ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
2025 Управление, вычислительная техника и информатика № 71
Tomsk State University Journal of Control and Computer Science

ОБРАБОТКА ИНФОРМАЦИИ

DATA PROCESSING

Original article
UDC 519.234
doi: 10.17223/19988605/71/6

On analysis of survival function estimators with its identifiability with model of right random censorship

Abdurakhim A. Abdushukurov[1], Sukhrob B. Bozorov[2]

[1] *Moscow State University, Branch in Tashkent, Tashkent, Uzbekistan*
[2] *Gulistan State University, Gulistan, Uzbekistan*
[1] *a_abdushukurov@rambler.ru*
[2] *suxrobbek_8912@mail.ru*

Abstract. It is considered random censorship from the right. In this model we investigate basically three types of estimators of exponential, product and power types. We illustrative study all these estimators including its presmothed modifications comparatively. We demonstrate confidence bands for considered estimators by using natural and artificial censored observation. It is showed that presmoothed power estimators have some peculiarities with respect to exponential and product-type estimators. It is shown that only power estimator has identifiability property with respect to right random censoring model in the case of a finite-size sample.

Keywords: model of right random censorship; survival function; exponential; product and power estimators; presmoothed kernel estmators.

For citation: Abdushukurov, A.A., Bozorov, S.B. (2025) On analysis of survival function estimators with its identifiability with model of right random censorship. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitelnaja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 71. pp. 58–70. doi: 10.17223/19988605/71/6

Научная статья
doi: 10.17223/19988605/71/6

Анализ оценок функции выживания с учетом их идентифицируемости с моделью случайного цензурирования справа

Абдурахим Ахмедович Абдушукуров[1], Сухроб Баходирович Бозоров[2]

[1] *Московский государственный университет, филиал в Ташкенте, Ташкент, Узбекистан*
[2] *Гулистанский государственный университет, Гулистан, Узбекистан*
[1] *a_abdushukurov@rambler.ru*
[2] *suxrobbek_8912@mail.ru*

Аннотация. Рассматривается модель случайного цензурирования справа, в которой исследуются сглаженные модификации экспоненциальной, множительной и степенной оценок функции выживания. Демонстрируются доверительные полосы для естественных и искусственных наблюдений для рассматриваемых оценок. Показано, что сглаженные степенные оценки обладают рядом преимуществ по сравнению с экспоненциальной и множественной оценками. Показано, что именно степенная оценка обладает свойством идентифицируемости с моделью случайного цензурирования справа в случае выборки конечного объема.

Ключевые слова: модель случайного цензурирования справа; функция выживания; экспоненциальная; множительная и степенные оценки; сглаженные ядерные оценки.

***Для цитирования*:** Абдушукуров А.А., Бозоров С.Б., Анализ оценок функции выживания с учётом их идентифицируемости с моделью случайного цензурирования справа // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. № 71. С. 58–70. doi: 10.17223/19988605/71/6

Introduction

Paper is devoted to consider random censorship model from the right. Suppose that $X_1, X_2, ...$ and $Y_1, Y_2, ...$ are two sequences of independent random variables (r.v.'s) with corresponding distribution function (d.f.) $F(t)$ and $G(t)$, $t \in R$, respectively. Both sequences of X_k's and Y_k's are mutually independent. Observation is available sample of size *n*:

$$\mathbb{C}^{(n)} = \{(Z_k, \delta_k), 1 \le k \le n\},$$

where $Z_k = \min(X_k, Y_k)$ and indicator $\delta_k = \mathrm{I}(X_k \le Y_k) = I(Z_k = X_k)$.

- If $X_k \le Y_k$, then $Z_k = \min(X_k, Y_k) = X_k$ is equal to $\delta_k = 1$ and in this case we can observe *X*.
- Otherwise if $Y_k \le X_k$, then $Z_k = \min(X_k, Y_k) = Y_k$ is equal to $\delta_k = 0$, and this can be censoring condition.

The main problem has been considered by many authors since 1950s. We comparatively study of survival function estimators taking into account its property of identifiability with model of right random censorship. Observe that in paper [1] authors proposed a new histogram type estimators for the distribution density using the kernel method by using presmoothed estimator for survival function. Remains relevant to considering identifiability of presmoothed power estimator with respect to exponential and product type estimators.

1. Proposed estimators

In sample $\mathbb{C}^{(n)}$, the number of observed X_k's is equal to summa $\nu(n) = \sum_{k=1}^{n} \delta_k$, which is binomial r.v. $Bi(n; p)$ with $\mathrm{P}(\delta_k = 1) = p$. The first, product-limit (PL) estimator has been proposed by E. Kaplan and P. Meier in [1] by formula

$$F_n^{KM}(t) = \begin{cases} 1 - \prod\limits_{\{k: Z_{(k)} \le t\}} \left(1 - \dfrac{\delta_{(k)}}{n-k+1}\right), & t \le Z_{(n)}, \\ 1, & t > Z_{(n)}, \ \delta_{(n)} = 1, \\ undefined, & t > Z_{(n)}, \ \delta_{(n)} = 0. \end{cases} \tag{1}$$

Here is $Z_{(1)} \le Z_{(2)} \le ... \le Z_{(n)}$-ordered sample of $\{Z_k, 1 \le k \le n\}$ and $\delta_{(k)}$ accompanies $Z_{(k)}$'s. Estimator (1) has a great influence in development of statistical theory of censored observation. In statistical literature one can meet several modificated variants of estimator (1) and they are used by many scientists [2–8]. In order to consider estimator we introduce d.f. of minima as

$$P(Z_k \le t) = H(t) = 1 - (1 - F(t))(1 - G(t)) = H_{(0)}(t) + H_{(1)}(t), \ t \in R, \tag{2}$$

where $H_{(m)}(t) = \mathrm{P}(Z_k \le t, \ \delta_k = m)$, $m = 0,1$ are subdistribution functions. It is not difficult to see that

$$H_{(0)}(t) = \int_{-\infty}^{t} (1 - F(u)) dG(u) \ \text{ and } \ H_{(1)}(t) = \int_{-\infty}^{t} (1 - G(u)) dF(u).$$

