The Korean Communications
in Statistics Vol. 7, No. 1, 2000
pp. 245-255

Asymptotic Normality of PL estimator for interval censored bivariate life-times[1)]

Shin-Soo Kang[2)]

Abstract

Large sample properties of Life-Table estimator are discussed for interval censored bivariate survival data. We restrict our attention to the situation where response times within pairs are not distinguishable, and the univariate survival distribution is the same for any individual within any pair.

1. Introduction

Survival data consisting of independent groups of correlated response times arise from a variety of situations, such as event times collected from husband and wife pairs, siblings, litter mates, distinct components of a machine, or repeated measurements on each individual subject. In this paper, we restrict our attention to situations where response times within groups are exchangeable, and the marginal survival distributions are same for all individuals within any group. Furthermore, we consider interval censored data, in which the exact event times are not observed, only the number of failures and the number of censored individuals are observed within a finite set of time intervals. We are interested in estimating marginal survival probabilities and their variances and covariances from the observed counts.

The life-table analysis has been used widely to summarize failure time or event time data without assuming any specific parametric distributions for response times. Kaplan and Meier (1958) introduced the product limit (PL) and the actuarial (AL) estimates of survival probabilities for univariate analysis with independent failure times. The PL estimate is consistent under the assumption that censoring only occurs at the end of time intervals. Breslow and Crowley (1974) showed that a necessary and sufficient condition for the consistency of the AL estimate of a survival probability is that F^0 satisfies

$$F^0(t) = 1 - [1/(1 + cH(t))]^{.5}, \quad (1)$$

for some constant $c > 0$, where $H(t)$ is a *cdf* of *iid* censoring times and absolutely continuous

1) This research was supported by Kwandong University grants in 1999
2) Associate Professor, Department of Information Statistics, Kwandong University, Kangnung, 210-701, Korea

with density *h(t)* on a finite time interval, and $F^0(t)$ is a *cdf* of *iid* failure times. They suggest that the uniform distribution is a good approximation for the censoring distribution in many cases that satisfy (1).

Suppose we have time intervals $A_h=(t_{h-1}, t_h]$, for $h=1,2,\cdots,m$. Then, define

q_h = Pr(an individual dies in A_h | he survives beyond A_{h-1})

P_h = Pr(an individual survives beyond A_h)

The estimator of P_h, $\widehat{P_h}$ is equal to $(1-\widehat{q_1})(1-\widehat{q_2})\cdots\cdots(1-\widehat{q_h})$, where $\widehat{q_h}$ is an estimate of q_h. The PL estimator, for example, uses $\widehat{q_h}=D_h/N_h$, where D_h is the observed number of failures in A_h and N_h is the number of observations ''at risk'' at time t_{h-1}. When the data contain correlated response times, we use the same life-table procedure to get estimates of the marginal survival probabilities that was described above for the case when all response times are independent. This provides appropriate estimates of survival probabilities and large sample properties are examined for this life-table estimator.

Turnbull's (1974) non-parametric likelihood function is based on the product limit estimate with censored observations on the left and some on the right. He introduced an algorithm to get a ''self-consistent estimator'' and showed that the estimator is unique consistent maximum likelihood estimate under fairly general assumptions with existence of failures during any time interval. A concept of a ''self-consistent estimator'' was defined by Efron (1967). Campbell (1981) and Hanley and Parnes (1983) studied non-parametric maximum likelihood estimation for a bivariate survival function when the response times are interval censored. These methods assume that there is a clear distinction between each member of a pair, such as male and female siblings, right and left eyes. Campbell (1981) showed that the resulting maximum likelihood estimator is a self-consistent estimator. He also examined the existence and uniqueness of the estimator, and showed that the matrix of the second partial derivatives of the loglikelihood function is non-positive definite. Consequently, the loglikelihood is unimodal and the MLE is unique up to possible flat spots. Horvath (1983) showed the consistency of a the multivariate PL estimator, computed from multivariate exact failure times, under the assumption that the joint survival function of multivariate failure times is continuous.

