Statistics and Its Interface Volume 10 (2017) 601–605

Empirical likelihood bivariate nonparametric maximum likelihood estimator with right censored data and continuous covariate

Jian-Jian Ren*

Recently, Ren and Riddlesworth (2014) derived the empirical likelihood-based *bivariate nonparametric maximum likelihood estimator* (BNPMLE) $\hat{F}_n(t,z)$ for the bivariate distribution function $F_0(t,z)$ of survival time T and covariate variable Z based on bivariate data where T is subject to right censoring. They showed that such BNPMLE $\hat{F}_n(t,z)$ is a consistent estimator of $F_0(t,z)$ when variable Z is discrete. Despite all nice properties of the BNPMLE $\hat{F}_n(t,z)$ shown in Ren and Riddlesworth (2014), in this article we show that surprisingly such $\hat{F}_n(t,z)$ is not a consistent estimator when the covariate variable Z is continuous. On the other hand, interestingly our simulation studies suggest that some remedy adjustments on $\hat{F}_n(t,z)$ based on the usual empirical likelihood treatments and the censoring mechanism may provide consistent estimators for $F_0(t,z)$ with continuous covariate Z.

AMS 2000 subject classifications: 62G05, 62N02.
Keywords and phrases: Bivariate data, Bivariate right censored data, Empirical likelihood, Maximum likelihood estimator, Right censored data.

1. INTRODUCTION

In the analysis of survival data, we often encounter situations where the response variable is the survival time T and is subject to right censoring, but the p-dimensional vector $\boldsymbol{Z}$ of covariates with components such as treatments, gender, age, etc., are completely observable. For simplicity of presentation, here we consider the case that covariate $\boldsymbol{Z}$ is a scalar rather than a vector, i.e., Z with dimension $p=1$, while it should be noted that with minor modifications, the generalization of the main results in this article to multivariate case with $p>1$ is straightforward. Specifically, suppose that

$$(1.1) \qquad (T_1,Z_1),(T_2,Z_2),\cdots,(T_n,Z_n)$$

is a random sample of (T,Z), but the actually observed survival data are the bivariate data with one coordinate subject to random right censoring as follows:

$$(1.2) \qquad (V_1,\delta_1,Z_1),(V_2,\delta_2,Z_2),\cdots,(V_n,\delta_n,Z_n),$$

where $V_i=\min\{T_i,C_i\}, \delta_i=I\{T_i\le C_i\}$, and C_i is the right censoring variable with distribution function (d.f.) F_C and is independent of (T_i,Z_i). In practice, if one wishes to use the nonparametric approach (i.e., without imposing any model assumptions) in the study of the relation between the right censored response variable T and the completely observable covariate variable Z, a natural thing to do is to estimate the bivariate distribution function $F_0(t,z)$ of (T,Z) based on observed survival data. Recently, Ren and Riddlesworth (2014) derived the *bivariate nonparametric maximum likelihood estimator* (BNPMLE) $\hat{F}_n(t,z)$ for $F_0(t,z)$ based on data (1.2) using the empirical likelihood method (Owen, 1988), and studied asymptotic properties of such $\hat{F}_n(t,z)$ under certain conditions.

As reviewed in Ren and Riddlesworth (2014), there have been some limited works on the bivariate d.f. estimation of $F_0(t,z)$ in the literature. For instance, Akritas (1994) constructed an estimator $\hat{F}_A$ for $F_0(t,z)$ with data (1.2) using the conditional survival distribution and kernel estimator approach, thus estimator $\hat{F}_A$ is kernel and bandwidth dependent, and is not a maximum likelihood estimator in any sense. Other related works include: Lin and Ying (1993) who considered the problem that both components T_i and Z_i in sample (1.1) are subject to the same univariate right censoring simultaneously; Dabrowska (1988, 1989) who considered the bivariate right censored data, i.e., each component of (T_i,Z_i) in sample (1.1) is subject to its own right censoring variable; van der Laan (1996) who considered a less related problem in which setting the bivariate right censoring vector is discrete and always observed. The work most closely related to the BNPMLE $\hat{F}_n(t,z)$ by Ren and Riddlesworth (2014) is that by Ren and Gu (1997) who constructed a bivariate distribution function estimator based on bivariate survival data which is subject to double censoring in one coordinate. Since right censoring is a special case of double censoring, our above data (1.2) is a special case of that considered in Ren and Gu (1997). The estimator by Ren and Gu

