The Annals of Statistics
1986, Vol. 14, No. 4, 1597–1605

ASYMPTOTIC PROPERTIES OF THE PRODUCT LIMIT ESTIMATE UNDER RANDOM TRUNCATION

By Mei-Cheng Wang, Nicholas P. Jewell and Wei-Yann Tsai[1]

Johns Hopkins University, University of California, Berkeley and Brookhaven National Laboratory

Many authors have considered the problem of estimating a distribution function when the observed data is subject to random truncation. A prominent role is played by the product limit estimator, which is the analogue of the Kaplan–Meier estimator of a distribution function under random censoring. Wang and Jewell (1985) and Woodroofe (1985) independently proved consistency results for this product limit estimator and showed weak convergence to a Gaussian process. Both papers left open the exact form of the covariance structure of the limiting process. Here we provide a precise description of the asymptotic behavior of the product limit estimator, including a simple explicit form of the asymptotic covariance structure, which also turns out to be the analogue of the covariance structure of the Kaplan–Meier estimator. Some applications are briefly discussed.

1. Introduction. Let $(x_1, t_1), \dots, (x_n, t_n)$ be a random sample from the joint cumulative distribution function

$$K(x,t) = \int_{-\infty}^{x}\int_{-\infty}^{t} I(u \le v)\, dG(v)\, dF(u)/P, \tag{1}$$

where F and G are arbitrary continuous distribution functions, $I(\cdot)$ is the usual indicator function and $P = \iint_{u \le v} dF(u)\, dG(v)$. From the sample a primary goal is to estimate the conditional distribution function $F^*(\cdot) = F(\cdot)/F(T^*)$ for any fixed T^* that is smaller than $\sup\{t: G(t) < 1\}$. Notice that if $\sup\{x: F(x) < 1\} < \sup\{t: G(t) < 1\}$ then we may choose T^* so that $F(\cdot)/F(T^*) = F(\cdot)$.

This is a model that describes observations on a random variable X subject to random truncation where the distribution of truncation values is independent of the distribution of X. For a general description and motivation of the problem with applications see Wang and Jewell (1985) and Woodroofe (1985). When the joint distribution of (X, T) is given by (1) it can be considered as if X and T are drawn independently from two arbitrary populations with distribution functions F and G, respectively, but the sampling mechanism is such that (X, T) is included in the sample if and only if $X \le T$.

An important application of the techniques discussed in this paper arises in the estimation of the slope parameters of a standard regression model $Y = \beta Z + e$ where the dependent variable Y is subject to truncation from above (or below) at

Received April 1985; revised April 1986.

[1]Supported in part by Air Force Grant AFOSR ISSA-85-00002.

AMS 1980 *subject classification*. Primary 62G05.

Key words and phrases. Product limit, truncated data, weak convergence, covariance structure.

a fixed value K. The distribution function F of the residuals e is unspecified. Given β, suppose one wishes to estimate F. This is important for two reasons: (i) one may wish to check potential parametric descriptions of F and (ii) one may wish to adaptively estimate β through an iterative scheme that estimates F at each stage. Estimation of β for simple linear regression is discussed by Bhattacharya, Chernoff and Yang (1983) using rank based methods. They also briefly discuss estimation of F at a finite set of values. Jewell (1985) introduces an iterative scheme for estimating β that uses an estimate of F at each stage of the iteration. The estimation of F is based on the set of observed residuals from a given line, $\{e_i\}$. Each e_i is then a random observation from a truncated version of F, the truncation point being $K - \beta Z_i$. Estimation of F under these circumstances is exactly the problem discussed in this paper with the truncation distribution in this case being generated by the marginal distribution of Z.

Based on the sample $(x_1, t_1), \dots, (x_n, t_n)$, Jewell (1985) and Woodroofe (1985) independently considered the following estimator for F^*:

$$F^*_{\mathrm{PL}}(x) = \prod_{x < x_j \leq T^*} \frac{m(x_j) - 1}{m(x_j)}, \tag{2}$$

where $m(x) = \sum_{j=1}^{n} I(x_j \leq x, t_j \geq x)$, and the product is taken to be 1 if there is no x_j with $x < x_j \leq T^*$. This estimator had previously been suggested by Lynden-Bell (1971) in the context of a problem in astronomy. The estimator is the analogue of the product limit estimator of Kaplan and Meier (1958) for randomly censored data.

