DOI: 10.17516/1997-1397-2022-15-3-292-307
УДК 519.2

On Approximation of Empirical Kac Processes under General Random Censorship Model

Abdurahim A. Abdushukurov*
Moscow State University named after M. V. Lomonosov, Tashkent Branch
Tashkent, Uzbekistan

Gulnoz S. Saifulloeva†
Navoi State Pedagogical Institute
Navoi, Uzbekistan

Received 07.11.2021, received in revised form 10.12.2021, accepted 20.02.2022

Abstract. A general random censorship model is considered in the paper. Approximation results are proved for empirical Kac processes. This model includes important special cases such as random censorship on the right and competing risks model. The obtained results use strong approximation theory and optimal approximation rates are built. Cumulative hazard processes are also investigated in a similar manner in the general setting. These results are also used for estimating of characteristic functions in random censorship model on the right.

Keywords: censored data, competing risks, empirical estimates, Kac estimate, strong approximation, Gaussian process, characteristic function.

Citation: A.A. Abdushukurov, G.S. Saifulloeva, On Approximation of Empirical Kac Processes under General Random Censorship Model, J. Sib. Fed. Univ. Math. Phys., 2022, 15(3), 292–307. DOI: 10.17516/1997-1397-2022-15-3-292-307.

1. Introduction and preliminaries

Following of ([3–5]) we define a general random censorship model in the following way: Let Z be a real random variable (r.v.) with distribution function (d.f.) $H(x) = P(Z \leqslant x)$, $x \in \mathbb{R}$. Let us assume that $A^{(1)}, \dots, A^{(k)}$ are pairwise disjoint random events for a fixed integer $k \geqslant 1$. Let us define the subdistribution functions $H(x;i) = P(Z \leqslant x, A^{(i)}), i \in \Im = \{1, \dots, k\}$. Suppose that when observing Z we are interested in the joint behaviour of the pairs $(Z, A^{(i)})$, $i \in \Im$. Let $\{(Z_j, A_j^{(1)}, \dots, A_j^{(k)}),\ j \geqslant 1\}$ be a sequence of independent replicas of $(Z, A^{(1)}, \dots, A^{(k)})$ defined on some probability space $\{\Omega, A, P\}$. We assume throughout that functions $H(x), H(x;1), \dots, H(x;k)$ are continuous. Let us denote the ordinary empirical d.f. of $Z_1, \dots, Z_n$ by $H_n(x)$ and introduce the empirical sub d.f. $H_n(x;i)$, $i \in \Im$

$$H_n(x;i) = \frac{1}{n}\sum_{j=1}^{n} \delta_j^{(i)} I(Z_j \leqslant x),\ (x;i) \in \overline{\mathbb{R}} \times \Im,$$

where $\overline{\mathbb{R}} = [-\infty;\infty]$, $\delta_j^{(i)} = I(A_j^{(i)})$ is the indicator of event $A_j^{(i)}$ and

*a-abdushukurov@rambler.ru
†sayfulloyevagulnoz@gmail.com

$$H_n(x;1)+\cdots+H_n(x;k)=\frac{1}{n}\sum_{j=1}^{n}I(Z_j\leqslant x)=H_n(x),\ x\in\overline{\mathbb{R}},$$

is the ordinary empirical d.f.. Properties of many biometric estimates depend on the limit behaviour of proposed empirical statistics. The following results are straightforward consequences of the Dvoretzky–Kiefer–Wolfowitz exponential inequality with constant $D=2$ [8, 12] :

For all $n=1,2,\dots$ and $\varepsilon>0$

$$P\left(\sup_{|x|<\infty}\bigl|H_n(x)-H(x)\bigr|>\left(\frac{(1+\varepsilon)}{2}\cdot\frac{\log n}{n}\right)^{1/2}\right)\leqslant 2n^{-(1+\varepsilon)},\tag{1}$$

and

$$P\left(\sup_{|x|<\infty}\bigl|H_n(x;i)-H(x;i)\bigr|>2\Bigl(\frac{(1+\varepsilon)}{2}\frac{\log n}{n}\Bigr)^{1/2}\right)\leqslant 4n^{-(1+\varepsilon)}.\tag{2}$$

Vector-valued empirical process $\Bigl\{a_n(t)=\Bigl(a_n^{(0)}(t_0),a_n^{(1)}(t_1),\dots,a_n^{(k)}(t_k)\Bigr),\ t=(t_0,\dots,t_k)\in\overline{\mathbb{R}}^{k+1}\Bigr\}$ plays a decisive role, where $a_n^{(0)}(x)=\sqrt{n}(H_n(x)-H(x))$, $a_n^{(i)}(x)=\sqrt{n}(H_n(x;i)-H(x;i))$, $i\in\Im$. The following Burke–Csörgő–Horváth theorem [3, 4] is an extended analogue of Komlós–Major–Tusnády's result [9–11]. ** **Theorem A**([3, 4]). *If the underlying probability space $\{\Omega,\mathcal{A},P\}$ is rich enough then one can define $k+1$ sequences of Gaussian processes $B_n^{(0)}(x),B_n^{(1)}(x),\dots,B_n^{(k)}(x)$ such that for $a_n(t)$ and $B_n(t)=(B_n^{(0)}(x_0),B_n^{(1)}(x_1),\dots,B_n^{(k)}(x_k))$, $t=(t_0,\dots,t_k)$ we have*

$$P\left\{\sup_{t\in\overline{\mathbb{R}}^{k+1}}\bigl\|a_n(t)-B_n(t)\bigr\|^{(k+1)}>n^{-\frac{1}{2}}(M(\log n)+z)\right\}\leqslant K\exp\left(-\lambda z\right),\tag{3}$$

for all real z, where $M=(2k+1)A_1$, $K=(2k+1)A_2$ and $\lambda=A_3/(2k+1)$ with A_1,A_2 and A_3 are absolute constants. Moreover, B_n is $(k+1)$-dimensional vector-valued Gaussian process that has the same covariance structure as the vector $a_n(t)$, namely, $EB_n^{(i)}(x)=0$, $(x,i)\in\overline{\mathbb{R}}\times\overline{\Im}=\Im\cup\{0\}$. We have for any $i,j\in\Im$, $i\neq j$, $x,y\in\overline{\mathbb{R}}$ that

$$\begin{aligned}
&EB_n^{(0)}(x)B_n^{(0)}(y)=\min\bigl\{H(x),H(y)\bigr\}-H(x)\cdot H(y),\\
&EB_n^{(i)}(x)B_n^{(i)}(y)=\min\bigl\{H(x;i),H(y;i)\bigr\}-H(x;i)\cdot H(y;i),\\
&EB_n^{(i)}(x)B_n^{(j)}(y)=-H(x;i)\cdot H(y;j),\\
&EB_n^{(0)}(x)B_n^{(i)}(y)=\min\bigl\{H(x;i),H(y;j)\bigr\}-H(x)\cdot H(y;i).
\end{aligned}\tag{4}$$

If we set $z=\left(\dfrac{(1+\varepsilon)}{\lambda}\log n\right)$ in (3) then

$$P\left\{\sup_{t\in\overline{\mathbb{R}}^{k+1}}\bigl\|a_n(t)-B_n(t)\bigr\|^{(k+1)}>Cn^{-\frac{1}{2}}\log n\right\}\leqslant Kn^{-(1+\varepsilon)},$$

where $C=(2k+1)\left(A_1+\dfrac{(1+\varepsilon)}{A_3}\right)$. Then

$$\Bigl\|a_n(t)-B_n(t)\Bigr\|^{(k+1)}\overset{a.s.}{=}O\Bigl(n^{-\frac{1}{2}}\log n\Bigr).$$

