The Annals of Statistics
1980, Vol. 8, No. 5, 1002–1016

LARGE SAMPLE THEORY FOR AN ESTIMATOR OF THE MEAN SURVIVAL TIME FROM CENSORED SAMPLES

BY V. SUSARLA AND J. VAN RYZIN

University of Wisconsin and Columbia University

This paper introduces and studies the large sample properties of an estimator for the mean survival time from censored samples. Let $X_1, \cdots, X_n$ be independent identically distributed random variables with $F(x) = P[X_1 > x]$. Let $Y_1, \cdots, Y_n$ be independent identically distributed (and independent of $X_1, \cdots, X_n$) censoring times with $G(y) = P[Y_1 > y]$. Based on observing only $Z_i = \min(X_i, Y_i)$ and which observations are censored (i.e., $X_i > Y_i$), we give a class of estimators of the mean survival time $\mu = \int_0^\infty F(x)\,dx$. The estimators are of the form $\hat{\mu} = \int_0^{M_n} \hat{F}(x)\,dx$, where $M_n \uparrow \infty$ as $n \uparrow \infty$ and $\hat{F}$ is an estimator of F depending on the Z_i's and the censoring pattern. Conditions of F, G and $\{M_n\}$ for the asymptotic normality of $\hat{\mu}$ are stated and proved in Section 2 based on approximations detailed in Section 3. Section 4 gives conditions for strong consistency of $\hat{\mu}$ with rates, while Section 5 examines the meaning of the conditions for the case of the negative exponential distributions for F and G.

1. Introduction and formulation of the problem. Let $X_1, \cdots, X_n$ be i.i.d. random variables with right sided continuous distribution function F (that is, $F(x) = P[X_1 > x]$) with $F(0) = 1$, and $Y_1, \cdots, Y_n$ be i.i.d. (independent also of $(X_1, \cdots, X_n)$) random variables with right sided continuous distribution G such that $G(0) = 1$. In several survival analysis models (for example, see Breslow and Crowley (1974), Gehan (1969), Gross and Clark (1975), and Kaplan and Meier (1958), and the references cited therein), we do not observe the true survival times $X_1, \cdots, X_n$; rather we observe only right censored times, censored by $Y_1, \cdots, Y_n$ respectively. That is, we observe only

$$(1.1) \qquad (\delta_1, Z_1), \cdots, (\delta_n, Z_n)$$

where for $i = 1, \cdots, n$,

$$(1.2) \qquad \delta_i = I_{[X_i \leq Y_i]}, \quad \text{and} \quad Z_i = \min\{X_i, Y_i\},$$

In the situation described above, an important characteristic is the mean survival time $\mu = -\int_0^\infty x\,dF(x) = \int_0^\infty F(x)\,dx$ which we shall assume is finite. The purpose of this paper is to obtain estimators of μ using $(\delta_1, Z_1), \cdots, (\delta_n, Z_n)$. One obvious estimator of μ is $\int_0^\infty \hat{F}(x)\,dx$ provided the stochastic process $\hat{F}$ is a good estimator of F. Two estimators $\hat{F}$ of F are the product limit estimator of Kaplan and Meier (1958) and its Bayesian generalization of Susarla and Van Ryzin (1976). Substitut-

Received April 1977; revised February 1979.

Research supported in part by DHEW, PHS, National Institutes of Health under Grant 5 R01 GM 23129 and by NSF under Grant No. MCS 76-05952, and revision supported by NSF Grant No. MCS77-26809 and Grant 7 R01 GM 28405.

AMS 1970 *subject classifications.* Primary 62G05; secondary 62E20.

Key words and phrases. Mean survival time estimation, randomly right censored samples, strong consistency, asymptotic normality, survival analysis, and *U*-statistics.

ing the estimator $\hat{F}_\alpha$ of Susarla and Van Ryzin (1978) for F in $\int_0^\infty F(x)\,dx$ gives an estimator (of μ) whose asymptotic properties are difficult to study since $\lim_{t\to\infty} \lim_{n\to\infty} \operatorname{Var}(\hat{F}_\alpha(t)) = \infty$. In this paper, we consider estimators of the form,

$$\hat{\mu} = \int_0^{M_n} \hat{F}(x)\,dx \tag{1.3}$$

where $M_n \uparrow \infty$ as $n \uparrow \infty$ and $\hat{F}$ is an estimator of F, and obtain conditions on F, G, and $\{M_n\}$ so that $\hat{\mu} \to \mu$ almost surely, and $\hat{\mu}$ is asymptotically normal. A method similar to the method given here for the almost sure convergence of $\hat{\mu}$ can be adopted to obtain a mean-square convergence result for $\hat{\mu}$.

We point out that Sander (1975) obtained estimators for $\int_0^T F(x)\,dx$, T fixed, which are asymptotically normal whenever $T < \infty$ and $F(T)G(T) > 0$, and indicated that it is extremely difficult to obtain the distribution theory for the estimators of $\int_0^T F(x)\,dx$ whenever $T = \infty$ or $F(T)G(T) = 0$; in particular, estimators for $\mu = \int_0^\infty F(x)\,dx$. See Theorem 1 and the remark following it in Sander (1975). Since it is impossible to estimate $F(x)$ (and, therefore, any functional of F involving $F(x)$) whenever $G(x) = 0$ without further assumptions it is assumed throughout this paper that

$$\begin{aligned} T &= \sup\{t \mid t \text{ is in the support of } F\} \\ &\leqslant \sup\{t \mid t \text{ is in the support of } G\}. \end{aligned} \tag{A1}$$

Since the results and the methods used to prove them are similar when $T = \infty$ and when $T < \infty$, we deal only with the case $T = \infty$ from here onward. Throughout, the arguments of functions are suppressed whenever they are clear from the context. Denote the indicator function of any set A by $[A]$, convergence in probability by $\to_P$, convergence almost surely by $\to_{\text{a.s.}}$, and convergence in law by $\to_{\mathcal{L}}$. Also the following notation is used throughout.

$$H = FG; \tag{1.4}$$

$$\tilde{H}(s) = P[\delta_1 = 0, Z_1 \leqslant s] = -\int_0^s F\,dG, \qquad s \geqslant 0 \tag{1.5}$$

$$\tilde{\tilde{H}}(s) = P[\delta_1 = 1, Z_1 \leqslant s] = -\int_0^s G\,dF, \qquad s \geqslant 0; \tag{1.6}$$

$$nH_n(\cdot) = \Sigma_{i=1}^n [Z_i > \cdot], \tag{1.7}$$

and

$$n\tilde{H}_n(\cdot) = \Sigma_{i=1}^n [\delta_i = 0, Z_i \leqslant \cdot]. \tag{1.8}$$

Throughout, F^{-1}, G^{-1} and H^{-1} stand for $1/F$, $1/G$ and $1/H$, respectively.

