Ann. Inst. Statist. Math.
Vol. 55, No. 2, 359-367 (2003)

BEZIER CURVE SMOOTHING OF THE KAPLAN-MEIER ESTIMATOR*

CHOONGRAK KIM[1], BYEONG U. PARK[2], WOOCHUL KIM[2] AND CHIYON LIM[1]

[1]*Department of Statistics, Pusan National University, Pusan 609-735, Korea*
[2]*Department of Statistics, Seoul National University, Seoul 151-747, Korea*

(Received April 15, 2002; revised June 3, 2002)

Abstract. Estimation of a survival function from randomly censored data is very important in survival analysis. The Kaplan-Meier estimator is a very popular choice, and kernel smoothing is a simple way of obtaining a smooth estimator. In this paper, we propose a new smooth version of the Kaplan-Meier estimator using a Bezier curve. We show that the proposed estimator is strongly consistent. Numerical results reveal that the proposed estimator outperforms the Kaplan-Meier estimator and its kernel weighted smooth version in the sense of mean integrated square error.

Key words and phrases: Bandwidth, censored data, kernel smoothing, strong consistency.

1. Introduction

Randomly right-censored data arise quite often in survival analysis for medical research, and estimating a distribution function is very important in this area. As an estimator of a survival function S, the Kaplan-Meier estimator (Kaplan and Meier (1958)) is most popular and has many desirable properties. First of all, the Kaplan-Meier estimator $\hat{S}(x)$ is self-consistent (Efron (1967), Miller (1981), Fleming and Harrington (1991)). Also, it is a generalized maximum likelihood estimator of $S(x)$ in the sense of Kiefer and Wolfowitz (1956). Peterson (1977) showed that $\hat{S}(x)$ is strongly consistent, and Breslow and Crowley (1974) established that it is asymptotically normal. However, it is a step-function which is undesirable as an estimator of a smooth survival function. This prompted many statisticians to find smooth versions of the Kaplan-Meier estimator. Among these, Blum and Susarla (1980) and Földes *et al.* (1981) suggested kernel methods based on the Kaplan-Meier estimator. A review of kernel density estimation from censored data is given by Padgett and McNichols (1984). Also, Padgett (1986) suggested a kernel-type estimator of a quantile function from right-censored data. Marron and Padgett (1987) discussed the bandwidth selection problem for the kernel density estimator with censored data. Note that all these results are based on kernel smoothing of the Kaplan-Meier estimator.

In this paper we suggest a new smooth version of the Kaplan-Meier estimator using a Bezier curve. Bezier smoothing is a very popular technique in computational graphics, especially for computer-aided-geometric-design. See Farin (2001) for a detailed discussion on Bezier curves. However, it appears to be virtually unknown to statisticians, and we

*This research is supported by the Korea Research Foundation (1998-015-d00047) made in the program year of 1998.

have found no other references to its use in statistics except Kim (1996) and Kim *et al.* (1999). In the latter two papers, it was shown that a density estimator using a Bezier curve has the same rate of convergence as the kernel estimator, but is superior to the latter for small sample sizes.

In Section 2, a kernel-type smooth version of the Kaplan-Meier estimator is introduced, which is a simple modification of the kernel-type distribution function estimator with complete data. In Section 3, the Bezier curve smoothing of the Kaplan-Meier estimator is suggested, and its strong consistency is established. Section 4 contains some simulation studies comparing the Bezier curve smoother with the kernel-type estimator by means of mean integrated square error.

2. The Kaplan-Meier estimator and kernel smoothing

Let $X_1^0, \ldots, X_n^0$ denote the true survival times from the unknown distribution function F, and let $C_1, \ldots, C_n$ denote the censoring random variables from the unknown distribution function G. Also, let S be the survival function, i.e., $S(x) = 1 - F(x)$. It is assumed that X^0 and C are independent. The randomly right-censored data are the pairs (X_i, δ_i), $i = 1, \ldots, n$ where $X_i = \min\{X_i^0, C_i\}$ and

$$\delta_i = \begin{cases} 1 & \text{if} \quad X_i^0 \le C_i \\ 0 & \text{if} \quad X_i^0 > C_i. \end{cases}$$

Here, δ_i is usually called a censoring indicator. For notational convenience, let $X_1 < X_2 < \cdots < X_n$ be the ordered survival times and δ_i be the censoring indicator corresponding to X_i. Also, it is assumed that there are no ties in survival times. Let $I(1) < I(2) < \cdots < I(N)$ be indices of the uncensored survival times, where $N = \sum_{i=1}^n \delta_i$ is the number of the uncensored survival times.

