Statistica Sinica **7**(1997), 139-148

ON TESTING OF LIFETIMES

Zheng Zukang

Fudan University and Hong Kong Baptist University

Abstract: In this paper we improve the traditional experiments of lifetime testing. A new estimator of survival function is suggested to get the full information of the individual's lifetime. It is unbiased with minimal variance among a certain class of the unbiased estimators. Simulation results are also given.

Key words and phrases: Censored data, Kaplan-Meier estimator.

1. Introduction

When experimental units are still in operation at the closing data of an investigation, and their subsequent times of failure are not known, the mortality data are incomplete. For economical consideration, the experiments can not last until all the lifetime data is to be obtained. We want to get the information about the lifetime distribution of the individual as much as possible under our budget. Termination of the experiment may be controlled in many ways. The following schemes are traditionally the two most common ways: (cf. Lawless (1982), Kalbfleisch and Prentice (1980))

Scheme 1. Run the experiments over a fixed time period in such a way that an individual's lifetime will be known exactly only if it is less than some predetermined value. In such situations the data are said to be Type 1 censored. Specifically let $X_1, \ldots, X_n$ be nonnegative independent identically distributed random variables and C a positive constant which always denotes the limit time that we could wait for. In such experiments we only observe

$$Z_i = \min(X_i, C),\ \delta_i = I(X_i \le C) \qquad (i = 1, \ldots, n),$$

where $I(X_i \le C)$ denotes the indicator function, equal to 1 if $X_i \le C$ or to 0 if $X_i > C$.

Scheme 2. Terminate taking observations when a preassigned number of deaths (say r) have occurred. Let $X_{(1)} \le \cdots \le X_{(n)}$ be the order statistics of $X_1, \ldots, X_n$. We only observe $X_{(1)} \le \cdots \le X_{(r)}$.

It is important to realize that in Scheme 1, the numbers of deaths are random variables, whereas in Scheme 2, the time of termination is a random variable. Although the two schemes provide information of the individual's lifetime, we still know nothing about the individual's lifetime after C in Scheme 1 and $X_{(r)}$ in Scheme 2. This is a serious problem. We shall fail in estimating the full lifetime distribution and the characteristics such as mean, variance in nonparametric models. The random censorship model perhaps can help us to get over the difficulties. Let $X_1, \ldots, X_n$ be nonnegative independent identically distributed random variables (lifetime) with distribution function F. Let $Y_1, \ldots, Y_n$ be nonnegative independent identically distributed random variables (censoring) with distribution function G. The two sequences are independent and we can only observe

$$Z_i = \min(X_i, Y_i),\ \ \delta_i = I(X_i \leq Y_i). \tag{1}$$

Denote $\tau_F = \inf\{t : F(t) = 1\}$, $\tau_G = \inf\{t : G(t) = 1\}$ and $\tau_H = \inf\{t : H(t) = 1\}$, where $H(t) = 1 - (1 - F(t))(1 - G(t))$ is the distribution function of Z_i. Assume that $\tau_G > \tau_F$, then we have a chance to detect the tail part of F. We also assume $E(X_i) < \infty$, $\mathrm{Var}\,(X_i) < \infty$.

Now our task is to design the distribution function G of censoring variables Y_i, then create Y_i by computer and use them instead of C in Scheme 1. First, one may think that we might pay the cost of $a \cdot n \cdot C$ in Scheme 1 where n is number of individuals and a is some constant. The new procedure stops at Z_i, with $E(Z_i) = E[\min(X_i, Y_i)] \leq E(Y_i)$. If the mean of Y_i is less or equal to C, it seems that we need not pay more in average. Secondly, to estimate the distribution function $F(t)$ or the survival function $S(t) = 1 - F(t)$ under the random censoring model, the Kaplan-Meier estimator (1958) is often suggested. It is defined by

$$\hat{S}(t) = \prod_{Z_{(i)} \leq t} \Big(1 - \frac{1}{n - i + 1}\Big)^{\delta_{(i)}}, \tag{2}$$

where $Z_{(i)}$ are the order statistics of Z_i and $\delta_{(i)}$ are the corresponding indicator functions. In our procedure, since G is known, a better estimator should be constructed. In Section 2 and Section 3, we develop a sum type estimator of $S(t)$ which is unbiased and consistent. It has minimal variance over a class of unbiased estimators. Simulation reports are also given. In Section 4, we give further results if some extra information is obtained.

