

УДК 519.7

СРАВНЕНИЕ МЕТОДОВ ОПТИМИЗАЦИИ РЕГРЕССИОННОЙ ОЦЕНКИ ПЛОТНОСТИ ВЕРОЯТНОСТИ ОДНОМЕРНОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

© А. В. Лапко^{1,2}, В. А. Лапко^{1,2}

¹Институт вычислительного моделирования СО РАН,
660036, г. Красноярск, Академгородок, 50, стр. 44

²Сибирский государственный университет науки и технологий им. академика
М. Ф. Решетнева,
660037, г. Красноярск, просп. им. газеты «Красноярский рабочий», 31
E-mail: lapko@icm.krasn.ru

Исследуются методы выбора коэффициента размытости ядерных функций регрессионной оценки плотности вероятности одномерной случайной величины. Регрессионная оценка плотности вероятности является модификацией статистики Розенблатта — Парзена и используется при обработке статистических данных большого объёма. Её синтез основан на сжатии исходной выборки путём декомпозиции области значений случайной величины. Элементами полученного массива данных являются центры интервалов дискретизации и частоты принадлежности им значений случайной величины из исходной выборки. Эта информация является достаточной для оценивания плотности вероятности случайной величины в виде непараметрической регрессии. Поэтому появляется возможность выбора коэффициента размытости ядерных функций регрессионной оценки из условия минимума её ошибки аппроксимации искомой плотности вероятности. Традиционный подход оптимизации непараметрической оценки плотности вероятности основан на минимизации её среднего квадратического отклонения. Анализируются аппроксимационные свойства регрессионной оценки плотности вероятности при использовании рассматриваемых методов её оптимизации.

Ключевые слова: регрессионная оценка плотности вероятности, одномерная случайная величина, ядерная оценка плотности вероятности, выбор коэффициентов размытости, формула Старджесса, формула Хайнкольда и Гаеде, выборки большого объёма.

DOI: 10.15372/AUT20240309

EDN: PRRFGH

Введение. Непараметрические оценки плотности вероятности типа Розенблатта — Парзена широко применяются при синтезе алгоритмов принятия решений в условиях априорной неопределённости. Вычислительная эффективность этих оценок плотностей вероятностей снижается при увеличении объёма исходной статистической информации. Для решения данной проблемы используются принципы декомпозиции статистических данных по их объёму и технологии параллельных вычислений [1].

Перспективное направление обхода проблемы больших объёмов выборок связано со сжатием исходной статистической информации на основе дискретизации области значений случайных величин. Полученный массив данных составляют центры интервалов дискретизации и частоты принадлежности им исходных статистических данных. Эта информация является достаточной для синтеза регрессионной оценки плотности вероятности, которая позволяет осуществлять выбор её коэффициента размытости ядерных функций с использованием двух подходов. Первый подход основан на традиционной методике оптимизации непараметрической оценки плотности вероятности ядерного типа из условия минимума её среднего квадратического отклонения [1–5]. Во втором подходе допускается выбор коэффициентов размытости ядерных функций регрессионной оценки плотности

вероятности с позиций оптимизации непараметрической регрессии на основе полученного массива сжатых исходных статистических данных.

Цель работы состоит в исследовании эффективности традиционного и нового методов оптимизации регрессионной оценки плотности вероятности на основе результатов вычислительных экспериментов при использовании известных методов дискретизации значений одномерной случайной величины.

Регрессионная оценка плотности вероятности одномерной случайной величины. Пусть имеется выборка $V = (x^i, i = \overline{1, n})$, составленная из n независимых наблюдений случайной величины x с неизвестной плотностью вероятности $p(x)$. Предполагается, что $p(x)$ допускает разложение в ряд Тейлора в каждой точке x .