The most famous and widely used estimator of the survival function S is the Kaplan-Meier estimator (Kaplan and Meier (1958)), which is defined by

$$\hat{S}(x) = \prod_{i:X_i \le x} \left(\frac{n-i}{n-i+1} \right)^{\delta_i}. \tag{2.1}$$

Note that if the last observation is censored, i.e., $\delta_n = 0$, then $S(x) \not\to 0$ as $x \to \infty$. To avoid technical difficulties arising from this, it is usually assumed that the last observation is uncensored, i.e., $\delta_n = 1$.

First, we introduce the kernel weighted smooth version of the Kaplan-Meier estimator $\hat{S}(x)$. Note that the kernel estimator of the distribution function F with the complete data $X_1^0, \ldots, X_n^0$ is given by

$$\hat{F}_K(x) = \int \frac{1}{h} K\left(\frac{x-y}{h}\right) F_n(y) dy = \frac{1}{n} \sum_{i=1}^n W\left(\frac{x - X_i^0}{h}\right) \tag{2.2}$$

where F_n is the empirical distribution function, $W(x) = \int_{-\infty}^x K(t)dt$, K is the kernel function, and h is the bandwidth to be chosen. The kernel weigted version of the Kaplan-Meier estimator $\hat{S}(x)$ is then obtained by replacing the empirical distribution function by $\hat{S}$, i.e.

$$\hat{S}_K(x) = \int \frac{1}{h} K\left(\frac{x-y}{h}\right) \hat{S}(y) dy = 1 - \sum_{i=1}^n s_i W\left(\frac{x - X_i}{h}\right) \tag{2.3}$$

where s_i is the jump size at X_i of the Kaplan-Meier estimator $\hat{S}$. By differentiating $\hat{S}_K(x)$ with respect to x, we get the kernel density estimator

$$\hat{f}_K(x) = \frac{1}{h}\sum_{i=1}^{n} s_i K\left(\frac{x - X_i}{h}\right)$$

which was studied by Földes *et al.* (1981) and McNichols and Padgett (1986).

3. Bezier curve smoothing

3.1 *Bezier curve*

Consider $k+1$ points in R^2, denoted by

$$\boldsymbol{b}_0 = (z_0, w_0)', \boldsymbol{b}_1 = (z_1, w_1)', \dots, \boldsymbol{b}_k = (z_k, w_k)'$$

where $z_0 \le z_1 \le \cdots \le z_k$. The Bezier curve based on the $k+1$ Bezier points (control points) $\boldsymbol{b}_0, \boldsymbol{b}_1, \dots, \boldsymbol{b}_k$ is defined by

$$\boldsymbol{b}(t) \equiv \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \sum_{i=0}^{k} \boldsymbol{b}_i B_{k,i}(t), \qquad t \in [0,1] \tag{3.1}$$

where $B_{k,i}(t)$ is the binomial density given by

$$B_{k,i}(t) = \binom{k}{i} t^i (1-t)^{k-i},$$

which is also called a Bernstein polynomial or a blending function.