*The author's research was partially supported by NSF grant DMS-1407461; *Mailing Address*: Department of Mathematics, University of Maryland, College Park, MD 20742.

(1997) was given by the product of the empirical distribution estimator for $F_Z(z)$ and the conditional NPMLE for $P\{T \le t \,|\, Z \le z\}$, which is the univariate version of the NPMLE for doubly censored data given by Mykland and Ren (1996).

In terms of applications, the problem with data (1.2) is encountered far more frequently in practical situations than other types of bivariate survival data mentioned above. In terms of methodology, except the BNPMLE $\hat{F}_n(t,z)$ by Ren and Riddlesworth (2014), most estimators proposed for bivariate data subject to censoring are ad hoc, and either are kernel and bandwidth dependent (e.g., Akritas, 1994) or contain negative probability masses; see discussions in van der Laan (1996). For instance, both bivariate distribution estimators by Dabrowska (1988) and Ren and Gu (1997) contain negative probability masses, thus not monotone in bivariate sense.

The BNPMLE $\hat{F}_n(t,z)$ for $F_0(t,z)$ based on data (1.2) by Ren and Riddlesworth (2014) has many nice properties and reveals many interesting discoveries. Ren and Riddlesworth (2014) showed the following: (i) $\hat{F}_n(t,z)$ has an explicit expression and is unique in the sense of empirical likelihood; (ii) the conditional distribution function of T given Z under $\hat{F}_n(t,z)$ is of the same form as the Kaplan-Meier estimator for the univariate case; (iii) $\hat{F}_n(t,z)$ is the sum of the products of the marginal NPMLE for $F_Z(z)$ and the conditional NPMLE (i.e., the Kaplan-Meier estimator) for $F(t|Z=z)$; (iv) $\hat{F}_n(t,z)$ has only nonnegative probability masses, thus is monotone in bivariate sense; (v) the marginal d.f. $\hat{F}_n(\infty,z)$ coincides with the empirical d.f. of the covariate sample Z_i's; (vi) $\hat{F}_n(t,z)$ coincides with the bivariate empirical distribution function when there is no censoring; (vii) the strong consistency and weak convergence of $\hat{F}_n(t,z)$ hold for discrete covariate Z.

However, despite all above nice properties of the BNPMLE $\hat{F}_n(t,z)$ (Ren and Riddlesworth, 2014), in this article we show that surprisingly such $\hat{F}_n(t,z)$ is not a consistent estimator when the covariate variable Z is continuous. On the other hand, interestingly our simulation studies suggest that some remedy adjustments on $\hat{F}_n(t,z)$ based on the usual empirical likelihood treatment and the censoring mechanism may provide consistent estimators for continuous covariate Z.

The main results of this article are organized as follows. In Section 2, we present the BNPMLE $\hat{F}_n(t,z)$ for $F_0(t,z)$ with data (1.2) which is derived and studied by Ren and Riddlesworth (2014), and with proofs deferred to Section 4 we show that $\hat{F}_n(t,z)$ is an inconsistent estimator for $F_0(t,z)$ with continuous covariate Z. At the end of Section 2, we provide discussions on the extension of our main results to the case of p-variate covariate $\boldsymbol{Z}$ with $p>1$. Section 3 discusses intuitive interpretation of the inconsistency results and some remedies for such an issue, then presents some simulation results along with some comparison and concluding remarks.