2. The Estimator of $S(t)$

As we see in Section 1, the random censorship model is suggested. Instead of constant C, we create a sequence of random variables $\{Y_i\}$, which are independent and identically distributed. And the observations are $Z_i = \min(X_i, Y_i)$, $\delta_i = I(X_i \leq Y_i)$. Now there are two things that should be done. One is to determine the distribution function G of Y_i. The other is to estimate the distribution function $F(t)$ or survival function $S(t)$ of X_i. We leave the first problem to the next sectoin. Here we study the second problem–finding the best estimator of $S(t)$ in some sense for known G.

Definition. Let Φ_1, Φ_2 be two functions such that

$$\begin{cases} [1-G(x)]I(x>t)\Phi_1(t,x)+\int_0^x \Phi_2(t,y)I(y>t)dG(y)=I(x>t) \\ \\ \Phi_1,\Phi_2 \text{ are independent of } F \text{ (but may dependent on } G) \end{cases} \tag{3}$$

where $I(\cdot)$ is indicator function. We say (Φ_1, Φ_2) belongs to Class $K^*((\Phi_1, \Phi_2) \in K^*)$.

It is easy to check that if we use

$$S_n(t)=\frac{1}{n}\sum_{i=1}^{n}[I(Z_i>t)\delta_i\Phi_1(t,Z_i)+I(Z_i>t)(1-\delta_i)\Phi_2(t,Z_i)] \tag{4}$$

to estimate $S(t)$, then we have

$$ES_n(t)=S(t). \tag{5}$$

A similar idea of constructing a class of unbiased estimators is also introduced to deal with censored data in many statistical problems such as linear regression, stochastic approximation etc. (cf. Zheng Zukang (1987)) as a useful tool. There are various elements in Class K^*. We recommend $\Phi_1(t,x)=\Phi_2(t,x)=[1-G(t)]^{-1}$ which does not depend on x or y. Thus,

$$S_n^*(t)=\frac{1}{n}\sum_{i=1}^{n}I(Z_i>t)[1-G(t)]^{-1} \tag{6}$$

with

$$\operatorname{Var} S_n^*(t)=E\Big[\frac{1}{n}\sum_{i=1}^{n}\frac{I(Z_i>t)}{1-G(t)}-S(t)\Big]^2=\frac{S(t)}{n}\Big[\frac{1}{1-G(t)}-S(t)\Big]. \tag{7}$$

If $G(t)\equiv 0$, $S_n^*(t)$ coincides with the empirical survival function. The reason why we recommend (6) of the Class K^* is the following:

Theorem 1.

$$\operatorname{Var} S_n^*(t)=\inf_{[\Phi_1,\Phi_2]\in K^*}\operatorname{Var} S_n(t). \tag{8}$$

Proof. Due to the unbiasedness of the elements in Class K^*, we only need to minimize the second moment. Notice that for any $t>0$, (omitting the subscript i)

$$\begin{aligned} A &\hat{=} E[I(x>t)\delta\Phi_1(t,Z)+I(y>t)(1-\delta)\Phi_2(t,Z)]^2 \\ &= E[I(x>t)I(x\le y)\Phi_1^2(t,Z)+I(y>t)I(y<x)\Phi_2^2(t,Z)] \end{aligned}$$

$$= \iint\limits_{y \geq x > t} \Phi_1^2(t,x)dG(y)dF(x) + \iint\limits_{x > y > t} \Phi_2^2(t,y)dG(y)dF(x)$$
$$= \int_t^\infty [1-G(x)]\Phi_1^2(t,x)dF(x) + \int_t^\infty \Big(\int_t^x \Phi_2^2(t,y)dG(y)\Big)dF(x).$$

Combining this with equation of (3), we obtain

$$A = \int_t^\infty [1-G(x)]^{-1}\Big[1-\int_t^x \Phi_2(t,y)dG(y)\Big]^2 dF(x) + \int_t^\infty \Big(\int_t^x \Phi_2^2(t,y)dG(y)\Big)dF(x)$$
$$= \int_t^\infty \Big\{[1-G(x)]^{-1}\Big[1-\int_t^x \Phi_2(t,y)dG(y)\Big]^2 + \int_t^x \Phi_2^2(t,y)dG(y)\Big\}dF(x).$$

Since $F(x)$ is arbitrary, minimizing A is equivalent to minimizing

$$B \hat{=} B(\Phi_2) = \Big[1-\int_t^x \Phi_2(t,y)dG(y)\Big]^2 + [1-G(x)]\int_t^x \Phi_2^2(t,y)dG(y),$$

where G is a known distribution function.