To define the estimator $\hat{\mu}$ considered in this paper, we first need to introduce some notation. Let

$$N^+(\cdot) = \text{number of } Z_j (j = 1, \cdots, n) > \cdot, \tag{1.9}$$

and

$$\hat{F}(u) = \frac{N^+(u)}{n} \Pi_{j=1}^n \left\{ \frac{2 + N^+(Z_j)}{1 + N^+(Z_j)} \right\}^{[\delta_j = 0,\, Z_j \leqslant u]} \tag{1.10}$$

for $u > 0$. The estimator $\hat{\mu}$ is defined by

$$\hat{\mu} = \int_0^M \hat{F}(u)\, du \tag{1.11}$$

where $M = M_n \uparrow \infty$ at an appropriate rate. The motivation for defining $\hat{\mu}$ as above comes from the facts that $\int_0^M F(u)\, du \to \int_0^\infty F(u)\, du = \mu$ as $n \to \infty$, and that $\hat{F}(u)$ defined by (1.10) can be seen to be a good estimator of $F(u)$ for each u in view of the large sample results obtained in Susarla and Van Ryzin (1978). The estimator (1.10), which is a slight variation of the Bayes estimate, allows us to follow through the proofs given in Susarla and Van Ryzin (1978). These proofs depend on the validity of certain logarithmic expansions, and such expansions are valid for the modified estimator given by (1.10).

The plan of the rest of the paper is as follows: Section 2 states the asymptotic normality result for the estimator $\hat{\mu}$ defined in (1.11), and discusses the conditions of the theorem. After presupposing the needed approximations which are proved in Section 3, we give a proof for the asymptotic normality result for $\hat{\mu}$ in Section 2. The asymptotic variance calculations are relegated to Appendix A. Section 4 states theorems concerning the almost sure consistency of $\hat{\mu}$, and indications of their proofs are provided in Appendix B. Section 5 elaborates the main results of the paper in the context of exponential distributions for F and G. The paper is concluded with a few remarks including a mean square result for $\hat{\mu}$.

2. Asymptotic normality of the proposed estimator $\hat{\mu}$ of (1.11). In this section, we state the following asymptotic normality result for $\hat{\mu}$.

THEOREM 2.1. (*Asymptotic Normality of $\hat{\mu}$ of* (1.11)). *Let F, G, and M satisfy the following requirements.*

$$\sigma^2 = \int_0^\infty H^{-2}\left(\int_s^\infty F\, du\right)^2 d\tilde{H} < \infty, \tag{A2}$$

$$n^{-\frac{1}{2}} H^{-3}(M) \to 0, \tag{A3}$$

and

$$n^{-\frac{1}{2}} \int_0^M H^{-4} G^{-1}\, du \to 0. \tag{A4}$$

Then $n^{\frac{1}{2}}(\hat{\mu} - \int_0^M F\, du) \to_{\mathcal{L}} N(0, \sigma^2)$ where σ^2 is defined in (A2).

The proof of Theorem 2.1 rests upon the fact (proved in the next section via various lemmas) that $n^{\frac{1}{2}}(\hat{\mu} - \int_0^M F\, du)$ has the same asymptotic distribution as that of $(S_n - E[S_n])$ where

$$S_n = n^{\frac{1}{2}}\left\{\int_0^M (2H^{-1} - H_n H^{-2})\mu_M\, d\tilde{H}_n + \int_0^M G^{-1} H_n\, du\right\} \tag{2.1}$$

with

$$\mu_M(\cdot) = \left(\int_\cdot^M F\, du\right)[\cdot \leq M]. \tag{2.2}$$

THEOREM 2.2. *$S_n - E[S_n] \to_{\mathcal{L}} N(0, \sigma^2)$ where σ^2 is defined in* (A2) *provided $n^{-1} M^2 H^{-2}(M) \to 0$.*

PROOF. The proof involves showing that S_n is asymptotically normal by using the method of proof of Theorem 7.1 of Hoeffding (1948). By the definitions of H_n and $\tilde{H}_n$ given by (1.7) and (1.8) respectively, we have

$$\begin{aligned} n^{-\frac{1}{2}}S_n &= 2n^{-1}\Sigma_{j=1}^{n}\frac{\mu_M(Z_j)}{H(Z_j)}[\delta_j = 0, Z_j \leqslant M] + n^{-1}\Sigma_{j=1}^{n}\textstyle\int_0^{Z_j\wedge M}G^{-1}\,du \\ &\quad - n^{-2}\Sigma_{j=1}^{n}\Sigma_{k=1}^{n}[Z_k > Z_j]H^{-2}(Z_j)\mu_M(Z_j)[\delta_j = 0, Z_j \leqslant M] \\ &= \binom{n}{2}^{-1}\Sigma'\Phi_n((\delta_j, Z_j), (\delta_k, Z_k)) \end{aligned} \tag{2.3}$$

where Σ' stands for summation over all (j,k) such that $1 \leqslant j < k \leqslant n$ and

$$\begin{aligned} 2\Phi_n((\delta_1, Z_1), (\delta_2, Z_2)) &= 2\mu_M(Z_1)H^{-1}(Z_1)[\delta_1 = 0, Z_1 \leqslant M] \\ &\quad + 2\mu_M(Z_2)H^{-1}(Z_2)[\delta_2 = 0, Z_2 \leqslant M] + \textstyle\int_0^{Z_1\wedge M}G^{-1}\,du \\ &\quad + \textstyle\int_0^{Z_2\wedge M}G^{-1}\,du - \dfrac{n-1}{n}\{[Z_1 > Z_2]\mu_M(Z_2)H^{-2}(Z_2)[\delta_2 = 0, Z_2 \leqslant M] \\ &\quad + [Z_2 > Z_1]\mu_M(Z_1)H^{-2}(Z_1)[\delta_1 = 0, Z_1 \leqslant M]\}. \end{aligned} \tag{2.4}$$

We now apply Hoeffding's (1948) method (his Section 7, and in particular, proof of his Theorem 7.1) of proving the asymptotic normality for the U-statistics since $n^{-\frac{1}{2}}S_n$ can be expressed as a sum of identically distributed random variables as described by (2.3) and (2.4) for each fixed n. For applying the method of Hoeffding (1948), let

$$\Psi_{n,1}((\delta_1, Z_1)) = E[\Phi_n|(\delta_1, Z_1)], \text{ and } \Psi_{n,2} = \Phi_n \tag{2.5}$$

for each n. Now a direct computation shows that

$$\begin{aligned} 2\Psi_{n,1}(\delta_1, Z_1) &= \textstyle\int_0^{Z_1\wedge M}G^{-1}\,du + n^{-1}(n+1)\mu_M(Z_1)H^{-1}(Z_1)[\delta_1 = 0, Z_1 \leqslant M] \\ &\quad - n^{-1}(n-1)\textstyle\int_0^{Z_1\wedge M}H^{-2}\mu_M\,d\tilde{H} + E[\textstyle\int_0^{Z_2\wedge M}G^{-1}\,du] \\ &\quad + 2E[\mu_M(Z_2)H^{-1}(Z_2)[\delta_2 = 0, Z_2 \leqslant M]]. \end{aligned} \tag{2.6}$$

Also, (5.13) of Hoeffding (1948) shows that

$$\operatorname{Var}(n^{-\frac{1}{2}}S_n) = \frac{4(n-2)}{n(n-1)}\operatorname{Var}(\Psi_{n,1}) + \frac{2}{n(n-1)}\operatorname{Var}(\Psi_{n,2}), \tag{2.7}$$

where a lengthy calculation shows (see Appendix A) that

$$4\operatorname{Var}(\Psi_{n,1}) = \textstyle\int_0^M H^{-2}(\int_u^\infty F\,ds)^2\,d\tilde{\tilde{H}} + O(n^{-1}H^{-1}(M)). \tag{2.8}$$

The rest of the proof involves showing that $\operatorname{Var}(\Psi_{n,2}) = \operatorname{Var}(\Phi_n) = O(M^2H^{-2}(M))$ which in turn shows (as in the proof of Theorem 7.1 of Hoeffding (1948)) that the asymptotic distribution of $n^{-\frac{1}{2}}S_n$ is the same as that of $n^{-1}\Sigma_{j=1}^{n}2\Psi_{n,1}((\delta_j, Z_j))$ in view of the condition $n^{-1}M^2H^{-2}(M) \to 0$. That $\operatorname{Var}(\Psi_{n,2}) = O(M^2H^{-2}(M))$ is the content of the following lemma.

