СБОРНИК **НАУЧНЫХ ТРУДОВ НГТУ. – 2015. – № 1(79).** – 33–44

УДК 519.24

# ИССЛЕДОВАНИЕ СВОЙСТВ НЕПАРАМЕТРИЧЕСКОЙ ОЦЕНКИ ФУНКЦИИ РАСПРЕДЕЛЕНИЯ ПО ИНТЕРВАЛЬНЫМ ДАННЫМ*

С.С. ВОЖОВ

630073, *РФ, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, аспирант кафедры теоретической и прикладной информатики. E-mail: vss920414@gmail.com*

Основные результаты прикладной математической статистики, связанные с развитием методов анализа интервальных наблюдений, получены при больших объемах выборок и малых погрешностях измерений. В связи с этим представляется актуальным исследование методами компьютерного моделирования статистических свойств оценок и критериев проверки статистических гипотез по интервальным данным ограниченного объема. Целью данной работы является исследование методами статистического моделирования свойств непараметрической оценки функции распределения по выборкам интервальных наблюдений, а также сравнение алгоритмов по времени вычисления этой оценки. Основная идея построения непараметрической оценки функции распределения по интервальным данным заключается в нахождении максимума логарифма функции правдоподобия по значениям функции распределения в граничных точках интервалов наблюдений при соблюдении условия монотонности функции распределения. В данной работе для этого используются алгоритмы Тёрнбулла и ICM. В результате сравнительного анализа алгоритмов по времени вычисления непараметрической оценки функции распределения при различных объемах выборок и длинах интервалов наблюдений показано, что время вычисления оценки по ICM-алгоритму существенно меньше, чем по алгоритму Тёрнбулла. Методами статистического моделирования проведены исследования свойств непараметрической оценки функции распределения по интервальным данным, а также получены оценки скорости сходимости непараметрической оценки к истинной функции распределения при различных длинах интервалов наблюдений. В качестве расстояния между непараметрической оценкой и истинной функцией распределения рассмотрены статистики Колмогорова и Крамера–Мизеса–Смирнова. Показано, что с ростом длин интервалов скорость сходимости падает.

**Ключевые слова:** интервальные данные, непараметрическая оценка функции распределения, алгоритм Тёрнбулла, ICM-алгоритм, миноранта, метод Монте-Карло, расстояние Колмогорова, расстояние Крамера–Мизеса–Смирнова

DOI: 10.17212/2307-6879-2015-1-33-44

* Статья получена 27 января 2015 г.

Исследование выполнено при поддержке Министерства образования и науки Российской Федерации в рамках государственного задания № 2014/138 (проект № 1689).

## ВВЕДЕНИЕ

Развитие статистических методов анализа интервальных данных является перспективным направлением исследований в области прикладной математической статистики. Природа возникновения интервальных данных многообразна. Например, вследствие погрешности измерительных приборов полученные наблюдения представляют собой интервалы фиксированной длины. В маркетинговых исследованиях наблюдения, полученные в результате анкетирования целевой группы потребителей, как правило, являются интервальными. В задачах анализа надежности и выживаемости данные типа времени жизни зачастую являются интервально-цензурированными. При этом длины интервалов могут быть бесконечными.

Для построения непараметрической оценки функции распределения по интервальным данным разработаны различные алгоритмы: алгоритм Тёрнбулла [1–3], ICM-алгоритм [4–6], гибридный ICM-EM-алгоритм [7]. В работе [8] алгоритм Тёрнбулла использовался для построения оценки функции распределения в маркетинговом исследовании спроса на биоэнергетические напитки. Исследование статистических свойств непараметрической оценки функции выживаемости по двустронне-цензурированным данным с использованием ICM-алгоритма проводилось в работах [9–11].

В данной работе проводится сравнительный анализ алгоритмов Тёрнбулла и ICM, а также исследование статистических свойств непараметрической оценки функции распределения по интервальным данным. Исследование проводится с использованием хорошо зарекомендовавшей себя методики компьютерного моделирования и исследования статистических закономерностей [12–15].

