

Научная статья

УДК 519.24

doi: 10.17223/19988605/61/6

Распределения второго порядка: построение, операции, приложения**Борис Станиславович Добронетс¹, Ольга Аркадьевна Попова²**^{1,2} *Сибирский федеральный университет, Красноярск, Россия*¹ *BDobronets@yandex.ru*² *OlgaArc@yandex.ru*

Аннотация. Рассмотрены методы исследования неопределенностей вероятностных оценок эмпирических функций распределения. Для этого предлагается новый подход, основанный на понятии распределения второго порядка. Для численных реализаций операций над распределениями второго порядка применяются вероятностные расширения их параметризованных представлений. Приведены примеры использования данного подхода для оценки надежности технических объектов в условиях малых выборок.

Ключевые слова: вычислительный вероятностный анализ; распределения второго порядка; неточные вероятности; вероятностные расширения

Для цитирования: Добронетс Б.С., Попова О.А. Распределения второго порядка: построение, операции, приложения // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2022. № 61. С. 61–68. doi: 10.17223/19988605/61/6

Original article

doi: 10.17223/19988605/61/6

Second-order distributions: construction, operations, applications**Boris S. Dobronets¹, Olga A. Popova²**^{1,2} *Siberian Federal University, Krasnoyarsk, Russia*¹ *BDobronets@yandex.ru*² *OlgaArc@yandex.ru*

Abstract. The article considers methods for studying the uncertainties of probabilistic estimates of empirical distribution functions. To this end, a new approach based on the concept of a second-order distribution is proposed. For numerical implementations of operations on second-order distributions, probabilistic extensions of their parametrized representations are used. Examples of using this approach to assess the reliability of technical objects under conditions of small samples are given.

Keywords: computational probabilistic analysis; second-order distributions; imprecise probabilities; probabilistic extensions

For citation: Dobronets, B.S. Popova, O.A. (2022) Second-order distributions: construction, operations, applications. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 61. pp. 61–68. doi: 10.17223/19988605/61/6

Построение эмпирических функций распределения практически не бывает точным; оно сопровождается ошибками, что приводит к неопределенности вероятностных оценок. Неточные вероятности – один из источников вероятностей второго порядка. Предположим, что у нас есть оценка вероятности P некоего события. Когда мы можем оценить границы $P_1 \leq P \leq P_2$, можно говорить об интервальной оценке вероятности P . Такие интервальные оценки для гистограмм приводят к интервальным гисто-

граммам. В тех случаях, когда для P существует вероятностная оценка, можно говорить о вероятности второго порядка. Таким образом, вероятности второго порядка – это вероятностные оценки самих вероятностей.

Для работы с неопределенными вероятностями существует несколько подходов [1]. Один из распространенных методов работы с неопределенными вероятностями – использование P-box [2]. P-box представляет собой интервальную функцию распределения ($F_1(x) \leq F(x) \leq F_2(x)$).

В работе [3] показано использование неопределенных вероятностей в инженерных расчетах. При этом предполагается, что при стохастическом моделировании для входных данных известны семейства распределений. Входные данные для моделирования выбираются случайным образом из этих семейств, далее методом Монте-Карло производится стохастическое моделирование. Важно отметить, что при работе с неопределенными вероятностями существенным препятствием является отсутствие эффективных арифметических операций.

Одним из подходов, моделирующих неопределенность вероятностных оценок, стало использование вероятностей второго порядка. В работе [4] для представления вероятностей второго порядка предложено новое понятие гистограммы второго порядка. Применение вычислительного вероятностного анализа позволило развить идею далее на основе применения кусочно-полиномиальных моделей [5, 6]. Для численных реализаций операций над распределениями второго порядка в статье рассматривается новый подход, основанный на использовании вероятностных расширений их параметризованных представлений.

Следует отметить давний и устойчивый интерес к вероятностям второго порядка в различных областях, включая принятие решений [7, 8]. Среди русскоязычных монографий, посвященных методам работы с неопределенностями, можно выделить работы О.И. Ужга-Реброва [9].

