ДОКЛАДЫ РОССИЙСКОЙ АКАДЕМИИ НАУК. МАТЕМАТИКА, ИНФОРМАТИКА, ПРОЦЕССЫ УПРАВЛЕНИЯ, 2024, том 518, с. 89–92

МАТЕМАТИКА

УДК 517.54

ДОПОЛНЕНИЕ К СТАТЬЕ “НОВАЯ СПЕКТРАЛЬНАЯ МЕРА СЛОЖНОСТИ И ЕЁ ВОЗМОЖНОСТИ ПО ОБНАРУЖЕНИЮ СИГНАЛОВ В ШУМЕ”

 А. А. Галяев[1, *], В. Г. Бабиков[1, **], П. В. Лысенко[1, *], Л. М. Берлин[1, ****]**

Поступило 05.08.2024 г.
После доработки 05.08.2024 г.
Принято к публикации 06.08.2024 г.

Заметке-дополнение к статье “Новая спектральная мера сложности и ее возможности по обнаружению сигналов в шуме”.

Ключевые слова: информационная энтропия, спектральная сложность, белый гауссовский шум

DOI: 10.31857/S2686954324040133, **EDN:** YYBQQQ

В данной заметке-дополнении к статье [1] не только устанавливается дополнительная связь между Леммами 1 и 2 оригинальной работы, но и приводится новый результат, задающий формулу для точного вычисления $\mathbb{E}[\eta_k(N)]$.

Итак, для решения Задачи 1 в общем случае требуется уметь находить

$$\mathbb{E}[\eta_k(N)] = \mathbb{E}\left[\frac{(\mathbf{T}I)_k}{E_X}\right]. \tag{1}$$

Справедлива следующая лемма.

Лемма Д1.

Пусть $z_1,\dots,z_N$ — результаты наблюдений случайной величины Z, имеющей экспоненциальное распределение $F(z) = 1 - \exp(-z)$. Рассмотрим значения последовательности $z_{(1)},\dots,z_{(N)}$ тех же результатов, но расположенных в порядке возрастания, где случайная величина $Z_{(k)}$ является неубывающей k-ой порядковой статистикой.

Тогда

$$\mathbb{E}\left[\frac{Z_{(k)}}{\sum_{i=1}^{N} Z_{(i)}}\right] = \frac{\int_0^\infty \int_0^{z_N} \dots \int_0^{z_3} \int_0^{z_2} \frac{z_k}{\sum_{i=1}^{N} z_i} \exp\left(-\sum_{i=1}^{N} z_i\right) dz_1 dz_2 \dots dz_{N-1} dz_N}{\int_0^\infty \int_0^{z_N} \dots \int_0^{z_3} \int_0^{z_2} \exp\left(-\sum_{i=1}^{N} z_i\right) dz_1 dz_2 \dots dz_{N-1} dz_N}. \tag{2}$$

Лемма проверяется непосредственно записью математического ожидания для порядковой статистики.

Отличие формулы (2) от известных заключается в вычислении среднего величины $\frac{z_k}{\sum_{i=1}^{N} z_i}$ по совместному распределению порядковых статистик $Z_{(1)},\dots,Z_{(N)}$. Затруднение вызывает вычисление интеграла в числителе (2). Однако, преодолеть это препятствие помогает следующее утверждение.

Утверждение Д1.

Искомое математическое ожидание равно

[1] *Институт проблем управления им. В.А. Трапезникова Российской академии наук, Москва, Россия*

*E-mail: galaev@ipu.ru

**E-mail: babikov@ipu.ru

***E-mail: pavellysen@ipu.ru

****E-mail: berlin.lm@phystech.edu

$$\frac{\int_0^\infty \int_0^{z_N} \ldots \int_0^{z_3} \int_0^{z_2} \frac{z_k}{\sum_{i=1}^N z_i} \exp\left(-\sum_{i=1}^N z_i\right) dz_1 dz_2 \ldots dz_{N-1} dz_N}{\int_0^\infty \int_0^{z_N} \ldots \int_0^{z_3} \int_0^{z_2} \exp\left(-\sum_{i=1}^N z_i\right) dz_1 dz_2 \ldots dz_{N-1} dz_N} = \frac{1}{N}\left(\sum_{i=1}^N \frac{1}{i} - \sum_{i=1}^{N-k} \frac{1}{i}\right), \tag{3}$$

