УДК 004.942:519.25

Приходько С. Б., Макарова Л. М.

ВИЗНАЧЕННЯ ДОВІРЧИХ ІНТЕРВАЛІВ СТАТИСТИЧНИХ МОМЕНТІВ ЧАСУ НАПРАЦЮВАННЯ МІЖ ВІДМОВАМИ ПРИСТРОЇВ ТЕРМІНАЛЬНОЇ МЕРЕЖІ

Виконано знаходження довірчих інтервалів вибіркового середнього і середньо-квадратичного відхилення часу напрацювання між відмовами пристроїв термінальної мережі на основі нормалізуючого перетворення Джонсона S_B. Проведено порівняння отриманих інтервальних оцінок з результатами непараметричного оцінювання.

***Ключові слова:** довірчий інтервал, вибіркове середнє, середньоквадратичне відхилення, нормалізуюче перетворення, перетворення Джонсона, параметричне оцінювання.*

Выполнено нахождение доверительных интервалов выборочного среднего и среднеквадратического отклонения времени наработки между отказами устройств терминальной сети на основе нормализирующего преобразования Джонсона S_B. Проведено сравнение полученных интервальных оценок с результатами непараметрического оценивания.

***Ключевые слова:** доверительный интервал, выборочное среднее, среднеквадратическое отклонение, нормализирующее преобразование, преобразование Джонсона, параметрическое оценивание.*

Finding confidence intervals of the sample mean and standard deviation of the time between failures of network terminal devices based on the normalizing transformation Johnson S_B is done. The comparison between the obtained interval estimates and the results of the non-parametric estimation is made.

***Key words:** confidence interval, sample mean, standard deviation, normalizing transformation, Johnson transformation, parametric estimation.*

Постановка проблемы. На сегодняшний день существует проблема, связанная с отказами в обслуживании устройств терминальной сети и последующим быстрым восстановлением их работоспособности, для чего необходимо иметь резерв запасных частей для всех элементов устройств терминальной сети, что экономически невыгодно. Для того, чтобы уменьшить такой резерв, необходимо иметь возможность прогнозирования отказов конкретных элементов устройств, для чего оценивать время наработки между отказами. Время наработки между отказами устройств терминальной сети является случайной величиной (СВ), закон распределения которой существенно отличается от нормального и не всегда удовлетворяет экспоненциальному [1; 2]. Оценка характеристик этой СВ является важной задачей для автоматизации системы управления терминальной сетью, т. к. в случае неэффективного использования и простоев терминальных устройств в состоянии отказа в обслуживании снижается качество обслуживания клиентов терминальных сетей, а владельцы этих сетей несут убытки.

При оценивании характеристик СВ, таких, как статистические моменты, можно использовать точечные или интервальные оценки. Однако интервальные оценки являются более надежными по сравнению с точечными.

В настоящее время существуют методы оценивания доверительных интервалов статистических моментов для СВ, имеющей нормальное распределение. Для СВ, распределение которой отличается от нормального, оценивание доверительных интервалов известно только для некоторых законов распределения, например, для экспоненциального [3]. Поэтому проблема нахождения доверительных интервалов статистических моментов времени наработки между отказами устройств терминальной сети является актуальной.

Анализ последних исследований и публикаций. В случае нормального закона распределения определение доверительных интервалов статистических моментов выполняют, используя традиционный способ: на основе *t*-распределения Стьюдента для математического ожидания и на основе распределения χ^2 Пирсона для среднеквадратического отклонения, приведенный, например в [4].

Достаточно провести оценку только первых двух моментов СВ – математического ожидания и дисперсии, оценка остальных моментов может быть получена из указанных. Точечными оценками математического ожидания и среднеквадратического отклонения являются выборочное среднее и выборочное среднеквадратическое отклонение соответственно.

В случае экспоненциального закона распределения определение доверительного интервала выполняют только для математического ожидания на основе распределения χ^2 Пирсона по формуле, приведенной в [3].

В случае, когда СВ не подчиняется нормальному или экспоненциальному закону распределения, для оценивания доверительных интервалов статистических моментов существует два подхода. На сегодняшний день традиционно используется непараметрический подход [5]. Суть непараметрического подхода состоит в принятии следующего предположения: при увеличении количества наблюдений закон распределения СВ стремится к нормальному, т. е. происходит замена эмпирического закона распределения на нормальный и дальнейшая оценка выполняется традиционным способом.

