УДК 519.222

ОБ ЭКСТРЕМАЛЬНЫХ СВОЙСТВАХ СРЕДНИХ ХАРАКТЕРИСТИК ВАРИАЦИОННЫХ РЯДОВ

В. Л. Хацкевич

Воронежский государственный университет

Поступила в редакцию 11.09.2014 г.

Аннотация. Исследованы экстремальные свойства средних характеристик вариационных рядов. Особое внимание уделяется структурным средним – медиане и среднему разделительному. Рассмотрен векторный случай. Установлены некоторые соотношения между средней арифметической, медианой и разделительным значением вариационных рядов.
Ключевые слова: вариационные ряды, средние характеристики, экстремальные свойства.
Annotation. Studied extremal properties of average characteristics variations-traditional series. Special attention is paid to structural mean, the median and the average separation. Considered the vector case. Established some relations between the arithmetic mean, the median and dividing by the value of a variational series.
Keywords: variational series, average characteristics, extreme properties.

ВВЕДЕНИЕ

Экстремальные свойства средней арифметической и медиа-ны конечных числовых последовательностей широко известны и используются в статистике [1]–[3]. Кроме того, известны экстремальные свойства математического ожидания и медианы случайной величины (см., напр., [4] гл. 5). Однако, для средних характеристик вариационных рядов, т.е. упорядоченных таблиц данных, соответствующие результаты менее известны или вообще не рассматривались. В настоящей работе обсуждаются некоторые известные результаты в этой области и устанавливаются новые экстремальные свойства аналитических и структурных средних вариационных рядов.

Пусть $x_1, x_2, \ldots, x_n$ – заданные числа (варианты), расположенные в порядке возрастания $x_1 < x_2 < \ldots < x_n$, , а $p_1, p_2, \ldots, p_n$ – заданные положительные числа (частоты вариантов). Рассмотрим вариационный или частотный ряд

$$\begin{matrix} x_1 & x_2 & \ldots & x_n \\ p_1 & p_2 & \ldots & p_n \end{matrix}. \qquad (1)$$

Средняя арифметическая взвешенная определяется равенством

$$\overline{x} = \frac{p_1 x_1 + p_2 x_2 + \ldots + p_n x_n}{p_1 + p_2 + \ldots + p_n}. \qquad (2)$$

Рассмотрим взвешенное среднеквадратическое отклонение чисел $x_1, x_2, \ldots, x_n$ от некоторой величины x

$$\rho^2(x) = \sum_{k=1}^{n} p_k (x_k - x)^2, \qquad (3)$$

где $p_1, p_2, \ldots, p_n$ – заданные положительные числа.

Как известно ([1], гл. I, § 5; [2], гл. 2), имеет место следующее характеристическое свойство взвешенных средних арифметических:

Теорема 1. *Взвешенное среднеквадратическое отклонение (3) достигает минимума при* $x = \overline{x}$, *определяемом формулой (2).*

Один из способов доказательства теоремы 1 состоит в применении формулы

$$\rho^2(x) = \rho^2(\overline{x}) + (\overline{x} - x)^2 \sum_{i=1}^{n} p_i.$$

Теорема 1 влечет экстремальные свойства других взвешенных средних (см., напр., рабо-

Работа выполнена при финансовой поддержке РФФИ, грант № 14-01-00253А

ту автора [5]). В частности, рассмотрим среднюю гармоническую вариационного ряда (1). В этом случае будем считать, что заданные числа $x_1, x_2, \dots, x_n$ положительны. Кроме того, рассмотрим положительные числа $\omega_1, \omega_2, \dots, \omega_n$. Взвешенная средняя гармоническая определяется формулой

$$\overline{x}_{гар.} = \left(\omega_1 + \omega_2 + \dots + \omega_n\right) \Big/ \left(\frac{\omega_1}{x_1} + \frac{\omega_2}{x_2} + \dots + \frac{\omega_n}{x_n}\right).$$

Положим $p_k = \dfrac{\omega_k}{x_k}$. В новых обозначениях $\overline{x}_{гар.}$ приобретает вид (2) взвешенной средней арифметической с весами p_k. Поэтому теорема 1 влечет

Следствие 1. *Взвешенная средняя гармоническая минимизирует взвешенное среднее квадратическое отклонение*

$$\sum_{k=1}^{n} \frac{\omega_k}{x_k} \left(x_k - x\right)^2.$$

Медианой монотонно возрастающей последовательности $x_1 \quad x_2 \quad \dots \quad x_N$, содержащей нечетное число членов, называют средний член последовательности с номером $m = \dfrac{n+1}{2}$. В случае четного числа членов n за медиану условились принимать среднюю арифметическую $\dfrac{1}{2}\left(x_m + x_{m+1}\right)$.