Let's consider integral hazard function corresponding to d.f.'s $F(t)$, $G(t)$ and $H(t)$ by formulas

$$\Lambda_{(1)}(t)=\int_{-\infty}^{t}\frac{dH_{(1)}(u)}{1-H(u)}=-\log\left(1-F(t)\right),\ \Lambda_{(0)}(t)=\int_{-\infty}^{t}\frac{dH_{(0)}(u)}{1-H(u)}=-\log\left(1-G(t)\right),$$
$$\Lambda(t)=\int_{-\infty}^{t}\frac{dH(u)}{1-H(u)}=\Lambda_{(0)}(t)+\Lambda_{(1)}(t)=-\log\left(1-H(t)\right) \tag{3}$$

and its estimators [5, 6, 9, 10]:

$$\Lambda_{(1)n}(t)=\int_{-\infty}^{t}\frac{dH_{(1)n}^{E}(u)}{1-H_{n}^{E}(u)+\frac{1}{n}},\ \Lambda_{(0)n}(t)=\int_{-\infty}^{t}\frac{dH_{(0)n}^{E}(u)}{1-H_{n}^{E}(u)+\frac{1}{n}},$$
$$\Lambda_{n}(t)=\int_{-\infty}^{t}\frac{dH_{n}^{E}(u)}{1-H_{n}^{E}(u)}=\Lambda_{(0)n}(t)+\Lambda_{(1)n}(t), \tag{4}$$

where summand $\frac{1}{n}$ in denominators of integrals (4) is adden in order to prevent dividing to zero. Here the empirical estimators of $H(t)$ and $H_{(m)}(t), m=0,1$ are

$$H_{n}^{E}(t)=\frac{1}{n}\sum_{k=1}^{n}I\left(Z_k\le t\right)=H_{(0)n}^{E}(t)+H_{(1)n}^{E}(t),\ t\in R \text{ and } H_{(m)n}^{E}(t)=\frac{1}{n}\sum_{k=1}^{n}I\left(Z_k\le t,\ \delta_k=m\right),\ m=0,1.$$

By formulas (3) and (4) one can see easily that natural estimator of d.f. $F(t)$ is following B. Altshuler [11] and N. Breslow [3] estimator of exponential form:

$$F_{n}^{AB}(t)=1-\exp\left(-\Lambda_{(1)n}(t)\right)=1-\exp\left\{-\frac{1}{n}\sum_{k=1}^{n}\frac{I\left(Z_k\le t,\ \delta_k=1\right)}{1-H_{n}^{E}(Z_k)+\frac{1}{n}}\right\}=$$
$$=1-\exp\left(-\sum_{\{k:Z_{(k)}\le t\}}\frac{\delta_{(k)}}{n-k+1}\right)=1-\left(\prod_{\{k:Z_{(k)}\le t\}}\exp\left(-\frac{1}{n-k+1}\right)\right)^{\delta_{(k)}},\ t\in R. \tag{5}$$

In order to compare with estimator (5) we use following modification and very popular form of Kaplan-Meier's estimator (1):

$$F_{n}^{KM}(t)=\begin{cases}1-\prod\limits_{\{k:Z_{(k)}\le t\}}\left(1-\frac{\delta_{(k)}}{n-k+1}\right), & t\le Z_{(n)},\\ 1, & t>Z_{(n)},\end{cases}=\begin{cases}\prod\limits_{\{k:Z_{(k)}\le t\}}\left(\frac{n-k}{n-k+1}\right)^{\delta_{(k)}}, & t\le Z_{(n)},\\ 1, & t>Z_{(n)}.\end{cases} \tag{6}$$

Because of approximate equality $e^{-x}\approx 1-x$ under $x\approx 0$ from (5) and (6) we have that

$$\Lambda_{(1)n}(t)\approx-\log\left(1-F_{n}^{KM}(t)\right),\ t<T_{H_{n}^{E}}=\inf\left\{t:H_{n}^{E}(t)=1\right\}. \tag{7}$$

Approximate equality (7) plays a key role in investigating of several asymptotic results for estimator (6) or (1). Author [12–14] proposed a relative risk power estimator $F_{n}^{RR}(t)$ by formula:

$$F_{n}^{RR}(t)=1-\left(1-H_{n}^{E}(t)\right)^{R_n(t)}=\begin{cases}0, & t<Z_{(1)},\\ 1-\left(\frac{n-k}{n}\right)^{R_n(t)}, & Z_{(k)}\le t<Z_{(k+1)},\quad 1\le k\le n-1,\\ 1, & t\ge Z_{(n)},\end{cases} \tag{8}$$

where fraction $R_n(t)=\Lambda_{(1)n}(t)\cdot\left(\Lambda_n(t)\right)^{-1}$ is estimator of relative risk function $R(t)=\Lambda_{(1)}(t)\cdot\left(\Lambda(t)\right)^{-1}$. It is not difficult to see that $0\le R(t), R_n(t)\le 1$ for all $t\in E.$

Now in order to introduce our next presmoothed estimator we suppose that both d.f. $F(t)$ and $G(t)$ are absolutely continuous with densities $f(t)$ and $g(t)$ respectively. Then d.f. of minima $H(t)$ has density $h(t)=(1-G(t))f(t)+(1-F(t))g(t)$. Consider presmoothing problem of estimators (5), (6) and (8).

It is not difficult to see that regression function of indicator δ_k given value of minima $Z_k=t$ is equal to conditional probability

$$p(t)=\mathrm{P}(\delta_k=1\,|\,Z_k=t)=\mathrm{E}[\delta_k\,|\,Z_k=t],\ \ t\in R. \tag{9}$$

Then the integral hazard function $\Lambda_{(1)}(t)$ in (3) can be represented as

$$\Lambda_{(1)}(t)=\int_{-\infty}^{t}p(u)d\Lambda(u),\ t\in R, \tag{10}$$

or equivalently for densities

$$h_{(1)}(t)=p(t)h(t),\ t\in R, \tag{11}$$

$$R(t)=\frac{\Lambda_{(1)}(t)}{\Lambda(t)}=\frac{\int_{-\infty}^{t}p(u)d\Lambda(u)}{\Lambda(t)},\ t\in R. \tag{12}$$

For probability (9) we can use any estimator for regression function. But we prefer Nadaraya-Watson estimator (see, [15])

$$p_n(t)=\frac{\frac{1}{nh(n)}\sum_{k=1}^{n}\delta_k k\left(\frac{t-Z_k}{h(n)}\right)}{\frac{1}{nh(n)}\sum_{k=1}^{n}k\left(\frac{t-Z_k}{h(n)}\right)}=\frac{\sum_{k=1}^{n}\delta_k k\left(\frac{t-Z_k}{h(n)}\right)}{\sum_{k=1}^{n}k\left(\frac{t-Z_k}{h(n)}\right)},\ t\in R, \tag{13}$$

obtaining by formula $p(t)=\frac{h_{(1)}(t)}{h(t)}$ (see, (11)). Here kernel $\{k(t),\,t\in R\}$ is well known function of density $\{h(n),\,n\geq 1\}$ is ≪window width≫ sequence, tending to zero at $n\to\infty$. By formulas (9)–(13) function (12) can be estimated as

$$R_n^p(t)=\frac{\Lambda_{(1)n}^p(t)}{\Lambda_n(t)}=\frac{\int_{-\infty}^{t}p_n(u)d\Lambda_n(u)}{\Lambda_n(t)},\ t\in R, \tag{14}$$

where $\Lambda_n(t)$ is estimator in (4), $\Lambda_{(1)n}^p(t)$ is presmoothed estimator of $\Lambda_{(1)}(t)$. Now replacing power $R_n(t)$ by $R_n^p(t)$ in (8) we obtain a new presmoothed relative risk power estimator

$$F_n^{pR}(t)=1-\left(1-H_n^E(t)\right)^{R_n^p(t)}=\begin{cases}0, & t<Z_{(1)},\\ 1-\left(\frac{n-k}{n}\right)^{R_n^p(t)}, & Z_{(k)}\leq t<Z_{(k+1)},\ 1\leq k\leq n-1,\\ 1, & t\geq Z_{(n)},\end{cases} \tag{15}$$