There is another approach based on counting processes to prove consistency of the PL estimator. Fleming and Harrington (1991) and Andersen et al. (1993) considered *N(t)*, the number of observed failures in $[0, t]$, as univariate counting process and *Y(t)* is the number at risk just prior to time *t*. They assumed that the survivor function of a failure time random variable *T* is absolutely continuous and the number at risk, *Y(t)*, converges to ∞ as *n* goes to ∞ to prove uniform consistency of PL estimator using the Lenglart (1977) inequality.

Independence of failure and censoring times was assumed in all papers reviewed above. This assumption is also used throughout this paper.

2. Life-Table analysis

Life-table analysis is one of the oldest statistical methods used to analyze survival data and it is widely used in medical, actuarial, and industrial reliability studies. Little consideration, however, has been given to multivariate life-table analysis.

We assume members of a group are not distinguishable from each other, and the marginal event time distribution is the same for each response. Under this situation, we consider the problem of estimating the common marginal survival probabilities through a non-parametric methods using life-tables.

2.1 Random Censorship Model

The bivariate case is mainly considered in this chapter for notational convenience, but more general multivariate cases involving groups with more than two response times, or groups with different numbers of response times can be derived in a similar manner.

Let $X_i^0 = \{(X_{i1}^0, X_{i2}^0)\}$, for $i=1,2,\cdots,n$ be independent pairs of true failure times with the joint survival distribution $S(s,t) = \Pr(X_{i1}^0 > s,\ X_{i2}^0 > t)$. Let $W_i = \{(W_{i1}, W_{i2})\}$, for $i=1,2,\cdots,n$, be independent pairs of censoring variables from the joint censoring distribution $C(s,t) = \Pr(W_{i1} > s,\ W_{i2} > t)$. The variables $X_{i1}, X_{i2}, \delta_{i1}, \delta_{i2}$ are observed, where

$$X_{i1} = \min(X_{i1}^0, W_{i1}), \qquad X_{i2} = \min(X_{i2}^0, W_{i2}), \text{ and}$$

$$\delta_{ij} = \begin{cases} 1 & \text{if } X_{ij} = X_{ij}^0 \\ 0 & \text{if } X_{ij} = W_{ij}, \end{cases} \quad \text{where } j = 1,2$$

Let $G(s,t) = \Pr(X_{i1} > s, X_{i2} > t)$ be the joint distribution of (X_{i1}, X_{i2}). We will assume that (X_{i1}^0, X_{i2}^0) is independent of (W_{i1}, W_{i2}). Most procedures for life-table analysis are based on this assumption.

This bivariate model was previously considered by Campbell (1981), Dabrowska (1988), and Pruitt (1991). Campbell studied non-parametric maximum likelihood estimation for the situation where there is a clear distinction between the first member and second member of each pair. Dabrowska introduced a bivariate analogue of the Kaplan-Meier estimator, but Pruitt describes conditions under which this estimator does not yield a proper survival function.

2.2 Life-Table quantities

Consider a study consisting of $2n$ subjects, where failure times are subject to right censoring. Furthermore, consider the situation where each subject is inspected at a finite set of m times $0 < t_1 < t_2 < \cdots < t_m < \infty$, so exact response times are not observed. Instead, the responses are interval censored, i.e., it is only known whether a subject failed or was censored between two adjacent inspection times. Denote the m time intervals by $A_h = (t_{h-1}, t_h], h = 1, 2, \cdots, m$, where $t_0 = 0$. An individual is said to be ''at risk'' at time t_h if the event has not yet occurred by time t_h and the individual was not censored before time t_h. The quantities used to construct a life-table are:

$N_h =$ Number of observations "at risk" at time t_{h-1}

$$= \sum_{i=1}^{n} \{ I(X_{i1} > t_{h-1}) + I(X_{i2} > t_{h-1}) \}$$

$D_h =$ Number of failures in $A_h = (t_{h-1}, t_h]$

$$= \sum_{i=1}^{n} \{ I(X_{i1} \in A_h, \delta_{i1} = 1) + I(X_{i2} \in A_h, \delta_{i2} = 1) \}$$

$C_h =$ Number of withdrawals in $A_h = (t_{h-1}, t_h]$

$$= \sum_{i=1}^{n} \{ I(X_{i1} \in A_h, \delta_{i1} = 0) + I(X_{i2} \in A_h, \delta_{i2} = 0) \}$$

where $I(x)$ is an indicator function, which is 1 if x is true, and zero otherwise.