There are several properties to note on Bezier curves. First, Bezier curves have endpoint interpolation property, i.e., $\boldsymbol{b}_0$ and $\boldsymbol{b}_k$ are always on the curve $\boldsymbol{b}(\cdot)$. In fact, $\boldsymbol{b}(0) = \boldsymbol{b}_0$ and $\boldsymbol{b}(1) = \boldsymbol{b}_k$. Next, $\boldsymbol{b}(\cdot)$ is invariant under reversal of numbering for the Bezier points, i.e., $\{\sum_{i=0}^{k} \boldsymbol{b}_i B_{k,i}(t) : 0 \le t \le 1\} = \{\sum_{i=0}^{k} \boldsymbol{b}_{k-i} B_{k,i}(t) : 0 \le t \le 1\}$. It is clear that it does not change when one reverses the indices of $\boldsymbol{b}_0, \boldsymbol{b}_1, \dots, \boldsymbol{b}_k$. Thirdly, it preserves linearity. Note that $\sum_{i=0}^{k}(i/k)B_{k,i}(t) = t$ for all $t \in [0,1]$ so that a straight line is reproduced. Finally, the first derivative of $\boldsymbol{b}(t)$ with respect to t is given by

$$\frac{d}{dt}\boldsymbol{b}(t) = k\sum_{i=0}^{k-1}(\boldsymbol{b}_{i+1} - \boldsymbol{b}_i)B_{k-1,i}(t). \tag{3.2}$$

3.2 *The Besier smoother*

We note that the Bezier curve depends heavily on the choice of Bezier points. Figure 1 shows three types of Bezier points based on the Kaplan-Meier estimator with the resulting Bezier curve. The Bezier points in Fig. 1(a) and (b) are located at the left-most and the right-most, respectively, on a Kaplan-Meier estimator $\hat{S}$, while those in Fig. 1(c) are located at both. The artificial data used in Fig. 1 are 9, 13^+, 18, 23, 28^+, 31, 34, 45^+, 48, 80, where the numbers with + denote the censored observation.

We suggest to choose the Bezier points as described in Fig. 1(a) since we found that they give the best numerical performance in mean integrated square error among the three types (see Section 4). Now, we need to select a value, denoted by A_n, on

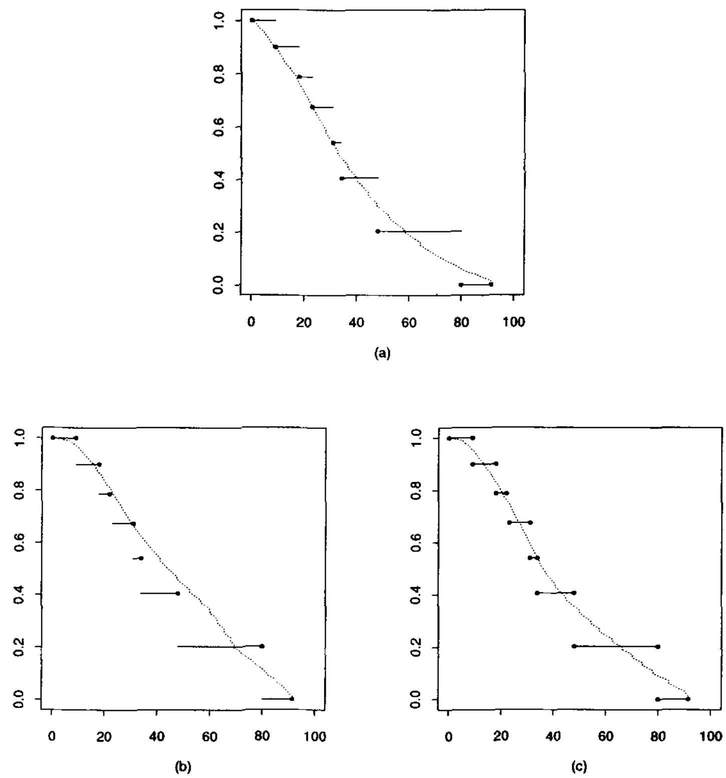

Fig. 1. Bezier curves (---) and Bezier points marked by • with $\hat{S}$ (—) where $\hat{S}$ is the Kaplan-Meier estimator.

the x-axis for the last Bezier point. First of all, A_n should be larger than the last uncensored observation $X_{I(N)}$. To see the influence of the last Bezier point, we performed a simulation study. Generating 100 pseudo values of $X \sim \text{Exp}(1)$ and $C \sim \text{Exp}(1)$, we computed mean integrated squared errors (MISE) for $A = 5(1)10$, and it turned out that MISE does not depend seriously on the choice of A_n. Based on our limited experience, we propose to choose $A_n = (1 + 1/N)X_{I(N)}$. For strong consistency of the Bezier estimator, it is sufficient to choose $A_n = (1 + r_n)X_{I(N)}$ for some random sequence r_n which converges to zero with probability one. See Theorem 1 below.

