Journal of Siberian Federal University. Mathematics & Physics 2020, 13(4), 422–430

DOI: 10.17516/1997-1397-2020-13-4-422-430
УДК 519.2

On Estimation of Bivariate Survival Function from Random Censored Data

Abdurakhim A. Abdushukurov*
Moscow State University Tashkent Branch
Tashkent, Uzbekistan
Rustamjon S. Muradov†
Namangan Institute of Engineering & Technology
Namangan, Uzbekistan

Received 11.03.2020, received in revised form 08.05.2020, accepted 16.06.2020

Abstract. At present there are several approaches to estimate survival functions of vectors of lifetimes. However, some of these estimators are either inconsistent or not fully defined in the range of joint survival functions. Therefore they are not applicable in practice. In this paper three types of estimates of exponential-hazard, product-limit and relative-risk power structures for the bivariate survival function are considered when the number of summands in empirical estimates is replaced with a sequence of Poisson random variables. It is shown that proposed estimates are asymptotically equivalent.

Keywords: bivariate survival function, Poisson random variables, empirical estimates.

Citation: A.A. Abdushukurov, R.S. Muradov, On Estimation of Bivariate Survival Function from Random Censored Data, J. Sib. Fed. Univ. Math. Phys., 2020, 13(4), 422–430.
DOI: 10.17516/1997-1397-2020-13-4-422-430.

Introduction

The problem of estimation of multivariate distribution (or survival) function from incomplete data was considered from the beginning of 1980's (Campbell (1981), Campbell & Földes (1982), Hanley & Parnes (1983), Horváth (1983), Tsay, Leurgang & Crowley (1986), Burke (1988), Dabrowska (1988, 1989), Gill (1992), Huang (2000), Abdushukurov(2004) etc.) (see, [1–20]). In the special bivariate case there are the numerous examples of paired data that represent life time of individuals (twins or married couples), the failure times of components of a system and others which are subject to random censoring. At present there are several approaches to estimate survival functions of vectors of life times. However, some of these estimators are either nconsistent or not fully defined in the range of joint survival functions. Hence they are not applicable in practice. In this work we present estimators for bivariate survival function and present some sample properties of estimators. We extend some results given in [1–4] to Poisson random summation. At the end of the paper we present consistent estimators of parameters of Marshall-Olkin exponential distribution.

1. Random right censoring model

Let $\mathbb{X}=\{X_i=(X_1 i, X_2 i)\}_{i=1}^{\infty}$ be a sequence of independent and identically distributed (i.i.d.) two-dimensional random vectors with a common continuous survival function $F(s,t)$ =

*abdushukurov1710@gmail.com
†rustamjonmuradov@gmail.com

$= P(X_{11} > s, X_{21} > t)$, $(s,t) \in \overline{R}^{+2} = [0,\infty) \times [0,\infty)$. This sequence is censored from the right by sequence $\mathbb{Y} = \{Y_i = (Y_{1i}, Y_{2i})\}_{i=1}^{\infty}$ of i.i.d. random vectors with survival function $G(s,t) = P(Y_{11} > s, Y_{21} > t)$, $(s,t) \in \overline{R}^{+2}$. Let us assume that there is the sample $\mathbb{V}^{(n)} = \{(Z_i, \Delta_i), 1 \leqslant i \leqslant n\}$, where $Z_i = (Z_{1i}, Z_{2i})$, $\Delta_i = (\delta_{1i}, \delta_{2i})$, $Z_{ki} = \min(X_{ki}, Y_{ki})$, $\delta_{ki} = I(Z_{ki} = X_{ki})$, $k = 1,2$, and $I(\cdot)$ is the indicator. The problem consist of estimating F from the sample $\mathbb{V}^{(n)}$. Let $H(s,t) = P(Z_{1i} > s, Z_{2i} > t)$, $(s,t) \in \overline{R}^{+2}$ and sequences $\mathbb{X}$ and $\mathbb{Y}$ are independent. Then $H(s,t) = F(s,t)\,G(s,t)$, $(s,t) \in \bar{R}^{+2}$. In this paper we use exponential-hazard, product-limit and relative-risk power types functionals in order to construct the corresponding estimates of three types for F. In the empirical estimates the upper index of summation n is replaced by the Poisson random variable (r.v.) μ_n with expectation $E\mu_n = n$. This arises in the insurance business as the size of group insurance payments by an insurance company to customers in connection with an insured event. Following [2], we introduce some auxiliary functionals for $(x,y) \in \bar{R}^{+2}$:

$$M(x,y) = P(Z_{11} \leqslant x,\, Z_{21} > y), \quad N(x,y) = P(Z_{11} > x,\, Z_{21} \leqslant y),$$

$$\bar{M}(x,y) = P(Z_{11} \leqslant x,\, Z_{21} > y,\, \delta_{11} = 1), \quad \bar{N}(x,y) = P(Z_{11} > x,\, Z_{21} \leqslant y,\, \delta_{21} = 1),$$

$$\begin{aligned} \Lambda_1(x,y) &= \int_0^x \frac{M(ds,y)}{H(s-,y)}, \quad \Lambda_2(x,y) = \int_0^y \frac{N(x,dt)}{H(x,t-)}, \\ \bar{\Lambda}_1(x,y) &= \int_0^x \frac{\bar{M}(ds,y)}{H(s-,y)}, \quad \bar{\Lambda}_2(x,y) = \int_0^y \frac{\bar{N}(x,dt)}{H(x,t-)}, \\ \Lambda(x,y) &= \Lambda_1(x,0) + \Lambda_2(x,y), \quad \bar{\Lambda}(x,y) = \bar{\Lambda}_1(x,0) + \bar{\Lambda}_2(x,y), \\ \Lambda^c(x,y) &= \Lambda_1^c(x,0) + \Lambda_2^c(x,y), \quad \bar{\Lambda}^c(x,y) = \bar{\Lambda}_1^c(x,0) + \bar{\Lambda}_2^c(x,y), \end{aligned} \tag{1.1}$$

where

$$\Lambda_1^c(x,y) = \Lambda_1(x,y) - \sum_{s \leqslant x} \Lambda_1(\triangle\, s, y), \quad \Lambda_1(\triangle\, s, y) = \Lambda_1(s,y) - \Lambda_1(s-,y),$$

$$\Lambda_2^c(x,y) = \Lambda_2(x,y) - \sum_{t \leqslant y} \Lambda_2(x, \triangle\, t), \quad \Lambda_2(x, \triangle\, t) = \Lambda_2(x,t) - \Lambda_2(x,t-),$$

and similarly defined $\bar{\Lambda}_1^c$ and $\bar{\Lambda}_2^c$. To construct estimates for F we estimate functionals (1.1). Firstly, we introduce the following empirical estimates of the first four probabilities in (1.1) from the sample $\mathbb{V}^{(n)}$:

$$\begin{aligned} H_n(x,y) &= \frac{1}{n}\sum_{i=1}^n I(Z_{1i} > x,\, Z_{2i} > y), \\ M_n(x,y) &= \frac{1}{n}\sum_{i=1}^n I(Z_{1i} \leqslant x,\, Z_{2i} > y), \\ N_n(x,y) &= \frac{1}{n}\sum_{i=1}^n I(Z_{1i} > x,\, Z_{2i} \leqslant y), \\ \bar{M}_n(x,y) &= \frac{1}{n}\sum_{i=1}^n I(Z_{1i} \leqslant x,\, Z_{2i} > y,\, \delta_{1i} = 1), \\ \bar{N}_n(x,y) &= \frac{1}{n}\sum_{i=1}^n I(Z_{1i} > x,\, Z_{2i} \leqslant y,\, \delta_{2i} = 1). \end{aligned} \tag{1.2}$$

Let $\{\mu_n,\, n \geqslant 1\}$ be a sequence of Poisson random variables (r.v-s.) with parameter $E\mu_n = n$, that is independent of the pair $(\mathbb{X}, \mathbb{Y})$. Along with estimates (1.2), we propose also their analogues

$H_n^*,\ M_n^*,\ N_n^*,\ \bar{M}_n^*,\ \bar{N}_n^*$ obtained from estimates (1.2) by replacing the upper limit of summation n by r.v. μ_n. However, it should be noted that these estimates have the disadvantage because they can be greater than 1. In fact, for example, for

$$H_n^*(x,y) = \frac{1}{n}\sum_{i=1}^{\mu_n} I(Z_{1i} > x,\ Z_{2i} > y),$$

we have

$$P(H_n^*(0,0) > 1) = P(\mu_n > n) = \sum_{m=n+1}^{\infty} \frac{n^m e^{-n}}{m!} > 0.$$