The conditional failure probability in time interval A_h is

$$q_h = \Pr(\text{an individual dies in } A_h \mid \text{he survives beyond } A_{h-1}).$$

The unconditional survival probability is

$$P_h = \Pr(\text{an individual survives beyond } A_h).$$

Kaplan and Meier (1958) first studied the properties of the PL(product limit) estimator and the AL(actuarial) estimator of P_h in the univariate case. The fundamental papers of Kaplan and Meier (1958) and Chiang (1961), Gilbert (1962), Efron (1967), Breslow (1969,1970), Thomas (1972), Breslow and Crowley (1974) contributed to the development of the theoretical properties of these estimators. Breslow and Crowley (1974) outlined a general theory for the large sample properties in the univariate case where any subject is assumed to respond independently of any other subject.

We will consider properties of the PL and AL estimators for P_h when the study consists of n pairs of subjects where each subject responds independently of any other subject from any other pair, but responses for subjects in the same pair can be correlated. The PL estimator ignoring pairs is defined as

$$\widehat{P}_h = (1 - \widehat{q}_1)(1 - \widehat{q}_2) \cdots\cdots (1 - \widehat{q}_h)$$

where $\widehat{q}_h = \dfrac{D_h}{N_h}$. This will be called the ''PL estimator ignoring pairs or groups'' or just the ''PL estimator''. In the next section the large sample properties of $\widehat{P}_h$ will be established.

3. Large sample properties of $\widehat{P}_h$

Let $X_{i1}, i=1,2,\cdots,n$ be n independent and identically distributed life times with survivor function $S(t)$, and let $X_{i2}, i=1,2,\cdots,n$ be *iid* life times with the same survival function $S(t)$. However X_{i1} is not necessarily independent of X_{i2}. Properties of $\widehat{P}_h$ will be examined for both the no censoring case and the censoring case.

3.1 The case with no censored observations

Establishing the consistency of $\widehat{P}_h$ is straightforward when there is no censoring. Define

$f=(f_{11},f_{12},\cdots,f_{1m},f_{21},\cdots,f_{2m},f_{31},\cdots,f_{3m},\cdots\cdots,f_{m1},\cdots,f_{mm})'$,

$\zeta=(\zeta_{11},\zeta_{12},\cdots,\zeta_{1m},\zeta_{21},\cdots,\zeta_{2m},\zeta_{31},\cdots,\zeta_{3m},\cdots\cdots,\zeta_{m1},\cdots,\zeta_{mm})'$,

where $f_{hh'}$ is the number of pairs where the first unit is observed to fail during time interval $(t_{h-1},t_h]$ and the second unit is observed to fail during time interval $(t_{h'-1},t_{h'}]$, and $\zeta_{hh'}$ is the joint failure probability corresponding to $f_{hh'}$. Then, f has a multinomial distribution with parameters n and ζ. It follows that as $n\to\infty$, the distribution of $\sqrt{n}^{-1}(f-n\zeta)$ converges to a multivariate normal distribution with mean 0 and covariance matrix, $diag(\zeta)-\zeta\zeta'$.

Now define the following quantities:

$$f_1=(u_1,u_2,u_3,\cdots,u_m);$$

where $u_h=\sum_{j=1}^{m} f_{hj}$ is the number of failures during time interval $(t_{h-1},t_h]$ experienced by the subjects listed first in the pairs, and

$$f_2=(v_1,v_2,v_3,\cdots,v_m);$$

where $v_h=\sum_{i=1}^{m} f_{ih}$ is the number of failures experienced during time interval $(t_{h-1},t_h]$ for the subjects listed second in the pairs. Also define a vector of marginal failure probabilities as