By the variational principle, if $\Phi_2 = \xi = \xi(t,y)$ minimizes B, we consider $B(\xi+\rho\Delta)$, where $\Delta = \Delta(t,y)$ and ρ is a parameter, That is

$$B(x+\rho\Delta)$$
$$= 1-2\int_t^\infty (\xi+\rho\Delta)dG(y) + \Big[\int_t^x (\xi+\rho\Delta)dG(y)\Big]^2 + [1-G(x)]\int_t^x (\xi+\rho\Delta)^2 dG(y)$$
$$= B(x) + \rho^2\Big\{\Big[\int_t^x \Delta dG(y)\Big]^2 + [1-G(x)]\int_t^x \Delta^2 dG(y)\Big\}$$
$$+2\rho\Big\{-\int_t^x \Delta dG(y) + \int_t^x \xi dG(y)\int_t^x \xi dG(y) + [1-G(x)]\int_t^x \xi\Delta dG(y)\Big\}.$$

The last expression is a quadratic form of ρ. It leads to

$$\frac{\partial B(\xi+\rho\Delta)}{\partial\rho}\Big|_{\rho=0}$$
$$= -2\int_t^x \Delta dG(y) + 2\int_t^x \xi dG(y)\int_t^x \Delta dG(y) + 2(1-G(x))\int_t^x \xi\Delta dG(y) = 0$$

or $\int_t^x \Delta\{-1+\int_t^x \xi dG(y) + [1-G(x)]\xi\}dG(y) = 0$.

By the arbitrariness of Δ, we have

$$-1 + \int_t^x \xi dG(y) + [1-G(x)]\xi = 0.$$

Thus ξ can not be a functon of y , i.e. $\xi = \xi(t)$ only. So

$$-1 + \xi(t)\int_t^x dG(y) + [1-G(x)]\xi(t) = 0,$$
$$\xi(t) = \frac{1}{1-G(t)}.$$

Therefore we conclude.

Remark. Since we know G, $S_n^*(t)$ has some advantages compared with the Kaplan-Meier estimator.

(i) It is a sum type estimator and easy to be computed, while the Kaplan-Meier estimator, a product type estimator, is complicated in calculation.

(ii) Although we can prove that $\mathrm{Var}\, S_n^*(t) \geq \mathrm{Var}\, \hat{S}_n(t)$, the order of $\mathrm{Var}\, S_n^*(t)$ is still $O(n^{-1})$. This guarantees the consistency of $S_n^*(t)$. On the other hand, since $nS_n^*(t)$ is a sum of i.i.d. random variables, the strong law of large number, the law of iterated logarithm and also the Strassen type (functional) LIL can be applied to it to obtain strong consistency with the best convergence rate. Also the normal approximation formulas of $\sqrt{n}(S_n^*(t) - S(t))$, such as the Edgeworth expansion, are available.

(iii) After we construct $S_n^*(t)$ of sample size n, a new datum is observed. The only thing we need to do is to add one term to $S_n^*(t)$ to get $S_{n+1}^*(t)$. But for $\hat{S}_n(t)$, we need to rearrange all the $n+1$ of samples, such as ordering Z_i, and readjust the factors of the product etc.

3. Discussion of Choosing G

How to choose the censoring distribution G is a difficult problem. Heavy censoring reduces the observing times but leads to the large variance of $S_n^*(t)$. Since $\mathrm{Var}\, S_n^*(t)$ depends on t and (7) shows that $\mathrm{Var}\, S_n^*(0) = 0$ and $\mathrm{Var}\, S_n^*(\tau_F) = 0$ by the assumption of $\tau_F < \tau_G$, we want to measure the total lost caused by the variance of $S_n^*(t)$ during $[0, \tau_F]$.

Definition. We call

$$M = \int_0^{\tau_F} [\mathrm{Var}\, S_n(t)] dt \tag{9}$$

the cumulative variance of $S_n(t)$. That is for $S_n^*(t)$

$$M^* = \frac{1}{n} \int_0^{\tau_F} S(t) \Big[\frac{1}{1 - G(t)} - S(t) \Big] dt. \tag{10}$$