Let us note that in proving Theorem A (Theorem 3.1 in [4]) the sequence of two-parametrical Gaussian processes $\mathbb{Q}^{(0)}(x,n),\mathbb{Q}^{(2)}(x,n),\dots,\mathbb{Q}^{(k)}(x,n)$ was constructed such that for $a_n(t)$ and $\mathbb{Q}(t;n)=\big(\mathbb{Q}^{(0)}(x;n),\dots,\mathbb{Q}^{(k)}(x;n)\big),\quad t\in\overline{\mathbb{R}}^{k+1}$ the following approximation was used

$$\Big\|a_n(t)-n^{-\frac{1}{2}}\mathbb{Q}(t,n)\Big\|^{(k+1)}\overset{a.s.}{=}O\Big(n^{-\frac{1}{2}}\log^2 n\Big),$$

where $\mathbb{Q}(t,n)$ is the $(k+1)$ dimensional vector-valued Gaussian process and $\mathbb{Q}(t;n)\overset{D}{=}n^{\frac{1}{2}}a_n(t)$. Hence

$$E\mathbb{Q}^{(i)}(x;n)=0,\quad (x,i)\in\overline{\mathbb{R}}\times\overline{\Im}$$

and we have for any $i,j\in\Im,\ i\neq j,\ x,y\in\overline{\mathbb{R}}$ that

$$\begin{aligned}
&E\mathbb{Q}^{(0)}(x;n)\mathbb{Q}^{(0)}(y;m)=\min(n,m)\big\{\min\{H(x),H(y)\}-H(x)H(y)\big\},\\
&E\mathbb{Q}^{(0)}(x;n)\mathbb{Q}^{(i)}(y;m)=\min(n,m)\big\{\min\{H(x;i),H(y;i)\}-H(x)H(y;i)\big\},\\
&E\mathbb{Q}^{(i)}(x;n)\mathbb{Q}^{(i)}(y;m)=\min(n,m)\big\{\min\{H(x;i),H(y;i)\}-H(x;i)H(y;j)\big\},\\
&E\mathbb{Q}^{(i)}(x;n)\mathbb{Q}^{(j)}(y;m)=-\min(n,m)H(x;i)\cdot H(y;j).
\end{aligned}\tag{5}$$

Let us observe that $\{\mathbb{Q}^{(i)},i\in\overline{\Im}\}$ are Kiefer processes and they satisfy the distributional equality

$$\mathbb{Q}^{(i)}(x;n)\overset{D}{=}W^{(i)}(H(x;i);n)-H(x;i)W^{(i)}(1;n),\tag{6}$$

where $\big\{W^{(i)}(y;n),0\leqslant y\leqslant 1,n\geqslant 1,i\in\Im\big\}$ are two-parametric Wiener processes with $EW^{(i)}(y;n)=0$ and

$$EW^{(i)}(y;n)W^{(i)}(u;m)=\min(n,m)\min(y,u),\quad i\in\Im.$$

It is important to note that though Kiefer processes $\big\{\mathbb{Q}^{(i)},\ i\in\Im\big\}$ are dependent processes, corresponding Wiener processes are independent. Indeed, it follows from the proof of Theorem A that

$$\begin{aligned}
&\mathbb{Q}^{(1)}(x;n)\overset{D}{=}\widetilde{K}(H(x;1);n),\\
&\mathbb{Q}^{(2)}(x;n)\overset{D}{=}\widetilde{K}(H(x;2)-H(+\infty;1);n)-\widetilde{K}(H(+\infty;1);n),\\
&\dots\dots\\
&\mathbb{Q}^{(i)}(x;n)\overset{D}{=}\widetilde{K}(H(x;i)+H(+\infty;1)+\dots+H(+\infty;i-1);n)-\\
&\qquad-\widetilde{K}(H(+\infty;1)+\dots+H(+\infty;i-1);n),\quad i\in\Im,
\end{aligned}$$

where $H(+\infty;i)=\lim\limits_{x\uparrow+\infty}H(x;i),\ H(+\infty;1)+\dots+H(+\infty;k)=1.$

The Kiefer processes $\big\{\widetilde{K}(y;n),0\leqslant y\leqslant 1,n\geqslant 1\big\}$ are represented in terms of two-parametrical Wiener processes $\big\{W(y;n),0\leqslant y\leqslant 1,n\geqslant 1\big\}$ by distributional equality

$$\big\{\widetilde{K}(y;n),\quad 0\leqslant y\leqslant 1,n\geqslant 1\big\}\overset{D}{=}\big\{W(y;n)-yW(1;n),0\leqslant y\leqslant 1,n\geqslant 1\big\}.\tag{7}$$

Then, taking into account (6) and (7), the Wiener process $\{W^{(i)},i\in\Im\}$ also admits the following representations for all $(x;i)\in\overline{\mathbb{R}}\times\Im$

$$\begin{aligned}
&W^{(1)}(H(x;1);n)\overset{D}{=}W(H(x;1);n),\\
&W^{(2)}(H(x;2);n)\overset{D}{=}W(H(x;2)+H(+\infty;1);n)-W^{(1)}(H(+\infty;1);n),\dots,\\
&W^{(i)}(H(x;i);n)\overset{D}{=}W(H(x;i)+H(+\infty;i-1);n)-W(H(+\infty;1)+\dots+H(+\infty;i-1);n).
\end{aligned}$$

Now performing direct calculations of covariances of processes $\{W^{(i)},\ i\in\Im\}$, it is easy to show that these processes are independent.

2. Kac processes under general censoring

Following [9] we introduce the modified empirical d.f. of Kac by the following way. Along with sequence $\{Z_j, j \geqslant 1\}$ on a probability space $\{\Omega, \mathcal{A}, P\}$ consider also a sequence $\{\nu_n,\ n \geqslant 1\}$ of r.v.-s that has Poisson distribution with parameter $E\nu_n = n,\ n = 1, 2, \dots$. Let us assume throughout that two sequences $\{Z_j, j \geqslant 1\}$ and $\{\nu_n, n \geqslant 1\}$ are independent. The Kac empirical d.f. is

$$H_n^*(x) = \begin{cases} \dfrac{1}{n}\displaystyle\sum_{j=1}^{\nu_n} I(Z_j \leqslant x) & \text{if} \quad \nu_n \geqslant 1 \quad a.s., \\ 0 \quad \text{if} \quad \nu_n = 0 \quad a.s., \end{cases}$$

while the empirical sub-d.f. is

$$H_n^*(x;i) = \begin{cases} \dfrac{1}{n}\displaystyle\sum_{j=1}^{\nu_n} I\big(Z_j \leqslant x, A_j^{(i)}\big), \quad i \in \Im \quad if \quad \nu_n \geqslant 1 \quad a.s., \\ 0, \quad i \in \Im \quad if \quad \nu_n = 0 \quad a.s., \end{cases}.$$

with $H_n^*(x;1) + \dots + H_n^*(x;k) = H_n^*(x)$ for all $x \in \overline{\mathbb{R}}$. Here we suppose that sequence $\{\nu_n,\ n \geqslant 1\}$ is independent of random vectors $\big\{\big(Z_j, \delta_j^{(1)}, \dots, \delta_j^{(k)}\big), j \geqslant 1\big\}$, where $\delta_j^{(i)} = I(A_j^{(i)})$. Let us note that statistics $H_n^*(x;i)$ (and also $H_n^*(x)$) are unbiased estimators of $H(x;i),\ i \in \Im$ (and also of $H(x)$)

$$E(H_n^*(x;i)) = \frac{1}{n}E\left\{\sum_{m=1}^{\infty} E\left[\sum_{k=1}^{n} \delta_k^{(i)} \cdot I(Z_k \leqslant x)\right], \nu_n = m\right\} =$$
$$= \frac{1}{n}E\left\{\sum_{m=1}^{\infty} E\left[\sum_{k=1}^{n} \delta_k^{(i)} \cdot I(Z_k \leqslant x)/\nu_n = m\right] \cdot P(\nu_n = m)\right\} =$$
$$= \frac{1}{n}\sum_{m=1}^{\infty} H(x;i) m P(\nu_n = m) = \frac{1}{n}H(x;i)\sum_{m=1}^{\infty} m \cdot \frac{n^m e^{-n}}{m!} =$$
$$= H(x;i)e^{-n}\sum_{m=0}^{\infty}\frac{n^m}{m!} = H(x;i), \quad (x;i) \in \overline{\mathbb{R}} \times \Im.$$

Consequently,

$$E\big[H_n^*(x)\big] = \sum_{i=1}^{k} E\big[H_n^*(x;i)\big] = \sum_{i=1}^{k} H(x;i) = H(x), \quad x \in \overline{\mathbb{R}}.$$

Let us define the empirical Kac processes $a_n^{(i)*}(x) = \sqrt{n}\big(H_n^*(x;i) - H(x;i)\big),\ i \in \Im$ and $a_n^{(0)*}(x) = \sqrt{n}\big(H_n^*(x) - H(x)\big)$.