$$\pi=(\pi_1,\pi_2,\cdots,\pi_m);$$

where

$$\pi_h=\int_{t_{h-1}}^{t_h}|dS(x)| = \Pr(\text{an individual dies in}\,(t_{h-1},t_h])$$

is the failure probability for time interval $(t_{h-1}, t_h]$. These quantities can be expressed as linear functions of f or ζ as follows:

$$\begin{aligned} f_1 &= (I_{m\times m} \otimes \underline{1}')f = Af, \\ f_2 &= (\underline{1}' \otimes I_{m\times m})f = Bf, \end{aligned}$$

where

$$A_{m\times m^2} = \begin{pmatrix} 11\cdots1 & 00\cdots0 & 00\cdots0 & \cdots\cdots & 00\cdots0 \\ 00\cdots0 & 11\cdots1 & 00\cdots0 & \cdots\cdots & 00\cdots0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 00\cdots0 & 00\cdots0 & 00\cdots0 & \cdots\cdots & 11\cdots1 \end{pmatrix},$$

$$B_{m\times m^2} = \begin{pmatrix} 100\cdots0 & 100\cdots0 & 100\cdots0 & \cdots\cdots & 100\cdots0 \\ 010\cdots0 & 010\cdots0 & 010\cdots0 & \cdots\cdots & 010\cdots0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 000\cdots1 & 000\cdots1 & 000\cdots1 & \cdots\cdots & 000\cdots1 \end{pmatrix}.$$

Here, $I_{m\times m}$ denotes the identity matrix of order m and $\underline{1}=(1,1,\cdots,1)'$ is a vector of order m.

Since either member of the pair could be arbitrarily designated as the first member, the restriction, $\pi=A\zeta=B\zeta$ is imposed. Let $(f_1', f_2')'$ be f^* and $(\pi', \pi')'$be π^*. It follows that as $n\rightarrow\infty$, the distribution of $\sqrt{n}^{-1}(f^*-n\pi^*)$ converges to a multivariate normal distribution with mean 0 and covariance matrix,

$$\begin{aligned} \Sigma &= \begin{bmatrix} A \\ B \end{bmatrix}(diag(\zeta)-\zeta\zeta')[A' \ B'] \\ &= \begin{bmatrix} diag(\pi)-\pi\pi' & diag(\zeta)-\pi\pi' \\ diag(\zeta)-\pi\pi' & diag(\pi)-\pi\pi' \end{bmatrix} \end{aligned} \tag{2}$$

Now consider $\widehat{q_h}=\frac{u_h+v_h}{N_h}$, where $N_h=\sum_{i=h}^{m} u_i+\sum_{i=h}^{m} v_i$ is the number of individuals ''at risk''. Each $\widehat{q_j}$ is a smooth function of f^*. Thus $\sqrt{n}(\hat{q}-q^*)\sim N(0, n\Sigma_{q^*})$ as $n\rightarrow\infty$ and Σ_{q^*} can be obtained from the Delta method.

It immediately follows that $\widehat{q_h}$ is a consistent estimate of

$$q^*_h = \frac{\pi_h/2+\pi_h/2}{\sum_{i=h}^{m}\pi_i/2+\sum_{i=h}^{m}\pi_i/2} = \frac{\pi_h}{\sum_{i=h}^{m}\pi_i} = \frac{S(t_{h-1})-S(t_h)}{S(t_{h-1})},$$

where $S(t_h)$ is the true unknown value of the underlying marginal survivor function. Finally, $\widehat{P_h}$ is also a consistent estimate of P_h because $\widehat{P_h}$ is a continuous function of q^*_h. Let us now turn to the censored case, which is more complicated than the no censoring case.