Wang and Jewell (1985) proved the uniform strong consistency of F^*_{PL} as an estimate of F^* over intervals $(-\infty, T^*]$. They also briefly discussed the weak convergence of F^*_{PL} to a Gaussian process, but left open the exact form of the covariance structure. Independently, Woodroofe (1985) obtained a weak consistency result for F^*_{PL} and showed weak convergence to a Gaussian process but also did not give the asymptotic covariance structure. Here we discuss the derivation of F^*_{PL} and describe precisely the asymptotic convergence of F^*_{PL} including a simple explicit form for the covariance structure.

In what follows, note that the derivation of the estimator F^*_{PL} and its asymptotic properties do not depend on any parametric assumption about the form of F or G. Also, the methods extend in a straightforward manner to truncation from below and thus estimation of G.

2. Preliminary comments. Consider the likelihood of the observed data points, written as a function of F and G:

$$L = \prod_{j=1}^{n} \left[dF(x_j)\, dG(t_j) \Big/ \iint_{x \leq t} dF(x)\, dG(t) \right].$$

This can be factorized as the product of the conditional likelihood of the x's given the observed values of the t's and the marginal likelihood of the t's. That is, with $F(t_j) = \int_{-\infty}^{t_j} dF(x)$, we have

$$L = \left[\prod_{j=1}^{n} dF(x_j)/F(t_j)\right]\left[\prod_{j=1}^{n} pr(t_j)\right]. \tag{3}$$

Turnbull (1976) considered maximization of the conditional likelihood, i.e., the first term in (3), with respect to F. Because he considered a more general form of truncation, he did not provide a closed form solution. Under the condition that $\sup\{x: F(x) < 1\} < \max\{t_1, \dots, t_n\}$, an algorithm, based on a self-consistency approach, was given for the estimation of F.

Writing $F^*(\cdot) = F(\cdot)/F(T^*)$ we can separate out from the conditional likelihood those terms that have information on F^*. That is, ordering the x values so that $x_1 \le x_2 \le \cdots \le x_n$, we have

$$CL = \prod_{j=1}^{n} dF(x_j)/F(t_j) = \left[\prod_{j=1}^{m} dF^*(x_j)/F^*(t_j)\right]\left[\prod_{j=m+1}^{n} dF(x_j)/F(t_j)\right]. \tag{4}$$

Here $m = \#\{x_j \le T^*\}$, and the second factor only contains information on $F(x)$ for $x > T^*$. Write $CL^* = \prod_{j=1}^{m} dF^*(x_j)/F^*(t_j)$.

Maximizing CL^* with respect to F^* is a special case of a general problem considered by Vardi (1985). He discusses the nonparametric maximum likelihood estimator F^*_{ML}, i.e., the choice of F^* that maximizes the value of CL^* subject to the constraint that F^* is a distribution function. Before describing his results we require some notation. For each i, $1 \le i \le m$, let $D_i = \{x_j: x_j < t_i\}$ and assume no ties amongst the observed t's. Then a straightforward application of Theorem 1 of Vardi (1985) yields the following result.

THEOREM 1 (Vardi). *A necessary and sufficient condition for the existence of a unique nonparametric maximum likelihood estimator F^*_{ML} is that for each proper subset B of $\{1, \dots, m\}$, the set $D_B \equiv \bigcup_{i \in B} D_i$ contains at least one x_j with $j \notin B$.*

It is easy to construct examples of data configurations that fail to satisfy the condition. If the existence condition is satisfied, Vardi (1985) gives an algorithm for computing the estimator, although no results on asymptotic convergence are available. No convergence results are proved in Turnbull (1976) either. Note that the theorem can easily be extended to cover the case where ties occur amongst the truncation values.

Recall that, for $x \le T^*$, $m(x) = \#\{t_j \ge x \text{ with } x_j \le x;\ j = 1, \dots, m\}$ and put $a_j = [m(x_j) - 1]/m(x_j)$ for $j = 1, \dots, m$. A careful analysis of the maximization of CL^* that yields F^*_{ML} establishes the following facts. First F^*_{PL} and F^*_{ML} coincide when F^*_{ML} exists. A necessary and sufficient condition for this is that $a_j > 0$ for $j = 2, \dots, m$. It is straightforward to show that this is equivalent to the condition of Theorem 1.