Theorem 1. *If the underlying probability space $\{\Omega, \mathcal{A}, P\}$ is rich enough then one can define $k+1$ sequences of Gaussian processes $W_n^{(0)}(x), W_n^{(1)}(x), \dots, W_n^{(k)}(x)$ such that for $a_n^*(t) = \big(a_n^{(0)*}(t_0), a_n^{(1)*}(t_1), \dots, a_n^{(k)*}(t_k)\big)$ and $W_n^*(t) = \big(W_n^{(0)}(t_0), W_n^{(1)}(t_1), \dots, W_n^{(k)}(t_k)\big)$, $t = (t_0, t_1, \dots, t_k)$ we have*

$$P\left\{\sup_{t \in \overline{\mathbb{R}}^{k+1}} \left\| a_n^*(t) - W_n^*(t) \right\|^{(k+1)} > C^* n^{-\frac{1}{2}} \log n\right\} \leqslant K^* n^{-r}, \tag{8}$$

where $r \geqslant 2$ is an arbitrary integer, $C^* = C^*(r)$ depends only on r, and K^* is an absolute constant. Moreover, $W_n^*(t)$ is $(k+1)$-dimensional vector-valued Gaussian process with expectation

$EW^{(i)}(x)=0,\ (x,i)\in\overline{\mathbb{R}}\times\overline{\Im}$. *We have for any* $i,j\in\Im,\ i\neq j,\ x,y\in\overline{\mathbb{R}}$ *that*

$$\begin{aligned} EW_n^{(0)}(x)W_n^{(0)}(y)&=\min\big\{H(x),H(y)\big\},\\ EW_n^{(i)}(x)W_n^{(j)}(y)&=\min\big\{H(x;i),H(y;j)\big\},\\ EW_n^{(i)}(x)W_n^{(0)}(y)&=\min\big\{H(x;i),H(y)\big\}.\end{aligned}\tag{9}$$

The basic relation between $a_n(t)$ and $a_n^*(t)$ is the following easily checked identity

$$a_n^*(x)=\sqrt{\frac{\nu_n}{n}}a_{\nu_n}^{(i)}(x)+H(x;i)\frac{(\nu_n-n)}{\sqrt{n}},\quad i\in\Im.\tag{10}$$

Hence, the approximating sequence have the form

$$W_n^{(i)}(x)=B_{\nu_n}^{(i)}(x)+H(x;i)\frac{W^*(n)}{\sqrt{n}},\quad i\in\Im,$$

where $B_{\nu_n}^{(i)}(x)$ is a Poisson indexed Brownian bridge type process of Theorem A and $\big\{W^{(*)}(x),\ x\geqslant 0\big\}$ is a Wiener process. It is easy to verify that $\big\{W_n^{(i)}(x),(x;i)\in\overline{\mathbb{R}}\times\overline{\Im}\big\}\overset{D}{=}\big\{W^*\big(H(x;i)\big),(x,i)\in\overline{\mathbb{R}}\times\overline{\Im}\big\}$. The proof of Theorem 1 is similar to the proof of Theorem 1 of Stute [6] and, it is omitted.

Since $\lim\limits_{x\uparrow+\infty}H_n^*(x)=H_n^*(+\infty)=\dfrac{\nu_n}{n}$ then using Stirlings formula, we obtain

$$P(\nu_n=n)=P\big(H_n^*(+\infty)=1\big)=\frac{n^ne^{-n}}{n!}=\frac{1}{\sqrt{2\pi n}}(1+o(1)),\quad n\to\infty,$$

and

$$P\big(H_n^*(+\infty)>1\big)=P(\nu_n>n)=\sum_{k=n+1}^{\infty}\frac{n^ke^{-n}}{k!}=o(1),\quad n\to\infty.$$

Thus $H_n^*(x)$ with positive probability is greater than 1. In order to avoid these undesirable property the following modifications of the Kac statistics is proposed

$$\begin{aligned}\widetilde{H}_n(x)&=1-\big(1-H_n^*(x)\big)I\big(H_n^*(x)<1\big),\quad x\in\overline{\mathbb{R}},\\ \widetilde{H}_n(x;i)&=1-\big(1-H_n^*(x;i)\big)I\big(H_n^*(x;i)<1\big),\quad (x;i)\in\overline{\mathbb{R}}\times\Im.\end{aligned}\tag{11}$$

The following inequalities are useful in studying the Kac processes.

Theorem 2. *Let* $\{\nu_n,n\geqslant 1\}$ *be a sequence of Poisson r.v.-s with* $E\nu_n=n$. *Then for any* $\varepsilon>0$ *such that*

$$\frac{n}{\log n}\geqslant\frac{\varepsilon}{8(1+\frac{e}{3})^2},\quad e=\exp(1),\tag{12}$$

we have

$$P\bigg(\big|\nu_n-n\big|>\frac{1}{2}\Big(\frac{\varepsilon}{2}n\log n\Big)^{\frac12}\bigg)\leqslant 2n^{-\varepsilon w},\tag{13}$$

$$P\bigg(\sup_{|x|<\infty}\big|H_n^*(x;i)-H(x;i)\big|>2\Big(\frac{\varepsilon\log n}{2n}\Big)^{\frac12}\bigg)\leqslant 4n^{-4\varepsilon w},\ i\in\Im,\tag{14}$$

$$P\bigg(\sup_{|x|<\infty}\big|\widetilde{H}_n(x;i)-H(x;i)\big|>2\Big(\frac{\varepsilon\log n}{2n}\Big)^{\frac12}\bigg)\leqslant 4n^{-4\varepsilon w},\ i\in\Im,\tag{15}$$

where $w=\left[16\Big(1+\dfrac{e}{3}\Big)\right]^{-1}$.

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

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

Using Taylor expansion for e^t, we obtain

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

where $\psi(t) = \dfrac{t^3}{6} \exp(\theta t), \quad 0 < \theta < 1$. Taking into account that $t^3 \leqslant t^2$ for $0 \leqslant t \leqslant 1$, we obtain the estimate for $\psi(t)$: $\psi(t) \leqslant \dfrac{t^3}{6} e \leqslant e \dfrac{t^2}{6}$. Thus, $E \exp(t\xi_k) = \exp\left\{\dfrac{t^2}{2}\left(1 + \dfrac{e}{3}\right)\right\}, \quad 0 \leqslant t \leqslant 1$.

The following result (from [13]) is necessary for further considerations.

Lemma 1 ([13]). *Let $\{\xi_n, n \geqslant 1\}$ be a sequence of independent r.v.-s with $E\xi_n = 0$, $n = 1, 2, \ldots$. Suppose that $U, \lambda_1, \ldots, \lambda_n$ are positive real numbers such that*

$$E \exp(t\xi_k) \leqslant \exp\left(\frac{1}{2}\lambda_k t_k^2\right) \quad \text{for} \quad k = 1, 2, \ldots, n \quad |t| \leqslant U. \tag{16}$$

Let $\Lambda = \lambda_1 + \cdots + \lambda_n$. Then

$$P\Big(\big|\xi_1 + \cdots + \xi_k\big| \geqslant z\Big) \leqslant \begin{cases} 2\exp\left(-\dfrac{z^2}{2\Lambda}\right) & if \quad o \leqslant z \leqslant \Lambda U, \\ 2\exp\left(-\dfrac{Uz}{2}\right) & if \quad z > \Lambda U. \end{cases}$$

Let us assume that $\lambda_k = 1 + \dfrac{e}{3}$, $U = 1$, $z = \dfrac{1}{2}\left(\frac{\varepsilon}{2} n \log n\right)^{1/2}$ in Lemma 1 then we obtain (13). Here $0 \leqslant z = \dfrac{1}{2}\left(\frac{\varepsilon}{2} n \log n\right)^{1/2} \leqslant \left(1 + \frac{e}{3}\right) n = \Lambda U$. Consider probability in (14). Using total probability formula, we have

$$P\left(\sup_{|x|<\infty} \big|H_n^*(x;i) - H(x;i)\big| > 2\left(\frac{\varepsilon \log n}{2n}\right)^{\frac{1}{2}}\right) =$$

$$= P\left(\sup_{|x|<\infty} \left|H_n(x;i) - H(x;i) + \frac{1}{n}\sum_{k=n+1}^{\nu_n} \delta_k^{(i)} I(Z_k \leqslant x)\right| > 2\left(\frac{\varepsilon \log n}{2n}\right)^{\frac{1}{2}} \Big/ \nu_n > n\right) \cdot P(\nu_n > n) +$$

$$+ P\left(\sup_{|x|<\infty} \left|H(x;i) - H(x;i) - \frac{1}{n}\sum_{k=\nu_n+1}^{n} \delta_k^{(i)} I(Z_k \leqslant x)\right| > 2\left(\frac{\varepsilon \log n}{2n}\right)^{\frac{1}{2}} \Big/ \nu_n \leqslant n\right) \cdot P(\nu_n \leqslant n) \leqslant$$