## 1. НЕПАРАМЕТРИЧЕСКАЯ ОЦЕНКА ФУНКЦИИ РАСПРЕДЕЛЕНИЯ

Пусть $X_1, X_2, ..., X_n$ – выборка независимых одинаково распределенных случайных величин из $F(x)$. Однако часто возникают ситуации, когда неизвестно точное значение $X_i$, но известно, что они попадают в некоторый интервал $(L_i, R_i)$, $i = \overline{1, n}$. Тогда исходную выборку можно представить в виде

$$\mathbb{X}_n = \{(L_1, R_1), (L_2, R_2), ..., (L_n, R_n)\}.$$

Основная идея построения непараметрической оценки функции распределения по интервальным данным заключается в том, что находится максимум логарифма функции правдоподобия:

$$\ln L\left(\mathbb{X}_n\right) = \sum_{i=1}^{n} \ln\left(F(R_i) - F(L_i)\right)$$

по значениям функции распределения в граничных точках интервалов наблюдений при соблюдении условия монотонности функции распределения. Однако решение данной оптимизационной задачи методом штрафных функций требует больших вычислительных ресурсов. Вместо этого целесообразно использование специальных алгоритмов, таких как алгоритмы Тёрнбулла и ICM.

Для оценивания функции распределения $F(x)$ рассмотрим разбиение $0<\tau_0<\tau_1<\ldots<\tau_m$, состоящее из всех неповторяющихся упорядоченных границ интервалов $L_i$ и $R_i$, $i=\overline{1,n}$. Заметим, что $m \le 2n-1$. При этом $m=2n-1$, если все левые и правые границы интервалов не совпадают друг с другом.

### 1.1. АЛГОРИТМ ТЁРНБУЛЛА

Для всех $j=\overline{1,m}$ и $i=\overline{1,n}$ определяются веса:

$$\alpha_{ij}=\begin{cases}1, \text{ если} (\tau_{j-1},\tau_j) \subseteq (L_i, R_i], \\ 0, \text{ если} (\tau_{j-1},\tau_j) \not\subset (L_i, R_i].\end{cases}$$

Алгоритм Тёрнбулла заключается в выполнении следующего итерационного процесса.

1. Положить $k=0$. Определить начальное приближение, например, следующим образом: $\hat{F}^{(0)}(\tau_j)=\sum_{l=1}^{m}\mathbf{1}\left\{\tau_l \le \tau_j\right\}$.

2. Вычислить вероятности попадания в интервал $(\tau_{j-1},\tau_j)$:

$$p_j^{(k)}=\hat{F}^{(k)}(\tau_j)-\hat{F}^{(k)}(\tau_{j-1}),\ j=\overline{1,m}.$$

3. Вычислить оценку числа отказов в момент $\tau_j$:

$$d_j=\sum_{i=1}^{n}\frac{a_{ij}\,p_j^{(k)}}{\sum_{l=1}^{m}a_{ij}\,p_l^{(k)}},\ j=\overline{1,m}.$$

4. Вычислить соответствующую оценку количества объектов, находящихся под наблюдением в момент $\tau_j$ :

$$Y_j = \sum_{l=j}^{m} d_l \ .$$

5. Используя вычисленные на шаге 2 и 3 величины $d_j$ и $Y_j$, пересчитать оценку по формуле Каплана–Майера:

$$\hat{F}^{(k+1)}(\tau_j) = 1 - \prod_{i=1}^{j} \left(1 - \frac{d_i}{Y_i}\right), \ \ j = \overline{1, m}.$$

6. Если для всех $j = \overline{1, m}$ выполняется условие останова

$$\left|\hat{F}^{(k)}(\tau_j) - \hat{F}^{(k+1)}(\tau_j)\right| \le \varepsilon,$$

то оценка Тёрнбулла найдена, иначе $k = k+1,$ перейти на шаг 1.

Полученная оценка функции распределения

$$\hat{F}(t) = \begin{cases} 0, \quad t < \tau_1, \\ \hat{F}(\tau_1), \ \ \tau_1 \le t < \tau_2, \\ \dots \\ \hat{F}(\tau_m), \ \ \tau_m \le t \end{cases}$$

является состоятельной, т. е.

$$P\left\{\sup_{t \ge 0} \left|\hat{F}(t) - F(t)\right| \to 0, n \to \infty\right\} = 1.$$

#### 1.2. ICM-АЛГОРИТМ

Идея алгоритма заключается в том, чтобы свести задачу максимизации функции правдоподобия к задаче последовательного вычисления изотонической регрессии [6]. При этом функция распределения максимальна в точках левой производной выпуклой миноранты, которая определяется по точкам $P_j = \left(G_j^{(k)}, V_j^{(k)}\right)$, при $P_0 = (0,0)$. Миноранта – функция, значение которой не больше соответствующих значений данной функции. На рис. 1 изображен пример такой функции.

