

УДК 519.248

Метод вычисления точных распределений статистик типа Колмогорова-Смирнова в случае нарушения однородности и независимости анализируемых выборок

Тянникова Н. Д.^{1,*}, Тимонин В. И.¹

¹МГТУ им. Н.Э. Баумана, Москва, Россия

В задачах непараметрической статистики часто возникает проблема сравнения нескольких выборок, о которых заранее известно, что они не принадлежат одной генеральной совокупности. Самой распространенной моделью, используемой для установления зависимости между теоретическими функциями распределения различных выборок, является модель Кокса. Кроме того, даже при проверке однородности нескольких выборок, эксперименты, необходимые для их получения, настолько сложны, что полученные выборки являются зависимыми. Во всех этих задачах необходима разработка новых непараметрических критериев проверки предполагаемых зависимостей. В силу того, что объемы выборок всегда малы, особую важность имеет знание точных распределений используемых статистик. В работе предлагается общий метод табулирования точных распределений (для конечных объемов выборок) для широкого класса статистик типа Колмогорова-Смирнова. Надвух примерах показано применение этого метода.

Ключевые слова: статистики типа Колмогорова-Смирнова, непараметрическая статистика, оценки Каплана-Мейера, степенные зависимости Лемана

Введение

В ряде задач теории надежности необходимо проверять непараметрические гипотезы о связях функции распределения нескольких выборок в случае, когда не выполняются классические условия независимости и одинаковости распределенности элементов выборок. Например, в работах [1,2,3] проверялась гипотеза о степенной зависимости функций распределения нескольких независимых выборок. В [4,5,6] проверялась однородность двух выборок, элементы которых являются зависимыми случайными величинами. Для вычисления точных распределений (для конечного объема выборок) статистик, предложенных в этих работах, были разработаны численные алгоритмы, позволяющие табулировать точные распределения для больших объемов выборок.

В настоящей работе предложен общий алгоритм вычисления распределений статистик типа Колмогорова-Смирнова, частным случаем которого являются алгоритмы,

описанные в [1-9]. В основе алгоритма лежит специальная модель случайного блуждания частицы по ячейкам, на множестве которых определена некоторая функция $h(x_1, x_2, \dots, x_k)$.

1. Модель случайного блуждания

Обозначим, $\Pi_{n_1, n_2, \dots, n_k}$ — k -мерное множество ячеек $a_{i_1, i_2, \dots, i_k}, 0 \leq i_j \leq n_j, j = \overline{1, k}$ (k -мерный параллелепипед). Пусть $A \subset \Pi_{n_1, n_2, \dots, n_k}$ — подмножество параллелепипеда, для которого выполняются следующие условия:

1. $a_{0,0,\dots,0}, a_{n_1, n_2, \dots, n_k} \in A, n = n_1 + n_2 + \dots + n_k$;
2. $\forall a_{i_1, i_2, \dots, i_k} \neq a_{0,0,\dots,0} \exists i_j : a_{i_1, i_2, \dots, i_{j-1}, i_k} \in A$.

Рассмотрим следующую модель случайного блуждания. Частица на первом шаге выходит из ячейки $a_{0,0,\dots,0}$ и на n -ом шаге она заканчивает блуждание в ячейке a_{n_1, n_2, \dots, n_k} . В дальнейшем для упрощения записи ячейке a_{i_1, i_2, \dots, i_k} будем ставить в соответствие вектор $\vec{i}_q = (i_1, i_2, \dots, i_k), i_1 + i_2 + \dots + i_k = q, 0 \leq q \leq n, 0 \leq i_j \leq n_j, j = \overline{1, k}$, состоящий из её индексов. Пусть $\vec{\Delta s}$ — вектор размерности k , состоящий из $(k-1)$ нуля и одной единицы на s -ом месте. На q -ом шаге частица переходит из ячейки \vec{i}_{q-1} в ячейку $\vec{i}_q = \vec{i}_{q-1} + \vec{\Delta s}_q, \vec{i}_{q-1}, \vec{i}_q \in A$.

Пусть $\omega = \omega_{i_0, i_1, \dots, i_n} = (\vec{i}_0, \vec{i}_1, \dots, \vec{i}_n)$ — возможная траектория блуждания частицы из начальной ячейки \vec{i}_0 в конечную $\vec{i}_n, \vec{i}_q \in A$. Обозначим множество всех таких траекторий через $\Omega(A)$.

