



<http://dx.doi.org/10.35596/1729-7648-2021-19-4-13-20>

Оригинальная статья  
Original paper

УДК 519.2

## ФИЛЬТРАЦИЯ ГИСТОГРАММНОЙ ОЦЕНКИ ПЛОТНОСТИ ВЕРОЯТНОСТИ НА ОСНОВЕ НЕЧЕТКОЙ ПРИНАДЛЕЖНОСТИ ДАННЫХ ИНТЕРВАЛУ ГРУППИРОВАНИЯ

А.В. ОВСЯННИКОВ<sup>1</sup>, В.М. КОЗЕЛ<sup>2</sup>

<sup>1</sup>Белорусский государственный университет (г. Минск, Республика Беларусь)

<sup>2</sup>Белорусский государственный университет информатики и радиоэлектроники (г. Минск, Республика Беларусь)

Поступила в редакцию 2 октября 2020

© Белорусский государственный университет информатики и радиоэлектроники, 2021

**Аннотация.** В работе предложена гистограммная оценка плотности вероятности на основе нечеткой принадлежности данных интервалу группирования. Приведена методика построения гистограммной оценки с применением гистограммного сглаживающего фильтра. Описана методика построения такого фильтра. Установлен основной параметр фильтра – коэффициент статистической взаимосвязи между количеством данных, попавших в интервал группирования при единичной функции включения и при подходе с использованием функции принадлежности. Применение итерационной процедуры для гистограммного фильтра позволяет обеспечить большую «сглаженность» гистограммы. Результаты моделирования показывают эффективность применения гистограммного фильтра для разных объемов данных. При этом становится не критичным выбор числа интервалов группирования для «правильного» распознавания плотности вероятности. Гистограммный фильтр является простым инструментом, который легко может быть встроен в любой алгоритм построения гистограммных оценок.

**Ключевые слова:** плотность вероятности, нечеткая принадлежность, взвешенная гистограммная оценка, гистограммный фильтр.

**Конфликт интересов.** Авторы заявляют об отсутствии конфликта интересов.

**Для цитирования.** Овсянников А.В., Козел В.М. Фильтрация гистограммной оценки плотности вероятности на основе нечеткой принадлежности интервалу группирования. Доклады БГУИР. 2021; 19(4): 13-20.

## FILTRATION OF HISTOGRAM EVALUATION OF PROBABILITY DENSITY BASED ON FUZZY DATA ACCESSIBILITY TO A GROUPING INTERVAL

ANDREI V. AUSIANNIKAU<sup>1</sup>, VICTOR M. KOZEL<sup>2</sup>

<sup>1</sup>Belarusian State University (Minsk, Republic of Belarus)

<sup>2</sup>Belarusian State University of Informatics and Radioelectronics (Minsk, Republic of Belarus)

Submitted 2 October 2020

© Belarusian State University of Informatics and Radioelectronics, 2021

**Abstract.** The paper proposes a histogram estimate of the probability density based on fuzzy data belonging to a grouping interval. A methodology for constructing a histogram estimate using a histogram smoothing filter

is presented. The technique of constructing such a filter is described. The main filter parameter is established – the coefficient of the statistical relationship between the amount of data falling into the grouping interval for a single inclusion function and when approaching to use the membership function. The use of an iterative procedure for a histogram filter allows for a greater “smoothness” of the histogram. The simulation results show the effectiveness of using a histogram filter for different data volumes. At the same time, the choice of the number of grouping intervals for the “correct” recognition of probability density becomes not critical. The histogram filter is a simple tool that can easily be built into any algorithm for constructing histogram estimates.

**Keywords:** probability density, fuzzy membership, weighted histogram estimate, histogram filter.

**Conflict of interests.** The authors declare no conflict of interests.

**For citation.** Ausiannikau A.V., Kozel V.M. Filtration of histogram evaluation of probability density based on fuzzy data accessibility to a grouping interval. Doklady BGUIR. 2021; 19(4): 13-20.