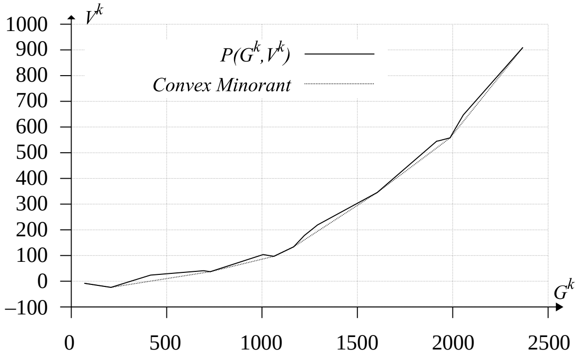


*Рис. 1.* Выпуклая миноранта

Для ICM-алгоритма для всех $j=\overline{1,m}$ и $i=\overline{1,n}$ определяются веса:

$$\alpha_{ji}=\begin{cases}1, \text{ если } \tau_j=L_i,\\ -1, \text{ если } \tau_j=R_i,\\ 0, \text{ если } \tau_j\not\subset(L_i,R_i].\end{cases}$$

Итерационный процесс ICM-алгоритма для произвольного шага алгоритма $k\geq 1$.

1. Для всех $j=\overline{1,m}$ найти точки $\left(G_j^{(k)},V_j^{(k)}\right)$ в соответствии со следующими выражениями:

$$G_j^{(k)}=G_{j-1}^{(k)}+\sum_{i=1}^{n}\frac{1}{\left|\alpha_{ji}\right|\left(\hat{F}^{(k)}(\tau_j)-\hat{F}^{(k)}(\tau_{j-1})\right)^2},$$

$$W_j^{(k)}=W_{j-1}^{(k)}+\sum_{i=1}^{n}\frac{1}{\alpha_{ji}\left(\hat{F}^{(k)}(L_i)-\hat{F}^{(k)}(R_i)\right)},$$

$$D_j^{(k)}=D_{j-1}^{(k)}+\sum_{i=1}^{n}\frac{\hat{F}^{(k)}(\tau_j)}{\left|\alpha_{ji}\right|\left(\hat{F}^{(k)}(\tau_j)-\hat{F}^{(k)}(\tau_{j-1})\right)^2},$$

$$V_j^{(k)}=W_j^{(k)}+D_j^{(k)}.$$

2. Установить $l = 0$.

3. Оценка функции распределения равна левой производной выпуклой миноранты:

$$\hat{F}^{(k)}(\tau_j) = \min_{l+1 \le s \le m} \left( \frac{V_s^{(k)} - V_l^{(k)}}{G_s^{(k)} - G_l^{(k)}} \right),$$

где $j = \overline{l+1, s}$, $s$ – индекс угловой точки выпуклой миноранты.

4. Изменить $l = s$. Если $l < m,$ переходим в пункт 3.

Алгоритм повторяем до тех пор, пока для всех $j = \overline{1, m}$ не выполнится условие

$$\left| \hat{F}^{(k)}(\tau_j) - \hat{F}^{(k-1)}(\tau_j) \right| \le 10^{-7}.$$

Полученная оценка функции распределения по ICM-алгоритму, так же как и оценка функции распределения, найденная по алгоритму Тёрнбулла, является непараметрической и равномерно строго состоятельной.

## 2. СРАВНИТЕЛЬНЫЙ АНАЛИЗ АЛГОРИТМОВ

Алгоритм Тёрнбулла и ICM-алгоритм вычисляют одинаковую оценку функции распределения в пределах заданной точности, однако время (*Time*) вычисления этой оценки у алгоритма ICM значительно меньше, что видно на рис. 2.

