ОПРЕДЕЛЕНИЕ ДОВЕРИТЕЛЬНОГО ИНТЕРВАЛА ТОЧЕЧНОЙ ОЦЕНКИ ПАРАМЕТРА ЭКСПОНЕНЦИАЛЬНОГО РАСПРЕДЕЛЕНИЯ

УДК 004.942:519.25

ПРИХОДЬКО Сергей Борисович
к.т.н., доцент, заведующий кафедрой программного обеспечения автоматизированных систем
Национального университета кораблестроения им. адмирала Макарова.
Научные интересы: математическое моделирование случайных процессов в информационных технологиях.
e-mail: sergiy.prykhodko@nuos.edu.ua

МАКАРОВА Лидия Николаевна
соискатель Национального университета кораблестроения им. адмирала Макарова./
Научные интересы: математическое моделирование случайных процессов в информационных технологиях.
e-mail: lidiya@ultra.mk.ua

ПОСТАНОВКА ПРОБЛЕМЫ

Экспоненциальный (показательный) закон распределения случайной величины (СВ) служит одним из наиболее часто используемых, наряду с нормальным. Применяется для описания процессов, обладающих свойством отсутствия последствий, в связи с чем является основным в теории стационарных марковских процессов и широко используется в теории массового обслуживания и в теории надежности [1].

В теории надежности экспоненциальный закон является одним из основных законов распределения и используется для описания распределения времени безотказной работы объектов при временных отказах или распределения времени между соседними отказами и времени восстановления. Популярность этого закона распределения объясняется тем, что он физически очень естественен, прост и удобен для использования. Задачи, решаемые на основе экспоненциального закона, оказываются на порядок проще, чем для произвольных законов распределения. Однако подчеркивается, что экспоненциальное распределение является приближенным вариантом аппроксимации эмпирических данных [1, 2, 3].

Существует проблема, связанная с оценкой параметра экспоненциального распределения. Как правило, параметр распределения определяется с помощью точечной оценки. Интервальное оценивание с помощью доверительного интервала является более надежным по сравнению с точечным. Для СВ, распределение которой отличается от нормального, определение доверительных интервалов известно только для некоторых законов распределения, одним из которых и является экспоненциальный [4].

Однако эмпирические данные не всегда могут быть в точности аппроксимированы экспоненциальным законом распределения и могут содержать некоторые отклонения, что потребует уточнения закона распределения СВ.

АНАЛИЗ ПОСЛЕДНИХ ИССЛЕДОВАНИЙ И ПУБЛИКАЦИЙ

Экспоненциальное распределение является непрерывным аналогом геометрического распределения. Это однопараметрическое распределением с параметром λ – постоянной интенсивностью отказов, при этом среднее время безотказной работы определяется как величина, обратная интенсивности отказов $t_{cp} = \frac{1}{\lambda}$ и является математическим ожиданием (начальным моментом первого порядка) экспоненциального распределения. Точечной оценкой математического ожидания является выборочное среднее.

На сегодняшний день существует методика определения доверительного интервала выборочного среднего в случае экспоненциального закона распределения на основе распределения χ^2 Пирсона по формуле, приведенной в [4]. Однако может потребоваться уточнение закона распределения эмпирических данных, содержащих значения, отклоняющиеся от экспоненциального закона. Подобное уточнение может быть получено с помощью применения нормализирующего преобразования Джонсона.

В работах [5, 6] предложен подход, основанный на применении нормализирующих преобразований. Суть этого подхода состоит в следующем: на основе нормализирующего преобразования получить СВ с нормальным законом распределения, определить доверительный интервал выборочного среднего этой СВ традиционным способом на основе *t*-распределения Стьюдента [7], а затем на основе обратного преобразования получить доверительный интервал выборочного среднего начальной СВ. Выбор конкретного нормализирующего преобразования необходимо выполнять в зависимости от эмпирических данных.

Для аппроксимации эмпирических распределений было предложено использовать подход на основе преобразования Джонсона. Использование указанного подхода показало хорошие результаты для различных областей применения [8, 9, 10].