## Введение

Гистограммная оценка плотности распределения вероятности (ПВ) – одна из самых распространенных, исторически первых и общепринятых процедур описательной, прикладной статистики, в частности, в сфере статистической радиотехники, задачах электромагнитной совместимости. Проблематика, связанная с гистограммной оценкой в прикладной статистике (ГОСТ Р 50.1.033-2001 Прикладная статистика), также хорошо известна [1, 2]. В стационарном случае, при исследовании стационарных временных рядов, эффективность гистограммной оценки связана лишь с наличием достаточного времени для ее получения. Если же исследуется нестационарный объект или нестационарный временной ряд и требуется более точное распознавание момента времени, когда состояние объекта или динамика ряда, характеризуемая распределением, значительно изменились, то построение простой, быстрой (на малых объемах данных) и эффективной гистограммной оценки ПВ является актуальной задачей.

Развитие непараметрических методов [3] и общая направленность тематики работ в области непараметрических процедур оценивания ПВ (ядерная, проекционная, сплайн и др.) [4] не исключают применения гистограммных оценок: например, переменных, к числу которых относятся и полиграммные. Одной из проблем «правильного» распознавания ПВ, наряду с выбором количества интервалов группирования данных, является возможная «изрезанность» гистограммы, что особенно характерно при относительно небольших наборах данных. Решение этой проблемы заключается в применении гистограммных фильтров, например, усредняющего, медианного, гауссовского и др. [5–10]. Однако их применение интуитивно и исходит в основном из практической целесообразности. В представляемой работе предлагается теоретически обоснованная методика построения гистограммного фильтра, учитывающая следующие соображения.

Во-первых, можно отказаться от строгой единичной функции включения данных в конкретный интервал группирования. Данные могут находиться вблизи границ интервала и при изменении числа интервалов оказаться в другом интервале. Расположение данных на интервале группирования (близость или удаленность от границ интервала) может интерпретироваться как нечеткая принадлежность данных конкретному интервалу.

Во-вторых, введение понятия нечеткой принадлежности данных интервалу группирования позволяет за счет весовых функций перегруппировать эти данные так, чтобы уменьшилась «изрезанность» гистограммы и тем самым обеспечивалась ее сглаженность.

Таким образом, цель статьи – получить методику построения простого и эффективного гистограммного фильтра, обеспечивающего получение при относительно небольших объемах данных сглаженной гистограммной оценки ПВ на основе нечеткой принадлежности данных интервалу группирования. В статье рассматривается одномерная задача построения гистограммного фильтра на примере гистограммной оценки ПВ.

### Общая методика построения гистограммной оценки на основе нечеткой принадлежности данных интервалу группирования

В самом общем виде построение гистограммной оценки неизвестной, непрерывной ПВ состоит в следующем. Пусть имеется выборка случайных данных  $\{x_i\}$ ,  $i = \overline{1, n}$ , и пусть определено разбиение числовой прямой на  $m$  непересекающихся и примыкающих друг к другу интервалов  $A_j$ ,  $j = \overline{1, m}$ , равной длины  $\Delta_x = X_{j+1} - X_j = (x_{\max} - x_{\min}) / m$ ,  $X_{m+1} = x_{\max} = \max_i x_i$ ,  $X_1 = x_{\min} = \min_i x_i$ , где  $X_j$  – границы интервалов. Тогда нормализованная гистограммная оценка ПВ будет определяться вектором  $\{g_j = v_j / \Delta_x n\}$ , где  $v_j = \sum_{x_i \in A_j} I_j(x_i)$  – целое число данных,

попавших в  $A_j$  интервал,  $\sum_j v_j = n$ ,  $I_j(x_i) = \begin{cases} 1, & \text{если } x_i \in A_j, \\ 0, & \text{если } x_i \notin A_j. \end{cases}$  – индикаторная функция.

Заменим индикаторную функцию  $I_j(x_i)$  (единичную функцию включения) весовой функцией  $\mu_j(x_i)$ ,  $0 \leq \mu_j(x) \leq 1$ , которая будет характеризовать нечеткую принадлежность данных  $\{x_i\}$ ,  $i = \overline{1, n}$ , интервалу группирования  $A_j$ . Заметим, что если область определения отдельной  $j$ -й индикаторной функции составляет величину  $\Delta_x$ , то область определения весовой функции может быть значительно шире  $\Delta_\mu > \Delta_x$  и зависит от ее конкретного выбора. Для конкретизации этого значения и из практических соображений будем полагать, что  $\Delta_\mu = x_{j+1} - x_{j-1}$ , где  $x_{j-1}$ ,  $x_{j+1}$  – точки середины интервалов группирования  $x_j = (X_{j+1} - X_j) / 2$ .