Приведём методику синтеза регрессионной оценки плотности вероятности:

1. Разбить область определения $p(x)$ на N равных непересекающихся интервалов длиной 2β . Для дискретизации области значений случайной величины x могут быть использованы, например, формула Хайнкольда — Гаеде [6]

$$\bar{N}_1 = \sqrt{n}, \quad (1)$$

формула Старджесса [7]

$$\bar{N}_2 = \log_2 n + 1, \quad (2)$$

формула Брукса — Каррузера [8]

$$\bar{N}_3 = 5 \lg n. \quad (3)$$

В работе [9] при анализе аппроксимационных свойств непараметрической оценки плотности вероятности ядерного типа, восстанавливаемой по статистическим данным объёма n , получена процедура выбора оптимального количества интервалов дискретизации:

$$\bar{N}_4 = \sqrt{\Delta \int_{-\infty}^{\infty} p^2(x) dx} \quad n = k_1 \sqrt{n} \quad (4)$$

области Δ значений одномерной случайной величины x с плотностью вероятности $p(x)$. Количество интервалов дискретизации N определяется как ближайшее целое число к $\bar{N}_1, \bar{N}_2, \bar{N}_3, \bar{N}_4$ при использовании одной из вышеприведённых формул. Аналитическая зависимость (4) по виду близка к формуле Хайнкольда — Гаеде (1) и совпадает с ней при оценивании плотности вероятности случайной величины с равномерным законом распределения. Установлено, что коэффициент k_1 определяется только видом плотности вероятности и не зависит от её параметров. Методика оценивания коэффициента k_1 по исходным статистическим данным V предложена в работе [10].

2. Определить оценки вероятностей принадлежности случайной величины j -му интервалу

$$\bar{P}^j = n_j/n, \quad j = \overline{1, N},$$

где $n_j, j = \overline{1, N}$ — количество значений случайной величины x из исходной выборки V в j -м интервале дискретизации.

3. Предположим, что в каждом j -м интервале имеет место равномерный закон распределения случайной величины x . Тогда оценка плотности вероятности в j -м интервале определяется значением

$$\bar{p}^j = \bar{P}^j/(2\beta), \quad j = \overline{1, N}.$$

4. Сформировать массив данных на основе информации из пунктов 2, 3:

$$\bar{V} = (z^j, \bar{p}^j, j = \overline{1, N}).$$

5. Восстановить плотность вероятности $p(x)$ по полученной информации как непараметрическую регрессию [11]

$$\bar{p}(x) = \frac{1}{c} \sum_{j=1}^N \bar{P}^j \Phi\left(\frac{x - z^j}{c}\right). \quad (5)$$

Здесь ядерные функции $\Phi(u)$ удовлетворяют условиям [12, 13]

$$\Phi(u) = \Phi(-u), \quad 0 \leq \Phi(u) < \infty, \quad \int_{-\infty}^{\infty} \Phi(u) du = 1, \quad \int_{-\infty}^{\infty} u^2 \Phi(u) du = 1,$$

$$\int_{-\infty}^{\infty} u^m \Phi(u) du < \infty, \quad 0 \leq m < \infty.$$

Нетрудно заметить, что статистика (5) обладает свойствами плотности вероятности. Она является асимптотически несмещённой и состоятельной оценкой плотности вероятности $p(x)$ [11].

Методы оптимизации регрессионной оценки плотности вероятности. Традиционный метод выбора коэффициентов размытости ядерных функций непараметрической оценки плотности вероятности $\bar{p}(x)$ основан на минимизации квадратического критерия

$$W(c) = \int_{-\infty}^{\infty} (\bar{p}(x) - p(x))^2 dx. \quad (6)$$

Нетрудно заметить, что этот критерий при оптимизации $\bar{p}(x)$ допускает преобразование к виду

$$\bar{W}(c) = \int_{-\infty}^{\infty} \bar{p}^2(x) dx - 2 \int_{-\infty}^{\infty} \bar{p}(x) p(x) dx, \quad (7)$$

так как $\int_{-\infty}^{\infty} p^2(x) dx$ не зависит от коэффициента размытости c ядерных функций. Интеграл во втором слагаемом в (7) является математическим ожиданием статистики $\bar{p}(x)$. Тогда критерий выбора оптимального коэффициента размытости статистики (5) после несложных преобразований запишется как

$$\bar{W}_1(c) = \frac{1}{c^2} \sum_{i=1}^N \sum_{j=1}^N \bar{P}^i \bar{P}^j \int_{-\infty}^{\infty} \Phi\left(\frac{x - z^i}{c}\right) \Phi\left(\frac{x - z^j}{c}\right) dx - \frac{2}{c} \sum_{i=1}^N \bar{P}^i \sum_{j=1, j \neq i}^N \bar{P}^j \Phi\left(\frac{z^i - z^j}{c}\right). \quad (8)$$

