Записки научных
семинаров ПОМИ
Том 384, 2010 г.

Е. А. Тимофеев

СМЕЩЕНИЕ НЕПАРАМЕТРИЧЕСКОЙ ОЦЕНКИ ЭНТРОПИИ ДЛЯ МАРКОВСКОЙ МЕРЫ

1. Введение

В настоящей статье изучается применение непараметрической оценки размерностей из работы [8] для задачи оценивания энтропии (на символ) по экспериментальным данным из информационного источника с конечным алфавитом.

Задача оценивания энтропии будет рассматриваться в следующей постановке.

Пусть даны $n+1$ слов $\boldsymbol{\xi}_0, \ldots, \boldsymbol{\xi}_n$ из информационного источника с конечным алфавитом A. Требуется оценить энтропию источника. В большинстве прикладных задач число возможных значений слов очень велико, поэтому применяются непараметрические оценки.

Большинство непараметрических оценок энтропии h основано либо на алгоритме сжатия Лемпеля–Зива, либо на методе ближайших соседей (см. обзор в [8]), и эти оценки имеют вид

$$\widehat{h}_n = \frac{\log n}{R_n},$$

где R_n – некоторая величина, вычисляемая по экспериментальным данным. Для таких оценок доказывается сходимость почти всюду (наиболее общие результаты получены в [15, 13]), но такие оценки очень трудно применять в прикладных задачах, поскольку $\log_2 n \leqslant 30-40$, а скорость сходимости таких оценок $O(1/\log n)$.

Подчеркнем, что для теоретического исследования оценок такого вида более удобно изучать обратную величину $\widehat{h}_n^{-1}$, поэтому далее в работе будут рассматриваться только оценки величины $1/h$ – обратной к энтропии.

Ключевые слова: энтропия, непараметрическая оценка, марковская мера, смещение.

Работа проводилась при финансовой поддержке Федерального агентства по науке и инновациям, ФЦП "Научные и научно-педагогические кадры инновационной России на 2009–2013 годы".

В работе [8] вместо деления на $\log n$ предлагается рассмотреть разности между средними логарифмами расстояний до k-й и $k+1$-й ближайших точек (см. (4)). При таком подходе сходимость доказывается при трудно проверяемых условиях, но скорость сходимости получается степенной – $O(n^{-c})$, где c – некоторая константа. Для задачи оценивания размерностей эти условия выполнены для абсолютно непрерывных мер (относительно лебеговской). Для задачи оценивания энтропии в работе [13] была проведена вычислительная проверка для нескольких одномерных динамических систем с известной энтропией. Проверка показала, что эта оценка имеет точность 0.01 для $n = 10^4$. В экспериментах применялась наиболее простая и часто используемая в теории информации метрика

$$\rho_0(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{\min\{k : x_k \neq y_k\}}, \tag{1}$$

где $\boldsymbol{x}, \boldsymbol{y} \in \Omega = A^{\mathbb{N}}$ – пространство правосторонних последовательностей, A – заданный конечный алфавит, $\mathbb{N} = \{1, 2, \dots\}$.

Эти результаты позволяли надеяться, что скорость сходимости будет степенной. Кроме того, в [13] было показано, что дисперсия оценки равна $O(n^{-1})$ (при фиксированных k). Однако в [14] найдено смещение для симметричной меры Бернулли, которое является периодической функцией с периодом пропорциональным $\log n$. Хорошее согласование результатов эксперимента в [13] с теоретическими значениями энтропии объясняется тем, что для небольших значений энтропии ($h < 3$) это смещение настолько мало, что его нельзя заметить в вычислительном эксперименте. Так, например, для $A = \{0, 1\}$ амплитуда смещения меньше, чем 10^{-6}. (Как будет видно из дальнейшего, для симметричной меры Бернулли достигается наибольшее смещение оценки среди всех марковских мер.)

Итак, задача нахождения смещения оценки представляет большой интерес для теоретической проверки ее качества.

В настоящей работе для метрики (1) будет показано, что смещение равно $O(1)$ для тех марковских мер, у которых логарифмы вероятностей перехода соизмеримы. Для остальных марковских мер оценка из [8] является асимптотически несмещенной. Таким образом, смещение является разрывной функцией от параметров марковской меры и похожей на функцию Римана, равную 0 в иррациональных точках и $1/q$ в рациональной точке p/q.

Наличие смещения для метрики (1) и некоторых марковских мер приводит к следующему вопросу: можно ли убрать смещение заменив метрику? Задача выбора метрики является, по-видимому, довольно трудной. В последнем разделе работы эта задача изучается для симметричных мер Бернулли.

1. Обозначения и постановка задачи

1.1. Обозначения

Информационный источник будем рассматривать как инвариантную относительно сдвига меру μ в пространстве правосторонних последовательностей $\Omega = A^{\mathbb{N}}$. Итак, бесконечную случайную последовательность будем представлять как случайную величину, принимающую значения в Ω и распределенную по мере μ, и будем обозначать ее как $\boldsymbol{\xi} = (\xi_1, \xi_2, \dots)$.

Цилиндр порядка s, содержащий точку $\boldsymbol{x} \in \Omega$, будем обозначать через

$$C_s(\boldsymbol{x}) = \{\boldsymbol{y} \in \Omega : y_1 = x_1, \dots, y_s = x_s)\}.$$

Будем считать, что $C_0(\boldsymbol{x}) = \Omega$.

Пусть $\boldsymbol{\xi} = (\xi_1, \xi_2, \dots)$ – случайная точка в Ω распределенная по мере μ. Напомним, (см., напр., [7]) что *энтропия* (энтропия на символ) определяется как

$$h = -\lim_{n\to\infty} \frac{1}{n} \mathrm{E} \ln \mu(C_n(\boldsymbol{\xi})). \tag{2}$$

Подчеркнем, что натуральное основание логарифма выбрано для того, чтобы не ставить дополнительные множители в исследуемой оценке.

Пусть на пространстве Ω задана метрика ρ, тогда через $B(\omega, r)$ будем обозначать шар радиуса r с центром в точке ω, а через $r = \nu(t, \omega)$ – обобщенную обратную функцию к мере шара $t = \mu(B(\omega, r))$ (при заданном ω), т.е.

$$\nu(t, \omega) \ = \sup\{r \ : \ \mu(B(\omega, r)) < t\}. \tag{3}$$

1.2. Непараметрическая оценка энтропии

В [8] предложены и исследованы статистики для нахождения размерностей мер в метрических пространствах. Для пространства последовательностей Ω с метрикой ρ эти статистики определяются следующим образом.