$$\leqslant P\left(\sup_{|x|<\infty} \big|H_n(x;i) - H(x;i)\big| > \left(\frac{\varepsilon \log n}{2n}\right)^{\frac{1}{2}}\right) + P\left(\sup_{|x|<\infty} \left|\frac{1}{n}\sum_{k=\min(n,\nu_n)+1}^{\max(n,\nu_n)} \delta_k^{(i)} I(Z_k \leqslant x)\right| >\right.$$

$$\left. > \left(\frac{\varepsilon \log n}{2n}\right)^{\frac{1}{2}}\right) \leqslant 2n^{-4\varepsilon} + P\left(\left|\frac{\nu_n - n}{n}\right| > \left(\frac{\varepsilon \log n}{2n}\right)^{\frac{1}{2}}\right) \leqslant 2n^{-4\varepsilon} + 2n^{-4w\varepsilon} \leqslant 4n^{-4w\varepsilon}, \quad i \in \Im,$$

where we applied (2) and (13) that proves (14). Let us define $T_n^{(i)} = \inf\big\{x : \widetilde{H}_n(x;i) = 1\big\}$, $i \in \Im$. If $x \geqslant \widetilde{T}_n^{(i)}$ and $\nu_n > n$ then $\widetilde{H}_n(x;i) = 1$ and $H_n^*(x;i) - H\,(x;i) \geqslant H_n^*(x;i) - \widetilde{H}(x;i) \geqslant 0$. Then assuming $\nu_n > n$, we obtain

$$\sup_{|x|<\infty}\left|\widetilde{H}_n(x;i)-H(x;i)\right| = \left\{\max\left[\sup_{x<\widetilde{T}_n^{(i)}}\left|H_n^*(x;i)-H(x;i)\right|, \sup_{x\geqslant\widetilde{T}_n^{(i)}}\left|\widetilde{H}_n(x;i)-H(x;i)\right|\right]\right\} \leqslant$$
$$\leqslant \left\{\max\left[\sup_{x<T_n^{(i)}}\left|H_n^*(x;i)-H(x;i)\right|, \sup_{x\geqslant T_n^{(i)}}\left|H_n^*(x;i)-H(x;i)\right|\right]\right\} =$$
$$= \sup_{|x|<\infty}\left|H_n^*(x;i)-H(x;i)\right|, \quad i\in\Im. \tag{17}$$

With $\nu_n \leqslant n$, it is obvious that $\widetilde{H}_n(x;i) = H_n^*(x;i)$ for all $(x;i)\in\overline{\mathbb{R}}\times\Im$.

Now taking into account the last two relations, total probability formula and (14), we obtain (15). Theorem 2 is proved. □

Let $\widetilde{a}_n(t) = \left(\widetilde{a}_n^{(0)}(t_0), \widetilde{a}_n^{(1)}(t_1), \ldots, \widetilde{a}_n^{(k)}(t_k)\right)$, where $\widetilde{a}_n^{(0)}(x) = \sqrt{n}\left(\widetilde{H}_n(x)-H(x)\right)$, $\widetilde{a}_n^{(i)}(x) = \sqrt{n}\left(\widetilde{H}_n(x;i)-H(x;i)\right)$, $(x;i)\in\overline{\mathbb{R}}\times\Im$. We will prove an approximation theorem of the vector-valued modified empirical Kac process $\widetilde{a}_n(t)$ by the appropriate Gaussian vector-valued process $W_n^*(t)$, $t\in\overline{\mathbb{R}}^{k+1}$ from Theorem 2.

Theorem 3. *Let $\{T_n, n\geqslant 1\}$ be a numerical sequence satisfying for each n the condition $T_n < T_H = \inf\{x : H(x)=1\} \leqslant \infty$ such that*

$$\min_{i\in\Im}\left\{P(A^{(i)}) - H(T_n,i)\right\} \geqslant 1-H(T_n) \geqslant 2\left(\frac{r\log n}{2wn}\right)^{1/2}. \tag{18}$$

If for any $\varepsilon>0$ condition (12) holds then on the probability space of Theorem 2 one can define $k+1$ sequences of mean zero Gaussian processes $W_n^{(0)}(x), W_n^{(1)}(x), \ldots, W_n^{(k)}(x)$ with the covariance structure (9) such that for $\widetilde{a}_n(t)$ and $W_n^*(t) = \left(W_n^{(0)}(t_0), W_n^{(1)}(t_1), \ldots, W_n^{(k)}(t_k)\right)$ we have

$$P\left\{\sup_{t\in(-\infty;T_n]^{(k+1)}}\|\widetilde{a}_n(t)-W_n^*(t)\|^{(k+1)} > \widetilde{C}n^{\frac{1}{2}}\log n\right\} \leqslant \widetilde{K}n^{-\beta}, \tag{19}$$

where $\widetilde{K}$ is an absolute constant, $\widetilde{C}=\widetilde{C}(\varepsilon)$ and $\beta=\min(r,\varepsilon w)$ for any $\varepsilon>0$.

Proof. It is easy to see that probability in (19) can be estimated by the sum

$$\begin{gathered}-P\left\{\sup_{x\leqslant T_n}\left|\widetilde{a}_n^{(0)}(x)-W_n^{(0)}(x)\right| > \widetilde{C}n^{\frac{1}{2}}\log n\right\} + \\ +\sum_{i=1}^{k} P\left(\sup_{x\leqslant T_n}\left|\widetilde{a}_n^{(i)}(x)-W_n^{(i)}(x)\right| > \widetilde{C}n^{\frac{1}{2}}\log n\right) = q_{1n}+q_{2n}.\end{gathered} \tag{20}$$

Taking into account that for any $x \leqslant T_n$, $H_n^*(x) \leqslant H_n^*(T_n)$, and if $H_n^*(T_n) \leqslant 1$ then $\widetilde{a}_n^{(0)}(x) = \widetilde{a}_n^{(0)*}(x)$. Using formula of total probability, we have

$$\begin{gathered} q_{1n} \leqslant P\left(\sup_{x\leqslant T_n}\left|\widetilde{a}_n^{(0)}(x)-W_n^{(0)}(x)\right| > C^* n^{-\frac{1}{2}}\log n \Big/ H_n^*(T_n)\leqslant 1\right) + P\left(H_n^*(T_n)>1\right) \leqslant \\ \leqslant P\left(\sup_{x\leqslant T_n}\left|a_n^{(0)*}(x)-W_n^{(0)}(x)\right| > C^* n^{-\frac{1}{2}}\log n\right) + P(H_n^*(T_n)>1) \leqslant \\ \leqslant Kn^{-r} + P\left(H_n^*(T_n)-H(T_n) > 1-H(T_n)\right) \leqslant \\ \leqslant K^* n^{-r} + P\left(\sup_{|x|<\infty}|H_n^*(x)-H(x)| > \left(\frac{r\log n}{2wn}\right)^{\frac{1}{2}}\right) \leqslant Ln^{-r}, \end{gathered} \tag{21}$$

where Theorem 1 and the analogue of (14) for $H_n^* - H,\ L = K^* + 4$ are used. Analogously,

$$
\begin{aligned}
q_{2n} \leqslant \sum_{i=1}^{k} P\left(\sup_{x\leqslant T_n}\left|\widetilde{a}_n^{(i)}(x) - W_n^{(i)}(x)\right| > C^* n^{\frac{1}{2}}\log n\right) + \sum_{i=1}^{k} P\left(H_n^*(T_n;i) > P\left(A^{(i)}\right)\right) \leqslant \\
\leqslant \sum_{i=1}^{k} P\left(\sup_{x\leqslant T_n}\left|a_n^{(i)*}(x) - W^{(i)}(x)\right| > C^* n^{-\frac{1}{2}}\log n\right) + \\
+ \sum_{i=1}^{k} P\left(\sup_{|x|<\infty}\left|a_n^{(i)*}(x) - W^{(i)}(x)\right| > C^* n^{-\frac{1}{2}}\log n\right) + \\
+ kP\left(\frac{|\nu_n - n|}{n} > \frac{1}{2}\left(\frac{4r\log n}{2wn}\right)^{\frac{1}{2}}\right) \leqslant kLn^{-r} + 2kn^{-4r},
\end{aligned}
\tag{22}
$$

where inequalities (13), (15) and Theorem 1 are used. Now (19) follows from (21) and (22). Theorem 3 is proved. □

3. Estimation of exponential-hazard function

In many practical situations when we are interested in the joint behaviour of the pairs $\left\{\left(Z, A^{(i)}\right), i \in \Im\right\}$ the so-called cumulative hazard functions $\left\{S^{(i)}(x) = \exp\left(-\Lambda^{(i)}(x)\right), i \in \Im\right\}$ plays a crucial role. Here $\Lambda^{(i)}(x)$ is the i-th hazard function $\left(\int\limits_{-\infty}^{x} = \int\limits_{(-\infty;x]}\right)$

$$
\Lambda^{(i)}(x) = \int_{-\infty}^{x} \frac{dH(u;i)}{1 - H(u)}, \quad i \in \Im,
$$

where $\Lambda^{(1)}(x) + \cdots + \Lambda^{(k)}(x) = \Lambda(x) = \int\limits_{-\infty}^{x} \dfrac{dH(u)}{1 - H(u)}$ is the corresponding hazard function of d.f. $H(x)$.