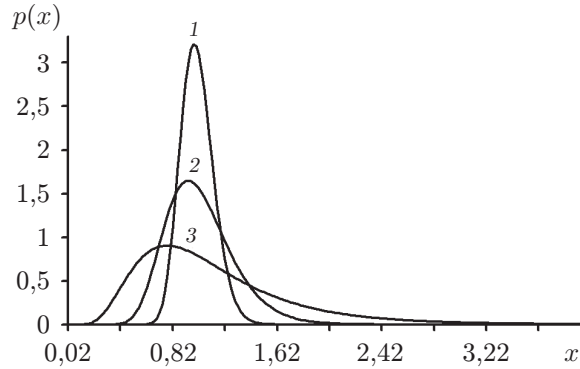


Рис. 1. Вид плотности вероятности $p(x)$ (10) одномерной случайной величины x при $a = 0$ и $s = 0,125$ (кривая 1), $s = 0,25$ (2), $s = 0,5$ (3)

Второй метод оптимизации регрессионной оценки плотности вероятности (5) с использованием массива данных $\bar{V} = (z^j, \bar{p}^j, j = \overline{1, N})$ основан на минимизации по коэффициенту размытости c критерия

$$\tilde{W}_1(c) = \frac{1}{N} \sum_{i=1}^N (\bar{p}^i - \bar{p}(x^i))^2, \quad (9)$$

где значение

$$\bar{p}(x^i) = \frac{1}{c} \sum_{j=1, j \neq i}^N \bar{p}^j \Phi\left(\frac{x^i - z^j}{c}\right).$$

Ситуация $(x^i, \bar{p}^i)$ из массива данных $\bar{V}$, которая подаётся на контроль при вычислении критерия (9), исключается из процесса формирования статистики (5).

Анализ результатов вычислительного эксперимента. Исследуем эффективность рассматриваемых методов оптимизации регрессионной оценки плотности вероятности (5) при восстановлении логнормальных законов распределения одномерной случайной величины [14]

$$p(x) = \frac{1}{sx\sqrt{2\pi}} \exp\left(-\frac{(\ln x - a)^2}{2s^2}\right) \quad (10)$$

при различных значениях параметра s , объёма n исходных статистических данных и методов дискретизации (1), (2), (3). В (10) параметры a и s меняют вид рассматриваемой плотности вероятности $p(x)$ (рис. 1). Подобный закон распределения, например, характерен для спектральных признаков данных дистанционного зондирования природных объектов.

Имеется выборка $V = (x^i, i = \overline{1, n})$ значений одномерной случайной величины объёма n с плотностью вероятности (10). При оценивании $p(x)$ статистикой (5) используется оптимальная ядерная функция В. А. Епанечникова

$$\Phi(u) = \begin{cases} \frac{3}{4\sqrt{5}} - \frac{3}{20\sqrt{5}} u^2, & \text{если } |u| < \sqrt{5}; \\ 0, & \text{если } |u| \geq \sqrt{5}, \end{cases}$$

которая минимизирует среднее квадратическое отклонение непараметрической оценки плотности вероятности Розенблатта — Парзена от $p(x)$ [13].

Таблица 1

Зависимость оценок ошибок аппроксимации плотности вероятности $p(x)$ (10) при $a = 0$ и $s = 0,125$

n	N	$\bar{c}_1$	$\bar{c}_2$	Δ	$\bar{R}(\bar{c}_1)$	$\bar{R}(\bar{c}_2)$
1000	32	0,068	0,061	0,0252	0,0879	0,0602
5000	71	0,053	0,048	0,0128	0,0371	0,0267
10000	100	0,049	0,041	0,0094	0,0278	0,0169
15000	122	0,046	0,041	0,0079	0,0236	0,0160
20000	141	0,042	0,040	0,0070	0,0173	0,0150
25000	158	0,040	0,040	0,0063	0,0142	0,0140
30000	173	0,040	0,036	0,0058	0,0140	0,0108
50000	224	0,040	0,031	0,0046	0,0138	0,0070