Пусть заданы $n+1$ независимых случайных точек $\boldsymbol{\xi}_0, \ldots, \boldsymbol{\xi}_n$ в пространстве Ω одинаково распределенных по мере μ, тогда изучаемые статистики $\eta_n^{(k)}$, $r_n^{(k)}$ задаются по формулам:

$$\eta_n^{(k)}(\rho) = k\left(r_n^{(k)}(\rho) - r_n^{(k+1)}(\rho)\right), \tag{4}$$

$$r_n^{(k)}(\rho) = -\frac{1}{(n+1)}\sum_{j=0}^{n} \ln\left(\min_{i:i\neq j}{}^{(k)} \rho(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j)\right), \tag{5}$$

где $\min^{(k)}\{X_1, \ldots, X_N\} = X_k$, если $X_1 \leqslant X_2 \leqslant \cdots \leqslant X_N$.

Отметим, что для упрощения вычислений, случайные величины $\boldsymbol{\xi}_0$, $\ldots, \boldsymbol{\xi}_n$ считаются бесконечными последовательностями. Нетрудно показать (см. [13]), что достаточно использовать только $O(\ln n)$ первых символов этих последовательностей, если мера удовлетворяет следующему условию

$$\exists\alpha > 0 \ : \ \mu(C_n(\omega)) = O(e^{-n\alpha}), \quad \text{для } \mu\text{-почти всех } \omega \in \Omega. \tag{6}$$

Поэтому для нахождения расстояний $\rho_0(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j)$ достаточно знать только первые $m = O(\ln n)$ символов этих последовательностей.

Для того, чтобы по этим статистикам оценить энтропию, нужно взять метрику ρ билипшицево эквивалентной метрике

$$\rho_1(\boldsymbol{x}, \boldsymbol{y}) = \theta^{-\min\{k: x_k \neq y_k\}}, \tag{7}$$

где $\theta > 1$ – некоторый параметр.

Из теоремы 1 и предложения 8 из [8] получаем

Утверждение 1. *Пусть μ – инвариантная относительно сдвига эргодическая борелевская вероятностная мера, а ρ – метрика билипшицево эквивалентная метрике ρ_1, тогда для $k = O(n^c)$, где $c < 1$,*

$$\lim_{n\to\infty} \frac{\mathrm{E} r_n^{(k)}(\rho)}{\ln n} = \frac{\ln\theta}{h}.$$

Рассматриваемая оценка обратной энтропии строится по этим статистикам как

$$\frac{1}{\ln\theta}\eta_n^{(k)}(\rho).$$

Дисперсия статистики $\eta_n^{(k)}(\rho)$ равна $O(n^{-c})$ для достаточно широкого класса мер и метрик [8]. Поэтому основной задачей является нахождение смещения этой статистики.

Отметим, что можно найти и $\mathrm{E}r_n^{(k)}(\rho_1)$, но у этой величины основной вклад в смещение является неизвестной константой (не зависящей от n), которая отлична от 0 (см. последний раздел статьи).

2. Основные результаты

На пространстве $\Omega = A^{\mathbb{N}}$ определим меру μ как эргодическую марковскую меру с начальным распределением $p_i^0 > 0$ и матрицей переходных вероятностей p_{ij}, где

$$\sum_{j \in A} p_j^0 p_{jk} = p_k^0, \quad \forall k \in A. \tag{8}$$

Для простоты будем считать, что $p_{ij} > 0$.

Для нахождения $\mathrm{E}\eta_n^{(k)}(\rho)$ применим прием “пуассонизации” [19]. Положим

$$L_k(x, \rho) = \sum_{m=k}^{\infty} \frac{x^m}{m!} e^{-x} \mathrm{E}\eta_m^{(k)}(\rho). \tag{9}$$

Нахождение асимптотики величины $L_k(x, \rho_1)$ при $x \to \infty$ и является основной задачей настоящей работы.

Стандартная техника [19] позволяет показать (утверждение 4), что

$$L_k(n, \rho_1) = \mathrm{E}\eta_n^{(k)}(\rho_1) + O(n^{-c}),$$

где $c > 0$ – некоторая константа.

Для нахождения $L_k(x, \rho_1)$ будем применять преобразование Меллина

$$\widetilde{L}_k(z, \rho) = \int_0^{\infty} L_k(x, \rho) x^{z-1} dx. \tag{10}$$

Отметим, что аналогичные подход применяется в задачах анализа алгоритмов [6, упр. 5.2.2.53], [19, 11].

В этом разделе будем рассматривать метрику ρ_1 с параметром $\theta = e$, поэтому статистика $\eta_n^{(k)}(\rho_1)$ будет оценкой обратной энтропии, а величины $L_k(x, \rho_1)$, $\widetilde{L}_k(z, \rho_1)$ при $\theta = e$ будем обозначать через $L_k(x)$, $\widetilde{L}_k(z)$.

Утверждение 2.

$$\widetilde{L}_k(z) = \frac{\Gamma(k+z)}{\Gamma(k)} \sum_{j\in A} \sum_{m\in A} p_m^0 \frac{D_{mj}(z)}{\det D(z)}, \tag{11}$$

где через $D(z)$ обозначается матрица с элементами

$$d_{ij} = \delta_{ij} - \left(\frac{p_{ij}p_i^0}{p_j^0}\right)^{-z+1}, \tag{12}$$

а через $D_{ij}(z)$ обозначается алгебраическое дополнение элемента d_{ij}.

Следствие 1. *Для меры Бернулли с вероятностями p_j*

$$\widetilde{L}_k(z) = \frac{\Gamma(k+z)}{\Gamma(k)} \frac{1}{1 - \sum\limits_{j\in A} p_j^{-z+1}}. \tag{13}$$

Функция $\widetilde{L}_k(z)$, определенная в полосе $-k < \operatorname{Re} z < 0$, по формуле (11) продолжается на всю комплексную плоскость.

Полюсами функции $\widetilde{L}_k(z)$ являются

- простые полюса функции $\Gamma(k+z)$ в точках $-k, -k-1, \dots;$
- корни уравнения
$$\det D(z) = 0, \tag{14}$$
которые обозначим через z_m, $m \in \mathbb{Z}$ и будем считать, что $z_0 = 0$.

Утверждение 3. *Все корни уравнения* (14) *лежат в вертикальной полосе*

$$0 \leqslant \operatorname{Re} z < \sigma_0, \tag{15}$$

где σ_0 – некоторая константа.

Пусть

$$\det D(z) = 1 + \sum_{j=1}^{N} c_j e^{b_j z}.$$

Если все числа $b_1, \dots, b_N$ рационально соизмеримы, то вещественные части корней принимают только конечное число значений, а корни на мнимой прямой образуют арифметическую прогрессию.