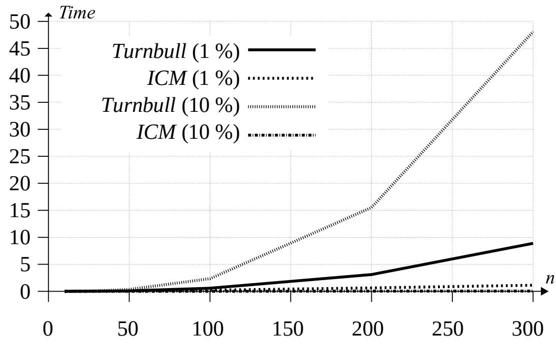


*Рис. 2.* Скорость вычисления оценки интервальными алгоритмами

С увеличением длин интервалов время вычисления оценки функции распределения алгоритмом Тёрнбулла увеличивается. Для ICM-алгоритма наоборот: чем больше длины интервалов наблюдений, тем за меньшее время вычисляется оценка функции распределения.

## 3. ИССЛЕДОВАНИЕ СВОЙСТВ НЕПАРАМЕТРИЧЕСКОЙ ОЦЕНКИ ФУНКЦИИ РАСПРЕДЕЛЕНИЯ

Проведем исследование методом Монте-Карло статистических свойств непараметрической оценки в зависимости от длины интервалов наблюдений. Интервальную выборку, в которой каждый элемент представляет собой интервал длины $\Delta$, будем моделировать в соответствии со следующим алгоритмом.

1. Методом обратной функции сгенерировать выборку $X_1, X_2, ..., X_n$ из распределения $F(x)$.

2. Вычислить границы интервальных наблюдений:

$$L_i = \begin{cases} X_i - \dfrac{\Delta}{2}, \text{ если } X_i \ge \dfrac{\Delta}{2}, \\ 0, \text{ иначе;} \quad R_i = X_i + \dfrac{\Delta}{2}, \quad i = \overline{1, n}, \end{cases}$$

и сформировать интервальную выборку $(L_1, R_1), ..., (L_n, R_n)$.

В качестве расстояний между непараметрической оценкой и соответствующей функцией распределения рассмотрим статистику Колмогорова вида

$$D_n = \sup_{0<t<\tau_m} \left| \hat{F}(x) - F(x) \right|$$

и статистику Крамера–Мизеса–Смирнова

$$\omega^2 = \int_0^{\tau_m} (\hat{F}(x) - F(x))^2 dF(x) .$$

На рис. 3 представлены графики полученных зависимостей значения $D_n$ от $n$, усредненного по 10 000 выборкам, при различной длине интервалов наблюдений, равной 1 %, 3 %, 5 % или 10 % от длины интервала $(0, F^{-1}(0.99))$, где $F^{-1}(\cdot)$ – функция, обратная к функции распределения отказов. В качестве функции распределения $F(x)$ взято распределение Вейбулла с параметром масштаба $\theta_1 = 2$ и параметром формы $\theta_2 = 2$.

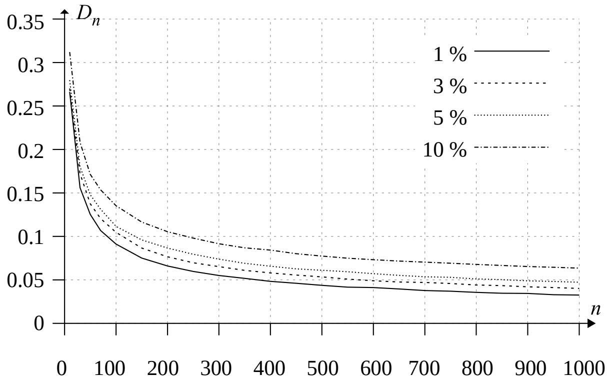


*Рис. 3.* Зависимость расстояния Колмогорова $D_n$ от объема выборки при различных $\Delta$

Как видно на рис. 3, с ростом объема выборки и уменьшением длины интервалов наблюдений расстояние $D_n$ от непараметрической оценки $\hat{F}(x)$ до истинной функции распределения $F(x)$ уменьшается. Аналогичная зависимость от $n$ и $\Delta$ наблюдалась и для расстояния Крамера–Мизеса–Смирнова.