Предположим, что вероятности $P(\omega_{\vec{i}_0, \vec{i}_1, \dots, \vec{i}_n})$ представимы в следующем виде

$$P(\omega_{\vec{i}_0, \vec{i}_1, \dots, \vec{i}_n}) = \prod_{q=1}^n \mu(\vec{i}_{q-1}, \vec{i}_q) = \prod_{q=1}^n \mu_{q-1, q}. \quad (1)$$

Вероятности (1) должны удовлетворять следующим условиям: их сомножители $\mu_{q-1, q}$ не зависит от траектории попадания частицы в ячейку $\vec{i}_{q-1}, \sum_{\omega} \prod_{q=1}^n \mu_{q-1, q} = 1$.

Пусть $h(i_1, i_2, \dots, i_k)$ - произвольная функция на множестве ячеек $(i_1, i_2, \dots, i_k) \in A$.

Теорема 1. Вероятность $P(\omega_{\vec{i}_0, \vec{i}_1, \dots, \vec{i}_n} | \max_{0 \leq q \leq n} h(\vec{i}_q) < x)$ равна величине $\pi_{\vec{i}_n}(x)$, которую можно получить повторным применением соотношения

$$\pi(\vec{i}_q) = \left(\sum_{1 \leq s \leq k, \vec{i}_q - \vec{\Delta s} \in A} (\mu_s(\vec{i}_q - \vec{\Delta s}, \vec{i}_q) \pi(\vec{i}_q - \vec{\Delta s})) \right) \chi(h(\vec{i}_q) < x), \quad (2)$$

где $\chi\left(h\left(\vec{i}_q\right) < x\right)$ – индикатор события $h\left(\vec{i}_q\right) < x$.

Доказательство: Обозначим через $\lambda\left(\vec{i}_q\right)$ множество «частичных» траекторий $\omega_{\vec{i}_0, \vec{i}_1, \dots, \vec{i}_q}$, оканчивающихся в ячейке \vec{i}_q . Пусть $\pi\left(\vec{i}_q\right) = \sum_{\lambda\left(\vec{i}_q\right)} \prod_{j=1}^q \mu_j\left(\vec{i}_{j-1}, \vec{i}_j\right)$ – сумма по всем таким «частичным» траекториям первых q сомножителей. Тогда соотношение (2) следует из того, что в ячейку \vec{i}_q за один скачок можно попасть только из ячеек $\vec{i}_q - \vec{\Delta s}, s = \overline{1, k}$. Множитель $\chi\left(h\left(\vec{i}_q\right) < x\right)$ обеспечивает обращение в нуль вероятностей тех траекторий, на которых значения $h\left(\vec{i}_q\right)$ превышает x . ▶

Главная задача при применении метода состоит в получении вероятностей (1). Ниже приведены примеры применения соотношения (2) для некоторых частных случаев.

2. Частные случаи

Пример 1. В [2] рассматривалась следующая задача. Пусть имеется k независимых выборок $\vec{\xi}_i = \left(\xi_1^i, \dots, \xi_{n_i}^i\right)$, $\xi_j^i \sim F_i(t)$, $i = \overline{1, k}; j = \overline{1, n_i}; n = \sum_{i=1}^k n_i$, где n_i – объем i -ой выборки. По этим данным проверялась гипотеза

$$H_0 : F_1^{r_1}(t) = \dots = F_k^{r_k}(t), \quad (3)$$

где $r_i \geq 1$ – заданные числа.

Для проверки гипотезы (3) предлагалась статистика типа Кифера-Гихмана:

$$T^2 = \max_t \frac{\sum_{i=1}^k n_i \left(\widehat{F}_i^{r_i} - \bar{F}\right)^2 + \bar{F} \cdot \left(\Phi \sqrt{k-1} - \Phi_1\right)}{\Gamma^2}. \quad (4)$$

Здесь $\widehat{F}_i(t)$ – эмпирическая функция распределения i -ой выборки; $\bar{F} = \sum_{i=1}^k \rho_i \cdot \widehat{F}_i^{r_i}(t)$;

