \delta_{[i]}): 1 \le i \le k\}, 1 \le k < n\}$ which are neither independent nor identically distributed. Furthermore the presence of random censorship and the nature of (1.1) leads to several technical complications which require a more subtle analysis.

2. Main results. Let X be an exponential r.v. with mean θ, $\theta \in (0, \infty)$, and let Y have the survival distribution $G(\cdot) = P(Y > \cdot)$. Both θ and G are unknown. Throughout the paper, we assumed that G has a continuous density g on its support $[0, y_0)$, $y_0 \le \infty$. Let $H = GF$ with $F(\cdot) = P(X > \cdot)$. In the random censorship model we observe (Z, δ), where $Z = X \wedge Y$ and $\delta = [X \le Y]$. The symbol $[A]$ denotes the indicator of the event A. For a random sample of n (≥ 1) units on test we have an underlying sequence of iid r.v.'s $\{(Z_i, \delta_i): 1 \le i \le n\}$. However, in view of the nature of our problem, at any intermediate stage k, we only have the data $\{(Z_{(i)}, \delta_{[i]}): 1 \le i \le k\}$ as described in the introduction. For each k, $1 \le k \le n$, we construct an estimator $\hat{\theta}_{n,k}$ of θ adapted to $\mathscr{B}_{n,k} = \sigma\{(Z_{(i)}, \delta_{[i]}): 1 \le i \le k\}$ by maximizing its likelihood over θ. This leads to a unique maximizer $V_{n,k}/\delta_{n,k}$, provided $\delta_{n,k} = \sum_{j=1}^{k}\delta_{[j]}$ is nonzero. So for $1 \le k \le n$, define $\hat{\theta}_{n,k}$ of θ by

$$\hat{\theta}_{n,k} = \begin{cases} \dfrac{V_{n,k}}{\delta_{n,k}} & \text{if } \delta_{n,k} > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

The risk incurred in estimation is given by (1.3) which we seek to minimize by an optimal choice k_n^0 of $k \in \{1, \ldots, n\}$. If k_n^0 $(\le n)$ is an optimal selection, the experiment is stopped at the k_n^0th stage and θ is estimated by $\hat{\theta}_{n,k_n^0}$ of (2.1) with corresponding minimum risk $R_n^0 = R_{n,k_n^0}$. In the generality considered here, if the weights a, b, c and sample size n are fixed, no explicit mathematical form in k for (1.3) is available and therefore we seek a solution to the problem determining k_n^0, when n is large. Formally, for a given per unit TTT cost c (> 0) we take $n = n(c)$ observations such that

$$\lim_{c \downarrow 0} cn^2(c) = a^* \quad \text{where } a^* \in (0, \infty). \tag{2.2}$$

For a justification of (2.2), see Sen (1980). We also assume that

$$b = \rho c \quad \text{where } \rho \in (0, \infty). \tag{2.3}$$

In the sequel all limits are taken as $c \downarrow 0$. We shall prove that, in this situation, an optimal choice of k_n^0 can be obtained by examining the behavior of (1.3) along sequences $\{k_n\}$ for which $n^{-1}k_n \to \lambda \in (0, 1]$. Then we show that (1.3) has the expansion

$$R_{n,k_n} = (a\theta^3/b_\lambda)k_n^{-1} + \rho cn + cb_\lambda k_n + o(k_n^{-1}), \tag{2.4}$$

where

$$b_\lambda = \lambda^{-1}\int_0^{H^{-1}(\lambda)} H(x)\,dx, \qquad H'(0)b(0) = -1, \tag{2.5}$$

and

$$H^{-1}(t) = \inf\{s > 0\colon 1 - H(s) \ge t\}. \tag{2.6}$$

From (2.4) an optimal choice is an integer k_n^* where

$$\operatorname{int}(a\theta^3/cb_\lambda^2)^{1/2} \le k_n^* \le \operatorname{int}(a\theta^3/cb_\lambda^2)^{1/2} + 1. \tag{2.7}$$

Here $\operatorname{int}(x)$ denotes the largest integer $\le x$. In view of (2.2), we also have $n^{-1}k_n^0 \to \lambda$ provided λ is a solution of

$$\int_0^{H^{-1}(\lambda)} H(x)\,dx = (a\theta^3/a^*)^{1/2}. \tag{2.8}$$

The left-hand side of (2.10) is a strictly increasing continuous function in λ whose value at $\lambda = 1$ is EZ. We can therefore obtain a unique solution in λ for (2.10) with $\lambda \in (0, 1]$ provided

$$\infty > a^* \ge a\theta^3/(EZ)^2. \tag{2.9}$$

However λ could be small if a^* is large. We are therefore led to define the optimal choice k_c^0 by

$$k_c^0 = \begin{cases} k_n^* & \text{if } a^* > a\theta^3(EZ)^{-2}, \\ n & \text{otherwise,} \end{cases} \tag{2.10}$$

where λ satisfies (2.8). Note that if $a^* > a\theta^3(EZ)^{-2}$ we have $\lambda < 1$ and so, asymptotically as $c \downarrow 0$, $k_c^0 < n$. Unfortunately both k_c^0 and R_c^0 still remain unspecified since θ and G are unknown. This motivates consideration of a sequential procedure defined via a random stopping number N_c and the associated estimator $\hat{\theta}_{n_c, N_c}$ $(= \hat{\theta}_c)$. This procedure is shown to be asymptotically risk efficient; that is, its risk $R_c^* = E[L_{n(c), N_c}]$ is such that $R_c^*/R_c^0 \to 1$ as $c \to 0$. Observe that $V_{n,k}$ of (1.2) can be expressed

$$V_{n,k} = n\int_0^{Z_{(k)}} H_n(x)\,dx, \tag{2.11}$$

where $nH_n(t) = \sum_{j=1}^n [Z_j > t]$. Then in view of (2.1) and (2.8), we define the stopping number N_c by

$$N_c = \begin{cases} \min\left\{k \le n-1 : \delta_{n,k}^3 \ge \left(\dfrac{a}{c}\right)V_{n,k}\right\}, \\ n \quad \text{if no such } k \text{ exists.} \end{cases} \tag{2.12}$$

We can now state the main results of the paper. All limits are taken as $c \downarrow 0$ in the rest of the paper.

