The Annals of Statistics
1979, Vol. 7, No. 4, 920–924

CONFIDENCE BANDS FOR THE KAPLAN-MEIER SURVIVAL CURVE ESTIMATE

BY MARY JO GILLESPIE AND LLOYD FISHER

University of Washington

The weak convergence of the Kaplan-Meier estimate for censored survival data to a Gaussian process is used to construct asymptotic confidence bands for the survival curves. The method involves transforming to obtain Brownian motion and using straight line boundaries and hitting probabilities for Brownian motion.

1. Introduction and background.

A. *The Kaplan-Meier estimate.* In medical follow-up studies of survival, data are usually analyzed with some fraction of the study-population still alive. The true survival time is thus right censored. An estimate of the underlying survival curve is given by Kaplan and Meier [6]. The Kaplan-Meier (K-M) estimate is often called the product limit estimate because it is the limit of the life-table method [2] of survival function estimation (which is based on the product of estimated conditional probabilities) as the partitioning of time intervals increases so that the maximum width approaches zero. It is assumed that there will be no ties, since the underlying density function is assumed to be continuous.

Given any sample of N individuals observed from time 0 to some time T, let the true survival time of each individual be X_i, with the distribution function $F^*(x) = P(X_i \leqslant x)$. Let the survival function be $F(x) = 1 - F^*(x)$, such that $F(0) \equiv 1$. Independent of the survival variables X_i, let the censoring variables be Y_i with distribution function $H^*(y) = P(Y_i \leqslant y)$. $H^*(y)$ is the probability that the individual is censored by time y. Let $H(y) = 1 - H^*(y)$.

Let the observed survival times be denoted as t_i, where $t_i = X_i$ when $X_i \leqslant Y_i$ and $X_i < T$. The Kaplan-Meier estimate $F_N(t)$ is computed as follows: given the values t_i, rank them in ascending order $t_1 \leqslant t_2 \leqslant \cdots \leqslant t_k$, where, without loss of generality, we can reindex them so that $t_i \leqslant t_{i+1}$ and where there are k observed deaths among the N individuals in the sample. Just prior to each death at time t_i there is some number of survivors $n_{t_i} = N$ minus $(i-1)$ deaths minus the number of censored individuals to this point. The estimated conditional probability of surviving beyond time t_i is $P_i = n_{t_{i-1}}/n_{t_i}$, i.e., the ratio of the number surviving beyond t_i to the number surviving just prior to t_i. The Kaplan-Meier estimate of the survival curve is the product of these conditional probabilities $F_N(t_i) = \prod_{j=1}^{i} P_j = F_N(t)$ for $t_i \leqslant t < t_{i+1}$.

Received August 1977; revised June 1978.

AMS 1970 *subject classifications.* Primary 62E20; secondary 62G15, 62E25.

Key words and phrases. Kaplan-Meier estimate, confidence band, survival curve, Monte Carlo study.

B. *Weak convergence to a Gaussian process.* Breslow and Crowley [3], Theorem 5, show that when F^* and H^* are continuous and $T < \infty$ satisfies the condition that the underlying distribution of the observed survival values is less than 1 at $t = T$, $N^{\frac{1}{2}}(F_N(t) - F(t))$ converges weakly to a Gaussian process G with mean 0 and covariance function

$$\operatorname{Cov}(G(s), G(t)) = F(s)F(t)a(s),\ s \leqslant t \qquad \text{where} \quad a(s) = \textstyle\int_0^s \frac{-dF}{F^2H}.$$

Efron [4] noted that

$$N^{\frac{1}{2}}(F_N(a^{-1}(s)) - F(a^{-1}(s)))/F(a^{-1}(s))$$

converged weakly to Brownian motion or the Weiner process, $W(s)$.