Thus, we consider $N + 2$ Bezier points which are given by

$$\boldsymbol{b}_0 = (0, 1)'; \quad \boldsymbol{b}_i = (X_{I(i)}, \hat{S}(X_{I(i)}))', \quad i = 1, \ldots, N; \quad \boldsymbol{b}_{N+1} = (A_n, 0)'.$$

The resulting Bezier curve is defined by

$$\boldsymbol{b}(t) \equiv \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \sum_{i=0}^{N+1} \boldsymbol{b}_i B_{N+1,t}(i),$$

where $B_{N+1,t}(i)$ is the binomial probability as defined in Subsection 3.1. Here for notational convenience, we used $B_{N,t}(i)$ instead of $B_{N,i}(t)$. Note that

$$x(t) \equiv x_n(t) = \sum_{i=0}^{N+1} X_{I(i)} B_{N+1,t}(i)$$

$$y(t) \equiv y_n(t) = \sum_{i=0}^{N+1} \hat{S}(X_{I(i)}) B_{N+1,t}(i)$$

with $X_{I(0)} = 0$, $X_{I(N+1)} = A_n$. Now, the Bezier estimator is defined by

$$\hat{S}_B(x) = y(t_n)$$

where t_n is the point such that $x(t_n) = x$.

There are several desirable properties of $\hat{S}_B$. First, $\hat{S}_B(0) = 1$ is guaranteed by the end-point interpolation property of the Bezier curve. Secondly, $\hat{S}_B$ is monotone which can be easily verified by using the first derivative of the Bezier curve given in (3.2). Thirdly, one does not need to choose a smoothing parameter which usually arises in other smoothing techniques. Fourthly, while the Bezier curve smoother is known to have some difficulties at boundaries in regression and density estimation settings, it does not have such problems in the current context. Finally, a Bezier curve estimator of the density with censored data may be obtained by differentiating $\hat{S}_B(x)$ with respect to t using (3.2).

To see consistency of $\hat{S}_B$, we define the sub-distribution function of the uncensored survival time by

$$F_u(x) = P(X_1 \le x, \delta_1 = 1).$$

A natural estimator of $F_u(x)$ is given by

$$\hat{F}_u(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \le x, \delta_i = 1).$$

Also, we define

$$F^{-1}(p) = \sup\{x : F(x) < p\}, \qquad 0 < p \le 1$$

where $F^{-1}(1)$ is the right end-point of the distribution F which may be $+\infty$. Let $f = F'$ and $g = G'$. We need the following assumptions:

(i) $F^{-1}(1) < G^{-1}(1) \le \infty$;
(ii) F is twice differentiable and G is differentiable on $(0, F^{-1}(1))$;
(iii) $\inf_{0<x<F^{-1}(1)} f(x) > 0$ and $\sup_{0<x<F^{-1}(1)} |f'(x)| < \infty$;
(iv) $\sup_{0<x<F^{-1}(1)} g(x) < \infty$.

THEOREM 1. *Let x be a fixed point in $(0, F^{-1}(1))$. Under the assumptions* (i) $\sim$ (iv), *$\hat{S}_B(x)$ converges to $S(x)$ as $n \to \infty$ with probability one if one takes $A_n = (1 + r_n)X_{I(N)}$ for some random sequence r_n which converges to zero with probability one.*

See the Appendix for a proof.