Статистическим эквивалентом числа  $v_j$  будет, в общем случае, вещественное число, равное

$$u_j = \sum_i \mu_j(x_i) = \sum_{x_i \in A_{j-1}} \mu_j(x_i) + \sum_{x_i \in A_j} \mu_j(x_i) + \sum_{x_i \in A_{j+1}} \mu_j(x_i), \quad j = \overline{1, m}. \quad (1)$$

Формула (1) учитывает вес некоторого числа данных, принадлежащих соседним с  $A_j$  участкам:  $A_{j-1}$  и  $A_{j+1}$ . Очевидно, что всегда выполняется неравенство  $v_j \geq u_j = \sum_{x_i \in A_j} \mu_j(x_i)$ , где

знак равенства будет соответствовать предельному случаю равенства весовой и индикаторной функций.

Для выполнения общих условий нормировки необходимо выполнение соотношений:

$$1) \sum_j \mu_j(x) = 1, \quad 2) \sum_j \sum_i \mu_j(x_i) = \sum_j u_j = n, \quad 3) 0 \leq \mu_j(x) \leq 1, \quad (2)$$

определяющих специальный вид функций  $\mu_j(x)$ . Так, в частности, первое условие системы (2) указывает на их взаимную симметрию на соприкасающихся интервалах и взаимозависимость, а также особый вид для конечных интервалов  $A_1$  и  $A_m$ . Таким образом, гистограммная оценка ПВ с весовой функцией, характеризующей нечеткую принадлежность данных интервалу группирования, будет следующей:

$$f_j^* = u_j / \Delta_x n, \quad j = \overline{1, m}. \quad (3)$$

На основании приведенных соображений дадим общее определение введенной гистограммной оценки (3).

**Определение.** Гистограммную оценку ПВ (3) будем называть *взвешенной гистограммной оценкой* (ВГО), полученной с использованием специальных весовых функций  $\mu_j(x)$ , характеризующих нечеткую принадлежность данных  $j$ -му интервалу группирования и для которых выполняются условия (2). Сделаем замечание. Оценка (3) рассматривается в статье только в контексте ее применения в гистограммном фильтре, поскольку сама по себе она не является асимптотически сходящейся по вероятности при  $n \rightarrow \infty$  к истинной.

### Реализация гистограммного фильтра

Для построения гистограммного сглаживающего фильтра определим коэффициент статистической взаимосвязи между числом  $v_j$  и числом  $u_{Aj} = \sum_{x_i \in A_j} \mu_j(x_i)$ , которое соответствует взвешенному числу данных, попавших в  $j$ -й интервал:

$$k = \frac{\bar{u}_{Aj}}{\bar{v}_j} = \frac{1}{\Delta_x} \int_{\Delta_x} \mu_j(x) dx. \quad (4)$$

Черта над символами в формуле (4) означает, что их отношение устойчиво для исследуемой выборки данных. Тот же результат можно получить в общем случае, используя усреднение функции принадлежности  $k = \int_{\Delta_x} \mu f dx / \int_{\Delta_x} f dx$ . Ограничиваясь первым членом разложения в ряд Тейлора ПВ в точке середины интервала группирования  $x_j$ , получим результат, совпадающий с (4). Формула (4) справедлива для внутренних интервалов гистограммы ( $j = \overline{2, m-1}$ ). Для конечных интервалов  $j = 1, m$ , исходя из условий нормировки (2) и в том случае, если  $\Delta_\mu = x_{j+1} - x_{j-1}$ , получим

$$k_e = \Delta_x^{-1} \left( \int_{\Delta_x/2} 1(x) dx + \int_{\Delta_x/2} \mu(x) dx \right) = (1+k)/2. \quad (5)$$

