

## ОЦЕНКА РАСПРОСТРАНЕННОСТИ ВИДОВ ЗАКОНОВ РАСПРЕДЕЛЕНИЙ НА ОСНОВЕ ВЕЛИЧИНЫ ЭНТРОПИЙНОГО КОЭФФИЦИЕНТА

В. Л. ЛАЗАРЕВ<sup>1</sup>, А. В. СПЕСИВЦЕВ<sup>2</sup>, В. В. ЗАХАРОВ<sup>3</sup>

<sup>1</sup>Университет ИТМО, 197101, Санкт-Петербург, Россия  
E-mail: info@npp-itb.spb.ru

<sup>2</sup>Военно-космическая академия им. А. Ф. Можайского, 197198, Санкт-Петербург, Россия

<sup>3</sup>Санкт-Петербургский институт информатики и автоматизации РАН, 199178, Санкт-Петербург, Россия  
E-mail: valeriov@yandex.ru

Исследована „распространенность“ законов распределений на практике. Исследования проводились на основе энтропийного коэффициента для параметров различных объектов и процессов. Обоснована несостоятельность подхода, основанного на „тотальной“ аппроксимации реальных законов распределений нормальным законом при проведении исследований и организации управления объектами различной природы. Учет изменчивости анализируемого параметра в виде величины энтропийного коэффициента реального закона распределения повышает адекватность описаний состояний неопределенности, что, в свою очередь, сказывается на эффективности исследований и принятия решений. В этом смысле весьма удобным и полезным инструментарием решения подобных задач являются методы и технологии теории энтропийных потенциалов. Возможности этой теории позволяют осуществлять мониторинг и организацию управления состояниями неопределенности сложных систем. Использование полученных данных позволяет оперативно оценивать доминирование конкретных типов законов распределений в различных ситуациях при решении задач мониторинга и управления на основе энтропийного подхода.

**Ключевые слова:** энтропия, состояние неопределенности, энтропийный коэффициент, законы распределения, мониторинг

**Введение.** Понятие энтропийного коэффициента  $K_e$  широко используется в теории энтропийных потенциалов (ТЭП), теории информации, мягких измерениях и вычислениях, которые получают все большее признание и распространение [1—3]. Использование величины  $K_e$  позволяет учесть влияние закона распределения на состояние неопределенности рассматриваемого параметра, что способствует повышению качества мониторинга и организации управления. Изначально существовала практика аппроксимации реальных законов распределения параметров нормальным законом. Она обосновывалась „изученностью“ свойств этого закона, простотой и удобством реализации вычислительных процедур для получения различных статистических оценок. Возрастание требований к повышению качества мониторинга и управления обусловило необходимость учета реальных законов распределения параметров как одного из важных факторов состояния неопределенности объектов различной природы.

Предпринимались различные попытки количественного описания свойств законов распределений. Например, классический подход основан на использовании величины вероятностной энтропии, требующей выборки наблюдений значительного объема [4—7]. Другой подход основан на введении специальных характеристик: коэффициента относительного рассеяния, практически предельного поля рассеяния, практически предельного отклонения и др. [8]. Особенности, достоинства, недостатки подходов приведены в [1]. Там же обоснована целесообразность использования величины  $K_e$  для описания свойств законов распределений

параметров. Дополнительные результаты исследований, подтверждающие состоятельность и эффективность предложенного выбора, приведены в [9].

**Постановка задачи исследований.** Энтропийный коэффициент позволяет установить связь между одномерными величинами:  $\sigma_x$  — средним квадратическим отклонением (СКО) какого-либо параметра  $x$  и  $\Delta_e$  — энтропийным потенциалом (ЭП) этого же параметра  $\Delta_e = K_e \sigma$ . Величина ЭП определяется как половина диапазона распределения по закону равномерной плотности, имеющего такую же энтропию  $H_x$ , как и у закона распределения этого параметра [1—3]. Величина ЭП, таким образом, является характеристикой состояния неопределенности параметра:

$$K_e = \frac{\Delta_e}{\sigma_x} = \frac{\frac{1}{2} e^{H_x}}{\sigma_x} = \frac{\frac{1}{2} e^{-\int_{-\infty}^{\infty} p(x) \ln p(x) dx}}{\sigma_x}, \quad (1)$$

где  $p(x)$  — функция плотности вероятностей (закон) распределения параметра.

