ФГБОУ ВО
Московский государственный университет имени М. В. Ломоносова
Механико-математический факультет

На правах рукописи
УДК 519.212.2

Хиль Елена Викторовна

Распределения функционалов от совокупностей локальных максимумов в последовательностях случайных величин

01.01.05 – теория вероятностей и математическая статистика

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата физико-математических наук

Москва 2016

Работа выполнена на кафедре математической статистики и случайных процессов механико-математического факультета Московского государственного университета имени М.В. Ломоносова.

Научный руководитель: доктор физико-математических наук
Зубков Андрей Михайлович

Официальные оппоненты:
доктор физико-математических наук, профессор
Гамкрелидзе Николай Георгиевич, профессор кафедры высшей математики ФГБОУ ВО «Российский государственный университет нефти и газа имени И.М. Губкина» (национальный исследовательский университет),

кандидат физико-математических наук
Меженная Наталья Михайловна, доцент кафедры прикладной математики ФГБОУ ВО «Московский государственный технический университет имени Н. Э. Баумана» (национальный исследовательский университет)

Ведущая организация:
ФГБУН «Институт прикладных математических исследований Карельского научного центра РАН»

Защита диссертации состоится «10» июня 2016 года в 16 часов 45 минут на заседании диссертационного совета Д 501.001.85 на базе МГУ имени М.В. Ломоносова по адресу: 119991, ГСП-1, Москва, Ленинские Горы, МГУ, механико-математический факультет, аудитория 16-24.

С диссертацией можно ознакомиться в Фундаментальной библиотеке ФГБОУ ВО «Московский государственный университет имени М. В. Ломоносова» по адресу Москва, Ломоносовский проспект, 27, сектор А, и на сайте механико-математического факультета: http://mech.math.msu.su/~snark/index.cgi.

Автореферат разослан «   » апреля 2016 года.

Ученый секретарь диссертационного
совета Д 501.001.85 на базе МГУ,
доктор физико-математических наук, профессор

Власов Виктор Валентинович

Общая характеристика работы

Диссертация посвящена изучению структуры последовательности случайных величин в терминах локальных максимумов.

Пусть $\{\xi_n, n \in \mathbb{Z}\}$ – последовательность одинаково распределённых случайных величин, $\chi_n = \mathrm{I}\{\xi_{n-1} < \xi_n > \xi_{n+1}\}$, $n \in \mathbb{Z}$, – индикаторы локальных максимумов (пиков) этой последовательности, $\{\tau_j,\, j \in \mathbb{Z}\} = \{n \in \mathbb{Z} : \chi_n = 1\}$ – возрастающая последовательность всех моментов появления локальных максимумов в $\{\xi_n\}$ и $\lambda_j = \tau_j - \tau_{j-1}$ – длины промежутков (расстояния) между соседними локальными максимумами.

Отметим, что для определения случайных величин τ_j (а, следовательно, и λ_j) не обязательно знать точные значения величин ξ_n, достаточно знать только отношения "больше-меньше" между соседними членами последовательности $\{\xi_n\}$, т.е. последовательности $\{\tau_j\}$ и $\{\lambda_j\}$ можно построить по сильно огрубленным данным.

Актуальность темы

Исследование различных характеристик последовательности случайных величин – классическая задача теории вероятностей. В частности, представляют интерес характеристики, связанные с отрезками монотонности в последовательности случайных величин: промежутки возрастания, убывания, локальные максимумы.

Вопрос о количестве локальных максимумов в последовательности случайных величин рассматривался при разных предположениях об их распределениях.

Например, А. М. Зубков и Д. В. Шуваев рассматривали[1] конечные последовательности перестановочных случайных величин (имеющих произвольное распределение), для которых нашли первые моменты количества пиков, впадин, возрастаний, убываний как выражения от простых характеристик совместных распределений.

Для простого симметричного случайного блуждания в статье[2] найдены количества траекторий с ровно k пиками (локальными максимумами), а так-

[1] Зубков А. М. , Шуваев Д. В. Вычисление моментов комбинаторных статистик от перестановочных случайных величин // Дискрет. Матем. 2005. Т. 17. Вып. 2. С. 3–18.

[2] Labarbe, J., Marckert, J. Asymptotics of Bernoulli random walks, bridges, excursions and meanders with a given number of peaks [Электронный ресурс] // Electron. J. Probab. 2007. Vol. 12. Pp. 229-261. Режим доступа: http://ejp.ejpecp.org/article/view/397.

же исследовано асимптотическое распределение количества пиков.

Хорошо известно, что при изучении таких функций от последовательности случайных величин, значения которых зависят только от сигнатуры этой последовательности (знания отношений “больше-меньше” между соседними членами последовательности), можно использовать переход к случайным перестановкам.