2. NONPARAMETRIC MAXIMUM LIKELIHOOD ESTIMATOR

To state empirical likelihood-based bivariate maximum likelihood estimator (BNPMLE) $\hat{F}_n(t,z)$ for $F_0(t,z)$ of (T,Z) with data (1.2) given by Ren and Riddlesworth (2014), we let

$$\text{(2.1)}\quad \begin{array}{l} U_1<\cdots<U_m \text{ be all distinct values among } V_1,\cdots,V_n \\ Y_1<\cdots<Y_q \text{ be all distinct values among } Z_1,\cdots,Z_n \end{array}$$

and we denote for $1\le i\le m, 1\le j\le q$:

$$\text{(2.2)}\quad \begin{cases} n_{ij}=\sum_{k=1}^{n} I\{V_k=U_i, Z_k=Y_j\} \\ \delta_{ij}=\sum_{k=1}^{n} I\{V_k=U_i, \delta_k=1, Z_k=Y_j\} \\ N_{ij}=n_{ij}+\cdots+n_{mj}=\sum_{k=1}^{n} I\{V_k\ge U_i, Z_k=Y_j\} \\ m_j=\max\{k \,|\, n_{kj}>0\}. \end{cases}$$

Ren and Riddlesworth (2014) show that in the sense of the empirical likelihood method, the BNPMLE $\hat{F}_n(t,z)$ for $F_0(t,z)$ is *uniquely* given by the following:

$$\text{(2.3)}\quad \begin{cases} \hat{F}_n(t,z)=\sum_{i=1}^{m}\sum_{j=1}^{q}\hat{p}_{ij} I\{U_i\le t, Y_j\le z\}, \\ \qquad \text{for } t\le U_m,\ z\in\mathbb{R} \\ \hat{p}_{ij}=\Big(\frac{\delta_{ij}}{N_{ij}}\Big)\Big(\frac{N_{1j}}{n}\Big)\prod_{k=1}^{i-1}\Big(1-\frac{\delta_{kj}}{N_{kj}}\Big), \\ \qquad \text{for } 1\le i\le m,\ 1\le j\le q \\ \hat{p}_{m+1,j}=P_{\hat{F}_n}\{T>U_m, Z=Y_j\} \\ \qquad =\Big(\frac{N_{1j}}{n}\Big)-\sum_{i=1}^{m}\hat{p}_{ij}, \quad \text{for } 1\le j\le q \end{cases}$$

where notation $\prod_{k=1}^{0} c_k \equiv 1$ is used, and $0/0$ is set as 0 whenever it occurs.

Note that (2.2) implies that for any $1\le j\le q$,

$$\text{(2.4)}\quad \begin{cases} n_{m_j,j}>0 \ \Rightarrow\ N_{1j}\ge N_{2j}\ge\cdots\ge N_{m_j,j}>0 \\ n_{ij}=\delta_{ij}=N_{ij}=0, \\ \qquad \text{for } m_j<i\le m \text{ when } m_j<m. \end{cases}$$

Thus, in (2.3) we have that for any $1\le j\le q$,

$$\text{(2.5)}\quad \begin{cases} \hat{p}_{ij}=0, \quad \text{for } m_j<i\le m \\ \hat{p}_{m_j+1,j}=P_{\hat{F}_n}\{T>U_{m_j}, Z=Y_j\} \\ \qquad =\hat{p}_{m+1,j}=P_{\hat{F}_n}\{T>U_m, Z=Y_j\}. \end{cases}$$

For discussions and proofs on the structure and those properties (i)-(vii) of BNPMLE $\hat{F}_n(t,z)$ listed in Section

1 as well as its relation to the estimator by Ren and Gu (1997), see Ren and Riddlesworth (2014). With the proofs deferred to Section 4, the following theorem shows that the BNPMLE $\hat{F}_n(t,z)$ for $F_0(t,z)$ is an inconsistent estimator when the covariate variable Z is continuous.