Из (1) видно, что энтропийный коэффициент определяется видом закона распределения, он характеризует „предсказуемость“ появления конкретных значений параметра. ЭП является уточненным вариантом описания состояния неопределенности по отношению к величине СКО, т.к. учитывает дополнительно свойства закона распределения.

Если закон распределения имеет аналитическое описание, то найти соответствующее значение  $K_e$  возможно с использованием выражения (1). В противном случае найти значение  $K_e$  возможно на основе результатов наблюдений с использованием дискретного „аналога“ (1), основанного на использовании ступенчатой гистограммы. Тогда процедура интегрирования функции  $p(x)$  заменяется суммированием по всем столбцам этой гистограммы.

Разработан ряд методов определения значений  $K_e$  для различных выборок исходных данных (результатов наблюдений) [1, 3]. Среди них стоит выделить метод робастного оценивания [1] на основе ограниченной выборки наблюдений. Метод позволяет получать искомые оценки величины  $K_e$  на основании анализа „размаха“ выборки с использованием специальных графических моделей.

Теоретически значения  $K_e$  являются функциональным параметром, т.е.  $K_e = [0; 2,07] \subset \mathbf{R}^1$ . Однако для подавляющего большинства случаев практических задач реальный диапазон изменения этой величины  $1 \leq K_e \leq 2,07$ , при этом верхний предел изменения  $K_e = 2,07$  соответствует нормальному закону распределения вероятностей. Другими словами, при одинаковой дисперсии параметра (или СКО) нормальный закон дает наибольшее значение величины ЭП, обеспечивая тем самым наибольший уровень его состояния неопределенности по отношению к любому другому закону.

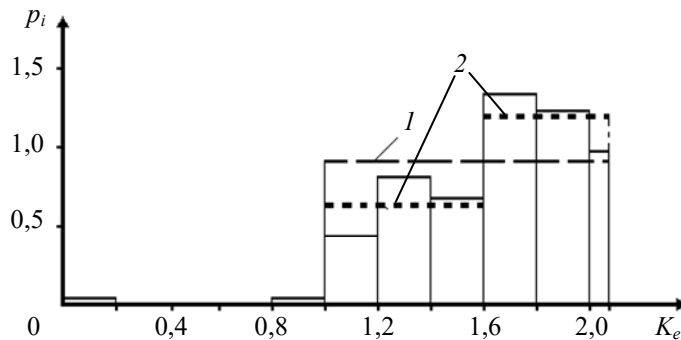
Такое „мажорантное“ свойство нормального закона, при его использовании в задачах аппроксимации, позволяет обеспечить своеобразный запас прочности в задачах оценивания состояний неопределенности при наличии различных просчетов, непредвиденного стечения обстоятельств и пр., но при этом чувствительность и качество соответствующих методов анализа, мониторинга и управления снижаются.

В качестве примера можно привести следующее: для нормального закона  $-\infty < x < \infty$ . С другой стороны, известно, что для подавляющего большинства параметров различных объектов и процессов диапазоны изменений ограничены. Очевидно, что аппроксимация реальных законов распределения нормальным законом исказит качество описания состояний неопределенности параметров.

**Методология описания вариативных свойств энтропийного коэффициента. Исходные данные.** Информативной и удобной формой описания вариативных свойств какой-

либо величины является ее закон распределения. Поскольку аналитического выражения такой зависимости для  $K_e$  не существует, то оценку плотности распределения можно осуществлять на основе имеющихся результатов наблюдений.

Гистограмма (см. рисунок) получена на основе значений  $K_e$  для законов распределения погрешностей средств измерений различных параметров [3], а также на основе значений  $K_e$  для законов распределений технологических параметров в различных производственных процессах [1, 10, 11]. В результате получена репрезентативная выборка по величине  $K_e$  суммарным объемом  $n=234$  значений. Весь диапазон изменения  $K_e=[0,0; 2,07]$  разбит на 11 интервалов, которые отложены на оси абсцисс: десять интервалов от 0,0 до 2,0 с шагом 0,2 и одиннадцатый — с шагом 0,07. На оси ординат показаны значения оценок средней плотности распределения  $K_e$  в этих интервалах  $p_i$  ( $i=1, 2, \dots, 11$ ), т.е.  $p_i = \frac{n_i}{nd_i}$ , где  $n_i$  — число точек, попадающих в  $i$ -й интервал,  $d_i$  — длина  $i$ -го интервала.