LEMMA 2.1. $\operatorname{Var}(\Psi_{n,2}) = \operatorname{Var}(\Phi_n) = \mathrm{O}(\mathrm{M}^2\mathrm{H}^{-2}(\mathrm{M}))$ as $n\uparrow\infty$.

PROOF. As will be shown below, each term in the right-hand side of (2.4) has a second moment bounded by a constant multiple ($c_1, c_2, \cdots$ are constants) of $M^2H^{-2}(M)$. In particular, the following inequalities hold.

$$E\big[[\delta_1 = 0, Z_1 \leq M]\mu_M^2(Z_1)H^{-2}(Z_1)\big] \leq c_1H^{-2}(M).$$

Also,

$$\begin{aligned} E\Big[\big(\textstyle\int_0^{Z_1\wedge M}G^{-1}\,du\big)^2\Big] &= -\textstyle\int_0^\infty\big(\int_0^{t\wedge M}G^{-1}\,du\big)^2\,dH(t) \\ &= -2\textstyle\int_0^M G^{-1}(v)\big(\int_0^v G^{-1}(u)\,du\big)\big(\int_v^\infty dH(t)\big)\,dv \\ &= 2\textstyle\int_0^M F(v)\big(\int_0^v G^{-1}(u)\,du\big)\,dv \\ &\leq 2\textstyle\int_0^M vF(v)/G(v)\,dv \leq c_2M^2H^{-1}(M). \end{aligned}$$

Finally, notice that

$$\begin{aligned} &E\big[[Z_1 > Z_2]\mu_M^2(Z_2)H^{-4}(Z_2)[\delta_2 = 0, Z_2 \leq M]\big] \\ &\qquad\leq -c_3\textstyle\int_0^{Z_1\wedge M}\mu_M^2H^{-3}\,d\tilde{H} \leq c_4H^{-2}(M). \end{aligned}$$

All three inequalities complete the proof of the lemma.

Returning to the proof of the theorem, we point out the intent of the above lemma is that by following the pattern of the proof of Theorem 7.1 of Hoeffding (1948) it can be shown that the asymptotic distribution of $n^{-\frac{1}{2}}S_n$ is the same as that of $n^{-1}\Sigma_{j=1}^n 2\Psi_{n,1}((\delta_j, Z_j))$, a sum of i.i.d. random variables for each fixed n. Since $4\operatorname{Var}(\Psi_{n,1}) \to \sigma^2$ (see Appendix A), by (2.8) and the condition $n^{-1}M^2H^{-2}(M) \to 0$, $n^{-1}\Sigma_{j=1}^n 2\Psi_{n,1}((\delta_j, Z_j))$ can be shown to be asymptotically normal by using the standard arguments.

In the following remarks, we discuss the conditions of Theorem 2.1 (see also the examples in Section 5).

REMARK 2.1. (A3) and (A4) of Theorem 2.1 are required in ascertaining that the asymptotic distribution of $n^{\frac{1}{2}}(\hat{\mu} - \int_0^M F\,du)$ is the same as that of $S_n - E[S_n]$ while (A2) and a condition implied by (A3) are required for Theorem 2.2 which gives the asymptotic distribution of $S_n - E[S_n]$. Observe that (A3) implies that $n^{-1}M^2H^{-2}(M) \to 0$ since $\mu = \int_0^\infty F(x)\,dx < \infty$ by assumption. The asymptotic variance σ^2 given in (A2) was conjectured by Breslow and Crowley (1974) (see (8.2) of their paper).

REMARK 2.2. It should be possible to obtain an analogue of Theorem 2.2 (and hence also Theorem 2.1) without (A2), that is, in the case when $\sigma^2 = \int_0^\infty H^{-2}(\int_s^\infty F\,du)^2\,d\tilde{H} = \infty$ since (A3) and (A4) imply that $n^{\frac{1}{2}}(\hat{\mu} - \int_0^M F\,du)$ has the same asymptotic distribution as the centered version of $n^{-1}\Sigma_{j=1}^n 2\Psi_{n,1}((\delta_j, Z_j))$ whose variance multiplied by $n = \sigma_M^2 \to \infty$. Consequently, we can show that under (A3) and (A4), $n^{-\frac{1}{2}}\Sigma_{j=1}^n\{\Psi_{n,1}((\delta_j, Z_j)) - E[\Psi_{n,1}((\delta_j, Z_j))]\}/\sigma_M \to_{\mathfrak{L}} N(0, 1)$.

REMARK 2.3. In general, the centering factor $\int_0^M F\,du$ in $n^{\frac{1}{2}}(\hat{\mu} - \int_0^M F\,du)$ of Theorem 2.1 *cannot* be replaced by $\mu = \int_0^\infty F\,du$ as the following example illustrates.

Let $F(u) = e^{-\theta u}$, $G(u) = e^{-\eta u}$, both for $u > 0$ and $\eta, \theta > 0$. Then (A2) holds iff $\theta > \eta$. Moreover (A4) $\Rightarrow$ (A3) and also that (A4) is equivalent to

$$n^{-1}e^{2(4\theta+5\eta)M} \to 0. \tag{2.9}$$

If the centering factor $\int_0^M F\,du$ is replaced by $\mu = \int_0^\infty F\,du$, we also need to have

$$ne^{-2\theta m} \to 0. \tag{2.10}$$

Obviously, if (2.9) holds, then (2.10) does not hold. Consequently, with $\theta > \eta$, the example is complete.

REMARK 2.4. In case we are interested in estimating $\int_0^{T^*} F\,du$ where $T^* < \infty$ and $F(T^*)G(T^*) > 0$, the asymptotic normality of $(\int_0^{T^*}\hat{F}\,du - \int_0^{T^*}F\,du)$ follows immediately from Theorem 2.1. This case, with $\hat{F}$ as the Kaplan-Meier estimator rather than as our (1.10), follows from Sander (1975, Theorem 1).

REMARK 2.5. In case $T < \infty$ but $F(T)G(T) = 0$ (a case not included in Remark 2.4), we chose M so that $M \uparrow T$, and (A2) through (A4) hold.

REMARK 2.6. $\hat{F}$ defined by (1.10) has all the large sample properties given in Susarla and Van Ryzin (1978, 1980) for the $\hat{F}_\alpha$ defined in that paper. In particular, for each u, $\hat{F}(u)$ is mean square consistent with rate $0(n^{-1})$, almost surely consistent with rate $o(\log n/n^{\frac{1}{2}})$ and additionally, the process $\{\hat{F}(t)|0 < t \leqslant u\}$ converges weakly to a Gaussian process, with all the results holding if $F(u)G(u) > 0$.

3. Approximation (2.1) to $n^{\frac{1}{2}}(\hat{\mu} - \int_0^M F\,du)$. We recall first that the proof of the main result of Section 2, namely Theorem 2.1, depends on the fact that $n^{\frac{1}{2}}(\hat{\mu} - \int_0^M F\,du)$ can be approximated by $S_n - E[S_n]$, with S_n defined by (2.1), for asymptotic distribution theory results. The purpose of this section is to justify rigorously this approximation or reduction. Throughout this section, we use (A3) or (A4) only.