Связь между последовательностью $\xi_1, \ldots, \xi_n$ независимых случайных величин, имеющих одно и то же непрерывное распределение, и случайной перестановкой задается следующим образом: сопоставим числу k номер случайной величины ξ_k в вариационному ряду $\xi_{(1)} < \xi_{(2)} < \cdots < \xi_{(n-1)} < \xi_{(n)}$, построенном по $\xi_1, \ldots, \xi_n$. То есть $\sigma_k = s : \xi_{(1)} < \cdots < \xi_{(s-1)} < \xi_k < \xi_{(s+1)} < \cdots < \xi_{(n)}$. Тогда случайная перестановка σ имеет равномерное распределение на множестве всех перестановок порядка n. Таким образом, вероятности событий, определяемых сигнатурой последовательности $\xi_1, \ldots, \xi_n$, можно вычислять, подсчитывая количество перестановок, обладающих соответствующими свойствами.

И наоборот, для вычисления количества перестановок, обладающих заданным свойством, может быть удобно[3] рассмотреть последовательность случайных величин и вычислить соответствующую вероятность.

Заметим, что количество локальных максимумов в перестановке $\sigma_1 \ldots \sigma_n$ тесно связано с максимальной длиной альтернирующей подпоследовательности (такой набор $i_1 < i_2 < \cdots$, что $\sigma_{i_1} < \sigma_{i_2} > \sigma_{i_3} < \cdots$), изучению которой (в том числе при дополнительных ограничениях на исходную перестановку) посвящено большое количество работ (например, статьи[4,5]).

Изучение не только количества, но и структуры множества локальных максимумов в перестановке – более сложная задача, отметим здесь работу[6], в которой изучается множество перестановок с фиксированным множеством значений локальных максимумов (например, перестановками порядка 5 с множеством значений локальных максимумов $\{4, 5\}$ являются

[3] Szpiro G. G. The number of permutations with a given signature, and the expectations of their elements // Discrete Math. 2001. Vol. 226. Is. 1–3. Pp. 423–430.

[4] Widom H. On the limiting distribution for the length of the longest alternating sequence in a random permutation [Электронный ресурс] // Electron. J. Combin. 2006. 13.1. #R25. Режим доступа: http://www.combinatorics.org/ojs/index.php/eljc/article/view/v13i1r25.

[5] Romik D. Local extrema in random permutations and the structure of longest alternating subsequences // DMTCS Proceedings, 23rd International Conference on Formal Power Series and Algebraic Combinatorics. 2011. Pp. 825–834.

[6] Bouchard P., Hungyung Chang, Jun Ma, Jean Yeh, Yeong-Nan Yeh. Value-Peaks of Permutations [Электронный ресурс] // Electron. J. Combin. 2010. Vol. 17. #R.46. Режим доступа: http://www.combinatorics.org/ojs/index.php/eljc/article/view/v17i1r46.

14253, 35241, 34152, …), а также статью[7], в которой показано, что количество перестановок порядка n, имеющих заданное множество S моментов появления локальных максимумов имеет вид $p(n)2^{n-|S|-1}$, где $p(n)$ – полином, зависящий от множества S (точные формулы для $p(n)$ получены только для некоторых множеств S).

Изучение локальных максимумов представляет интерес и с точки зрения практических приложений, например, авторы статьи[8] рассматривают применение локальных максимумов в качестве точек доступа для разметки текста. Строка (текст) заменяется последовательностью локальных максимумов и расстояний между ними, что позволяет сравнивать строки быстрее и использовать меньше памяти (математическое ожидание числа максимумов зависит от мощности алфавита, но не превышает 1/3 длины текста, при этом задача минимизации количества локальных максимумов в тексте за счет выбора порядка букв в алфавите NP-сложна). Авторы применяют свои результаты к поиску фиксированной последовательности в ДНК.

В статье А. Кукетаева[9] использовано несколько стандартных статистик, основанных на расстояниях λ_j между соседними локальными максимумами в последовательности независимых случайных величин $\{\xi_n, n \in \mathbb{Z}\}$, имеющих одно и то же непрерывное распределение, для тестирования некоторых генераторов случайных чисел, при этом неявно предполагается, что $\{\lambda_j\}$ – последовательность независимых одинаково распределённых случайных величин.

В дипломной работе Н. А. Харитоновой[10] было показано, что предположение о независимости случайных величин λ_j неверно. С помощью комбинаторных методов в этой работе (см. также [1]) была доказана следующая

Теорема A. *Если случайные величины $\{\xi_n, n \in \mathbb{Z}\}$ независимы и имеют одно и то же непрерывное распределение, то*

$$\mathbf{P}\{\lambda = k\} = \frac{3(k+2)(k-1)2^k}{(k+3)!}, \quad k = 2, 3, \ldots, \tag{1}$$

$$\mathbf{M}\lambda = 3, \quad \mathbf{D}\lambda = 3(e^2 - 7) \approx 1,167, \quad \mathbf{P}\{\lambda > 6\} \approx 0.0074.$$

[7]Billey S., Burdzy K., Sagan B. E. Permutations with Given Peak Set // J. Integer Seq. 2013. Vol. 16. Is. 6. Article 13.6.1.