To avoid this disadvantage we consider the following truncated versions of estimates $H_n^*,\ M_n^*,\ N_n^*,\ \bar{M}_n^*, \bar{N}_n^*$:

$$H_n^0(x,y) = 1 - (1 - H_n^*(x,y))\, I(H_n^*(x,y) \leqslant 1) = \begin{cases} H_n^*(x,y) & \text{if} \quad H_n^*(x,y) \leqslant 1, \\ 0 & \text{if} \quad H_n^*(x,y) > 1, \end{cases}$$

and similarly constructed estimates $M_n^0,\ N_n^0,\ \bar{M}_n^0,\ \bar{N}_n^0$. In similar way we construct the corresponding estimates for functionals in (1.1):

$$\begin{gathered} \Lambda_{1n}(x,y) = \int_0^x \frac{M_n^0(ds,y)}{H_n^0(s-,y)}, \quad \Lambda_{2n}(x,y) = \int_0^y \frac{N_n^0(x,dt)}{H_n^0(x,t-)}, \\ \bar{\Lambda}_{1n}(x,y) = \int_0^x \frac{\bar{M}_n^0(ds,y)}{H_n^0(s-,y)}, \quad \bar{\Lambda}_{2n}(x,y) = \int_0^y \frac{\bar{N}_n^0(x,dt)}{H_n^0(x,t-)}, \\ \Lambda_n(x,y) = \Lambda_{1n}(x,0) + \Lambda_{2n}(x,y), \quad \bar{\Lambda}_n(x,y) = \bar{\Lambda}_{1n}(x,0) + \bar{\Lambda}_{2n}(x,y). \end{gathered} \tag{1.3}$$

The relative-risk function is

$$R(x,y) = \frac{\bar{\Lambda}(x,y)}{\Lambda(x,y)}$$

and its estimator is

$$R_n(x,y) = \frac{\bar{\Lambda}_n(x,y)}{\Lambda_n(x,y)}.$$

Using estimates (1.3), we propose the following three estimates of $F(x,y)$ for exponential, product and power structures

$$\begin{aligned} F_{1n}(x,y) &= \exp\left\{-\bar{\Lambda}_n(x,y)\right\} = \exp\left\{-\left(\bar{\Lambda}_{1n}(x,0) + \bar{\Lambda}_{2n}(x,y)\right)\right\}, \\ F_{2n}(x,y) &= \prod_{s \leqslant x}\left(1 - \bar{\Lambda}_{1n}(\triangle\, s, 0)\right) \prod_{t \leqslant y}\left(1 - \bar{\Lambda}_{2n}(x, \triangle\, t)\right), \\ F_{3n}(x,y) &= \left[H_n(x,y)\right]^{R_n(x,y)}. \end{aligned} \tag{1.4}$$

Let $\Delta_n = \left[0, Z_1^{(n)}\right] \times \left[0, Z_2^{(n)}\right] \cap \Delta$, where $Z_k^{(n)} = \max(Z_{k1}, \ldots, Z_{kn})$, $\Delta = \left[0, T_Z^{(1)}\right] \times \left[0, T_Z^{(2)}\right]$, $T_Z^{(k)} = \inf\{t \geqslant 0 : P(Z_{k1} \leqslant t) = 1\},\ k = 1,2$. The following theorem states the asymptotic equivalence of estimates (1.4).

Theorem 1.1. *For all* $(x,y) \in \Delta_n$:

$$(I) \qquad 0 \leqslant F_{1n}(x,y) - F_{2n}(x,y) = \mathrm{O}_p\left(\frac{1}{n}\right).$$

If the survival function G is also continuous on Δ_n then

$$(II) \qquad |F_{1n}(x,y) - F_{3n}(x,y)| = \mathrm{O}_p\left(\left(\frac{\log n}{n}\right)^{1/2}\right).$$

One can also obtain from (I) and (II) that

$$|F_{3n}(x,y) - F_{2n}(x,y)| = \mathrm{O}_p\left(\left(\frac{\log n}{n}\right)^{1/2}\right).$$

To prove Theorem 1.1 we need the following auxiliary statements.

Lemma 1.1. *Let* $\{\mu_n, n \geqslant 1\}-$ *be a sequence of Poisson r.v-s. with expectation* n*. Then for any number* $\varepsilon > 0$ *and for* n *such that*

$$\frac{n}{\log n} \geqslant \frac{\varepsilon}{8\left(1+\dfrac{e}{3}\right)^2}, \quad e = \exp(1), \tag{1.5}$$

the inequality

$$P\left(\frac{|\mu_n - n|}{n} > \frac{1}{2}\left(\frac{\varepsilon}{2}\cdot\frac{\log n}{n}\right)^{1/2}\right) \leqslant 2n^{-c_0}, \tag{1.6}$$

is true, where $c_0 = c_0(\varepsilon) = \varepsilon/16\,(1+e/3)$.