$\rho_i = n_i/n$; $\Phi_1 = \sum_{i=1}^k r_i^2 (1 - \rho_i) \bar{F}^{1-r_{0i}} (1 - \bar{F}^{r_{0i}})$; $\Gamma = \bar{F} + \Phi / \sqrt{k-1}$; $r_{0i} = 1/r_i$;

$$\Phi = \left[\left(\sum_{i=1}^k r_i^2 \rho_i \bar{F}^{1-r_{0i}} (1 - \bar{F}^{r_{0i}}) \right)^2 + \sum_{i=1}^k r_i^4 \bar{F}^{2(1-r_{0i})} (1 - \bar{F}^{r_{0i}})^2 (1 - 2\rho_i) \right]^{1/2}.$$

Метод вычисления точных распределений статистики, предложенный в [2], является частным случаем алгоритма, предложенного в настоящей статье. Для этой задачи $A = \Pi_{n_1, n_2, \dots, n_k}$. Функция $h(i_1, \dots, i_k) = h(\vec{i}_q)$ определяется следующим образом.

$$\omega'_{i_1, \dots, i_k} = \sum_{j=1}^k n_j \left[\left(\frac{i_1}{n_1} \right)^{r_1} + \dots + \left(\frac{i_k}{n_k} \right)^{r_k} - z_{i_1, \dots, i_k} \right]^2,$$

где $z_{i_1, \dots, i_k} = \frac{1}{n} \sum_{j=1}^k n_j \left(\frac{i_j}{n_j} \right)^{r_j}$.

Обозначим

$$T_{i_1, i_2, \dots, i_k} = \left(\sum_{j=1}^k r_j^2 \rho_j \left(\sum_{v=1}^k \rho_v \left(\frac{i_v}{n_v} \right)^{r_v} \right)^{1-r_{0j}} \left(1 - \left(\sum_{v=1}^k \rho_v \left(\frac{i_v}{n_v} \right)^{r_v} \right)^{r_{0j}} \right) \right)^2 + \sum_{v=1}^k r_v^4 \left(1 - \left(\sum_{v=1}^k \rho_v \left(\frac{i_v}{n_v} \right)^{r_v} \right)^{r_{0j}} \right)^2 (1 - 2\rho_v) \left(\sum_{v=1}^k \rho_v \left(\frac{i_v}{n_v} \right)^{r_v} \right)^{2-2r_{0j}};$$

$$A_{i_1, i_2, \dots, i_k} = \omega'_{i_1, \dots, i_k} + \left(\sum_{v=1}^k \rho_v \left(\frac{i_v}{n_v} \right)^{r_v} \right) \times \left(\left(\sum_{j=1}^k r_j^2 (1 - \rho_j) \left(\frac{i_j}{n_j} \right)^{r_j-1} \left(1 - \frac{i_j}{n_j} \right) \right) - \sqrt{(k-1) \cdot T_{i_1, i_2, \dots, i_k}} \right);$$

$$B_{i_1, i_2, \dots, i_k} = \left(\sum_{j=1}^k \rho_j \left(\frac{i_j}{n_j} \right)^{r_j} + \sqrt{\frac{T_{i_1, i_2, \dots, i_k}}{k-1}} \right)^2.$$

Тогда

$$h(i_1, i_2, \dots, i_k) = A_{i_1, i_2, \dots, i_k} / B_{i_1, i_2, \dots, i_k}$$

Множители $\mu_s(\vec{i}_q - \Delta \vec{s}, \vec{i}_q)$ имеют вид

$$\mu_s(\vec{i}_q - \Delta \vec{s}, \vec{i}_q) = \frac{i_s r_{0s}}{\sum_{j=1}^k i_j r_{0j}}, s = 1, k.$$

В табл. 1 для данной модели в случае $k = 3$ рассчитаны вероятности $V_1 = P\left(\omega_{\vec{i}_0, \vec{i}_1, \dots, \vec{i}_n} | \max_{0 \leq q \leq n} h(\vec{i}_q) < x\right)$. На рис.1 показаны графики изменения этих вероятностей в зависимости от x и r_1, r_2, r_3 .