THEOREM 2.1. *Under* (2.2) (i) $N_c/k_c^0 \to 1$ *a.s. and* (ii) $E(N_c/k_c^0)^r \to 1$ *for any* $r > 0$.

THEOREM 2.2. *Under* (2.4) (i) $\hat{\theta}_c \to \theta$ *a.s. and* (ii) $\mathscr{L}\{N_c^{1/2}(\hat{\theta}_c - \theta)/\hat{\sigma}_c\} \to \mathscr{N}(0,1)$, *where* $V_{n(c), N_c}\hat{\sigma}_c = N_c\hat{\theta}_c^3$.

THEOREM 2.3. *Under* (2.4) *and* (2.5) $R_c^*/R_c^0 \to 1$ *as* $c \to 0$.

3. Representation of risk $R_{n,k}$. Before we commence the proofs of the theorems of Section 2, we shall provide here arguments leading to the risk expansion (2.4) under the restriction to the sequences $\{k_n\}$ such that $k_n/n \to \lambda \in (0,1]$. We then show that for the minimization (of risk $R_{n,k}$) problem described in Section 2, the optimal sequence cannot be such that $k_n/n \to 0$. The argument for any other type of sequence $\{k_n\}$ is given right before (3.16). Throughout, $k_n \in \{1, \dots, n\}$ and the dependence of k_n on n is suppressed for convenience. All the inequalities involving conditional expectations are taken to hold almost surely. Let $\|\xi\|_p^p = (E[|\xi|^p])$ for any $p \ge 1$.

In view of (2.3), we write for each k,

$$\hat{\theta}_{n,k} - \theta = (V_{n,k} - \theta\delta_{n,k})\delta_{n,k}^{-1}[\delta_{n,k} \ge 1] - \theta[\delta_{n,k} = 0]. \tag{3.1}$$

Setting $U_{n,k} = V_{n,k} - \theta\delta_{n,k}$, and $kd_{n,k} = n\int_0^{H^{-1}(p_{n,k})} H(x)\,dx$, $(n+1)p_{n,k} = k$, (3.1) takes the form

$$\hat{\theta}_{n,k} - \theta = \theta(1/d_{n,k})\{(U_{n,k}/k) - W_{n,k}\}, \tag{3.2}$$

where

$$W_{n,k} = \left(k^{-1}U_{n,k}\right)\left\{k^{-1}\delta_{n,k} - \theta^{-1}d_{n,k}\right\} \times \left\{k^{-1}\delta_{n,k}\right\}^{-1}[\delta_{n,k} \geq 1] - \theta[\delta_{n,k} = 0]. \tag{3.3}$$

Our goal below is then to obtain the behavior of the second moment of $(\hat{\theta}_{n,k} - \theta)$ via representation (3.2) and the moments of $U_{n,k}$, $\delta_{n,k}$, and $\delta_{n,k}^{-1}$.

By Lemma 4.1, we have that $\{U_{n,k}^2 - \sum_{i=1}^k U_{n,i}^{*2}; \mathscr{B}_{n,k}: 1 \leq k \leq n\}$ is a zero mean martingale where, for each k,

$$U_{n,k}^* = (n - k + 1)\left(Z_{(k)} - Z_{(k-1)}\right) - \theta\delta_{[k]} = V_{n,k}^* - \theta\delta_{[k]}. \tag{3.4}$$

Therefore,

$$E\left[U_{n,k}^2\right] = E\left[\sum_{i=1}^k E\left[U_{n,i}^{*2}|\mathscr{B}_{n,i-1}\right]\right]. \tag{3.5}$$

Using the conditional distribution of $Z_{(k)}$ given $B_{n,k-1}$ and (3.4), we obtain that $E(U_{n,i}^{*2}|\mathscr{B}_{n,i-1}) = \theta E(V_{n,i}^*|\mathscr{B}_{n,i-1})$ which, via (3.5) and Lemma 4.4, implies that

$$k^{-1}E\left(U_{n,k}^2\right) \to \theta\lambda^{-1}\int_0^{H^{-1}(\lambda)} H(x)\,dx = \theta b_\lambda \tag{3.6}$$

under the assumption $k/n \to \lambda \in [0,1]$, which is assumed until otherwise stated. Therefore, in view of (2.5), we have from (3.2) and (3.6),

$$E\left(\hat{\theta}_{n,k} - \theta\right)^2 = \left(a\theta^3/b_\lambda\right)k^{-1} + o(k^{-1}) + \theta^2 d_{n,k}^{-2}\left\{E\left(W_{n,k}\left\{W_{n,k} - 2k^{-1}U_{n,k}\right\}\right)\right\}. \tag{3.7}$$