Таблица 2

Зависимость оценок ошибок аппроксимации плотности вероятности $p(x)$ (10) при $a = 0$ и $s = 0,25$

n	N	$\bar{c}_1$	$\bar{c}_2$	Δ	$\bar{R}(\bar{c}_1)$	$\bar{R}(\bar{c}_2)$
1000	32	0,139	0,118	0,0540	0,0277	0,0182
5000	71	0,107	0,093	0,0276	0,0110	0,0075
10000	100	0,095	0,085	0,0206	0,0074	0,0058
15000	122	0,090	0,081	0,0175	0,0066	0,0052
20000	141	0,087	0,081	0,0157	0,0055	0,0050
25000	158	0,084	0,079	0,0140	0,0049	0,0047
30000	173	0,081	0,078	0,0130	0,0043	0,0041
50000	224	0,074	0,069	0,0104	0,0034	0,0031

Объём n исходных статистических данных выбирался из интервала $n \in [1000; 50000]$. В этих условиях определялись регрессионные оценки плотности вероятности (5) при оптимальных коэффициентах размытости ядерных функций $\bar{c}_1$, $\bar{c}_2$, минимизирующие критерии (7) и (9) соответственно. При синтезе регрессионной оценки плотности вероятности использовались формулы дискретизации (1), (2), (3) области значений случайной величины.

Критерии аппроксимации плотности вероятности (10) статистикой (5) при оптимальных значениях коэффициентов размытости $\bar{c}_1$, $\bar{c}_2$ запишем в виде

$$R(\bar{c}_j) = \frac{1}{n} \sum_{i=1}^n (p(x^i) - \bar{p}(x^i))^2, \quad j = 1, 2.$$

При конкретном объёме n определялись значения оптимальных коэффициентов размытости $\bar{c}_1$, $\bar{c}_2$ и соответствующие им значения $R(\bar{c}_1)$, $R(\bar{c}_2)$ вычислялись 50 раз. Полученные результаты вычислительного эксперимента усреднялись. Обозначим средние значения коэффициентов размытости $\bar{c}_1$, $\bar{c}_2$ и критериев $R(\bar{c}_1)$, $R(\bar{c}_2)$ через $\bar{\bar{c}}_1$, $\bar{\bar{c}}_2$, $\bar{\bar{R}}(\bar{\bar{c}}_1)$, $\bar{\bar{R}}(\bar{\bar{c}}_2)$.

Результаты вычислительных экспериментов при синтезе статистики (5) с использованием формулы дискретизации Хайнкольда и Гаеде (1) приведены в табл. 1, 2, 3, которые отличаются значениями параметра s логнормального закона распределения (10).

Метод оптимизации регрессионной оценки плотности вероятности $p(x)$ (10) при значениях параметра $s = 0,125$, $s = 0,25$ из условия минимума критерия (9) по коэффициенту размытости s обладает преимуществом по сравнению с традиционным методом [1–5].

Таблица 3

**Зависимость оценок ошибок аппроксимации плотности вероятности $p(x)$ (10)
при $a = 0$ и $s = 0,5$**

n	N	$\bar{c}_1$	$\bar{c}_2$	Δ	$\bar{R}(\bar{c}_1)$	$\bar{R}(\bar{c}_2)$
1000	32	0,266	0,238	0,1454	0,0130	0,0098
5000	71	0,211	0,209	0,0801	0,0063	0,0069
10000	100	0,190	0,207	0,0605	0,0047	0,0071
15000	122	0,179	0,195	0,0521	0,0038	0,0062
20000	141	0,175	0,211	0,0482	0,0035	0,0078
25000	158	0,165	0,201	0,0424	0,0029	0,0067
30000	173	0,162	0,212	0,0399	0,0027	0,0078
50000	224	0,147	0,187	0,0316	0,0019	0,0056

Этот вывод сохраняется для объёма исходных статистических данных $n \in [1000; 50000]$ в условиях использования формулы дискретизации области значений случайной величины (1). Например, при $s = 0,125$ и $n = 1000$ значения критериев $\bar{R}(\bar{c}_1) = 0,0879$, а $\bar{R}(\bar{c}_2) = 0,0602$. При увеличении объёма статистических данных до $n = 50000$ значение критериев $\bar{R}(\bar{c}_1) = 0,0138$, $\bar{R}(\bar{c}_2) = 0,007$. В этих условиях количество интервалов дискретизации области значений случайной величины $N \in [32; 224]$. Причём средние значения оптимальных коэффициентов размытости $\bar{c}_1 \in [0,068; 0,04]$, а $\bar{c}_2 \in [0,061; 0,031]$ (см. табл. 1). С ростом объёма n исходных статистических данных и увеличением количества интервалов дискретизации значения оптимальных коэффициентов размытости ядерных функций статистики (5) уменьшаются, что согласуется с результатами исследования асимптотических свойств регрессионной оценки плотности вероятности [11]. Этот вывод сохраняется и в последующем анализе результатов вычислительных экспериментов при $s = 0,25$ и $s = 0,5$.