Proof. Let $\gamma_1, \gamma_2, \ldots$ be a sequence of Poisson r.v.-s with expectation $E(\gamma_k) = 1$ for all $k = 1, 2, \ldots$. Then $\mu_n - n = \sum\limits_{k=1}^{n}(\gamma_k - 1) = \sum\limits_{k=1}^{n}\xi_k$, where

$$Ee^{t\xi_k} = e^{-t}Ee^{t\gamma_1} = e^{-t}e^{-1}\sum_{k=0}^{\infty}\frac{(e^t)^k}{k!} = \exp(e^t - (t+1)).$$

Using Taylor expansion of e^t, we have

$$Ee^{t\xi_k} = \exp\left(1 + t + \frac{t^2}{2} + \Psi(t) - (t+1)\right) = \exp\left(\frac{t^2}{2} + \Psi(t)\right),$$

where $\Psi(t) = \dfrac{t^3}{6}\exp(\theta t)$, $0 < \theta < 1$. For $0 \leqslant t \leqslant 1$,we have $t^3 \leqslant t^2$ and consequently $\Psi(t) \leqslant \dfrac{t^3}{6}\cdot e \leqslant e\cdot\dfrac{t^2}{6}$. From here, for $0 \leqslant t \leqslant 1$ we obtain

$$Ee^{t\xi_k} \leqslant \exp\left(\frac{t^2}{2}\left(1+\frac{e}{3}\right)\right) = \exp\left(\frac{\lambda_k}{2}\cdot t^2\right), \quad \lambda_k = 1 + \frac{e}{3}.$$

Then using following exponential inequality for nonidentical distributed r.v.-s of Petrov ([22])

$$P\left(\left|\sum_{k=1}^{n}\xi_k\right| > u\right) \leqslant 2\exp\left(-\frac{u^2}{2}\right), \quad 0 \leqslant u \leqslant N,$$

under $0 \leqslant u = \dfrac{1}{2}\left(\dfrac{\varepsilon}{2}n\log n\right)^{\frac{1}{2}} \leqslant \lambda_k n = N$, we obtain (1.6). □

The following inequality for two-dimensional empirical estimates from [21, p. 292] is used below. Let $C = C(H) = H\left(T_Z^{(1)}, T_Z^{(2)}\right) > 0$.

Lemma 1.2 ([21])**.** *For all real* $z > 0$

$$P\left(\sup_{(x,y)\in\bar{R}^{+2}}|H_n(x,y) - H(x,y)| > zC^2\right) \leqslant \mathrm{V}_z\cdot\left(1+n^2\right)^2\exp\left(-2nz^2\cdot C^4\right), \tag{1.7}$$

where $\mathrm{V}_z = \mathrm{V}_z(H) = 4\exp\left(4zC^2 + 4z^2C^4\right)$.

Corollary 1.1. Let $z = z_0 = \left(\frac{4+\varepsilon}{2} \cdot \frac{\log n}{n}\right)^{1/2} \cdot C^{-2}$ in (2.7). Then

$$P\left(\sup_{(x,y)\in\bar{R}^{+2}} |H_n(x,y) - H(x,y)| > \left(\frac{4+\varepsilon}{2} \cdot \frac{\log n}{n}\right)^{1/2} \cdot C^{-2}\right) \leqslant q_n(\varepsilon), \qquad (1.8)$$

where

$$q_n(\varepsilon) = 4\exp\left(4\left(\frac{4+\varepsilon}{2n} \cdot \log n\right)^{1/2}\left[1 + \left(\frac{4+\varepsilon}{2n} \cdot \log n\right)^{1/2}\right]\right) \cdot \left(n^2+1\right)^2 n^{-(4+\varepsilon)} = \mathrm{O}\left(n^{-\varepsilon}\right).$$

Therefore, for $\varepsilon > 1$ from (1.8) we have by Borel-Cantelli lemma that

$$\sup_{(x,y)\in\Delta_n} |H_n(x,y) - H(x,y)| \overset{a.s.}{=} \mathrm{O}\left(\left(\frac{\log n}{n}\right)^{1/2}\right). \qquad (1.9)$$

In the next lemma we establish an analogue of (1.7) for an empirical estimate H_n^0. Let $q_n^0(\varepsilon)$ be obtained from $q_n(\varepsilon)$ by replacing $4+\varepsilon$ with $(4+\varepsilon)/4$.