Since $V_{n,k} = \sum_{i=1}^k V_{n,i}^*$, Lemma 4.4 implies

$$k^{-1}E(V_{n,k}) \to b_\lambda. \tag{3.8}$$

Thus the expansion (2.4) will hold in view of (2.2), (3.7), and (3.8) provided we show

$$E\left(W_{n,k}\left\{W_{n,k} - 2k^{-1}U_{n,k}\right\}\right) = o(k^{-1}). \tag{3.9}$$

To obtain this, note that from (3.3) and Hölder's inequality,

$$\|W_{n,k}\|_2^2 \leq \|k^{-1}U_{n,k}\|_6^2 \|k^{-1}\delta_{n,k} - \theta^{-1}d_{n,k}\|_6^2 \times \|k^{-1}\delta_{n,k}[\delta_{n,k} \geq 1]\|_6^2 + \theta^2 P[\delta_{n,k} = 0]. \tag{3.10}$$

Therefore, applying Lemmata 4.2, 4.4, and 4.6 we obtain

$$0 \leq E\left(W_{n,k}^2\right) - \theta^2 P[\delta_{n,k} = 0] = O(k^{-2}). \tag{3.11}$$

When $n^{-1}k \to \lambda \in (0,1]$, we show $P[\delta_{n,k} = 0] = o(n^{-\alpha})$ for any $\alpha > 0$, whence in this case $E(W_{n,k}^2) = O(k^{-2})$. Also

$$E\left(W_{n,k}\left(k^{-1}U_{n,k}\right)\right) \leq \|W_{n,k}\|_2 \|k^{-1}U_{n,k}\|_2 = O(k^{-3/2}). \tag{3.12}$$

Combining (3.10), (3.11), and (3.12) will yield the desired (3.9). Thus we are left with proving

$$P[\delta_{n,k} = 0] = o(n^{-\alpha}) \tag{3.13}$$

provided $n^{-1}k \to \lambda \in (0,1]$. Consider first the case $\lambda \in (0,1)$. Since $[\delta_{n,k} = 0] = \bigcap_{i=1}^{k}[\delta_{[i]} = 0]$, we obtain that

$$P[\delta_{n,k} = 0] = \{n!/(n-k)!\}\int\left\{\prod_{i=1}^{k} F(y_i)g(y_i)\right\}H^{n-k}(y_k), \tag{3.14}$$

where the integration is over $\{(y_1,\dots,y_k): 0 < y_1 < \cdots < y_k < \infty\}$. On writing $Y_{(1)},\dots,Y_{(n)}$ for the order statistics corresponding to a random sample of size n from G, (3.14) may be written as

$$P[\delta_{n,k} = 0] = E\left\{\left(\prod_{i=1}^{k-1} F(Y_{(i)})\right)F^{n-k+1}(Y_{(k)})\right\}. \tag{3.15}$$

For arbitrary $\varepsilon > 0$ consider the expectation in (3.15) separately over the events $[Y_{(k)} \le \varepsilon]$ and $[Y_{(k)} > \varepsilon]$. Then we obtain exponential bounds for the probability of these two events; for the first one by a judicious choice of ε $(k/n - \bar{G}(\varepsilon) > 0$ for large $n)$ and Hoeffding's (1963) inequality, and the second one is bounded by $(F(\varepsilon))^{n-k+1}$. These two bounds lead to (3.13), and (3.13) holds also for $\lambda = 1$ since $\delta_{n,k}$ is nondecreasing in k.

Observe that when $n^{-1}k \to 0$, $R_{n,k} \ge (a\theta^3/b_0)k^{-1} + o(k^{-1})$. However, $R_{n,n} = (a\theta^3/EZ + a^*\rho + a^*EZ)n^{-1} + o(n^{-1})$ and therefore as $n \to \infty$, we have $R_{n,k} > R_{n,n}$. Thus the optimal sequence (k) cannot satisfy $n^{-1}k \to 0$. Also note that whenever $n^{-1}k \to \lambda \in [0,1]$, $kR_{n,k} \ge \inf\{a\theta^3/b_\lambda + a^*\lambda(\rho + \lambda b_\lambda)\} + o(1) = C + o(1)$, with $C = a\theta^3/\sup b_\lambda$. Now the usual arguments utilizing subsequences reveal that for all sequences (k_n) we have $\liminf kR_{n,k} \ge C + o(1)$. It now follows that we may restrict attention to sequences for which $n^{-1}k \to \lambda \in (0,1]$, in order to obtain the optimal choice k_c^0.

We now turn to the proofs of the theorems of Section 2. Recall the definition of the stopping number N_c of (2.14). We also define τ_c by

$$\tau_c = \begin{cases} \min\{k \le n-1: \delta_{n,k}^3 \ge (a/c)n^{-\gamma}\}, \\ n \quad \text{if no such } k \text{ exists.} \end{cases} \tag{3.16}$$

Here $\gamma > 0$ is a constant to be selected later. We prove

LEMMA 3.1. *As $c \downarrow 0$, almost surely, $\tau_c < n$ and $\tau_c \ge k_{1c}$ where* $k_{1c} = \operatorname{int}[(a/c)^{1/3}n^{-\gamma/3}]$.

PROOF. Note that $n^{-1}k_{1c} \le (a/cn^2)^{1/3}n^{-(1+\gamma)/3}$ and so from (2.4) as $c \downarrow 0$ we have $k_{1c} < n$. Also, the inequality $\tau_c \ge k_{1c}$ is immediate once we established $\tau_c < n$ a.s. as $c \downarrow 0$. To this end observe that from (3.16), $[\tau_c = n] \subset [n^{-1}\delta_{n,n} < (a/c)^{1/3}n^{-\gamma/3} + n^{-1}]$. Hence since $E(n^{-1}\delta_{n,n}) = \theta^{-1}EZ$ and $n(c) \sim c^{-1/2}$, we obtain that for any sequence $(c_m)\downarrow 0$ and n large enough $P[\tau_{c_m} = n] \le P[\{n^{-1}\delta_{n,n} - \theta^{-1}EZ\} < -(2\theta)^{-1}EZ]$. The result now follows from Hoeffding's (1963) inequality and the Borel–Cantelli lemma.