Аппроксимации зависимости расстояний Колмогорова и Крамера–Мизеса–Смирнова от объема выборки в виде степенных функций с параметрами, оцененными по методу наименьших квадратов, и коэффициентом детерминации, близким к единице, имеют следующий вид:

- при $\Delta = 0.01F^{-1}(0.99)$: $D_n = 0.76n^{-0.46}$, $\omega^2 = 0.16n^{-0.98}$;
- при $\Delta = 0.03F^{-1}(0.99)$: $D_n = 0.70n^{-0.42}$, $\omega^2 = 0.14n^{-0.93}$;
- при $\Delta = 0.05F^{-1}(0.99)$: $D_n = 0.68n^{-0.39}$, $\omega^2 = 0.14n^{-0.91}$;
- при $\Delta = 0.1F^{-1}(0.99)$: $D_n = 0.70n^{-0.35}$, $\omega^2 = 0.15n^{-0.86}$.

Как видим, скорость сходимости непараметрической оценки $\hat{F}(x)$ к истинной функции распределения $F(x)$ с ростом длины интервалов наблюдений уменьшается. При этом необходимо отметить, что при малой длине интервальных наблюдений скорость сходимости близка к скорости сходимости эмпирической функции распределения $F_n(x)$ по полным выборкам к истинной функции распределения $F(x)$: $O(n^{-0.5})$ для расстояния $D_n$ и $O(n^{-1})$ для расстояния $\omega^2$.

## ЗАКЛЮЧЕНИЕ

В результате сравнительного анализа алгоритмов Тёрнбулла и ICM показано, что точность вычисления оценки функции распределения у алгоритмов одинаковая, однако время вычисления оценки по ICM-алгоритму существенно меньше, чем у алгоритма Тёрнбулла. Алгоритмы Тёрнбулла и ICM для вычисления непараметрической оценки функции распределения по интервальным данным реализованы на базе программной системы LiTiS [16].

Методами статистического моделирования проведено исследование свойств непараметрической оценки функции распределения по интервальным данным, получены оценки скорости сходимости непараметрической оценки к истинной функции распределения при различных длинах интервалов наблюдений. Показано, что с ростом длин интервалов скорость сходимости падает.

## СПИСОК ЛИТЕРАТУРЫ

1. *Turnbull B.W.* Nonparametric estimation of a survivorship function with doubly-censored data // Journal of the American Statistical Association. – 1974. – Vol. 69, iss. 345. – P. 169–173. – doi: 10.1080/01621459.1974.10480146.

2. *Turnbull B.W.* The empirical distribution function with arbitrarily grouped, censored and truncated data // Journal of the Royal Statistical Society. Series B. – 1976. – Vol. 38, N 3. – P. 290–295. – doi: 10.2307/2984980.

3. *Efron B.* The two sample problem with censored data // Procceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probality. – New York: Prentice Hall, 1967. – Vol. 4. – P. 831–853.

4. *Groeneboom P.* Asymptotics for interval censored observations: technical report 87-18 / University of Amsterdam, Department of Mathematics. – Amsterdam, 1987. – 69 p.

5. *Groeneboom P.* Nonparametric maximum likelihood estimators for interval censoring and deconvolution: technical report no. 378 / Stanford University, Department of Statistics. – Stanford, California, 1991. – 87 p.

6. *Groeneboom P., Wellner J.A.* Information bounds and nonparametric maximum likelihood estimation. – Basel: Birkhauser Verlag, 1992. – 126 p. – doi: 10.1007/978-3-0348-8621-5.

7. *Wellner J.A., Zhan Y.* A hybrid algorithm for computation of the nonparametric maximum likelihood estimator from censored data // Journal of the American Statistical Association. – 1997. – Vol. 92, iss. 439. – P. 945–959.

8. *Зенкова Ж.Н., Краковецкая И.В.* Непараметрическая оценка Тернбулла для интервально-цензурированных данных в маркетинговом исследовании спроса на биоэнергетические напитки // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2013. – № 3 (24). – С. 64–69.

9. *Chang M.N.* Weak convergence of a self-consistent estimator of the survival function with doubly censored data // The Annals of Statistics. – 1990. – Vol. 18, N 1. – P. 391–404. – doi: 10.1214/aos/1176347506.

10. *Chang M.N., Yang G.L.* Strong Consistency of a nonparametric estimator of the survival function with doubly censored data // The Annals of Statistics. – 1987. – Vol. 15, N 4. – P. 1536–1547. – doi: 10.1214/aos/1176350608.