Помимо этого подхода, в работе [6] предложен подход, основанный на применении нормализирующих преобразований. Суть этого подхода: на основе нормализирующего преобразования получить СВ с нормальным законом распределения, вычислить доверительные интервалы точечных оценок статистических моментов для этой СВ традиционным способом, а затем на основе обратного преобразования получить доверительные интервалы статистических моментов начальной СВ. Выбор конкретного нормализирующего преобразования необходимо выполнять в зависимости от эмпирических данных.

Для аппроксимации эмпирических распределений, которые нельзя описать нормальным или экспоненциальным законом, было предложено использовать подход на основе четырехпараметрического преобразования Джонсона. Использование указанного подхода показало хорошие результаты для различных областей применения, например, [7].

Целью данной статьи является нахождение доверительных интервалов выборочного среднего и среднеквадратического отклонения времени наработки между отказами устройств терминальной сети в случае эмпирических распределений, которые нельзя описать экспоненциальным законом, на основе нормализирующего преобразования Джонсона.

Изложение основного материала. Известна СВ x – время наработки между отказами устройств терминальной сети, закон распределения которой существенно отличается от нормального. Данное распределение можно нормализировать с помощью преобразования Джонсона. Для этого по эмпирическому распределению СВ x находим оценки асимметрии и эксцесса, по значениям которых выбираем соответствующее семейство распределений Джонсона. Как было показано в [8], таким семейством является семейство S_B. Далее находим параметры распределения γ, η, λ, φ и выполняем нормализацию СВ x [6]:

$$z = \gamma + \eta q(x, \varphi, \lambda); \eta > 0; -\infty < \gamma < \infty; \lambda > 0; -\infty < \varphi < \infty, \quad (1)$$

где z – нормированная нормально распределенная СВ, $q(x, \varphi, \lambda)$ – нелинейная функция, зависящая от выбранного семейства распределений Джонсона; γ, η, λ, φ – параметры распределения; x – СВ, которая нормализируется.

Находим (1-α)% доверительные интервалы точечных оценок выборочного среднего $[\hat{\alpha}_1(z)]$ и среднеквадратического отклонения $[S_z]$ СВ z для нормальной генеральной совокупности [5]:

$$[\hat{\alpha}_1(z)] = [\bar{z} - t_{n-1} S_z / \sqrt{n}, \bar{z} + t_{n-1} S_z / \sqrt{n}], \quad (2)$$

где $\bar{z}$ – выборочное среднее СВ z; n – количество значений в выборке, S_z – среднеквадратическое отклонение СВ z; t_{n-1} – квантиль t-распределения Стьюдента.

$$[S_z] = \left[S_z \sqrt{n / \chi^2_{\alpha/2}}, S_z \sqrt{n / \chi^2_{1-\alpha/2}} \right], \quad (3)$$

где $\chi^2_{\alpha/2}$ и $\chi^2_{1-\alpha/2}$ – верхние 100α% точки распределения χ^2.

Нахождение доверительных интервалов выборочного среднего $[\alpha_1(x)]$ и среднеквадратического отклонения $[S_x]$ СВ x для всех функций q из преобразования (1) основано на разложении функции $q(x, \varphi, \lambda)$ в ряд Тейлора:

$$q(x, \varphi, \lambda) = \tilde{x} + R,$$

где $\tilde{x} = (x - \varphi) / \lambda$ – член ряда, линейно зависимый от x; R – сумма остаточных членов ряда, и приведено в [9].

Рассмотрим случай семейства Джонсона S_B, для которого $q(x, \varphi, \lambda) = \ln\left(\frac{x - \varphi}{\lambda + \varphi - x}\right), \varphi < x < \varphi + \lambda$, а плотность вероятности задается формулой:

$$f_B(x) = \frac{\eta\lambda}{\sqrt{2\pi}(x - \varphi)(\lambda + \varphi - x)} \exp\left\{ -\frac{1}{2}\left[\gamma + \eta \ln\left(\frac{x - \varphi}{\lambda + \varphi - x} \right) \right]^2 \right\}.$$

Находим (1-α)% доверительные интервалы точечных оценок выборочного среднего $[\hat{\alpha}_1(x)]$ и среднеквадратического отклонения $[S_x]$ СВ x:

$$[\hat{\alpha}_1(x)] = \varphi + \frac{\lambda}{2}\{1 + [\hat{\alpha}_1(\tilde{x})]\}, \quad (4)$$