3.2 The case with censored observations

Let X_{i1}, $i=1,2,\cdots,n$ be n independent and identically distributed failure times with survivor function $S(t)$ and let X_{i2}, $i=1,2,\cdots,n$ be *iid* failure times with the same survivor function $S(t)$. X_{i1} and X_{i2} need not to be independent. Let W_{i1}, $i=1,2,\cdots,n$ be n independent and identically distributed censoring times with survivor function $C(t)$, and let W_{i2}, $i=1,2,\cdots,n$ be *iid* censoring times with the same survivor function $C(t)$. Independence of W_{i1} and W_{i2} is not necessary but it is assumed that the censoring mechanism does not affect the true life times. Define

$$f=(f_{11},f_{12},\cdots,f_{1m},f_{21},\cdots,f_{2m},f_{31},\cdots,f_{3m},\cdots\cdots,f_{m1},\cdots,f_{mm})'$$

and

$$\zeta=(\zeta_{11},\zeta_{12},\cdots,\zeta_{1m},\zeta_{21},\cdots,\zeta_{2m},\zeta_{31},\cdots,\zeta_{3m},\cdots\cdots,\zeta_{m1},\cdots,\zeta_{mm})',$$

where $f_{hh'}$ is the number of pairs such that the first unit is observed to fail during time interval $(t_{h-1},t_h]$ and the second unit is observed to fail during time interval $(t_{h'-1},t_{h'}]$, and $\zeta_{hh'}$ is the joint failure probability corresponding to $f_{hh'}$. Define

$$d=(d_{11},d_{12},\cdots,d_{1m},d_{21},\cdots,d_{2m},d_{31},\cdots,d_{3m},\cdots\cdots,d_{m1},\cdots,d_{mm})'$$

and

$$\nu=(\nu_{11},\nu_{12},\cdots,\nu_{1m},\nu_{21},\cdots,\nu_{2m},\nu_{31},\cdots,\nu_{3m},\cdots\cdots,\nu_{m1},\cdots,\nu_{mm})',$$

where $d_{hh'}$ is the number of pairs such that the first unit is observed to fail during time interval $(t_{h-1},t_h]$ and the second unit is censored during time interval $(t_{h'-1},t_{h'}]$, and $\nu_{hh'}$ is the joint probability corresponding to $d_{hh'}$. Define

$$c=(c_{11},c_{12},\cdots,c_{1m},c_{21},\cdots,c_{2m},c_{31},\cdots,c_{3m},\cdots\cdots,c_{m1},\cdots,c_{mm})'$$

and

$$\gamma=(\gamma_{11},\gamma_{12},\cdots,\gamma_{1m},\gamma_{21},\cdots,\gamma_{2m},\gamma_{31},\cdots,\gamma_{3m},\cdots\cdots,\gamma_{m1},\cdots,\gamma_{mm})',$$

where $c_{hh'}$ is the number of pairs such that the first unit is censored during time interval $(t_{h-1},t_h]$ and the second unit is observed to fail during time interval $(t_{h'-1},t_{h'}]$, and $\gamma_{hh'}$ is the joint probability corresponding to $c_{hh'}$. Finally, define

$$g=(g_{11},g_{12},\cdots,g_{1m},g_{21},\cdots,g_{2m},g_{31},\cdots,g_{3m},\cdots\cdots,g_{m1},\cdots,g_{mm})'$$

and

$$\eta=(\eta_{11},\eta_{12},\cdots,\eta_{1m},\eta_{21},\cdots,\eta_{2m},\eta_{31},\cdots,\eta_{3m},\cdots\cdots,\eta_{m1},\cdots,\eta_{mm})',$$

where $g_{hh'}$ is the number of pairs such that the first unit is censored during time interval $(t_{h-1},t_h]$ and the second unit is censored during time interval $(t_{h'-1},t_{h'}]$, and $\eta_{hh'}$ is the joint probability corresponding to $g_{hh'}$. These vectors are combined into large vectors

$$V=(f',d',c',g')',$$

and

$$\Theta=(\zeta',\nu',\gamma',\eta')'.$$

Then, V has a multinomial distribution with parameters n and Θ. It follows from the multivariate central limit theorem that as $n \to \infty$, the distribution of $\sqrt{n}^{-1}(V - n\Theta)$ converges to multivariate normal distribution with mean 0 and covariance matrix, $diag(\Theta) - \Theta\Theta'$.