Таблица 1. Значения вероятностей V_1

$n_1 = n_2 = n_3$	$x = 2$		$x = 3$	
	$r_1 = 1, r_2 = 2, r_3 = 3$	$r_1 = 1, r_2 = 2.5, r_3 = 3$	$r_1 = 1, r_2 = 2, r_3 = 3$	$r_1 = 1, r_2 = 2.5, r_3 = 3$
50	0,90697	0,90896	0,98468	0,98561
100	0,89624	0,89898	0,98204	0,98260
150	0,89362	0,89477	0,98079	0,98140
200	0,89146	0,89224	0,98024	0,98070
250	0,88968	0,89058	0,97986	0,98020
300	0,88855	0,88945	0,97952	0,97987
400	0,88682	0,88752	0,97907	0,97939
500	0,88569	0,88651	0,97879	0,97908
∞	0,87857	0,87857	0,97653	0,97653

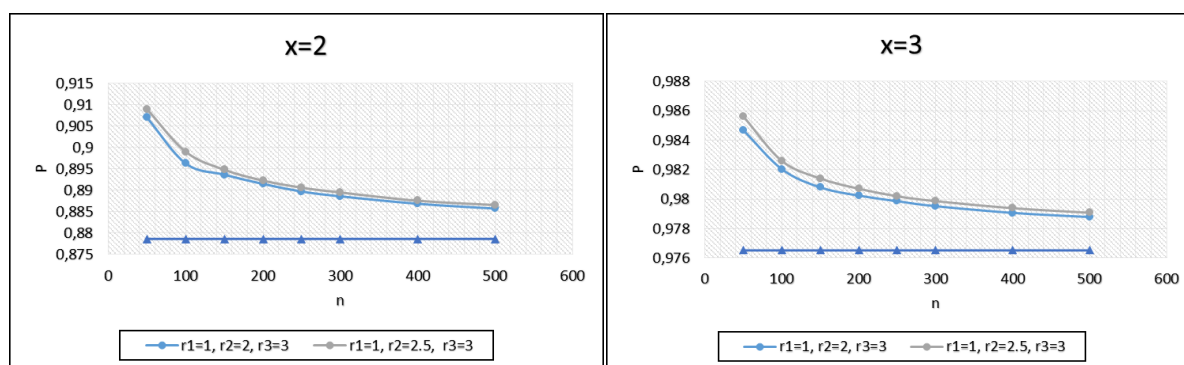


Рис.1. Графики вероятностей V_1 в зависимости от $n = n_1 = n_2 = n_3$

Пример 2. В работе [5] рассматривалась следующая задача. Имеется N одинаковых систем, каждая из которых состоит из m идентичных параллельно соединенных элементов. Все элементы испытываются до отказа и функция надежности элементов оценивается двумя способами: по полной выборке Q из отказов всех элементов ($\hat{P}_q(t)$) и по прогрессивно цензурированной выборке Θ , образованной первыми порядковыми статистиками наработок до отказа каждой системы ($\hat{P}_\theta(t)$). Они имеют вид

$$\hat{P}_\theta(t) = \begin{cases} 1, d_1(t) = 0, \\ \prod_{i=1}^{d_1(t)} \left(1 - \frac{1}{m(N-i+1)} \right), 1 \leq d_1(t) \leq (N-1), \\ 0, d_1(t) = N, \end{cases} \quad (5)$$

$$\hat{P}_q(t) = 1 - \frac{d_2(t)}{mN},$$

где $d_1(t), d_2(t)$ — количество элементов выборок Θ и Q соответственно, меньших t .

Для проверки того, что обе оценки оценивают одну и ту же функцию надежности $P_0(t)$, в [5] рассматривалась статистика

$$T_{mN} = m\sqrt{N} \max_t \frac{(\hat{P}_q(t))^{m-1}}{1 - m \cdot (1 - \hat{P}_q(t)) \cdot (\hat{P}_q(t))^{m-1}} \cdot |\hat{P}_\theta(t) - \hat{P}_q(t)|. \quad (6)$$

Метод вычисления точных распределений статистики (6), предложенный в [5], может быть получен из общего алгоритма (2), предложенного в настоящей статье. Заметим, что в оригинальной статье рассматривался алгоритм, в котором блуждание осуществлялось в обратном порядке.

В данном примере размерность A равна двум ($k = 2$). Множество A имеет вид

$$A = \{\vec{i} = (i_1, i_2), i_1 = \overline{0, N}; i_2 = \overline{0, i_1(m-1)}\}.$$

Функция $h(\vec{i}_q)$ равна

$$h(i_1, i_2) = h(\vec{i}_q) = m\sqrt{N} \cdot \frac{mN(mN - i_1 - i_2)^{m-1}}{(mN)^m - m(mN - i_1 - i_2)^{m-1}(i_1 + i_2)} \cdot \left| \frac{mN - i_1 - i_2}{mN} - \prod_{l=1}^{i_1} \left(1 - \frac{1}{m(N-l+1)} \right) \right|$$