C. *Brownian hitting probabilities.* Anderson [1] gives a series expansion for the probability that $W(t)$ hits one of the nonintersecting straight lines $c_1 + d_1t$, $c_2 + d_2t$ (where $c_1 < 0$ and $c_2 > 0$) before time T. The probability is denoted by $P(c_1, d_1, c_2, d_2, T)$. Let Φ be the $N(0, 1)$ cumulative distribution function. If $c = -c_1 = c_2$ and $d = -d_1 = d_2$ the probability is: $P(c_1, d_1, c_2, d_2, T) = P(c, d, T)$

$$= 1 + \Sigma_{s=-\infty}^{\infty}(-1)^{s+1}e^{-2cds^2}\left[\Phi\left(\frac{dT + 2sc + c}{T^{\frac{1}{2}}}\right) - \Phi\left(\frac{-dT + 2sc - c}{T^{\frac{1}{2}}}\right)\right]$$

$$= 2\Phi\left(\frac{-dT - c}{T^{\frac{1}{2}}}\right) + 2\Sigma_{s=1}^{\infty}(-1)^{s+1}e^{-2cds^2}\left[\Phi\left(\frac{dT + 2sc + c}{T^{\frac{1}{2}}}\right) - \Phi\left(\frac{-dT + 2sc - c}{T^{\frac{1}{2}}}\right)\right].$$

2. Asymptotic confidence bands. The results of Section 1 give the asymptotic confidence bands almost immediately.

THEOREM. *Let $c_1 < 0$, $c_2 > 0$. Let F and H be continuous with $F(T) > 0$ and $H(T) > 0$. Then*

$$\lim_{N\to\infty} P\left[\frac{F_N(t)}{1 + \frac{c_2 + d_2a_N(t)}{N^{\frac{1}{2}}}} \leqslant F(t) \leqslant \frac{F_N(t)}{1 + \frac{c_1 + d_1a_N(t)}{N^{\frac{1}{2}}}},\ 0 \leqslant t \leqslant T\right]$$

$$= P(c_1 + d_1t \leqslant W(t) \leqslant c_2 + d_2t,\ 0 \leqslant t \leqslant a(T))$$

$$= 1 - P(c_1, d_1, c_2, d_2, a(T))$$

where

$$a_N(t) \equiv \int_0^t \frac{-dF_N}{F_N^2 H_N}.$$

PROOF. Note that

$$\begin{aligned}(1)\qquad \sup_{0\leqslant t\leqslant T}|a_N(t) - a(t)| &= \sup_{0\leqslant t\leqslant T}\left|\int_0^t \frac{dF}{F^2H} - \int_0^t \frac{dF_N}{F_N^2H_N}\right| \\ &\leqslant \sup_{0\leqslant t\leqslant T}\left|\int_0^t dF_N\left(\frac{1}{F_N^2H_N} - \frac{1}{F^2H}\right)\right| \\ &\quad + \sup_{0\leqslant t\leqslant T}\left|\int_0^t \frac{d(F-F_N)}{F^2H}\right|\end{aligned}$$

which converges to zero almost surely since F_N and H_N converge to F and H almost surely in the sup norm and $F^2(T)H(T) > 0$.

The probability whose limit is being taken in the theorem is algebraically equal to

$$(2)\quad P\left(c_1 + d_1a_N(t) \leqslant N^{\frac{1}{2}}(F_N(t) - F(t))/F(t) \leqslant c_2 + d_2a_N(t), 0 \leqslant t \leqslant T\right).$$

From (1) it follows that the probability (2) is asymptotically between

$$P\left(c_1 + d_1(a(t) \mp \varepsilon) \leqslant N^{\frac{1}{2}}(F_N(t) - F(t))/F(t) \leqslant c_2 + d_2(a(t) \pm \varepsilon), 0 \leqslant t \leqslant T\right)$$

for each $\varepsilon > 0$. Thus the limit in (2) is the same as the limit in (2) with $a_N(t)$ replaced by $a(t)$. This limit is (changing variables $t = a^{-1}(s)$) equal to

$$\lim_{N\to\infty} P\left(c_1 + d_1s \leqslant N^{\frac{1}{2}}(F_N(a^{-1}(s)) - F(a^{-1}(s)))\right.$$
$$\left./F(a^{-1}(s)) \leqslant c_2 + d_2s, 0 \leqslant s \leqslant a(T)\right).$$

The weak convergence result of Efron implies that this converges to $P(c_1 + d_1s \leqslant W(s) \leqslant c_2 + d_2s, 0 \leqslant s \leqslant a(T)) = 1 - P(c_1, d_1, c_2, d_2, a(T))$ which completes the proof.