Now define the following quantities:

$$f_1 = (u_1, u_2, \cdots, u_m, w_1, w_2, \cdots, w_m)',$$

where u_h is the number of failures and w_h is the number of censored units during time interval $(t_{h-1}, t_h]$ experienced by the subjects listed first in the pairs, and

$$f_2 = (v_1, v_2, \cdots, v_m, z_1, z_2, \cdots, z_m)',$$

where v_h is the number of failures and z_h is the number of censored units experienced during time interval $(t_{h-1}, t_h]$ for the subjects listed second in the pairs.

$$\phi = (\pi_1, \pi_2, \cdots, \pi_m, \tau_1, \tau_2, \cdots, \tau_m)',$$

where π_h and τ_h are the failure probability and censoring probability, respectively such that

$$\pi_h = \Pr(t_{h-1} < X_{ij} \le t_h, X_{ij} \le W_{ij}) = \int_{t_{h-1}}^{t_h} C(x)|dS(x)|$$

and

$$\tau_h = \int_{t_{h-1}}^{t_h} S(x)|dC(x)|$$

These quantities can be expressed as linear functions of V or Θ as follows:

$$\begin{aligned} f_1 &= QV, \\ f_2 &= RV, \\ \phi &= Q\Theta = R\Theta, \end{aligned}$$

where

$$Q_{2m \times 4m^2} = \begin{pmatrix} A & A & 0 & 0 \\ 0 & 0 & A & A \end{pmatrix},$$

$$R_{2m \times 4m^2} = \begin{pmatrix} B & 0 & B & 0 \\ 0 & B & 0 & B \end{pmatrix},$$

$$A_{m \times m^2} = \begin{pmatrix} 11\cdots1 & 00\cdots0 & 00\cdots0 & \cdots\cdots & 00\cdots0 \\ 00\cdots0 & 11\cdots1 & 00\cdots0 & \cdots\cdots & 00\cdots0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 00\cdots0 & 00\cdots0 & 00\cdots0 & \cdots\cdots & 11\cdots1 \end{pmatrix},$$

$$B_{m \times m^2} = \begin{pmatrix} 100\cdots0 & 100\cdots0 & 100\cdots0 & \cdots\cdots & 100\cdots0 \\ 010\cdots0 & 010\cdots0 & 010\cdots0 & \cdots\cdots & 010\cdots0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 000\cdots1 & 000\cdots1 & 000\cdots1 & \cdots\cdots & 000\cdots1 \end{pmatrix}.$$

Let $(f_1{}', f_2{}')'$ be $f^\star$ and $(\phi', \phi')'$ be $\phi^\star$. It follows that as $n \to \infty$, the distribution of $\sqrt{n}^{-1}(f^\star - n\phi^\star)$ converges to a multivariate normal distribution with mean 0 and covariance matrix,

$$\Sigma = \begin{bmatrix} Q \\ R \end{bmatrix} (diag(\Theta) - \Theta\Theta')[Q' \ R']. \tag{3}$$

As in the no censoring case, $\widehat{q_h} = \frac{u_h + v_h}{N_h}$, where $N_h = \sum_{i=h}^{m}(u_i + v_i + w_i + z_i)$, and the $\widehat{q_j}$'s are smooth function of $f^\star$. Thus, $\sqrt{n}(\widehat{q} - q^\star) \sim N(0, n\Sigma_{q^\star})$ as $n \to \infty$and $\Sigma_{q^\star}$ can be obtained from (3) by the Delta method based.

It follows that $\widehat{q_h}$ is a consistent estimate of

$$q_h^\star = \frac{\pi_h/2 + \pi_h/2}{\sum_{i=h}^{m}(\pi_i + \pi_i + \tau_i + \tau_i)/2} = \frac{\pi_h}{\sum_{i=h}^{m}(\pi_i + \tau_i)}.$$

Since

$$\sum_{i=h}^{m}(\pi_i + \tau_i) = S(t_{h-1})C(t_{h-1}),$$

then

$$q_h^\star = \frac{\int_{t_{h-1}}^{t_h} C(x) | dS(x) |}{S(t_{h-1})C(t_{h-1})}. \tag{4}$$

In general $q_h^\star$ may not be equal to $\frac{S(t_{h-1}) - S(t_h)}{S(t_{h-1})}$, so $\widehat{q_h}$ is not necessarily a consistent estimate of q_h. Furthermore, if we have $2n$ independent observations that are subject to random censorship as described in section 2.1, $E(\widehat{q_h})$ has exactly the same formula as equation (4).