Широко известно характеристическое экстремальное свойство медианы: *сумма абсолютных значений отклонений* $\sum_{i=1}^{n} \left|x - x_i\right|$ *величин ряда от медианы является наименьшей.*

Доказательство этого утверждения приведено, например, в [1], гл. I, § 5, [2], гл. 2.

1. ХАРАКТЕРИСТИЧЕСКОЕ ЭКСТРЕМАЛЬНОЕ СВОЙСТВО МЕДИАНЫ ВАРИАЦИОННОГО РЯДА

Рассмотрим, как выглядит это свойство для частотных рядов вида (1), когда частоты $p_1, p_2, \dots, p_n$ – натуральные числа, а сумма частот $\sum_{i=1}^{n} p_i = N$. В случае нечетного N значение медианы x_m определяется следующими неравенствами

$$p_1 + p_2 + \dots + p_{m-1} < \frac{N+1}{2};$$

$$p_1 + p_2 + \dots + p_m \geq \frac{N+1}{2}. \qquad (4)$$

В случае четного N есть две возможности. Первая состоит в том, что для некоторого k выполнены два условия

$$p_1 + p_2 + \dots + p_{k-1} < \frac{N}{2}, \quad p_1 + p_2 + \dots + p_k > \frac{N}{2}. \qquad (5)$$

В этом случае значение медианы $x_m = x_k$.

Вторая возможность состоит в выполнении для некоторого k соотношений (случай неопределенной медианы)

$$p_1 + p_2 + \dots + p_{k-1} < \frac{N}{2}, \quad p_1 + p_2 + \dots + p_k = \frac{N}{2}. \qquad (6)$$

В этом случае в качестве медианы можно брать любое число из интервала $\left(x_k, x_{k+1}\right)$. Обычно берут $\dfrac{1}{2}\left(x_k + x_{k+1}\right)$, хотя есть и другие подходы. Ниже в замечании 1 мы предлагаем свой вариант выбора медианы в неопределенном случае.

Теорема 2. (Экстремальное свойство медианы вариационного ряда). *Взвешенная сумма абсолютных значений отклонений* $\sum_{i=1}^{n} \left|x - x_i\right| p_i$ *величин ряда от медианы является наименьшей, т.е.*

$$\sum_{i=1}^{n} \left|x_m - x_i\right| p_i \leq \sum_{i=1}^{n} \left|x_s - x_i\right| p_i \quad \left(\forall s = 1, 2, \dots n\right). \qquad (7)$$

При этом равенство в формуле (7) возможно только в случае неопределен-ной медианы, когда $s = k$, *либо* $s = k+1$.

Доказательство этого утверждения не является широко доступным. Например, в известных книгах [1], [2] оно формулируется, но доказывается только для числовых последовательностей. Мы предлагаем свое доказательство, причем подробно останавливаемся на нетривиальном случае неопределенной (многозначной) медианы. Подчеркнем, что наш метод доказательства позволяет исследовать экстремальные свойства других средних, в частности, разделительного значения.

Доказательство теоремы 2. Чтобы не различать случаи четного и нечетного N в доказательстве теоремы 2 свойства медианы (4), (5) можно давать в виде

$$p_1 + p_2 + \ldots + p_{m-1} < p_m + \ldots + p_n \text{ и}$$

$$p_1 + p_2 + \ldots + p_m > p_{m+1} + \ldots + p_n. \qquad (8)$$

Либо в случае неопределенной медианы (6) в виде

$$p_1 + p_2 + \ldots + p_{k-1} < p_k + \ldots + p_n \text{ и}$$

$$p_1 + p_2 + \ldots + p_k = p_{k+1} + \ldots + p_n. \qquad (9)$$

Пусть выполнено (8) и x_m – медиана. Зафиксируем член ряда с номером s. Рассмотрим отклонение

$$\Delta_s = \sum_{i=1}^{n} |x_s - x_i| p_i =$$

$$= \sum_{i=1}^{s} (x_s - x_i) p_i + \sum_{i=s+1}^{n} (x_i - x_s) p_i.$$

Пусть $s > m$, (т.е. $x_s > x_m$). Тогда

$$\sum_{i=1}^{s} (x_s - x_i) p_i = \sum_{i=1}^{m} (x_s - x_i) p_i + \sum_{i=m+1}^{s} (x_s - x_i) p_i$$