Следует отметить, что представленные результаты получены на основании ограниченной выборки значений  $K_e$  на объектах различной природы и не претендуют на исчерпывающую полноту. При изменении объема выборки или перечня анализируемых объектов вид гистограммы может претерпевать изменения. Тем не менее результаты отражают имеющиеся тенденции проявления вариативных свойств величины  $K_e$ .

Особенности представленной гистограммы состоят в следующем. Информативная часть гистограммы определяется диапазоном реальных значений энтропийного коэффициента. В первом интервале ( $0 \leq K_e < 0,2$ ) находятся значения, соответствующие т.н. экстремальным законам распределения, например, распределению Коши с  $K_e=0$ . Представленность этих законов в общей массе наблюдений относительно невелика. Площадь первого столбца не превышает 0,03, общая площадь равна единице. В следующих трех последовательных интервалах изменений величины  $K_e$  ( $0 \leq K_e < 0,8$ ) результатов не зафиксировано. Пятый интервал изменений величины  $K_e$  ( $0,8 \leq K_e < 1,0$ ) также характеризуется незначительным количеством результатов попаданий в реальных задачах. Площадь соответствующего столбца не превышает 0,05. Таким образом, подтверждается вывод о том, что „рабочая“ часть гистограммы по оси абсцисс определяется интервалом  $1 \leq K_e \leq 2,07$ .

Для удобства использования гистограммы необходимо выполнить ее адекватную аппроксимацию. В основу выбора аппроксимирующей функции могут быть положены следующие соображения. Можно полагать, что средняя плотность распределения величин  $K_e$  в рассмотренных пяти интервалах равна нулю, т.е.  $p_i=0$  ( $i=1, 2, \dots, 5$ ). Следовательно, соответствующий фрагмент аппроксимирующей функции в интервале  $0,0 \leq K_e < 1,0$  будет совпадать с осью абсцисс. На следующей части гистограммы возможны различные варианты аппроксимации (см. рисунок, 1). Первый, самый простой и удобный, основан на допущении (исходя из вида рабочей части гистограммы), что средние плотности распределения в интервале  $[1,0; 2,07]$

подчиняются закону равномерной плотности с  $p = \frac{1}{2,07-1} \approx 0,93$ . Другими словами, осуществляется усреднение по всему рабочему интервалу.

Второй, более адекватный полученным результатам, вариант предполагает использование двухступенчатой аппроксимации (см. рисунок, 2). В этом случае осуществляется последовательное усреднение по двум участкам, состоящим из 6, 7 и 8 интервалов и 9, 10 и 11 интервалов. Соответствующее выражение для средних плотностей распределения в указанных интервалах будет иметь вид

$$p \approx \begin{cases} 0,0, & \text{при } 0,0 \leq K_e < 1,0, \\ 0,6, & \text{при } 1,0 \leq K_e < 1,6, \\ 1,3, & \text{при } 1,6 \leq K_e < 2,07. \end{cases} \quad (2)$$

Возможны трехступенчатый вариант аппроксимации и другие.

**Обсуждение результатов.** Анализ полученной гистограммы позволяет сделать вывод, что на практике наиболее распространены законы, значения энтропийного коэффициента которых находятся в девятом интервале:  $1,6 \leq K_e < 1,8$ . „Знаковым“ представителем этого интервала является широко распространенный закон равномерной плотности с  $K_e \approx 1,73$ . Здесь следует иметь в виду, что множество законов распределения по параметру  $K_e$  обладают свойством сюръекции [1—3]. Другими словами, могут существовать подмножества законов, имеющих одинаковое значение  $K_e$ . В данном случае, например, это могут быть разновидности симметричных двумодальных законов и ряд других. При этом вид кривой классического закона равномерной плотности распределения будет являться неким „собирательным образом“ этого подмножества. В любом случае вид кривых плотностей распределения таких законов будет отличаться от соответствующей кривой нормального закона распределения. Это обстоятельство дополнительно подтверждает целесообразность учета вида закона распределения (в данном случае — по величине энтропийного коэффициента) при исследовании состояний неопределенности и организации управления этими состояниями.