4. Discussion

The large sample properties of the actuarial(AL) estimate can be derived in a similar way. For the AL estimator, $\widehat{q_h}$ is equal to $\frac{u_h + v_h}{N_h - (w_h + c_h)/2}$ and $q_h^\star$ is

$$\frac{\int_{t_{h-1}}^{t_h} C(x) | dS(x) |}{S(t_{h-1})C(t_{h-1}) - 0.5 \int_{t_{h-1}}^{t_h} S(x) | dC(x) |}.$$

Thus, consistency properties of the PL and AL estimators in the univariate case are the same as those for the PL and AL estimators when pairs are ignored. Variances of the PL and AL estimators are affected by correlation within pair and this is derived by Kang & Koehler(1997).

Breslow and Crowley (1974) proposed a necessary and sufficient condition, relating the survival and censoring distributions, for the consistency of the AL estimator for the univariate case. They examined finite sample bias of the AL estimates for estimating the survival probability through simulation study and concluded that the bias will not be serious unless the number of intervals is fewer than ten.

Reference

[1] Andersen, P. K., Borgan, O., Gill, R. O., & Keiding, N. (1993). *Statistical models based on counting processes.* New York: Wiley.

[2] Breslow, N. E. (1969). On large sample sequential analysis with applications to survivorship data. *Journal of Applied Probability*, 6, 261-274.

[3] Breslow, N. E. (1970). A generalized Kruskal-Wallis test for comparing K samples subject to unequal patterns of censorship. *Biometrika*, 57, 579-594.

[4] Breslow, N. E., & Crowley, J. (1974). A large sample study of the life-table and product limit estimates under random censorship. *Annals of Statistics*, 2, 437-453.

[5] Campbell, G. (1981). Non-parametric bivariate estimation with randomly censored data. *Biometrika*, 68, 417-422.

[6] Chiang, C. L. (1961). A stochastic study of the life-table and its applications: The follow-up study with consideration of competing risks. *Biometrics*, 17, 57-78.

[7] Dabrowska, D. M. (1988). Kaplan-Meier estimate on the plane. *Annals of Statistics*, 16, 1475-1489.

[8] Efron, B. (1967). The two sample problem with censored data. *Proceedings of the Fifth Berkeley Symposium*, 4, 831-853.

[9] Fleming, T. R., & Harrington, D. P. (1991). *Counting processes and survival analysis.* New York: Wiley.

[10] Gilbert, J. P. (1962). Random censorship. *Ph.D. thesis, Univ. of Chicago.*

[11] Hanley, J. A. & Parnes, M. N. (1983). Non-parametric estimation of a multivariate distribution in the presence of censoring. *Biometrics,* 39, 129-139.

[12] Horvath, L. (1983). The rate of strong uniform consistency for the multivariate product-limit estimator. *Journal of Multivariate Analysis*, 13, 202-209.

[13] Kaplan, E. L., & Meier, P. (1958). Non-parametric estimation from incomplete

observations. *Journal of the American Statistical Association*, 53, 457-481.

[14] Kang, S. S. & Koehler, K. (1997). Modification of the Greenwood formula for correlated response times. *Biometrics*, 53, 885-899

[15] Lenglart, E. (1977). Relation de domination entre deux processus. *Annals of Institute Henri Poincar'e*, 13, 171-179.

[16] Pruitt, R. C. (1991). On negative mass assigned by the bivariate Kaplan-Meier estimator. *Annals of Statistics*, 19, 433-453.

[17] Thomas, D. R. (1972). On a generalized Savage statistic for comparing two arbitrarily censored samples. *Technical Report. Oregon State Univ.*

[18] Turnbull, B. W. (1974). Non-parametric estimation of a survivorship function with doubly censored data. *Journal of the American Statistical Association*, 69, 169-173.