и

$$\sum_{i=s+1}^{n} (x_i - x_s) p_i =$$

$$= \sum_{i=m+1}^{n} (x_i - x_s) p_i - \sum_{i=m+1}^{s} (x_i - x_s) p_i.$$

Поэтому

$$\Delta_s = \sum_{i=1}^{m} (x_s - x_i) p_i + \sum_{i=m+1}^{n} (x_i - x_s) p_i +$$

$$+ 2 \sum_{i=m+1}^{s} (x_s - x_i) p_i.$$

Здесь первое и второе слагаемое можно представить соответственно в виде

$$\sum_{i=1}^{m} (x_s - x_i) p_i = \sum_{i=1}^{m} (x_m - x_i) p_i + (x_s - x_m) \sum_{i=1}^{m} p_i$$

и

$$\sum_{i=m+1}^{n} (x_i - x_s) p_i = \sum_{i=m+1}^{n} (x_i - x_m) p_i +$$

$$+ (x_m - x_s) \sum_{i=m+1}^{n} p_i.$$

Тогда

$$\Delta_s = \sum_{i=1}^{n} |x_m - x_i| p_i + 2 \sum_{i=m+1}^{s} (x_s - x_i) p_i +$$
$$+ (x_s - x_m) \left(\sum_{i=1}^{m} p_i - \sum_{i=m+1}^{n} p_i \right). \qquad (10)$$

По условию второе слагаемое неотрицательно (если $s = m+1$, то оно равно нулю). Последнее слагаемое положительно в силу предположения $x_s > x_m$ и второго неравенства формулы (8). Таким образом, при $s > m$ выполнено соотношение (7).

Для $s < m$ аналогично предыдущему имеем

$$\Delta_s = \sum_{i=1}^{n} |x_m - x_i| p_i + 2 \sum_{i=s+1}^{m} (x_i - x_s) p_i +$$

$$+ (x_m - x_s) \left(\sum_{i=m+1}^{n} p_i - \sum_{i=1}^{m} p_i \right).$$

Перепишем это выражение в виде

$$\Delta_s = \sum_{i=1}^{n} |x_m - x_i| p_i + 2 \sum_{i=s+1}^{m-1} (x_i - x_s) p_i +$$
$$+ (x_m - x_s) \left(\sum_{i=m}^{n} p_i - \sum_{i=1}^{m-1} p_i \right). \qquad (11)$$

По условию второе слагаемое в правой части (11) неотрицательно (если $s+1 = m$, то оно отсутствует). Положительность третьего слагаемого влечет предположение $x_m - x_s > 0$ и первое неравенство формулы (8).

Пусть для некоторого k выполнены соотношения (9) (случай неопределенной медианы). Рассмотрим случай $m = k$. Тогда при $s > m$ оценка (7) следует из (10) и правой части формулы (9). Отметим, что для $s = k+1$ неравенство (7) превращается в равенство. При $s < m$ оценка (7) вытекает из (11) и левой части формулы (9).

В случае $m = k+1$, при $s > m$ оценка (7) следует из (10) и правой части формулы (9), а при $s < m$ из (11) и правой части формулы (9). Отметим, что для $s = k$ неравенство (7) превращается в равенство. Теорема доказана.

Продолжим изучать случай неопределенной медианы.

Следствие 2. *Пусть для некоторого натурального k выполнены соотношения (9) и x – произвольное число из интервала (x_k, x_{k+1}). Тогда для него выполнено экстремальное соотношение (7) в виде*

$$\sum_{i=1}^{n} |x - x_i| p_i < \sum_{i=1}^{n} |x_s - x_i| p_i \ \ (\forall s \neq k, k+1).$$

В случае $s = k$, либо $s = k+1$ это неравенство превращается в равенство.