Целью данной статьи является определение доверительного интервала точечной оценки выборочного среднего на основе как экспоненциального закона распределения, так и нормализирующего преобразования Джонсона, и сравнение полученных результатов с существующими.

ИЗЛОЖЕНИЕ ОСНОВНОГО МАТЕРИАЛА

Имеется СВ x, которая подчиняется экспоненциальному закону распределения, плотность вероятности которого задается формулой [2]:

$$f(x) = \lambda e^{-\lambda x}, \quad (1)$$

где λ – параметр распределения.

Параметр распределения λ может быть выражен через величину, обратную математическому ожиданию: $\lambda = \frac{1}{m_x}$. На практике для задания экспоненциального закона распределения можно применять либо интенсивность отказов λ, либо время наработки между отказами t, и тогда параметром распределения будет служить математическое ожидание m_x. Точечной оценкой математического ожидания является выборочное среднее $\hat{m}_x$.

Доверительный интервал выборочного среднего $[\hat{\alpha}_1(x)]$ СВ x в случае экспоненциального закона распределения получим следующим образом [4]:

$$[\hat{\alpha}_1(x)] = \left[2n\hat{m}_x / \chi^2_{2n,(1-\beta)/2}, 2n\hat{m}_x / \chi^2_{2n,(1+\beta)/2}\right], \quad (2)$$

где n – количество значений в выборке, $\hat{m}_x$ – выборочное среднее СВ x; $\chi^2_{2n,(1-\beta)/2}$ и $\chi^2_{2n,(1+\beta)/2}$ – верхние 100α% точки распределения χ^2, β – доверительная вероятность.

При использовании в качестве параметра распределения λ, получим следующий доверительный интервал:

$$[\hat{\lambda}] = \begin{bmatrix} \chi^2_{2n,(1+\beta)/2} \cdot \lambda^* / 2(n-1), \\ \chi^2_{2n,(1-\beta)/2} \cdot \lambda^* / 2(n-1) \end{bmatrix}, \quad (3)$$

где $\lambda^* = \frac{n-1}{\sum_{i=1}^{n} x_i}$ – несмещенная точечная оценка λ.

При использовании подхода, основанного на применении нормализирующих преобразований, значения СВ можно нормализировать с помощью преобразования Джонсона [5]:

$$\begin{aligned} & z = \gamma + \eta h(x, \phi, \lambda); \\ & \eta > 0; \\ & -\infty < \gamma < \infty; \\ & \lambda > 0; \\ & -\infty < \phi < \infty. \end{aligned} \quad (4)$$

Преобразование (4) имеет обратное преобразование:

$$x = \phi + \lambda h^{-1}(z, \gamma, \eta);\quad \eta > 0;\quad -\infty < \gamma < \infty;\quad \lambda > 0;\quad -\infty < \phi < \infty, \tag{5}$$

Здесь z – нормально распределенная случайная величина с математическим ожиданием ноль и дисперсией единица; x – случайная величина с распределением Джонсона; γ, η, λ, φ – параметры преобразования или распределения Джонсона; h и h^{-1} – функции определенного семейства:

$$h = \begin{cases} \ln(\tilde{x}), & x > \varphi, \quad \text{для семейства } S_L; \\ \ln[\tilde{x}/(1-\tilde{x})], & \varphi < x < \varphi + \lambda, \quad \text{для семейства } S_B; \\ \operatorname{Arsh}(\tilde{x}), & -\infty \le x \le +\infty, \quad \text{для семейства } S_U, \end{cases}$$

$$h^{-1} = \begin{cases} e^{\zeta}, & \text{для семейства } S_L; \\ 1/(1+e^{-\zeta}), & \text{для семейства } S_B; \\ (e^{\zeta} - e^{-\zeta})/2, & \text{для семейства } S_U. \end{cases}$$

Конкретное семейство распределений Джонсона выбирается исходя из значений оценок квадрата асимметрии A^2 и эксцесса ε исходной выборки [11].