Точные распределения статистики T_{mN} вычисляются согласно алгоритму (2), где

$$\mu_s(\vec{i}_q - \Delta s, \vec{i}_q) = \frac{mi_s(-1)^{s-1} + (s-1)((m-1)q+1)}{q}, s = 1, 2.$$

В табл.2 для данного примера рассчитаны вероятности $V_2 = P\left(\omega_{\vec{i}_1, \vec{i}_2, \dots, \vec{i}_{mN}} \mid \max_{0 \leq q \leq mN} h(\vec{i}_q) < x\right)$

для $m = 3, 4$; $x = 1, 22; 1, 36$. На рис.2 показаны графики изменения этих вероятностей для тех же значений m, x .

Таблица 2. Значения вероятностей V_2

n	$x = 1.22$		$x = 1.36$	
	$m = 3$	$m = 4$	$m = 3$	$m = 4$
100	0,88617	0,85564	0,93747	0,91323
500	0,90421	0,90245	0,95312	0,95199
1000	0,90271	0,90227	0,95270	0,95239
2000	0,90182	0,90151	0,95232	0,95212
5000	0,90059	0,90151	0,95178	0,95166
10000	0,89989	0,89976	0,95145	0,95138
20000	0,89938	0,89931	0,95120	0,95115
30000	0,89915	0,89909	0,95107	0,95104
40000	0,89901	0,89897	0,95100	0,95097
50000	0,89892	0,89888	0,95095	0,95093
60000	0,89885	0,89881	0,950952	0,95089
∞	0,89810	0,89810	0,95051	0,95051

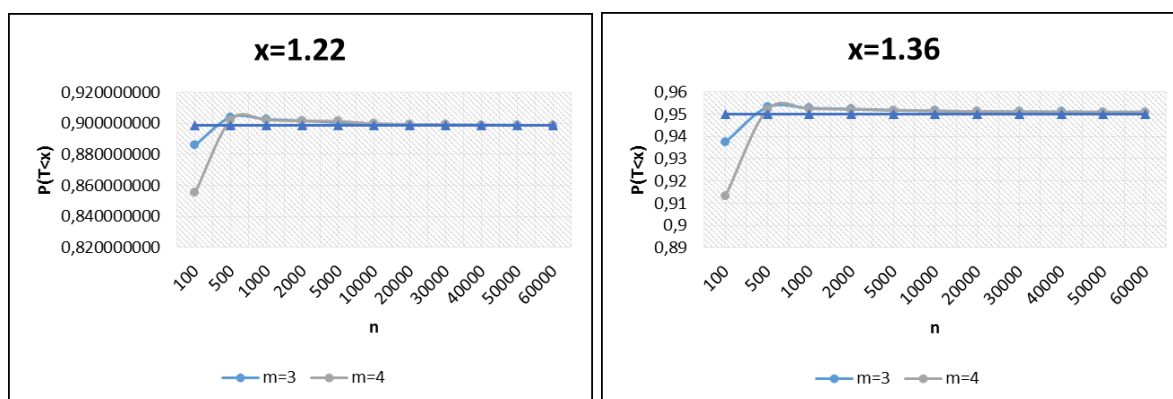


Рис.2. Графики вероятностей V_2 в зависимости от n

Предложенный метод применим и при расчётах точных распределений статистик типа Реньи [10,11,12,13], применяемых для проверок аналогичных гипотез при цензурированных данных.

Заключение

В работе разработан метод вычисления распределений статистик типа Колмогорова-Смирнова, когда анализируемые выборки не обязательно принадлежат одной совокупности. Метод основан на модели случайного блуждания специально вида по элементам

k -мерного параллелепипеда, причем условные вероятности перехода не зависят от предыстории блуждания. Распределение статистик Колмогорова-Смирнова определяется через вероятности невыхода траекторий блуждания из подмножества параллелепипеда.