Now we consider the total cost of the experiment. It has the form

$$T = c \sum_{i=1}^{n} Z_i + aM^*, \tag{11}$$

where c, a are some positive constants. T is a random variable and we consider the expectation

$$\begin{aligned} E(T) &= c \sum_{i=1}^{n} EZ_i + aM^* \\ &= cn \int_0^{\tau_F} u dH(u) + \frac{a}{n} \int_0^{\tau_F} S(u) \Big[\frac{1}{1 - G(u)} - S(u) \Big] du \end{aligned}$$

$$= cn\int_0^{\tau_F}(1-F(u))(1-G(u))du + \frac{a}{n}\int_0^{\tau_F} S(u)\Big[\frac{1}{1-G(u)} - S(u)\Big]du$$
$$= \int_0^{\tau_F} S(u)\Big\{cn(1-G(u)) + \frac{a}{n}\frac{1}{1-G(u)}\Big\}du - \int_0^{\tau_F}\frac{aS^2(u)}{n}du. \tag{12}$$

Let $a = n^2c\alpha$. Then the integrand in the first term of (12) becomes

$$cn\Big\{[1-G(u)] + \frac{a}{1-G(u)}\Big\} \tag{13}$$

(i) If $0 < \alpha \leq 1$,

$$cn\Big\{[1-G(u)] + \frac{a}{1-G(u)}\Big\} \geq 2cn\sqrt{n}.$$

The minimum is attained at $1-G(u) \equiv \sqrt{\alpha}$ for any $u \in [0,\tau_F]$ and

$$E(T) = cn\Big[2\sqrt{\alpha}\int_0^{\tau_F} S(u)du - \alpha\int_0^{\tau_F} S^2(u)du\Big]$$

(ii) If $\alpha > 1$, (13) is a decreasing functoin of $1-G(u)$ which attains the minimum at $G(u) \equiv 0$ for any $u \in [0,\tau_F]$ and

$$E(T) = cn(1+\alpha)\int_0^{\tau_F} S(u)du - cn\alpha\int_0^{\tau_F} S^2(u)du.$$

It seems that if a is very large ($a > cn^2$) we should let the experiment continue until all the lifetime data can be obtained, otherwise only use $[\sqrt{\alpha}n]$ units on the experiment and also wait for the last lifetime datum.

One may think that the cost should also depend on the length of the period which the experiment lasts. Hence the total cost is

$$T_1 = c\sum_{i=1}^n Z_i + b\max_{1\leq i\leq n} Z_i + aM^*$$

with

$$\begin{aligned} E(T_1) &= cn\int_0^{\tau_F}(1-F(u))(1-G(u))du + b\int_0^{\tau_F} u\,dH^n(u) + aM^* \\ &= cn\int_0^{\tau_F}(1-F(u))(1-G(u))du + b\Big(\tau_F - \int_0^{\tau_F} H^n(u)du\Big) + aM^* \\ &= b\tau_F - \frac{a}{n}\int_0^{\tau_F} S^2(u)du + \int_0^{\tau_F}\Big\{cn(1-F(u))(1-G(u)) \\ &\qquad -b[1-(1-F(u))(1-G(u))]^2 + \frac{a}{n}\frac{S(u)}{1-G(u)}\Big\}du. \end{aligned} \tag{14}$$

Obviously the solution of minimizing ET_1 is complicated and the unknown $F(u)$ is involved.

However, the simplest idea is to choose G such that $EY_i = C$ of Scheme 1 guarantees the total observing time not to exceed that of Scheme 1 in average. The advantage here is to get information of the distribution after time C. The exponential distribution $G(y) = 1 - e^{-y/c}$ with $\tau_G = \infty$ is always suitable.

Table 1. Comparisons between $\hat{S}_{n,E}(u)$ and $S_n^*(u)$

u	$\hat{S}_{n,E}(u)$ (m.s.e.)	$S_n^*(u)$ (m.s.e.)
0.00	1.0000	1.0000
0.05	0.9482 (4.9235×10^{-5})	0.9480 (1.0010×10^{-4})
0.10	0.8985 (9.3178×10^{-5})	0.8986 (1.7998×10^{-4})
0.15	0.8507 (1.1566×10^{-4})	0.8524 (2.1514×10^{-4})
0.20	0.8011 (1.4205×10^{-4})	0.8011 (3.1780×10^{-4})
0.25	0.7501 (1.8421×10^{-4})	0.7511 (5.4101×10^{-4})
0.30	0.7011 (1.7249×10^{-4})	0.7013 (6.4012×10^{-4})
0.35	0.6515 (2.0623×10^{-4})	0.6536 (4.8947×10^{-4})
0.40	0.6008 (2.4100×10^{-4})	0.6024 (6.4763×10^{-4})
0.45	0.5518 (2.4442×10^{-4})	0.5524 (5.3862×10^{-4})
0.50	0.5034 (2.1210×10^{-4})	0.5034 (5.5453×10^{-4})
0.55	0.4525 (1.6319×10^{-4})	0.4520 (5.9636×10^{-4})
0.60	0.4031 (1.8335×10^{-4})	0.4002 (7.6347×10^{-4})
0.65	0.3531 (2.1739×10^{-4})	0.3496 (7.4137×10^{-4})
0.70		0.2969 (6.9940×10^{-4})
0.75		0.2469 (5.8715×10^{-4})
0.80		0.1969 (5.1338×10^{-4})
0.85		0.1487 (4.4569×10^{-4})
0.90		0.0980 (3.1918×10^{-4})
0.95		0.0472 (1.6906×10^{-4})