Theorem 1. *For right censored bivariate survival data (1.2), assume that covariate variable Z and variable $V = \min\{T, C\}$ are continuous. Then, under the usual conditions on the empirical distribution function, we have*

$$\sup_{t,\, z\in\mathbb{R}} |\hat{F}_n(t,z) - F_1(t,z)| \overset{a.s.}{\to} 0, \qquad \text{as } n \to \infty \tag{2.6}$$

where $F_1(t,z) = P\{V \le t, \delta = 1, Z \le z\} \ne F_0(t,z)$.

Remark 1 (Meaning of Theorem 1)**.** It is interesting to note that Theorem 1 is saying that with censored bivariate data (1.2), the bivariate maximum likelihood distribution estimator $\hat{F}_n(t,z)$ constructed in the sense of ordinary likelihood is inconsistent for continuous variable Z, although $\hat{F}_n(t,z)$ coincides with the usual bivariate empirical d.f. when there is no censoring (see Corollary 3 of Ren and Riddlesworth (2014)). But from Ren and Riddlesworth (2014), this same distribution estimator $\hat{F}_n(t,z)$ is consistent when variable Z is discrete.

Remark 2 (Extension to p-Variate $\boldsymbol{Z}$ with $p > 1$)**.** If $\boldsymbol{Z}_i$'s in (1.2) is p-variate with $p > 1$ containing at least one component that is continuous, then we let $\boldsymbol{Y}_j \in \mathbb{R}^p$ in (2.1), now without ordering, represent all distinct vectors of $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$. Since p-variate $\boldsymbol{Z}_i$ contains at least one component that is continuous, we know that with probability 1, vectors $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$ are all distinct, thus $q = n$ in (2.1). Following the proofs of Theorem 1 given in Section 4, with some minor modifications we can show that above theorem holds for p-Variate $\boldsymbol{Z}$ with $p > 1$, where $Y_j \le z$ and $Z = Y_j$ are replaced by $\boldsymbol{Y}_j \le \boldsymbol{z}$ and $\boldsymbol{Z} = \boldsymbol{Y}_j$, respectively, with the use of the following notations: $(\boldsymbol{Y}_j \le \boldsymbol{z}) \equiv (Y_{1j} \le z_1, \ldots, Y_{pj} \le z_p)$.

3. REMEDIES AND SIMULATIONS

Theorem 1 shows that the BNPMLE $\hat{F}_n(t,z)$ for $F_0(t,z)$ given by Ren and Riddlesworth (2014) is inconsistent estimator when variable Z in right censored bivariate survival data (1.2) is continuous. This section considers the issues of remedies for such inconsistency and presents some simulation results.