To arrive at the desired approximation, we write $\hat{F}$ of (1.10) using (1.9) as

$$\hat{F} = H_n W_n \quad \text{where} \quad nH_n(u) = N^+\ (u). \tag{3.1}$$

Consequently, we have

$$\hat{F} - F = H_n(W_n - G^{-1}) + G^{-1}(H_n - H) \tag{3.2}$$

where H is defined by (1.4). The integral on $(0, M]$ of the second term of the right-hand side of (3.2) is easy to deal with since it is only a constant multiple of the difference between an empiric distribution function and the true distribution function.

Now consider $H_n(W_n - G^{-1})$. Note from a logarithmic expansion that

$$\begin{aligned} H_n(W_n - G^{-1}) &= H_n(e^{\ln W_n} - e^{\ln G^{-1}}) \\ &= G^{-1}\{H_n(\ln W_n - \ln G^{-1}) \\ &\quad + 2^{-1}H_n(\ln W_n - \ln G^{-1})^2 e^c\} \end{aligned} \tag{3.3}$$

for some c between 0 and $\ln W_n - \ln G^{-1}$. Hence,

$$2G|H_n(W_n - G^{-1}) - H_nG^{-1}(\ln W_n - \ln G^{-1})| = H_ne^c(\ln W_n - \ln G^{-1})^2. \tag{3.4}$$

LEMMA 3.1. *For $u \leqslant M$, $H_ne^c \leqslant 2n^{-1}(n+1)$.*

PROOF. Let α be any probability measure with support in $[M, \infty)$. Then (1.3) of Susarla and Van Ryzin (1976) shows that $(N^+(u) + 1)W_n(u) \leqslant (n+1)$ for $u \leqslant M$ since the left-hand side is $(n+1)$ times the Bayes estimate of F under squared error loss function. Therefore, $H_n(u)W_n(u) \leqslant (n+1)N^+(u)/\{n(1+N^+(u))\} \leqslant n^{-1}(n+1)$ for $u \leqslant M$. Now observe that $G^{-1}H_ne^c \leqslant H_nW_n + G^{-1}H_n \leqslant n^{-1}(n+1)(1+G^{-1})$ on $(0, M]$, since $H_n \leqslant 1$. $G \leqslant 1$ completes the proof.

In view of the above lemma, and (3.4), one has, on $(0, M]$,

$$\left|\int_0^M\left\{H_n(W_n - G^{-1}) - H_nG^{-1}(\ln W_n - \ln G^{-1})\right\} du\right| \leqslant n^{-1}(n+1)\int_0^M G^{-1}(\ln W_n - \ln G^{-1})^2\, du. \tag{3.5}$$

We now state conditions on F, G, and M under which $n^{\frac{1}{2}}$(right-hand side of (3.5)) $\to_P 0$.

LEMMA 3.2. *If*

$$n^{-\frac{1}{2}}\int_0^M H^{-4}G^{-1}\, du \to 0, \tag{A4}$$

then $n^{\frac{1}{2}}\int_0^M G^{-1}(\ln W_n - \ln G^{-1})^2\, du \to_P 0$, and hence

$$n^{\frac{1}{2}}\left|\int_0^M\left\{H_n(W_n - G^{-1}) - H_nG^{-1}(\ln W_n - \ln G^{-1})\right\} du\right| \to_P 0. \tag{3.6}$$

PROOF. The result follows immediately by making the following changes in the three lemmas of Section 2 of Susarla and Van Ryzin (1978). (1) Assume that $u \leqslant M$, and (2) take α to be a probability measure with support in $[M, \infty)$.

In view of (3.2) and (3.6), we need to consider only

$$n^{\frac{1}{2}}\int_0^M\left\{H_nG^{-1}(\ln W_n - \ln G^{-1}) + G^{-1}(H_n - H)\right\} du \tag{3.7}$$

for finding the asymptotic distribution of $n^{\frac{1}{2}}(\hat{\mu} - \int_0^M F\, du)$ where $\hat{\mu}$ is defined by (1.11). The next step in our further reduction is to replace H_n and $\ln W_n(u) = \sum_{j=1}^n[\delta_j = 0, Z_j \leqslant u]\ln\{(2 + N^+(Z_j))/(1 + N^+(Z_j))\}$ in the first term of (3.7) by H and

$$\int_0^u\left\{2H^{-1}(s) - H_n(s)H^{-2}(s)\right\} d\tilde{H}_n(s) \tag{3.8}$$

for $u \leqslant M$ respectively. Replacing H_n by H follows easily since $\sup\{n^{\frac{1}{2}}|H_n(u) - H(u)|\,|0 < u < \infty\}$ converges to a random variable in probability, and since (A4) implies that $\int_0^M G^{-1}(\ln W_n - \ln G^{-1})\, du \to_P 0$ as in the proof of Lemma 3.2, the first part of the above plan is completed thus reducing (3.7) to

$$\begin{aligned} &n^{\frac{1}{2}}\int_0^M\left\{G^{-1}H(\ln W_n - \ln G^{-1}) + G^{-1}(H_n - H)\right\} du \\ &\qquad = n^{\frac{1}{2}}\int_0^M\left\{F(\ln W_n - \ln G^{-1}) + G^{-1}(H_n - H)\right\} du. \end{aligned} \tag{3.9}$$

By making the changes suggested in the proof of Lemma 3.2, it follows from Section 4 of Susarla and Van Ryzin (1978) that the right-hand side of the above equality has the same asymptotic distribution as that of

$$(3.10)\quad n^{\frac{1}{2}}\textstyle\int_0^M G^{-1}(H_n - H)\,du + n^{\frac{1}{2}}\int_0^M F(x)\left(\int_0^x\left(\frac{n}{1+nH_n}d\tilde{H}_n - \frac{d\tilde{H}}{H}\right)\right)dx.$$

Since the assumption of finite mean implies that $\int_0^M F(x)\,dx \leqslant \int_0^\infty F(x)\,dx < \infty$, we close this section by proving the following lemma which allows replacing the integral $\int_0^x n(1+nH_n)^{-1}\,d\tilde{H}_n$ by $\int_0^x(2H^{-1} - H_nH^{-2})\,d\tilde{H}_n$ in (3.10) yielding (2.1) + constants.

LEMMA 3.3. *If* (A3) *holds, then*

$$n^{\frac{1}{2}}\sup_{0\leqslant x\leqslant M}\left(\int_0^x\left|\frac{n}{1+nH_n} - \frac{2}{H} + \frac{H_n}{H^2}\right|d\tilde{H}_n\right) \to_P 0.$$

PROOF. By algebraic manipulation, one has

$$\begin{aligned}&|n(1+nH_n(u))^{-1} - H^{-1}(u) + H^{-2}(u)(H_n(u) - H(u))|\\&\qquad\leqslant \wedge_n(1+n\wedge_n)H^{-2}(u)(1+nH_n(u))^{-1} + H^{-1}(u)(1+nH_n(u))^{-1}\end{aligned}$$

where $\wedge_n = \sup\{|H_n(u) - \mathrm{H}(u)|\,|0 < u < \infty\}$. Hence $n^{\frac{1}{2}}|n(1+nH_n(u))^{-1} - 2H^{-1}(u) + H_n(u)H^{-2}(u)| \leqslant n^{\frac{1}{2}}\wedge_n(1+n\wedge_n)H^{-2}(M)(1+nH_n(M))^{-1} + n^{\frac{1}{2}}H^{-1}(M)(1+nH_n(M))^{-1}$ for $u \leqslant M$. $E[(1+nH_n(M))^{-1}] = \sum_{l=0}^n\binom{n}{l}(l+1)^{-1}H^l(M)(1-H(M))^{n-l} \leqslant c(nH(M))^{-1}$ for some constant c, the result follows from the above inequality and the fact that $n^{\frac{1}{2}}\wedge_n$ converges in probability to a random variable since $n^{-\frac{1}{2}}H^{-3}(M) \to 0$ when (A3) holds.