По аналогии приведём количественные характеристики рассматриваемых методов оптимизации статистики (5) при значениях параметра плотности вероятности $s = 0,25$. Для $n = 1000$ значения критериев $\bar{R}(\bar{c}_1) = 0,0277$, а $\bar{R}(\bar{c}_2) = 0,0182$. С увеличением объёма статистических данных до $n = 50000$ значения критериев $\bar{R}(\bar{c}_1) = 0,0034$, а $\bar{R}(\bar{c}_2) = 0,0031$. Оптимальные значения коэффициентов размытости $\bar{c}_1 \in [0,139; 0,074]$, а $\bar{c}_2 \in [0,118; 0,069]$ (см. табл. 2).

Если плотность вероятности $p(x)$ (10) характеризуется параметром $s = 0,5$, то при $n = 1000$ значения критериев $\bar{R}(\bar{c}_1) = 0,013$, $\bar{R}(\bar{c}_2) = 0,0098$. При $n = 5000$ значения $\bar{R}(\bar{c}_1) = 0,0063$, $\bar{R}(\bar{c}_2) = 0,0069$, что определяет преимущество традиционного метода оптимизации по сравнению с методикой, основанной на минимизации критерия (9). Эта тенденция наблюдается и при значениях $n > 5000$. Например, при $n = 50000$ значения критериев $\bar{R}(\bar{c}_1) = 0,0019$, а $\bar{R}(\bar{c}_2) = 0,0056$. Этот результат можно объяснить ростом асимметрии $p(x)$ с увеличением её параметра s и сокращением области значений случайной величины со сложным видом плотности вероятности (см. рис. 1). Иллюстрации обнаруженных закономерностей при анализе табл. 1, 2, 3 представлены на рис. 2, 3, 4.

Сравним эффективность методов оптимизации регрессионной оценки плотности вероятности при использовании различных методов дискретизации области значений случайной величины. Пусть плотность вероятности $p(x)$ характеризуется параметром $s = 0,125$, при котором она близка к нормальному закону распределения.

В табл. 4 приведены средние значения оценок ошибок аппроксимации $\bar{R}(\bar{c}_1)$, $\bar{R}(\bar{c}_2)$ регрессионной оценкой плотности вероятности $p(x)$ при использовании формул дискретизации Хайнкольда — Гаеде (1), Старджесса (2), Брукса — Каррузера (3).

Нетрудно заметить, что использование формул дискретизации Хайнкольда — Гаеде при синтезе регрессионной оценки плотности вероятности (5) обеспечивает значительно

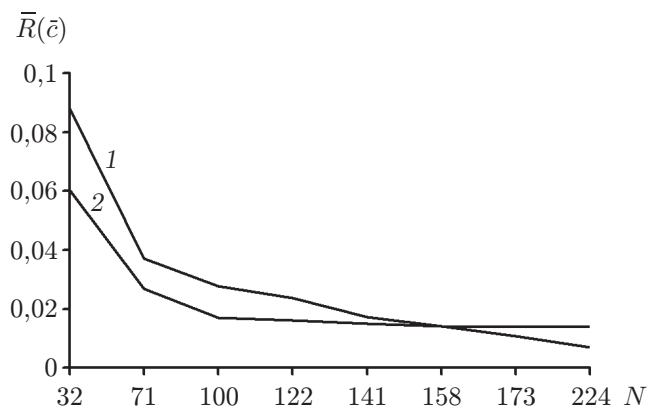


Рис. 2. Зависимость критериев оценки ошибок аппроксимации $p(x)$ (10) при $a = 0$ и $s = 0,125$. Кривая 1 соответствует традиционному методу оптимизации $\bar{p}(x)$, а 2 — предлагаемому