Если числа $b_1, \ldots, b_N$ рационально не соизмеримы, то точка $z_0 = 0$ является единственным корнем на мнимой прямой, а вещественные части корней принимают бесконечное число значений.

Для нахождения асимптотики $L_k(x)$ применяется обратное преобразование Меллина

$$L_k(x) = \frac{1}{2\pi i} \int\limits_{-\sigma-i\infty}^{-\sigma+i\infty} \widetilde{L}_k(z, \rho_1) x^{-z}\, dz, \tag{16}$$

где $0 < \sigma < 1$.

Теорема 1. *Пусть γ такое, что на прямой $\operatorname{Re} z = \gamma$ нет корней уравнения (14) и для некоторой константы $c > 0$ на этой прямой*

$$|\det D(z)| > c > 0,$$

тогда

$$L_k(x) = -\sum_{z_m} \operatorname{Re} s \left(\frac{\Gamma(k+z) \sum\limits_{j\in A} \sum\limits_{n\in A} p_n^0 D_{nj}(z) x^{-z}}{\Gamma(k) \det D(z)}, z_m \right) + O(x^{-\gamma}). \tag{17}$$

где суммирование идет по всем корням уравнения (14), лежащим в полосе $0 \leqslant \operatorname{Re} z_m < \gamma$.

Отметим, что условие теоремы выполнено для $\gamma > \sigma_0$.

Следствие 2. *При $x \to \infty$*

$$L_k(x) = -\sum_{z_m} \operatorname{Re} s \left(\frac{\Gamma(k+z) \sum\limits_{j\in A} \sum\limits_{n\in A} p_n^0 D_{nj}(z) x^{-z}}{\Gamma(k) \det D(z)}, z_m \right) + o(1), \tag{18}$$

где суммирование идет по всем корням уравнения (14), лежащим на мнимой оси.

Следствие 3. *Пусть*

$$\det D(z) = 1 + \sum_{j=1}^{N} c_j e^{b_j z}$$

и величины $b_1, \ldots, b_N$ рационально несоизмеримы, тогда при $x \to \infty$

$$L_k(x) = \frac{1}{h} + o(1). \tag{19}$$

Утверждение 4. *Для любого $c < 1$*

$$L_k(n) = \mathrm{E}\eta_n^{(k)}(\rho_1) + O(n^{-c}). \tag{20}$$

3. Вспомогательные результаты

В этом разделе рассматривается произвольная инвариантная относительно сдвига эргодическая борелевская вероятностная мера μ. Для произвольной метрики ρ и метрики ρ_1 будут приведены явные формулы для нахождения $L_k(x, \rho)$ и необходимых вспомогательных величин.

3.1. Формулы для произвольной метрики ρ

При доказательстве теоремы 1 из [8] получена формула (4.8), которую приведем в следующем виде.

Лемма 1.

$$\mathrm{E}r_n^{(k)}(\rho) = \frac{n!}{(k-1)!(n-k)!}\int\limits_0^1 \chi(t,\rho)\, t^{k-1}(1-t)^{n-k}\, dt. \tag{21}$$

Здесь через $\chi(t,\rho)$ обозначается функция

$$\chi(t,\rho) = -\int\limits_\Omega \ln\nu(t,\omega)\, d\mu(\omega), \tag{22}$$

а функция $r = \nu(t,\omega)$ определена в (3).

Проинтегрировав по частям, перепишем выражение (21), а затем заменим обратную функцию $\nu(t,\omega)$. В результате получим

Лемма 2. *Пусть выполнено условие*

$$\exists\alpha > 0 \ : \ \mu(B(\omega, r)) = O(r^\alpha), \ \text{для } \mu\text{-почти всех } \omega\in\Omega,$$

тогда

$$\mathrm{E}r_n^{(k)}(\rho) = \int\limits_0^1\int\limits_\Omega \left[1 - \sum_{i=0}^{k-1}\binom{n}{i}\mu(B(\omega,r))^i\,(1-\mu(B(\omega,r)))^{n-i}\right]\frac{dr}{r}\, d\mu(\omega). \tag{24}$$

Отметим, что выполнение условия (6) для метрик билипшицево эквивалентных метрике (7) следует из условия (23).

Следствие 4. *Пусть выполнены условия леммы 2, тогда*

$$\mathrm{E}\eta_n^{(k)}(\rho) = k\binom{n}{k}\int\limits_0^1\int\limits_\Omega \mu(B(\omega,r))^k\,(1-\mu(B(\omega,r)))^{n-k}\,\frac{dr}{r}\,d\mu(\omega). \quad (25)$$

Следствие 5. *Пусть выполнены условия леммы 2, тогда*

$$L_k(x,\rho) = \int\limits_0^1\int\limits_\Omega \phi_k(x\mu(B(\omega,r))\,\frac{dr}{r}\,d\mu(\omega), \quad (26)$$

где через $\phi_k(x)$ обозначается функция

$$\phi_k(x) = \frac{x^k}{(k-1)!}\,e^{-x}. \quad (27)$$

Вычисляя интегралы через гамма-функцию, найдем преобразование Меллина.

Лемма 3. *В полосе* $-k < \operatorname{Re} z < 0$

$$\widetilde{L}_k(z,\rho) = \frac{\Gamma(k+z)}{\Gamma(k)}\int\limits_0^1\int\limits_\Omega \mu(B(\omega,r))^{-z}\,\frac{dr}{r}\,d\mu(\omega). \quad (28)$$

3.2. Формулы для метрики ρ_1 Для метрики ρ_1 полученные формулы можно упростить.

Лемма 4. *Пусть выполнено условие* (6), *тогда*

$$\mathrm{E}\eta_n^{(k)}(\rho_1) = k\ln\theta\binom{n}{k}\sum_{m=0}^{\infty}\int\limits_\Omega \mu(C_m(\omega))^k\,(1-\mu(C_m(\omega)))^{n-k}\,d\mu(\omega), \quad (29)$$

$$L_k(x,\rho_1) = \ln\theta\sum_{m=0}^{\infty}\int\limits_\Omega \phi_k(x\mu(C_m(\omega))\,d\mu(\omega), \quad (30)$$

в полосе $-k < \operatorname{Re} z < 0$

$$\widetilde{L}_k(z,\rho_1) = \ln\theta\frac{\Gamma(k+z)}{\Gamma(k)}\sum_{n=0}^{\infty}\int\limits_\Omega \mu(C_n(\omega))^{-z}\,d\mu(\omega). \quad (31)$$

Доказательство. Ясно, что для метрики (7)

$$C_n(\omega) = B(\omega, \theta^{-n}) = B(\omega, r), \quad \theta^{-n-1} < r \leqslant \theta^{-n}. \tag{32}$$