4. Almost sure consistency of $\hat{\mu}$ of (1.11). In this section, we state a result concerning the almost sure (a.s.) consistency of $\hat{\mu}$ as an estimator of μ. We state here the main result and the needed lemmas for it, with the proofs of various lemmas relegated to Appendix B. Throughout, denote $\sup\{|f(x)|\,|0 < x \leqslant M\}$ by $\|f\|_M$ for any function f on $(0,\infty)$, and $c_1, c_2, \cdots$ denote constants.

We start with the following basic inequality.

(4.1)

$$\begin{aligned}|\hat{\mu} - \textstyle\int_0^M F\,du| &= |\textstyle\int_0^M \hat{F}\,du - \int_0^M F\,du| = |\int_0^M(H_nW_n - HG^{-1})\,du|\\&\leqslant MG^{-1}(M)\|H_n - H\|_M + M\|H_n(W_n - G^{-1})\|_M = I + II\end{aligned}$$

where the inequality follows by a triangle inequality after adding and subtracting the integral $\int_0^M G^{-1}H_n\,du$. Now observe that

$$(4.2)\qquad I = O\big(M(\log\log n)/n^{\frac{1}{2}}G(M)\big)$$

by the law of iterated logarithm. To deal with II, we observe that, as in Section 3,

$$II \leqslant c_1M\|G^{-1}H_n(\ln W_n - \ln G^{-1})\|_M + c_2M\|G^{-1}H_n(\ln W_n - \ln G^{-1})^2\|M$$

which can be weakened to

(4.3) $II \leqslant c_1 M \|G^{-1}(\ln W_n - \ln G^{-1})\|_M + c_2 M \|G^{-1}(\ln W_n - \ln G^{-1})^2\|_M$

since $H_n \leqslant 1$. A rate for the a.s. convergence of II to zero can be obtained by obtaining a rate for

$$M\|G^{-1}(\ln W_n - \ln G^{-1})\|_M \to 0 \quad \text{a.s.} \tag{4.4}$$

By the definition of W_n in (3.1), we have

$$\ln W_n(u) = \Sigma_{j=1}^n \Sigma_{l=1}^\infty [\delta_j = 0, Z_j \leqslant u] l^{-1}(2 + N^+(Z_j))^{-l}, \tag{4.5}$$

and also

$$\ln G^{-1}(u) = E[[\delta_1 = 0, Z_1 \leqslant u] H^{-1}(Z_1)] = \int_0^u H^{-1}\, d\tilde{H}. \tag{4.6}$$

LEMMA 4.1. *If a_n are positive constants such that*

$$\Sigma_{n=1}^\infty a_n^2 / n^2 H^4(M) < \infty, \tag{B1$'$}$$

then $\|\Sigma_{j=1}^n \Sigma_{l=2}^\infty [\delta_j = 0, Z_j \leqslant u] l^{-1}(2 + N^+(Z_j))^{-l}\|_M = O(a_n^{-1})$ a.s.

LEMMA 4.2. *If* $0 < 2\alpha < 1$, *and* $0 < 2\beta \leqslant 1 < p$,

$$\Sigma_{n=1}^\infty a_n^p / H^{2p}(M) n^{\beta p} < \infty, \tag{B1}$$

and

$$\liminf n^\alpha H(M) > 0, \tag{B2}$$

then $\|\Sigma_{j=1}^n [\delta_j = 0, Z_j \leqslant u](2 + N^+(Z_j))^{-1} - \int_0^u H^{-1}\, d\tilde{H}\|_M$

$$= 0(\max\{a_n^{-1}, \log n / n^{(1-2\alpha)/2}\}) \text{ a.s.}$$

Considering (4.5), (4.6), and the above two lemmas, we obtain that

(4.7) $\|G^{-1}(\ln W_n - \ln G^{-1})\|_M$

$$= O(G^{-1}(M) \max(a_n^{-1}, (\log n)/n^{(1-2\alpha)/2})) \text{ a.s.}$$

under (B1$'$), (B1) and (B2). Hence, we have the following theorem.

THEOREM 4.1. (*Strong consistency of $\hat{\mu}$ with rates*). *Let* (B1$'$), (B1) *and* (B2) *hold. Then*

$$\hat{\mu} - \int_0^M F\, du = O(G^{-1}(M) M \max(a_n^{-1}, (\log n)/n^{(1-2\alpha)/2})) \text{ a.s.} \tag{4.8}$$

The factor $\int_0^M F\, du$ can be replaced by $\mu = \int_0^\infty F\, du$ resulting in

THEOREM 4.2. *If* (B1$'$), (B1), *and* (B2) *hold, then*

$$\hat{\mu} - \mu = 0(\int_M^\infty F\, du, G^{-1}(M) M \max(a_n^{-1}, \log n / n^{(1-2\alpha)/2}) \text{ a.s.} \tag{4.9}$$

REMARK 4.1. (B1) with $\beta p \geqslant 2$ and $p \geqslant 4$ implies (B1$'$). We will use this implication in the examples given in the next section. (B1) and (B2) are not directly comparable to the conditions (A3) and (A4). Examples satisfying (B1$'$), (B1), and (B2) are given in the next section.

REMARK 4.2. If we are only interested in estimating $\mu^* = \int_0^{T^*} F\,du$ where $T^* < \infty$ and $F(T^*)G(T^*) > 0$, then $\hat{\mu}^* = \int_0^{T^*} \hat{F}\,du$, with $\hat{F}$ defined by (1.11), can be shown to be an a.s. consistent estimator of μ^* with $o((\log n)/n)^{\frac{1}{2}}$ by using the methods of Section 3 of Susarla and Van Ryzin (1978). This was the situation considered by Sander (1975) in her asymptotic distribution theory considerations for estimators of μ^*, with $\hat{F}$ taken to be the Kaplan-Meier (1958) estimator.

5. Choice of $M_n(= M)$ and an example. M satisfying (A3) and (A4) of Section 2, and (B1) and (B2) of Section 4 exist as will be shown here. Such a choice of M will obviously depend on F and G. But if F and G belong to exponential or gamma families of distributions, M can be chosen to be independent of any unknown parameters rather easily.