где $[\hat{\alpha}_1(\tilde{x})] = \frac{[\hat{\alpha}_1(z)] - \gamma}{2\eta} + \hat{\alpha}_1(R)$;

$$\hat{\alpha}_1(R) = \frac{\hat{\alpha}_1(z) - \gamma}{2\eta} - \hat{\alpha}_1(\tilde{x}); \qquad \hat{\alpha}_1(\tilde{x}) = \frac{2}{\lambda}\{\hat{\alpha}_1(x) - \varphi\} - 1.$$

$$[S_x] = \left[\sqrt{\alpha_2(x) - \alpha_1^2(x)} \right], \quad (5)$$

где

$$[\alpha_2(x)] = \frac{\lambda^2}{4}\{[\alpha_2(\tilde{x})] - 1\} + (\lambda - 2\varphi)[\alpha_1(x)] - \varphi(\lambda + \varphi);$$

$$[\alpha_2(\tilde{x})] = [\alpha_2(\tilde{z})] - \alpha_2(R_2);$$

$$[\alpha_2(\tilde{z})] = \frac{1}{4\eta^2}\{[\alpha_2(z)] - 2\gamma[\alpha_1(z)] + \gamma^2\};$$

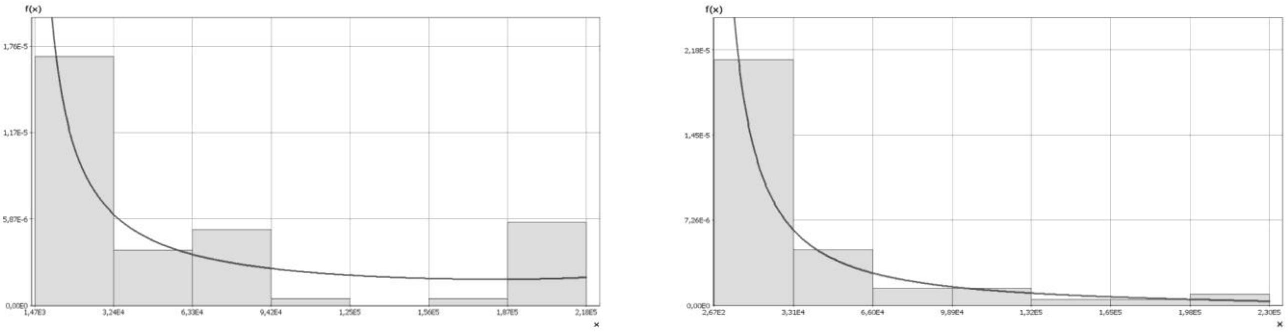
$$[\alpha_2(z)] = [S_z^2] + [\alpha_1^2(z)];$$

$$\alpha_2(R_2) = \alpha_2(\tilde{z}) - \alpha_2(\tilde{x});$$

$$\alpha_2(\tilde{z}) = \frac{1}{4\eta^2}\{\alpha_2(z) - 2\gamma\alpha_1(z) + \gamma^2\};$$

$$\alpha_2(\tilde{x}) = \frac{4}{\lambda^2}\{\alpha_2(x) - 2\varphi\alpha_1(x) + \varphi^2\} - \frac{4}{\lambda}\{\alpha_1(x) - \varphi\} + 1.$$

Для проверки данного подхода был выполнен расчет доверительных интервалов точечных оценок выборочного среднего и среднеквадратического отклонения времени наработки между отказами устройств терминальной сети. В качестве исходных данных рассмотрены 28106 случаев отказа в обслуживании 278 устройств терминальной сети за период 36 месяцев, приведенные в работе [8]. Время наработки между отказами устройств терминальной сети является СВ x. Расчет был проведен для 284 выборок эмпирических данных, гистограммы наиболее типичных из которых представлены на рис. 1. На этом же рисунке приведены и распределения Джонсона S_B, а параметры распределений – в табл. 1.