Значения оценок неизвестных параметров преобразования можно найти с помощью решения задачи математического программирования [5]:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \left\{ A_z^2 + (\varepsilon_z - 3)^2 + \bar{z}^2 + (S_z^2 - 1)^2 \right\}, \tag{6}$$

где $\boldsymbol{\theta}$ – вектор неизвестных параметров, $\boldsymbol{\theta} = \{\gamma, \eta, \varphi, \lambda\}$; A_z – точечная оценка асимметрии нормализированной выборки, $A_z = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} \left(\frac{z_i - \bar{z}}{S_z} \right)^3$; ε_z – точечная оценка эксцесса нормализированной выборки, $\varepsilon_z = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^{n} \left(\frac{z_i - \bar{z}}{S_z} \right)^4$; $\bar{z}$ – выборочное среднее нормализированной выборки, $\bar{z} = \frac{1}{n} \sum_{i=1}^{n} z_i$; S_z^2 – выборочная дисперсия нормализированной выборки, $S_z^2 = \frac{1}{n-1} \sum_{i=1}^{n} (z_i - \bar{z})^2$; n – количество значений в выборке.

Алгоритм определения (1-α)% доверительного интервала выборочного среднего $[\hat{\alpha}_1(x)]$ СВ x определяется следующим образом [6]:

1) для выборочного среднего $\bar{x}$ определяем соответствующее значение $\bar{z}$ по формуле (4);

2) находим $\Delta z = t_{n-1} S_z / \sqrt{n}$;

3) получаем границы доверительного интервала выборочного среднего СВ *z*: $z_{min} = \bar{z} - \Delta z$ и $z_{max} = \bar{z} + \Delta z$;

4) для значений z_{min} и z_{max} по преобразованию (5), получаем границы доверительного интервала выборочного среднего СВ x: x_{min} и x_{max}.

Для проверки приведенного подхода выполним расчет примера 1 из [4], количество значений в выборке n = 12, выборочное среднее $\bar{x} = 547{,}25$, выборочная дисперсия $S_x^2 = 158586{,}19$. Расчет доверительных интервалов выборочного среднего $[\hat{\alpha}_1(x)]$ и точечной оценки параметра $[\hat{\lambda}]$ для экспоненциального закона распределения выполнялся по формулам (2) и (3). Расчет с применением нормализирующего преобразования Джонсона выполнен по приведенному выше алгоритму. По значениям оценок $A^2 = 0{,}9357$ и $\varepsilon = 2{,}7256$ в соответствии с диаграммой, приведенной в [11], было выбрано семейство распределений Джонсона S_B. Параметры распределения γ, η, λ, φ были найдены в результате решения задачи (6): $\gamma = 3{,}1525$, $\eta = 1{,}0853$, $\varphi = 34{,}1672$, $\lambda = 7337{,}26$. Также было выполнено непараметрическое оценивание согласно [12]. Результаты расчета приведены в табл. 1.

Таблица 1 –

Доверительные интервалы выборочного среднего и точечной оценки параметра экспоненциального закона распределения

Оценка	Значения из [4]	Расчет по формулам (2) и (3)	Расчет с применением преобразования Джонсона	Непараметрическое оценивание
$[\hat{\alpha}_1(x)]$	не рассчитывалось	[297,51; 944,34]	[294,02; 810,59]	[235,00; 740,82]
$[\hat{\lambda}]$	$[1{,}15 \cdot 10^{-3}; 3{,}67 \cdot 10^{-3}]$	$[1{,}155 \cdot 10^{-3}; 3{,}667 \cdot 10^{-3}]$	$[1{,}234 \cdot 10^{-3}; 3{,}401 \cdot 10^{-3}]$	$[1{,}350 \cdot 10^{-3}; 4{,}255 \cdot 10^{-3}]$

Проверка адекватности нормализации выполнена с помощью критерия согласия Колмогорова-Смирнова [13, 14], т.к. приведенная выборка – малая ($n < 30$). Для критерия Колмогорова $\lambda = 0,544$ при критическом значении $\lambda = 0,895$, для критерия Смирнова $W^2 = 0,058$ при критическом значении $W^2 = 0,126$. С доверительной вероятностью 0,95 гипотеза о соответствии преобразованной выборки нормальному закону распределения СВ принимается.