To choose $\{M_n\}$, let $l_k = k + 1$ for $k = 1, 2, \cdots, \varepsilon_k \downarrow 0$ as $k \uparrow \infty$ and $0 < 2\alpha, 2\beta < 1 < p$. The first observation we make here is that both (A3) and (A4) are implied by

$$n^{-\alpha} M_n H^{-(4 \vee 2p)}(M_n) G^{-1}(M_n) \to 0. \tag{5.1}$$

Find $n_k(> n_{k-1})$ so that

$$n_k^{-\alpha}(k+1) H^{-(4 \vee 2p)}(k+1) G^{-1}(k+1) < \varepsilon_k, \quad k = 1, 2, \cdots.$$

Such a choice of n_k is possible for each k since $N^{-\alpha} a H^{-(4 \vee 2p)}(a) G^{-1}(a) \to 0$ as $N \to \infty$ for each fixed a. Now define

$$M_n^* = k + 1 \quad \text{for} \quad n_k < n \leqslant n_{k+1}. \tag{5.2}$$

$\{M^*\}$ as defined above satisfies (5.1), and with $a = 0(n^\gamma)$, with $0 < \gamma < \beta < 1$, and $p(\beta - \gamma) > 1 + \alpha$ also satisfies (B1) of Section 4 since (5.1) implies that $H^{2p}(M_n^*)$ can not go to zero any slower than $n^{-\alpha}$. Now for each fixed M_l^*, define $N_l(> N_{l-1})$ so that for a fixed $\delta > 0$,

$$n^\alpha H(M_l^*) > \delta \quad \text{for all} \quad n \geqslant N_l. \tag{5.3}$$

Now we define our required sequence $\{M_n\}$ as follows.

$$M_n = M_l^* \quad \text{for} \quad N_l \leqslant n < N_{l+1}, \qquad l = 1, 2, \cdots. \tag{5.4}$$

It can now be checked that $\{M_n\}$ defined by (5.4) satisfies (A3) and (A4), and also satisfies (B1) with $p \geqslant 4$ and $\beta p > 2$ and (B2).

Now we consider negative exponential distributions for F and G. Let $F(u) = e^{-\theta u}$, $G(u) = e^{-\eta u}$ for $u > 0$ and $\theta, \eta > 0$. (See also Remark 2.3). It can be seen that in this situation, (A4) implies (A3) and that (A4) is equivalent to $n^{-\frac{1}{2}} e^{M(5\eta + 4\theta)} \to 0$ while (B1) with $p \geqslant 4$ and (B2) are equivalent to

$$\textstyle\sum_{n=1}^{\infty} a_n^p e^{2Mp(\theta+\eta)} / n^{\beta p} < \infty$$

with $\beta p \geqslant 2$ and for $0 < 2\alpha < 1$,

$$\liminf n^\alpha e^{-(\theta+\eta)M} > 0$$

respectively. Now by taking $M = (\ln n)^\delta$ with $0 < \delta < 1$, we see that all the above three conditions are satisfied if $a_n = 0(n^\gamma)$ with $0 < 2\gamma < 1$ and $\gamma < \beta < 1$ since $\exp\{-(\theta + \eta)p(\ln n)^\delta\} \to 0$ slower than n^{-a} for each $a > 0$. For this choice of M, Theorem 4.1 shows that $\hat{\mu} - \int_0^M F\,du = 0(n^{-a})$ a.s. for any $0 < 2a < 1$. If, on the other hand, we want to use Theorem 4.2 to obtain a rate for $\hat{\mu} - \int_0^\infty F\,du (= \hat{\mu} - \theta^{-1}) \to 0$, a.s. then M should be chosen so that

$$\int_M^\infty F\,du = e^{-\theta M} = G^{-1}(M) M \max\{a_n^{-1}, n^{(2\alpha-1)/2}\log n\} = Me^{\eta M} n^{-\gamma}$$

where $0 < 2\gamma < 1$, and α and γ are chosen so that the extreme right-hand side holds with $a_n = n^\gamma$. This gives $\ln M + (\theta + \eta)M = \gamma \ln n$. This equation is difficult to solve. If $B^* \geqslant \theta$, $\eta \geqslant B > 0$, then we take $M = \gamma(\ln n)/(B + B^*)$ in which case

$$\hat{\mu} - \theta^{-1} = \hat{\mu} - \int_0^\infty F\,du = 0(\ln n / n^{\gamma B/(B+B^*)}) \text{ a.s.} \tag{5.5}$$

If bounds on θ or η are not available, then take $M = (\ln n)/l_n$ where $l_n \to \infty$ slower than $\ln n$. In this case,

$$\hat{\mu} - \theta^{-1} = \hat{\mu} - \int_0^\infty F\,du = 0(e^{-\ln\{n^{(B/l_n^\gamma)}\}}) \text{ a.s.} \tag{5.6}$$

where $\gamma > 1$. This last rate might be too slow to be of any practical significance.

6. Concluding remarks. We first remark that one can obtain rates for the mean square consistency of $\hat{\mu}$. The proof of the following theorem involves the details of Section 2 of Susarla and Van Ryzin (1978), and proceeds as described below. We write $\hat{\mu} - \int_0^M F\,du$ as $\int_0^M (\hat{F} - F)\,du$. Consequently, $E[(\hat{\mu} - \int_0^M F\,du)^2] \leqslant M\int_0^M E[(\hat{F}(u) - F(u))^2]\,du$. Hence, by a triangle inequality, notice that $E[(\hat{\mu} - \int_0^M F\,du)^2] \leqslant 2M\int_0^M \{E[H_n^2(u)(W_n(u) - G^{-1}(u))^2 + G^{-2}(u)(H_n(u) - H(u))^2]\}\,du$. Upper bounds for $E[H_n^2(u)(W_n(u) - G^{-1}(u))^2]$ and $E[G^{-2}(u)(H_n(u) - H(u))^2]$ can now be obtained by proceeding as in Section 2 of Susarla and Van Ryzin (1978).

THEOREM 6.1.

$$nE\left[\left(\hat{\mu} - \int_0^M F\,du\right)^2\right] = 0\left(M\int_0^M H^{-4}G^{-2}\,du\right).$$

For illustrating the above theorem, consider the example of Section 5. We obtain from the above theorem that

$$E\left[\left(\hat{\mu} - \int_0^M F\,du\right)^2\right] = E\left[\left(\hat{\mu} - \frac{(1 - e^{-\theta M})}{\theta}\right)^2\right]$$
$$= 0(n^{-1}Me^{M(6\theta + 2\eta)})$$

with $M(= M_n)\uparrow\infty$. Also, if $B^* \geqslant \theta$, $\eta \geqslant B > 0$, then

$$E\left[\left(\hat{\mu} - \theta^{-1}\right)^2\right] = E\left[\left(\hat{\mu} - \int_0^\infty F\,du\right)^2\right] = 0(n^{-B/\{B+(4+\gamma)B^*\}})$$

where $2M = \ln n/\{B + (4 + \gamma)B^*\}$ with $\gamma > 0$.

Our second remark concerns σ^2 of (A2). The form of σ^2 of (A2) was conjectured by Breslow and Crowley (1974) (see (8.2) of their paper). Thus this paper rigorously shows that the asymptotic form for σ^2 is true for a particular $\hat{\mu}$, given here by (1.11).