[8]Crescenzi P., Lungo A. D., Grossi R., Lodi E., Pagli L., Rossi G. Text sparsification via local maxima // Theor. Comput. Sci. 2003. Vol. 304. Is. 1–3. Pp. 341–364.

[9]Kuketayev A. Probability distribution of distances between local extrema of random number series // Вестник Караганд. ун-та, сер. Физика. 2011. Т. 62. № 2. С. 21–34; arXiv:math/0611130, 2006.

[10]Харитонова Н. А. Распределения расстояний между локальными максимумами случайной последовательности. — Дипл. работа, мех.-мат. ф-т МГУ имени М. В. Ломоносова. 2009. 15 с.

Кроме того,

$$\mathbf{P}\{\lambda_1 = k, \lambda_2 = l\} = \mathbf{P}\{\lambda_1 = l, \lambda_2 = k\} = \tag{2}$$
$$= \frac{3}{(k+l+3)!} \sum_{m=1}^{k+l+2} \sum_{n=1}^{k+2} C_{m-1}^{n-1} C_{k+l+3-m}^{k+3-n} P(k,n) Q(l, m-n+1),$$

где

$$P(k,n) = \begin{cases} k2^k - (2k-n+3)2^{n-3}, & \textit{если} \;\; n \geq 3, \\ k2^k - (k+1), & \textit{если} \;\; n \in \{1,2\}; \end{cases}$$
$$Q(l,m) = \begin{cases} (2l-m-1)2^{m-3}, & \textit{если} \;\; m \geq 3, \\ l-1, & \textit{если} \;\; m \in \{1,2\}. \end{cases}$$

Формулы (2) можно упростить, если одна из длин равна 2 или 3:

$$\mathbf{P}\{\lambda_1 = k, \lambda_2 = 2\} = \frac{(k^2+9k+12)(k+4)(k-1)2^k}{(k+5)!},$$
$$\mathbf{P}\{\lambda_1 = k, \lambda_2 = 3\} = \frac{(k+2)(k-1)2^k}{(k+3)!} = \mathbf{P}\{\lambda_1 = k\}\mathbf{P}\{\lambda_2 = 3\}.$$

Из упрощенных формул видно, что события $\{\lambda_j = k\}$ и $\{\lambda_{j+1} = 2\}$ зависимы, таким образом случайные величины λ_j зависимы. При этом событие $\{\lambda_j = 3\}$ не зависит от значений λ_{j-1} и λ_{j+1}.

В настоящей работе уточнён характер зависимости между случайными величинами λ_j, получены точные формулы для совместных вероятностей (2) и для совместных распределений соседних локальных максимумов и длин промежутков между ними. Описана асимптотическая структура промежутка между соседними локальными максимумами при условии, что его длина стремится к бесконечности.

Также показана асимптотическая нормальность совместных распределений количеств появления промежутков заданных длин и предложен статистический критерий проверки гипотезы о вероятностной структуре исходной последовательности.

Кроме того, в работе найдены вероятности появления локальных максимумов и распределение расстояний между ними для некоторых последовательностей зависимых или не одинаково распределённых случайных величин.

Цели работы

Исследование структурных свойств последовательностей случайных величин в терминах расстояний между моментами появления локальных максимумов.

Изучение возможности использования этих характеристик для проверки гипотез о вероятностной структуре исходной последовательности.

Научная новизна

Основные результаты являются новыми, получены автором самостоятельно и состоят в следующем:

1. Найдены совместные распределения соседних локальных максимумов и длин промежутков между ними.

2. Описана асимптотическая структура промежутка между соседними локальными максимумами при условии, что его длина стремится к бесконечности.

3. Доказана асимптотическая нормальность совместных распределений количеств появления промежутков заданных длин. Предложен статистический критерий проверки гипотезы о вероятностной структуре исходной последовательности.

4. Найдены вероятности появления локальных максимумов и распределение расстояний между ними для некоторых последовательностей зависимых или не одинаково распределённых случайных величин.

Основные методы исследования

В работе используются аналитические и комбинаторные методы теории вероятностей.

Теоретическая и практическая ценность работы

Работа имеет теоретический характер, в частности, обнаружены новые свойства последовательности независимых одинаково распределённых случайных величин. Полученные результаты могут использоваться при статистической обработке огрублённых наблюдений (достаточно, чтобы были известны только отношения "больше-меньше" между соседними наблюдениями), а также при тестировании датчиков случайных чисел.

Апробация работы

Результаты диссертации докладывались автором на научном семинаре кафедры математической статистики и случайных процессов механико-математического факультета МГУ имени М. В. Ломоносова (Москва, 2009–2015 гг.); Десятом Всероссийском симпозиуме по прикладной и промышленной математике (осенняя сессия) (Сочи, 2009г.); 9-й международной конференции «Computer Data Analysis and Modeling: Theoretical and Applied Stochastics» (Минск, 2010 г.); Восемнадцатой Всероссийской школе-коллоквиуме по стохастическим методам (Казань, 2011г.); Международной конференции «Теория вероятностей и ее приложения», посвященной 100-летию со дня рождения Б.В. Гнеденко (Москва, 2012г.); 10-й международной конференции «Computer Data Analysis and Modeling: Theoretical and Applied Stochastics» (Минск, 2013 г.).