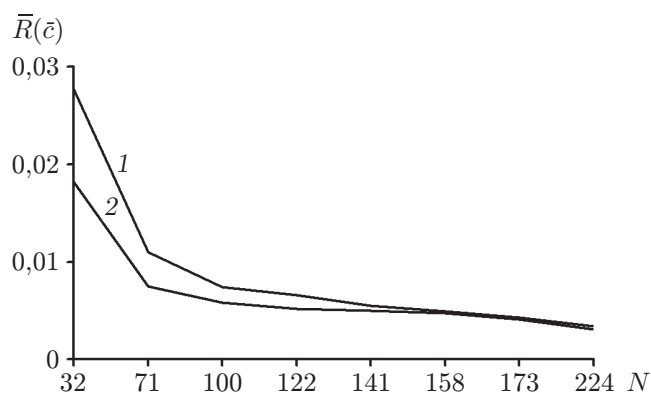


Рис. 3. Зависимость критериев оценки ошибок аппроксимации $p(x)$ (10) при $a = 0$ и $s = 0,25$. Кривая 1 соответствует традиционному методу оптимизации $\bar{p}(x)$, а 2 — предлагаемому

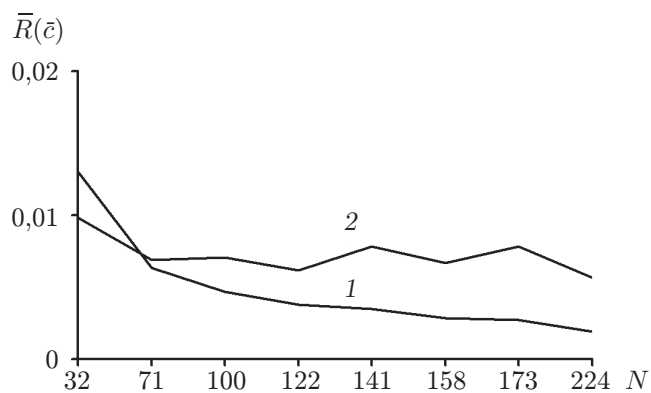


Рис. 4. Зависимость критериев оценки ошибок аппроксимации $p(x)$ (10) при $a = 0$ и $s = 0,5$. Кривая 1 соответствует традиционному методу оптимизации $\bar{p}(x)$, а 2 — предлагаемому

Таблица 4

Зависимость оценок ошибок аппроксимации плотности вероятности $p(x)$ (10) при $a = 0$ и $s = 0,125$

n	N для формул дискретизации			$\bar{R}(\bar{c}_1)$			$\bar{R}(\bar{c}_2)$		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
1000	32	11	15	0,0879	0,2250	0,1695	0,0602	0,1750	0,1441
5000	71	13	18	0,0371	0,1727	0,1435	0,0267	0,1552	0,1271
10000	100	14	20	0,0278	0,1862	0,1287	0,0169	0,1408	0,1139
15000	122	15	21	0,0236	0,2564	0,1374	0,0160	0,1307	0,1036
20000	141	15	22	0,0173	0,2661	0,1325	0,0150	0,1299	0,1035
25000	158	16	22	0,0142	0,2403	0,1370	0,0143	0,1241	0,0984
30000	173	16	22	0,0140	0,2476	0,1426	0,0108	0,1155	0,0963
50000	224	17	23	0,0138	0,2215	0,1590	0,0070	0,1153	0,0906

меньшую ошибку аппроксимации плотности вероятности $p(x)$ (10) по сравнению с методами дискретизации Старджесса и Брукса — Каррузера в рассматриваемых условиях вычислительных экспериментов. Причём формула дискретизации Старджесса менее эффективна по сравнению с формулой Брукса — Каррузера. Обнаруженная закономерность сохраняется для традиционного и нового методов оптимизации регрессионной оценки плотности вероятности. Например, при использовании традиционного метода оптимизации и $n = 1000$ применение формулы дискретизации (1) обеспечивает значение $\bar{R}(\bar{c}_1) = 0,0879$. В этих условиях при использовании формул дискретизации (2) и (3) ошибки аппроксимации $p(x)$ равны 0,225 и 0,1695. Применение нового метода оптимизации $\bar{p}(x)$ и формулы дискретизации (1) позволяет получить значение $\bar{R}(\bar{c}_2) = 0,0602$. Ошибки аппроксимации $p(x)$, соответствующие методам дискретизации (2), (3), определяются значениями 0,175 и 0,1441. Эффективность метода дискретизации Хайнкольда — Гаеде (1) сохраняется и для значений $n > 1000$ (см. табл. 4).