Let us consider two important special cases of the considered generalized censorship model:

1. Let $\{X_1, X_2, \dots\}$ be a sequence of independent r.v.-s with common continuous d.f. F. They are censored on the right by a sequence $\{Y_1, Y_2, \dots\}$ of independent r.v.-s. They are independent of the X-sequence with common continuous d.f. G. One can only observe the sequence of pairs $\left\{(Z_k, \delta_k), k = \overline{1,n}\right\}$, where $Z_j = \min(X_j, Y_j)$ and $\delta_j = \delta_j^{(1)}$ is the indicator of event $A_j = A_j^{(1)} = \{Z_j = X_j\}$. In this case $k = 2,\ 1 - H(x) = (1 - F(x))(1 - G(x))$, $H(x;1) = \int\limits_{-\infty}^{x}(1 - G(u))dF(u)$. Thus $S^{(1)}(x) = S(x) = 1 - F(x)$. The useful special case is $1 - G(x) = (1 - F(x))^{\beta},\ \beta > 0$ which corresponds to independence of r.v.-s Z_j and $\delta_j, j \geqslant 1$.

2. Let us assume that $k > 1$ and consider independent sequences $\left\{Y_1^{(i)}, Y_2^{(i)}, \dots\right\}$ $(i = 1, \dots, k)$ of independent r.v.-s with common continuous d.f. F. Let $Z_j = \min\left(Y_j^{(1)}, \dots, Y_j^{(k)}\right)$. Let us observe the sequences $\left\{\left(Z_j, \delta_j^{(i)}\right),\ i = \overline{1,k}\right\}_{j=1}^{n}$, where $\delta_j^{(i)}$ is the indicator of the event $A_j^{(i)} =$
$= \left\{Z_j = Y_j^{(i)}\right\}$. This is the competing risks model with $S^{(i)}(x) = 1 - F^{(i)}(x),\ i \in \Im$.

Let us define the natural Kac-type estimator

$$\widetilde{\Lambda}_n^{(i)}(x) = \int_{-\infty}^{x} \frac{d\widetilde{H}(u;i)}{1-\widetilde{H}_n(u)}, \quad i \in \Im$$

of $\Lambda^{(i)}(x),\ i \in \Im$. Let $w_n^{(i)}(x) = \sqrt{n}\left(\widetilde{\Lambda}_n^{(i)}(x) - \Lambda^{(i)}(x)\right),\ i \in \Im$, is an Kac-type hazard process, $w_n(t) = \left(w_n^{(1)}(t_1), \dots, w_n^{(k)}(t_k)\right),\ t = (t_1, \dots, t_k)$, and $Y_n(t) = \left(Y_n^{(1)}(t_1), \dots, Y_n^{(k)}(t_k)\right)$ is the corresponding vector process with

$$Y_n^{(i)}(x) = \int_{-\infty}^{x} \frac{W_n^{(0)}(u)dH(u;i)}{(1-H(u))^2} + \frac{W_n^{(i)}(x)}{1-H(x)} - \int_{-\infty}^{x} \frac{W_n^{(i)}(u)dH(u)}{(1-H(u))^2}, \quad i \in \Im$$

and $\left\{W_n^{(0)}(x), W_n^{(1)}(x), \dots, W_n^{(k)}(x)\right\}$ are Wiener processes with the covariance structure (9). Then for $i \in \Im$, $EY_n^{(i)} = 0$ and

$$EY_n^{(i)}(x)Y_n^{(i)}(y) = C(x,y),$$

where $x, y \leqslant T_H = \inf\{x : H(x) = 1\} \leqslant \infty$.

Theorem 4. *Let $\{T_n, n \geqslant 1\}$ be a numerical sequence satisfying for each n the condition $T_n < T_H$ such that*

$$\frac{n}{\log n} \geqslant \max\left\{32\varepsilon w^2, \frac{2rb_n^2}{w}, \frac{2\varepsilon b_n^2}{w}\right\}, \tag{23}$$

where $b_n = (1-H(T_n))^{-1},\ \varepsilon > 0,\ r \geqslant 2$. Then

$$P\left(\sup_{t \in (-\infty;T_n]^{(k)}} \|w_n(t) - Y_n(t)\|^{(k)} > r(n)\right) \leqslant k\Phi_1 n^{-\beta}, \tag{24}$$

on a probability space of Theorem 2, where $r(n) = \Phi_0 b_n^2 n^{-\frac{1}{2}} \log n$, $Phi_0 = \Phi_0(\varepsilon, r)$, Φ_1- are absolute constants.

Proof. It is sufficient to prove that for each $i \in \Im$

$$P\left(\sup_{x \leqslant T_n}\left(w_n^{(i)}(x) - Y_n^{(i)}(x)\right) > r(n)\right) \leqslant \Phi_1 n^{-\beta}. \tag{25}$$

We have representation for each $i \in \Im$ for difference

$$w_n^{(i)}(x) - Y_n^{(i)}(x) = \int_{-\infty}^{x} \frac{\left(\widetilde{a}_n^{(0)}(u) - W_n^{(0)}(u)\right)dH(u;i)}{(1-H(u))^2} + \frac{\widetilde{a}_n^{(i)}(x) - W_n^{(i)}(x)}{1-H(x)} -$$
$$- \int_{-\infty}^{x} \frac{\left(a_n^{(i)}(u) - W_n^{(i)}(u)\right)dH(u)}{(1-H(u))^2} + n^{-\frac{1}{2}} \int_{-\infty}^{x} \frac{\left(\widetilde{a}_n^{(0)}(u)\right)^2 dH(u;i)}{(1-H(u))^2\left(1-\widetilde{H}_n(u)\right)} +$$
$$+ n^{-\frac{1}{2}} \int_{-\infty}^{x} \frac{\widetilde{a}_n^{(0)}(u)d\widetilde{a}_n^{(i)}(u)}{(1-H(u))\left(1-\widetilde{H}_n(u)\right)} = \sum_{m=1}^{4} R_{mn}^{(i)}(x).$$

Using (15) and (19), we have for sum $R_{1n}^{(i)}(x) + R_{2n}^{(i)}(x) + R_{3n}^{(i)}(x)$

$$P\left(\sup_{x \leqslant T_n}\left|\sum_{m=1}^{4} R_{mn}^{(i)}(x)\right| > 3\widetilde{C}n^{-\frac{1}{2}}\log n + \varepsilon n^{-\frac{1}{2}} b_n^3 \log n\right) \leqslant$$
$$\leqslant 3\widetilde{K}n^{-\beta} + 2Ln^{-w\varepsilon} \leqslant (3\widetilde{K} + 2L)n^{-\beta}, \quad i \in \Im. \tag{26}$$

Rewrite $R_{4n}^{(i)}$ in the form

$$R_{4n}^{(i)}(x) = n^{-\frac{1}{2}} \int_{-\infty}^{x} \frac{\left(\widetilde{a}_n^{(0)}(u)\right)^2 d(H(u;i) - H(u;i))}{(1-H(u))^2 \left(1 - \widetilde{H}_n(u)\right)} + \\ + n^{-\frac{1}{2}} \int_{-\infty}^{x} \frac{\widetilde{a}_n^{(0)}(u) d\widetilde{a}_n^{(i)}(u)}{(1-H(u))^2} = \overline{R}_{4n}^{(i)}(x) + \overline{\overline{R}}_{4n}^{(i)}(x). \tag{27}$$

Then taking into account (15), we obtain for $i \in \Im$

$$P\left(\sup_{x \leqslant T_n} \left|\overline{R}_{4n}^{(i)}(x)\right| > 2\varepsilon n^{-\frac{1}{2}} b_n^3 \log n\right) \leqslant 2Ln^{-w\varepsilon} \leqslant 2Ln^{-\beta}. \tag{28}$$