Note that authors [15] have been investigated presmoothed variants of estimator (5) and (6).

$$F_n^{pAB}(t)=1-\exp\left(-\Lambda_{(1)n}^p(t)\right)=1-\exp\left(-\sum_{\{k:Z_{(k)}\leq t\}}\frac{p_n(Z_{(k)})}{n-k+1}\right), \tag{16}$$

$$F_n^{pKM}(t)=1-\prod_{\{k:Z_{(k)}\leq t\}}\left(1-\frac{p_n(Z_{(k)})}{n-k+1}\right),\ t\in R, \tag{17}$$

for estimator (16) and (17) we have asymptotic relation

$$F_n^{pKM}(t)-F(t)=(1-F(t))\left(\Lambda_{(1)n}^{p}(t)-\Lambda(t)\right)+\mathrm{O}_p\left(\frac{1}{n}\right),\ \ n\to\infty.$$

From here we get analogue of (7) for presmoothed estimators (16) and (17)

$$\Lambda_{(1)n}^{p}(t)\approx-\log\left(1-F_n^{pKM}(t)\right),\ t<T_{H_n^E}. \tag{18}$$

In particular, in case of independence of indicator δ_k from minima Z_k, from (9) we have $p(t)=p=E\delta_k$ for all $t\in R,$ consequently by characterization of simple Proportional Hazards Model (PHM) $1-F(t)=(1-H(t))^p$, $p=\frac{1}{1+\beta}$, $\beta>0$ is some number. Then estimator p by $p_n=\frac{1}{n}\sum_{k=1}^{n}\delta_k$ we obtain well known power estimator of the from

$$\hat{F}_n(t)=1-\left(1-H_n^E(t)\right)^{p_n}=\begin{cases}0, & t<Z_{(1)},\\ 1-\left(\frac{n-k}{n}\right)^{p_n}, & Z_{(k)}\le t<Z_{(k+1)},\ 1\le k\le n-1,\\ 1, & t\ge Z_{(n)},\end{cases} \tag{19}$$

for details, see [16–19]. At the end of section let us remind some theoretical that is asymptotic properties of estimator (15) from [20]. According to the theorem 1 in [20] authors approximate centered and normed estimator (15) by sum of independent and identically distributed random function, in theorem 2 authors get strong uniform consistency result for estimator (15) and in theorem 3 they proved asymptotically normality results for estimator (15).

3. Simulation study

In papers [12–14] authors investigated several asymptotic properties of estimators (1), (5) and (8) at $n\to\infty$ and showed that all three estimators are asymptotically equivalent in sense of convergence to same centered Gaussian process. Suppose that $G_n^{KM}(t)$, $G_n^{AB}(t)$, $G_n^{RR}(t)$ and $\hat{G}_n(t)$ are estimator of d.f. $G(t)$ defined by analogues of formulas (1), (5), (8) and (19) as follows:

$$G_n^{KM}(t)=\begin{cases}1-\prod_{\{k:Z_{(k)}\le t\}}\left(1-\frac{1-\delta_{(k)}}{n-k+1}\right), & t\le Z_{(n)},\\ 1, & t>Z_{(n)},\delta_{(n)}=0,\\ \mathit{undefined}, & t>Z_{(n)},\ \delta_{(n)}=1,\end{cases} \tag{20}$$

$$G_n^{AB}(t)=1-\exp\left(-\Lambda_{(0)n}(t)\right)=1-\exp\left\{-\frac{1}{n}\sum_{k=1}^{n}\frac{I(Z_k\le t,\delta_k=0)}{1-H_n^E(Z_k)+\frac{1}{n}}\right\}, \tag{21}$$

$$G_n^{RR}(t)=1-\left(1-H_n^E(t)\right)^{1-R_n(t)}=\begin{cases}0,\ t<Z_{(1)},\\ 1-\left(\frac{n-k}{n}\right)^{1-R_n(t)}, Z_{(k)}\le t<Z_{(k+1)}, 1\le k\le n-1,\\ 1,\ \ t\ge Z_{(n)},\end{cases} \tag{22}$$

and

$$\hat{G}_n(t)=1-\left(1-H_n^E(t)\right)^{1-p_n}=\begin{cases}0,\ t<Z_{(1)},\\ 1-\left(\frac{n-k}{n}\right)^{1-p_n}, Z_{(k)}\le t<Z_{(k+1)}, 1\le k\le n-1,\\ 1,\ \ t\ge Z_{(n)}.\end{cases} \tag{23}$$

In considered right random censorship model we have for d.f. of minima Z_i that from (2)

$$1-H(t)=\big(1-F(t)\big)\big(1-G(t)\big),\ t\in R. \qquad (24)$$

Which is easy fulfilled for two power estimator with its analogues: (8) with (22) and (19) with (23):

$$1-H_n^E(t)=\big(1-F_n^{RR}(t)\big)\big(1-G_n^{RR}(t)\big),\ t\in R,$$

$$1-H_n^E(t)=\big(1-\hat{F}_n(t)\big)\big(1-\hat{G}_n(t)\big),\ t\in R.$$

But for other two estimators (1) with (20) and (5) with (21) we have relations:

1) $\big(1-F_n^{KM}(t)\big)\big(1-G_n^{KM}(t)\big)\neq 1-H_n^E(t)$ for all $t\in R$, so for $t>Z_{(n)}$ both estimators $F_n^{KM}(t)$ and $G_n^{KM}(t)$ are not defined;

2) $\big(1-F_n^{AB}(t)\big)\big(1-G_n^{AB}(t)\big)=\exp\big(-\Lambda_{(1)n}(t)\big)\neq 1-H_n^E(t)$ and $\max\{F_n^{AB}(t),G_n^{AB}(t)\}<1.$

Hence, only power type estimator (8) and (19) are identifiable with considered right censorship model defined by equality (24).