The salient properties of the stopping number NC are summarized in the following lemma. For $0 < \varepsilon < 1$, write $k_{2c} = \text{int}(k_c^0(1-\varepsilon))$ and $k_{3c} = \text{int}(k_c^0(1+\varepsilon))$, where k_c^0 was defined in (2.10). Notice that with c sufficiently small, $k_{1c} < k_{2c} < k_{3c}$. In the sequel all limits are as $c \downarrow 0$, unless otherwise stated.

LEMMA 3.2. *As $c \downarrow 0$, almost surely* (i) $N_c \ge \tau_c$ *and* (ii) $N_c < n$ *if* $a^*(EZ)^2 > a\theta^3$.

Furthermore, for $p \ge 2$, (iii) $P[N_c > k_{3c}] = O(c^{p/4})$, *and* (iv) $P[N_c \le k_{2c}] = O(c^{p(1-2\gamma)/12})$.

PROOF. (i) Observe that in view of Lemma 3.1,

$$(3.17)\quad [N_c < \tau_c] = [N_c < k_{1c}] \cup [N_c \ge k_{1c}]\left[\bigcap_{k=k_{1c}}^{N_c} [\delta_{n,k}^3 < (a/c)n^{-\gamma}]\right].$$

Now $[N_c < k_{1c}] = [\delta_{n,k}^3 \ge (a/c)V_{n,k}$, for some $k \in \{1,\dots,k_c - 1\}] \subseteq [Z_{(1)} \le n^{-(1+\gamma)}]$. Similarly the second event on the r.h.s. of (3.17) is contained in $[Z_{(1)} \le n^{-(1+\gamma)}]$. Since the series $\Sigma\{P[X \le n^{-(1+\gamma)}] + P[Y \le n^{-(1+\gamma)}]\}$ is convergent, it now follows from the above set of inequalities and Theorem 4.3.3 of Galambos (1978) that for any sequence (c_m) with $c_m \downarrow 0$ as $m \to \infty$, $\Sigma_m P[N_{c_m} < \tau_{c_m}] < \infty$. This gives (i) of the lemma.

(ii) By definition $[N_c = n] = [\delta_{n,n-1}^3 < (a/c)V_{n,n-1}]$. In view of Lemma 4.4, $(n-1)^{-1}V_{n,n-1} \to EZ$ a.s. and $(n-1)^{-1}\delta_{n,n-1} \to \theta^{-1}EZ$ a.s. Hence from (2.4) and our hypothesis $[N_c = n] \to [a^* \le a\theta^3/EZ^2] = 0$ a.s. which entails $N_c < n$ a.s.

To establish (iii) note that $P[N_c > k_{3c}] = P[V_{n,k_{3c}} > (c/a)\delta_{n,k_{3c}}^2(\delta_{n,k_{3c}}^2 - (a/c)\theta)]$. Write K for a generic constant not depending on c. From Lemma 4.4, $k_{3c}^{-1}\delta_{n,k_{3c}} \to \theta^{-1}b_{\lambda(1+\varepsilon)}$ a.s. and further since $a\theta^3 = (\lambda b_\lambda)^2 a^*$ we can take $\{(k_{3c}^{-1}\delta_{n,k_{3c}})^2 - (a/c)\theta k_{3c}^{-2}\} > K$ (>0) by taking c sufficiently small. Hence we obtain, for small c, $P[N_c > k_{3c}] \le K^{-p}k_{3c}^{-p}E|U_{n,k_{3c}}|^p = O(c^{p/4})$, by Lemma 4.2, provided $p \ge 2$, completing the proof of (iii).

(iv) We first observe that from definition $n^{-1}k_{2c} \to \lambda(1-\varepsilon)$ where $\lambda < 1$ or $\lambda = 1$ according as $a^* > a\theta^3/(EZ)^2$ or $a^* \le a\theta^3/(EZ)^2$. For either case the proof of (iii) is the same. Observe that

$$\begin{aligned}[N_c \le k_{2c}] &= [\delta_{n,k}^3 \ge (a/c)V_{n,k}, \text{for some } k \in \{\tau_c,\dots,k_{2c}\}]\\ &\subseteq \left[U_{n,k} \le (c/a)\delta_{n,k}\left(\delta_{n,k_{2c}}^2 - (a/c)\theta\right), \text{for some } k \in \{\tau_c,\dots,k_{2c}\}\right].\end{aligned}$$

Therefore we have

$$(3.18)\quad P[N_c \le k_{2c}] \le P\left[\bigcup_{k=\tau_c}^{k_{2c}} \left\{U_{n,k} \le (c/a)\delta_{n,k}\left(\delta_{n,k_{2c}}^2 - (a/c)\theta\right)\right\}\right].$$

In view of Lemma 4.4, and the fact that $(\lambda b_\lambda)^2 = (a\theta^3/a^*)$, we may take $\{(k_{2c}^{-1}\delta_{n,k_{2c}})^2 - (a/c)\theta k_{2c}^{-2}\} < -K$, with $K > 0$ for sufficiently small c. Thus in (3.18) we obtain for small c

$$(3.19)\quad P[N_c \le k_{2c}] \le P\left[\max_{k_{1c} \le k \le k_{2c}} |U_{n,k}| \ge K k_{2c}^{(2-\gamma)/3}\right].$$

Applying the maximal inequality to the martingale $\{U_{n,k}\}$ and Lemma 4.2 yields

$$P[N_c \le k_{2c}] \le K(k_{2c})^{p(\gamma-2)/3}(k_{2c})^{p/2} = O(c^{p(1-2\gamma)/12}), \tag{3.20}$$

completing the proof of (iv).