Поэтому для каждой меры μ и функции $f(x)$ имеем

$$\int_0^1 f(\mu(B(\omega,r)))\,\frac{dr}{r} = \sum_{m=0}^{\infty} f(\mu(C_m(\omega))) \int_{\theta^{-m-1}}^{\theta^{-m}} \frac{dr}{r}$$
$$= \ln\theta \sum_{m=0}^{\infty} f(\mu(C_m(\omega))).$$

Ясно, что условие (6) для метрики (7) эквивалентно условию (23), поэтому можно применить лемму 2 и следствия из нее. Подставляя равенство для соответствующей функции в (25), (26), (28), получим требуемые формулы. □

Отметим, что из приведенных формул видно, что для метрики (7) вспомогательная статистика

$$\frac{r_n^{(k)}(\rho_1)}{\ln\theta},$$

не зависит от θ и ее можно более просто записать через метрику (1)

$$\frac{r_n^{(k)}(\rho_1)}{\ln\theta} = \frac{1}{(n+1)} \sum_{j=0}^{n} \left(\min_{i:i\neq j}{}^{(k)} \rho_0(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j) \right)^{-1}. \tag{33}$$

Лишний параметр θ оставлен для того, чтобы изучить его влияние для метрик билипшицево эквивалентных метрике (7).

4. Доказательство основных результатов

В этом разделе для марковской меры будут доказаны утверждения 2–4 и теорема 1.

Из определения (8) видно, что мера цилиндров $C_n(\omega)$ находится следующим образом

$$\mu(C_n(jk\omega)) = \frac{p_{jk}p_j^0}{p_k^0}\mu(C_{n-1}(k\omega)), \quad \mu(C_1(j\omega)) = p_j^0 \quad \forall j,k \in A. \tag{34}$$

Здесь через $j\omega$ обозначается последовательность $(j, \omega_1, \omega_2, \dots)$.

4.1. Доказательство утверждения 2

Введем вспомогательные переменные

$$\widetilde{L}_j(z) = \sum_{n=0}^{\infty} \int\limits_{\Omega} \mu(C_n(j\omega))^{-z}\, d\mu(j\omega), \quad j \in A.$$

Применяя определение марковской меры (34), получим

$$\begin{aligned} \widetilde{L}_j(z) &= p_j^0 + \sum_{m\in A} \sum_{n=1}^{\infty} \int\limits_{\Omega} \mu(C_n(jm\omega))^{-z}\, d\mu(jm\omega) \\ &= p_j^0 + \sum_{m\in A} \sum_{n=1}^{\infty} \left(\frac{p_{jm}p_j^0}{p_m^0}\right)^{-z+1} \int\limits_{\Omega} \mu(C_{n-1}(m\omega))^{-z}\, d\mu(m\omega) \\ &= p_j^0 + \sum_{m\in A} \left(\frac{p_{jm}p_j^0}{p_m^0}\right)^{-z+1} \widetilde{L}_m(z). \end{aligned}$$

Обозначая через $\widetilde{\mathbf{l}}(z)$ вектор с компонентами $\widetilde{l}_j(z)$, а через $\mathbf{p}^0$ вектор с компонентами p_j^0, эти уравнения можно переписать в матричном виде

$$D(z)\widetilde{\mathbf{l}}(z) = \mathbf{p}^0,$$

где $D(z)$ – матрица, заданная в (12). Записав обратную матрицу через алгебраические дополнения

$$d_{ij}^{(-1)} = \frac{D_{ji}(z)}{\det D(z)},$$

получим

$$\widetilde{l}_j(z) = \sum_{m\in A} p_m^0 \frac{D_{mj}(z)}{\det D(z)}, \quad j \in A.$$

Через вспомогательные переменные $\widetilde{l}_j(z)$ функция $\widetilde{L}_k(z)$, заданная в (31), переписывается следующим образом

$$\widetilde{L}_k(z) = \frac{\Gamma(k+z)}{\Gamma(k)} \sum_{j\in A} \widetilde{L}_j(z).$$

Подставляя найденные значения $\widetilde{l}_j(z)$, получим (11).

Лемма 5.
$$L_k(x) = \sum_{j \in A} l_j(x), \tag{35}$$

где функции $l_j(x)$ удовлетворяют системе функциональных уравнений
$$l_j(x) = p_j^0 \phi_k(x) + \sum_{i \in A} \frac{p_{ji} p_j^0}{p_i^0}\, l_i\left(\frac{p_{ji} p_j^0}{p_i^0} x\right), \quad j \in A. \tag{36}$$

Доказательство. Введем вспомогательные переменные
$$l_j(x) = \sum_{n=0}^{\infty} \int_{\Omega} \phi_k(x\mu(C_n(j\omega)))\, d\mu(j\omega), \quad j \in A.$$

Повторяя выкладки доказательства утверждения 2 с подстановкой определения марковской меры (34), но к равенству (30), получим (36). □

Эта лемма показывает необходимость применения преобразования Меллина и будет нужна в дальнейшем для оценки функции $L_k(x)$.

4.3. Доказательство утверждения 3

Из определения (12) матрицы $D(z)$ видно, что
$$\det D(z) = 1 + \sum_{j=1}^{N} c_j e^{b_j z}$$

и $0 < b_1 < b_2 < \cdots < b_N$. Поэтому при достаточно большом значении $\operatorname{Re} z$
$$|\det D(z)| \approx |c_N| e^{b_N \operatorname{Re} z}.$$

Отсюда следует существование константы σ_0.

Покажем, что в области $\operatorname{Re} z < 0$ нет корней уравнения (14). Для этого достаточно показать, что в этой области
$$\| E - D(z) \| < 1.$$

В качестве нормы матрицы B выберем октаэдрическую норму, которая задается следующим образом [2, гл. XIV]
$$\| B \| = \max_j \sum_k |b_{kj}|.$$

Для матрицы $E - D(z)$ имеем

$$\| E - D(z) \| = \max_{j \in A} \sum_{k \in A} \left| \left(\frac{p_{kj} p_k^0}{p_j^0} \right)^{-z+1} \right| < \max_{j \in A} \sum_{k \in A} \frac{p_{kj} p_k^0}{p_j^0} = 1,$$

поскольку $\left| \left(\dfrac{p_{kj} p_k^0}{p_j^0} \right)^{-z} \right| < 1$ при $\operatorname{Re} z < 0$.