В самом деле, по условию найдется $\alpha \in (0,1)$ такое, что $x = \alpha x_k + (1-\alpha) x_{k+1}$. Поэтому

$$\sum_{i=1}^{n} |x - x_i| p_i =$$

$$= \sum_{i=1}^{n} |\alpha (x_k - x_i) + (1-\alpha)(x_{k+1} - x_i)| p_i \le$$

$$\le \alpha \sum_{i=1}^{n} |x_k - x_i| p_i + (1-\alpha) \sum_{i=1}^{n} |x_{k+1} - x_i| p_i.$$

Тогда согласно предыдущему для всякого $s \neq k, k+1$ можем записать

$$\sum_{i=1}^{n} |x - x_i| p_i <$$

$$< \alpha \sum_{i=1}^{n} |x_s - x_i| p_i + (1-\alpha) \sum_{i=1}^{n} |x_s - x_i| p_i =$$

$$= \sum_{i=1}^{n} |x_s - x_i| p_i,$$

что и влечет высказанное утверждение.

Следствие 3. *Пусть задан произвольный* $y \in (x_1, x_n)$, *тогда экстремальное свойство медианы* x_m *имеет вид*

$$\sum_{i=1}^{n} |x_m - x_i| p_i \leq \sum_{i=1}^{n} |y - x_i| p_i. \qquad (12)$$

Действительно, по условию найдется натуральное s такое, что $1 < s < n$ и $y \in (x_s, x_{s+1})$. Поэтому для некоторого $\beta \in (0,1)$ имеем $y = \beta x_s + (1-\beta) x_{s+1}$. Рассмотрим выражение

$$\sum_{i=1}^{n} |y - x_i| p_i = \sum_{i=1}^{s} (y - x_i) p_i + \sum_{i=s+1}^{n} (x_i - y) p_i.$$

Подставляя вместо y соответствующую выпуклую комбинацию элементов x_s, x_{s+1}, получим

$$\sum_{i=1}^{n} |y - x_i| p_i = \beta \sum_{i=1}^{n} |x_s - x_i| p_i + (1-\beta) \sum_{i=1}^{n} |x_{s+1} - x_i| p_i.$$

Используя для каждого из слагаемых правой части неравенство (7), получим формулу (12).

Замечание 1. Если N четное и выполнен случай неопределенности медианы (6) в частотном ряду, то из интервала (x_k, x_{k+1}) в качестве медианы автором предлагается выбирать $\frac{1}{p_k + p_{k+1}} (p_k x_k + p_{k+1} x_{k+1})$. Это выражение превращается в $\frac{1}{2}(x_k + x_{k+1})$ при $p_k = p_{k+1} = 1$. Возможен другой подход, учитывающий расстояние до начального x_1 и конечного x_N элементов. А именно, в качестве элемента из интервала (x_k, x_{k+1}) брать число $\frac{1}{r_1 + r_2}(r_1 x_k + r_2 x_{k+1})$, где $r_1 = x_k - x_1$, а $r_2 = x_N - x_{k+1}$. Либо наоборот $\frac{1}{r_1 + r_2}(r_2 x_k + r_1 x_{k+1})$.

Замечание 2. В предположении выполнения соотношения (8) либо (9) все рассуждения теоремы 2 годятся для произвольных положительных (не обязательно целых) весов p_i.

2. ПРИЛОЖЕНИЯ ТЕОРЕМЫ 2.

Рассмотрим дискретную случайную величину ξ, принимающую значения x_i с вероятностями p_i $(i = 1, 2, ..., n)$. Пусть $F(x)$ – функция распределения этой случайной величины. Как известно, [6], гл. 6 медианой распределения $F(x)$ называется такое значение аргумента x_m, для которого выполнены не-равенства

$$F(x_m) \leq \frac{1}{2} \leq F(x_m + 0). \qquad (13)$$

Как следствие теоремы 2 получим

Утверждение 1. *Абсолютный момент дискретной случайной величины* $M|\xi - c| = \sum_{i=1}^{n} |x_i - c| p_i$ *принимает минимальное значение, если* c *выбрано равным медиане распределения.*

Действительно, для дискретной случайной величины имеем $F(x) = \sum_{x_i < x} p_i$. Поэтому условие (13) означает выполнение неравенств $\sum_{i=1}^{m-1} p_i \leq \frac{1}{2} \leq \sum_{i=1}^{m} p_i$. Так как $\sum_{i=1}^{n} p_i = 1$, то отсюда следует, что выполнены условия (7), (8). Поэтому теорема 2 влечет высказанное утверждение.

Этот результат нетрудно обобщить на случай, когда дискретная слу-чайная величина принимает счетное число значений $x_1, x_2, ..., x_n, \ldots$, при-чем $\sum_{i=1}^{\infty} p_i = 1$.