Lemma 1.3. *Under the conditions of Lemma 1.1*

$$P\left(\sup_{(x,y)\in\Delta_n} \left|H_n^0(x,y) - H(x,y)\right| > \left(\frac{4+\varepsilon}{2} \cdot \frac{\log n}{n}\right)^{1/2} \cdot C^{-2}\right) \leqslant 2n^{-c_0(\varepsilon+4)} + q_n^0(\varepsilon). \qquad (1.10)$$

Proof. For $\mu_n \leqslant n : H_n^0(x,y) = H_n^*(x,y)$ for all $(x,y) \in \bar{R}^{+2}$ and for $\mu_n > n$ we have

$$\sup_{(x,y)\in\bar{R}^{+2}} |H_n^0(x,y) - H(x,y)| \leqslant \sup_{(x,y)\in\bar{R}^{+2}} |H_n^*(x,y) - H(x,y)|.$$

Using the formula of complete probability, we obtain

$$P\left(\sup_{(x,y)\in\Delta_n} \left|H_n^0(x,y) - H(x,y)\right| > z_0 C\right) \leqslant P\left(\sup_{(x,y)\in\Delta_n} |H_n^*(x,y) - H(x,y)| > z_0 C^2\right) \leqslant$$

$$\leqslant P\left(\sup_{(x,y)\in\Delta_n} \left|H_n(x,y) - H(x,y) + \frac{1}{n}\sum_{i=n+1}^{\mu_n} I(Z_{1i} > x, Z_{2i} > y)\right| > z_0 C^2/\mu_n > n\right) \cdot P(\mu_n > n) +$$

$$+P\left(\sup_{(x,y)\in\Delta_n} \left|H_n(x,y) - H(x,y) - \frac{1}{n}\sum_{i=n+1}^{\mu_n} I(Z_{1i} > x, Z_{2i} > y)\right| > z_0 C^2/\mu_n > n\right) P(\mu_n > n) \leqslant$$

$$\leqslant P\left(\sup_{(x,y)\in\Delta_n} |H_n(x,y) - H(x,y)| > \frac{1}{2} z_0 C^2\right) +$$

$$+P\left(\sup_{(x,y)\in\Delta_n} \left|\frac{1}{n}\sum_{i=n\wedge\mu_n+1}^{n\vee\mu_n} I(Z_{1i} > x, Z_{2i} > y)\right| > \frac{1}{2} z_0 C^2\right) \leqslant$$

$$\leqslant q_n^0(\varepsilon) + P\left(\frac{|\mu_n - n|}{n} > \frac{1}{2} z_0 C^2\right) \leqslant 2n^{-c_0(\varepsilon+4)} + q_n^0(\varepsilon),$$

where (1.6) and (1.8) are used. □

Proof of Theorem 1.1. From inequalities (2.4.2) in [2] applied to estimates F_{1n} and F_{2n} we have

$$0 \leqslant F_{1n}(x,y) - F_{2n}(x,y) \leqslant \frac{1}{2}\sum_{i=1}^{\mu_n-1}\left[\left(q_{1i}^{(n)}(x,0)\right)^2 + \left(q_{2i}^{(n)}(x,y)\right)^2\right] =$$

$$= \frac{1}{2n^2}\left\{\sum_{i=1}^{\mu_n-1}\left[\frac{\delta_{1(i)}\, I\left(Z_1^{(i)} \leqslant x\right)}{\left(S_{1n}^{0Z}\left(Z_1^{(i)}-\right)\right)^2}\right] + \sum_{i=1}^{\mu_n-1}\left[\frac{\delta_{2(i)}\, I\left(Z_1^{(i)} \leqslant x, Z_{2i} \leqslant y\right)}{\left(H_n^0\left(x, Z_2^{(i)}-\right)\right)^2}\right]\right\} \leqslant \tag{1.11}$$

$$\leqslant \frac{\mu_n}{2n^2}\left[S_{1n}^{0Z}\left(Z_1^{(\mu_n-1)}-\right)^{-2} + \left(H_n^0\left(x, Z_2^{(\mu_n-1)}-\right)\right)^{-2}\right],$$