Intuitive interpretation First, there is not an obviously intuitive interpretation for the inconsistency result we obtain in Theorem 1 of Section 2, which is why the discovery of such a fact is surprising. On the other hand, the following facts and observations may be helpful for us to see some light in this regard and find the remedies. It is well known that in the univariate data case, the Kaplan-Meier estimator $\hat{F}_{KM}(t)$ based on right censored data (V_i, δ_i)'s in (1.2) is an NPMLE for d.f. $F_T(t)$ of T, but it is not always a proper d.f. because $\hat{F}_{KM}(\infty) < 1$ if the largest value among observed V_i's is right censored (Kaplan and Meier, 1958; Shorack and Wellner, 1986). In practice, we usually compute the adjusted version of $\hat{F}_{KM}(t)$ by setting 1 as the value of the NPMLE at the largest observation of the V_i's; this kind of adjustment of the NPMLE is a generally adopted convention for right censored data (Efron, 1967; Miller, 1976). For bivariate survival data (1.2), BNPMLE $\hat{F}_n(t,z)$ given in (2.3) for $F_0(t,z)$ was obtained based on the empirical likelihood method which restricts all possible candidates to those bivariate d.f.'s that assign all their probability masses to points (U_i, Y_j) in (2.1) and line segments $L_j = \{(t, Y_j) \in \mathbb{R}^2 ; U_m < t < \tau_T\}$ for $1 \le i \le m, 1 \le j \le q$, where τ_T is from the support $(0, \tau_T)$ of survival time T and may be a finite number or ∞; see Ren and Riddlesworth (2014) for details. Such a method is similar to the derivation of the Kaplan-Meier estimator $\hat{F}_{KM}(t)$ for the univariate survival data case. In the case of univariate data, the Kaplan-Meier estimator $\hat{F}_{KM}(t)$ places $[1 - \hat{F}_{KM}(U_m)]$ as the estimated probability mass on line segment $L = (U_m, \tau_T)$, and $[1 - \hat{F}_{KM}(U_m)]$ converges to 0 as $n \to \infty$ because $U_m = V_{(n)}$ converges to τ_T in usual situation. Thus, $\hat{F}_{KM}(t)$ is a consistent estimator of F_T. In the case of discrete Z in (1.2) with a constant q in (2.1), we have $q < n$, thus as $n \to \infty$, for any $1 \le j \le q$ the number of V_i's observed on each line segment $\{(t, Y_j) \in \mathbb{R}^2 ; 0 < t < U_m\}$ goes to ∞. Hence, as shown in Ren and Riddlesworth (2014), the BNPMLE $\hat{F}_n(t,z)$ is a consistent estimator of $F_0(t,z)$ for discrete variable Z. In the case of continuous Z in (1.2), with probability 1 we know that $q = n$ in (2.1), and that if for certain (U_i, Y_j) there exists k such that $(V_k, Z_k) = (U_i, Y_j)$ with $\delta_k = 1$, it means $U_i = V_k = T_k$ is not a censored observation and the BNPMLE $\hat{F}_n(t,z)$ gives $1/n$ as the probability mass at point $(V_k, Z_k) = (U_i, Y_j)$; however, if $\delta_k = 0$, it means $U_i = V_k = C_k$ is a right censored observation and the BNPMLE $\hat{F}_n(t,z)$ assigns $1/n$ as the probability on line segment $\{(t, Y_j) \in \mathbb{R}^2 ; U_{m_j} < t < \tau_T\}$ due to (2.2)-(2.5), because there is no other $l \ne k$ such that $V_l \ne U_i$ and $Z_l = Y_j$, noting that just like the Kaplan-Meier estimator, the BNPMLE $\hat{F}_n(t,z)$ does not actually specify how the probability $1/n$ is assigned on line segment $\{(t, Y_j) \in \mathbb{R}^2 ; U_{m_j} < t < \tau_T\}$. In other words, with continuous variable Z in (1.2), we know that with probability 1, for any $1 \le j \le q = n$ there is exactly one of V_i's observed on line segment $\{(t, Y_j) \in \mathbb{R}^2 ; 0 < t < U_m\}$, and if such an observation is censored, the BNPMLE $\hat{F}_n(t,z)$ assigns $1/n$ as the probability on line segment $\{(t, Y_j) \in \mathbb{R}^2 ; U_{m_j} < t < \tau_T\}$, which contains no observed data points (V_i, Z_i)'s. With the understanding on these points, in the rest of this section we consider the remedies for the inconsistency issue with continuous Z and some simulation results for the bivariate survival data (1.2), which serves as explorative studies

Table 1. Comparison of $\hat{F}_{0n}, \hat{F}_n, \hat{F}_{nn}$, F_0 with Right Censored Samples