PROOF OF THEOREM 2.1. In view of (iii) and (iv) of Lemma 3.2, we have immediately $(k_c^0)^{-1}N_c \to 1$ a.s., by selecting $\gamma < \frac{1}{2}$ a priori and p (≥ 2) large enough.

Furthermore $(k_c^0)^{-1}N_c = (n^{-1}N_c)(n^{-1}k_c^0)^{-1} \le (n^{-1}k_c^0)^{-1}$ a.s. as $c \downarrow 0$, $n^{-1}k_{0c} \to \lambda > 0$. Thus $(k_c^0)^{-1}N_c$ is a.s. bounded for sufficiently small c. It then follows from the dominated convergence theorem that $E(k_c^0)^{-r}N_c^r \to 1$ for any $r > 0$.

PROOF OF THEOREM 2.2. (i) In view of Lemma 4.4 we have that $k^{-1}V_{n,k} \to b_\lambda$ a.s. whenever $n^{-1}k \to \lambda \in (0,1]$. Therefore from (2.10) and part (i) of Theorem 2.1 the result obtains.

(ii) Notice that $\hat{\sigma}_c^2 \to \theta^3/b_\lambda$ a.s. We first show

$$L\left[(k_c^0)^{1/2}(\hat{\theta}_{n,k_c^0} - \theta)\right] \to N(0, \theta^3/b_\lambda). \tag{3.21}$$

If $a^* \le a\theta^3(EZ)^{-2}$, then $k_c^0 = n$ and $\|W_{n,n}\| = O(n^{-2})$. Then (3.21) follows from (3.2) and an application of the ordinary central limit theorem. If $a^* > a\theta^3(EZ)^{-2}$ then $n^{-1}k_c^0 \to \lambda \in (0,1)$ with λ satisfying (2.8). Now (3.21) obtains once we establish

$$\mathscr{L}\left[(k_c^0)^{-1/2}U_{n,k_c^0}\right] \to \mathscr{N}(0, \theta b_\lambda). \tag{3.22}$$

To this end we apply the martingale central limit theorem of MacLeish (1974, Corollary 3.8). From Lemma 4.6 and the arguments preceding (3.6) we get

$$(k_c^0)^{-1}\sum_{i=1}^{k_c^0} E(U_{n,i}^{*2}|\mathscr{B}_{n,i-1}) \to \theta b_\lambda \quad \text{in probability.} \tag{3.23}$$

Furthermore, for each $\varepsilon > 0$ and $\eta > 0$, $E(U_{n,i}^{*2}[|U_{n,i}^*| > \varepsilon(k_c^0)^{1/2}])$ is bounded by $(\varepsilon^2 k_c^0)^{-\eta/2}E(|U_{n,i}^*|^{2+\eta})$ and, from the proof of Lemma 4.2 we have that for $p \ge 1$, $(k_c^0)^{-1}\sum_{i=1}^{k_c^0} E(|U_{n,i}^*|^p)$ is bounded in c. It now follows that $(k_c^0)^{-1}\sum_{i=1}^{k_c^0}E(U_{n,i}^{*2}[|U_{n,i}^*| > \varepsilon(k_c^0)^{1/2}]) \to 0$ and so together with (3.23) we get (3.22). Finally Theorem 2.2 will be established once we prove $E(\hat{\theta}_{n,N_c} - \hat{\theta}_{n,k_c^0})^2 = o(c^{1/2})$ which follows from (3.25) and (3.26) below.

PROOF OF THEOREM 2.3. The proof of the theorem follows along lines analogous to those of Theorem 1 of Gardiner and Susarla (1984) and therefore only an outline of the details is provided here. Notice that

$$\begin{aligned} R_c^*/R_c^0 - 1 &= a_0 c^{-1/2}\left\{E(\hat{\theta}_{n,N_c} - \theta)^2 - E(\hat{\theta}_{n,k_c^0} - \theta)^2\right\} \\ &\quad + a_1 c^{1/2}\{EV_{n,N_c} - EV_{n,k_c^0}\} \\ &= A_0 + A_1 \quad \text{(say)}, \end{aligned} \tag{3.24}$$

where a_0, a_1 are constants independent of c. Since $E(\hat{\theta}_{n,k_c^0} - \theta)^2 = (a\theta^3)^{1/2}c^{1/2} + o(c^{1/2})$ and utilizing the Hölder inequality we can establish $A_0 = o(1)$ in (3.24) once we show

$$E\Big\{\big(\hat{\theta}_{n,N_c} - \hat{\theta}_{n,k_c^0}\big)^2[k_{2c} < N_c \le k_{3c}]\Big\} = o(c^{1/2}), \tag{3.25}$$

$$E\Big\{\big(\hat{\theta}_{n,N_c} - \theta\big)^2[N_c \le k_{2c} \text{ or } N_c > k_{3c}]\Big\} = o(c^{1/2}), \tag{3.26}$$