Публикации

Основные результаты диссертации опубликованы в 8 работах (из них 2 в журналах из перечня ВАК), список которых приведён в конце автореферата.

Структура и объём работы

Диссертация состоит из введения, трёх глав, заключения и списка литературы из 43 наименований. Общий объём диссертации составляет 98 страниц.

Краткое содержание диссертации

Во **введении** приведён обзор литературы по тематике работы, изложены цели исследования, а также перечислены основные полученные результаты.

Первая глава диссертации посвящена изучению последовательности локальных максимумов и расстояний $\{\lambda_j\}$ между ними для независимых случайных величин $\{\xi_n\}$, имеющих одно и то же непрерывное распределение.

В параграфе 1.1 содержатся основные определения и обозначения.

В параграфе 1.2 показано, что теорему A (и ряд других результатов) можно обобщить на случай перестановочных случайных величин $\{\xi_n\}$, удовлетворяющих условию $\mathbf{P}\{\xi_1 = \xi_2\} = 0$.

Лемма 1. *Пусть* $\{\xi_n, n \in \mathbb{Z}\}$ *– такая последовательность перестановочных случайных величин, что* $\mathbf{P}\{\xi_1 = \xi_2\} = 0$.

Тогда последовательность $\{\tau_j,\ j \in \mathbb{Z}\}$ *моментов появления локальных максимумов в* $\{\xi_n\}$ *имеет такое же распределение, как последовательность моментов появления локальных максимумов в последовательности независимых случайных величин, имеющих равномерное распределение на отрезке* $[0, 1]$.

В параграфе 1.3 доказана следующая

Теорема 1. *Пусть случайные величины* ξ_n, $n \in \mathbb{Z}$, *независимы и имеют одно и то же непрерывное распределение.*

Тогда последовательность пар $(\lambda_j, \xi_{\tau_j})$, $j \in \mathbb{Z}$, *образует однородную по времени цепь Маркова.*

Параграф 1.4 посвящен исследованию характеристик цепи $(\lambda_j, \xi_{\tau_j})$ в случае, когда исходные случайные величины $\{\xi_n\}$ имеют равномерное распределение на отрезке $[0, 1]$.

Рассмотрим два соседних локальных максимума $\xi_{\tau_{j-1}}, \xi_{\tau_j}$. Введём обозначения $\xi_{\min} = \min\{\xi_{\tau_{j-1}},\ \xi_{\tau_j}\}$, $\xi_{\max} = \max\{\xi_{\tau_{j-1}}, \xi_{\tau_j}\}$.

Лемма 2. *Условная плотность* $g_k(x, y)$ *распределения вектора* $(\xi_{\min}, \xi_{\max})$ *при условии, что расстояние* $\lambda_j = k$, *имеет вид*

$$g_k(x, y) = \frac{k(k+1)(k+3)}{(k-1)2^k} xy \left((y+x)^{k-1} - (y-x)^{k-1}\right), 0 \le x \le y \le 1. \quad (3)$$

Эта лемма позволила получить следующие теоремы 2 и 3.

Теорема 2. *Пусть случайные величины* $\{\xi_n, n \in \mathbb{Z}\}$ *независимы и равномерно распределены на отрезке* $[0, 1]$.

Тогда переходная плотность цепи $(\lambda_j, \xi_{\tau_j})$ *равна*

$$p_{\left(\lambda_{j+1}, \xi_{\tau_{j+1}}\right) | \left(\lambda_j, \xi_{\tau_j}\right)} (k, y \,|\, l, x) = \frac{1}{2(k-1)!} \frac{y\left((y+x)^{k-1} - |y-x|^{k-1}\right)}{x},$$
$$0 \le x, y \le 1.$$

В теореме А формула для совместного распределения длин соседних промежутков представлена в виде двойной суммы (это представление получено комбинаторными методами). С помощью прямых вероятностных методов (искомая вероятность представляется в виде интеграла от условной вероятности), в диссертации получена явная формула.

Теорема 3. *Пусть случайные величины* $\{\xi_n, n \in \mathbb{Z}\}$ *независимы и имеют одно и то же непрерывное распределение.*

Тогда для $k = 2, 3, \ldots,\ l = 2, 3, \ldots$

$$\mathbf{P}\{\lambda_1 = k, \lambda_2 = l\} = \mathbf{P}\{\lambda_1 = l, \lambda_2 = k\} =$$
$$= \frac{3 \cdot 2^{k+l-1}\left(C_{k+l}^{l}(kl(k+l+1) + 1 - k^2 - l^2) - (k+l-1)(k+1)(l+1)\right)}{(l+1)(k+1)(k+l+3)(k+l+1)!}, \tag{4}$$

$$\mathbf{P}\{\lambda_{j-1} = k, \lambda_j = 2, \lambda_{j+1} = l\} = \mathbf{P}\{\lambda_{j-1} = l, \lambda_j = 2, \lambda_{j+1} = k\} =$$
$$= \frac{3 \cdot 2^{k+l}}{(l+1)!(k+1)!}\left(\frac{kl(k+l+5)^2 - (k+l+3)((k+l)^2 + 6(k+l) - 3)}{(k+3)(l+3)(k+l+3)(k+l+5)} + \frac{2(k+l+1)}{(k+l+3)(k+l+5)C_{k+l+2}^{l+1}}\right). \tag{5}$$