4.3. Доказательство теоремы 1

Перепишем формулу (11) в следующем виде

$$\widetilde{L}_k(z) = \frac{\Gamma(k+z) f(z)}{\det D(z)}, \tag{37}$$

где функция $f(z)$ – взвешенная сумма экспонент, поэтому она аналитична на всей комплексной плоскости и ограничена в полосе $-\sigma \leqslant \operatorname{Re} z_m < \gamma$.

Для нахождения $L_k(x)$ по формуле (16) рассмотрим в комплексной области прямоугольник

$$R_\beta = \{z : -\sigma \leqslant \operatorname{Re} z \leqslant \gamma;\ -\beta \leqslant \operatorname{Im} z \leqslant \beta\},$$

где параметр β определим позднее.

По теореме о вычетах

$$\begin{aligned} &\frac{1}{2\pi i} \oint\limits_{\partial R_\beta} \frac{\Gamma(k+z) f(z) x^{-z}}{\det D(z)}\, dz \\ &= \sum_{z_m \in R_\beta} \operatorname{Res} \left(\frac{\Gamma(k+z) f(z) x^{-z}}{\det D(z)},\ z_m \right). \end{aligned} \tag{38}$$

Здесь интеграл по контуру берется в направлении против часовой стрелки.

Оценим интегралы по границе прямоугольника.

Лемма 6. *Пусть β таково, что на прямой $\operatorname{Im} z = \beta$ нет корней уравнения* (14), *пусть*

$$I_1(x, \beta) = \int\limits_{-\sigma + i\beta}^{\gamma + i\beta} \frac{\Gamma(k+z) f(z) x^{-z}}{\det D(z)}\, dz,$$

тогда

$$|I_1(x,\beta)| \leqslant C_1 e^{-\pi|\beta|/2} x^{\sigma}. \tag{39}$$

Доказательство. Для оценки гамма-функции применим формулу [3, 8.328.1], которую запишем в следующем виде

$$|\Gamma(x+iy)| \leqslant C e^{-\pi|y|/2} |y|^{-x+1/2}.$$

В нашем случае

$$|\Gamma(k+z)| \leqslant C e^{-\pi|\beta|/2} |\beta|^{\sigma+1/2-k} \leqslant C e^{-\pi|\beta|/2}. \tag{40}$$

Подставляя эту оценку в $I_1(x,\beta)$, получим (39) с константой

$$C_1 = CC_0(\gamma+\sigma)/m,$$

где m – минимум $|\det D(z)|$ на отрезке интегрирования, C_0 – максимум $|f(z)|$ в полосе $-\sigma \leqslant \operatorname{Re} z_m < \gamma$. □

Лемма 7. *Пусть*

$$I_2(x,\beta) = \int\limits_{\gamma-i\beta}^{\gamma+i\beta} \frac{\Gamma(k+z) f(z) x^{-z}}{\det D(z)}\, dz,$$

тогда

$$|I_2(x,\beta)| \leqslant C_2 x^{-\gamma}. \tag{41}$$

Доказательство. Внесем модуль под интеграл

$$\begin{aligned} |I_2(x,\beta)| &\leqslant \frac{x^{-\gamma}}{c} \int\limits_{-\beta}^{\beta} |f(\gamma+iy)\Gamma(k+\gamma+iy)|\, dy \\ &< \frac{x^{-\gamma}}{c} C_0 \int\limits_{-\infty}^{\infty} |\Gamma(k+\gamma+iy)|\, dy, \end{aligned}$$

где C_0 – максимум $|f(z)|$ в полосе $-\sigma \leqslant \operatorname{Re} z_m < \gamma$. Последний интеграл существует, что очевидно из оценки (40). □

Лемма 8. *Пусть* $\beta > 0$,

$$I_3(x,\beta) = \int\limits_{-\sigma+i\beta}^{-\sigma+i\infty} \frac{\Gamma(k+z)f(z)x^{-z}}{\det D(z)}\,dz,$$

тогда

$$|I_3(x,\beta)| \leqslant C_3 e^{-\pi\beta/2} x^{\sigma}. \tag{42}$$

Доказательство. Внесем модуль под интеграл

$$|I_3(x,\beta)| \leqslant x^{\sigma} \int\limits_{\beta}^{\infty} \left| \frac{\Gamma(k-\sigma+iy)f(-\sigma+iy)}{\det D(-\sigma+iy)} \right| dy$$

$$\leqslant \frac{C_0 x^{\sigma}}{d} \int\limits_{\beta}^{\infty} |\Gamma(k-\sigma+iy)|\,dy.$$

где C_0 – максимум $|f(z)|$ в полосе $-\sigma \leqslant \operatorname{Re} z_m < \gamma$, а

$$d = \min_y |D(-\sigma+iy)|.$$

При доказательстве леммы 3 показано, что

$$d \geqslant 1 - \max_{j,k\in A} \left| \left(\frac{p_{kj}p_k^0}{p_j^0} \right)^{\sigma} \right| > 0.$$

Для оценки гамма-функции применим формулу (40). Подставляя эту оценку и интегрируя, получим (42). □

Вернемся к доказательству теоремы.

Подставляя в (16) формулу вычетов (38) и оценивая оставшиеся интегралы по леммам 6–8, получим

$$L_k(x) = -\sum_{z_m\in R_\beta} \operatorname{Res}\left(\frac{\Gamma(k+z)f(z)x^{-z}}{\det D(z)}, z_m \right) + O(e^{-\pi\beta/2}x^{\sigma}) + O(x^{-\gamma}).$$

Знак минус здесь ставится потому, что интеграл по контуру берется в направлении по часовой стрелке.

Возьмем

$$\beta = \frac{2}{\pi}(\gamma+\sigma)\ln x,$$

тогда оценки погрешностей интегралов совпадают, а вычеты корней вне прямоугольника R_β имеют порядок $O(x^{-\gamma})$. Подставив вместо $f(z)$ ее явный вид, получим (17).

Лемма 9.

$$\operatorname{Res}\left(\frac{\Gamma(k+z)\sum_{j\in A}\sum_{m\in A}p_m^0 D_{mj}(z)x^{-z}}{\Gamma(k)\det D(z)},\ 0\right)=-\frac{1}{h}. \tag{43}$$

Доказательство. По определению алгебраического дополнения имеем

$$\frac{d}{dz}\det D(z)=\sum_{k,j\in A}D_{kj}(z)d'_{kj}(z)=\sum_{k,j\in A}D_{kj}(z)\left(\frac{p_{kj}p_k^0}{p_j^0}\right)^{-z+1}\ln\frac{p_{kj}p_k^0}{p_j^0}.$$

Покажем, что

$$D_{kj}(0)=Cp_j^0,\quad k,j\in A.$$

Действительно, поскольку $\det D(0)=0$, то алгебраические дополнения $D_{kj}(0)$ удовлетворяют однородной системе уравнений

$$\sum_{k\in A}d_{km}(0)D_{kj}(0)=0,$$

а из определения (8) величин p_j^0 видно, что $D_{kj}(0)=Cp_j^0$ удовлетворяет этой системе уравнений.