There exists an absolute constant A such that

$$P\left(\sup_{x \leqslant T_n} \left|\overline{\overline{R}}_{4n}^{(i)}(x)\right| > 3An^{-\frac{1}{2}} b_n^2 \log n\right) \leqslant P\left(H_n^*(T_n) > 1\right) + \\ + P\left(\sup_{x \leqslant T_n} n^{-\frac{1}{2}} \left|\int_{-\infty}^{x} \frac{a_n^{(0)*}(u) da_n^{(i)*}(u)}{(1-H(u))^2}\right| > 3An^{-\frac{1}{2}} b_n^2 \log n\right) \leqslant Ln^{-r} + p_n, \tag{29}$$

so that for any $x \leqslant T_n, H_n^*(x) \leqslant H_n^*(T_n)$ and if $H_n^*(T_n) \leqslant 1$ then $H_n^*(x;i) \leqslant H_n^*(T_n)$ and hence $\widetilde{a}_n^{(i)}(x) = a_n^{(i)*}(x)$ for $i \in \Im$. It is sufficient to estimate probability p_n. According to proof of Theorem 1 in [6], supposing $a_{\nu_n}^{(0)}(x) = \sqrt{\nu_n}(H_{\nu_n}^*(x) - H(x))$, $a_{\nu_n}^{(i)}(x) = \sqrt{\nu_n}(H_{\nu_n}^*(x;i) - H(x;i))$, $i \in \Im$ and using representation (10), we have $p_n = p_{1n} + p_{2n} + p_{3n} + p_{4n}$, where

$$p_{1n} = P\left(\frac{\nu_n}{n} \sup_{x \leqslant T_n} \left|\int_{-\infty}^{x} \frac{a_{\nu_n}^{(0)}(u) da_{\nu_n}^{(i)}(u)}{(1-H(u))^2}\right| > 3An^{-\frac{1}{2}} b_n^2 \log n\right),$$

$$p_{2n} = P\left(\sqrt{\frac{\nu_n}{n}} \cdot \frac{|\nu_n - n|}{n} \sup_{x \leqslant T_n} \left|\int_{-\infty}^{x} \frac{a_{\nu_n}^{(0)}(u) dH(u;i)}{(1-H(u))^2}\right| > \frac{\varepsilon}{2}\left(\frac{3}{2}\right)^{-\frac{1}{2}} n^{-\frac{1}{2}} b_n^2 \log n\right),$$

$$p_{3n} = P\left(\sqrt{\frac{\nu_n}{n}} \cdot \frac{|\nu_n - n|}{n} \sup_{x \leqslant T_n} \left|\int_{-\infty}^{x} \frac{H(u) da_{\nu_n}(u)}{(1-H(u))^2}\right| > \frac{\varepsilon}{2}\left(\frac{3}{2}\right)^{-\frac{1}{2}} n^{-\frac{1}{2}} b_n^2 \log n\right),$$

$$p_{4n} = P\left(\cdot \frac{|\nu_n - n|^2}{\sqrt{n}} \sup_{x \leqslant T_n} \left\{\int_{-\infty}^{x} \frac{H(u) dH(u;i)}{(1-H(u))^2}\right\} > \frac{\varepsilon}{8} n^{-\frac{1}{2}} b_n^2 \log n\right).$$

Taking into account Lemma in [5], we have

$$P\left(\sup_{x \leqslant T_n} \left|\int_{-\infty}^{x} \frac{a_n^{(0)}(u) da_n^{(i)}(u)}{(1-H(u))^2}\right| > Ab_n^2 \log n\right) \leqslant Bn^{-\varepsilon}, \tag{30}$$

where $A = A(\varepsilon)$ and B is an absolute constant. Moreover, using (13), we have

$$P\left(\frac{|\nu_n - n|}{n} > \frac{1}{2}\right) \leqslant 2n^{-\frac{2nw}{\log n}}. \tag{31}$$

It follows from (30) and (31) that

$$p_{1n} = P\left(\sup_{x \leqslant T_n} \left|\int_{-\infty}^{x} \frac{a_{\nu_n}^{(0)}(u) da_{\nu_n}^{(i)}(u)}{(1-H(u))^2}\right| > 2Ab_n^2 \log \nu_n \frac{\log n}{\log \nu_n},\ \frac{n}{2} \leqslant \nu_n \leqslant \frac{3n}{2}\right) + P\left(\frac{|\nu_n - n|}{n} > \frac{1}{2}\right) \leqslant \\ \leqslant 2n^{-\frac{2nw}{\log n}} + P\left(\sup_{x \leqslant T_n} \left|\int_{-\infty}^{x} \frac{a_{\nu_n}^{(0)}(u) da_{\nu_n}^{(i)}(u)}{(1-H(u))^2}\right| > Ab_n^2 \log \nu_n\right) + 2n^{-\frac{2nw}{\log n}} \leqslant$$

$$\leqslant e^{-n} + \sum_{m=1}^{\infty} P\left(\sup_{x\leqslant T_n}\left|\int_{-\infty}^{x} \frac{a_m^{(0)}(u)da_m^{(i)}(u)}{(1-H(u))^2}\right| > Ab_n^2 \log m\right) P(\nu_n = m) + 2n^{-\frac{2nw}{\log n}} \leqslant$$
$$\leqslant e^{-n} + 2n^{-\frac{2nw}{\log n}} + B\sum_{m=1}^{\infty} m^{-\varepsilon}\cdot\frac{n^m}{m!}e^{-n} \leqslant e^{-n} + 2n^{-\frac{2nw}{\log n}} + \widetilde{B}n^{-\varepsilon}. \quad (32)$$

Analogously, using (31) and (1), we obtain

$$p_{2n} = P\left(\sqrt{\frac{\nu_n}{n}}\frac{|\nu_n - n|}{n}\sup_{x\leqslant T_n}\left|\int_{-\infty}^{x}\frac{a_{\nu_n}^{(0)}(u)dH(u;i)}{(1-H(u))^2}\right| > \frac{\varepsilon}{2}\left(\frac{3}{2}\right)^{\frac{1}{2}} n^{-\frac{1}{2}}b_n^2\log n,\ \frac{n}{2}\leqslant\nu_n\leqslant\frac{3n}{2}\right) +$$
$$+P\left(\frac{|\nu_n - n|}{n} > \frac{1}{2}\right) \leqslant 2n^{-\frac{2nw}{\log n}} + 2n^{-w\varepsilon} + P\left(\sup_{|x|<\infty}\left|a_{\nu_n}^{(0)}(x)\right| > \left(\frac{\varepsilon}{2}\log\nu_n\right)^{\frac{1}{2}}\right) \leqslant$$
$$\leqslant 2n^{-\frac{2nw}{\log n}} + 2n^{-w\varepsilon} + e^{-n} + \widetilde{D}n^{-\varepsilon}. \quad (33)$$

Integrating by parts and using (2), we obtain

$$p_{3n} \leqslant 2n^{-\frac{2nw}{\log n}} + 2n^{-w\varepsilon} + P\left(\sup_{|x|<\infty}\left|a_{\nu_n}^{(i)}(x)\right| > (2\varepsilon\log\nu_n)^{\frac{1}{2}}\right) \leqslant$$
$$\leqslant 2n^{-\frac{2nw}{\log n}} + 2n^{-w\varepsilon} + e^{-n} + 2Dn^{-\varepsilon}. \quad (34)$$

Finally, using (13), we have

$$p_{4n} \leqslant P\left(\frac{|\nu_n - n|}{n^{\frac{1}{2}}} > \frac{1}{2}\left(\frac{\varepsilon}{2}\log n\right)^{\frac{1}{2}}\right) \leqslant 2n^{-w\varepsilon}. \quad (35)$$

Now combining (26)-(29) and (32)-(35), we obtain (25). Theorem 4 is proved.

Corollary 1. *It follows from* (24) *that for suitable* $r \geqslant 2$ *and* $\varepsilon > 0$ *one can obtain an approximation on* $(-\infty;T]^{(k)}$ *with* $b^{-1} = 1 - H(T) > 0$:

$$\sup_{t\in(-\infty;T]^{(k)}}\|w_n(t) - Y_n(t)\|^{(k)} \overset{a.s.}{=} O\left(n^{-\frac{1}{2}}\log n\right), \quad n\geqslant 2. \quad (36)$$

Now we consider joint estimation of exponential-hazard functions $\{S^{(i)}x) = \exp\left(-\Lambda^{(i)}(x)\right),\ i\in\Im\}$. Let us consider hazard function estimate

$$\Lambda_n(x) = \int_{-\infty}^{x}\frac{d\widetilde{H}_n(u)}{1-\widetilde{H}_n(u)}$$

and corresponding hazard process $w^{(0)}(x) = \sqrt{n}\,(\Lambda_n(x) - \Lambda(x))$. In the next Theorem 5 we approximate $w_n^{(0)}(x)$ by sequence of Gaussian processes $Y_n^{(0)}(x) = \dfrac{W_n^{(0)}(x)}{1-H(x)}$.