что в точности до множителя $\frac{1}{N}$ совпадает со значением другого математического ожидания

$$\frac{\int_0^\infty \int_0^{z_N} \ldots \int_0^{z_3} \int_0^{z_2} z_k \exp\left(-\sum_{i=1}^N z_i\right) dz_1 dz_2 \ldots dz_{N-1} dz_N}{\int_0^\infty \int_0^{z_N} \ldots \int_0^{z_3} \int_0^{z_2} \exp\left(-\sum_{i=1}^N z_i\right) dz_1 dz_2 \ldots dz_{N-1} dz_N} = \sum_{i=1}^N \frac{1}{i} - \sum_{i=1}^{N-k} \frac{1}{i}. \tag{4}$$

Формула (4) является известным результатом для нахождения математического ожидания порядковых статистик [2]. Результат утверждения получен с использование символьного пакета вычислений Maple.

Остается заметить, что в работе

$$\mathbb{E}\left[\frac{Y_{(N-k+1)}}{\sum_{i=1}^N Y_{(i)}}\right] = \mathbb{E}\left[\frac{Z_{(k)}}{\sum_{i=1}^N Z_{(i)}}\right],$$

или

$$\tilde{n}_k(N) = \mathbb{E}\left[\frac{Y_{(k)}}{\sum_{i=1}^N Y_{(i)}}\right] = \mathbb{E}\left[\frac{Z_{(N-k+1)}}{\sum_{i=1}^N Z_{(i)}}\right] = \frac{1}{N}\left(\sum_{i=1}^N \frac{1}{i} - \sum_{i=1}^{k-1} \frac{1}{i}\right). \tag{5}$$

Используя формулу (5), получаем, что

$$\sum_{k=1}^N \mathbb{E}\left[\frac{Y_{(k)}}{\sum_{i=1}^N Y_{(i)}}\right] = N \cdot \frac{1}{N} = 1.$$

Это означает, что условие нормировки выполнено.

Далее, изначально в работе Лемма 1 была выбрана в качестве базовой для перехода в Лемме 2 и к основной формуле работы (6) вида

$$n_k(N) = -\frac{1}{N \cdot K_N} \ln \frac{k}{N+1}, \quad \text{где } K_N = -\frac{1}{N}\sum_{k=1}^N \ln \frac{k}{N+1}, \tag{6}$$

поскольку, как указано в Лемме 1, $\mathbb{E}(\exp(-Y_{(k)})) = \frac{k}{N+1}$, а соответствующая дисперсия, быстро стремится к нулю при $N \to \infty$. Это означает, что в окрестности порядковой статистики $Y_{(k)}$ (под окрестностью понимается область, равная нескольким стандартным отклонениям), соответствующее математическое ожидание – это линейная функция, зависящая от номера отсчета, в то время как исходное распределение близко к ядерному. Поэтому имеем $\mathbb{E}\exp(Y_{(k)}) \approx \exp(\mathbb{E}Y_{(k)})$, исключение составляют наименьшие значения $Y_{(k)}$.