1. Вероятностные расширения

Основа вычислительного вероятностного анализа – численные арифметические операции над функциями плотностей вероятностей и вероятностные расширения (построение законов распределений функций случайных аргументов). Функции плотности вероятности случайных величин x, y, z будем обозначать полужирным фонтон $\mathbf{x}, \mathbf{y}, \mathbf{z}$. Тогда функция плотности вероятности $\mathbf{z}$ случайной величины $z = x * y$, $*$ $\in \{+, -, \cdot, /\}$, записывается в виде $\mathbf{z} = \mathbf{x} * \mathbf{y}$.

Рассмотрим задачу определения закона распределения функции нескольких случайных аргументов.

Пусть $(x_1, x_2, \dots, x_n)$ – система случайных непрерывных переменных с совместной функцией плотности вероятности $p(x_1, x_2, \dots, x_n)$. Случайная переменная z

$$z = f(x_1, x_2, \dots, x_n),$$

где функция $f : R^n \rightarrow R$.

Определение 1. Будем говорить, что случайная функция $\mathbf{f} : R^n \rightarrow R$ является вероятностным продолжением вещественной функции $f : R^n \rightarrow R$ на множестве $D \subset R^n$, если $\mathbf{f}(x) = f(x)$ для всех точечных аргументов $x \in D$.

Определение 2. Случайная функция $\mathbf{f} : R^n \rightarrow R$ называется вероятностным расширением вещественной функции $f : R^n \rightarrow R$ на множестве $D \subset R^n$, если она

(i) является вероятностным продолжением f на D ,

(ii) функция плотности вероятности $\mathbf{f}$ совпадает с функцией плотности вероятности $\mathbf{z}$ случайной величины z :

$$\mathbf{z} = \mathbf{f}(x_1, x_2, \dots, x_n).$$

Таким образом, мы можем записать

$$\mathbf{z} = \mathbf{f}(\mathbf{x}_1, \dots, \mathbf{x}_n).$$

В тех случаях, когда надо указать непосредственно значение плотности вероятности f в некоторой точке ξ , будем использовать обозначение

$$z(\xi) = f(x_1, \dots, x_n)(\xi).$$

Теорема 1 [10]. Пусть $(x_1, \dots, x_n)$ – случайные величины и $f(x_1, \dots, x_n)$ – вероятностное расширение $f(x_1, x_2, \dots, x_n)$, и для всех вещественных t функция $f(t, x_2, \dots, x_n)$ – вероятностное расширение $f(t, x_2, \dots, x_n)$. Тогда

$$f(x_1, \dots, x_n)(\xi) = \int_{x_1}^{\bar{x}_1} x_1(t) f(t, x_2, \dots, x_n)(\xi) dt.$$

2. Распределения второго порядка

В вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$ случайный процесс представляет собой набор случайных величин

$$\{a(x, \omega), x \in D, \omega \in \Omega\}.$$

Термин «случайное поле» обычно относится к случайному процессу, принимающему значения в евклидовом пространстве R^d , $d = 1, 2, 3$. Случайное поле можно посмотреть двумя способами:

- для фиксированного $x \in D$, $a(x, \cdot)$ является случайной величиной в Ω ;
- для фиксированного $\omega \in \Omega$, $a(\cdot, \omega)$ является реализацией случайного поля в D .

Определение 3. Распределение второго порядка $f^{(2)}$ – случайное поле $f(x, \omega)$, $x \in D, \omega \in \Omega$, заданное на $D \subset R$, где $(\Omega, \mathcal{F}, \mathbb{P})$ – вероятностное пространство. Обладает следующими свойствами: для фиксированного $\omega \in \Omega$, $f(\cdot, \omega)$ является функцией распределения f_ω .

В тех случаях, когда для распределений второго порядка f_i , $f_i(\cdot, \omega)$ – функции плотности вероятности, для распределений второго порядка можно определить арифметические операции.