С учетом определенных формулами (1), (4), (5) коэффициентов  $k, k_e$  можем записать соотношение между числами  $\bar{v}_j$  и  $\bar{u}_j$ , которое будет представлять собой гистограммный фильтр нулевого порядка (априорные данные о предполагаемой ПВ отсутствуют) в одномерном случае:

$$\begin{cases} \bar{u}_j = \alpha \bar{v}_{j-1} + k \bar{v}_j + \alpha \bar{v}_{j+1}, & j = \overline{2, m-1}, \\ \alpha = \bar{u}_{Aj-1} / \bar{v}_{j-1} = \bar{u}_{Aj+1} / \bar{v}_{j+1} = (1-k)/2, \\ \bar{u}_1 = k_e \bar{v}_1 + (1-k_e) \bar{v}_2, & \bar{u}_m = (1-k_e) \bar{v}_{m-1} + k_e \bar{v}_m. \end{cases} \quad (6)$$

Заменой переменных  $\bar{v}_j, \bar{u}_j$  на  $g_j = v_j / \Delta_x n$  и  $f_j^*$  получим гистограммный фильтр относительно значений ПВ на интервале группирования данных:

$$\begin{cases} f_j^* = \alpha g_{j-1} + k g_j + \alpha g_{j+1}, & j = \overline{2, m-1} \\ f_1^* = k_e g_1 + (1-k_e) g_2, & f_m^* = (1-k_e) g_{m-1} + k_e g_m. \end{cases} \quad (7)$$

Если в формуле  $k = \int_{\Delta_x} \mu f dx / \int_{\Delta_x} f dx$  не ограничиваться первым членом разложения в ряд Тейлора ПВ, можем получить гистограммные фильтры высших порядков, требующих априорных знаний о предполагаемой ПВ. В данной работе рассматривается только гистограммный фильтр нулевого порядка. Введением итеративной процедуры для формул (6), (7) можно добиться большей степени сглаживания:

$$\begin{cases} \bar{u}_j^{q+1} = \alpha \bar{u}_{j-1}^q + k \bar{u}_j^q + \alpha \bar{u}_{j+1}^q, & j = \overline{2, m-1} \\ \bar{u}_1^{q+1} = k_e \bar{u}_1^q + (1-k_e) \bar{u}_2^q, & \bar{u}_m^{q+1} = (1-k_e) \bar{u}_{m-1}^q + k_e \bar{u}_m^q, \end{cases} \quad (8)$$

$$\begin{cases} f_j^{*q+1} = \alpha f_{j-1}^q + k f_j^{*q} + \alpha f_{j+1}^{*q}, & j = \overline{2, m-1} \\ f_1^{*q+1} = k_e f_1^{*q} + (1-k_e) f_2^{*q}, & f_m^{*q+1} = (1-k_e) f_{m-1}^{*q} + k_e f_m^{*q}, \end{cases} \quad (9)$$

где  $q$  – порядковый номер итерации,  $q = 1$  соответствует процедуре (6), (7). Заметим, что используя аналогичный подход в фильтрации многомерных данных, в частности, гистограмм изображений, получаем теоретически обоснованные результаты [5–9].

Перейдем к определению значения коэффициента  $k$ . При отсутствии априорных данных о предполагаемом виде ПВ, можно воспользоваться «информационным подходом» и показать, что оптимальным значением весовых коэффициентов фильтра будет  $k = 1/3$ . Это следует из условия максимизации энтропии  $-\sum_{i=j-1}^{j+1} K_i \log(K_i) \rightarrow \max_k$ , где  $K_i$  – информационный вклад (весовой коэффициент)  $i$ -го интервала:  $K_i = \alpha = (1-k)/2$  для  $i = j-1, j+1$  и  $K_j = k$ ,  $0 \leq K_i \leq 1$ ,  $\sum_{i=j-1}^{j+1} K_i = 1$ .

С другой стороны, для упрощения решения задачи и последующего моделирования, функции  $\mu_j(x)$  можно выбрать ступенчатыми и симметричными. В этом случае, на основании формулы (4), коэффициент  $k$  для каждого внутреннего интервала группирования ( $j = \overline{2, m-1}$ ) будет постоянен и одинаков:  $\mu_j(x) = \{k, \text{ для } A_j; \alpha \text{ для } A_{j-1}, A_{j+1}\}$ , а для конечных интервалов соответственно  $\mu_1(x) = \{k_e \text{ для } A_1; \alpha \text{ для } A_2\}$  и  $\mu_m(x) = \{k_e \text{ для } A_m; \alpha \text{ для } A_{m-1}\}$ .