Дальнейшим „приближением“ к нормальному закону являются десятый и начало одиннадцатого интервала. Здесь представлен т.н. треугольный закон, или распределение Симпсона, которое характеризуется  $K_e \approx 2,02$ . При перемещении вправо по оси абсцисс проявляется тенденция к уменьшению наклонов „боковых фронтов“ кривых законов распределения и возрастания размаха параметра, который в пределе, при  $K_e \approx 2,07$ , стремится к бесконечности, что характерно для нормального закона распределения [1, 9, 12]. Граница перехода строго не определена, но таковой можно считать  $K_e \approx 2,04—2,05$ . Исходя из этого представляется возможным оценить „долю“ законов, находящихся правее этой границы и условно относящихся к классу нормального закона. Соответствующая площадь на гистограмме не будет превышать 0,08, это также свидетельствует о том, что нормальный закон не доминирует над реально существующими законами.

Такие выводы необходимо учитывать при проведении мониторинга и организации управления состояниями неопределенности [13—15].

**Заключение.** В работе исследованы вариации энтропийного коэффициента, являющегося одной из важнейших характеристик законов распределений различных параметров. Показано, что аппроксимация реальных законов распределений нормальным зачастую может исказить описание состояний неопределенности. Для повышения качества проводимых исследований и принятия решений в условиях неопределенности необходимо учитывать свойства законов распределений параметров в конкретных ситуациях, которые предлагается оце-

нивать по значениям энтропийных коэффициентов. Удобным инструментом проведения исследований в рассматриваемом направлении являются методы и технологии ТЭП.

Исследования, выполненные В. В. Захаровым по данной тематике, проводились при частичной финансовой поддержке грантов РФФИ (17-29-07073-офи-м, 18-07-01272, 18-08-01505, 19-08-00989), Госзадания Министерства образования и науки РФ №2.3135.2017/4.6, в рамках бюджетной темы №№0073-2019-0004.

#### СПИСОК ЛИТЕРАТУРЫ

1. Лазарев В. Л. Теория энтропийных потенциалов. СПб: Изд-во Политехн. ун-та, 2012. 127 с.
2. Прокопчина С. В., Тарасов В. Б., Лазарев В. Л., Горохов В. Л., Барышев Ю. В., Витковский В. В., Куприянов М. С., Логинская Л. Г., Шичкина Ю. А. Мягкие вычисления и измерения. Т. 1. Теоретические основы и методы / Под ред. С. В. Прокопчиной. М.: Научная библиотека, 2017. 420 с.
3. Туричин А. М., Новицкий П. В., Левшина Е. С. Электрические измерения неэлектрических величин / Под ред. П. В. Новицкого. Л.: Энергия, 1975. 576 с.
4. Прангшвили И. В. Энтропийные и другие системные закономерности: Вопросы управления сложными системами. М.: Наука, 2003. 428 с.
5. Колмогоров А. Н. Теория информации и теория алгоритмов. М.: Наука, 1987. 304 с.
6. Шеннон К. Работы по теории информации и кибернетике. М.: Изд-во иностранной лит-ры, 1963. 830 с.
7. Кузнецов Н. А. Фундаментальные основы инфокоммуникаций // Мехатроника, автоматизация, управление. 2007. № 4. С. 21—31.
8. Точность производства в машиностроении и приборостроении / Под ред. А. Н. Гаврилова. М.: Машиностроение, 1973. 567 с.
9. Лазарев В. Л. Исследование вариаций параметров на основе значений энтропийного коэффициента // Сб. докл. XXI Междунар. конф. по мягким вычислениям и измерениям SCM-2018. 23—25 мая 2018. СПб: Изд-во СПбГЭТУ „ЛЭТИ“, 2018. Т. 1. С. 7—10.
10. Лазарев В. Л., Богданов П. А., Петров М. М., Томсон К. Ю. К вопросу оценки неоднородности свойств и состава сред // Вестн. Междунар. академии холода. 2018. № 1. С. 74—80.
11. Лазарев В. Л. Информационное шкалирование вариаций законов распределения параметров в приложениях к задачам мониторинга и управления // Науч.-техн. ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2019. Т. 12, № 2. С. 7—15. DOI: 10.18721/JCSTCS.12201.
12. Прокопчина С. В., Шестопалов М. Ю., Уткин Л. В., Куприянов М. С., Лазарев В. Л., Имаев Д. Х., Горохов В. Л., Жук Ю. А., Спесивцев А. В. Управление в условиях неопределенности. СПб: СПбГЭТУ „ЛЭТИ“, 2014. 304 с.
13. Благовецкая М. М., Злобин Л. А. Информационные технологии систем управления технологическими процессами. М.: Высш. шк., 2005. 768 с.
14. Уткин Л. В. Анализ риска и принятие решений при неполной информации. СПб: Наука, 2007. 404 с.
15. Zadeh L. A. Toward a Generalized Theory of Uncertainty (GTU): an Outline // Information Sciences – Information and Computer Science. 2005. Vol. 172, N 1—2. P. 1—40.