Now we consider presmoothed estimators (16), (17) and (15) with corresponding variants of estimators of $G(t)$. Here we observe analogue satiation:

3) $\big(1-F_n^{pKM}(t)\big)\big(1-G_n^{pKM}(t)\big)\neq 1-H_n^E(t)$ for all $t\in R$, where

$$G_n^{pKM}(t)=1-\prod_{\{k:Z_{(k)}\le t\}}\left(1-\frac{1-\delta_{(k)}}{n-k+1}\right),\ t\in R.$$

4) $\big(1-F_n^{pAB}(t)\big)\big(1-G_n^{pAB}(t)\big)=\exp\big(-\Lambda_n^p(t)\big)\neq 1-H_n^E(t),\ t\in R,$ where $\Lambda_n^p(t)=\Lambda_{(0)n}^p(t)+\Lambda_{(1)n}^p(t).$

5) $\big(1-F_n^{pR}(t)\big)\big(1-G_n^{pR}(t)\big)=1-H_n^E(t)$ for all $t\in\mathbb{R}.$

For demonstrating properties of estimators, at first, we consider following ordered natural censored sample $\{(Z_{(k)},\delta_{(k)}),\ k=1,...,97\}$ of size *n* = 97 from paper [21] (well known Chenning Hause data) (table 1):

Table 1

Ordered natural censored sample $\{(Z_{(k)},\delta_{(k)}),\ k=1,...,97\}$ of size n = 97

k	$(Z_{(k)},\delta_{(k)})$	k	$(Z_{(k)},\delta_{(k)})$	k	$(Z_{(k)},\delta_{(k)})$	k	$(Z_{(k)},\delta_{(k)})$	k	$(Z_{(k)},\delta_{(k)})$
9	(777,1)	21	(940,0)	41	(977,0)	61	(1018,0)	81	(1059,1)
2	(781,0)	22	(942.5,0)	42	(983,1)	62	(1022,1)	82	(1060,1)
3	(843,0)	23	(943,0)	43	(984;0)	63	(1023,0)	83	(1060,0)
4	(866,0)	24	(945,1)	44	(985,1)	64	(1025,1)	84	(1064,0)
5	(869,1)	25	(945,0)	45	(989,1)	65	(1027,0)	85	(1070,0)
6	(872,1)	26	(948,1)	46	(992.5,1)	66	(1029,1)	86	(1073,0)
7	(876,1)	27	(951,0)	47	(993,1)	67	(1031,1)	87	(1080,1)
8	(893,1)	28	(953,0)	48	(996,1)	68	(1031,0)	88	(1085,1)
9	(894,1)	29	(956,0)	49	(998,1)	69	(1031.5,0)	89	(1093,0)
10	(895,0)	30	(957,1)	50	(1001,0)	70	(1033,1)	90	(1093.5,1)
11	(898,1)	31	(957,0)	51	(1002,0)	71	(1036,1)	91	(1094,1)
12	(906,0)	32	(959,0)	52	(1005,0)	72	(1043,1)	92	(1106,0)
13	(907,1)	33	(960,0)	53	(1006,0)	73	(1043,0)	93	(1107,0)
14	(909,1)	34	(966,1)	54	(1009,1)	74	(1044,1)	94	(1118,0)
15	(911,1)	35	(966,0)	55	(1011.5,1)	75	(1044,0)	95	(1128,1)
16	(911,0)	36	(969,1)	56	(1012,1)	76	(1045,0)	96	(1139,1)
17	(914,0)	37	(970,0)	57	(1012,0)	77	(1047,0)	97	(1153,0)
18	(927,1)	38	(971,1)	58	(1013,0)	78	(1053,1)		
19	(932,1)	39	(972,0)	59	(1015,0)	79	(1055,1)		
20	(936,0)	40	(973,0)	60	(1016,0)	80	(1058,0)		

These data in months, where “1”-corresponds to uncensoring and otherwise “0” to censoring. In the next figures 1–4 we demonstrate estimators $1-F_n^{AB}(t)$, $1-F_n^{KM}(t)$, $1-F_n^{RR}(t)$ and $1-F_n^{pR}(t)$ separately and in picture 5 all estimators together.

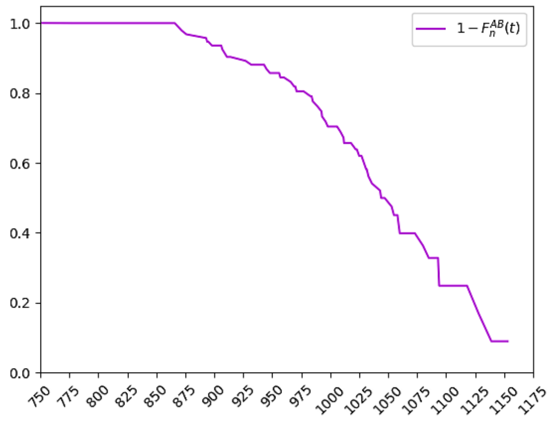

Fig. 1. Estimator $1-F_n^{AB}(t)$

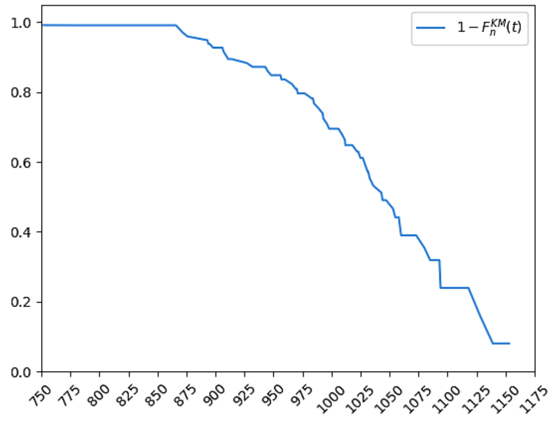

Fig. 2. Estimator $1-F_n^{KM}(t)$

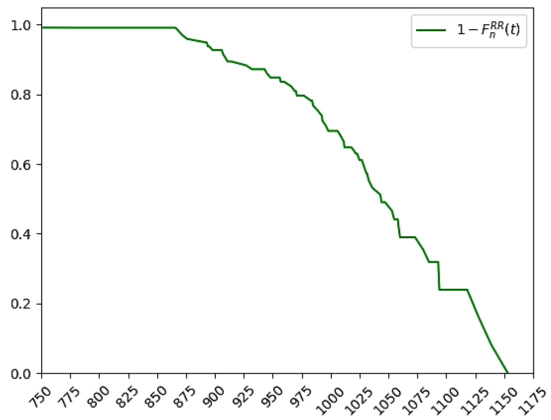

Fig. 3. Estimator $1-F_n^{RR}(t)$

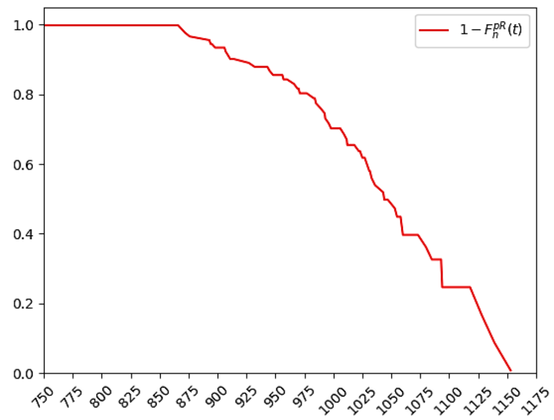

Fig. 4. Estimator $1-F_n^{pR}(t)$

In figures 3, 4 one can see that only power estimators $F_n^{RR}(t)$ and $F_n^{pR}(t)$ well defined in the whole line.