Despite the independence of X and T, the marginal likelihood of the t's, i.e., the second factor of (3), does depend on F in contrast to the situation with randomly censored data. The marginal likelihood of T is given by

$$ML(t) = dG(t)F(t) / \iint_{x \le t} dF(x)\, dG(t).$$

Thus the observed values of T do carry some information about the distribution of X; however, the information is small. In fact if we observe $T = t$, all we can infer regarding F, in the absence of knowledge of G, is that $\inf\{x\colon F(x) > 0\} \le t$. Nevertheless, this suggests maximizing the full likelihood (3) simultaneously with respect to F and G. This turns out not to make any difference, i.e., when the unique nonparametric conditional maximum likelihood estimate, F^*_{ML}, exists it is also the unique nonparametric maximum likelihood estimator of the full likelihood (3) (Wang, (1986)).

All of the above has implicitly assumed that there are no ties amongst the observed x's. It is easy to see that the estimator directly extends when ties occur. Explicitly we have in this case:

$$F^*_{\mathrm{PL}}(x) = \prod_{x < x_j \le T^*} \frac{m(x_j) - d(x_j)}{m(x_j)},$$

where the product is taken over distinct x_j's and $d(x_j)$ is the number of tied values at $x = x_j$.

3. A representation for F^*. In this section we describe a representation for F^* in terms of the sampling distributions of observed random variables. It is this representation that allows us to develop the large sample properties of F^*_{PL}. For simplicity we assume that F is continuous. The extension of the results to arbitrary F is discussed at the end of the section.

Let $A = \{X \le \min(T, T^*)\}$, and let $P^* = \mathrm{pr}(A)$. Define

$$\begin{aligned} H(x,t) &= \mathrm{pr}(X \le x, T \ge t \mid (X,T) \in A) \\ &= \int_{-\infty}^{\min(x,T^*)} \int_t^{\infty} I(u \le v)\, dG(v)\, dF(u)/P^*, \end{aligned} \tag{5}$$

and let $R(x) = H(x,x)$.

Also, set

$$C(x) = \mathrm{pr}(X \le x \mid (X,T) \in A) = \int_{-\infty}^{\min(x,T^*)} \int_{-\infty}^{\infty} I(u \le v)\, dG(v)\, dF(u)/P^*$$

and $\Lambda(x) = \int_x^{T^*} (dF(u)/F(u))$. Note that $C(T^*) \equiv 1$.

It is easily seen from (5) that

$$\Lambda(x) = \int_x^{T^*} \frac{dC(u)}{R(u)}. \tag{6}$$

Furthermore,

$$F^*(x) = \exp[-\Lambda(x)]. \tag{7}$$

A natural estimator of F^* can be constructed by substituting into (7) the estimator of Λ given by replacing R and C in (6) by their empirical counterparts, i.e., we can estimate Λ by

$$\Lambda_m(x) = \int_x^{T^*} \frac{dC_m(u)}{R_m(u)}, \tag{8}$$

where $C_m(u) = m^{-1}\Sigma_{j=1}^m I(x_j \le u)$ and $R_m(u) = m^{-1}\Sigma_{j=1}^m I(x_j \le u \le t_j) = m(u)/m$. The estimator of F^* given by $\exp[-\Lambda_m]$ is asymptotically equivalent to F_{PL}^* but is easier to study. The representation of F^* given by (6) and (7) and the estimator Λ_m were derived in Wang and Jewell (1985) using slightly different notation. Woodroofe ((1985), Theorem 1) also obtained this representation, again with different notation. This representation is analogous to one obtained for a survival distribution function when the underlying survival and censoring distributions are continuous. This representation is discussed in Breslow and Crowley (1974) and Peterson (1977).

An immediate application of the representation is that it provides a straightforward method of proving consistency properties of F_{PL}^* as an estimator of F^*. In particular the following is true.

THEOREM 2. *Assume that F is continuous and $T^* < \sup\{t: G(t) < 1\}$. Then $\sup_{u \le T^*}|F_{\mathrm{PL}}^*(u) - F^*(u)| \to 0$ with probability one as $n \to \infty$.*

Details of the proof of Theorem 2 are given in Wang and Jewell (1985). Woodroofe ((1985), Theorem 2) gives an alternative proof and extends the result to cover the case $T^* = \sup\{t: G(t) < 1\}$ at the expense of substituting weak consistency for strong consistency.