#### Сведения об авторах

- Виктор Лазаревич Лазарев** — канд. техн. наук, доцент; Университет ИТМО, факультет низкотемпературной энергетики; E-mail: vlazarev@itmo.ru
- Александр Васильевич Спесивцев** — канд. техн. наук, доцент; ВКА им. А. Ф. Можайского; E-mail: sav2050@gmail.com
- Валерий Вячеславович Захаров** — Санкт-Петербургский институт информатики и автоматизации РАН, лаборатория информационных технологий в системном анализе и моделировании; младший научный сотрудник; E-mail: valeriov@yandex.ru

Поступила в редакцию  
04.10.19 г.

**Ссылка для цитирования:** Лазарев В. Л., Спесивцев А. В., Захаров В. В. Оценка распространенности видов законов распределений на основе величины энтропийного коэффициента // Изв. вузов. Приборостроение. 2020. Т. 63, № 2. С. 113—119.

## ESTIMATION OF THE PREVALENCE OF TYPES OF DISTRIBUTION LAWS BASED ON THE ENTROPY COEFFICIENT VALUE

V. L. Lazarev<sup>1</sup>, A. V. Spesivtsev<sup>2</sup>, V. V. Zakharov<sup>3</sup>

<sup>1</sup>ITMO University, 197101, St. Petersburg, Russia  
E-mail: info@npp-itb.spb.ru

<sup>2</sup>Mozhaisky Military Space Academy, 197198, St. Petersburg, Russia

<sup>3</sup>St. Petersburg Institute for Informatics and Automation of the RAS, 199178, St. Petersburg, Russia  
E-mail: valeriov@yandex.ru

Results of studies of the "prevalence" of distribution laws in practice are presented. The basis of the studies is the entropy coefficient for parameters of various objects and processes. Inconsistency of the approach based on "total" approximation of the real distribution laws by the normal law when conducting research and organizing the management of objects of different nature is substantiated. It is argued that accounting for variability of the analyzed parameter in the form of the entropy coefficient of a real distribution law increases the adequacy of descriptions of uncertainty states, which, in turn, affects the effectiveness of research and decision-making. It makes the methods and technologies of the theory of entropy potentials very convenient and useful "tools" for solving such problems in question. The theory potentialities allow monitoring and organization of control of uncertainty states of complex systems. The use of obtained results makes it possible to quickly assess the dominance of specific types of distribution laws in different situations when solving monitoring and control problems using the entropy approach.