Полученные результаты объясняются значительно бóльшим количеством интервалов дискретизации области значений случайной величины при использовании формулы (1) по сравнению с формулами (2), (3). С увеличением количества интервалов дискретизации N улучшаются условия сходимости регрессионной оценки плотности вероятности [11].

Заключение. Регрессионная оценка плотности вероятности является основой синтеза алгоритмов обработки статистических данных большого объёма. Её построение предполагает сжатие исходной статистической информации с использованием формул дискретизации и методов выбора оптимальных коэффициентов размытости ядерных функций. Выбор коэффициентов размытости ядерных функций регрессионной оценки плотности вероятности, основанный на минимизации критерия (9), обладает преимуществом над традиционным методом оптимизации при значениях параметра логнормального закона распределения одномерной случайной величины $s \leq 0,25$ с применением формулы дискретизации Хайнкольда — Гаеде. При $s = 0,5$ установлено преимущество традиционного метода оптимизации регрессионной оценки плотности вероятности по сравнению с предлагаемым подходом. Этот вывод можно объяснить увеличением асимметрии логнормального закона распределения и сокращением области значений со сложным видом плотности вероятности. Использование формулы дискретизации Хайнкольда — Гаеде при синтезе регрессионной оценки имеет значительное преимущество по сравнению с методами Старджесса, Брукса — Каррузера.

Полученные результаты исследования являются основой развития при оптимизации регрессионной оценки плотности вероятности многомерной случайной величины.

СПИСОК ЛИТЕРАТУРЫ

1. **Rudemo M.** Empirical Choice of Histograms and Kernel Density Estimators // Scand. Journ. Statist. 1982. **9**, N 2. P. 65–78.
2. **Bowman A. W.** A Comparative Study of Some Kernel-Based Nonparametric Density Estimators // Journ. Statist. Comput. Simul. 1982. **21**, Iss. 3–4. P. 313–327.
3. **Hall P.** Large Sample Optimality of Least Squares Cross-Validation in Density Estimation // Ann. Statist. 1983. **11**, N 4. P. 1156–1174.
4. **Jiang M., Provost S. B.** A hybrid bandwidth selection methodology for kernel density estimation // Journ. Statist. Comput. Simul. 2014. **84**, Iss. 3. P. 614–627. DOI: 10.1080/00949655.2012.721366.
5. **Dutta S.** Cross-validation Revisited // Commun. Statist. — Simul. Comput. 2016. **45**, Iss. 2. P. 472–490. DOI: 10.1080/03610918.2013.862275.
6. **Heinhold V. J., Gaede K.-W.** Ingenieur-Statistik. München-Wien: Springer-Verlag, 1968. 346 p.
7. **Sturges H. A.** The Choice of a Class Interval // Journ. Amer. Statist. Assoc. 1926. **21**, Iss. 153. P. 65–66. DOI: 10.1080/01621459.1926.10502161.
8. **Шторм Р.** Теория вероятностей. Математическая статистика. Статистический контроль качества. Пер. с нем. Н. Н. и М. Г. Федоровых; под ред. Н. С. Райбмана. М.: Мир, 1970. 368 с.
9. **Лапко А. В., Лапко В. А.** Оптимальный выбор количества интервалов дискретизации области изменения одномерной случайной величины при оценивании плотности вероятности // Измерительная техника. 2013. № 7. С. 24–27.
10. **Лапко А. В., Лапко В. А.** Оценивание интеграла от квадрата плотности вероятности одномерной случайной величины // Измерительная техника. 2020. № 7. С. 22–28. DOI: 10.32446/0368-1025it.2020-7-22-28.
11. **Лапко А. В., Лапко В. А.** Регрессионная оценка многомерной плотности вероятности и её свойства // Автометрия. 2014. **50**, № 2. С. 50–56.
12. **Parzen E.** On Estimation of a Probability Density Function and Mode // Ann. Math. Statistic. 1962. **33**, Iss. 3. P. 1065–1076.
13. **Епанечников В. А.** Непараметрическая оценка многомерной плотности вероятности // Теория вероятностей и её применения. 1969. **14**, вып. 1. С. 156–161.
14. **Градов В. М., Овечкин Г. В., Овечкин П. В., Рудаков И. В.** Компьютерное моделирование. М.: ИНФРА-М (КУРС), 2019. 264 с.

Поступила в редакцию 16.06.2023

После доработки 26.03.2024

Принята к публикации 17.04.2024