Формулы (4) и (5) позволяют найти отношение условных вероятностей

$$\frac{\mathbf{P}\{\lambda_3 = l \,|\, \lambda_1 = k, \lambda_2 = 2\}}{\mathbf{P}\{\lambda_3 = l \,|\, \lambda_2 = 2\}} = \frac{6}{5}\left(1 + O\left(\frac{1}{k} + \frac{1}{l}\right)\right), \quad k, l \to \infty,$$

которое показывает, что последовательность расстояний $\{\lambda_j\}$ не является цепью Маркова.

Во **второй главе** изучается структура последовательности независимых случайных величин ξ_n, имеющих одно и то же непрерывное распределение, с выделенными локальными максимумами.

В параграфе 2.1 описана структура "оврага" (последовательно идущие убывающий и возрастающий участки последовательности случайных величин) и его асимптотическое поведение.

Теорема 4. *Пусть случайные величины* $\xi_n, n \in \mathbb{N}$, *независимы и имеют равномерное распределение на отрезке* $[0, 1]$, *функции* $y_k(x) = |1 - 2x/(k+1)|$, $V_1^k = \{\exists\, t : \xi_1 > \cdots > \xi_t < \cdots < \xi_k,\ 1 \le t \le k\}$. *Тогда*

$$\forall\, \varepsilon > 0 \qquad \mathbf{P}\{\max_{n=1,\ldots,k} |\xi_n - y_k(n)| > \varepsilon \,|\, V_1^k\} \to 0, \quad k \to \infty.$$

Аналогичная теорема справедлива и для участка последовательности между соседними локальными максимумами.

Теорема 5. *Пусть случайные величины* $\xi_n, n \geq 0$, *независимы и имеют равномерное распределение на отрезке* $[0,1]$, *функции* $y_k(x) = |1 - 2x/(k+1)|$. *Тогда*

$$\forall \varepsilon > 0 \qquad \mathbf{P}\{\max_{n=1,\ldots,k} |\xi_n - y_k(n)| > \varepsilon \,|\, W_1^k\} \to 0, \quad k \to \infty,$$

$$W_1^k = \{\exists\, t : \xi_0 < \xi_1 > \xi_2 > \cdots > \xi_t < \cdots \xi_{k-1} < \xi_k > \xi_{k+1}\}.$$

Следствие 1. *Если случайные величины* $\xi_n, n \geq 0$, *независимы и имеют одно и то же непрерывное распределение с функцией распределения* $F(x)$, *то*

$$\forall \varepsilon > 0 \qquad \mathbf{P}\{\max_{n=1,\ldots,k} |F(\xi_n) - y_k(n)| > \varepsilon \,|\, W_1^k\} \to 0, \quad k \to \infty.$$

Если, кроме того, функция $F(x)$ *строго монотонна, то*

$$\forall \varepsilon > 0 \qquad \mathbf{P}\{\max_{n=1,\ldots,k} |\xi_n - F^{-1}(y_k(n))| > \varepsilon \,|\, W_1^k\} \to 0, \quad k \to \infty.$$

В параграфе 2.2 показано, что событие $\{\lambda = 3\}$ является рекуррентным событием с запаздыванием как для исходной последовательности $\{\xi_n\}_{n=1}^{\infty}$, так и для последовательности $\{\chi_n\}_{n=2}^{\infty}$ индикаторов локальных максимумов, найдены производящие функции для вероятностей f_k, b_k того, что рекуррентное событие произошло впервые при k-м испытании (в регулярной части и в начальном отрезке) и вероятности u_k того, что рекуррентное событие произошло при k-м испытании

$$F(s) = \sum_{k=0}^{\infty} s^k f_k = 1 - (1-s)\left(1 - s + \frac{s^3}{3} - \frac{s^4}{3} + \frac{2s^5}{15} - \frac{s^6}{45}\right)^{-1},$$

$$B(s) = \sum_{k=0}^{\infty} s^k b_k = \frac{1}{9} s^6 \left(1 - s + \frac{s^3}{3} - \frac{s^4}{3} + \frac{2s^5}{15} - \frac{s^6}{45}\right)^{-1},$$

$$U(s) = \sum_{k=0}^{\infty} s^k u_k = \frac{1}{9} \sum_{k=6}^{\infty} s^k = \frac{s^6}{9(1-s)}.$$