**Keywords:** entropy, state of uncertainty, entropy coefficient, distribution laws, monitoring

### REFERENCES

1. Lazarev V.L. *Teoriya entropiynykh potentsialov* (Theory of Entropy Potentials), St. Petersburg, 2012, 127 p. (in Russ.)
2. Prokopchina S.V., Tarasov V.B., Lazarev V.L., Gorokhov V.L., Baryshev Yu.V., Vitkovskiy V.V., Kupriyanov M.S., Loginskaya L.G., Shichkina Yu.A. *Myagkiye vychisleniya i izmereniya. T. 1. Teoreticheskiye osnovy i metody* (Soft Calculations and Measurements. Vol. 1. Theoretical Foundations and Methods), Moscow, 2017, 420 p. (in Russ.)
3. Turichin A.M., Novitskiy P.V., Levshina E.S. *Elektricheskiye izmereniya neelektricheskikh velichin* (Electrical Measurements of Non-Electrical Quantities), Leningrad, 1975, 576 p. (in Russ.)
4. Prangishvili I.V. *Entropiynyye i drugie sistemnyye zakonomernosti: Voprosy upravleniya slozhnyimi sistemami* (Entropic and Other Systemic Laws: Issues of Managing Complex Systems), Moscow, 2003, 428 p. (in Russ.)
5. Kolmogorov A.N. *Teoriya informatsii i teoriya algoritmov* (Information Theory and Algorithm Theory), Moscow, 1987, 304 p. (in Russ.)
6. Shannon C. *Raboty po teorii informatsii i kibernetike* (Works on Information Theory and Cybernetics), Moscow, 1963, 830 p. (in Russ.)
7. Kuznetsov N.A. *Mekhatronika, avtomatizatsiya, upravleniye*, 2007, no. 4, pp. 21–31. (in Russ.)
8. Gavrilov A.N., ed., *Tochnost' proizvodstva v mashinostroyenii i priborostroyenii* (Precision in Mechanical Engineering and Instrumentation), Moscow, 1973, 567 p. (in Russ.)
9. Lazarev V.L. *Sbornik dokladov XXI Mezhdunarodnoy konferentsii po myagkim vychisleniyam i izmereniyam SCM-2018* (Collection of Reports of the XXI International Conference on Soft Computing and Measurements SCM-2018), St. Petersburg, May 23–25, 2018, vol. 1, pp. 7–10. (in Russ.)
10. Lazarev V.L., Bogdanov P.A., Petrov M.M., Thomson K.Yu. *Journal of International Academy of Refrigeration*, 2018, no. 1, pp. 74–80. (in Russ.)
11. Lazarev V.L. *St. Petersburg State Polytechnical University Journal. Computer Science. Telecommunications and Control Systems*, 2019, no. 2(12), pp. 7–15. DOI: 10.18721/JCSTCS.12201. (in Russ.)
12. Prokopchina S.V., Shestopalov M.Yu., Utkin L.V., Kupriyanov M.S., Lazarev V.L., Imaev D.Kh., Gorokhov V.L., Zhuk Yu.A., Spesivtsev A.B. *Upravleniye v usloviyakh neopredelennosti* (Uncertainty Management), St. Petersburg, 2014, 304 p. (in Russ.)
13. Blagoveshchenskaya M.M., Zlobin L.A. *Informatsionnyye tekhnologii sistem upravleniya tekhnologicheskimi protsessami* (Information Technology Process Control Systems), Moscow, 2005, 768 p. (in Russ.)

14. Utkin L.V. *Analiz riska i prinyatiye resheniy pri nepolnoy informatsii* (Risk Analysis and Decision Making with Incomplete Information), St. Petersburg, 2007, 404 p. (in Russ.)
15. Zadeh L.A. *Information Sciences – Information and Computer Science*, 2005, no. 1–2(172), pp. 1–40.

**Data on authors**

- Victor L. Lazarev** — PhD, Associate Professor; ITMO University, Faculty of Cryogenic Engineering; E-mail: vllazarev@itmo.ru
- Alexander V. Spesivtsev** — PhD, Associate Professor; Mozhaisky Military Space Academy; E-mail: sav2050@gmail.com
- Valery V. Zakharov** — St. Petersburg Institute for Informatics and Automation of the RAS, Laboratory of Information Technologies in System Analysis and Modeling; Junior Researcher; E-mail: valeriov@yandex.ru

**For citation:** Lazarev V. L., Spesivtsev A. V., Zakharov V. V. Estimation of the prevalence of types of distribution laws based on the entropy coefficient value. *Journal of Instrument Engineering*. 2020. Vol. 63, N 2. P. 113–119 (in Russian).

DOI: 10.17586/0021-3454-2020-63-2-113-119