Theorem 5. *Let* $\{T_n, n\geqslant 1\}$ *be a numerical sequence that satisfies the condition* $T_n < T_H$ *for each* n *such that* (23) *holds. Then on a probability space of Theorem 2 we have*

$$P\left(\sup_{x\leqslant T_n}\left|w_n^{(0)}(x) - Y_n^{(0)}(x)\right| > r_0(n)\right) \leqslant \Psi_1 n^{-\beta}, \quad (37)$$

where $r_0(n) = \Phi_0 b_n^2 n^{-\frac{1}{2}}\log n$ *and* $\Phi_0 = \Phi_0(\varepsilon, r)$, Ψ_1 *are absolute constants.*

Proof. It is easy to verify that

$$w_n^{(0)}(x) - Y_n^{(0)}(x) = \frac{\left(\widetilde{a}_n^{(0)}(x) - W_n^{(0)}(x)\right)}{1 - H(x)} + n^{-\frac{1}{2}} \int_{-\infty}^{x} \frac{\left(\widetilde{a}_n^{(0)}(u)\right)^2 dH(u)}{(1 - H(u))^2 \left(1 - \widetilde{H}_n(u)\right)} +$$

$$+ n^{-\frac{1}{2}} \int_{-\infty}^{x} \frac{\widetilde{a}_n^{(0)}(u) da_n^{(0)}(u)}{(1 - H(u)) \left(1 - \widetilde{H}_n(u)\right)}.$$

Now further proof of (37) is similar to the proof of Theorem 4 and hence details are omitted. Theorem 5 is proved. □

One can obtain from Theorems 4 and 5 the following theorem on deviations of processes $w_n^{(0)}$ and $w_n^{(i)}, \quad i \in \Im$.

Theorem 6. *Let $\{T_n, n \geqslant 1\}$ be a numerical sequence that satisfies for each n the condition $T_n < T_H$ such that (23) holds. Then*

$$P\left(\sup_{x \leqslant T_n} \left|w_n^{(0)}(x)\right| > r_0(n) + 2b_n(\varepsilon \log n)^{\frac{1}{2}}\right) \leqslant \Psi_1 n^{-\beta} + n^{-\varepsilon}, \tag{38}$$

and for $i \in \Im$

$$P\left(\sup_{x \leqslant T_n} \left|w_n^{(i)}(x)\right| > r_0(n) + 6b_n^2(\varepsilon \log n)^{\frac{1}{2}}\right) \leqslant \Psi_1 n^{-\beta} + 3n^{-\varepsilon}. \tag{39}$$

Proof. It is easy to verify that for any $n \geqslant 1$

$$W_n^{(0)}(x) \stackrel{D}{=} W\left(H(x)\right) \quad \text{and} \quad W_n^{(i)}(x) \stackrel{D}{=} W\left(H(x;i)\right), \quad (x;i) \in \overline{\mathbb{R}} \times \Im,$$

where $\{W(y), 0 \leqslant y \leqslant 1\}$ is a standard Wiener process on $[0,1]$. Then probability in (38) is not greater than

$$\begin{gathered} P\left(\sup_{x \leqslant T_n} \left|w_n^{(0)}(x) - Y_n^{(0)}(x)\right| > r_0(n)\right) + P\left(\sup_{x \leqslant T_n} \left|Y_n^{(0)}(x)\right| > 2b_n\left(\varepsilon \log n\right)^{\frac{1}{2}}\right) \leqslant \\ \leqslant \Psi_1 n^{-\beta} + P\left(|W(1)| > 2\left(\varepsilon \log n\right)^{\frac{1}{2}}\right) \leqslant \Psi_1 n^{-\beta} + n^{-\varepsilon}, \end{gathered} \tag{40}$$

where inequality (37) and well-known exponential inequality for Wiener process (see [14], Eq. (29.2)) are used. Analogously, (39) follows from (25) and the second estimate in (40). Theorem 6 is proved. □

To estimate the exponential hazard functions $\left\{S^{(i)}(x) = \exp\left(-\Lambda^{(i)}(x)\right), i \in \Im\right\}$ we use the following exponential of Altshuler-Breslow, product-limit of Kaplan-Meier and relative risk power estimates of Abdushukurov ([1–3]):

$$\begin{aligned} S_{1n}^{(i)}(x) &= \exp\left(-\Lambda_n^{(i)}(x)\right), \\ S_{2n}^{(i)}(x) &= \textstyle\prod_{u \leqslant x}\left(1 - \Delta\Lambda_n^{(i)}(x)\right), \\ S_{3n}^{(i)}(x) &= \left[1 - H_n(x)\right]^{R_n^{(i)}(x)}, \end{aligned} \tag{41}$$

where $R_n^{(i)}(x) = \Lambda_n^{(i)}(x)(\Lambda_n(x))^{-1}, \ i \in \Im$.

It follows from the proof of Theorem 1.4.1 in [3] that for all $(x;i) \in (-\infty, Z_{(n)}) \times \Im$, $Z_{(n)} = \max(Z_1, \dots, Z_n)$

$$0 \leqslant S_{1n}^{(i)}(x) - S_{2n}^{(i)}(x) \leqslant \frac{1}{2n}\int_{-\infty}^{x} \frac{d\widetilde{H}_n(u;i)}{\left(1-\widetilde{H}_n(u)\right)^2} \overset{a.s.}{=} O\left(\frac{1}{n}\right),$$
$$0 \leqslant S_{1n}^{(i)}(x) - S_{3n}^{(i)}(x) \leqslant \frac{1}{2n}\int_{-\infty}^{x} \frac{d\widetilde{H}_n(u;i)}{\left(1-\widetilde{H}_n(u)\right)^2} \overset{a.s.}{=} O\left(\frac{1}{n}\right). \tag{42}$$

Hence it is sufficient to consider only estimator $S_{1n}^{(i)}$. Let us introduce vector-processes $\mathbb{Q}_n(t) = \left(\mathbb{Q}_n^{(1)}(t_1), \ldots, \mathbb{Q}_n^{(k)}(t_k)\right)$ and $\mathbb{Q}_n^*(t) = \left(\mathbb{Q}_n^{(1)*}(t_1), \ldots, \mathbb{Q}_n^{(k)*}(t_k)\right)$, where $\mathbb{Q}_n^{(i)}(x) = \sqrt{n}\left(S_{1n}^{(i)}(x) - S^{(i)}(x)\right)$ and $\mathbb{Q}_n^{(i)*}(x) = S^{(i)}(x)Y_n^{(i)}(x),\ i \in \Im$.

In the next theorem vector-valued process $Q_n(t)$ is approximated by Gaussian vector-valued process $\mathbb{Q}_n^*(t),\ t \in \mathbb{R}^k$.

Theorem 7. *Let $\{T_n,\, n \geqslant 1\}$ be a numerical sequence that satisfies for each n the condition $T_n < T_H$ such that inequality* (23) *holds. Then we have on a probability space of Theorem 2*

$$P\left(\sup_{t\in(-\infty;T_n]^{(k)}} \|\mathbb{Q}_n(t) - \mathbb{Q}_n^*(t)\|^{(k)} > r^*(n)\right) \leqslant kR^* n^{-\beta}, \tag{43}$$

where $r^*(n) = \left\{r_0(n) + \frac{1}{2}n^{-\frac{1}{2}}\left(r(n) + 6b_n^2(\varepsilon\log n)^{\frac{1}{2}}\right)^2\right\}$ *and R^* is an absolute constant.*

Proof. Using Taylor expansion for each $i \in \Im$, we obtain

$$\mathbb{Q}_n^{(i)}(x) = S^{(i)}(x)w_n^{(i)}(x) + \frac{1}{2}n^{-\frac{1}{2}}\exp\left(-\theta_n^{(i)}(x)\right)\left(w_n^{(i)}(x)\right)^2,$$

where $\theta_n^{(i)}(x) \in \left[\min\left(\Lambda_n^{(i)}(x), \Lambda^{(i)}(x)\right), \max\left(\Lambda_n^{(i)}(x), \Lambda^{(i)}(x)\right)\right]$. Now using (24), (38) and (39), we obtain the required result. Theorem 7 is proved. □