Следовательно,

$$\begin{aligned}\frac{d}{dz}\det D(0)&=\sum_{k,j\in A}Cp_j^0\frac{p_{kj}p_k^0}{p_j^0}\ln\frac{p_{kj}p_k^0}{p_j^0}\\&=C\sum_{k,j\in A}p_{kj}p_k^0\left(\ln p_{kj}+\ln p_k^0-\ln p_j^0\right)\\&=C\sum_{k,j\in A}p_{kj}p_k^0\ln p_{kj}=-Ch.\end{aligned}$$

Итак, точка $z=0$ является простым полюсом. Подставляя $z=0$ в числитель (43) и используя тот факт, что значение суммы в числителе (43) равно C, получим (43). $\square$

4.4. Доказательство утверждения 4

Для доказательства равенства (20) применим теорему 10.3 из [19], в которой оно выводится из следующих двух условий:

- для любого $\beta > 0$ и $|z| \to \infty$

$$|L_k(z)| = O(|z|^\beta), \quad z \in S_\alpha; \tag{44}$$

- для некоторого $\gamma < 1$

$$|L_k(z)e^z| = O(e^{\gamma|z|}), \quad z \notin S_\alpha; \tag{45}$$

здесь

$$S_\alpha = \{z \in \mathbb{C} \ : \ |\operatorname{Im} z| \leqslant \alpha \operatorname{Re} z\}.$$

Для доказательства этих неравенств применим лемму 5 и введенные в ней функции $l_j(z)$.

Для доказательства (44) достаточно показать, что для некоторой константы C_1

$$|l_j(z)| \leqslant C_1 p_j^0 |z|^\beta, \quad z \in S_\alpha, \quad j \in A. \tag{46}$$

Ясно, что это неравенство выполнено при $z \in S_\alpha$, $|z| \leqslant R$. Покажем, что из справедливости (46) при некотором R следует его справедливость (с той же константой C_1) и для $|z| \leqslant BR$, где $B > 1$. Возьмем

$$B = \min_{i,j \in A} \frac{p_i^0}{p_{ji} p_j^0}, \tag{47}$$

и применим (36) и оценку (46) при $|z| \leqslant R$, тогда для $|z| \leqslant BR$ имеем

$$\begin{aligned}|l_j(z)| &\leqslant p_j^0 |\phi_k(z)| + \sum_{i \in A} \frac{p_{ji} p_j^0}{p_i^0} \left| l_i\left(\frac{p_{ji} p_j^0}{p_i^0} z\right)\right| \\ &\leqslant p_j^0 \left(|\phi_k(z)| + C_1 |z|^\beta \sum_{i \in A} p_{ji} \left(\frac{p_{ji} p_j^0}{p_i^0}\right)^\beta\right).\end{aligned}$$

Поскольку

$$\sum_{i \in A} p_{ji} \left(\frac{p_{ji} p_j^0}{p_i^0}\right)^\beta \leqslant \max_{i,j \in A} \left(\frac{p_{ji} p_j^0}{p_i^0}\right)^\beta < 1,$$

а в конусе S_α функция $\phi_k(z)$ ограничена, так как $|z| \leqslant (1+\alpha)\operatorname{Re} z$, то неравенство (46) выполняется с той же константой C_1 и для $|z| \leqslant BR$.

Для доказательства (45) достаточно показать, что для некоторой константы C_2 и $\dfrac{\alpha}{\sqrt{1+\alpha^2}} < \gamma < 1$

$$|l_j(z)e^z| \leqslant C_2 p_j^0 e^{\gamma|z|}, \quad z \notin S_\alpha, \quad j \in A. \tag{48}$$

Ясно, что это неравенство выполнено при $z \in S_\alpha$, $|z| \leqslant R$. Покажем, что из справедливости (48) при некотором R следует его справедливость (с той же константой C_2) и для $|z| \leqslant BR$, где B определено в (47). Для краткости положим

$$q_{ji} = \frac{p_{ji}p_j^0}{p_i^0}$$

и применим (36) и оценку (48) при $|z| \leqslant R$, тогда для $|z| \leqslant BR$ имеем

$$\begin{aligned}|l_j(z)e^z| &\leqslant p_j^0|\phi_k(z)e^z| + \sum_{i\in A}\frac{p_{ji}p_j^0}{p_i^0}\ |l_i(q_{ji}z)e^{q_{ji}z}|\left|e^{(1-q_{ji})z}\right| \\ &\leqslant p_j^0\left(\frac{|z|^k}{(k-1)!} + C_2\sum_{i\in A}p_{ji}\ e^{\gamma q_{ji}|z|}e^{(1-q_{ji})\operatorname{Re} z}\right).\end{aligned}$$

Поскольку вне конуса S_α выполняется неравенство $\operatorname{Re} z \leqslant \dfrac{\alpha}{\sqrt{1+\alpha^2}}|z|$, и $\dfrac{\alpha}{\sqrt{1+\alpha^2}} < \gamma$, то

$$\gamma q_{ji}|z| + (1-q_{ji})\operatorname{Re} z < \gamma|z|.$$

Итак, неравенство (48) выполняется и для $|z| \leqslant BR$ с той же константой C_2. □

5. Новая метрика

В этом разделе будет рассмотрена метрика

$$\rho_\theta(\boldsymbol{x},\boldsymbol{y}) = \sum_{k=1}^{\infty}(1-\delta_{x_k,y_k})\,\theta^{-k}, \tag{49}$$

где $\theta > 1$ – параметр метрики.

Для этой метрики, так же как и для метрики (7), справедлива лемма 1. Однако для метрики (49) в явном виде неизвестно математическое ожидание оценок даже при $n = 2$.

Чтобы показать основные трудности и возникающие преимущества, рассмотрим только случай $|A| = 2$ и μ – мера Бернулли с равновероятными символами. Подчеркнем, что даже в этом простом случае задача до конца не решена.

Обозначим

$$F(r) = \mu(B(\boldsymbol{x}, r)). \tag{50}$$

Ясно, что для меры Бернулли с равновероятными символами F не зависит от $\boldsymbol{x}$ и удовлетворяет уравнению

$$F(r) = \frac{1}{2}F(\theta r) + \frac{1}{2}F(\theta r - \theta + 1). \tag{51}$$

Функция $F(r)$ является функцией распределения случайной величины, которая обычно называется *бесконечной сверткой Бернулли* и изучается с 1930-ых годов. Начальное изучение функции $F(r)$ связано с вопросами гармонического анализа [5, 12.11], [1, 14.20].