Рассмотрим теперь основные свойства выборочной статистики  $\chi_{Aj}^2 = \sum_{j=1}^m [\bar{u}_{Aj} - np_j]^2 / np_j$  по отношению к стандартной величине  $\chi^2 = \sum_{j=1}^m [\bar{v}_j - np_j]^2 / np_j$  с  $m-1$  степенями свободы, плотностью вероятности  $f(x) = \left[ 2^{\frac{m-1}{2}} \Gamma\left(\frac{m-1}{2}\right) \right]^{-1} x^{\frac{m-3}{2}} e^{-\frac{x}{2}}$ ,  $x > 0$ , параметрами математического ожидания  $\mathbf{M}\chi^2 = m-1$  и дисперсии  $\mathbf{D}\chi^2 = 2\mathbf{M}\chi^2 = 2(m-1)$ .

**Теорема.** В условиях нормального распределения выборочной совокупности при  $n \rightarrow \infty$  распределение величины  $\chi_{Aj}^2$  определяется ПВ:

$$f_{Aj}(x) = \left[ 2^{\frac{m-1}{2}} k_2 \Gamma\left(\frac{m-1}{2}\right) \right]^{-1} \left( \frac{x-C}{k^2} \right)^{\frac{m-3}{2}} e^{-\frac{x-C}{2k^2}}, \quad C = n(1-k)^2, \quad x > 0 \quad (10)$$

с параметрами математического ожидания и дисперсии, равными

$$\mathbf{M}\chi_{Aj}^2 = k^2(m-1) + 2n(1-k)^2, \quad \mathbf{D}\chi_{Aj}^2 = k^4\mathbf{D}\chi = 2k^4(m-1). \quad (11)$$

**Доказательство.** Подставим  $\bar{u}_{Aj} = k\bar{v}_j$  в выражение для статистики  $\chi_{Aj}^2$ . Получим  $\chi_{Aj}^2 = \sum_{j=1}^m [k\bar{v}_j - np_j]^2 / np_j = k^2 \sum_{j=1}^m \bar{v}_j^2 / np_j - n(2k-1)$ . Поскольку  $\mathbf{M} \sum_{j=1}^m \bar{v}_j^2 / np_j = n+m-1$ , то математическое ожидание статистики  $\chi_{Aj}^2$  получается следующим:

$$\mathbf{M}\chi_{Aj}^2 = k^2\mathbf{M}\chi^2 + n(1-k)^2. \quad (12)$$

Линейное преобразование случайной величины  $\chi^2$  приводит к ПВ величины  $\chi_{Aj}^2$  следующего вида:  $f_{Aj}(x) = f((x-C)/k^2)/k^2$ , что соответствует формуле (10) с параметрами математического ожидания и дисперсии, определяемыми формулами (11). Таким образом, теорема доказана.

*Следствие.* Линейная зависимость статистик (12) позволяет определить значение параметра  $k_{Aj}$ , при котором значение  $\mathbf{M}\chi_{Aj}^2$  стремится к минимуму:  $k_{Aj} = [1 + (m-1)/n]^{-1}$ . При этом значении коэффициента  $k_{Aj}$  математическое ожидание статистики  $\chi_{Aj}^2$  равно  $\mathbf{M}\chi_{Aj}^2 = \mathbf{M}\chi^2 / 2$ , т. е. качество, уверенность «правильного распознавания» неизвестной ПВ возрастает вдвое.

### Моделирование гистограммного фильтра

Далее приведены примеры некоторых результатов моделирования и применения гистограммного фильтра ( $k = 1/3$ ) к генерируемым случайным данным (число выборок 5),

распределенным с нормальной плотностью  $f(x) = (2\pi\sigma_x^2)^{-1/2} \exp(-x^2 / 2\sigma_x^2)$ ,  $\sigma_x^2 = 1$ .

Сравнивались критерии:  $\chi^2$  – стандартный и  $\chi_{\text{ВГО}}^2 = \sum_{j=1}^m [\bar{u}_{Aj} - np_j]^2 / np_j$  (см. табл. 1, 2).