where $Z_k^{(1)} \leqslant \ldots \leqslant Z_k^{(n)}$ order statistics are constructed from Z_{ki}, $k = 1, 2$, $\delta_{k(i)}$ corresponds to $Z_k^{(i)}$ and $S_{1n}^{0Z}(x) = H_n^0(x; 0)$. It is known that for $n \to \infty$, $Z_k^{(n)} \xrightarrow{p} T_Z^{(k)}$, $k = 1, 2$. We show that $Z_k^{(\mu_n)} \xrightarrow{p} T_Z^{(k)}$, $k = 1, 2$ when $n \to \infty$. For $\varepsilon > 0$, $0 < \delta < 1$ and $k = 1, 2$ we have

$$P\left(\left|Z_k^{(\mu_n)} - T_Z^{(k)}\right| > \varepsilon\right) \leqslant$$

$$\leqslant P\left(\left|Z_k^{(\mu_n)} - T_Z^{(k)}\right| > \varepsilon,\ \left|\frac{\mu_n}{n} - 1\right| < \delta\right) + P\left(\left|\frac{\mu_n}{n} - 1\right| \geqslant \delta\right) \leqslant$$

$$\leqslant P\left(\left|Z_k^{(\mu_n)} - T_Z^{(k)}\right| > \varepsilon,\ n(1-\delta) < \mu_n < n(1+\delta)\right) + P\left(\left|\frac{\mu_n}{n} - 1\right| \geqslant \delta\right) \leqslant$$

$$\leqslant P\left(\left|Z_k^{(n)} - T_Z^{(k)}\right| > \varepsilon\right) + P\left(\left|\frac{\mu_n}{n} - 1\right| \geqslant \delta\right).$$

For arbitrary $\eta > 0$ there are numbers n_1 and ε such that for $n \geqslant n_1$

$$P\left(\left|Z_k^{(n)} - T_Z^{(k)}\right| > \varepsilon\right) < \frac{\eta}{2},\ \ k = 1, 2. \tag{1.12}$$

Since $P\left(\left|\frac{\mu_n}{n} - 1\right| \geqslant \delta\right) \to 0$ when $n \to \infty$ then for $n \geqslant n_2$

$$P\left(\left|\frac{\mu_n}{n} - 1\right| \geqslant \delta\right) < \frac{\eta}{2}. \tag{1.13}$$

Then for $n \geqslant n_0 = \max(n_1, n_2)$ we obtain from (1.12) and (1.13) that

$$P\left(\left|Z_k^{(\mu_n)} - T_Z^{(k)}\right| > \varepsilon\right) < \eta, \tag{1.14}$$

which is required result. Thus, taking into account (1.13) and (1.14), for $n \to \infty$ with probability close to 1 we have

$$\begin{gathered} Z_k^{(\mu_n-1)} \approx Z_k^{(n-1)},\ \ k = 1, 2, \\ \frac{\mu_n}{n^2} = \mathrm{O}_p\left(\frac{1}{n}\right). \end{gathered} \tag{1.15}$$

Taking into account (1.15) and the following relations obtained from (1.10) for $(x, y) \in \Delta_n$

$$\frac{1}{S_{1n}^{0Z}(x)} \leqslant \frac{\left|S_{1n}^{0Z}(x) - S_1^Z(x)\right|}{S_{1n}^{0Z}(x)\, S_1^Z(x)} + \frac{1}{S_1^Z(x)} = \frac{1}{S_1^Z(x)} + \mathrm{O}_p\left(\left(\frac{\log n}{n}\right)^{1/2}\right),$$

$$\frac{1}{H_n^0(x,y)} \leqslant \frac{\left|H_n^0(x,y) - H(x,y)\right|}{H_n^0(x,y)\, H(x,y)} + \frac{1}{H(x,y)} = \frac{1}{H(x,y)} + \mathrm{O}_p\left(\left(\frac{\log n}{n}\right)^{1/2}\right),$$

we obtain the right estimate in (I). Now according to the inequality $|u-v| \leqslant |\log u - \log v|$, for $0<u,\ v\leqslant 1,\ 0\leqslant R_n(x,y)\leqslant 1$ and $(x,y)\in\Delta_n$ we have

$$\begin{aligned} |F_{1n}(x,y)-F_{3n}(x,y)| &\leqslant \bar{\Lambda}_n(x,y)\left|-1+\frac{\left(-\log H_n^0(x,y)\right)}{\Lambda_n(x,y)}\right| = \\ &= R_n(x,y)\left|\left(-\log H_n^0(x,y)\right)-\Lambda_n(x,y)\right| \leqslant \\ &\leqslant \left|-\log H_n^0(x,y)+\log H(x,y)\right| + \left|\left(-\log H(x,y)\right)-\Lambda_n(x,y)\right|. \end{aligned} \tag{1.16}$$