Table 1 shows the results of simulations (50 runs) where X_i are i.i.d. uniformly distributed on $[0, 1]$ and $n = 1000$, $C = 0.7$. We also let Y_i be i.i.d. with $G(y) = 1 - e^{-y/c}$. The second column lists the mean values of the empirical survival function in Scheme 1 with mean square errors while the third column lists the values of $S_n^*(t)$. The average experimental time of each sample is 0.4558 unit in Scheme 1 and 0.3276 unit in our procedure. We save 0.1282 unit time for each sample in average. The important thing is that we obtain the estimations of $S(u)$ for $u > C$.

It is conservative to choose $G(y)$ with $E(Y) \le C$. In fact

$$\begin{aligned} E(X \Lambda C) &= \int_0^C x dF(x) + \int_C^\infty C dF(x) \\ &= -x(1 - F(x))\Big|_0^C + \int_0^C [1 - F(x)]dx + C[1 - F(C)] \end{aligned}$$

$$= \int_0^C [1 - F(x)]dx$$

and

$$E(X \Lambda Y) = \int_0^\infty [1 - F(x)][1 - G(x)]dx$$
$$= \int_0^\infty [1 - F(x)]dx - \int_0^\infty [1 - F(x)]G(x)dx.$$

We only need to have

$$\int_C^\infty [1 - F(x)]dx < \int_0^\infty [1 - F(x)]G(x)dx. \tag{15}$$

Since the unknown distribution function F is involved, (15) may not reveal very much in practice.

4. Further Studies with Extra Information

In this section we study the function

$$S_A(t) = \frac{1}{n} \sum_{i=1}^n [\delta_i I(X_i > t) + (1 - \delta_i) P(X_i > t | X_i > Y_i)]. \tag{16}$$

It looks like an estimator of $S(t)$. Notice that Y_i are the censoring variables. When $\delta_i = 0$, the conditional probability $P(X_i > t | X_i > Y_i)$ is introduced. In fact, we have

Theorem 2.

(i) $ES_A(t) = S(t)$ (17)

(ii) $$\operatorname{Var} S_A(t) = \frac{1}{n} \Big\{ S(t) - S^2(t) - \Big[\int_t^\infty G(x)dF(x)\Big]\Big[\int_0^t G(x)dF(x)\Big] \Big[\int_0^\infty G(x)dF(x)\Big]^{-1} \Big\} \tag{18}$$

Remark. The variance of $S_A(t)$ is less than that of the empirical survival function. The reason is that the extra information $P(X_i > t | X_i > Y_i)$ is given.

Proof. Since

$$E[\delta_i I(X_i > t) + (1 - \delta_i) P(X_i > t | X_i > Y_i)]$$
$$= \iint\limits_{x \le y} I(x > t) dF(x) dG(x) + \iint\limits_{x > y} P(X_i > t | X_i > Y_i) dF(x) dG(x)$$
$$= \int_t^\infty [1 - G(x)]dF(x) + \int_t^\infty G(x)dF(x) = S(t),$$

the proof of (i) is complete. For the second assertion,

$$\begin{aligned}
&E[S_A(t)-S(t)]^2\\
&=\frac{1}{n^2}E\Big\{\sum_{i=1}^{n}[\delta_i I(X_i>t)+(1-\delta_i)P(X_i>t|X_i>Y_i)-S(t)]\Big\}^2\\
&=\frac{1}{n}\Big\{E[\delta_i I(X_i>t)+(1-\delta_i)P^2(X_i>t|X_i>Y_i)]-S^2(t)\Big\}\\
&=\frac{1}{n}\Big\{\int_t^\infty[1-G(x)]dF(x)+\Big[\int_t^\infty G(x)dF(x)\Big]^2\Big[\int_0^\infty G(x)dF(x)\Big]^{-1}-S^2(t)\Big\}\\
&=\frac{1}{n}\Big\{\int_t^\infty[1-G(x)]dF(x)+\int_t^\infty G(x)dF(x)\\
&\qquad-\Big[\int_t^\infty G(x)dF(x)\Big]\Big[\int_0^t G(x)dF(x)\Big]\Big[\int_0^\infty G(x)dF(x)\Big]^{-1}-S^2(t)\Big\}\\
&=\frac{1}{n}\Big\{S(t)-S^2(t)-\Big[\int_t^\infty G(x)dF(x)\Big]\Big[\int_0^t G(x)dF(x)\Big]\Big[\int_0^\infty G(x)dF(x)\Big]^{-1}\Big\}.
\end{aligned}$$