Утверждение 1. *Пусть* ρ *– расстояние между моментами появления событий* $\{\lambda = 3\}$, *а* ρ_0 *– время ожидания первого момента появления события* $\{\lambda = 3\}$ *в последовательности* $\{\xi_n,\, n \in \mathbb{N}\}$. *Тогда*

$$\mathbf{M}\rho = 9, \quad \mathbf{D}\rho = 57,6; \qquad \mathbf{M}\rho_0 = 13,2, \quad \mathbf{D}\rho_0 = 32,04.$$

В параграфе 2.3 рассматриваются количество локальных максимумов и количества расстояний заданных длин между соседними локальными максимумами последовательности $\{\xi_i\}_{i=1}^T$,

$$N_0(T) = \sum_{i=1}^{T-2} \chi_i = \sum_{i=1}^{T-2} \mathrm{I}\{\xi_{i-1} < \xi_i > \xi_{i+1}\},$$

$$N_k(T) = \sum_{i=1}^{T-2-k} \mathrm{I}\{\chi_i = \chi_{i+k} = 1, \chi_j = 0, j = i+1, \ldots, i+k-1\}, \quad k \geq 2,$$

$$N^{(s)} = \left(N_0(T), N_2(T), N_3(T), \ldots, N_s(T)\right).$$

Для вектора $N^{(s)}$ справедлива следующая предельная теорема.

Теорема 6. *Для любого натурального s вектор $N^{(s)}$ асимптотически нормален при $T \to \infty$ с параметрами $(A^{(s)}T, C_sT)$, где*

$$A^{(s)} = \frac{1}{3}(1, \mathbf{P}\{\lambda = 2\}, \ldots, \mathbf{P}\{\lambda = s\}), \tag{6}$$

а элементы матрицы ковариаций $C_s = \|c_{km}\|_{k,m=1}^s$ определяются формулами

$$c_{11} = \frac{2}{45},$$
$$c_{1k} = c_{k1} = \frac{2}{3}\mathbf{P}\{\lambda_1 = 2, \lambda_2 = k\} - \frac{k-1}{9}\mathbf{P}\{\lambda = k\}, \quad k \geqslant 2,$$
$$c_{kk} = \frac{1}{3}\mathbf{P}\{\lambda = k\} - \frac{2k+5}{9}\mathbf{P}^2\{\lambda = k\} + \frac{2}{3}\mathbf{P}\{\lambda_1 = \lambda_2 = k\}$$
$$+\frac{2}{3}\mathbf{P}\{\lambda_1 = k, \lambda_2 = 2, \lambda_3 = k\}, \quad k \geqslant 2,$$
$$c_{mk} = c_{km} = -\frac{k+m+5}{9}\mathbf{P}\{\lambda = k\}\mathbf{P}\{\lambda = m\} + \frac{2}{3}\mathbf{P}\{\lambda_1 = k, \lambda_2 = m\}$$
$$+\frac{2}{3}\mathbf{P}\{\lambda_1 = k, \lambda_2 = 2, \lambda_3 = m\}, \quad m > k \geqslant 2.$$

Зависимость случайных величин $\{\lambda_j\}$ делает применение стандартных статистических критериев в статье А. Кукетаева[11] некорректным, однако теорема 6 позволяет строить критерии согласия для гипотезы

H_0: $\{\xi_j\}$ – последовательность перестановочных случайных величин, таких что $\mathbf{P}\{\xi_1 = \xi_2\} = 0$.

[11] Kuketayev A. Probability distribution of distances between local extrema of random number series // Вестник Караганд. ун-та, сер. Физика. 2011. Т. 62. № 2. С. 21–34; arXiv:math/0611130, 2006.

Например, для критерия согласия с гипотезой H_0, имеющего асимптотический уровень значимости α, можно использовать следующее правило:

$$H_0 \text{ отклоняется } \iff \frac{1}{T}\left(C_s^{-1}\left(N^{(s)}(T)-A^{(s)}T\right), N^{(s)}(T)-A^{(s)}T\right) > u_{1-\alpha},$$

где u_α – α-квантиль распределения хи-квадрат с s степенями свободы, $(\cdot,\cdot)$ – скалярное произведение в $\mathbb{R}^s$.

Сформулируем также частный случай теоремы 6.

Теорема 7. *Вектор* $(N_0(T), N_2(T), N_3(T))$ *при* $T \to \infty$ *асимптотически нормален с параметрами*

$$A^{(3)} = T\left(\frac{1}{3}, \frac{2}{15}, \frac{1}{9}\right), \quad C_3 = T\begin{pmatrix} \frac{2}{15} & \frac{20}{315} & \frac{2}{135} \\ \frac{20}{315} & \frac{1772}{14175} & -\frac{22}{945} \\ \frac{2}{135} & -\frac{22}{945} & \frac{32}{405} \end{pmatrix}.$$

В **третьей главе** рассмотрено несколько видов последовательностей случайных величин, не являющих перестановочными.

В параграфе 3.1 изучаются распределения расстояний между соседними локальными максимумами и количества расстояний заданных длин между соседними локальными максимумами в последовательности скользящих сумм.