Sample Size	Average $\|\hat{F}_{0n} - F_0\|$ (s.d.)	Average $\|\hat{F}_n - F_0\|$ (s.d.)	Average $\|\hat{F}_{nn} - F_0\|$ (s.d.)
$n = 50$	0.2149 (.0741)	0.1286 (.0386)	0.1339 (.0399)
$n = 100$	0.2320 (.0603)	0.0988 (.0284)	0.1032 (.0293)
$n = 200$	0.2508 (.0445)	0.0764 (.0222)	0.0801 (.0226)
$n = 500$	0.2711 (.0279)	0.0553 (.0142)	0.0573 (.0147)
$n = 1000$	0.2827 (.0218)	0.0451 (.0106)	0.0469 (.0108)
Distributions	$F_{T\|Z} = \text{Exp}(Z)$, $F_C = \text{Exp}(3)$, $F'_Z(z) = 2z^{-2}I\{1 < z < 2\}$		
Censoring %	31.4% − 31.6%		

before getting into the situation with multidimensional covariate $\boldsymbol{Z}$.

No adjustment If we do not do any usual adjustment above mentioned (like the one for the Kaplan-Meier estimator $\hat{F}_{KM}(t)$) for the BNPMLE $\hat{F}_n(t,z)$ in its computation, then it is not a proper d.f. and we denote it as $\hat{F}_{0n}(t,z)$.

Naive adjustment As pointed out above and in Ren and Riddlesworth (2014), one practical issue in the actual computation of BNPMLE $\hat{F}_n(t,z)$ that needs to be noted is that in (2.3) we have $\hat{p}_{m+1,j} > 0$ for some j's, which is the same issue as that with the Kaplan-Meier estimator in the univariate case. Due to equation (2.5), a natural thing to do is to evenly distribute the probability mass $\hat{p}_{m+1,j}$, whenever positive, to points $(U_{m_j+1}, Y_j), \cdots, (U_m, Y_j)$, because in the line segment $\{(t, Y_j) \in \mathbb{R}^2; U_{m_j} < t < \tau_T\}$ we at the best only observe $U_{m_j+1}, \cdots, U_m$ on T from the available data (1.2). We call this the ***naive adjusted*** BNPMLE, still denoted by $\hat{F}_n(t,z)$. All simulation results presented in Ren and Riddlesworth (2014) are conducted for such naive adjusted BNPMLE $\hat{F}_n(t,z)$ with discrete or continuous Z, and they compare well with the bivariate distribution estimator $\hat{F}_{RG}(t,z)$ by Ren and Gu (1997) for right censored data (1.2).

Neighborhood adjustment However, it should be noted that the above method of naive adjustment of the BNPMLE $\hat{F}_n(t,z)$ implies the assumption that the support of $F_0(t,z)$ is rectangular, which may not always be the case in practical situations. A natural way to avoid such issue for continuous Z is to evenly distribute the probability mass $\hat{p}_{m+1,j}$ for each given j, whenever positive, to points V_i's which are greater than U_{m_j} and have corresponding Z_i's fall in the neighborhood of Y_j: $(Y_j - d, Y_j + d)$, where d is the radius of the neighborhood. We call this the ***neighborhood adjusted*** BNPMLE, denoted by $\hat{F}_{nn}(t,z)$.

To compare the performance of $\hat{F}_{0n}(t,z), \hat{F}_n(t,z)$ and $\hat{F}_{nn}(t,z)$ with continuous covariate variable Z, we conduct simulation studies on the example presented in Table 3 of Ren and Riddlesworth (2014). Let $\text{Exp}(\mu)$ represent the exponential distribution with mean μ. Our simulation studies consider right censored data (1.2) with $F_C = \text{Exp}(3)$ as the d.f. of right censoring variable C, $F_{T|Z} = \text{Exp}(Z)$ as the conditional d.f. of T given Z, where Z is a continuous r.v. with p.d.f. $F'_Z(z) = f_Z(z) = 2/z^2$ if $1 < z < 2$; 0, elsewhere. To compare the performance of $\hat{F}_{0n}(t,z), \hat{F}_n(t,z)$ and $\hat{F}_{nn}(t,z)$ with the d.f. $F_0(t,z)$ of (T,Z), we generate 1000 such samples (1.2) with $n = 50, 100, 200, 500, 1000$, respectively. For each n, Table 1 includes the right censoring percentage of the generated samples, and includes the simulation average of $\|\hat{F}_{0n} - F_0\|, \|\hat{F}_n - F_0\|$ and $\|\hat{F}_{nn} - F_0\|$ with the simulation standard deviation (s.d.) given in the parenthesis, where the uniform norm $\|\cdot\|$ is taken over all sample points $(V_i, Z_i), i = 1, \ldots, n$. For the neighborhood adjusted BNPMLE $\hat{F}_{nn}(t,z)$, the radius of the neighborhood we used in our simulation studies is $d = [Z_{(n)} - Z_{(1)}]/\sqrt{n}$.