The results presented here can be extended to estimators of the form $\int_0^M \hat{F}_\alpha(u)\,du$ where

$$\hat{F}_\alpha(u) = \frac{\alpha(u,\infty)}{\alpha(R^+) + n} \Pi_{j=1}^n \left\{ \frac{\alpha[Z_j,\infty) + N^+(Z_j) + 1}{\alpha[Z_j,\infty) + N^+(Z_j)} \right\}^{[\delta_j = 0,\, Z_j < u]}$$

is the Bayes estimator derived under squared error loss (see (1.3) of Susarla and Van Ryzin (1976)) under appropriate changes in conditions (A3), (A4), (B1), and (B2). (A3) and (A4) will be replaced by the following three conditions:

$$(1) \qquad n^{-\frac{1}{2}} \int_0^M \alpha(u,\infty) G^{-1}(u)\,du \to 0,$$

$$(2) \qquad n^{-\frac{1}{2}} H^{-3}(M) \max\{1, \alpha^{-1}(M,\infty)\} \to 0, \quad \text{and}$$

$$(3) \qquad \max\left\{ n^{-\frac{1}{2}} \int_0^M H^{-4} G^{-1}\,du,\ n^{-1} \int_0^M \alpha^{-2}(u,\infty) H^{-4}(u) G^{-1}(u)\,du \right\} \to 0.$$

(B) and (B2) will be replaced by

$$(1) \qquad \Sigma_{n=1}^\infty a_n^p n^{-\beta p} \{ H^{-2p}(M) \alpha^{-p}(M,\infty) \} < \infty \quad \text{with} \quad p \geq 4,$$

and

$$(2) \qquad \liminf n^\alpha \alpha^{-1}(M,\infty) H(M) > 0.$$

The estimator $\int_0^M \hat{F}_\alpha(u)\,du$ will also be an admissible estimator for μ under squared-error loss if α is assumed to satisfy (additionally) that its support equals $(0,\infty)$. The estimator $\hat{\mu}$ defined by (1.11) *cannot* be obtained as a specialization of the above estimator $\int_0^M \hat{F}_\alpha(u)\,du$. Two important advantages of $\hat{\mu}$ over this last estimator are: (1) its simplicity for computational purposes, and (2) the fact that weaker conditions are needed to obtain its asymptotic properties.

APPENDIX A

In this appendix, we want to obtain (2.8) where $2\Psi_{n,1}$ is defined by (2.6). Observe that

$$\begin{aligned} 4 \times \operatorname{Var}(\Psi_{n,1}) &= \operatorname{Var}\Big\{ \int_0^{Z_1 \wedge M} G^{-1}\,du + \mu_M(Z_1) H^{-1}(Z_1)[\delta_1 = 0, Z_1 \leq M] \\ &\quad - \int_0^{Z_1 \wedge M} \mu_M H^{-2}\,d\tilde{H} \\ &\quad + \frac{1}{n}\Big[\mu_M(Z_1) H^{-1}(Z_1)[\delta_1 = 0, Z_1 \leq M] + \int_0^{Z_1 \wedge M} \mu_M H^{-2}\,d\tilde{H} \Big] \Big\} \\ &= \operatorname{Var}\{A + B - C + D\}. \end{aligned}$$

Now we see that $\operatorname{Var}(D) = O(n^{-2} H^{-2}(M))$ since μ_M is bounded, and that

$$\int_0^{Z_1 \wedge M} H^{-2}\,d\tilde{H} \leq 2F^{-1}(M) G^{-1}(M) = 2H^{-1}(M)$$

where we have used the equality $d\tilde{H} = -F\,dG$. Consequently,

$$\begin{aligned} (\text{A.1}) \qquad & 4 \times \operatorname{Var}(\Psi_{n,1}) \\ &= \operatorname{Var}(A + B - C) + O(n^{-2} H^{-2}(M)) + O(n^{-1} H^{-1}(M)) \end{aligned}$$

provided $\mathrm{Var}(A + B - C)$ is bounded for large n. Observe that the boundedness of $\mathrm{Var}(A + B - C)$ and the fact that $\mathrm{Var}(D) = O(n^{-2}H^{-2}(M))$ imply that the $\mathrm{Cov}(A + B - C, D) = O(n^{-1}H^{-1}(M))$. Thus, (A2) will imply (2.8) if

$$\text{(A.2)} \qquad \mathrm{Var}(A + B - C) = \int_0^M H^{-2}\left(\int_u^\infty F\, ds\right)^2 d\tilde{\tilde{H}}$$

where $\tilde{\tilde{H}}(u) = P[\delta_1 = 1, Z_1 \leqslant u]$.

Since it can be easily seen that

$$E[B] = E[C] = \int_0^M \mu_M H^{-1}\, d\tilde{H},$$

$$\text{(A.3)} \qquad \mathrm{Var}(A + B - C) = \mathrm{Var}(A) + E[B^2] + E[C^2] + 2E[AB] - 2E[AC] - 2E[BC].$$

We calculate below each of the terms in the right-hand side of (A.3). All the calculations use integration by parts, and the integral equality $(\int_0^\infty g(t)\, dt)^2 = 2\int_0^\infty g(t)[\int_0^t g(s)\, ds]\, dt$.

$$\text{(A.4)} \qquad \mathrm{Var}(A) = -\int_0^\infty\left(\int_0^{t\wedge M} G^{-1}\, du\right)^2 dH(t) - \left(\int_0^\infty\left(\int_0^{t\wedge M} G^{-1}\, du\right) dH(t)\right)^2$$
$$= 2\int_0^M F(v)\int_0^v G^{-1}(u)\, du\, dv - 2\int_0^M F(v)\int_0^v F(u)\, du\, dv$$

$$\text{(A.5)} \qquad E[B^2] = \int_0^M \mu_M^2 H^{-2}\, d\tilde{H}$$

$$\text{(A.6)} \qquad E[C^2] = -\int_0^\infty\left(\int_0^{t\wedge M} \mu_M(s) H^{-2}(s)\, d\tilde{H}(s)\right)^2 dH(t)$$
$$= 2\int_0^M \mu_M(v) H^{-1}(v) \int_0^v \mu_M(u) H^{-2}(u)\, d\tilde{H}(u)\, d\tilde{H}(v)$$

$$\text{(A.7)} \qquad 2E[AB] = 2\int_0^M \mu_M(t)\left(\int_0^{t\wedge M} G^{-1}(u)\, du\right) H^{-1}(t)\, d\tilde{H}(t)$$
$$= 2\int_0^M\left(\int_0^{t\wedge M} G^{-1}(u)\, du\right)\left(\int_t^M F(v)\, dv\right) H^{-1}(t)\, d\tilde{H}(t)$$
$$= 2\int_0^M F(v)\int_0^v G^{-1}(u)\int_u^v H^{-1}(t)\, d\tilde{H}(t)\, du\, dv$$

$$\text{(A.8)} \qquad -2E[BC] = -2\int_0^M \mu_M(t) H^{-1}(t)\int_0^t \mu_M(s) H^{-2}(s)\, d\tilde{H}(s)\, d\tilde{H}(t)$$

$$\text{(A.9)} \qquad -2E[AC] = +2\int_0^\infty\left(\int_0^{t\wedge M} G^{-1}(u)\, du\right)\left(\int_0^{t\wedge M} \mu_M(s) H^{-2}(s)\, d\tilde{H}(s)\right) dH(t)$$
$$= +2\int_{0<u<s<M} \mu_M(s) G^{-1}(u) H^{-2}(s) \int_s^\infty dH(t)\, du\, d\tilde{H}(s)$$
$$+ 2\int_{0<s<u<M} \mu_M(s) G^{-1}(u) H^{-2}(s)\int_u^\infty dH(t)\, du\, d\tilde{H}(s)$$
$$= -2\int_{0<u<s<M} \mu_M(s) G^{-1}(u) H^{-1}(s)\, du\, d\tilde{H}(s)$$
$$- 2\int_{0<s<M} \mu_M^2(s) H^{-2}(s)\, d\tilde{H}(s)$$
$$= -2\int_{0<u<v<M} F(v) G^{-1}(u)\left(\int_u^v H^{-1}(s)\, d\tilde{H}(s)\right) du\, dv$$
$$- 2\int_{0<s<M} \mu_M^2(s) H^{-2}(s)\, d\tilde{H}(s).$$

By adding (A.4) through (A.9), we obtain that

$$\text{(A.10)}\quad \begin{aligned} &\operatorname{Var}(A + B - C)\\ &\quad = 2\int_{0<u<v<M}\{F(v)G^{-1}(u) - F(u)F(v)\}\,du\,dv - \int_0^M \mu_M^2 H^{-2}\,d\tilde{H}. \end{aligned}$$

Since

$$\begin{aligned} -\int_0^M \mu_M^2 H^{-2}\,d\tilde{H} &= -2\int_0^M\Big(\int_s^M F(u)\int_u^M F(v)\,dv\,du\Big)H^{-2}(s)\,d\tilde{H}(s)\\ &= 2\int_{0<u<v<M} F(u)F(v)\int_0^u(-H^{-2})\,d\tilde{H}\\ &= 2\int_{0<u<v<M} F(u)F(v)\Big\{1 - F^{-1}(u)G^{-1}(u) + \int_0^u H^{-2}\,d\tilde{\tilde{H}}\Big\}\,du\,dv, \end{aligned}$$

(A.10) gives (A.2) completing the proof for (2.8).