Определение 4. Операции над распределениями второго порядка. Пусть $f^{(2)}, g^{(2)}$ – распределения второго порядка, $(\Omega_f, \mathcal{F}_f, \mathbb{P}_f), (\Omega_g, \mathcal{F}_g, \mathbb{P}_g)$ – соответственно их вероятностные пространства. Тогда результат операции $f^{(2)} * g^{(2)}$, $*$ $\in \{+, -, \cdot, /\}$, – распределение второго порядка $F^{(2)}$:

$$F(\cdot, \omega_*) = \{f(\cdot, \omega_f) * g(\cdot, \omega_g) | (\omega_f, \omega_g) \in \Omega_*\},$$

где $(\Omega_*, \mathcal{F}_*, \mathbb{P}_*)$ – вероятностное пространство, $\Omega_* = \Omega_f \times \Omega_g$.

Рассмотрим пример построения кусочно-полиномиального представления для распределения второго порядка. Пусть $(\xi_1, \dots, \xi_n)$, $n = 9$, – выборка случайной величины X с функцией распределения $F(t), t \in [0, 2]$. Далее $z_i = F(\xi_i), i = 1, \dots, n$. Заметим, что $z_i, i = 1, \dots, n$ – равномерно распределенные случайные величины на отрезке $[0, 1]$. Если $z_1 \leq z_2 \leq \dots \leq z_n$, тогда z_k – k -я порядковая статистика и математическое ожидание $[Ez_k] = k/(n+1)$.

На рис. 1 представлена функция распределения второго порядка, аппроксимирующая распределение кусочно-линейных интерполяций распределений Ирвина–Холла третьей степени, построенная на выборке случайной величины размерности 9. Голубые линии – плотности вероятности случайной кусочно-линейной функции. Красная линия – точная функция распределения. Зеленые линии – границы 95%-ной доверительной области.

Функция плотности вероятности есть производная от функции распределения

$$f(x) = F'(x),$$

следовательно, производная от функции распределения второго порядка будет функцией плотности вероятности второго порядка.

Производная от кусочно-линейной функции – кусочно-постоянная функция. На рис. 2 приведен пример кусочно-постоянной функции – функции плотности вероятности второго порядка (надежной

оценки ф.п.в.). Красная линия – точная функция плотности вероятности. Синие линии – функции плотности вероятности надежной оценки, построенной на выборке случайной величины размерности 9.

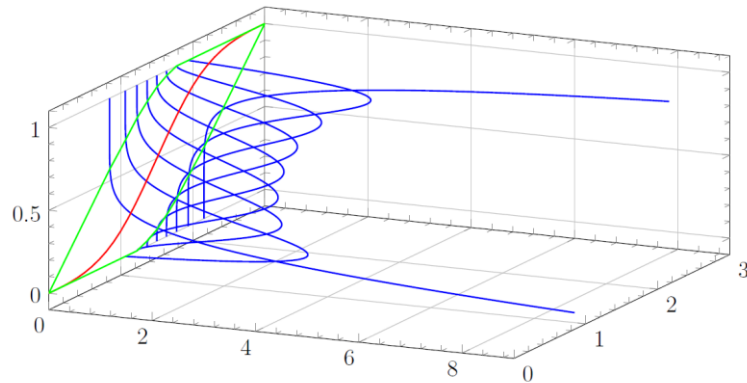


Рис. 1. Распределение второго порядка, аппроксимирующее распределение Ирвина–Холла третьей степени
Fig.1. Second order distribution, approximating the Irwin-Hall distribution of the third degree

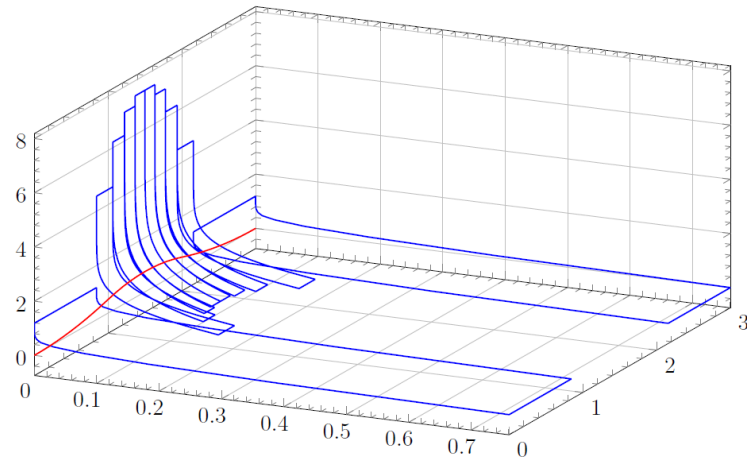


Рис. 2. Функция плотности вероятности второго порядка
Fig. 2. Second order probability density function

