Моделирование и анализ информационных систем. Т. 24, № 5 (2017), с. 521–536
Modeling and Analysis of Information Systems. Vol. 24, No 5 (2017), pp. 521–536

DOI: 10.18255/1818-1015-2017-5-521-536

УДК 519.17

Существование несмещенной оценки энтропии для специальной меры Бернулли

Тимофеев Е.А.

получена 10 июля 2017

Аннотация.

Пусть $\Omega = \mathcal{A}^{\mathbb{N}}$ – пространство правосторонних бесконечных последовательностей символов из алфавита $\mathcal{A} = \{0, 1\}$, $\mathbb{N} = \{1, 2, \dots\}$,

$$\rho(\boldsymbol{x}, \boldsymbol{y}) = \sum_{k=1}^{\infty} |x_k - y_k| 2^{-k}$$

– метрика на Ω и μ – вероятностная мера на Ω. Пусть $\boldsymbol{\xi_0}, \boldsymbol{\xi_1}, \dots, \boldsymbol{\xi_n}$ – независимые случайные точки на Ω, распределенные по мере μ. Будем изучать оценку $\eta_n^{(k)}(\gamma)$ величины обратной к энтропии $1/h$, которая определяется следующим образом:

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

где

$$r_n^{(k)}(\gamma) = \frac{1}{n+1} \sum_{j=0}^{n} \gamma \left(\min_{i:i \neq j}{}^{(k)} \rho(\boldsymbol{\xi_i}, \boldsymbol{\xi_j}) \right),$$

$\min^{(k)}\{X_1, \dots, X_N\} = X_k$, if $X_1 \le X_2 \le \dots \le X_N$. Число k и функция $\gamma(t)$ – вспомогательные параметры. Основной результат работы:

Теорема. Пусть μ – мера Бернулли с вероятностями $p_0, p_1 > 0$, $p_0 + p_1 = 1$, $p_0 = p_1^2$, тогда существует функция $\gamma(t)$ такая, что

$$\mathsf{E}\eta_n^{(k)}(\gamma) = \frac{1}{h}.$$

Ключевые слова: мера, метрика, энтропия, оценка, несмещенность, самоподобие, мера Бернулли

Для цитирования: Тимофеев Е.А., "Существование несмещенной оценки энтропии для специальной меры Бернулли", *Моделирование и анализ информационных систем*, **24**:5 (2017), 521–536.

Об авторе:
Тимофеев Евгений Александрович, orcid.org/0000-0002-0980-2507, доктор физ.-мат. наук, профессор,
Ярославский государственный университет им. П.Г. Демидова,
ул. Советская, 14, г. Ярославль, 150003 Россия, e-mail: timofeevEA@gmail.com

В работе [5] предложена непараметрическая оценка энтропии, которая зависит от семейства метрик. В [8] эта оценка модифицирована так, что она зависит от конкретной метрики и некоторой функции $\gamma(t)$. Эта модификация более удобна для анализа, поэтому именно она и будет рассматриваться в настоящей работе.

Основным показателем качества оценки является ее точность, которая складывается из несмещенности и дисперсии оценки. Нахождение порядка убывания дисперсии оценки является более простой задачей. Почти оптимальный порядок убывания дисперсии доказан в [6]. Нахождение математического ожидания оценки в общем случае является очень трудной задачей, найти решение которой, по-видимому, невозможно. Поэтому находить ее будем для мер Бернулли. Для мер Бернулли асимтотическая несмещенность доказана для почти всех значений параметров и конкретной функции $\gamma(t)$ [9]. Для бинарных мер Бернулли с параметрами p_0, p_1 ($p_0 + p_1 = 1$) исключительные случаи, в которых асимптотическая сходимость не доказана, описываются как

$$\frac{\log p_0}{\log p_1} - \text{рационально.}$$

Отсутствие сходимости в исключительных случаях является принципиальной трудностью, поскольку при простейшем выборе метрики и функции $\gamma(t)$ (оценка Грассбергера [3]) в случае $p_0 = p_1 = 1/2$ показано [7], что оценка является смещенной, хотя смещение очень мало (порядка 10^{-6}). В [5] показано, что при метрике, заданной в (1), и функции $\gamma(t) = -\log_2 t$ в случае $p_0 = p_1 = 1/2$ оценка будет несмещенной.

В настоящей работе будет показано, что и для следующего исключительного случая

$$\frac{\log p_0}{\log p_1} = 2$$

для метрики, заданной в (1), можно найти такую функцию $\gamma(t)$, для которой оценка энтропии будет несмещенной.

