On copula type estimators of survival and mean residual life functions under random right censored observations at fixed covariate values

R. S. Muradov [1]

[1]Namangan Institute of Engineering and Technology, Namangan, Uzbekistan, rustamjonmuradov@gmail.com

Let's consider the case when the support of covariate C is the interval $[0,1]$ and we describe our results on fixed design points $0 \leq x_1 \leq x_2 \leq \cdots \leq x_n \leq 1$ at which we consider responses (survival or failure times) $X_1, ..., X_n$ and censoring times $Y_1, ..., Y_n$ of identical objects, which are under study. These responses are independent and nonnegative random variables (r.v.-s) with conditional distribution function (d.f.) at x_i, $F_{x_i}(t) = P(X_i \leq t/C_i = x_i)$. They are subjected to random right censoring, that is for X_i there is a censoring variable Y_i with conditional d.f. $G_{x_i}(t) = P(Y_i \leq t/C_i = x_i)$ and at n-th stage of experiment the observed data is

$$S^{(n)} = \{(Z_i, \delta_i, C_i), 1 \leq i \leq n\},$$

where $Z_i = min(X_i, Y_i), \delta_i = I(X_i \leq Y_i)$ with $I(A)$ denoting the indicator of event A. Note that in sample $S^{(n)}$ r.v. X_i is observed only when $\delta_i = 1$. Commonly, in survival analysis to assume independence between the r.v.-s X_i and Y_i conditional on the covariate C_i. But, in some practical situations, this assumption does not hold. Therefore, in this article we consider a dependence model in which dependence structure is described through copula function. So let

$$S_x(t_1, t_2) = P(X_x > t_1, Y_x > t_2), t_1, t_2 \geq 0,$$

the joint survival function of the response X_x and the censoring variable Y_x at x. Then the marginal survival functions are $S_x^X(t) = 1 - F_x(t) = S_x(t, 0)$ and $S_x^Y(t) = 1 - G_x(t) = S_x(0, t), t \geq 0$. We suppose that the marginal d.f.-s F_x and G_x are continuous. Then according to the Theorem of Sklar [7], the joint survival function $S_x(t_1, t_2)$ can be expressed as

$$S_x(t_1, t_2) = C_x(S_x^X(t_1), S_x^Y(t_2)), t_1, t_2 \geq 0, \tag{1}$$

where $C_x(u, v)$ is a known copula function depending on x, S_x^X and S_x^Y in a general way.

Assume that at the fixed design value $x \in (0,1)$, C_x in (2) is Archimedean copula, i.e.

$$S_x(t_1, t_2) = \varphi_x^{[-1]}(\varphi_x(S_x^X(t_1)) + \varphi_x(S_x^Y(t_2))), t_1, t_2 \geq 0, \tag{2}$$

where copula generator function φ_x is strict, i.e. $\varphi_x(0)=\infty$ and hence $\varphi_x^{[-1]}=\varphi_x^{-1}$. From (3), it follows that

$$P(Z_x>t)=1-H_x(t)=\overline{H_x(t)}=S_x^Z(t)=S_x(t,t)=$$
$$=\varphi_x^{-1}(\varphi_x(S_x^X(t))+\varphi_x(S_x^Y(t))),\ \ t\geq 0, \tag{3}$$

Let $H_x^{(1)}(t)=P(Z_x\leq t,\delta_x=1)$ be a subdistribution function and $\Lambda_x(t)$ is crude hazard function of r.v. X_x subjecting to censoring by Y_x, that is

$$\Lambda_x(dt)=\frac{P(X_x\in dt, X_x\leq Y_x)}{P(X_x\geq t, Y_x\geq t)}=\frac{H_x^{(1)}(dt)}{S_x^Z(t-)}. \tag{4}$$

From (4) and (5) one can obtain following expression of survival function S_x^X :

$$S_x^X(t)=\varphi_x^{-1}[-\int_0^t S_x^Z(u-)\varphi_x'(S_x^Z(u))d\Lambda_x(u)]=$$
$$=\varphi_x^{-1}[-\int_0^t \varphi_x'(S_x^Z(u))dH_x^{(1)}(u)],\ \ t\geq 0, \tag{5}$$