11. *Samuelsen S.O.* Asymptotic theory of nonparametric estimator from doubly-censored data // Scandinavian Journal of Statistics. – 1989. – Vol. 16. – P. 1–21.

12. *Лемешко Б.Ю., Чимитова Е.В., Ведерникова М.А.* Модифицированные критерии согласия Колмогорова, Крамера-Мизеса-Смирнова и Андерсона-Дарлинга для случайно цензурированных выборок. Ч. 2 // Научный вестник НГТУ. – 2013. – № 1 (50). – С. 3–16.

13. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б.Ю. Лемешко, С.Б. Лемешко, С.Н. Постовалов, Е.В. Чимитова. – Новосибирск: Изд-во НГТУ, 2011. – 888 с. – (Серия «Монографии НГТУ»).

14. Компьютерное моделирование и исследование вероятностных закономерностей / Б.Ю. Лемешко, А.А. Горбунова, С.Б. Лемешко, С.Н. Постовалов, А.П. Рогожников, Е.В. Чимитова // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2013. – № 1 (22). – С. 74–85.

15. *Демин В.А., Чимитова Е.В.* Выбор оптимального параметра сглаживания для непараметрической оценки регрессионной модели надежности // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2013. – № 1 (22). – С. 59–65.

16. Система статистического анализа данных типа времени жизни «LiTiS 1.2»: [свидетельство о государственной регистрации программы для ЭВМ] № 2015610901 / Е.В. Чимитова, А.В. Румянцев, М.А. Семенова, Н.С. Галанова, В.А. Демин; правообладатель: Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Новосибирский государственный технический университет». – Заявка № 2014661905; заявл. 24.11.2014; опубл. 20.01.2015. – 1 с.

***Вожов Станислав Сергеевич***, аспирант кафедры теоретической и прикладной информатики Новосибирского государственного технического университета. Основное направление научных исследований – статистические методы анализа интервальных данных. E-mail: vss920414@gmail.com

## Investigation of the properties of nonparametric estimate for distribution function with interval data*

**S.S. Vozhov**

*Novosibirsk State Technical University, 20 pr. Karla Marksa, Novosibirsk, 630073, Russian Federation, post-graduate student of the department of Theoretical and Applied Informatics. E-mail: vss920414@gmail.com*

The evolution of statistical methods for the analysis of interval data is a promising area of research in the field of applied mathematical statistics. The nature of interval data is various. For example, due to instrument accuracy obtained observations are intervals of fixed length. In marketing research, observations obtained from the survey of the target group of consumers are usually interval. In reliability and survival analysis problems, lifetime data are often interval-censored. Interval lengths may be infinite. Various algorithms are developed to construct a non-parametric estimate of the distribution function on interval data: the Turnbull algorithm, the ICM algorithm. The concept of constructing a non-parametric estimate of the distribution function on interval data is to find a maximum of logarithm of likelihood function by the values of the distribution function at the observation intervals boundary points subject to the condition of the distribution function monotonicity. The Turnbull algorithm and the ICM algorithm have been compared in terms of the time, spent for the calculation of the nonparametric estimate of the distribution function. It has been shown that the time of calculation by the ICM algorithm is much less than by the Turnbull algorithm. Estimates of a convergence rate of a non-parametric estimate to the real distribution function for different lengths of observable intervals are obtained as a result of the investigation of a non-parametric estimate of the distribution function on interval data properties carried out by statistical modeling methods. The Kolmogorov and Cramer-von Mises-Smirnov statistics have been considered as the distance between non-parametric estimates and the true distribution function.It has been shown that the convergence rate decreases when the interval length increases.

**Keywords:** interval data, nonparametric estimate of the distribution function, Turnbull algorithm, ICM algorithm, minorant, Monte Carlo method, Kolmogorov statistic, Cramer–von Mises–Smirnov statistic

DOI: 10.17212/2307-6879-2015-1-33-44

## REFERENCES

1. Turnbull B.W. Nonparametric estimation of a survivorship function with doubly-censored data. *Journal of the American Statistical Association*, 1974, vol. 69, iss. 345, pp. 169–173. doi: 10.1080/01621459.1974.10480146

2. Turnbull B.W. The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society. Series B*, 1976, vol. 38, no. 3, pp. 290–295. doi: 10.2307/2984980

3. Efron B. The two sample problem with censored data. *Procceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probality*, New York, Prentice Hall, 1967, vol. 4, pp. 831–853.