Далее будем использовать эрмитовы кубические сплайны S и точки $(x_i, i/(n+1))$ для построения аппроксимации функции распределения F :

$$F(x) = \begin{cases} x^2 / 2, & x \in [0, 1], \\ x(2 - x / 2) - 1, & x \in [1, 2], \end{cases}$$

$$S(0) = 0, \quad S'(0) = 0, \quad S(1) = 1, \quad S'(1) = 0,$$

$$S(x) = fv(x-1) + mw(x-1) + v(x-2).$$

$$v(x) = (|x| - 1)^2(2|x| + 1), \quad w(x) = x(|x| - 1)^2,$$

$$\sum_{i=1}^n (s(\xi_i) - z_i)^2 + \alpha \|s''\|_2^2 \rightarrow \min,$$

$$Ay = b,$$

где вектор $y = (f, m)^T$. Значение параметра α выбиралось таким образом, чтобы константы f, m определяли монотонный сплайн; легко видеть, что точка (f, m) должна лежать внутри треугольной области с вершиной в точке $(0.5; 1.5)$ и основанием $[0; 1]$. На рис. 3 приведены функции плотности вероятности констант m и f . Распределение второго порядка S можно представить в параметризованном виде

$$S(x, f, m) = fv(x-1) + mw(x-1) + v(x-2).$$

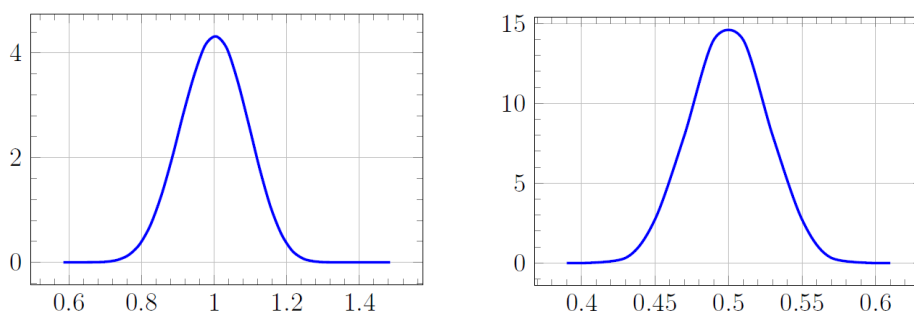


Рис. 3. Функции плотности вероятности констант m и f
Fig. 3. Probability density functions of constants m and f

Для нахождения функции плотности вероятности распределения второго порядка продифференцируем S :

$$S'(x, f, m) = fv'(x-1) + mw'(x-1) + v'(x-2).$$

На рис. 4 представлена функция плотности вероятности второго порядка. Красная линия – математическое ожидание ES , зеленые линии – верхние и нижние границы доверительного интервала, синие линии – функции плотности вероятности при фиксированных значениях аргумента.

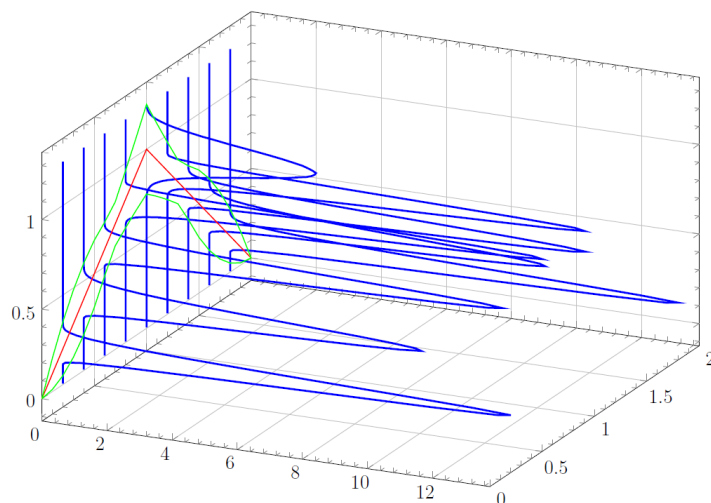


Рис. 4. Функция плотности вероятности второго порядка
Fig. 4. Second order probability density function