APPENDIX B

In this appendix, we provide proofs for Lemmas 4.1 and 4.2 of Section 4.

PROOF OF LEMMA 4.1. Since

$$\text{(B.1)}\quad \begin{aligned} &\Big\|\sum_{j=1}^n\sum_{l=2}^\infty[\delta_j = 0, Z_j \leqslant u]\,l^{-1}(2 + N^+(Z_j))^{-l}\Big\|_M\\ &\quad\leqslant \sum_{j=1}^n\big\|[\delta_j = 0, Z_j \leqslant u](1 + N^+(Z_j))^{-1}(2 + N^+(Z_j))^{-1}\big\|_M \end{aligned}$$

by a weakening leading to a geometric series. We have

$$\text{(B.2)}\qquad \text{left-hand side of (B.1)} \leqslant n(1 + n^+(M))^{-2}.$$

Now let ε be a fixed positive number. Then

$$\text{(B.3)}\quad \begin{aligned} P[a_n(\text{left-hand side of (B.1)}) \geqslant \varepsilon] &\leqslant P[a_n(\text{right-hand side of (B.2)}) \geqslant \varepsilon]\\ &\leqslant \varepsilon^{-2}(na_n)^2 E\big[(1 + N^+(M))^{-4}\big] \end{aligned}$$

by Markov's inequality. Since

$$\begin{aligned} &E\big[(1 + N^+(M))^{-4}\big]\\ &\quad = \sum_{k=0}^n\binom{n}{k}(k+1)^{-4}H^k(M)(1 - H(M))^{n-k} \leqslant C(nH(M))^{-4} \end{aligned}$$

for some constant C, (B.3) gives the result in Lemma 4.1 in view of (B1′).

PROOF OF LEMMA 4.2. Observe that

$$\text{(B.4)}\quad \sum_{j=1}^n[\delta_j = 0, Z_j \leqslant u](2 + N^+(z_j))^{-1} = \int_0^u n(2 + nH_n(s))^{-1}\,d\tilde{H}_n(s)$$

where $nH_n(S) = \sum_{j=1}^n[Z_j > s]$, and $n\tilde{H}_n(s) = \sum_{j=1}^n[\delta_j = 0, Z_j \leqslant s]$. Hence, by a triangle inequality,

$$\text{(B.5)}\quad \begin{aligned} &\Big\|\sum_{j=1}^n[\delta_j = 0, Z_j \leqslant u](2 + N^+(Z_j))^{-1} - \int_0^u H^{-1}\,d\tilde{H}\Big\|_M\\ &\quad\leqslant \|n(2 + nH_n)^{-1} - H^{-1}\|_M + \Big\|\int_0^u H^{-1}\,d(\tilde{H}_n - \tilde{H})\Big\|_M\\ &\quad\leqslant \frac{n\|H_n - H\|_M + 2}{H(M)(2 + nH_n(M))} + \frac{2\|\tilde{H}_n - \tilde{H}\|_M}{H(M)} = I + II \end{aligned}$$

where the second inequality follows by applying integration by parts to the second term.

Since $E[\| H_n - H \|_M^q] = O(n^{-q/2})$ for any $q > 0$ due to Lemma 2 of Dvoretzky, Kiefer, and Wolfowitz (1956), and since $E[(2 + nH_n(M))^{-q}] \leq c_q(nH(M))^{-q}$ for some constant c_q for any $q > 0$, (B1) implies $I = O(a_n^{-1})$ a.s. because $a_n^{-p}\varepsilon^p P[a_n I \geq \varepsilon] \leq E[I^p] \leq c/n^{p/2}H^{2p}(M)$ for some constant c. Also, by following the proof of Theorem 1 of Singh (1975), it can be shown that $\| \tilde{H}_n - \tilde{H} \|_M = O(\log n/n^{\frac{1}{2}})$ a.s. Consequently, $II = O(\log n/n^{(1-2\alpha)/2})$ under (B2). In view of inequality (B.5), the proof of the lemma is complete.

Acknowledgments. The authors wish to express their appreciation to Professor Herman Rubin for a discussion relating to the proof of Theorem 2.2 via the asymptotic distribution theory for U-statistics. The authors are extremely grateful to the referee for the highly critical examination of the original draft of the paper. In particular, the referee's constructive criticism, detailed comments, suggestions for improvements, and pointing out the missing terms in the proofs of Lemmas 3.3 and 4.1 which required rewriting of Sections 4 and 5, are gratefully acknowledged.

REFERENCES

[1] BRESLOW, N. and CROWLEY, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *Ann. Statist.* **2** 437–453.

[2] DVORETZKY, A., KIEFER, J. and WOLFOWITZ, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Statist.* **27** 642–669.

[3] EFRON, B. (1967). The two sample problem with censored data. *Proc. Fifth Berkeley Symp. Math. Statist. Probability* **4** 831 – 853. Univ. California Press.

[4] GEHAN, E. (1969). Estimating survival function from the life table. *J. Chron. Dis.* **21** 629–644.

[5] GROSS, ALAN J. and CLARK, VIRGINIA A. (1975). *Survival Distributions: Reliability Applications in Biomedical Sciences*. Wiley, New York.

[6] HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.

[7] KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481.

[8] SANDER, J. (1975). Asymptotic normality of linear combinations of functions of order statistics with censored data. Tech. Report No. 8, Div. Biostatist., Stanford Univ.

[9] SINGH, RADHEY S. (1975). On the Glivenko-Cantelli theorem for weighted empiricals based on independent random variables. *Ann. Probability* **3** 371–374.

[10] SUSARLA, V. and VAN RYZIN, J. (1976). Nonparametric Bayesian estimation of survival curves from incomplete observations. *J. Amer. Statist. Assoc.* **61** 897–902.

[11] SUSARLA, V. and VAN RYZIN, J. (1978). Large sample theory for a Bayesian nonparametric survival curve estimator based on censored samples. *Ann. Statist.* **6** 755–768.

[12] SUSARLA, V. and VAN RYZIN, J. (1980). Addendum to Large sample theory for a Bayesian nonparametric survival curve estimator based on censored samples. *Ann. Statist.* **8** 693.

DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF WISCONSIN
P.O. BOX 413
MILWAUKEE, WISCONSIN 53201

DIVISION OF BIOSTATISTICS
DEPARTMENT OF MATHEMATICAL STATISTICS
COLUMBIA UNIVERSITY
NEW YORK, NEW YORK 10032