Приведем простые и легко проверяемые свойства этой функции.

- $F(r)$ либо абсолютно непрерывна, либо сингулярна, в зависимости от θ.
- $F(r)$ сингулярна при $\theta > 2$.
- $F(r) = r$ $(0 \leqslant r \leqslant 1)$ при $\theta = 2$.

При $\theta = 2$ выполняется условие гладкости из [8]) и смещение оценки (4) для метрики (7) равно $O(n^{-1})$. Отметим, что это утверждение легко проверить (для рассматриваемой метрики) прямым вычислением по формуле (21).

А. Винтнер [20] нашел еще ряд значений θ, при которых функция $F(r)$ является степенной в положительной окрестности 0. Это значения $\theta = 2^{1/k}$ при $k \geqslant 1$. Например, при $\theta = \sqrt{2}$ функция $F(r)$ задается как

$$F(x) = \begin{cases} \dfrac{(1+\sqrt{2})^2}{2\sqrt{2}} x^2, & 0 \leqslant x \leqslant -1+\sqrt{2}; \\[2ex] \dfrac{1+\sqrt{2}}{\sqrt{2}} x - \dfrac{1}{2\sqrt{2}}, & -1+\sqrt{2} \leqslant x \leqslant 2-\sqrt{2}; \\[2ex] 1 - \dfrac{(1+\sqrt{2})^2}{2\sqrt{2}} (1-x)^2, & 2-\sqrt{2} \leqslant x \leqslant 1. \end{cases}$$

5.1. Абсолютная непрерывность функции $F(r)$

Основная часть исследований по функции $F(r)$ связана с изучением более слабого свойства гладкости – абсолютной непрерывности. Хотя эти результаты и не нужны в настоящей работе, приведем их небольшой обзор, который покажет сложность задачи изучения функции $F(r)$.

П. Эрдеш [10] показал, что $F(r)$ является сингулярной, если $1 < \theta < 2$ является PV-числом (Pisot–Vijayaraghavan). Более того, преобразование Фурье $\phi(t)$ функции $F(r)$ не стремится к 0 при $t \to \infty$.

Напомним, что PV-числом называется такое алгебраическое целое число $\theta > 1$, у которого все сопряженные числа по модулю меньше 1 (сопряженные числа – это остальные корни минимального многочлена). Примеры таких чисел – решения уравнений $x^2 - x - 1 = 0$ или $x^3 - x - 1 = 0$. Р. Салем [17] доказал обратный результат к методу П. Эрдеша для доказательства сингулярности. Если $\phi(t)$ не стремится к 0 при $t \to \infty$, то θ является PV-числом.

На следующий год П. Эрдеш [9] получил результат в противоположном направлении: существует $t_0 > 1$ такое, что для почти всех θ в интервале $1 < \theta \leqslant t_0$

$$|\phi(t)| = O(|t|^{-\gamma}), \tag{52}$$

где $\gamma > 0$ не зависит от θ. В [16] вычислено, что $\gamma > 6 \cdot 10^{-4}$ и $t_0 > 0.99933$.

А. Гарсиа [12] нашел новые примеры алгебраических чисел $1 < \theta < 2$, для которых функция $F(r)$ абсолютно непрерывна. Найденные им числа являются корнями в интервале $(1, 2)$ минимальных полиномов с коэффициентами -2, 0, 2. Примерами таких многочленов являются $x^3 - x - 2$, $x^3 - 2x - 2$, $x^4 - x^3 - 2$.

Б. Соломяк [18] доказал, что функция $F(r)$ абсолютно непрерывна и имеет плотность из L^2 для почти всех $\theta \in (1, 2)$.

5.2. Нахождение смещения

Подставляя функцию $F(r)$ в (21), получим

$$\mathrm{E}r_n^{(k)}(\rho_\theta) = -\frac{n!}{(k-1)!(n-k)!}\int_0^1 \ln F^{-1}(t)\, t^{k-1}(1-t)^{n-k-1}\, dt. \tag{53}$$

Наиболее важное свойство функции F, на котором будет основано вычисление, это ее подобие:

$$F(r) = \frac{1}{2}F(\theta r), \quad 0 \leqslant r \leqslant \frac{\theta-1}{\theta}.$$

Введем вспомогательную функцию $g(x)$, положив

$$F^{-1}(t) = t^{\log_2\theta} e^{-g(-\ln t)}. \tag{54}$$

Условие подобия функции F переписывается следующим образом:

$$g(x+\ln 2) = g(x), \quad x \geqslant -\ln F(\theta-1).$$

Таким образом, функция $g(x)$ является периодической, начиная с некоторого значения. Обозначим через $\widehat{g}(x)$ периодическую часть функции $g(x)$ и разложим ее в ряд Фурье

$$\widehat{g}(x) = \sum_{m=-\infty}^{\infty} g_m e^{2\pi i m x/\ln 2}. \tag{55}$$

Подставляя это разложение в (53) и учитывая, что основное значение интеграла сосредоточено в окрестности 0, получим

$$\mathrm{E}r_n^{(k)}(\rho_\theta) = -\frac{n!}{(k-1)!(n-k)!}\int_0^1 \left[\log_2\theta \ln t - \sum_{m=-\infty}^{\infty} g_m t^{-2\pi i m/\ln 2}\right] \times t^{k-1}(1-t)^{n-k-1}\, dt + O(\gamma^{-n}),$$

где $0 < \gamma < 1$ – некоторая константа.

Вычисляя интегралы [3, 4.253.1], получим

$$\mathrm{E}r_n^{(k)}(\rho_\theta) = \frac{\ln\theta}{\ln 2}(H_n - H_{k-1}) + \sum_{m=-\infty}^{\infty} g_m \frac{\Gamma\left(k+\frac{2\pi m i}{\ln 2}\right)\Gamma(n+1)}{\Gamma(k)\Gamma\left(n+1+\frac{2\pi m i}{\ln 2}\right)} + O(\gamma^{-n}),$$

где через H_s обозначаются гармонические числа

$$H_s = \sum_{i=1}^{s} \frac{1}{i}.$$

Применяя формулу Стирлинга, получим асимптотическое выражение

$$\mathrm{E}r_n^{(k)}(\rho_\theta) = \frac{\ln\theta}{\ln 2}(H_n - H_{k-1}) + \sum_{m=-\infty}^{\infty} g_m \frac{\Gamma\left(k + \frac{2\pi m i}{\ln 2}\right)}{\Gamma(k)} n^{-2\pi m i/\ln 2} + O(n^{-1}).$$

Подставляя в (1), получим

$$\mathrm{E}\eta_n^{(k)}(\rho_\theta) = \frac{\ln\theta}{\ln 2} + \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} g_m \frac{2\pi m i\Gamma\left(k + \frac{2\pi m i}{\ln 2}\right)}{\ln 2\Gamma(k)} n^{-2\pi m i/\ln 2} + O(n^{-1}).$$

Полученная асимптотическая формула является обобщением аналогичной формулы, полученной в [14] для $\theta = \infty$. Однако, здесь остались неизвестными величины g_m.