This completes the proof.

Unfortunately, $S_A(t)$ is not a statistics and we need extra information $P(X_i > t|X_i > Y_i)$. If we continue the testing after censoring and pursue the true lifetime data, the estimation of $P(X > t|X > Y)$ could be obtained as follows.

Definition. For any $t \geq 0$

$$\hat{P}(X>t|X>Y)=\begin{cases}\dfrac{\sum_{j=1}^{n} I(X_j>Y_j,X_j>t)}{\sum_{j=1}^{n} I(X_j>Y_j)}, & \text{if } \sum_{j=1}^{n} I(X_j>Y_j)>0,\\ 0, & \text{if } \sum_{j=1}^{n} I(X_j>Y_j)=0,\end{cases}$$

and

$$\begin{aligned}
\hat{S}_A(t)&=\frac{1}{n}\sum_{i=1}^{n}[\delta_i I(Z_i>t)+(1-\delta_i)\hat{P}(X>t|X>Y)]\\
&=\frac{1}{n}\sum_{i=1}^{n}\delta_i I(Z_i>t)+\Big[\frac{1}{n}\sum_{i=1}^{n}(1-\delta_i)\Big]\hat{P}(X>t|X>Y)
\end{aligned}$$

Theorem 3.

$$E\hat{S}_A(t)=S(t)\qquad \operatorname{Var}\hat{S}_A(t)=\frac{1}{n}[S(t)-S^2(t)]$$

Proof. For any $t \geq 0$

$$E\hat{S}_A(t)=E\Big[\frac{1}{n}\sum_{i=1}^{n}\delta_i I(Z_i>t)\Big]+E\Big\{\Big[\frac{1}{n}\sum_{i=1}^{n}(1-\delta_i)\Big]\hat{P}(X>t|X>Y)\Big\}$$

$$= \int_t^\infty [1 - G(x)]dF(x) + E\Big[\frac{1}{n}\sum_{j=1}^n I(X_j > Y_j, X_j > t)\Big]$$
$$= \int_t^\infty [1 - G(x)]dF(x) + \iint_{x>y,x>t} dF(x)dG(y) = S(t)$$

and

$$\begin{aligned}
&E[\hat{S}_A(t) - S(t)]^2 \\
&= E\Big\{\frac{1}{n}\sum_{i=1}^n \delta_i I(Z_i > t) + \Big[\frac{1}{n}\sum_{i=1}^n (1-\delta_i)\Big]\hat{P}(X > t | X > Y) - S(t)\Big\}^2 \\
&= E\Big\{\frac{1}{n}\sum_{i=1}^n \delta_i I(Z_i > t) + \frac{1}{n}\sum_{i=1}^n I(X_i > Y_i, X_i > t) - S(t)\Big\}^2 \\
&= E\Big\{\frac{1}{n}\sum_{i=1}^n [\delta_i I(Z_i > t) + I(X_i > Y_i, X_i > t) - S(t)]\Big\}^2 \\
&= \frac{1}{n} E[\delta_i I(Z_i > t) + I(X_i > Y_i, X_i > t) - S(t)]^2 \\
&= \frac{1}{n}\{E[\delta_i I(Z_i > t) + I(X_i > Y_i, X_i > t)] - S^2(t)\} \\
&= \frac{1}{n}[S(t) - S^2(t)].
\end{aligned}$$

Acknowledgements

The author is very grateful to the referees for their many helpful comments and suggestions.

References

Kalbfleish, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data.* John Wiley, New York.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53**, 457-481.

Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data.* John Wiley, New York.

Zheng Zukang (1987). A class of estimators for the parameters in linear regression with censored data. *Acta Math. Appl. Sinica* **3**, 131-141.

Department of Statistics and O.R., School of Management, Fudan University, Shanghai, 200433.

(Received July 1995; accepted January 1996)