According to Lemma 1.3 and the mean value theorem for $(x,y)\in\Delta_n$ we obtain

$$\left|-\log H_n^0(x,y)+\log H(x,y)\right| = \mathrm{O}_p\left(\left(\frac{\log n}{n}\right)^{1/2}\right). \tag{1.17}$$

Taking into account continuity of G, Lemma 3.4.3, the proof of Theorem 2.4.3 and Remark 2.4.4 in [2] we obtain for $(x,y)\in\Delta_n$ that

$$\left|-\log H(x,y)-\Lambda_n(x,y)\right| = \mathrm{O}_p\left(\left(\frac{\log n}{n}\right)^{1/2}\right). \tag{1.18}$$

Now (II) follows from relations (2.16)–(2.18). □

It was shown in Theorem 2.4.3 in [2] that in the case of continuity of F and G both exponential-hazard and relative-risk power functionals coincide with the estimated survival function F. Then, taking into account Theorem 1.1, we can state that all three estimates (1.4) are consistent estimates of F (see, also [5]).

2. Estimation of parameters of Marshall-Olkin exponential distribution

Let us consider survival function $F(s,t)=P(X_{11}>s, X_{21}>t)$, $(s,t)\in\overline{R}^{+2}$ of Marshall-Olkin exponential form with unknown parameters $\lambda_1,\lambda_2,\lambda_{12}$:

$$F(s,t)=\exp\left(-\lambda_1 s-\lambda_2 t-\lambda_{12}\max(s,t)\right), \quad (s,t)\in\overline{R}^{+2}. \tag{2.1}$$

Then corresponding cumulative hazard function is

$$\Lambda(s,t)=-\log F(s,t)=\lambda_1 s+\lambda_2 t+\lambda_{12}\max(s,t). \tag{2.2}$$

Nonparametric estimator of $\Lambda(s,t)$ from (2.4) is $\overline{\Lambda}_n(s,t)=-\log F_{1n}(s,t)=\overline{\Lambda}_{1n}(s,0)+\overline{\Lambda}_{2n}(s,t)$. It is easy to verify from (2.2) that we have the system of equations for $s>0$

$$\begin{cases} \Lambda(s,0)=\lambda_1 s+\lambda_{12}s, \\ \Lambda(0,s)=\lambda_2 s+\lambda_{12}s, \\ \Lambda(s,s)=\lambda_1 s+\lambda_2 s+\lambda_{12}s. \end{cases} \tag{2.3}$$

From (2.3) we find expressions for unknown parameters λ_1,λ_2 and λ_{12} for a fixed point $s=s_0>0$:

$$\begin{cases} \lambda_1=\dfrac{1}{s_0}\left(\Lambda(s_0,s_0)-\Lambda(0,s_0)\right), \\ \lambda_2=\dfrac{1}{s_0}\left(\Lambda(s_0,s_0)-\Lambda(s_0,0)\right), \\ \lambda_{12}=\dfrac{1}{s_0}\left(\Lambda(s_0,0)+\Lambda(0,s_0)-\Lambda(s_0,s_0)\right). \end{cases} \tag{2.4}$$

Now we obtain estimators of parameters from (2.4) by replacing Λ with $\overline{\Lambda}_n$:

$$\begin{cases} \lambda_1^{(n)} = \dfrac{1}{s_0}\left(\overline{\Lambda}_n\left(s_0, s_0\right) - \overline{\Lambda}_n\left(0, s_0\right)\right), \\ \lambda_2^{(n)} = \dfrac{1}{s_0}\left(\overline{\Lambda}_n\left(s_0, s_0\right) - \overline{\Lambda}_n\left(s_0, 0\right)\right), \\ \lambda_{12}^{(n)} = \dfrac{1}{s_0}\left(\overline{\Lambda}_n\left(s_0, 0\right) + \overline{\Lambda}_n\left(0, s_0\right) - \overline{\Lambda}_n\left(s_0, s_0\right)\right). \end{cases} \tag{2.5}$$

It follows from Theorem 1.1 that $\overline{\Lambda}_n(s,t)$ is consistent estimator of $\Lambda(s,t)$. Consequently, relations (2.5) give consistent estimators of corresponding parameters (2.4) of distribution (2.1).

References

[1] A.A.Abdushukurov, Nonparametrical estimators of survival function on the plane by censored observations, *Uzbek mathematical journal*, **2**(2004), 3–11 (in Russian).