Remark 3. Note that in the example we consider for Table 1, T and Z are dependent. The simulation results in Table 1 show that the naive adjusted BNPMLE $\hat{F}_n(t,z)$ and the neighborhood adjusted BNPMLE $\hat{F}_{nn}(t,z)$ have similar performances, while non-adjusted BNPMLE $\hat{F}_{0n}(t,z)$ performs poorly. In our additional simulation studies, we also considered different choices of the radius d for $\hat{F}_{nn}(t,z)$, but the results do not appear to be significantly different from what's presented in Table 1. In practice, the neighborhood adjusted BNPMLE $\hat{F}_{nn}$ requires the choice of radius d, thus not so desirable, but it is applicable in situations when the support of $F_0(t,z)$ is not rectangular. The proofs on the consistency of these adjusted BNPMLE's need further studies, and will be considered in a separate paper. In this context, we note that other types of adjustments, such as kernel method, etc., may also be considered as possible remedies for the inconsistency issue of BNPMLE $\hat{F}_n(t,z)$ with continuous covariate Z.

4. PROOFS

Proof of Theorem 1. Without loss of generality, we consider the case that there are no ties among V_i's and there are no ties among Z_i's. Then, we have $m = q = n$ in (2.1), and we may assume:

$$V_1 < V_2 < \cdots < V_n, \tag{4.1}$$

in turn, we have in (2.1):

$$U_i = V_i \quad \text{and} \quad Y_j = Z_{(j)}, \qquad \text{for } i, j = 1, 2, \cdots, n \tag{4.2}$$

where $Z_{(j)}$'s are the order statistics among $Z_1, \cdots, Z_n$. Thus, if we denote

$$k_j = \{\, k \,|\, Z_k = Z_{(j)} \,\}, \qquad j = 1, 2, \cdots, n \tag{4.3}$$

then from (2.2) we obtain

$$\begin{cases} \delta_{ij} = \sum_{k=1}^{n} I\{V_k = V_i, \delta_k = 1, Z_k = Z_{(j)}\} \\ \quad\;\; = \delta_i I\{Z_i = Z_{(j)}\} \\ N_{ij} = \sum_{k=1}^{n} I\{V_k \geq V_i, Z_k = Z_{(j)}\} \\ \quad\;\; = \sum_{k=i}^{n} I\{Z_k = Z_{(j)}\} = I\{k_j \geq i\}. \end{cases} \tag{4.4}$$

Notice that $N_{1j} \equiv 1$ and

$$\frac{I\{Z_i = Z_{(j)}\}}{I\{k_j \geq i\}} = \begin{cases} 1 & \text{if } Z_i = Z_{(j)} \\ 0 & \text{otherwise}\,. \end{cases} \tag{4.5}$$

Hence, $\hat{p}_{ij}$'s in (2.3) can be simplified as:

$$\begin{aligned} \hat{p}_{ij} &= \Big(\frac{\delta_{ij}}{N_{ij}}\Big)\Big(\frac{N_{1j}}{n}\Big)\prod_{k=1}^{i-1}\Big(1 - \frac{\delta_{kj}}{N_{kj}}\Big) \\ &= \frac{\delta_i I\{Z_i = Z_{(j)}\}}{n I\{k_j \geq i\}} \prod_{k=1}^{i-1}\Big(1 - \frac{\delta_k I\{Z_k = Z_{(j)}\}}{I\{k_j \geq k\}}\Big) \\ &= \frac{\delta_i I\{Z_i = Z_{(j)}\}}{n} \prod_{k=1}^{i-1}\Big(1 - \delta_k I\{Z_k = Z_{(j)}\}\Big) \\ &= \frac{\delta_i}{n} I\{Z_i = Z_{(j)}\}. \end{aligned} \tag{4.6}$$