Выборка 1 **Выборка 2**

Рис. 1. Гистограммы типичных распределений эмпирических данных

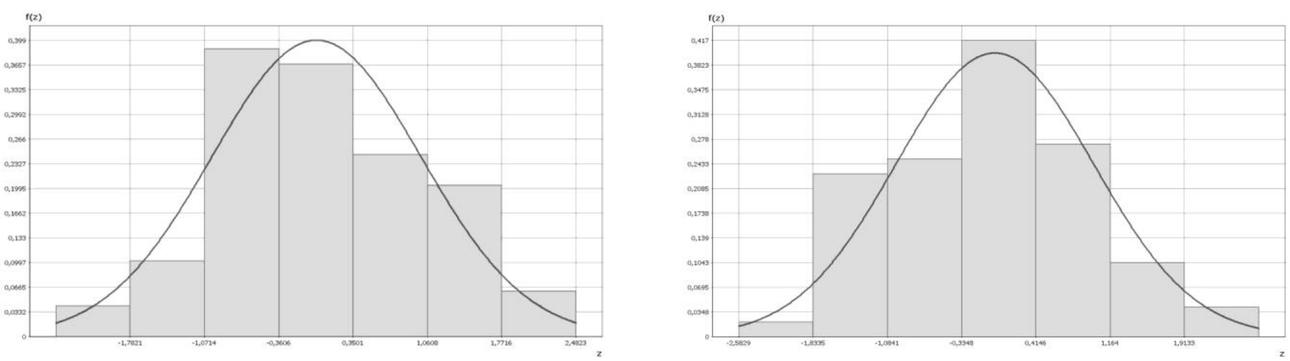
Закон распределения СВ x отличается как от нормального, так и от экспоненциального. Для проверки соответствия эмпирического распределения экспоненциальному закону применялся критерий согласия χ^2 Пирсона. Для выборки 1 $\chi^2 = 6{,}48$, для выборки 2 $\chi^2 = 6{,}62$ при критическом значении $\chi^2 = 5{,}99$. С доверительной вероятностью 0,95 гипотеза о соответствии эмпирической выборки экспоненциальному закону распределения СВ отклоняется. Для СВ x по гистограмме были найдены оценки асимметрии A и эксцесса ε, значения которых приведены в табл. 1. По значениям оценок A^2 и ε в соответствии с диаграммой, приведенной в [9], было выбрано семейство распределений Джонсона S_B.

Параметры распределения γ, η, λ, φ были найдены в результате решения следующей задачи [9]:

$$\hat{\theta} = \arg\min_{\theta}\{A_z^2 + (\varepsilon_z - 3)^2 + m_z^2 + (D_z - 1)^2\},$$

где θ – вектор неизвестных параметров, θ = {γ,η,λ,φ}.

Для выборки 1: γ = 0,7034, η = 0,4078, λ = 219212,8, φ = 1383,5; для выборки 2: γ = 1,4051, η = 0,5081, λ = 249622,9, φ = 169,6. Согласно преобразованию (1) была проведена нормализация СВ x, в результате которой получены значения нормированной нормально распределенной СВ z, эмпирическое и теоретическое распределения которой представлены на рис. 2, а статистические параметры – в табл. 1.



Выборка 1 **Выборка 2**

Рис. 2. Нормализированные распределения эмпирических данных

Таблица 1

Статистические параметры эмпирических и нормализированных выборок данных

Параметр	Выборка 1		Выборка 2	
	эмпирическая	нормализированная	эмпирическая	нормализированная
n	69	69	64	64
m	61679,0	-0,0006	36342,2	-0,00004
D	4774231753,5	0,9998	2661643089,2	0,9999
σ	69095,8	0,9999	51591,1	0,9999
A	1,11	0,1406	2,20	0,2035
ε	2,76	3,0707	7,49	3,1437

Проверка адекватности нормализации выполнена с помощью критерия согласия χ^2 Пирсона. Для выборки 1 $\chi^2 = 3{,}76$, для выборки 2 $\chi^2 = 3{,}98$ при критическом значении $\chi^2 = 9{,}49$. С доверительной вероятностью 0,95 гипотеза о соответствии преобразованной выборки нормальному закону распределения СВ принимается. При этом значения целевой функции составили $2{,}4767 \cdot 10^{-2}$ и $6{,}2071 \cdot 10^{-2}$ соответственно.

По формулам (2) и (3) были найдены 95% доверительные интервалы точечных оценок выборочного среднего и среднеквадратического отклонения СВ x при непараметрическом оценивании. По формулам (4) и (5) были найдены 95% доверительные интервалы точечных оценок выборочного среднего и среднеквадратического отклонения СВ x с использованием преобразования Джонсона. Указанные доверительные интервалы приведены в табл. 2.