Отметим, что в книге [4], гл.5 и других источниках приведен результат об экстремальном свойстве медианы для непрерывных (и именно непрерывных) случайных величин. Соответствующее свойство для дискретных случайных величин в силу их специфики, по-видимому, ранее не отмечалось.

Рассмотрим случай векторных средней арифметической и медианы. Пусть дана совокупность m-мерных векторов $X^1, X^2, ..., X^n$ и положительных чисел $p_1, p_2, ... p_n$. Средним арифметическим взвешенным этих векторов является вектор $\overline{X} = \frac{1}{N} \sum_{j=1}^{n} p_j X^j$, где $N = \sum_{j=1}^{n} p_j$.

Рассмотрим в пространстве R^m нормы вектора X

$$\|X\|_1 = \sum_{i=1}^{m} |x_i| \text{ и } \|X\|_2 = \left(\sum_{i=1}^{m} |x_i|^2 \right)^{1/2}.$$

Поскольку $\| \ \|_2$ эвклидова норма, то согласно работе автора [5] справедливо

Утверждение 2. *Вектор $\bar{X}$ минимизирует взвешенное отклонение*

$$\delta_2(X) = \left(\sum_{j=1}^{n} p_j \left\|X - X^j\right\|_2^2 \right)^{1/2}.$$

Рассмотрим отклонение $\delta_1(X) = \sum_{j=1}^{n} p_j \left\|X - X^j\right\|_2$. В книге [1] гл. III предлагается определение медианного вектора, как вектора, минимизирующего это отклонение. Рассмотрим отклонение $\tilde{\delta}_1(X) = \sum_{j=1}^{n} p_j \left\|X - X^j\right\|_1$. Оказывается, его минимизирует вектор с координатами из одномерных медиан.

В самом деле, пусть векторы X^j при $j = 1, 2, ..., n$ имеют вид $X^j = \left(x_1^j, x_2^j, ..., x_m^j\right)$. Расположим числа $x_1^1,\ x_1^2,\ \ldots,\ x_1^n$ в возрастающем порядке и выберем среди них медиану $\tilde{x}_1$. Аналогично поступим с координатами $x_2^1,\ x_2^2,\ \ldots,\ x_2^n$. Обозначим соответствующую медиану $\tilde{x}_2$ и так далее до $x_m^1,\ x_m^2,\ \ldots,\ x_m^n$ с медианой ранжированной совокупности $\tilde{x}_m$. Вектор $\tilde{X} = \left(\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_m\right)$ назовем медианным.

Утверждение 3. *Медианный вектор $\tilde{X}$ минимизирует отклонение*

$$\tilde{\delta}_1(X) = \sum_{j=1}^{n} p_j \left\|X - X^j\right\|_1.$$

Это предложение справедливо в силу экстремального свойства медианы, примененного покоординатно, поскольку

$$\tilde{\delta}_1(X) = \sum_{j=1}^{n} p_j \sum_{i=1}^{m} \left|x_i - x_i^j\right| = \sum_{i=1}^{m} \sum_{j=1}^{n} p_j \left|x_i - x_i^j\right|.$$

Вернемся к изучению частотного ряда (1). Рассмотрим случай, когда все члены ряда положительны (как это принято в статистике). В этом случае разделительным значением частотного ряда (1) называют такое значение $x_k = r$, которое удовлетворяет следующим неравенствам см. [1] гл. I, § 3

$$x_1 p_1 + x_2 p_2 + ... + x_{k-1} p_{k-1} < x_k p_k + ... + x_n p_n \quad (14)$$

и

$$x_1 p_1 + x_2 p_2 + ... + x_k p_k > x_{k+1} p_{k+1} + ... + x_n p_n. \quad (15)$$

Если вместо второго неравенства справедливо равенство

$$x_1 p_1 + x_2 p_2 + ... + x_k p_k = x_{k+1} p_{k+1} + ... + x_n p_n, \quad (16)$$

то разделительным значением называют любое число из интервала $\left(x_k, x_{k+1}\right)$, в частности $\frac{1}{2}\left(x_k + x_{k+1}\right)$.

Замечание 3. Если выполнен случай неопределенности разделительного значения в частотном ряду, то из интервала $\left(x_k, x_{k+1}\right)$ в качестве разделительного значения, аналогично замечанию 2, автором рекомендуется выбирать взвешенную с учетом весов выпуклую комбинацию элементов x_k и x_{k+1}.