Для наглядности покажем график изменения амплитуды функции $g(x)$ – величины

$$\sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} |g_m|^2.$$

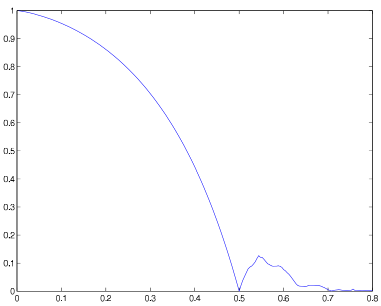

Рис. 1. Нормированная амплитуда функции $g(x)$ в зависимости от параметра $\theta^{-1} \in [0, 0.8]$ (при $\theta = \infty$ амплитуда равна 1).

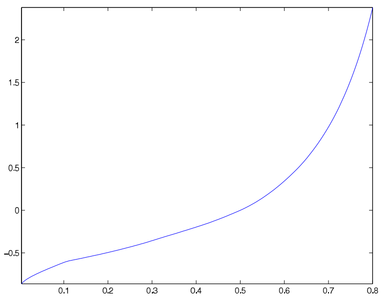

Рис. 2. Значение g_0 в зависимости от параметра $\theta^{-1} \in [0, 0.8]$.

Функция $F(r)$ строилась как аттрактор итерационной системы функций по отображениям

$$S_1(x) = \frac{x}{\theta}, \quad S_2(x) = \frac{x}{\theta} + \frac{\theta - 1}{\theta},$$

а по ней находилась функция $g(x)$ по формуле (54). Строилось 2^{20} точек и 2000 интервалов для задания $F(r)$.

Как видно из графика, нулевые амплитуды достигаются только при $\theta = 2$ и $\theta = \sqrt{2}$ (остальные значения не попали на график). Отметим, что небольшой пик при $\theta = 0.755$ соответствует обратному значению наименьшего PV-числа – корня уравнения $x^3 - x - 1 = 0$.

Значение g_0, показанное на втором графике, имеет достаточно большое значение. Теоретически определить значение θ, при котором это значение нулевое, не удается. Этот график, в частности, показывает, что оценка обратной энтропии

$$\frac{r_n^{(k)}(\rho_\theta)}{\ln n \ln \theta}$$

имеет смещение (и скорость сходимости) $O(1/\ln n)$. Наличие такого достаточно большого смещения и является основанием для изучения оценки $\eta_n^{(k)}(\rho_\theta)$.

Литература

1. Н. К. Бари, *Тригонометрические ряды*, ГИФМЛ, М., 1961.
2. Ф. Р. Гантмахер, *Теория матриц*, Наука, М., 1988.
3. И. С. Градштейн, И. М. Рыжик, *Таблицы интегралов, сумм, рядов и произведений*, Наука, М., 1971.

4. Р. Л. Добрушин, *Упрощенный метод экспериментальной оценки энтропии стационарной последовательности*, — Теория вероятн. и ее примен. **III** (1958), No. 4, 462–464.
5. А. Зигмунд, *Тригонометрические ряды*, Мир, М., 1965.
6. Д. Кнут, *Искусство программирования для ЭВМ*. Том 3. Мир, М., 1978.
7. Н. Мартин, Дж. Ингленд, *Математическая теория энтропии*, Мир, М., 1988.
8. Е. А. Тимофеев, *Статистически оцениваемые инварианты мер*. — Алгебра и анализ **17** (2005) No. 3, 204–236.
9. P. Erdös, *On the smoothness properties of a family of Bernoulli convolutions*. — Amer. J. Math. **62** (1940) No. 1, 180–186.
10. P. Erdös, *On a family of symmetric Bernoulli convolutions*. — Amer. J. Math. **61** (1939) No. 4, 974–976.
11. P. Flajolet, R. Sedgewick, *Analytic Combinatorics*. Cambridge Univ. Press, 2008.
12. A. M. Garsia, *Arithmetic properties of Bernoulli convolutions*. — Trans. Amer. Math. Soc, **102** (1962) No. 3, 409–432.
13. A. Kaltchenko, N. Timofeeva, *Entropy estimators with almost sure convergence and an $O(n^{-1})$ variance*. — Adv. Math. Comm. **2** (2008) No 1, 1–13.
14. A. Kaltchenko, N. Timofeeva, *Rate of convergence of the nearest neighbor entropy estimator*, — AEU - Intern. J. Electr. Comm. **64** (2010) No. 1, 75–79.
15. I. Kontoyiannis, Yu. M. Suhov, *Prefixes and the entropy rate for long-range sources*. — In: Probability Statistics and Optimization (F. P. Kelly, ed.). Wiley, New York, 1994.
16. Y. Peres, W. Schlag, B. Solomyak, *Sixty years of Bernoulli convolutions*. — Fractals and Stochastics II, (C. Bandt, S. Graf, M. Zähle, eds.), Progress in Probability **46**, Birkhauser (2000), pp. 39–65.
17. R. Salem, *Sets of uniqueness and sets of multiplicity*. — Trans. Amer. Math. Soc. **54** (1943) No. 2, 218–228.
18. B. Solomyak, *On the random series $\sum \pm\lambda^n$ (an Erdös problem)*. — Ann. Math. 2nd Ser. **142** (1995) No. 3, 611–625.
19. W. Szpankowski, *Average Case Analysis of Algorithms on Sequences*. John Wiley & Sons, New York, 2001.
20. A. Wintner, *On convergent Poisson convolutions*. — Amer. J. Math. **57** (1935) No. 4, 827–838.

Timofeev E. A. Bias of the nonpametric entropy estimator for Markov measures.

We consider the problem of the nonparametric entropy (entropy rate) estimation. We investigate the technique of the nonparametric entropy estimation based on so-called “nearest neighbor distances” and obtain a closed-form expression of the bias for Markov measures.

This bias is a discontinuous function of transition probabilities.

Ярославский государственный
университет им. П. Г. Демидова,
ул. Советская 14, 150040, Ярославль Россия
E-mail: TimofeevEA@gmail.com

Поступило 18 октября 2010 г.