[2] A.A.Abdushukurov, Estimates of unknown distributions from incomplete observations and their properties, LAMBERT Academic Publishing, 2011, (in Russian).

[3] A.A.Abdushukurov, R.S.Muradov Estimation of two-dimensional survival function by random right censored Data, Proc. of XI-th Intern. Conf. Computer Data Analysis and Modelling, Minsk, Belorus, 2016, 85–86.

[4] A.A.Abdushukurov, Estimation of joint survival function from censored observations, *Industrial Lab. Diagn. Mater.*, **82**(2016), 80–84 (in Russian).

[5] A.A.Abdushukurov, R.S.Muradov On the estimates of the distribution function in a random censorship model, *Industrial Lab. Diagn. Mater.*, **80**(2014), 62–67 (in Russian).

[6] A.A.Abdushukurov, R.S.Muradov On Estimation of Conditional Distribution Function under Dependent Random Right Censored Data, *Journal of Siberian Federal University*, **7**(2014), 409–416.

[7] A.A.Abdushukurov, R.S.Muradov Estimation of survival functions from dependent random right censored data, *Intern. J. Innovation in Science and Mathematics*, **2**(2014), 280–287.

[8] A.A.Abdushukurov, R.S.Muradov, Some algebraic properties of Archimedean copula functions and their applications in the statistical estimation of the survival function, *New Trends in Math. Sci.*, **4**(2016), 88–98. DOI:10.20852/ntmsci.2016318808

[9] M.D.Burke, Estimation of a bivariate distribution function under random censorship, *Biometrika*, **75**(1988), 379–382.

[10] G.Campbell, Nonparametric bivariate estimation with randomly censored data, *Biometrika*, **68**(1981), 417–422.

[11] G.Campbell, A.Földes Lange sample properties of nonparametric bivariate estimators with censored data, *Colloquia Mathematica-Societatis János Bolyai*, **32**(1982), 103–122.

[12] D.M.Dabrowska, Kaplan–Meier estimate on the plane, *Ann. Statist.*, **16**(1988), 1475–1489.

[13] D.M.Dabrowska, Kaplan–Meier estimate on the plane: weak convergence, LIL and the bootstrap, *J. Multivar. Anal.*, **29**(1989), 308–325.

[14] R.D.Gill, Multivariate Survival Analysis I, *Theory Probab. Appl.*, **37**(1992), 19–35 (in Russian).

[15] R.D.Gill, Multivariate Survival Analysis II, *Theory Probab. Appl.*, **37**(1992), 307–328 (in Russian).

[16] J.A.Hanley, M.N.Parnes, Nonparametric estimation of a multivariate distribution in the presence of censoring, *Biometrica*, **39**(1983), 129–139.

[17] L.Horváth, The rate of strong uniform consistency for the multivariate product-limit estimator, *J. Multivar. Anal.*, **13**(1983), 202–209.

[18] Y.Huang, Two–sample multistate accelerated sojourn times model, *J.A.S.A.*, **95**(2000), 619–627.

[19] R.S.Muradov, A.A.Abdushukurov, Estimation of multivariate distributions and there mixes by incomplete data, LAMBERT Academic Publishing, Germany, 2011 (in Russian).

[20] W.Y.Tsai, S.Leurgans, J.Crowley, Nonparametric estimation of a bivariate survival function in the presence of censoring, *Ann. Statist.*, **14**(1986), 1351–1365.

[21] J.D.Fermanian, Multivariate hazard rates under random censorship, *J. Multivar. Anal.*, **62**(1997), 273–309.

[22] V.V.Petrov, Limit Theorems of Probability Theory: Sequence of independent r.v.-s., Oxford Studies in Probability-4, Vol. 4, Clerandon Press, Oxford, 1995.

Об оценивании двумерной функции выживания по случайно цензурированным данным

Абдурахим А. Абдушукуров
МГУ, Ташкентский филиал
Ташкент, Узбекистан
Рустамжон С. Мурадов
Наманганский инженерно-технический институт
Наманган, Узбекистан

Аннотация. В настоящее время существует несколько подходов к оценке функций выживания векторов времени жизни. Однако некоторые из этих оценок либо являются несостоятельными, либо не полностью определены в области функций совместного выживания и поэтому не применимы на практике. В работе авторами предложены состоятельные оценки совместной функции выживания экспоненциальной, множительной и степенной структур при случайном пуассоновском объёме выборки. Показано, что эти оценки асимптотически эквивалентны.

Ключевые слова: двумерная функция выживания, пуассоновские случайные величины, эмпирические оценки.