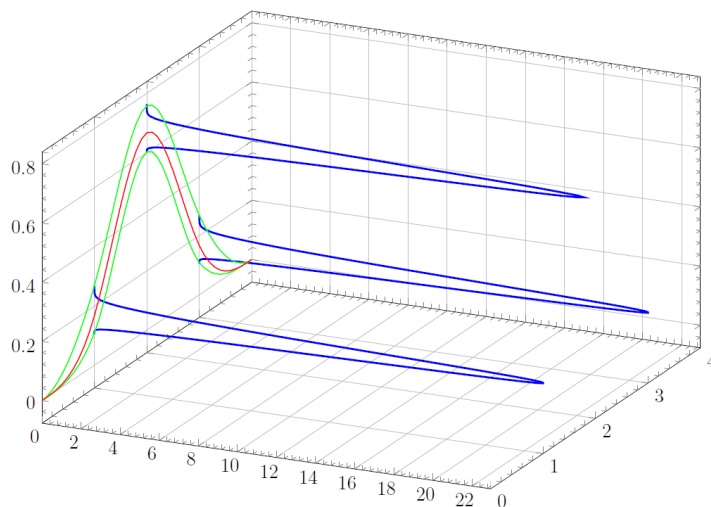


Рис. 5. Сумма распределений второго порядка
Fig. 5. The sum of the second order distributions

В силу того, что распределения второго порядка мы представляем в параметризованном виде, для операций над распределениями второго порядка воспользуемся техникой вероятностных расширений. Таким образом, арифметические операции можно представить как вероятностные расширения соответствующих функций от параметров. На рис. 5 показан результат сложения двух распределений второго порядка, представленных случайными эрмитовыми сплайнами [6]. Красная линия – математическое ожидание распределения второго порядка, зеленые линии – верхние и нижние 95%-ные доверительные границы, синие линии – функции плотности вероятности при фиксированных значениях аргумента.

3. Достоверные оценки показателей надежности оборудования

В качестве примера использования распределений второго порядка рассмотрим построение достоверных оценок показателей надежности оборудования в условиях малых выборок статистических данных об отказах. Подход основан на использовании порядковых статистик и случайных интерполяционных многочленов для построения достоверных оценок функций распределения.

Вероятность безотказной работы $P(t)$ – это вероятность того, что в течение указанного времени работы не произойдет отказа. Время работы – это продолжительность или объем работы. Частота отказов – мера отказов за единицу времени. Частота отказов зависит от распределения отказов, которое представляет собой совокупную функцию распределения, описывающую вероятность отказов до момента времени t :

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq \xi < t + \Delta t | t \leq \xi)}{\Delta t} = \frac{f(t)}{P(t)} = \frac{f(t)}{1 - F(t)}.$$

Заметим, что $f(t) = P'(t)$ и

$$\lambda(t) = -\frac{P'(t)}{P(t)},$$

где $P(t)$ – вероятность безотказной работы.

Пусть $(\xi_1, \xi_2, \dots, \xi_n)$ – статистика отказов, полученная опытным путем. Тогда

$$-\ln(z_i) = \int_0^{\xi_i} \lambda(\xi) d\xi,$$

где $z_i = P(\xi_i)$.