Therefore, in (2.3) the proof follows from

$$\begin{aligned} &\hat{F}_n(t, z) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\delta_i}{n} I\{Z_i = Z_{(j)}\} I\{V_i \leq t, Z_{(j)} \leq z\} \\ &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i I\{V_i \leq t, Z_i = Z_{(j)} \leq z\} \\ &= \frac{1}{n}\sum_{i=1}^{n} \delta_i I\{V_i \leq t\}\Big(\sum_{j=1}^{n} I\{Z_i = Z_{(j)} \leq z\}\Big) \qquad (4.7) \\ &= \frac{1}{n}\sum_{i=1}^{n} \delta_i I\{V_i \leq t, Z_i \leq z\}\Big(\sum_{j=1}^{n} I\{Z_i = Z_{(j)}\}\Big) \\ &= \frac{1}{n}\sum_{i=1}^{n} \delta_i I\{V_i \leq t, Z_i \leq z\} \\ &= \frac{1}{n}\sum_{i=1}^{n} I\{V_i \leq t, \delta_i = 1, Z_i \leq z\}. \end{aligned}$$

□

ACKNOWLEDGMENT

The author is grateful to the Editor and two referees for their comments and suggestions on the earlier draft of this article.

Received 3 July 2016

REFERENCES

Akritas, M. A. (1994). Nearest neighbor estimation of a bivariate distribution under random censoring. *The Annals of Statistics*, Vol. 22, 1299–1327. MR1311977

Dabrowska, D. M. (1988). Kaplan-Meier estimate on the plane. *The Annals of Statistics*, Vol. 16, 1475–1489. MR0964934

Dabrowska, D. M. (1989). Kaplan-Meier estimate on the plane: weak convergence, LIL, and the bootstrap. *Journal of Multivariate Analysis*, Vol. 29, 308–325. MR1004341

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

Lin, D. Y. and Ying, Z. L. (1993). A simple nonparametric estimator of the bivariate survival function under univariate censoring. *Biometrika*, Vol. 80, 573–581. MR1248022

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of American Statistical Association*, Vol. 53, 457–481. MR0093867

Miller, R. G. (1976). Least squared regression with censored data. *Biometrika*, Vol. 63, 449–464. MR0458737

Mykland, P. A. and Ren, J. (1996). Self-consistent and maximum likelihood estimation for doubly censored data. *The Annals of Statistics*, Vol. 24, 1740–1764. MR1416658

Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, Vol. 75, 237–249. MR0946049

Ren, J. and Gu, M. G. (1997). Regression M-estimators for doubly censored data. *The Annals of Statistics*, Vol. 25, 2638–2664. MR1604432

Ren, J. and Riddlesworth, T. (2012). Empirical likelihood ratio confidence intervals for conditional survival probabilities with right censored data. *Statistics and Its Interface*, Vol. 5, 339–343. MR2979213

Ren, J. and Riddlesworth, T. (2014). Empirical likelihood bivariate nonparametric maximum likelihood estimator with right censored data. *The Annals of Institute of Statistical Mathematics*, Vol. 66, 913–930. MR3250822

Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley. MR0838963

van der Laan, M. J. (1996). Efficient estimation in the bivariate censoring model and repairing NPMLE. *The Annals of Statistics*, Vol. 24, 596–627. MR1394977

Jian-Jian Ren
University of Maryland
College Park
USA
E-mail address: jjren@umd.edu