В табл. 1, 2 также приведены вероятности возможного превышения полученного значения статистики при истинности нулевой гипотезы  $Pv$  и  $Pu$ . Гипотезу о согласии не следует отвергать, если  $Pv(\chi_{\text{кр}}^2 > \chi^2) > \alpha^*$ ,  $Pu(\chi_{\text{кр}}^2 > \chi_{\text{ВГО}}^2) > \alpha^*$ , где  $\alpha^* = 0,05$  – заданный уровень значимости,  $\chi_{\text{кр}}^2$  – критическое значение при заданном уровне значимости.

Для сопоставления результатов моделирования с коэффициентом  $k = 1/3$  вычислялись выборочные коэффициенты, использующие априорные сведения о теоретических вероятностях:  $k_{\text{выб}} = 1 + 2 \left[ \sum_{j=1}^m \bar{V}_j^2 / np_j \right]^{-1} \left[ \sum_{j=1}^m (\bar{v}_j - np_j) \bar{V}_j / np_j \right]$  (второй столбец табл. 1 и табл. 2),

найденные из условия  $\chi_{\text{ВГО}}^2 \rightarrow \min_k$ . Значения величин, входящих в формулу:

$\bar{V}_j = \bar{v}_{j-1} - 2\bar{v}_j + \bar{v}_{j+1}$  для  $j = \overline{2, m-1}$  и  $\bar{V}_1 = -\bar{v}_1 + \bar{v}_2$ ,  $\bar{V}_m = -\bar{v}_m + \bar{v}_{m-1}$ . Вычисляемые в результате

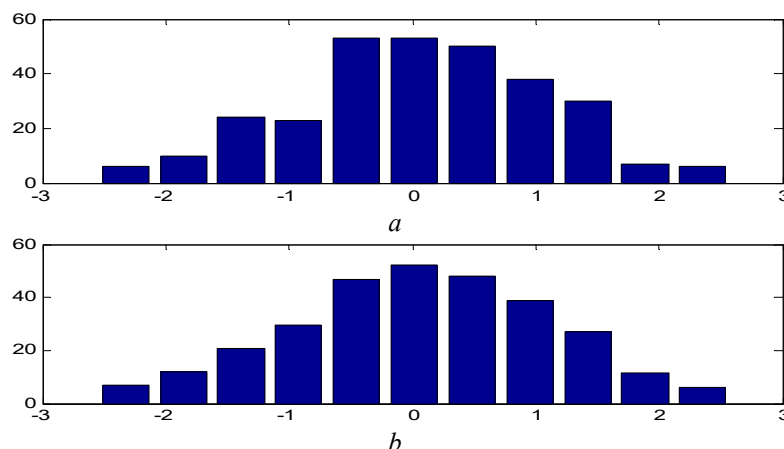
моделирования коэффициенты  $k_{\text{выб}}$  оказываются меньше единицы, а не равны ей, как при стандартном подходе построения гистограммы. Этот эффект свидетельствует о возможности применения гистограммного фильтра (формулы (6)–(9)) при относительно небольших объемах выборок. Многочисленные результаты моделирования показывают, что с увеличением объема выборки коэффициент  $k_{\text{выб}}$  стремится к единице, и применение гистограммного фильтра становится нецелесообразным.

В табл. 1 приведены результаты моделирования, полученные при  $n = 300$ ,  $m = 11$ . На рис. 1 приведены столбцовые гистограммы векторов  $v$  (рис. 1, а) и  $u$  (рис. 1, б), полученные в результате моделирования (строка № 3 табл. 1).

В табл. 2 приведены результаты моделирования, полученные при  $n = 1000$ ,  $m = 19$ . На рис. 2 приведены столбцовые гистограммы векторов  $v$  (рис. 2, а) и  $u$  (рис. 2, б), полученные в результате моделирования (строка № 1 табл. 2).