Для нахождения $\lambda(t)$ будем использовать метод наименьших квадратов. Представим аппроксимацию $\lambda(t)$ в виде случайной кусочно-линейной функции $l(t)$

$$\lambda(t) \approx l(t) = \sum_{i=1}^m a_i \psi_i(t), t \in [0, T],$$

где $\{\psi_i, i = 1, \dots, m\}$ – базис в пространстве кусочно-линейных функций. Следуя работе [5], получаем приближение

$$\lambda(t) \approx \sum_{l=1}^n \left(\sum_{j=1}^m \psi_j(t) \right) \gamma_l \ln(z_l).$$

Поскольку известна совместная функция плотности вероятности $p(z_1, z_2, \dots, z_n)$, можем построить вероятностное расширение $l(t)$:

$$l(t) = \sum_{l=1}^n \left(\sum_{j=1}^m \psi_j(t) \right) \gamma_l \ln(z_l).$$

Пример. Имеем $(\xi_1, \xi_2, \dots, \xi_n)$, $n = 29$. Предположим, $\lambda(t)$ имеет вид:

$$\lambda(t) = \begin{cases} 0,2, & t \in [0; 0,7], \\ 0,2 + 12(t - 0,7) & t > 0,7. \end{cases}$$

На рис. 6 приведена достоверная оценка интенсивности отказов модельной задачи. Оттенками серого показаны плотности вероятности распределения второго порядка. На рис. 6 показана оценка функций плотности вероятности $P(t)$ в момент времени $t = 1$.

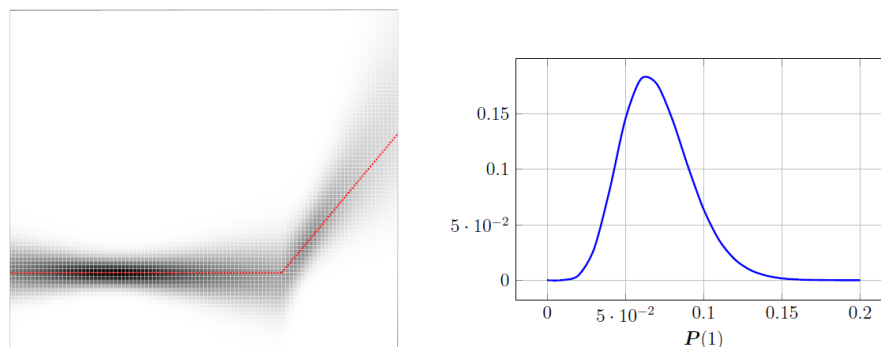


Рис. 6. Достоверная оценка интенсивности отказов и вероятности безотказной работы

Fig. 6. Reliable assessment of failure rate and probability of failure-free operation

Используя вероятностное расширение λ , можно вычислить значения оценки функции плотности вероятности $P(t)$ в любое время. Используя оценки функции плотности вероятности $P(t)$ в виде сплайнов, мы можем, например, оценить риск того, что вероятность $P(1) > 0,2$ или $P(1) < 0,05$.

Заключение

Применение кусочно-полиномиальных моделей распределений второго порядка позволяет строить надежные оценки эмпирических функций распределений. Рассмотренные примеры численных операций над распределениями второго порядка в задачах построения оценок надежности оборудования в условиях малых выборок подтверждают этот вывод. Дальнейшее использование распределений второго порядка может быть направлено на оценки рисков, принятие решений, стохастическое моделирование в условиях эпистемической неопределенности.