Теорема 3. *В случае положительности всех чисел $x_1, x_2, ..., x_n$ взвешенная сумма абсолютных значений отклонений $\sum_{i=1}^{n} \left|x - x_i\right| x_i p_i$ величин ряда от разделительного значения является наименьшей.*

Этот результат, насколько нам известно, ранее не отмечался даже в стандартном (все $p_i = 1$) случае. Это свойство является характеристическим свойством разделительного значения. Оно может быть использовано в задачах логистики подобно тому, как это делается в случае медианы (см., напр., [3]).

Доказательство теоремы 3 с учетом замечания 2 повторяет доказательство теоремы 2, если ввести обозначения $x_i p_i = q_i$ $\left(i = 1, 2, ..., n\right)$ и во всех формулах теоремы 2 вместо p_i писать q_i.

3. НЕКОТОРЫЕ СООТНОШЕНИЯ МЕЖДУ СРЕДНИМИ

Расстояние между средней арифметической $\bar{x}$ частотного ряда (1) и его медианой Me оценивается в следующем, интересном на наш взгляд, утверждении.

Теорема 4. *Пусть дан частотный ряд (1). Имеет место оценка*

$$\left|\bar{x} - Me\right| \le \frac{1}{2} \max\left\{x_n - Me, Me - x_1\right\}, \quad (17)$$

где x_1 и x_n – первый и последний член ряда соответственно.

Действительно, рассмотрим случай нечетного $N=\sum_{j=1}^{n} p_j$, когда медиана потеряется однозначно. Пусть $Me = x_k$. Согласно определению (2) для $\overline{x}$ имеем

$$\overline{x}-x_k=\frac{1}{N}\left[\sum_{j=1}^{k}\left(x_j-x_k\right)p_j+\sum_{j=k+1}^{n}\left(x_j-x_k\right)p_j\right].$$

Отсюда,

$$\overline{x}-x_k\le\frac{1}{N}\sum_{j=k+1}^{n}\left(x_j-x_k\right)p_j\le\frac{1}{N}\left(x_n-x_k\right)\sum_{j=k+1}^{n}p_j. \quad (18)$$

Заметим, что из определения медианы (4) следует, что

$$\sum_{j=k+1}^{n}p_j=N-\sum_{j=1}^{k}p_j\le N-\frac{N+1}{2}=\frac{N-1}{2}<\frac{N}{2}.$$

Тогда неравенство (18) влечет соотношение $\overline{x}-x_k\le\frac{1}{2}\left(x_n-x_k\right)$.

Аналогично предыдущему

$$x_k-\overline{x}=$$

$$=\frac{1}{N}\left[\sum_{j=1}^{k}\left(x_k-x_j\right)p_j+\sum_{j=k+1}^{n}\left(x_k-x_j\right)p_j\right]\le$$

$$\le\frac{1}{N}\sum_{j=1}^{k}\left(x_k-x_j\right)p_j.$$

Отсюда,

$$x_k-\overline{x}\le\frac{1}{N}\sum_{j=1}^{k-1}\left(x_k-x_j\right)p_j\le\frac{1}{N}\left(x_k-x_1\right)\sum_{j=1}^{k-1}p_j. \quad (19)$$

Полагая $N=2m+1$, из соотношения (4) получим $\sum_{j=1}^{k-1}p_j<\frac{N+1}{2}=m+1$. Поскольку, $\sum_{j=1}^{k-1}p_j$ – целое неотрицательное число, то отсюда $\sum_{j=1}^{k-1}p_j\le m$. Тогда неравенство (19) влечет соотношение $x_k-\overline{x}\le\frac{1}{2}\left(x_k-x_1\right)$. Таким образом, в случае нечетного N неравенство (17) установлено. Если N четное и выполнено соотношение (5), то медиана по прежнему определяется однозначно и доказательство неравенства (17) совершенно аналогично предыдущему.

В случае выполнения (6) (случай неопределенной медианы) соотношение (17) аналогично предыдущему будет выполнено в виде

$$\left|\overline{x}-x_m\right|\le\frac{1}{2}\max\left\{x_n-x_m,x_m-x_1\right\}\quad\left(m=k,k+1\right).$$

Представляя любой элемент $x\in\left(x_k,x_{k+1}\right)$ как выпуклую комбинацию x_k и x_{k+1}, а также используя отмеченные выше соотношения получим неравенство (17) в случае неопределенной медианы. Теорема доказана.

Отметим взаимосвязь средней арифметической вариационного ряда (1) и разделительного значения этого ряда. При этом будем рассматривать ряды с положительными членами, которые обычно встречаются в статистике.