**Таблица 1.** Статистики и вероятности при  $n = 300$ ,  $m = 11$   
**Table 1.** Statistics and probabilities for  $n = 300$ ,  $m = 11$

№	$k_{\text{выб}}$	$\chi_{\text{кр}}^2$	$\chi^2$	$\chi_{\text{ВГО}}^2$	$Pv$	$Pu$
1	0,547	18,307	19,98	12,91	0,029	0,228
2	0,5857		10,97	5,36	0,359	0,865
3	0,541		22,85	13,87	0,011	0,178
4	0,498		6,56	1,12	0,766	0,999
5	0,767		19,75	17,45	0,031	0,064

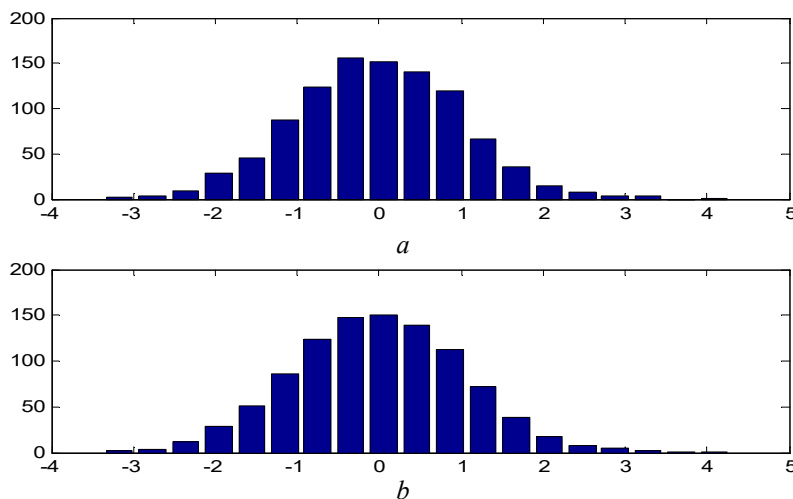


**Рис. 1.** Столбцовые гистограммы векторов  $v$  (а) и  $u$  (б), полученные в результате моделирования (строка № 3 табл. 1)

**Fig. 1.** Column histograms of vectors (a) and (b) obtained as a result of modeling (line No. 3 of table 1)

**Таблица 2.** Статистики и вероятности при  $n = 1000, m = 19$   
**Table 2.** Statistics and probabilities for  $n = 1000, m = 19$

№	$k_{\text{выб}}$	$\chi_{\text{сп}}^2$	$\chi^2$	$\chi_{\text{ВГО}}^2$	$Pv$	$Pu$
1	0,389	28,869	37,87	15,38	0,004	0,635
2	0,458		20,75	9,20	0,291	0,954
3	0,431		8,21	3,74	0,975	0,999
4	0,411		21,20	8,68	0,269	0,966
5	0,546		30,74	20,82	0,031	0,288



**Рис. 2.** Столбцовые гистограммы векторов  $v(a)$  и  $u(b)$ , полученные в результате моделирования (строка № 1 табл. 2).

**Fig. 2.** Column histograms of vectors  $(a)$  and  $(b)$  obtained as a result of modeling (line No. 1 of table 2)

### Заключение

Рассмотренная в работе методика построения и применения одномерного гистограммного фильтра является простым и вместе с тем эффективным инструментом статистического анализа данных при относительно небольших их объемах, что подтверждается результатами моделирования. Из анализа табл. 1, 2 можно сделать вывод, что  $\chi_{\text{КР}}^2 > \chi_{\text{ВГО}}^2$  и  $\chi^2 > \chi_{\text{ВГО}}^2$  во всех пяти экспериментах, в то время как стандартная статистика может «ошибаться» (строки № 1, 3, 5 табл. 1 и строки № 1, 5 табл. 2).

Гистограммный фильтр является простым инструментом, который легко может быть встроен в любой алгоритм построения гистограммных оценок.