Asymptotic confidence intervals for the probability of survival at *one* fixed time using Greenwood's formula have been widely used. Thomas and Grunkemeir [7] present a Monte Carlo study which compares the Greenwood method with the likelihood ratio approach and a modification of the Greenwood approach.

3. A Monte Carlo study of the method. To investigate the small sample validity of the method, a Monte Carlo study was done. The lines of the theorem were chosen to by symmetric: $-c - dt, c + dt$ where $c > 0$. Anderson's probabilities were used to find c and d so that the true confidence $1 - \alpha'$ differed from the nominal confidence $1 - \alpha$ by less than 0.001; c and d were chosen so that if there were not censoring, the widths of the confidence band at $t = 0$ and $t_{\frac{1}{2}}$ $(F(t_{\frac{1}{2}}) = \frac{1}{2})$

were approximately equal. T was chosen to be equal to 1. Details of the computation are available in Gillespie [5]. The survival curve was $F(t) = e^{-t}$ and the censoring was uniform on the internal $[0, \tau]$. Table 1 presents the results for several combinations of N and τ. Each combination was simulated 1000 times. Each simulation was evaluated for $\alpha = 0.01, 0.05$ and 0.10. The number of curves that were outside the upper boundary of the curve are given in parentheses.

TABLE 1
Number of curves outside confidence bands.

		$\tau = 1.5$	2	4	∞	$\bar{\alpha}$	S.E.
$N = 50$	$\alpha = .01$	13(0)	26(0)	21(0)	22(0)	.0205	.0022
	.05	42(0)	61(0)	51(0)	51(1)	.0513	.0035
	.10	58(1)	91(2)	92(6)	84(4)	.0813	.0043
$N = 100$	$\alpha = .01$	20(0)	18(0)	24(0)	21(0)	.0208	.0023
	.05	57(3)	53(1)	56(2)	55(3)	.0553	.0036
	.10	93(8)	96(7)	94(6)	99(15)	.0955	.0046
$N = 200$	$\alpha = .01$	23(0)	20(0)	23(0)	24(0)	.0225	.0022
	.05	74(13)	64(10)	60(3)	70(10)	.0670	.0035
	.10	116(24)	126(29)	117(20)	132(34)	.1228	.0043

The asymmetry of the bands is obvious in that the majority of the curves which are not enclosed by the confidence bands are outside at the lower boundary. Of the total number of curves outside the bands, only 9.4% are out at the upper boundary. Notable, in particular, is the asymmetry for $\alpha = .01$, where none of the curves is out at the upper boundary. The percentage increases as α increases, 6.6% at $\alpha = .05$, 13.0% at $\alpha = .10$. The asymmetry decreases with larger α and also with larger sample size N. For $N = 50$, 2.3% of those curves outside the bands are out at the upper boundary, for $N = 100$, 6.6%, and $N = 200$, 16.8%. Limiting asymptotic theory would give 50% out at each boundary.

The observed $\bar{\alpha}$'s obtained when α is set at .01 are consistently high, while those for α set at .10 appear to be increasing with larger N. Despite the theoretically asymptotic nature of the confidence bands, it does not appear that $N = 200$ is large enough for the asymptotic theory to hold with high precision.

Gillespie [5] contains more information on the properties of the curves including information about the maximum and average widths of the curves.

REFERENCES

1. ANDERSON, T. W. (1960). A modification of the sequential probability ratio test to reduce the sample size. *Ann. Math. Statist.* **31** 165–197.
2. ARMITAGE, P. (1971). *Statistical Methods in Medical Research.* Wiley, New York.
3. 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.

4. EFRON, B. (1967). The two-sample problem with censored data. *Proc. Fifth Berkeley Symp. Math. Statist. Probability* **4** 831–853.
5. GILLESPIE, MARY JO. (1975). An investigation of the small sample behavior of asymptotic confidence bands for censored survival data. Master of Sciences Thesis, Univ. Washington.
6. KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481.
7. THOMAS, D. R. and GRUNKEMEIR, G. L. (1975). Confidence interval estimation of survival probabilities for censored data. *J. Amer. Statist. Assoc.* **70** 865–871.

DEPARTMENT OF BIOSTATISTICS, SC-32
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195