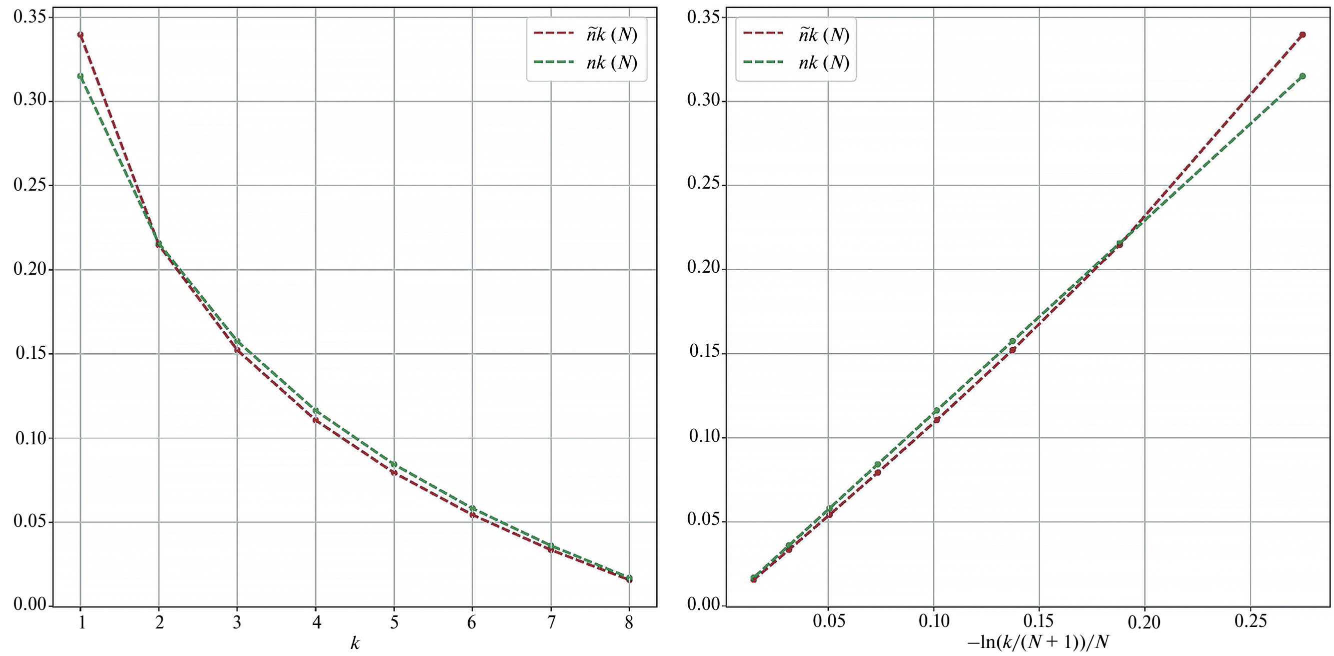
Теперь сравним дискретные распределения: точное, задаваемое формулой (5), и приближенное, определяемое по формуле (6).

При больших значения N и k выражение (5), которое является разностью гармонических рядов, можно приближать различными способами. Выберем следующую оценку

$$\sum_{i=1}^N \frac{1}{i} - \sum_{i=1}^{k-1} \frac{1}{i} \approx -\ln \frac{k}{N+1},$$

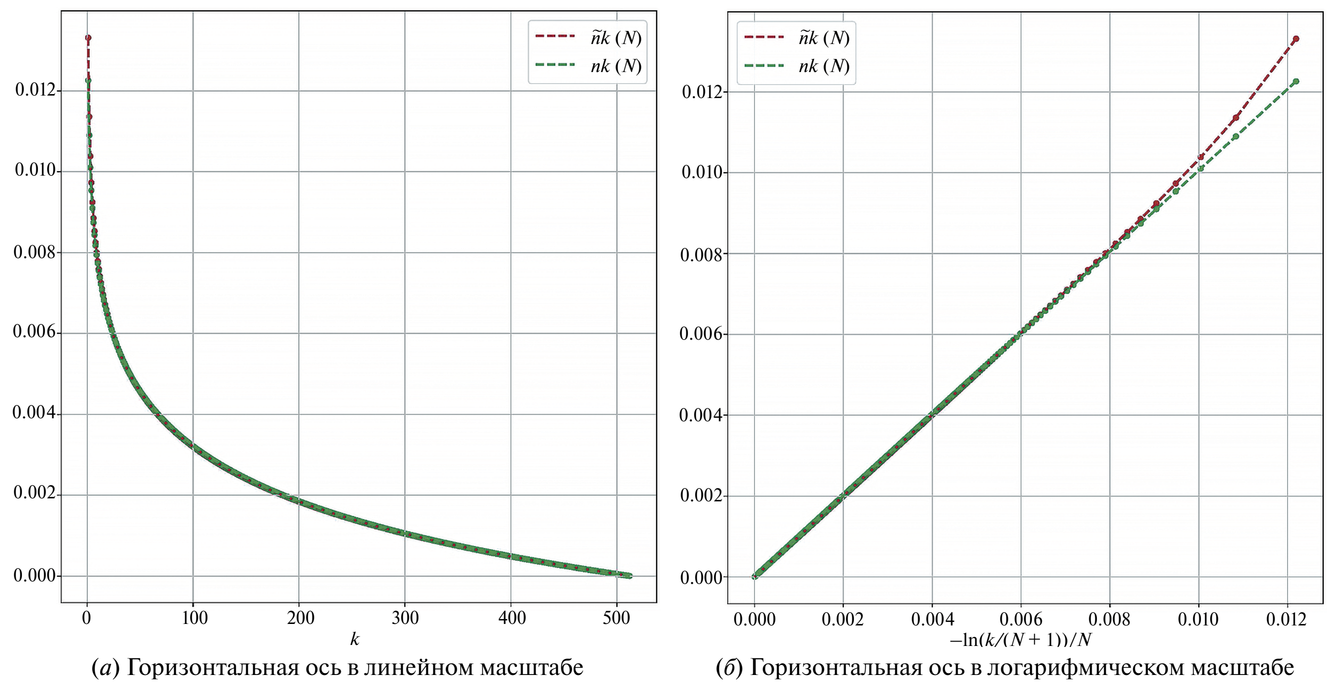
которая определена и имеет смысл при всех N и $k \in [1, \ldots, N]$, а также входит в формулу (6).