4. Groeneboom P. *Asymptotics for interval censored observations*. Technical report 87-18. University of Amsterdam, Department of Mathematics, 1987. 69 p.

5. Groeneboom P. *Nonparametric maximum likelihood estimators for interval censoring and deconvolution*. Technical report no. 378. Stanford University, Department of Statistics, 1991. 87 p.

* *Received on 27 January 2015.*

*This research has been supported by the Russian Ministry of Education and Science as part of the state task No 2014/138 (project No 1689).*

6. Groeneboom P., Wellner J.A. *Information bounds and nonparametric maximum likelihood estimation*. Basel, Birkhauser Verlag, 1992. 126 p. doi: 10.1007/978-3-0348-8621-5

7. Wellner J.A., Zhan Y. A hybrid algorithm for computation of the nonparametric maximum likelihood estimator from censored data. *Journal of the American Statistical Association*, 1997, vol. 92, iss. 439, pp. 945–959.

8. Zenkova Zh.N., Krakovetskaya I.V. Neparametricheskaya otsenka Ternbulla dlya interval'no-tsenzurirovannykh dannykh v marketingovom issledovanii sprosa na bioenergeticheskie napitki [Nonparametric Turnbull estimator for interval-censored data in the marketing research of the demand of bio-energy drinks]. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk state university journal of control and computer science*, 2013, no. 3 (24), pp. 64–69.

9. Chang M.N. Weak convergence of a self-consistent estimator of the survival function with doubly censored data. *The Annals of Statistics*, 1990, vol. 18, no. 1, pp. 391–404. doi: 10.1214/aos/1176347506

10. Chang M.N., Yang G.L. Strong consistency of a nonparametric estimator of the survival function with doubly censored data. *The Annals of Statistics*, 1987, vol. 15, no. 4, pp. 1536–1547. doi: 10.1214/aos/1176350608

11. Samuelsen S.O. Asymptotic theory for nonparametric estimators from doubly censored data. *Scandinavian Journal of Statistics*, 1989, vol. 16, no. 1, pp. 1–21.

12. Lemeshko B.Yu., Chimitova E.V., Vedernikova M.A. Modificirovannye kriterii soglasija Kolmogorova, Kramera-Mizesa-Smirnova i Andersona-Darlinga dlja sluchajno cenzurirovannyh vyborok. Ch. 2 [Modified goodness-of-fit tests of Kolmogorov, Cramer-von Mises-Smirnov and Anderson-Darling for randomly censored samples. Pt. 2]. *Nauchnyi vestnik Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta – Science Bulletin of the Novosibirsk State Technical University*, 2013, no. 1 (50), pp. 3–16.

13. Lemeshko S.B., Lemeshko B.Yu., Postovalov S.N., Chimitova E.V. *Statisticheskii analiz dannykh, modelirovanie i issledovanie veroyatnostnykh zakonomernostei. Komp'yuternyi podkhod* [Statistical data analysis, simulation and study of probability regularities. Computer approach]. Novosibirsk, NSTU Publ., 2011. 888 p.

14. Lemeshko B.Yu., Gorbunova A.A., Lemeshko S.B., Postovalov S.N., Rogozhnikov A.P., Chimitova E.V. Komp'yuternoe modelirovanie i issledovanie veroyatnostnykh zakonomernostei [Computer simulations and research of probabilistic regularities]. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk state university journal of control and computer science*, 2013, no. 1 (22), pp. 74–85.

15. Demin V.A., Chimitova E.V. Vybor optimal'nogo parametra sglazhivaniya dlya neparametricheskoi otsenki regressionnoi modeli nadezhnosti [Choice of optimal smoothing parameter for nonparametric estimation of regression reliability model]. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk state university journal of control and computer science*, 2013, no. 1 (22), pp. 59–65.

16. Chimitova E.V., Rumyantsev A.V., Semenova M.A., Galanova N.S., Demin V.A. Sistema statisticheskogo analiza dannykh tipa vremeni zhizni «LiTiS 1.2» [Software system of statistical lifetime data analysis «LiTiS 1.2»]. The Certificate on official registration of the computer program. No. 2015610901, 2015. (In Russian, unpublished).