and a corresponding statement of (3.26) with $\hat{\theta}_{n,N_c}$ replaced by $\hat{\theta}_{n,k_c^0}$. To establish (3.26) consider first $E\{(\hat{\theta}_{n,N_c} - \theta)^2[N_c \le k_{2c}]\}$. (The proof of the remaining part of (3.26) is entirely analogous.) From (3.2),

$$E\Big\{\big(\hat{\theta}_{n,N_c} - \theta\big)^2[N_c \le k_{2c}]\Big\} \le E\Big\{\big(N_c^{-1}U_{n,N_c}\big)^2[N_c \le k_{2c}]\Big\} + E\big(W_{n,N_c}^2[N_c \le k_{2c}]\big), \tag{3.27}$$

where here and in the sequel, we have suppressed constants not depending on c. Now the first term on the r.h.s. of (3.27) is bounded by $E^{1/p}\{\max_{k_{1c} \le k \le k_{2c}}(k^{-1}U_{n,k})\}^{2p}(P[N_c \le k_{2c}])^{1/q}$ with $p^{-1} + q^{-1} = 1$. On applying the maximal inequality, Lemmata 3.2 and 4.2 we find that the term is of order $O(c^{(2h+1)/2})$ where $h = \frac{1}{3}\{(s/4q) - 1 - ((s/2q) + 1)\gamma\}$ and $s \ge 2$. Taking $\gamma < \frac{1}{2}$ and (s/q) appropriately we have $h > 0$. This yields

$$E\Big\{\big(N_c^{-1}U_{n,N_c}\big)^2[N_c \le k_{2c}]\Big\} = o(c^{1/2}). \tag{3.28}$$

For the second term in (3.27), use Lemma 3.2 and the definition τ_c to drop the term $[\delta_{n,k} = 0]$ in $W_{n,k}$ of (3.3). Then using the fact that $W_{n,k}\|_p = o(k^{-1})$ for $p \ge 1$, the Hölder inequality and Lemma 3.2 we obtain $E\{W_{n,N_c}^2[N_c \le k_{2c}]\} = O(c^{1/2})$ whence in view of (3.28) the proof of (3.26) may be terminated.

To show (3.25) we again use (3.2) and consider the two parts involving $U_{n,k}$ and $W_{n,k}$ separately. Then bound $E(\{N_c^{-1}U_{n,N_c} - (k_c^0)^{-1}U_{n,k_c^0}\}^2[k_{2c} < N_c \le k_{3c}])$ by $E(\max_{k_{2c} < k \le k_{3c}}(k^{-1}U_{n,k} - (k_c^0)^{-1}U_{n,k_c^0}))^2$. Notice that for $k_c^0 < k \le k_{3c}$

$$\Big(k^{-1}U_{n,k} - \big(k_c^0\big)^{-1}U_{n,k_c^0}\Big)^2 \le 2\big(k_c^0\big)^{-2}\big(U_{n,k} - U_{n,k_c^0}\big)^2 + 2\Big\{\big(k_c^0\big)^{-1} - k_{3c}^{-1}\Big\}^2 U_{n,k_c^0}^2. \tag{3.29}$$

For the first term in (3.29) use the maximal inequality of the martingale $\{(U_{n,k} - U_{n,k_c^0})\colon k_c^0 < k \le k_{3c}\}$ and for the second term use Lemma 4.2. Then since ε can be arbitrarily chosen and $(k_c^0)^{-2} \sim c$ we get

$$E\bigg(\max_{k_c^0 < k \le k_{3c}}\Big(k^{-1}U_{n,k} - \big(k_c^0\big)^{-1}U_{n,k_c^0}\Big)\bigg)^2 = o(c^{1/2}). \tag{3.30}$$

For the term involving $W_{n,k}$ in (3.25) follow the same argument using the fact that $E(W_{n,k}^2) = O(k^{-2})$. Then together with (3.30) this establishes (3.25). We are left with the term A_1 in (3.24). Once again consider separately the expectation

$E(V_{n,N_c} - V_{n,k_c^0})$ restricted to each of the events $[N_c \le k_{2c}]$, $[N_c > k_{3c}]$, and $[k_{2c} < N_c \le k_{3c}]$. For the first term use Lemma 3.2 with $\gamma < \frac{1}{2}$ to show that it is of order $O(c^{(1-2\gamma)/12})$ and likewise the second term is of order $o(c^{1/2})$. Finally for the last term we have a bound $\varepsilon O(c^{-1/2})$. This establishes $A_1 = o(1)$ and so from (3.24) the theorem is proven.

4. Auxiliary lemmata. We present here the proofs of several auxiliary results utilized in the proofs of Section 3 some of which are of interest in themselves. In particular, Lemma 4.1 gives a general martingale result and Lemma 4.3 gives a moment bound on a useful functional of centered empiricals. If $\xi_1, \ldots, \xi_n$ are n uniform (0,1) r.v.'s, $n\Gamma_n(t) = \sum_{i=1}^n [\xi_i \le t]$ will denote its empirical d.f. Also if I is the identity function on (0,1) and $q(t) = \{t(1-t)\}^{1/r}$, $r > 2$, $t \in (0,1)$, we write $\rho_q(\Gamma_n, I) = \sup\{|\Gamma_n(t) - t|/q(t): 0 < t < 1\}$. $c_1, c_2, \ldots$ are constants independent of n and of any k in $\{1, \ldots, n\}$.

LEMMA 4.1. *For any* $n \ge 1$, $\{U_{n,k}, \mathscr{B}_{n,k}: 1 \le k \le n\}$ *is a zero-mean martingale.*

LEMMA 4.2. *For any* $p \ge 2$ *and* n, $\|U_{n,k}\|_p \le c_1 k^{1/2}$.