Проведенный расчет показал следующее. Подход на основе нормализирующего преобразования Джонсона дает адекватный результат: относительные погрешности для левой и правой границы $\left[\hat{\lambda}\right]$ составляют 7,3%, а относительно выборочного среднего $\overline{\lambda} = 2{,}050 \cdot 10^{-3}$ границы расположены несимметрично как и в [4]. Использование непараметрического оценивания согласно [12] приводит к качественно неправильному результату: относительно выборочного среднего $\overline{\lambda}$ границы расположены симметрично. При этом относительные погрешности для левой и правой границы $\left[\hat{\lambda}\right]$ составляют соответственно 17,4% и 15,9%.

ВЫВОДЫ

Выполнено определение доверительного интервала точечной оценки выборочного среднего на основе как экспоненциального закона распределения, так и нормализирующего преобразования Джонсона. Проведено сравнение полученных результатов с существующими данными. Подход на основе нормализирующего преобразования Джонсона является более универсальным, т.к. его можно использовать в случае, когда эмпирические данные отличаются от экспоненциального закона распределения.

Для выполнения расчетов было доработано соответствующее программное обеспечение на языке программирования Java. Проведенный расчет показал его работоспособность и адекватность.

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. Острейковский В.А. Теория надежности: Учеб. для вузов /В.А. Острейковский. – М.: Высш. шк., 2003. – 463 с.
2. Гнеденко Б.В. Математические методы в теории надежности /Б.В. Гнеденко, Ю.К. Беляев, А.Д. Соловьев. – М.: Наука, 1965. – 524 с.
3. Джонсон Н. Статистика и планирование эксперимента в технике и науке /Н. Джонсон, Ф. Лион. – М.: Мир, 1980. – 610 с.
4. 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.
5. Приходько С.Б. Інтервальне оцінювання статистичних моментів негаусівських випадкових величин на основі нормалізуючих перетворень /С.Б. Приходько //Математичне моделювання: науковий журнал. – 2011. – №1 (24). – С.9-13.
6. Приходько С.Б. Метод побудови нелінійних рівнянь регресії на основі нормалізуючих перетворень /С.Б. Приходько //Тези доповідей міждерж. науково-методич. конф. «Проблеми математичного моделювання» (13-15 червня 2012 р., м. Дніпродзержинськ). – Дніпродзержинськ: ДДТУ, 2012. – С.31-33.
7. 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.
8. Приходько С.Б. Інтервальне оцінювання математичного сподівання часу затримок виконання програмних проектів на основі перетворення Джонсона /С.Б. Приходько, А.В. Пухалевич //Вестник Херсонского национального технического университета. – 2010. – №2 (38). – С.402-404.
9. Приходько С.Б. Выбор нормализующего преобразования для оценки продолжительности работ при управлении временем в проектах разработки конструкторской документации судна /С.Б. Приходько, О.А. Кудин //Зб. наук. праць НУК. – 2011. – №4 (439). – С.140-145.
10. 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.
11. Коваленко І.І. Сучасні методи статистичного аналізу даних: Навчальний посібник /І.І. Коваленко, С.Б. Приходько, Л.О. Латанська. – Миколаїв: НУК, 2011. – 192 с.
12. Орлов А.И. Прикладная статистика /А.И. Орлов. – М.: Экзамен, 2004. – 656 с.
13. Степнов М.Н. Статистические методы обработки результатов механических испытаний: Справочник /М.Н. Степнов. – М.: Машиностроение, 1985. – 232 с.
14. Тюрин Ю.Н. Непараметрические методы статистики /Ю.Н. Тюрин. – М.: Знание, 1978. – 64 с.