4. Numerical study

To evaluate the numerical performance of $\hat{S}_B$, we conducted a simulation study comparing the mean integrated squared errors (MISE) of several estimators of S. These estimators include the Kaplan-Meier estimator ($\hat{S}$), its kernel smoother ($\hat{S}_K$), and the three types of the Bezier curve smoothers ($\hat{S}_{B_1}, \hat{S}_{B_2}, \hat{S}_{B_3}$). Here, B_1, B_2 and B_3 indicate the types of the Bezier points used, corresponding to the left-most, the right-most and the case where both are taken, respectively. We generated survival times from Exp(1), i.e., $S(t) = e^{-t}$, and censoring times from Exp(λ), i.e., $1 - G(t) = e^{-\lambda t}$ with $\lambda = 1$ (50% censoring) and $\lambda = 3/7$ (30% censoring). Sample sizes considered are n = 30, 50 and

Table 1. Monte Carlo estimates of the mean integrated squared error for $\hat{S}$, $\hat{S}_K$, $\hat{S}_{B_1}$, $\hat{S}_{B_2}$ and $\hat{S}_{B_3}$.

n	λ	$\hat{S}$	$\hat{S}_K$	$\hat{S}_{B_1}$	$\hat{S}_{B_2}$	$\hat{S}_{B_3}$
30	1	.0414731	.0241248	.0120058	.0655141	.0192931
		(.0034031)	(.0025061)	(.0012955)	(.0066889)	(.0023042)
	3/7	.0233780	.0178458	.0114956	.0308447	.0159091
		(.0018285)	(.0018032)	(.0013272)	(.0031426)	(.0018397)
50	1	.0279322	.0177510	.0096526	.0311106	.0130147
		(.0024091)	(.0018981)	(.0010202)	(.0036589)	(.0015957)
	3/7	.0151211	.0117405	.0086413	.0151438	.0105112
		(.0011062)	(.0010559)	(.0009123)	(.0016240)	(.0010962)
100	1	.0147137	.0109856	.0055154	.0156707	.0073325
		(.0011710)	(.0010692)	(.0005718)	(.0016815)	(.0008267)
	3/7	.0086773	.0075782	.0053548	.0087745	.0064684
		(.0006986)	(.0007075)	(.0004768)	(.0008417)	(.0006067)

Note: Results are based on 100 pseudo samples. Standard errors of the Monte Carlo estimates are given in the parentheses.

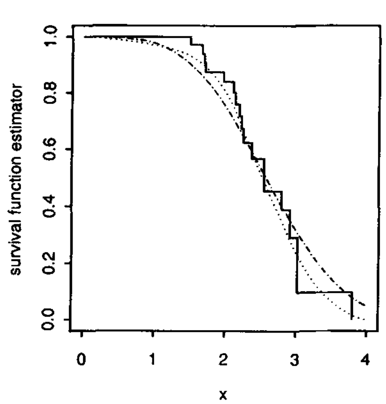

Fig. 2. Three types of survival estimates, Kaplan-meier estimate $\hat{S}$ (—), kernel estimate $\hat{S}_K$ (-·-), and the Bezier estimator $\hat{S}_{B_1}$ (···) with the switch life data.

100. For each sample size, 100 pseudo samples are generated. For the kernel estimator $\hat{S}_K$, we used the Epanechnikov kernel and the global bandwidth which minimizes the MISE.

The simulation results are summarized in Table 1. First, the MISE decreases as the sample size increases. Second, the first type of the Bezier curve smoother $\hat{S}_{B_1}$ outperforms the other two. Also, the kernel estimator $\hat{S}_K$ is inferior to the Bezier estimator $\hat{S}_{B_1}$ even though we used the optimal bandwidth for $\hat{S}_K$. We see a slight improvement of $\hat{S}$ with $\hat{S}_K$. Finally, the MISE decreases as the censoring proportion decreases.

As an illustrative example, we considered the switch life data with $n = 40$, which was also used in quantile estimation by Padgett (1986). The data set consists of 17 uncensored observations and 23 censored ones, i.e. there is 57.5% censoring. The three estimators $\hat{S}$, $\hat{S}_K$ and $\hat{S}_{B_1}$ are depicted in Fig. 2. The bandwidth for $\hat{S}_K(x)$ we used is 1.3. We see that the Bezier smoother is much closer to the Kaplan-Meier estimator than the kernel estimator.