Таблица 2

Доверительные интервалы статистических моментов СВ *x*

Параметр	Выборка 1		Выборка 2	
	непараметрическое оценивание	преобразование Джонсона	непараметрическое оценивание	преобразование Джонсона
$[\hat{\alpha}_1(x)]$	[45080,2; 78277,7]	[29220,3; 94137,6]	[23454,9; 49229,6]	[3457,3; 69227,2]
Длина	33197,5	64917,3	25774,7	65769,9
$[S_x]$	[59775,9; 84034,9]	[9601,5; 74415,5]	[44419,6; 63304,0]	[57582,5; 59906,4]
Длина	24259,0	64814,0	18884,4	2323,9

Для выборки 1 ($\varepsilon = 2{,}76 < 3$) непараметрическое оценивание приводит к уменьшению длин доверительных интервалов выборочного среднего и среднеквадратического отклонения в 1,96 и 2,67 раз соответственно по сравнению с подходом на основе преобразования Джонсона. Для выборки 2 ($\varepsilon = 7{,}49 > 3$) непараметрическое оценивание приводит к увеличению длины доверительного интервала среднеквадратического отклонения в 8,13 раз, а длина доверительного интервала выборочного среднего уменьшается в 2,55 раза по сравнению с подходом на основе преобразования Джонсона.

Из представленных расчетов следует, что использование непараметрического оценивания в случае, когда эмпирическое распределение существенно отличается от нормального, может привести к заметно искаженным результатам. В данном случае предпочтительным является применение подхода на основе нормализирующего преобразования Джонсона, однако в случае большого эксцесса эмпирического распределения необходимо уточнить расчет доверительного интервала точечной оценки выборочного среднего другим методом.

Выводы. Найдены доверительные интервалы выборочного среднего и среднеквадратического отклонения времени наработки между отказами устройств терминальной сети в случае эмпирических распределений, которые нельзя описать экспоненциальным законом, на основе нормализирующего преобразования Джонсона.

Для выполнения расчетов было доработано соответствующее программное обеспечение на языке программирования Java.

Планируется использование полученных результатов для дальнейшего развития вероятностной модели распределения времени наработки между отказами устройств терминальной сети на основе нормализирующего преобразования Джонсона и построения информационной технологии автоматизации системы управления терминальной сетью.

ЛИТЕРАТУРА

1. Гнеденко Б. В. Математические методы в теории надежности / Б. В. Гнеденко, Ю. К. Беляев, А. Д. Соловьев. – М. : Наука, 1965. – 524 с.
2. Острейковский В. А. Теория надежности : [учеб. для вузов] / В. А. Острейковский. – М. : Высш. шк., 2003. – 463 с.
3. Монсик В. Б. Оценивание параметра показательного распределения по усеченной выборке / В. Б. Монсик, А. А. Скрынников // Научный вестник Московского государственного технического университета гражданской авиации. Серия Прикладная математика. Информатика. – 2006. – № 105. – С. 134–140.

4. Поллард Дж. Справочник по вычислительным методам статистики / Дж. Поллард ; пер. с англ. В. С. Занадворова; под ред. и с предисл. Е. М. Четыркина. – М. : Финансы и статистика, 1982. – 344 с.
5. Орлов А. И. Прикладная статистика / А. И. Орлов. – М. : Экзамен, 2004. – 656 с.
6. Приходько С. Б Інтервальне оцінювання статистичних моментів негаусівських випадкових величин на основі нормалізуючих перетворень / С. Б. Приходько // Математичне моделювання: науковий журнал. – Дніпродзержинськ : ДДТУ, 2011. – № 1 (24). – С. 9–13.
7. Єременко В. С. Дослідження перетворення Джонсона для задач підвищення точності метрологічних характеристик стандартних зразків / В. С. Єременко, В. М. Мокійчук, О. В. Самойліченко // Системи обробки інформації: зб. наук. праць ХУПС. – 2010. – № 4 (85). – С. 36–42.
8. Приходько С. Б. Выбор аналитической модели закона распределения времени наработки между отказами устройств терминальной сети / С. Б. Приходько, Л. Н. Макарова // Наукові праці: науково-методичний журнал. – Вип. 179. Т. 191. Комп`ютерні технології. – Миколаїв : Вид-во ЧДУ ім. Петра Могили, 2012. – С. 42–45.
9. Коваленко І. І. Сучасні методи статистичного аналізу даних: [навчальний посібник] / І. І. Коваленко, С. Б. Приходько, Л. О. Латанська. – Миколаїв : НУК, 2011. – 192 с.

Рецензенти: Кондратенко Ю. П., д.т.н., професор;
Гожий О. П., к.т.н., доцент.

 Дата надходження статті до редколегії 05.05.2013 р.

ПРИХОДЬКО С. Б., к.т.н., доцент, Національний університет кораблебудування імені адмірала Макарова, м. Миколаїв, Україна.

МАКАРОВА Л. М., пошукувач, м. Миколаїв, Україна.