In the preceding we have assumed, for simplicity, that F is continuous. Theorem 2 can be extended to the general case by using the following representation, which is the extension of the one previously obtained:

$$F^*(x) = \exp\left[-\int_x^{T^*} dC_c(u)/R(u)\right] \cdot \prod_{z \in D}\left[1 - q(z)\left(R(z)^{-1}\right)\right]^{I(z \ge x)},$$

where C_c is the continuous component of C, D is the set of discontinuity points of C, and $q(z) = \mathrm{pr}(X = z|(X,T) \in A)$. This representation can be proved in the same manner as Corollary 2.1 in Beran (1982).

4. Weak convergence of F_{PL}^*. Before stating our main results we need a lemma, which follows immediately from the representation of Section 3 using the techniques of Breslow and Crowley (1974) and the observation that $F_{\mathrm{PL}}^*(x) = \exp\{\Sigma_{x < x_j \le T^*}\log[1 - m(x_j)^{-1}]\}$. Details of the proof are given in Wang and Jewell (1985).

LEMMA 3. *Let* $n(x) = \#\{x_j \le x \text{ with } t_j \ge T^*\}$. *Let* $\inf\{z\colon F(z) > 0\} < x \le T^* < \sup\{t\colon G(t) < 1\}$. *Then, with probability one,*

$$0 < -\log F_{\mathrm{PL}}^*(x) - \Lambda_m(x) < [m - n(x)]/[n(x)(n(x) - 1)] \quad \text{as } n \to \infty.$$

Standard results concerning convergence of empirical processes yield the next lemma.

LEMMA 4. *Define* $Y_m = m^{1/2}(C_m - C)$ *and* $Z_m = m^{1/2}(R_m - R)$. *Then* (Y_m, Z_m) *converges weakly to the bivariate Gaussian process* (Y, Z), *which has mean* **0** *and covariance structure given by*

$$\begin{aligned} \operatorname{cov}(Y(x), Y(y)) &= C(x \wedge y) - C(x)C(y), \\ \operatorname{cov}(Z(x), Z(y)) &= H(x \wedge y, x \vee y) - R(x)R(y), \\ \operatorname{cov}(Y(x), Z(y)) &= H(x \wedge y, y) - C(x)R(y), \end{aligned} \tag{9}$$

where $x \wedge y = \min(x, y)$ *and* $x \vee y = \max(x, y)$.

Theorem 5 describes the weak convergence of Λ_m to a Gaussian process. Woodroofe ((1985), Theorem 3) independently obtained the weak convergence of Λ_m using a related argument but did not describe the asymptotic covariance structure explicitly.

THEOREM 5. *Let* $R(T^*) \ne 0$ *and suppose* F *and* G *are continuous. Suppose* $\inf\{z\colon F(z) > 0\} < a \le x \le T^* < \sup\{t\colon G(t) < 1\}$. *Then the random function* $\sqrt{m}\,(\Lambda_m(x) - \Lambda(x))$ *converges weakly to the Gaussian process* $W = A - B$ *defined by*

$$A(x) = \left(\frac{Y(T^*)}{R(T^*)} - \frac{Y(x)}{R(x)}\right) + \int_x^{T^*} \frac{Y}{R^2}\,dR \quad \text{and} \quad B(x) = \int_x^{T^*} \frac{Z}{R^2}\,dC, \tag{10}$$

where (Y, Z) *is the bivariate mean* **0** *Gaussian process satisfying* (9). *Furthermore, the covariance structure of the limiting process* W *is given by*

$$\operatorname{cov}(W(x), W(y)) = \int_{x \vee y}^{T^*} \frac{dC}{R^2}. \tag{11}$$

PROOF. Using Lemmas 3 and 4, the establishment of weak convergence follows exactly the arguments in the proof of Theorem 4 of Breslow and Crowley (1974) and so the details are omitted. What remains is to evaluate the covariance structure of the limiting process W.

For $x \le y$, write $\operatorname{cov}(W(x), W(y)) = \operatorname{cov}(A(x), A(y)) + \operatorname{cov}(B(x), B(y)) - \operatorname{cov}(A(x), B(y)) - \operatorname{cov}(A(y), B(x))$, where A and B are the processes defined by (10). To evaluate each of these terms we use integration of parts repeatedly together with the relationships (6) and (9) and the fact that $C(T^*) = 1$. Alternatively these calculations are special cases of Lemma A.1 of Tsai ((1982), page