Теорема 8. *Пусть* $\{\xi_n, n \in \mathbb{Z}\}$ *– последовательность независимых случайных величин, имеющих одно и то же непрерывное распределение, и случайные величины* $\xi_n^* = \xi_n + \xi_{n+1}$, $n \in \mathbb{Z}$. *Пусть* λ^* *– длина промежутка между соседними локальными максимумами последовательности* $\{\xi_n^*\}$.

Тогда для $l = 1, 2, \ldots$

$$\mathbf{P}\{\lambda^* = 2l\} = \frac{4}{[(l+2)!]^2}\left(l^2 - \frac{2}{l+2}\right) C_{2l+1}^l,$$
$$\mathbf{P}\{\lambda^* = 2l+1\} = \frac{8l(l+1)}{(l+2)!(l+3)!} C_{2l+1}^l.$$

Компьютерные вычисления показывают, что

$$\mathbf{M}\lambda^* = 4, \quad \mathbf{D}\lambda^* \approx 2,117, \quad \mathbf{P}\{\lambda^* > 8\} \approx 0.0081.$$

Пусть $N_k^*(T)$ – случайные величины, аналогичные $N_k(T)$,

$$N_0^*(T) = \sum_{i=1}^{T-2} \chi_i^* = \sum_{i=1}^{T-2} \mathrm{I}\{\xi_{i-1}^* < \xi_i^* > \xi_{i+1}^*\},$$

$$N_k^*(T) = \sum_{i=1}^{T-2-k} \mathrm{I}\{\chi_i^* = \chi_{i+k}^* = 1, \chi_j^* = 0, j = i+1, \ldots, i+k-1\}, \quad k \geq 2.$$

Теорема 9. *Вектор* $(N_0^*(T), N_2^*(T), N_3^*(T))$ *при* $T \to \infty$ *асимптотически нормален с параметрами*

$$T\left(\frac{1}{4}, \frac{1}{36}, \frac{1}{12}\right), \qquad T\begin{pmatrix} \frac{5}{144} & \frac{5}{288} & \frac{5}{144} \\ \frac{5}{288} & \frac{52}{2025} & \frac{1}{720} \\ \frac{5}{144} & \frac{1}{720} & \frac{259}{3456} \end{pmatrix}.$$

Сравнение теорем 7 и 9 показывает, что с помощью рассмотренных в них статистик можно построить критерий различения гипотезы H_0 и описанной в теореме 8 альтернативы.

В параграфе 3.2 рассматривается последовательность взвешенных сумм $\widehat{\xi}_n^c = \xi_n + c\xi_{n+1}$, $n \in \mathbb{Z}$, $c = const \neq 0$, где случайные величины $\{\xi_n, n \in \mathbb{Z}\}$ независимы и имеют одно и то же непрерывное распределение (при $c = 0$ $\widehat{\xi}_n^c = \xi_n$, и этот случай подробно изучен в первой главе).

При $c = 1$ получаем последовательность $\{\xi_i^*\}$, рассмотренную в параграфе 3.1. По теореме 8 распределение расстояния между соседними локальными максимумами в этой последовательности не зависит от вида непрерывного распределения случайных величин ξ_n, то же справедливо и для вероятности появления локального максимума в фиксированный момент времени. Однако при $c \neq 1$ это не верно и вероятность $\mathbf{P}\{\widehat{\xi}_{n-1}^c < \widehat{\xi}_n^c > \widehat{\xi}_{n+1}^c\}$ зависит от распределения случайных величин ξ_n.

Утверждение 2. *Если случайные величины* ξ_n *имеют равномерное распределение на отрезке* $[0,1]$, *то для любого целого* n

$$\mathbf{P}\{\widehat{\xi}_{n-1}^c < \widehat{\xi}_n^c > \widehat{\xi}_{n+1}^c\} = \begin{cases} \frac{1}{3} - \frac{c}{12(c^2 - c + 1)}, & c < 0, \\ \frac{1}{3} - \frac{c + c^2 - c^3}{12}, & 0 \leq c < 1, \\ \frac{1}{3} - \frac{c^2 + c - 1}{12c^3}, & c \geq 1. \end{cases} \tag{7}$$

Если случайные величины ξ_n имеют стандартное нормальное распределение, то для любого целого n

$$\mathbf{P}\{\widehat{\xi}^c_{n-1} < \widehat{\xi}^c_n > \widehat{\xi}^c_{n+1}\} = \frac{1}{\pi}\operatorname{arctg}\sqrt{\frac{3c^2-4c+3}{c^2+1}}. \tag{8}$$

В параграфе 3.3 найдена вероятность появления локального максимума в фиксированный момент времени в частном случае последовательности независимых, но не одинаково распределённых величин.