Список литературы

1. Тимонин В.И. О предельном распределении статистики одного непараметрического критерия // Теория вероятностей и её применение. 1987. Т. 32, № 4. С. 790-792.
2. Ермолаева М.А., Тимонин В.И. Многовыборочный аналог критерия Смирнова проверок степенных гипотез Лемана // Электромагнитные волны и электронные системы. 2011. № 11. С. 6-11.
3. Тимонин В.И., Черномордик О.М. Метод вычисления точного распределения статистик типа Колмогорова-Смирнова при альтернативах Лемана // Теория вероятностей и ее применение. 1985. Т. 30, № 3. С. 572-573.
4. Тимонин В.И. Оптимизация проведения предварительных исследований в теории форсированных испытаний // Вестник МГТУ им. Н.Э. Баумана. Сер. Естественные науки. 2003. № 2. С. 28-41.
5. Тимонин В.И., Ермолаева М.А. Оценки Каплана-Мейера в статистиках типа Колмогорова-Смирнова при проверке гипотез в испытаниях с переменной нагрузкой // Электромагнитные волны и электронные системы. 2010. Т.15, № 7. С. 18-26.
6. Crowder M.J. Multivariate Survival Analysis and Competing Risks. CRC Press; Chapman and Hall, 2012. 417 p. (Ser. Texts in Statistical Science).
7. May S., Hosmer D.W. A simplified method of calculating an overall goodness-of-fit test for the Cox proportional hazards model // Lifetime Data Analysis. 1998. Vol. 4, no. 2. P. 109-120. DOI: [10.1023/A:1009612305785](https://doi.org/10.1023/A:1009612305785)
8. Ермолаева М.А. Непараметрический анализ зависимости между распределениями наработок до отказа изделий и устройств в разных условиях эксплуатации // Труды российского научно-технического общества радиотехники, электроники и связи имени А.С. Попова. Сер. Акустооптические и радиолокационные методы измерений и обработки информации. Вып. 3. М.: РНТОРЭС им. А.С. Попова, 2009. С. 227-230.
9. Corder G.W., Foreman D.I. Nonparametric statistics: A step-by-step approach. New Jersey: Wiley, 2014. 288 p.
10. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. М.: Наука, 1983. 416 с.
11. Hajek J., Sidak Z. Theory of rank tests. London: Academic Press, 2004. 438 p.
12. Gao J., Ozturk O. Two-sample distribution-free inference based on partially rank-ordered set samples // Statistics and Probability Letters. 2012. Vol. 82, iss. 5. P. 876–884. DOI: [10.1016/j.spl.2012.01.021](https://doi.org/10.1016/j.spl.2012.01.021)
13. McLain A.C., Ghosh S.K. Nonparametric estimation of the conditional mean residual life function with censored data // Lifetime Data Analysis. 2011. Vol. 17, no. 4. P. 514-532. DOI: [10.1007/s10985-011-9197-x](https://doi.org/10.1007/s10985-011-9197-x)

The Method of Calculating the Exact Distributions of the Kolmogorov-Smirnov Statistics in Case of Violation of Homogeneity and Independence of the Analyzed Samples

N.D. Tiannikova^{1,*}, V.I. Timonin¹

[*tiannikova@yandex.ru](mailto:tiannikova@yandex.ru)

¹Bauman Moscow State Technical University, Moscow, Russia

Keywords: the Kolmogorov-Smirnov statistics, non-parametric statistics, Kaplan-Meier estimates, power Lehmann dependences

To establish the relationship between the distribution functions of the experimental results for different values of the external factors are most commonly used parametric models in which the parameters of the distribution functions depend on factors, and their views do not change. Meanwhile, when we have a small amount of data (and this is more common in practice), the distribution function is often unknown, and it is difficult to determine. Hence, it is of great importance to evaluate different relationships between the distribution laws without specifying a particular form of the distribution (these issues are handled by non-parametric statistics). The most common model, used to establish the relationship between the theoretical distribution functions of different samples, is the Cox model. Furthermore, even for testing the homogeneity of multiple samples, experiments, which are necessary to obtain them, are so complex that the obtained samples are dependent. So, all of these tasks requires the development of new non-parametric tests for dependency. Due to the fact, that the volume of the sample is always small, knowledge of exact distributions of statistics, which are used, is of special importance. The paper develops a general method for tabulating the exact distributions (for finite volumes of samples) of a wide class of statistics of the Kolmogorov-Smirnov test. With the appropriate specialization of the proposed algorithm, it allows us to calculate the distribution of various statistics of the specified type. In particular, it is applicable for calculating the distribution of statistics such as Kiefer-Gikhman used to check the dependencies between Lehmann distribution functions of several samples. With small modifications it allows us to tabulate the distribution statistics of the Kolmogorov-Smirnov used for checking the homogeneity of dependent samples. Along with the fact that the method has great generality, it also allows us to calculate the exact distribution for very large volumes of samples. This fact allows us to estimate the volume of the sample, in which the asymptotic distribution can be applied.