### Список литературы

1. Орлов Ю.Н. Оптимальное разбиение гистограммы для оценивания выборочной плотности функции распределения нестационарного временного ряда. *Препринты ИПМ им. М.В. Келдыша*. 2013;14:26.
2. Chong Gu, Yongho Jeon and Yi Lin. Nonparametric density estimation in high-dimensions. *Statistica Sinica*. 2013;23:1131-1153.
3. Орлов А.И. Новая парадигма прикладной статистики. *Заводская лаборатория. Диагностика материалов*. 2012;78(11).
4. Devroye L. and Györfi L. *Nonparametric Density Estimation: The L1 View*. New York: John Wiley Sons; 1985.
5. Gonzalez R. *Digital Image Processing*, 3rd. Pearson Hall; 2008.
6. Solomon C.J., Breckon T.P. *Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab*. Wiley-Blackwell; 2010. DOI:10.1002/9780470689776.
7. Gonzalez R. *Digital image processing*. New York, NY: Pearson; 2018.
8. Овсянников А.В. *Применение стохастических формирующих фильтров*. Germany, KG, Saarbrücken: LAP LAMBERT Academic Publishing, OmniScriptum GmbH & Co; 2017.
9. Bernd J. *Digital Image Processing* (6 ed.). Springer Berlin Heidelberg New York; 2005.
10. Chakravorty Pragnan. What is a Signal? [Lecture Notes]. *IEEE Signal Processing Magazine*. 35 (5):175-177. DOI: 10.1109/MSP.2018.2832195.

## References

1. Orlov Yu.N. [Optimal histogram partitioning for estimating the sample density of the distribution function of an unsteady time series]. *Preprinty IPM im. M.V. Keldysha = Preprints of the Institute of Applied Mathematics M.V. Keldysh RAS.* 2013;14:26. (in Russ.)
2. Chong Gu, Yongho Jeon and Yi Lin. Nonparametric density estimation in high-dimensions. *Statistica Sinica.* 2013;23:1131-1153.
3. Orlov A.I. [The new paradigm of applied statistics]. *Zavodskaya laboratoriya. Diagnostika materialov = Industrial Laboratory. Diagnostics of Materials.* 2012;78(11). (in Russ.)
4. Devroye L. and Györfi L. *Nonparametric Density Estimation: The L1 View.* New York: John Wiley Sons; 1985.
5. Gonzalez R. *Digital Image Processing*, 3rd. Pearson Hall; 2008.
6. Solomon C.J., Breckon T.P. *Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab.* Wiley-Blackwell; 2010. DOI:10.1002/9780470689776.
7. Gonzalez R. *Digital image processing.* New York, NY: Pearson; 2018.
8. Ovsyannikov A.V. [The use of stochastic shaping filters]. Germany, KG, Saarbrücken: LAP LAMBERT Academic Publishing, OmniScriptum GmbH & Co; 2017. (in Russ.)
9. Bernd J. *Digital Image Processing* (6 ed.). Springer Berlin Heidelberg New York; 2005.
10. Chakravorty Pragnan. What is a Signal? [Lecture Notes]. *IEEE Signal Processing Magazine.* 35 (5):175-177. DOI: 10.1109/MSP.2018.2832195.

## Вклад авторов

Овсянников А.В. разработал общую методику построения гистограммной оценки на основе нечеткой принадлежности данных интервалу группирования, предложил алгоритм гистограммной фильтрации данных.

Козел В.М. выполнил моделирование работы гистограммного фильтра.

## Authors' contribution

Ausiannikau A.V. developed a general methodology for constructing a histogram estimate based on fuzzy data belonging to the grouping interval, proposed an algorithm for histogram data filtering.

Kozel V.M. performed modeling of the histogram filter.

## Сведения об авторах

Овсянников А.В., к.т.н., доцент, доцент кафедры информационных технологий Белорусского государственного университета.

Козел В.М., к.т.н., доцент, доцент кафедры информационных радиотехнологий Белорусского государственного университета информатики и радиоэлектроники.

## Information about the authors

Ausiannikau A.V., PhD, Associate Professor, Associate Professor at the Information Technologies Department of Belarusian State University.

Kozel V.M., PhD, Associate Professor, Associate Professor at the Information Radiotechnologies Department of the Belarusian State University of Informatics and Radioelectronics.

## Адрес для корреспонденции

220030, Республика Беларусь,  
г. Минск, пр. Независимости, 4,  
Белорусский государственный университет;  
тел. +375-17-209-58-94;  
e-mail: andovs@mail.ru  
Овсянников Андрей Витальевич

## Address for correspondence

220030, Republic of Belarus,  
Minsk, Nezavisimosti ave., 4,  
Belarusian State University;  
tel. +375-17-209-58-94;  
e-mail: andovs@mail.ru  
Ausiannikau Andrei Vital'evich