Утверждение 4. *Пусть* $x_k=r$ *– разделительное значение вариационного ряда (1) с положительными членами, а* $\overline{x}$ *– средняя арифметическая этого ряда. Тогда имеют место неравенства*

$$\frac{1}{N}\sum_{j=k+1}^{n}x_jp_j\le\frac{\overline{x}}{2},\quad\frac{1}{N}\sum_{j=k}^{n}x_jp_j>\frac{\overline{x}}{2} \quad (20)$$

и

$$\frac{1}{N}\sum_{j=1}^{k-1}x_jp_j<\frac{\overline{x}}{2},\quad\frac{1}{N}\sum_{j=1}^{k}x_jp_j\ge\frac{\overline{x}}{2}, \quad (21)$$

где $N=\sum_{j=1}^{n}p_j$.

Действительно, по определению средней арифметической и на основании (15) или (16) можем записать

$$\overline{x}=\frac{1}{N}\sum_{j=1}^{k}x_jp_j+\frac{1}{N}\sum_{j=k+1}^{n}x_jp_j\ge\frac{2}{N}\sum_{j=k+1}^{n}x_jp_j.$$

Следовательно, справедливо левое неравенство в (20). С другой стороны, на основании (14) можем записать следующее соотношение

$$\overline{x}=\frac{1}{N}\sum_{j=1}^{k-1}x_jp_j+\frac{1}{N}\sum_{j=k}^{n}x_jp_j<\frac{2}{N}\sum_{j=k}^{n}x_jp_j.$$

Поэтому, имеет место правое неравенство в (20).

Кроме того, согласно левому неравенству из (20) получим соотношение

$$\frac{1}{N}\sum_{j=1}^{k}x_jp_j=\overline{x}-\frac{1}{N}\sum_{j=k+1}^{n}x_jp_j\ge\frac{\overline{x}}{2},$$

которое влечет правую часть неравенства (21). Аналогично, используя правую часть неравенства (20) получим

$$\frac{1}{N}\sum_{j=1}^{k-1}x_jp_j=\overline{x}-\frac{1}{N}\sum_{j=k}^{n}x_jp_j<\frac{\overline{x}}{2},$$

т.е. левую часть неравенства (21).

Отметим, что каждое из условий (20), либо (21) может быть принято за определение разделительного значения в случае частотного ряда с положительными членами.

Следствие 4. *Для вариационных рядов с положительными членами справедливо неравенство* $\overline{x} < 2r$.

В самом деле, согласно правой части (21) имеет место соотношение

$$\frac{\overline{x}}{2} \le \frac{1}{N}\sum_{j=1}^{k} x_j p_j < \frac{1}{N} x_k \sum_{j=1}^{k} p_j < \frac{1}{N} x_k \sum_{j=1}^{n} p_j = x_k,$$

которое и влечет указанное неравенство.

Отметим, что неравенство $\overline{x} < 2r$ довольно грубое, зато наглядное. Ниже укажем более тонкое соотношение, связанное с результатом следующей теоремы 5.

Теорема 5. *Медиана* Me *вариационного ряда (1) с положительными членами и его разделительное значение* r *связаны соотношением* $Me \le r$.

Доказательство. Предположим противное. Пусть медиана определяется однозначно и $Me > r$. Если $Me = x_m$, а $r = x_k$, неравенство $Me > r$ означает, что $m > k$, следовательно $m \ge k+1$. Тогда из левой части неравенства (20) вытекает

$$\frac{x_m}{N}\sum_{j=m}^{n} p_j < \frac{1}{N}\sum_{j=m}^{n} x_j p_j \le \frac{1}{N}\sum_{j=k+1}^{n} x_j p_j \le \frac{\overline{x}}{2}. \quad (22)$$

При этом, согласно (8) $\sum_{j=m}^{n} p_j > \sum_{j=1}^{m-1} p_j$. Следовательно $2\sum_{j=m}^{n} p_j > N$. Тогда (22) влечет неравенство $\frac{1}{2}x_m = \frac{1}{N}x_m\frac{N}{2} < \frac{1}{N}x_m\sum_{j=m}^{n} p_j < \frac{\overline{x}}{2}$. Следовательно, $x_m < \overline{x}$.