Appendix: Proof of Theorem 1

We prove $\hat{F}_B(x) \to F(x)$ with probability 1 where $\hat{F}_B(x) = 1 - \hat{S}_B(x)$. Write $\hat{F}(x) = 1 - \hat{S}(x)$. For the sub-distribution function F_u, define

$$F_u^{-1}(p) = \sup\{x : F_u(x) < p\}, \qquad 0 < p \leq 1.$$

Likewise, define $\hat{F}_u^{-1}(p)$ for $0 < p \leq 1$. Let $F_u^{-1}(0) = \hat{F}_u^{-1}(0) = 0$. Note that $F_u^{-1}(F_u(+\infty)) = F^{-1}(1)$. We may write

$$x(t) = \sum_{i=0}^{N} \hat{F_u}^{-1}\left(\frac{i}{n}\right) B_{N+1,t}(i) + A_n t^{N+1}$$

$$y(t) = \sum_{i=0}^{N} \hat{F}\left(\hat{F_u}^{-1}\left(\frac{i}{n}\right)\right) B_{N+1,t}(i) + \hat{F}(A_n) t^{N+1}.$$

We claim that

$$x(t) = F_u^{-1}(F_u(+\infty)t) + R_n(t) \tag{A.1}$$

where $\sup_{0 \leq t \leq 1} |R_n(t)| \to 0$ as $n \to \infty$ with probability 1. Suppose (A.1) holds. Let t_n be a sequence of random variables such that $x(t_n) = x$. By (A.1) we get $F_u{}^{-1}(F_u(+\infty)t_n) \to x$ so that

$$t_n \to \frac{F_u(x)}{F_u(+\infty)} \tag{A.2}$$

with probability 1.

Next, we consider $y(t)$. By Csörgo and Horváth (1983) or Földes and Rejtö (1981), we may write

$$y(t) = \sum_{i=0}^{N} F\left(\hat{F}_u^{-1}\left(\frac{i}{n}\right)\right) B_{N+1,t}(i) + F(A_n) t^{N+1} + \tilde{R}_n(t) \tag{A.3}$$

where $\sup_{0\le t\le 1}|\tilde{R}_n(t)| = O(n^{-1/2}(\log\log n)^{1/2})$ as $n\to\infty$ with probability 1. Using (A.3) and applying the same arguments leading to (A.1) along with a Taylor expansion we obtain

$$y(t) = F(F_u^{-1}(F_u(+\infty)t)) + R_n{}^*(t), \tag{A.4}$$

where $\sup_{0\le t\le 1}|R_n{}^*(t)| \to 0$ as $n\to\infty$ with probability 1. Putting (A.2) and (A.4) together and noting that F_u is strictly increasing on $(0, F^{-1}(1))$, we get $\hat{F}_B(x) = y(t_n) \to F(x)$ with probability 1.

It remains to prove (A.1). Let $R_n(t) = R_{n1}(t) + R_{n2}(t) + R_{n3}(t)$ where

$$\begin{aligned}
R_{n1}(t) &= \sum_{i=0}^{N}\left\{\hat{F}_u^{-1}\left(\frac{i}{n}\right) - F_u^{-1}\left(\frac{i}{n}\right)\right\}B_{N+1,t}(i)\\
&\quad + \left\{A_n - F_u^{-1}\left(\frac{N+1}{n}\right)\right\}t^{N+1};\\
R_{n2}(t) &= \sum_{i=0}^{N+1}\left\{F_u^{-1}\left(\frac{i}{n}\right) - F_u^{-1}\left(\frac{(N+1)t}{n}\right)\right\}B_{N+1,t}(i);\\
R_{n3}(t) &= F_u^{-1}\left(\frac{(N+1)t}{n}\right) - F_u^{-1}(F_u(+\infty)t).
\end{aligned}$$