Now we used Gillespie-Fisher’s result about 95%-confidence bands of estimators $1-F_n^{AB}(t)$, $1-F_n^{KM}(t)$, $1-F_n^{RR}(t)$ and $1-F_n^{pR}(t)$ considering its asymptotically equivalence. Propose band in the form (see, [22]):

$$B_n^*(t,\eta_1,\eta_2)=\left[B_n^{*-}(t,\eta_1,\eta_2);B_n^{*+}(t,\eta_1,\eta_2)\right],$$

where

$$B_n^{*-}(t,\eta_1,\eta_2)=F_n^*(t)-n^{-\frac{1}{2}}\cdot\left(1-F_n^*(t)\right)\cdot\left(\eta_1\cdot d_n^{\frac{1}{2}}(M)+\eta_2\cdot d_n(t)\cdot d_n^{-\frac{1}{2}}(M)\right),$$

$$B_n^{*+}(t,\eta_1,\eta_2)=\frac{F_n^*(t)+n^{-\frac{1}{2}}\cdot\left(\eta_1\cdot d_n^{\frac{1}{2}}(M)+\eta_2\cdot d_n(t)\cdot d_n^{-\frac{1}{2}}(M)\right)}{1+n^{-\frac{1}{2}}\cdot\left(\eta_1\cdot d_n^{\frac{1}{2}}(M)+\eta_2\cdot d_n(t)d_n^{-\frac{1}{2}}(M)\right)},$$

$$M=1128;\ \eta_1=1;\ \eta_2=1{,}37\ \text{and}\ d_n(t)=\int_{-\infty}^{t}\frac{dH_{(1)n}^E(u)}{\left(1-H_n^E(u)+1/n\right)^2}.$$

Here $F_n^*(t)$ one of there estimators $F_n^{AB}(t)$, $F_n^{KM}(t)$, $F_n^{RR}(t)$ and $F_n^{pR}(t)$. These confidence bands we demonstrate in figures 5–8.

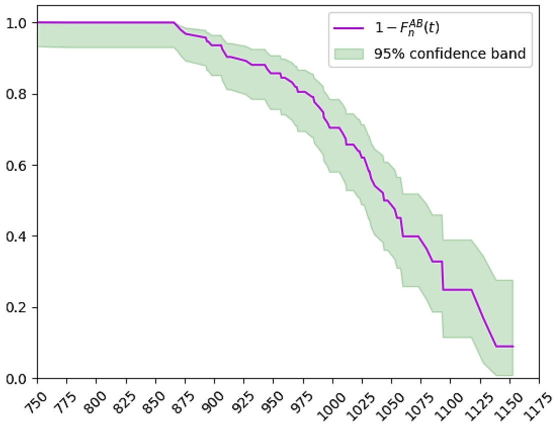

Fig. 5. Estimator $1-F_n^{AB}(t)$

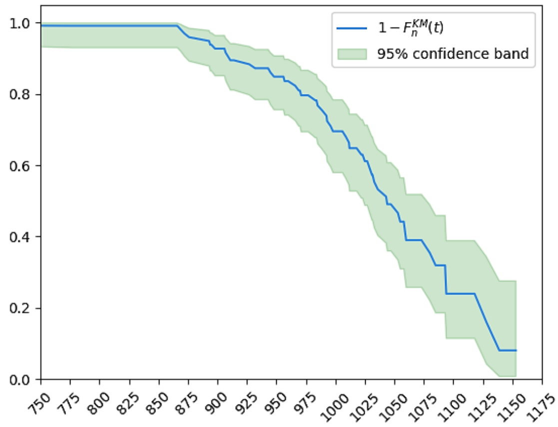

Fig. 6. Estimator $1-F_n^{KM}(t)$

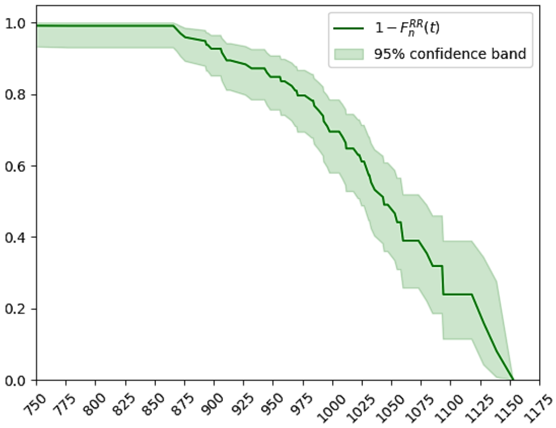

Fig. 7. Estimator $1-F_n^{RR}(t)$

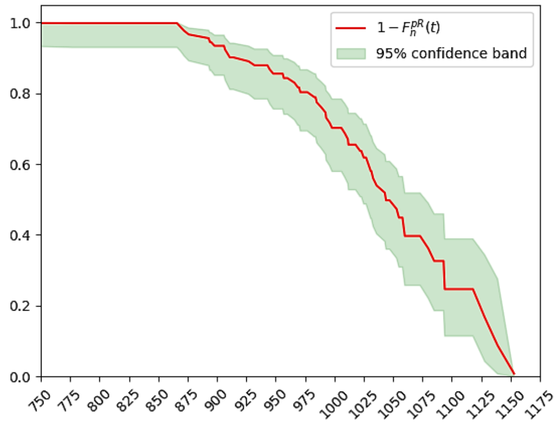

Fig. 8. Estimator $1-F_n^{pR}(t)$

Now we consider statistics $R_n(t), F_n^{RR}(t)$ and presmoothed estimators $R_n^p(t)$, $F_n^{pKM}(t)$ and $F_n^{pR}(t)$. Artificial data are modelled by exponential distribution $F(t)=1-e^{-t}, t\geq 0$ for several sample sizes *n*. In order to investigate estimator $F_n^{KM}(t)$ we choose *n* = 500 with censoring degree 30% (figure 9; table 2). In order to investigate estimator $R_n(t)$ we choose *n* = 500 with censoring degree 30% (figure 10). In the figure 11 we demonstrate estimator $F_n^{RR}(t)$ for artificial data of size *n* = 500 degree of censoring equal 30% (table 3).