(see, for example, [3,4]). In order to constructing the estimator of S_x^X according to representation (6), we introduce some smoothed estimators of $S_x^Z, H_x^{(1)}$ and regularity conditions for them. Similarly to Breakers and Veraverbeke [5], we will also use the Gasser-Müller weights

$$\omega_{ni}(x,h_n)=\frac{1}{q_n(x,h_n)}\int_{x_{i-1}}^{x_i}\frac{1}{h_n}\pi(\frac{x-z}{h_n})dz,\ \ i=1,...,n,$$

with

$$q_n(x,h_n)=\int_0^{x_n}\frac{1}{h_n}\pi(\frac{x-z}{h_n})dz,$$

where $x_0=0$, π is a known probability density function(kernel) and $\{h_n, n\geq 1\}$ is a sequence of positive constants, tending to zero as $n\to\infty$, called bandwidth sequence. Let's introduce the weighted estimators of H_x, S_x^Z and $H_x^{(1)}$ respectively as

$$H_{xh}(t)=\sum_{i=1}^n \omega_{ni}(x,h_n)I(Z_i\leq t),$$
$$S_{xh}^Z(t)=1-H_{xh}(t), \tag{6}$$
$$H_{xh}^{(1)}(t)=\sum_{i=1}^n \omega_{ni}(x,h_n)I(Z_i\leq t,\delta_i=1).$$

Then pluggin in (6) estimators (7) we get corresponding estimator of $S_x^X(t)$ as

$$S_{xh}^X(t)=1-F_{xh}(t)=\varphi_x^{-1}[-\int_0^t \varphi_x'(S_{xh}^Z(u))dH_{xh}^{(1)}(u)],\ \ t\geq 0, \tag{7}$$

Let

$$E(t) = E(t; S^X) = E\left(X_1 - t/X_1 > t\right) = (S^X(t))^{-1} \cdot \int_t^{+\infty} S^X(y)dy, \ t \in [0, T_X] \, ,$$

is mean residual life function of r.v. X_1. Consider estimate of $E\left(t\right)$:

$$E_{xh}(t) = \begin{cases} E\left(t; S_{xh}^X\right), \ x \in [0,1)\,, \\ \quad 0, \quad t \geq 0. \end{cases} \tag{8}$$

Remark that in the case of no covariate, estimator (7) reduces to estimator first obtained by Zeng and Klein [9]. In the case of the independent copula $\varphi(y) = -logy$, Zeng and Klein estimate reduces to a exponential-hazard estimate (see, [3,8,9]). Also it is well-known that under independent censoring case Kaplan-Meier's [6] product-limit estimator and exponential-hazard estimators are asymptotical equivalent. Therefore, we will show that estimator (7) and copula-graphic estimator of Breakers and Veraverbeke [5] have the same asymptotic behaviours(see, [3]).In this article we also state our result on consistency of estimator (8).

We note that in paper Abdushukurov [4] author proposed a new another estimator of survival function in the presence of covariate and studied its large sample properties. This estimator is an extended analogue of relative-risk power estimator introduced in Abdushukurov [4].

Acknowledgements. The author is supported by the grant UZB-Ind-2021-97.

References

1. A. A. Abdushukurov, Nonparametric estimation of the distribution function based on relative risk function, *Communications in Statistics : Theory and Methods.* (1998), **27**:8, 1991-2012.

2. A. A. Abdushukurov, On Nonparametric Estimation of reliability indices by censored samples, *Theory of Probability and its applications.* (1999), **43**:1, 3-11.

3. A. A. Abdushukurov, R. S. Muradov, On estimation of conditional distribution function under dependent random right censored observations. *Journal of Siberian Federal University. Mathematics & Physics.* (2014), **7**:4, p.409-416.

4. A. A. Abdushukurov, Estimation of conditional survival function under dependent right random censored data, *Austrian Journal of Statistics.* (2020), **49**, 1-8.

5. R. Breakers,N. Veraverbeke, A copula-graphic estimator for the conditional survival function under dependent censoring, *The Canadian Journal of Statistics*, (2005), **33**:3, 429-447.

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

7. R. B. Nelsen, *An Introduction to Copulas*, Springer, New York,1999.

8. L. P. Rivest, M. T. Wells, A martingale approach to the copula-graphic estimator for the survival function under dependent censoring, *Journal of Multivariate Analysis*, (2001),**79**, 138-155.

9. M. Zeng, J. P. Klein Estimates of marginal survival for dependent competing risks based on an assumed copula, *Biometrika*,(1995),**82**, 127-138.