117). Such calculations yield

$$\begin{aligned}\operatorname{cov}(A(x), A(y)) &= \int_x^{T*}\int_y^{T*} \frac{d\{C(u \wedge v) - C(u)C(v)\}}{R(u)R(v)} \\ &= \int_y^{T*} \frac{dC(u)}{R^2(u)} - \Lambda(x)\Lambda(y);\end{aligned} \tag{12}$$

$$\begin{aligned}\operatorname{cov}(B(x), B(y)) &= \int_x^{T*}\int_y^{T*} \frac{[H(u \wedge v, u \vee v) - R(u)R(v)]}{R^2(u)R^2(v)}\, dC(u)\, dC(v) \\ &= \int_y^{T*}\int_x^{v} \frac{H(u,v)}{R^2(u)R^2(v)}\, dC(u)\, dC(v) \\ &\quad + \int_y^{T*}\int_v^{T*} \frac{H(v,u)}{R^2(u)R^2(v)}\, dC(u)\, dC(v) \\ &\quad - \int_y^{T*}\int_x^{T*} \frac{dC(u)\, dC(v)}{R(u)R(v)} \\ &= \int_y^{T*}\int_x^{v} \frac{H(u,v)}{R^2(u)R^2(v)}\, dC(u)\, dC(v) \\ &\quad + \int_y^{T*}\int_y^{v} \frac{H(u,v)}{R^2(u)R^2(v)}\, dC(u)\, dC(v) - \Lambda(x)\Lambda(y);\end{aligned} \tag{13}$$

$$\begin{aligned}-\operatorname{cov}(A(x), B(y)) &= -\int_y^{T*}\int_x^{T*} \frac{d_u\{H(u \wedge v, v) - C(u)R(v)\}\, dC(v)}{R(u)R^2(v)} \\ &= -\int_y^{T*}\int_x^{v} \frac{d_u H(u,v)\, dC(v)}{R(u)R^2(v)} + \Lambda(x)\Lambda(y);\end{aligned} \tag{14}$$

$$\begin{aligned}-\operatorname{cov}(A(y), B(x)) &= -\int_x^{T*}\int_y^{T*} \frac{d_u\{H(u \wedge v, v) - C(u)R(v)\}\, dC(v)}{R(u)R^2(v)} \\ &= -\int_y^{T*}\int_y^{v} \frac{d_u H(u,v)\, dC(v)}{R(u)R^2(v)} + \Lambda(x)\Lambda(y).\end{aligned} \tag{15}$$

Furthermore, for $u \le v \le T^*$,

$$\begin{aligned}d_u H(u,v) &= (1 - G(v))\, dF(u)/P^* = \frac{(1 - G(v))F(u)}{P^*}\frac{dF(u)}{F(u)} \\ &= H(u,v)\frac{dC(u)}{R(u)}.\end{aligned}$$

Then $\operatorname{cov}(W(x), W(y))$ is the sum of terms (12) through (15), which is equal to $\int_y^{T*} dC(u)/R^2(u)$, completing the proof of Theorem 5. □

The following theorem is an immediate consequence of Theorem 5 using the delta method and Lemma 3.

THEOREM 6. *Assume the conditions of Theorem* 5 *hold. Then the random function* $\sqrt{m}\,(F^*_{\mathrm{PL}} - F^*)$ *converges weakly to the Gaussian process* W^* *with mean* **0** *and covariance given by*

$$\operatorname{cov}(W^*(x), W^*(y)) = F^*(x)F^*(y)\int_{x\vee y}^{T^*} \frac{dC}{R^2}. \tag{16}$$

REMARKS. (1) We can easily obtain a consistent estimate for $V(x)$, the asymptotic variance of $F^*_{\mathrm{PL}}(x)$. For $dC_m(u) = m^{-1}$, $R_m(u) = m(u)/m$ and $R_m(u^-) = (m(u) - 1)/m$. Substituting C_m for C and $R_m(u)R_m(u^-)$ for $R^2(u)$ in the expression (16) yields the following estimate for $V(x)$:

$$\hat{V}(x) = F^*_{\mathrm{PL}}(x)^2 \sum_{T^* \geq x_j \geq x} \left[m(x_j)\left(m(x_j) - 1\right)\right]^{-1}.$$

This formula is the analogue of Greenwood's formula, an estimate of the asymptotic variance of the Kaplan–Meier estimate of a survival function.