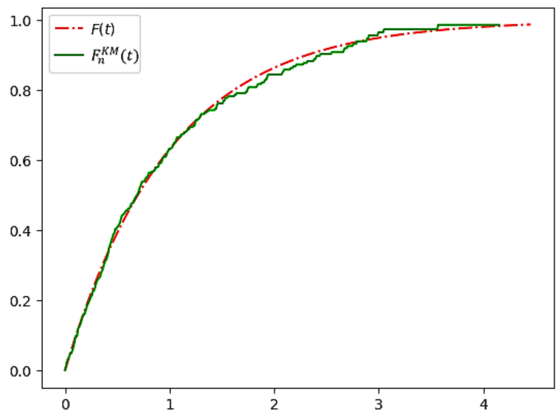

Fig. 9. Estimator $F_n^{KM}(t)$

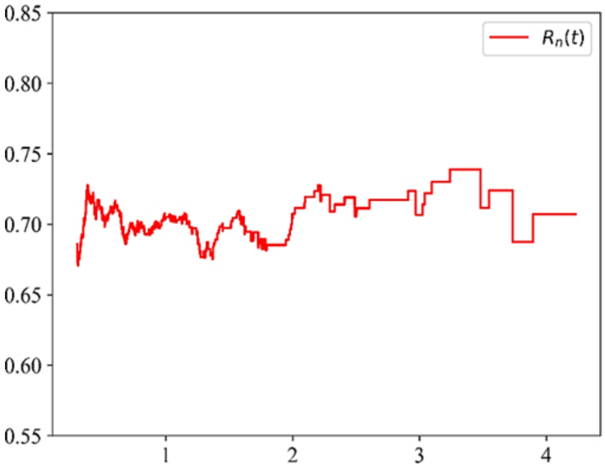

Fig. 10. Estimator $R_n(t)$

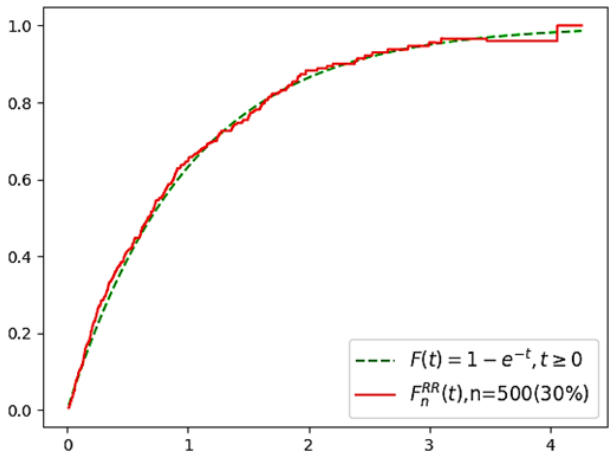

Fig. 11. Estimator $F_n^{RR}(t)$

In figure 12 we demonstrate estimator $R_n^p(t)$ for several smoothing parameter and in figure 13 estimator $F_n^{pKM}(t)$.

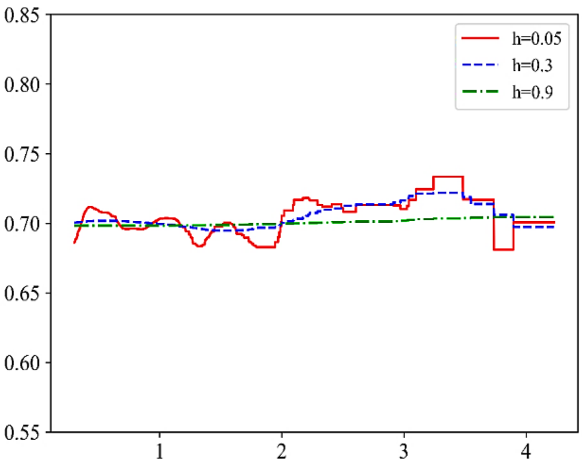

Fig. 12. Estimator $R_n^p(t)$

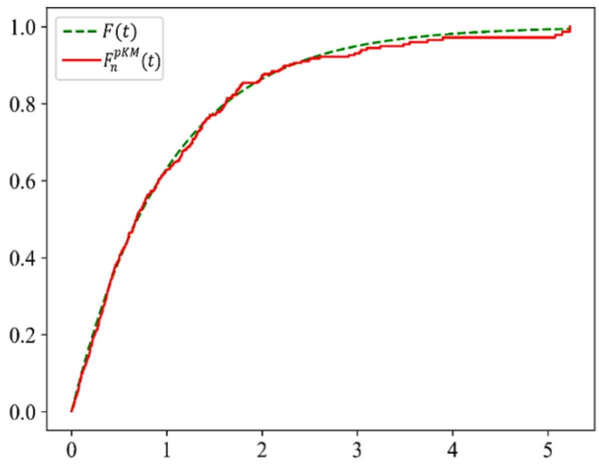

Fig. 13. Estimator $F_n^{pKM}(t)$

In figure 14 we demonstrate presmoothed relative risk power estimator $F_n^{pR}(t)$ for artificial data size n = 500 with degree of censoring 30%, (table 4).

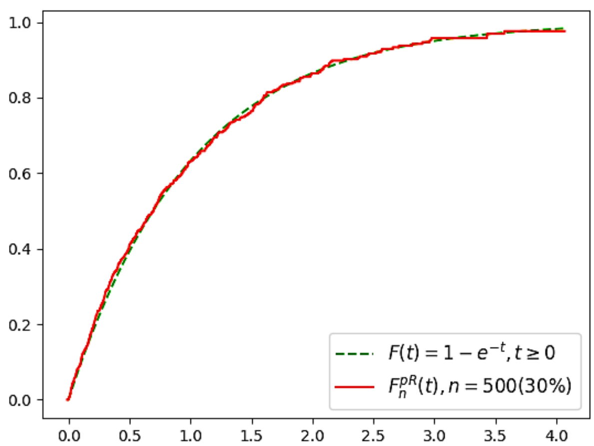

Fig. 14. Estimator $F_n^{pR}(t)$

Table 2

$F_n^{KM}(t)$ **deviation of the price from the theoretical distribution**

censoring \ n	n = 100	n = 500	n = 1 000	n = 2 000
10 %	0,08703	0,03948	0,02793	0,01977
20 %	0,09014	0,04079	0,02889	0,02041
30 %	0,09482	0,04317	0,03052	0,02166
40 %	0,10512	0,04846	0,03456	0,02453
50 %	0,12387	0,06106	0,04431	0,03233
60 %	0,15258	0,08502	0,06528	0,05020
70 %	0,23185	0,14288	0,11633	0,09412
80 %	0,36522	0,26567	0,23120	0,20212