Утверждение 3. *Пусть $\{\xi_n,\, n \in \mathbb{Z}\}$ – последовательность независимых случайных величин, имеющих равномерное распределение на отрезке $[0,1]$, и пусть $\widetilde{\xi}_n = \xi_n + vn$, $n \in \mathbb{Z}$, $v = const$.*

Тогда для любого целого n

$$\mathbf{P}\{\widetilde{\xi}_{n-1} < \widetilde{\xi}_n > \widetilde{\xi}_{n+1}\} = \begin{cases} \dfrac{1}{3} - v^2\dfrac{9-8|v|}{6}, & |v| < \dfrac{1}{2}, \\ \dfrac{(1-|v|)^2}{2}, & \dfrac{1}{2} \le |v| < 1, \\ 0, & |v| \ge 1. \end{cases}$$

Заключение

Основные результаты диссертации состоят в следующем:

1. Найдены совместные распределения соседних локальных максимумов и длин промежутков между ними.

2. Описана асимптотическая структура промежутка между соседними локальными максимумами при условии, что его длина стремится к бесконечности.

3. Доказана асимптотическая нормальность совместных распределений количеств появления промежутков заданных длин. Предложен статистический критерий проверки гипотезы о вероятностной структуре исходной последовательности.

4. Найдены вероятности появления локальных максимумов и распределение расстояний между ними для некоторых последовательностей зависимых или не одинаково распределённых случайных величин.

Дальнейшие исследования могут быть связаны с изучением не рассмотренных в работе классов последовательностей зависимых или не одинаково распределённых случайных величин.

Также представляет интерес исследование расстояний между появлением тех или иных комбинаций в последовательностях случайных величин, имеющих дискретные распределения.

Благодарности. Автор глубоко благодарен своему научному руководителю д.ф.-м.н. Андрею Михайловичу Зубкову за постановку интересных задач, плодотворные обсуждения, ценные замечания и постоянное внимание к работе.

Публикации автора по теме диссертации

[1] Зубков А. М., Харитонова Н. А., Хиль Е. В. Расстояния между локальными максимумами в последовательностях случайных величин // Теория вероятн. и ее примен. — 2011. — Т. 56. — Вып. 4. — С. 690–703.

English transl.: Zubkov A. M., Kharitonova N. A., and Khil E. V. Gaps between Local Maxima in Sequences of Random Variables // Theory Probab. Appl. — 2012. — V. 56. — Is. 4. — P. 590–601.

[А. М. Зубкову принадлежат постановка задачи и научное руководство исследованиями, Н. А. Харитоновой принадлежит теорема 1 (случай независимых случайных величин), Е. В. Хиль принадлежат теорема 3 (асимптотическая нормальность вектора частот), теорема 4 и утверждения 1 и 2 (частные случаи зависимых случайных величин).]

[2] Хиль Е. В. Марковские зависимости в последовательности локальных максимумов и промежутков между ними // Теория вероятн. и ее примен. — 2013. — Т. 58. — Вып. 3. — С. 472–485.

English transl.: Khil E. V. Markov Structure of the Sequence of Local Maxima and the Gaps between Them // Theory Probab. Appl. — 2014. — Vol. 58. — No. 3. — P. 430–441.

[3] Хиль Е. В. Структура последовательности случайных величин с выделенными локальными максимумами. — М., 2016. — Деп. в ВИНИТИ РАН 19.02.2016, № 35-В2016.

[4] Зубков А. М., Харитонова Н. А., Хиль Е. В. Формулы для распределений расстояний между соседними локальными максимумами // Обозрение прикл. и промышл. матем. — 2009. — Т. 16. — Вып. 4. — С. 658–659. [А. М. Зубкову принадлежат постановка задачи и научное руководство исследованиями, Н. А. Харитоновой принадлежит теорема 1, Е. В. Хиль принадлежит теорема 2.]

[5] Kharitonova N. A., Khil E. V., Zubkov A. M. Goodness-of-fit test based on local maxima, // Proceedings of 9th International Conference "Computer Data Analysis and Modeling: Complex Stochastics Data and Systems" — Minsk. — 2010. — V. 1. — P. 74–76. [А. М. Зубкову принадлежат постановка задачи и научное руководство исследованиями, Н. А. Харитоновой принадлежит теорема 1, Е. В. Хиль принадлежат теоремы 2 и 3.]

[6] Зубков А. М., Хиль Е. В. Марковские зависимости в последовательности локальных максимумов // Обозрение прикл. и промышл. матем. — 2011. — Т. 18. — Вып. 1. — С. 83–84. [А. М. Зубкову принадлежит постановка задачи и научное руководство исследованиями, Е. В. Хиль принадлежат основные результаты.]

[7] Хиль Е. В. Расстояния между локальными максимумами в последовательностях случайных величин // Международная конференция "Теория вероятностей и ее приложения", посвященная 100-летию со дня рождения Б.В. Гнеденко (Москва, 26–30 июня 2012 года): Тезисы докладов. — М.: Ленанд. — 2012. — С. 71–72.

[8] Khil E. V. Markov properties of gaps between local maxima in a sequence of independent random variables // Proceedings of 10th International Conference "Computer Data Analysis and Modeling: Theoretical and Applied Stochastics" — Minsk. — 2013. — V. 2. — P. 13–16.