(2) Since the covariance structure (16) has a similar form to that for the Kaplan–Meier estimator, the techniques of Hall and Wellner (1980) can be adapted directly in order to construct confidence bands for F^*. In particular, with $\bar{K}(t) = \{1 + \int_t^{T^*} dC/R^2\}^{-1}$ and $K(t) = 1 - \bar{K}(t)$, the process Z, convolved with $\bar{K}/F^*$, is a rescaled Brownian bridge with K acting as a natural time scale. Consideration of this process in light of Theorem 6 leads to large-sample confidence bands for F^* in terms of $\bar{K}$ that can be estimated from data in the manner described for $V(x)$ above.

(3) Alternatives to Theorems 5 and 6 can be stated in terms of convergence as $n \to \infty$ rather than as $m \to \infty$, if we alter the definitions of R and T appropriately.

(4) Two sample tests under random truncation can be constructed by using the idea that Efron (1967) applied to the two sample problem with censored data. This yields the test statistic $W = \int_{-\infty}^{T^*} F^*_{\mathrm{PL1}}(u)\, dF^*_{\mathrm{PL2}}(u)$, where $F^*_{\mathrm{PL1}}, F^*_{\mathrm{PL2}}$ are the product limit estimators of the distributions underlying sample 1 and sample 2, respectively. The asymptotic variance of W is straightforward to deduce using Theorem 6 and can be estimated in the standard way. Other two sample tests may be derived from (i) considering Lehmann alternatives or (ii) linear rank methods as for censored data.

(5) Using techniques similar to Meier (1975) should allow our results to be extended to the case of fixed truncation values rather than random truncation.

Finally, as noted earlier, if $\sup\{x\colon F(x) < 1\} < \sup\{t\colon G(t) < 1\}$, then Theorem 6 yields the asymptotic distributional properties of F itself rather than F^* by choosing T^* suitably. Woodroofe (1985) points out that if

$$\sup\{x\colon F(x) < 1\} = \sup\{t\colon G(t) < 1\}$$

the limiting process of F^*_{PL} may not be defined if we take $T^* = \sup\{x: F(x) < 1\}$. However, in this case, a limiting distribution exists if $\int(1 - G)^{-1} dF < \infty$. For further details see Woodroofe ((1985), Section 6).

REFERENCES

BERAN, R. (1982). Nonparametric regression with randomly censored data. Unpublished report, Univ. California, Berkeley.

BHATTACHARYA, P. K., CHERNOFF, H. and YANG, S. S. (1983). Nonparametric estimation of the slope of a truncated regression. *Ann. Statist.* **11** 505–514.

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.

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

HALL, W. J. and WELLNER, J. A. (1980). Confidence bands for a survival curve from censored data. *Biometrika* **67** 133–143.

JEWELL, N. P. (1985). Least squares estimation of the slope of a truncated regression. Technical Report, Program in Biostatistics, Univ. California, Berkeley.

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

LYNDEN-BELL, D. (1971). A method for allowing for known observational selection in small samples applied to 3CR quasars. *Mon. Nat. Royal Astr. Soc.* **155** 95–118.

MEIER, P. (1975). Estimation of a distribution function from incomplete observations. *Perspectives in Prob. and Statist.* Applied Prob. Trust, Sheffield, 67–87.

PETERSON, A. V., JR. (1977). Expressing the Kaplan–Meier estimator as a function of empirical subsurvival functions. *J. Amer. Statist. Assoc.* **72** 854–858.

TSAI, W.-Y. (1982). Bivariate survival time and censoring. Ph.D. thesis, Univ. Wisconsin, Madison.

TURNBULL, B. W. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *J. Roy. Statist. Soc. Ser. B* **38** 290–295.

VARDI, Y. (1985). Empirical distributions in selection bias models. *Ann. Statist.* **13** 178–203.

WANG, M.-C. and JEWELL, N. P. (1985). The product limit estimate of a distribution function under random truncation. Technical Report, Program in Biostatistics, Univ. California, Berkeley.

WANG, M.-C. (1986). Product-limit estimates: A generalized maximum likelihood study. Unpublished.

WOODROOFE, M. (1985). Estimating a distribution function with truncated data. *Ann. Statist.* **13** 163–177.

MEI-CHENG WANG
DEPARTMENT OF BIOSTATISTICS
JOHNS HOPKINS UNIVERSITY
BALTIMORE, MARYLAND 21205

NICHOLAS P. JEWELL
DEPARTMENT OF STATISTICS AND
PROGRAM IN BIOSTATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720

WEI-YANN TSAI
APPLIED MATHEMATICS DEPARTMENT
BROOKHAVEN NATIONAL LABORATORY
UPTON, NEW YORK 11973