We prove all the three terms tend to zero uniformly in $t (0\le t\le 1)$ with probability 1. Consider R_{n1} first. Divide the summation into two parts : the one for $0\le i\le nF_u(+\infty)$ and the other for $i > nF_u(+\infty)$. Now, we can get an analogue of Theorem 3 of Kiefer (1970) with F_u taking the role of F there:

$$\sup_{0\le p\le F_u(+\infty)} f_u(F_u^{-1}(p))|\hat{F}_u^{-1}(p) - F_u^{-1}(p)| = O(n^{-1/2}(\log\log n)^{1/2})$$

with probability 1 where $f_u = F_u'$. Thus, the first summation has an order of magnitude $n^{-1/2}(\log\log n)^{1/2}$ uniformly in t $(0\le t\le 1)$ with probability 1. Next, the second summation plus $\{A_n - F_u^{-1}(\frac{N+1}{n})\}t^{N+1}$ is bounded (uniformly in t) by $|\hat{F}_u^{-1}(F_u(+\infty)) - F^{-1}(1)| + r_n X_{I(N)}$ which converges to zero with probability 1. Consider R_{n2} now. We note that

$$\sum_{i=0}^{N+1}\{i-(N+1)t\}B_{N+1,t}(i) = 0$$

and $(F_u^{-1})''$ is bounded. Thus a Taylor expansion shows that

$$\sup_{0\le t\le 1}|R_{n2}(t)| = O(n^{-1})$$

with probability 1. Finally, it follows that $\sup_{0\le t\le 1}|R_{n3}(t)|\to 0$ with probability 1 since $\frac{N}{n}\to F_u(+\infty)$ with probability 1 and $(F_u^{-1})'$ is bounded.

REFERENCES

Blum, J. R. and Susarla, V. (1980). Maximum deviation theory of density and failure rate function estimates based on censored data, *J. Multivariate Anal.*, **5**, 213–222.

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.

Csörgo, S. and Horváth, L. (1983). The rate of strong uniform consistency for the product-limit estimator, *Z. Wahrsch. Verw. Gebiete*, **62**, 411–426.

Efron, B. (1967). The two sample problem with censored data, *Proc. Fifth Berkeley Symp. on Math. Statist. Probab.*, Vol. 4, 831–853.

Farin, G. (2001). *Curves and Surfaces for Computer Aided Geometric Design*, 5th ed., Morgan-Kaufmann, San Francisco.

Fleming, T. R. and Harrington, D. P. (1991). *Counting Processes and Survival Analysis*, Wiley, New York.

Földes, A. and Rejtö, L. (1981). Strong uniform cosistency for nonparametric survival curve estimators from randomly censored data, *Ann. Statist.*, **9**, 122–129.

Földes, A., Rejtö, L. and Winter, B. B. (1981). Strong consistency properties of nonparametric estimator for randomly censored data, II: Estimation of density and failure rate, *Periodica Mathematica Hungarica*, **12**, 15–19.

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

Kiefer, J. (1970). Deviations between the sample quantile process and the sample D.F., *Non-parametric Techniques in Statistical Inference* (ed. M. L. Puri), 299–319, Cambridge University Press, Cambridge.

Kiefer, J. and Wolfowitz, J. (1956). Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters, *Ann. Math. Statist.*, **27**, 887–906.

Kim, C. (1996). Nonparametric density estimation via the Bezier curve, *Proceeding of the Section on Statistical Graphics*, 25–28, American Statistical Association, Alexandria, Virginia.

Kim, C., Kim, W., Hong, C., Park, B. U. and Jeong, M. (1999). Smoothing techniques via the Bezier curve, *Comm. Statist. Theory Methods*, **28**, 1577–1596.

Marron, J. S. and Padgett, W. J. (1987). Asymptotically optimal bandwidth selection for kernel density estimators from randomly right-censored samples, *Ann. Statist.*, **15**, 1520–1535.

McNichols, D. T. and Padgett, W. J. (1986). Mean and variance of a kernel density estimator under the Koziol-Green model of random censorship, *Sankhyā Ser. A*, **48**, 150–168.

Miller, R. G., Jr. (1981), *Survival Analysis*, Wiley, New York.

Padgett, W. J. (1986). A kernel-type estimation of a quantile function from right-censored data, *J. Amer. Statist. Assoc.*, **81**, 215–222.

Padgett, W. J. and McNichols, D. T. (1984). Nonparametric density estimation from censored data, *Comm. Statist. Theory Methods*, **13**, 1581–1611.

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