Список источников

1. Augustin T., Coolen F., Cooman G., Troffaes M. Introduction to Imprecise Probabilities. John Wiley & Sons, 2014. 448 p.
2. Ferson S., Kreinovich V., Ginzburg L., Myers D.S., Sentz K. Constructing Probability Boxes and Dempster–Shafer Structures. Sandia National Laboratories, SAND2002-4015, 2003.
3. Swiler L.P., Giunta A.A. Aleatory and Epistemic Uncertainty Quantification for Engineering Applications. Sandia Technical Report, SAND2007-2670C, 2007.
4. Добронец Б.С., Попова О.А. Численный вероятностный анализ для исследования систем в условиях неопределенности // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2012. № 4 (21). С. 39–46.
5. Добронец Б.С., Попова О.А. Вычислительный вероятностный анализ: модели и методы. Красноярск : Сиб. федер. ун-т, Ин-т космических и информ. технологий, 2020. 236 с.
6. Попова О.А. Применение численного вероятностного анализа в задачах интерполяции // Вычислительные технологии. 2017. Т. 22, № 2. С. 99–114.
7. Sahlin N., Goldsmith R. The role of second-order probabilities in decision making // Advances in Psychology. 1983. V. 14. P. 455–467.
8. Utkin L., Augustin T. Decision Making with Imprecise Second-Order Probabilities // Proc. of the Third International Symposium on Imprecise Probabilities and Their Applications, Lugano, Switzerland, July 14–17, 2003. P. 547–561.
9. Ужга-Ребров О.И. Управление неопределенностями. Rēzekne, 2014. Ч. 4: Комбинирование неопределенностей. 408 с.
10. Dobronets B.S., Popova O.A. Computational Aspects of Probabilistic Extensions // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2019. № 47. С. 41–48.

References

1. Augustin, T., Coolen, F., Cooman, G. & Troffaes, M. (2014) *Introduction to Imprecise Probabilities*. John Wiley & Sons.
2. Ferson, S., Kreinovich, V., Ginzburg, L., Myers, D.S. & Sentz, K. (2003) *Constructing Probability Boxes and Dempster–Shafer Structures*. Sandia National Laboratories, SAND2002-4015.

3. Swiler, L.P. & Giunta, A.A. (2007) *Aleatory and Epistemic Uncertainty Quantification for Engineering Applications*. Sandia Technical Report, SAND2007-2670C
4. Dobronets, B.S. & Popova, O.A. (2012) Numerical probabilistic analysis for the study of systems with uncertainty. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitel'naya tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 4(21). pp. 39–46.
5. Dobronets, B.S. & Popova, O.A. (2020) *Vychislitel'nyy veroyatnostnyy analiz: modeli i metody* [Computational Probabilistic Analysis: Models and Methods]. Krasnoyarsk: Siberian Federal University, Institute of Space and Information Technologies.
6. Popova, O.A. (2017) Application of numerical probabilistic analysis in interpolation problems. *Vychislitel'nye tekhnologii – Computational Technologies*. 22(2). pp. 99–114.
7. Sahlin, N. & Goldsmith, R. (1983) The role of second-order probabilities in decision making. *Advances in Psychology*. 14. pp 455–467. DOI: 10.1016/S0166-4115(08)62250-5
8. Utkin, L. & Augustin, T. (2003) Decision Making with Imprecise Second-Order Probabilities. *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications*. Lugano, Switzerland, July 14–17. pp. 547–561.
9. Uzhga-Rebrov, O. (2014) *Upravlenie neopredelennostyami* [Uncertainty Management]. Vol. 4. Rezekne: [s.n.].
10. Dobronets, B.S. & Popova, O.A. (2019) Computational aspects of probabilistic extensions. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitel'naya tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 47. pp. 41–48. DOI: 10.17223/19988605/47/5

Информация об авторах:

Добронет Борис Станиславович – профессор, доктор физико-математических наук, профессор кафедры систем искусственного интеллекта Института космических и информационных технологий Сибирского федерального университета (Красноярск, Россия). E-mail: BDobronets@yandex.ru

Попова Ольга Аркадьевна – доцент, кандидат технических наук, доцент кафедры систем искусственного интеллекта Института космических и информационных технологий Сибирского федерального университета (Красноярск, Россия). E-mail: OlgaArc@yandex.ru

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Information about the authors:

Dobronets Boris S. (Doctor of Physics and Mathematics, Professor, Siberian Federal University, Institute of Space and Information Technologies, Krasnoyarsk, Russian Federation). E-mail: BDobronets@yandex.ru

Popova Olga A. (Candidate of Technical Sciences, Associate Professor, Siberian Federal University, Institute of Space and Information Technologies, Krasnoyarsk, Russian Federation). E-mail: OlgaArc@yandex.ru

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Поступила в редакцию 07.07.2021; принята к публикации 29.11.2022

Received 07.07.2021; accepted for publication 29.11.2022