The limits of this method applicability are also given. It assumes the validity of a special model of random movement of particle on a multidimensional lattice in which the future behavior of the particle trajectory at presently given is independent of its past.

References

1. Timonin V.I. On the Limit Distribution of Statistics of a Nonparametric Test. *Teoriia veroiatnostoni i ee primeneniia*, 1987, vol. 32, no. 4, pp. 790-792. (English translation: *Theory of Probability and Its Applications*, 1988, vol. 32, no. 4, pp. 721-724. DOI: [10.1137/1132108](https://doi.org/10.1137/1132108)).
2. Ermolaeva M.A., Timonin V.I. A multi-Sample Analogue to the Smirnov Test Criterion for the Lehmann Power Hypothesis. *Elektromagnitnye volny i elektronnye sistemy = Electromagnetic Waves and Electronic Systems*, 2011, no. 11, pp. 6-11. (in Russian).
3. Timonin V.I., Chernomordik O.M. A Method for Calculating the Exact Distribution of Kolmogorov–Smirnov Statistics under Lehmann Alternatives. *Teoriya veroyatnostey i ee primeneniye*, 1985, vol. 30, no. 3, pp. 572-573. (English translation: *Theory of Probability and Its Applications*, 1986, vol. 30, no. 3, pp. 608-610. DOI: [10.1137/1130077](https://doi.org/10.1137/1130077)).
4. Timonin V.I. Optimization of Preliminary Studies in Theory of Forced Testing. *Vestnik MGTU im. N.E. Baumana. Ser. Estestvennyye nauki = Herald of the Bauman MSTU. Ser. Natural science*, 2003, no. 2, pp. 28-41. (in Russian).
5. Timonin V.I., Ermolaeva M.A. About Kaplan-Meyer Estimators in Statistics Similar to Kolmogorov-Smirnov for Testing the Hypothesis in Variable Load Tests. *Elektromagnitnye volny i elektronnye sistemy = Electromagnetic Waves and Electronic Systems*, 2010, vol. 15, no. 7, pp. 18-26. (in Russian).
6. Crowder M.J. *Multivariate Survival Analysis and Competing Risks*. CRC Press; Chapman and Hall, 2012. 417 p. (Ser. *Texts in Statistical Science*).
7. May S., Hosmer D.W. A simplified method of calculating an overall goodness-of-fit test for the Cox proportional hazards model. *Lifetime Data Analysis*, 1998, vol. 4, no. 2, pp. 109-120. DOI: [10.1023/A:1009612305785](https://doi.org/10.1023/A:1009612305785)
8. Ermolaeva M.A. Non-parametric analysis of the relationship between the distributions of operating time to failure of products and devices in different operating conditions. *Trudy rossiyskogo nauchno-tekhnicheskogo obshchestva radiotekhniki, elektroniki i svyazi imeni A.S. Popova. Ser. Akustoopticheskie i radiolokatsionnye metody izmereniy i obrabotki informatsii. Vyp. 3* [Proc. of the Russian Scientific and Technical Society of Radio Engineering, Electronics and Communication named after A.S. Popov. Ser. Acoustooptical and Radar

- Methods for Information Measurements and Processing. Iss. 3]. Moscow, RNTORES Publ., 2009, pp. 227-230. (in Russian).
9. Corder G.W., Foreman D.I. *Nonparametric statistics: A step-by-step approach*. New Jersey, Wiley, 2014. 288 p.
 10. Bol'shev L.N., Smirnov N.V. *Tablitsy matematicheskoy statistiki* [Tables of Mathematical Statistics]. Moscow, Nauka Publ., 1983. 416 p. (in Russian).
 11. Hajek J., Sidak Z. *Theory of rank tests*. London, Academic Press, 2004. 438 p.
 12. Gao J., Ozturk O. Two-sample distribution-free inference based on partially rank-ordered set samples. *Statistics and Probability Letters*, 2012, vol. 82, iss. 5, pp. 876–884. DOI: [10.1016/j.spl.2012.01.021](https://doi.org/10.1016/j.spl.2012.01.021)
 13. McLain A.C., Ghosh S.K. Nonparametric estimation of the conditional mean residual life function with censored data. *Lifetime Data Analysis*, 2011, vol. 17, no. 4, pp. 514-532. DOI: [10.1007/s10985-011-9197-x](https://doi.org/10.1007/s10985-011-9197-x)