С другой стороны, т.к. по предположению $m-1 \ge k$, то из правой части неравенства (21) следует, что $\frac{1}{N}\sum_{j=1}^{m-1} x_j p_j > \frac{\overline{x}}{2}$. Тогда имеет место

$$\frac{x_{m-1}}{N}\sum_{j=1}^{m-1} p_j > \frac{\overline{x}}{2}. \quad (23)$$

При этом, согласно (4) в случае нечетного $N = 2m+1$ имеем $\sum_{j=1}^{m-1} p_j < \frac{N+1}{2} = m+1$. Тогда $\sum_{j=1}^{m-1} p_j \le m < \frac{N}{2}$. А в случае четного N согласно (5) также $\sum_{j=1}^{m-1} p_j < \frac{N}{2}$. Поэтому из (23) с учетом неравенства $x_m > x_{m-1}$ получим $\frac{x_m}{N}\cdot\frac{N}{2} > \frac{\overline{x}}{2}$, т.е. $x_m > \overline{x}$. Вместе с предыдущим это дает противоречие, которое в случае однозначности медианы доказывает нашу теорему.

Пусть медиана определяется неоднозначно (в случае выполнения (6)). Тогда рассматривая элемент $x \in (x_k, x_{k+1})$ как выпуклую комбинацию x_k и x_{k+1}, а также используя отмеченные выше соотношения для x_k и x_{k+1} получим утверждение теоремы в случае неопределенной медианы. Теорема доказана.

Следствие 5. *В условиях теоремы 5 справедлива оценка*

$$\overline{x} - x_k \le \frac{2}{N} x_k \sum_{j=m}^{k} p_j.$$

В самом деле, справедливо следующее соотношение

$$\frac{\overline{x}}{2} \le \frac{1}{N}\sum_{j=1}^{k} x_j p_j \le \frac{x_k}{N}\sum_{j=1}^{k} p_j =$$

$$= \frac{1}{N} x_k \left(\sum_{j=1}^{m-1} p_j + \sum_{j=m}^{k} p_j\right).$$

Отсюда, с учетом неравенства $\sum_{j=1}^{m-1} p_j \le \frac{N}{2}$, вытекающего из (4), либо (5) получим требуемую оценку.

ЗАКЛЮЧЕНИЕ

В работе исследованы экстремальные свойства средних величин. Они представляют определенный интерес, поскольку проясняют природу средних и используются в практических приложениях. В п. 1 предложен новый метод доказательства экстремального свойства медианы вариационного ряда. На основании этого метода в п. 2 доказано экстремальное свойство медианы дискретной случайной величины. Проведен сравнительный анализ, с точки зрения экстремальных свойств, различных определений медианы векторного вариационного ряда. Указанные в п. 3 соотношения между средними характеристиками вариационных рядов нам ранее не встречались. Их сложнее было увидеть, чем доказать.

Отметим еще близкую по тематике работу автора [7], в которой обсуждается связь между медианами и математическими ожиданиями непрерывных случайных величин.

СПИСОК ЛИТЕРАТУРЫ

1. *Джини К.* Средние величины / Пер. с итал. М. : Статистика, 1970. – 447 с.

2. *Венецкий И. Г.* Вариационные ряды и их характеристики. М. : Ста-тистика, 1970. – 160 с.

3. Общая теория статистики / под ред. А. Я. Боярского, Г. Л. Громыко. – М. : МГУ, 1985.

4. *Гнеденко Б. В.* Курс теории вероятностей. М. : Наука, 2005. – 448 с.

5. *Хацкевич В. Л.* О некоторых экстремальных свойствах средних значений и математических ожиданий случайных величин // Вестник Воронежского государственного технического университета. – 2013, т. 9, № 3.1. – С. 39–44.

6. *Королюк В. С., Портенко М. И., Скороход А. В., Турбин А. Ф.* Справочник по теории вероятностей и математической статистике. М. : Наука, 1985. – 640 с.

7. *Хацкевич В. Л., Агранович Ю. Я.* Об экстремальных взаимосвязях медиан и средних // Вестник Воронежского государственного технического университета. – 2013, т. 9, № 6.1. – С. 31–33.

Хацкевич Владимир Львович – заведующий кафедрой математики и информатики Института заочного экономического образования Воронежского государственного университета, д.т.н., профессор. E-mail: vlkhats@mail.ru

Khatskevich Vladimir Lvovich – head of the Department of mathematics and information mathematics Institute of correspondence education in Economics, Voronezh state University, doctor of technical Sciences, Professor.
E-mail: vlkhats@mail.ru