Проиллюстрируем на рис. 1, 2 графики зависимостей распределений (5) и (6) от номера отсчета.



(*а*) Горизонтальная ось в линейном масштабе (*б*) Горизонтальная ось в логарифмическом масштабе

Рис. 1. Дискретные распределения $\tilde{n}_k(N)$ и $n_k(N)$ для ряда размера $N = 8$.



(*а*) Горизонтальная ось в линейном масштабе (*б*) Горизонтальная ось в логарифмическом масштабе

Рис. 2. Дискретные распределения $\tilde{n}_k(N)$ и $n_k(N)$ для ряда размера $N = 512$.

Рис. 1 и 2 иллюстрируют, что точные значения математического ожидания (5) для малого количества точек (членов спектрального ряда) незначительно отличаются от распределения (6), предложенного в оригинальной работе, для остальных отсчетов они практически совпадают. С ростом N доля точек отклоняющихся от оценки быстро падает.

Таким образом, дополнительно обосновывается введение спектрального нормированного упорядоченного распределения, предложенного в Лемме 2, посредством которого определяется понятие спектральной сложности. Теперь спектральную сложность также можно задавать на основе распределения (5) по формуле

$$C_{SS}(p) = -\frac{1}{4\log_2 N}\left(\sum_{k=1}^{N} p_k \log_2 p_k\right) \times \times \left(\sum_{k=1}^{N} \left|p_k - \tilde{n}_k(N)\right|\right)^2. \quad (7)$$

В результате получен очень чувствительный, превышающий известные, метод обнаружения сигнала в белом шуме по наблюдениям в одном прямоугольном окне, основанный на вычислении спектральной сложности двух видов.

СПИСОК ЛИТЕРАТУРЫ

1. *Галяев А.А., Бабиков В.Г., Лысенко П.В., Берлин, Л.М.* Новая спектральная мера сложности и её возможности по обнаружению сигналов в шуме // Доклады Российской академии наук. Математика, информатика, процессы управления. 2024. Т. 518. № 4.
2. David Whitmer, Lecture 11: Probability, Order Statistics and Sampling, February 10, 2017, 15-750: Graduate Algorithms, Spring 2017, Carnegie Mellon University, School of Computer Science.

ADDITION TO THE ARTICLE “A NEW SPECTRAL MEASURE OF COMPLEXITY AND ITS CAPABILITIES FOR DETECTING SIGNALS IN NOISE”

Corresponding Member of the RAS A. A. Galyaev[a], V. G. Babikov[a], P. V. Lysenko[a], L. M. Berlin[a]

[a] *Institute of Control Sciences of RAS, Moscow, Russian Federation*

Note-addition to the article “A new spectral measure of complexity and its capabilities for detecting signals in noise”.

Keywords: information entropy, spectral complexity, additive white Gaussian noise