4. Estimation of characteristic function under random right censoring

Let $X_1, X_2, \ldots$ be independent identically distributed r.v.-s with common continuous d.f. F. They are interpreted as an infinite sample of the random lifetime X. Another sequence of independent and identically distributed r.v.-s $Y_1, Y_2, \ldots$ with common continuous d.f. G censors on the right is introduced. This sequence is independent of $\{X_j\}$. Then the observations available at the n-th stage consist of the pairs $\{(Z_j,\delta_j), 1 \leqslant j \leqslant n\} = \mathbb{C}^{(n)}$, where $Z_j = \min(X_j, Y_j)$ and δ_j is the indicator of the event $A_j = \{Z_j = X_j\} = \{X_j \leqslant Y_j\}$. Let

$$C(t) = \int_{-\infty}^{\infty} e^{itx} dF(x)$$

be the characteristic function of d.f. F. The problem consists in estimating of d.f. F from censored sample $\mathbb{C}^{(n)}$. In some situations it is more desirable to estimate $C(t)$ rather then F. We consider estimator for $C(t)$ in this model as Fourier–Stieltjes transform of estimator $F_n(x) = 1 - S_{1n}(x) = 1 - \exp\left(-\Lambda_n^{(1)}(x)\right)$:

$$C_n(t) = \int_{-\infty}^{\infty} e^{itx} dF_n(x), \quad t \in \mathbb{R}.$$

It follows from (39) that when $n \to \infty$

$$\sup_{x \leqslant T_n} |F_n(x) - F(x)| \stackrel{a.s.}{=} O\left(b_n^2 \left(\frac{\log n}{n}\right)^{\frac{1}{2}}\right), \tag{44}$$

where $b_n^{-1} = 1 - H(T_n)$. It also follows from (44) that when $n \to \infty$

$$1 - F_n(T_n) \stackrel{a.s.}{=} O\left(1 - F(T_n)\right), \quad F_n(-T_n) \stackrel{a.s.}{=} O\left(F(-T_n)\right). \tag{45}$$

It is obvious that $\Delta_n(\tau) \stackrel{a.s.}{\to} 0$ when $n \to \infty$ for any $\tau < \infty$, where $\Delta_n(\tau) = \sup\limits_{|t| \leqslant \tau} |C_n(t) - C(t)|$. Let us consider quantity $\Delta_n(\tau_n)$ for some special numerical sequence τ_n that tends to $+\infty$ when $n \to \infty$.

In the following theorem we prove result of uniform convergence for the empirical characteristic function.

Theorem 8. *Let $\{\tau_n, n \geqslant 1\}$ be a numerical sequence that tends to $+\infty$ slowly when $n \to \infty$. Then, $\Delta_n(\tau_n) \stackrel{a.s.}{\to} 0$ when $n \to \infty$.*

Proof. Let us choose a sequence $\{\tau_n, n \geqslant 1\}$ such that when $n \to \infty$

$$\gamma_n = \max\left\{1 - F(T_n), F(-T_n), b_n^2 \tau_n T_n \left(\frac{\log n}{n}\right)^{\frac{1}{2}}\right\} \to 0, \tag{46}$$

where $\{T_n, \quad n \geqslant 1\}$ is a sequence that satisfies condition (23). Introducing the truncated integrals

$$b_n(t) = \int_{|x| \leqslant T_n} e^{itx} dF_n(x), \quad \widetilde{b}_n(t) = \int_{|} x| \leqslant T_n e^{itx} dF(x)$$

and introducing $d_n(t) = b_n(t) - \widetilde{b}_n(t)$, we have that

$$\Delta_n(\tau_n) \leqslant \sup_{|t| \leqslant \tau_n} |d_n(t)| + \sup_{|t| \leqslant \tau_n} |b_n(t) - C_n(t)| + \sup_{|t| \leqslant \tau_n} \left|\widetilde{b}_n(t) - C(t)\right|. \tag{47}$$

Integrating by parts, we obtain

$$\begin{gathered}\sup_{|t| \leqslant \tau_n} |d_n(t)| \leqslant \sup_{|t| \leqslant \tau_n} \left|\int_{|t| \leqslant T_n} e^{itx} d\left(F_n(x) - F(x)\right)\right| \leqslant \\ \leqslant \sup_{|t| \leqslant \tau_n} \left[\left|e^{itx}\right| |F_n(x) - F(x)|_{-T_n}^{T_n}\right] + \sup_{|t| \leqslant \tau_n} \left|it \int_{|x| \leqslant T_n} e^{itx} d\left(F_n(x) - F(x)\right)\right| dx \leqslant \\ \leqslant 2(1 + 2\tau_n T_n) \sup_{|x| \leqslant T_n} |F_n(x) - F(x)|.\end{gathered} \tag{48}$$

On the other hand,

$$\sup_{|t| \leqslant \tau_n} |b_n(t) - C_n(t)| \leqslant \sup_{|t| \leqslant \tau_n} \int_{|x| > T_n} \left|e^{itx}\right| dF_n(x) \leqslant 1 - F_n(T_n) + F_n(-T_n) \tag{49}$$

and

$$\sup_{|t| \leqslant \tau_n} \left|\widetilde{b}_n(t) - C(t)\right| \leqslant \sup_{|t| \leqslant \tau_n} \int_{|x| > T_n} \left|e^{itx}\right| dF(x) \leqslant 1 - F(T_n) + F(-T_n). \tag{50}$$

Now adding (44)–(50), we have that $\Delta_n(\tau_n) \stackrel{a.s.}{=} O(\gamma_n), \quad n \to \infty$. Theorem 8 is proved. □

References

[1] A.A.Abdushukurov, Nonparametric estimation of the distribution function based on relative risk function, *Commun. Statis. Th. Meth.*, **27**(1998), no. 8, 1991–2012.
DOI: 10.1080/03610929808832205

[2] A.A.Abdushukurov, On nonparametric estimation of reliability indices by censored samples, *Theory Probab. Appl.*, **43**(1999), no. 1, 3–11.

[3] A.A.Abdushukurov, Statistics of incomplete observations. University Press, Tashkent, 2009 (in Russian).

[4] M.D.Burke, Csörgő, L.Horváth, Strong approximations of some biometric estimates under random censorship, *Z. Wahrschein. Verw. Gebiete,* **56**(1981), 87–112.

[5] M.D.Burke, S.Csörgő, L.Horváth, A correction to and improvement of "Strong approximations of some biometric estimates under random censorship", *Probab. Th. Ret. Fields.*, **79**(1988), 51–57.

[6] S.Csörgő, Strong approximation of empirical Kac processes, *Carleton Math. Lect. Note.*, **26**(1980), 71–86.

[7] S.Csörgő, L.Horvath, On random censorship from the right, *Acta. Sci. Math.*, **44**(1982), 23–34.

[8] A.Dvoretzky, J.Kiefer, J.Wolfowitz, Asymptotic minimax character of the sample distribution function and of the multinomial estimator, *Ann. Math. Statist.*, **27**(1956), 642–669.

[9] M.Kac, On denations between theoretical and empirical distributions, *Proc. Nat. Acad. Sci. USA*, **35**(1949), 252–257.

[10] D.M.Mason, Classical empirical process theory and wighted approximation, Comunicaciones del CIMAT, N.I-15-03/09-11-2015/ (PE/CIMAT).

[11] D.M.Mason, Selected Definition and Results from Modern Empirical Theory, Comunicaciones del CIMAT, N.I-17-01, 16.03.2017 (PE),.

[12] P.Massart, The tight constant in the Dvoretzky-Kiefer-Wolfowits inequality, *Ann. Probab.*, **18**(1990), no. 3, 1269–1283.

[13] W.W.Petrov, Limit theorems for sum of random variables, Moscow, Nauka, 1987 (in Russian).

[14] A.W.Skorokhod, Random processes of independent increments, Moscow, Nauka, 1964.

Об аппроксимации эмпирических процессов Каца в общей модели случайного цензурирования

Абдурахим А. Абдушукуров
Филиал Московского государственного университета в Ташкенте
Ташкент, Узбекистан
Гульназ С. Сайфуллоева
Навоийский государственный педагогический институт
Навои, Узбекистан

Аннотация. В статье рассматривается общая модель случайного цензурирования и доказываются результаты аппроксимации для эмпирических процессов Каца. Эта модель включает в себя такие важные специальные случаи, как случайное цензурирование справа и модель конкурирующих рисков. Наши результаты включают в себя теорию сильной аппроксимации, и нами построены оптимальные скорости аппроксимации. Также исследованы кумулятивные процессы риска. Эти результаты использованы для оценивания характеристической функции в модели случайного цензурирования справа.

Ключевые слова: цензурированные данные, конкурирующие риски, эмпирические оценки, оценка Каца, сильная аппроксимация, гауссовские процессы, характеристическая функция.