Table 3

$F_n^{RR}(t)$ **deviation of the price from the theoretical distribution**

censoring \ n	n = 100	n = 500	n = 1 000	n = 2 000
10 %	0,08576	0,03891	0,02778	0,01946
20 %	0,08620	0,03941	0,02832	0,01977
30 %	0,08721	0,04036	0,03006	0,02124
40 %	0,08845	0,04257	0,03195	0,02207
50 %	0,10239	0,05093	0,03661	0,02596
60 %	0,14121	0,07887	0,05906	0,03708
70 %	0,18716	0,12348	0,10134	0,07319
80 %	0,21807	0,16330	0,14907	0,12187

Table 4

$F_n^{pR}(t)$ **deviation of the price from the theoretical distribution**

censoring \ n	n = 100	n = 500	n = 1 000	n = 2 000
10 %	0,08561	0,03812	0,02767	0,01932
20 %	0,08617	0,03871	0,02830	0,01956
30 %	0,08712	0,03945	0,02902	0,02017
40 %	0,08785	0,04105	0,03112	0,02197
50 %	0,09034	0,04789	0,03538	0,02487
60 %	0,12056	0,07012	0,05423	0,02718
70 %	0,17120	0,10256	0,09558	0,07018
80 %	0,19823	0,14993	0,13554	0,11956

At the end of section we consider other extension of power estimator (8). In the model of right random censorship we consider other sample $\tilde{\mathbb{C}}^{(n)} = \{(Z_k, \delta_k, C_k),\ k = 1,...,n\}$ is available for observation, where $Z_k = \min(X_k, Y_k)$, $\delta_k = I(X_k \le Y_k)$. Here r.v.'s X_k and Y_k are independent given covariate $C_k = x$ with values in the interval $(0,1]$. For example, this corresponds to experiment, where for tested object (individual or physical system) *x* denote dose of drug in medicine or temperature, press or voltage in industry. Now the problem consists estimating of conditional d.f.

$$F_x(t) = P(X_k \le t \mid C_k = x)$$

under nuisance d.f. $G_x(t) = P(Y_k \le t \mid C_k = x)$ and

$$1 - H_x(t) = P(Z_k > t \mid C_k = x) = (1 - F_x(t))(1 - G_x(t)),\ t \in R.$$

Consider relative risk power estimator (8) for $F_x(t)$ in presence of covariate in the form

$$F_{xh}(t)=1-\left(1-H_{xh}^{E}(t)\right)^{R_{xh}(t)}=\begin{cases}0, & t<Z_{(1)},\\ 1-\left(\dfrac{n-k}{n}\right)^{R_{xh}(t)}, & Z_{(k)}\le t<Z_{(k+1)},\ 1\le k\le n-1,\\ 1, & t\ge Z_{(n)},\end{cases} \tag{25}$$

where

$$H_{xh}^{E}(t)=\sum_{k=1}^{n}\omega_{nk}(x,h)I(Z_k\le t)=H_{(0)xh}^{E}(t)+H_{(1)xh}^{E}(t),\ \ H_{(m)xh}^{E}(t)=\sum_{k=1}^{n}\omega_{nk}(x,h)I(Z_k\le t,\ \delta_k=m),\ m=0{,}1.$$

are Stoun's type estimators with Gasser-Miiller's wights:

$$\omega_{nk}(x,h)=\frac{1}{C_n(x;h)}\int_{x_{k-1}}^{x_k}\frac{1}{h}k\left(\frac{x-u}{h}\right)du,\ k=1,\dots,n,$$

and

$$C_n(x;h)=\int_{0}^{x_n}\frac{1}{h}k\left(\frac{x-u}{h}\right)du.$$

Usually for large enough n, $C_n(x;h)\approx 1$. For our purposes we put $C_n=1$. Here $R_{xh}(t)=\Lambda_{(1)xh}(t)\cdot\left(\Lambda_{xh}(t)\right)^{-1}$, $\Lambda_{xh}(t)=\Lambda_{(0)xh}(t)+\Lambda_{(1)xh}(t)$, and $\Lambda_{(m)xh}(t)=\int_{-\infty}^{t}\frac{dH_{(m)xh}^{E}(u)}{1-H_{xh}^{E}(u)}$, $m=0{,}1.$

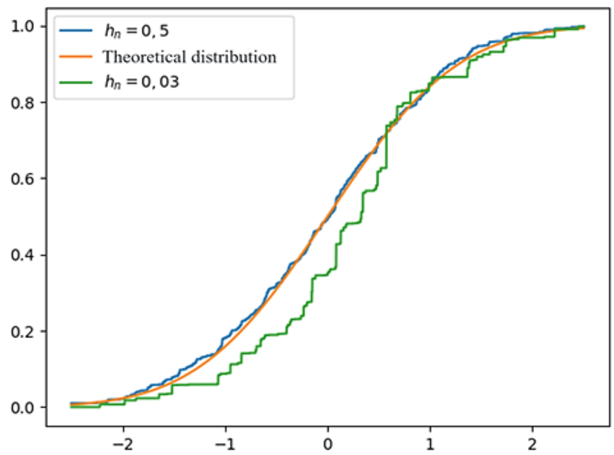

Fig. 15. Estimator

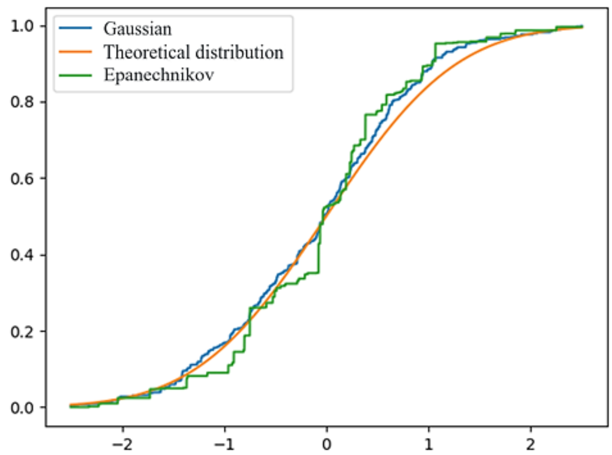

Fig. 16. Estimator

Table 5

Deviation of the estimate from the theoretical distribution

$h(n)$ \ u	0,1	0,2	0,3	0,5	0,7	0,9	1
Gaussian kernel function							
0,1	0,07232	0,06524	0,06251	0,06360	0,06343	0,07257	0,08896
0,3	0,04902	0,04486	0,04345	0,04075	0,04268	0,04891	0,05133
0,5	0,04095	0,03994	0,03871	0,03813	0,03927	0,04075	0,04200
0,7	0,03933	0,03849	0,03831	0,03781	0,03804	0,03835	0,03913
0,9	**0,03826**	**0,03770**	**0,03786**	**0,03704**	**0,03707**	**0,03806**	**0,03832**
Epanechnikov kernel function							
0,1	0,09185	0,09226	0,09216	0,09190	0,09182	0,09273	0,12820
0,3	0,06519	0,05647	0,05496	0,05360	0,05267	0,06499	0,07467
0,5	0,05275	0,04857	0,04487	0,04106	0,04551	0,05322	0,05875
0,7	0,04627	0,04323	0,04032	0,03832	0,04076	0,04547	0,04897
0,9	**0,04100**	**0,03923**	**0,03844**	**0,03819**	**0,03865**	**0,04137**	**0,04353**

In proportional hazards model with covariate we consider conditional d.f. as $F_x(t) = 1 - \exp\{-(1-2\log x)t\}, t \ge 0, 0 < x \le 1$, and let $\varepsilon_x(h) = \sup_{t\ge 0} |F_x(t) - F_{xh}(t)|$.