PROOF. By the theorem of Dharmadhikari, Fabian, and Jogdeo (1968) and Lemma 4.1, $\|U_{n,k}\|_p \le C_p \|(\sum_{i=1}^k U_{n,i}^{*2})^{1/2}\|_p$ for $p > 1$ with c_p depending only on p. Now an application of Hölder's inequality followed by c_r-inequality obtains that the right-hand side of the above inequality is $O(k^{p/2})$ provided that $E[k^{-1}\sum_{i=1}^k V_{n,i}^{*p})] < M < \infty$ with M independent of k and n. But this follows since

$$(4.1)\quad E\left[V_{n,i}^{*p} | B_{n,i-1}\right] = p\int_0^\infty x^{p-1}\left\{H\left(Z + x(n-i+1)^{-1}\right)/H(Z)\right\}^{n-i+1} dx,$$

where $Z_{(i-1)}$ has been abbreviated to Z and since the integrand can be dominated by the integrable function $x^{p-1}\exp(-x/\theta)$.

LEMMA 4.3. *For each* $p > 0$, $\|\rho_q(\Gamma_n, I)\|_p = O(n^{-1/2})$, *provided* $r > 2 \vee p$.

PROOF. We obtain the result by showing that

$$A = n^{p/2} E\left[\left\{\sup\{|\Gamma_n(t) - t|/q(t): 0 \le t \le \tfrac{1}{2}\}\right\}^p\right] = O(1).$$

To get this result, we take $\theta = \frac{1}{2}$ and q as defined earlier in Theorem 1 of Wellner (1977b). Then with Y_i as in Wellner's result, we have $E[|Y_1|^p] < \infty$ for $p < r$. Now let $p \ge 2$ until otherwise stated. Now Wellner's theorem obtains that $A \le c_2 E[|T_n|^p]$ where $nT_n = \sum Y_i$ with $Y_i = q^{-1}(\xi_i)(0 < \xi_i \le \frac{1}{2}] - \int_0^{\xi_i}(1/(1-x)q_{1/2}(x))\,dx$. Hence an application of Burkholder's inequality followed by Jensen's inequality (need $p \ge 2$ here) to $E[|T_n|^p]$ shows that $A \le c_3 E[|Y_1|^p] < \infty$, completing the proof for $p \ge 2$. For $p < 2$, $A \le 1 + p\int_1^\infty \lambda^{p-1} P[\sup\{\sqrt{n}|\Gamma_n(t) - t|/q(t): 0 \le t \le \frac{1}{2}\} > \lambda] d\lambda$ and again by Wellner's theorem, the last integral is at most $c_4(\int_1^\infty \lambda^{p-3}\, d\lambda) E[Y_1^2] < \infty$, completing the proof of the lemma.

LEMMA 4.4. *For any* $p > 0$, $\|k^{-1}V_{n,k} - d_{n,k}\|_p = O(k^{-1/2}$ *and* $\|k^{-1}\delta_{n,k} - \theta^{-1}d_{n,k}\| = O(k^{-1/2})$.

PROOF. Since the second result follows from the first result and Lemma 4.2, we prove the first result only. From (2.11), we have

$$k^{-1}V_{n,k} - d_{n,k} = \left(\frac{n}{k}\right)\int_0^{Z_{(k)}}(H_n - H) + \frac{n}{k}\left(\int_0^{Z_{(k)}}H - \int_0^{H^{-1}(p_{n,k})}h\right) \tag{4.2}$$
$$= I + II \quad (\text{say}).$$

Define the function k on $[0,1]$ by $k(\cdot) = \int_0^{H^{-1}(\cdot)}H(x)\,dx$. By our assumptions, k is differentiable on $[0,1)$ with $k'(t) \le \theta$. So by the mean value theorem,

$$\left\|\int_0^{Z_{(k)}}H - \int_0^{H^{-1}(p_{n,k})}H\right\|p \le \theta\|\xi_{(k)} - p_{n,k}\|_p, \tag{4.3}$$

where $\xi_{(k)}$ is the kth order statistic corresponding to a random sample size n from the uniform distribution on $(0,1)$. By Lemma 2 of Wellner (1977a), $\|\xi_{(k)} - p_{n,k}\|_p = O(k^{1/2}/n)$, whence in view of (4.3) we have $II = O(k^{-1/2})$. To handle I of (4.2) note that with q as in Lemma 4.3, we have

$$\left|\int_0^{Z_{(k)}}(H_n - H)\right| \le \rho_q(\Gamma_n, I)\int_0^{Z_{(k)}}q(H). \tag{4.4}$$

Since $\int_0^\infty q(H) < \infty$, Lemma 4.3 and (4.4) yield that $I = O(k^{-1/2})$ provided $\lambda = \liminf k/n > 0$. If $\lambda = 0$, there exists a subsequence $\{n_l\}$, such that $k_{n_l}/n_l \to 0$, along which we show $I = O(k^{-1/2})$. It then follows from the usual subsequence arguments and the first part of our proof, that this same order for I obtains for all sequences $\{k_n\}$. Thus in the sequel to show $I = O(k^{-1/2})$ we assume $n^{-1}k \to 0$. Note that

$$\left|\int_0^{Z_{(k)}}(H_n - H)\right| \le \rho_q(I_n, I)\left[\xi_{(k)} > \tfrac{1}{2}\right]\left(\int_0^\infty q(H)\right) + \rho^*(\Gamma_n, I)\left(\int_0^{Z_{(k)}}H\right), \tag{4.5}$$

where $\rho^*(\Gamma_n, I) = \sup\{|\Gamma_n(t) - t|/(1-t)\colon 0 < t \le \frac{1}{2}\}$ and ρ_q as in Lemma 4.3. For the first term in (4.5), apply Lemma 4.3 together with the fact the $P[\xi_{(k)} > \frac{1}{2}]$ has an exponential rate of convergence to zero. This yields

$$\|\rho_q(\Gamma_n, I)\left[\xi_{(k)} > \tfrac{1}{2}\right]\|_p = O(n^{-1}k^{1/2}). \tag{4.6}$$