For generating of *n* = 1 000 artificial data with degree of censoring 30% for Gaussian and Epanechnikov kernels, we demonstrate estimator (25) for several smoothing parameters h (fig. 15, 16; table 5). The nearest to zero values of distance $\varepsilon_x(h)$, we see for Gaussian kernel.

Conclusion

In this paper authors are comparatively considering in right random censorship model three types of survival function estimators: presmoothed power, exponential (Altshuler-Breslow) and product-types (Kaplan-Meier). In simulation study they used natural sample of size n=97 and artificial data for several sample size. It is well know that in right random censoring model survival function of observed minima is the product of survival functions of censored and censoring random variables. Authors are demonstrating that this property is easy fulfilled for two presmoothed power type estimators. But for other two estimators: of exponential and product-limit it is not full filled, because of their uncertainty in whole line. Hence, only presmoothed power-type estimators are identifiably with considered right censorship model.

References

1. Abdushukurov, A.A., Bozorov, S.B., & Mansurov, D.R. (2024) Optimal kernel density estimation by statistical modeling with random censoring of observations. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravleniye, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science.* 67. pp. 49–59. DOI: 10.17223/19988605/67/6
2. Akritas, M.G. (2000) The central limit theorem under censoring. *Bernulli.* 6(6). pp. 1109–1120.
3. Breslow, N. (1972). Discussion of professor Cox's paper. *Journal of Royal Statistics Society*. 34(A). pp. 216–217.
4. Breslow, N. & Crowley, J. (1974) A large sample study of the life table and product-limit estimates under random censorship. *Annual Statistics*. 2. pp. 437–453.
5. Burke, M.D., Csörgő, S. & Horváth, L. (1981) Strong approximations some biometric estimates under random censorship. *Z. Wahrscheinlich. Verw. Gebiete*. 56. pp. 87–112.
6. Burke, M.D., Csörgő, S. & Horváth, L. (1988) A correction toand improvement of "Strong approximations of some biometric estimates under random censorship." *Probability Theory and Related* Fields. 79. pp. 51–57.
7. Csörgő, S. (1996). Universial Gaussian approximations under random censorship. *Annual Statistics.* 24(6). pp. 2744–2778.
8. Csörgő, S., & Horváth, L. (1982). On random censorship from the right. *Acta Sci. Math*. 44. pp. 23–34.
9. Kaplan, E.L. & Meier, P. (1958) Nonparametric estimation from incomplete observations. *Journal of American Statistics Association*. 53. pp. 457–481.
10. Abdushukurov, A.A., & Mansurov, D.R. (2023) Estimating the distribution function using parametric methods in informative model of random censorship from both siodes. In: Dudin, A., Nazarov, A. & Moiseev, A. (eds) *Communication in Computer and Information Sciences (CCIS).* Springer. pp. 2025–2037.
11. Altshuler, B. (1970) Theory of measurements for competing risks in animal expriments. *Math. Bioscieces*. 6. pp. 1–11.
12. Abdushukurov, A.A. (1998) Nonparametric estimation of the distribution function based on relative risk function. *Communication in Statististics – Theory & Methods.* 27(8). pp. 1991–2012.
13. Abdushukurov, A.A. (1999) On nonparametric estimation of probability indices by censored samples. *Theory of Probability and its Applications*. 43(1). pp. 3–11.
14. Abdushukurov, A.A. (2011) *Estimation of Unknown Distributions by Incomplete Observations.* LAMBERT Academic Publishing (LAP).
15. Jacome, M.A. & Cao, R. (2004) Presmoothed kernel density estimator for censored data. *Journal of Nonparam. Statist*. 16. pp. 289–309.
16. Abdushukurov, A.A. (1987) Nonparametric estimation in Proportional Hazards Model. *Technical Repert* N3448-B -87.
17. Abdushukurov, A.A. (1987) Estimating of probability density and intensity functions of the Koziol–Green model of random censoring. *Sankhya Indiren Journal of Statistics*. 48(1). pp. 150–168.
18. Cheng, P.E. & Ling, G.D. (1987) Maximum likelihood estimation of a survival function under Koziol-Green proportional hazards model. Statist. *Probability Letters*. 5. pp. 75–80.
19. Csörgő, S. (1989) Testing for the proportional hazards model of random censorship. *Proceedings IV Prague Symp. Asump. Statist*. Charles Univ. Press. Prague. pp. 41–53.
20. Abdushukurov, A.A., Nurmukhamedova, N.S. & Bozorov, S.B. (2021) Nonparametric estimation of the distribution function under right random censoring based on presmoothed relative risk function. *Lobachevskii Journal of Mathematics*. 42(2). pp. 257–268.
21. Efron, B. (1981) Censored data and the Bootstrap. *The Journal of the Acoustical Society of America*. 76(374). pp. 312–319.
22. Csörgő, S. & Horváth, L. (1986) Confidence bands from censored samples. *Canadian Journal of Statistics*. 14(2). pp. 131–144.

Information about the authors:
Abdushukurov Abdurakhim A. (Doctor of Physical and Mathematical Sciences, Professor, Moscow State University, Branch in Tashkent, Tashkent, Uzbekistan). E-mail: a_abdushukurov@rambler.ru
Bozorov Sukhrob B. (PhD Student, Gulistan State University, Gulistan, Uzbekistan.). E-mail: suxrobbek_8912@mail.ru

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Информация об авторах:
Абдушукуров Абдурахим Ахмедович – доктор физико-математических наук, профессор Московского государственного университета, филиала в Ташкенте (Ташкент, Узбекистан). E-mail: a_abdushukurov@rambler.ru
Бозоров Сухроб Баходирович – докторант Гулистанского государственного университета (Гулистан, Узбекистан). E-mail: suxrobbek_8912@mail.ru

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Received 22.12.2024; accepted for publication 02.06.2025

Поступила в редакцию 22.12.2024; принята к публикации 02.06.2025