For the second term in (4.5) on noting that $\{(1-t)^{-1}(\Gamma_n(t) - t)\colon 0 < t \le \frac{1}{2}\}$ is a martingale yields $\|\rho^*(\Gamma_n, I)\|_p = O(n^{-1/2})$. Treat $\int_0^{Z_{(k)}}H$ by triangulation using (4.3) and the fact that $(n/k)\int_0^{H^{-1}(k/n)}H \to (h(0))^{-1}$. We will get $\|\int_0^{Z_{(k)}}H\|_p = O(k/n)$ and so in view of (4.5) and (4.6) the proof may be terminated.

LEMMA 4.5. *For* $p > 0$, $\|k\delta_{n,k}^{-1}[\delta_{n,k} \ge 1]\|_p = O(1)$.

PROOF. Write

$$E\left[(k^{-1}\delta_{n,k})^{-p}[\delta_{n,k} \ge 1]\right] = \sum_{j=1}^{k}(k^{-1}j)^{-p}P[\delta_{n,k} = j] \tag{4.7}$$
$$= \sum_{j < \epsilon k} + \sum_{j < \epsilon k} + \sum_{j > \epsilon k} = I + II,$$

where ε (> 0) will be selected later on. Whenever $n^{-1}k \to \lambda$ we have $d_{n,k} \to b_\lambda$ with b_λ defined in (2.5). Therefore, by the usual subsequence arguments we have $\liminf d_{n,k} \geq A(> 0)$ where $A = \inf\{b_\lambda: \lambda \in [0,1]\}$. Now choose ε, a priori such that $\theta\varepsilon < A$. Then for the sufficiently large n, $P[\delta_{n,k} < \varepsilon k] \leq P[|k^{-1}\delta_{n,k} - \theta^{-1}d_{n,k}| > d]$ for a $d > 0$. The lemma now follows from (4.7) and Lemma 4.4 upon observing that $I \leq k^p P[\delta_{n,k} < \varepsilon k] \leq k^{-p}0(k^p)$ and $II \leq \varepsilon^{-p}$.

LEMMA 4.6. *If $k/n \to \lambda \in (0,1)$, then for $p > 0$*

$$k^{-1}\sum_{i=1}^{k} E(V_{n,i}^{*p}|\mathscr{B}_{n,i-1}) \to_{L_1} \Gamma(p+1)\int_0^{H^{-1}(\lambda)}\{H(x)/h(x)\}^p h(x)\,dx.$$

PROOF. Follows from arguments similar to those in Lemma 4.1 of Gardiner (1982) on first showing that

$$\sup_{1\leq i\leq k}|E(V_{n,i}^{*p}|\mathscr{B}_{n,i-1}) - \Gamma(p+1)\{H(Z_{(i-1)})/h(Z_{(i-1)})\}^p| \to_{L_1} 0.$$

Acknowledgments. The authors wish to thank the associate editor (especially in connection with the proof of Lemma 4.3) and the referees for carefully reading the paper and for suggestions which led to an improved version of the paper.

REFERENCES

DHARMADHIKARI, S. W., FABIAN, V. and JOGDEO, K. (1968). Bounds on moments of martingales. *Ann. Math. Statist.* **39** 1719–1723.

GALAMBOS, J. (1978). *The Asymptotic Theory of Extreme Order Statistics.* Wiley, New York.

GARDINER, J. (1982). Local asymptotic normality of progressively censored likelihood ratio statistics and applications. *J. Multivariate Anal.* **12** 230–247.

GARDINER, J. C. and SUSARLA, V. (1984). Risk efficient estimation of the mean exponential survival time under random censorship. *Proc. Nat. Acad. Sci.* **81** 5906–5909.

GARDINER, J., SUSARLA, V. and VAN RYZIN, J. (1984). Time sequential estimation of the exponential mean under random withdrawals. Technical Report No. B-39, Department of Statistics, Columbia University.

HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30.

MACLEISH, D. L. (1974). Dependent central limit theorems and invariance principles. *Ann. Probab.* **2** 620–628.

SEN, P. K. (1980). On time-sequential point estimation of the mean of an exponential distribution. *Comm. Statist.* A **9** 27–38.

WELLNER, J. A. (1977a). A law of the iterated logarithm for functions of order statistics. *Ann. Statist.* **5** 481–494.

WELLNER, J. A. (1977b). A martingale inequality for empirical processes. *Ann. Probab.* **5** 303–308.

J. C. GARDINER
DEPARTMENT OF STATISTICS AND PROBABILITY
MICHIGAN STATE UNIVERSITY
EAST LANSING, MICHIGAN 48824

V. SUSARLA
DEPARTMENT OF MATHEMATICAL SCIENCES
STATE UNIVERSITY OF NEW YORK
BINGHAMTON, NEW YORK 13901

J. VAN RYZIN
DEPARTMENT OF STATISTICS
COLUMBIA UNIVERSITY
NEW YORK, N.Y. 10025