Таким образом, предложенный в [5, 8] метод адаптивного оценивания энтропии (предварительное нахождение вспомогательной функции) работает и в исключительных случаях.

1. Непараметрическая оценка энтропии

Обозначим через $\Omega = \mathcal{A}^{\mathbb{N}}$ пространство правосторонних бесконечных последовательностей символов из конечного алфавита $\mathcal{A} = \{0, 1\}$.

Пусть даны $\boldsymbol{\xi_0}, \boldsymbol{\xi_1}, \ldots, \boldsymbol{\xi_n}$ – независимые случайные точки в Ω, одинаково распределенные по вероятностной мере μ.

Зададим метрику ρ на Ω, положив

$$\rho(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{\infty} |x_i - y_i| 2^{-i}. \qquad (1)$$

Пусть заданы:

1. k – вспомогательный параметр, который k будем считать небольшим и фиксированным. Этот параметр служит для контроля применимости (оценки, полученные для различных значений k, являются оценками одной и той же величины);

2. $\gamma(t)$ – функция, определенная на полуинтервале $(0,1]$.

Оценка $\eta_n^{(k)}(\gamma)$ величины обратной к энтропии $1/h$ определяется следующим образом:

$$\eta_n^{(k)}(\gamma) = k\left(r_n^{(k)}(\gamma) - r_n^{(k+1)}(\gamma)\right), \tag{2}$$

где

$$r_n^{(k)}(\gamma) = \frac{1}{n+1}\sum_{j=0}^{n} \gamma\left(\min_{i:i\neq j}{}^{(k)} \rho(\boldsymbol{\xi_i}, \boldsymbol{\xi_j})\right), \tag{3}$$

и $\min^{(k)}\{X_1,\dots,X_N\} = X_k$, если $X_1 \le X_2 \le \dots \le X_N$.

Подчеркнем, что логарифм в определении энтропии выбирается натуральным (в этом случае в формуле (3) нет дополнительных множителей).

Ясно, что оценка линейно зависит от функции $\gamma(t)$.

Покажем, что ранее рассмотренные оценки укладываются в приведенную схему.

Возьмем функцию

$$\gamma_0(2^{-k_1} + 2^{-k_2} + 2^{-k_3} + \dots) = k_1, \quad 0 < k_1 < k_2 < k_3 < \dots \tag{4}$$

В этом случае статистика $r_n^{(1)}(\gamma_0)$ совпадает со статистикой в оценке Грассбергера [3]).

Отметим, что

$$\rho_0(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{\gamma_0(\rho(\boldsymbol{x},\boldsymbol{y}))}$$

– метрика, основанная на первом несовпадении.

В [6] показано,

$$\mathsf{E}r_n^{(k)}(\gamma_0) = \frac{H_n - H_{k-1}}{h} + \mathcal{O}(1),$$

где через

$$H_n = \sum_{m=1}^{n} \frac{1}{m} \tag{5}$$

обозначаются гармонические числа.

Для дисперсии статистики $r_n^{(k)}(\gamma_0)$ в [6] показано, что $\mathsf{Var}\, r_n^{(k)}(\gamma_0) = \mathcal{O}(n^{-1}\ln n)$.

В работе [5] аналогичные оценки доказаны для произвольной монотонной функции $\gamma(t)$, имеющей порядок $\log_2 t$.

Смещение оценки $\eta_n^{(k)}(\gamma_0)$ изучено в [9], где доказана асимптотическая несмещенность, если логарифмы некоторых вероятностей рационально не соизмеримы. Для меры Бернулли с вероятностями $p_0, p_1, \dots, p_m$ асимптотическая несмещенность выполняется, если $\log p_0, \log p_1, \dots, \log p_m$ рационально не соизмеримы.

2. Построение несмещенной оценки в случае $p_0 = p_1^2$

Теорема 1. *Пусть μ – мера Бернулли с параметрами $p_0 = q^2$, $p_1 = q$, тогда существует такая функция $\gamma(t)$, что*

$$\mathsf{E}\eta_n^{(k)}(\gamma) = \frac{1}{h}.$$

Доказательство. В [10] показано, что

$$\mathsf{E}r_n^{(k)}(\gamma) = n\binom{n-1}{k-1}\int_0^1 \chi(t)t^{k-1}(1-t)^{n-k}\,dt, \tag{6}$$

где

$$\chi(t) = \int_\Omega \gamma(\nu(t,\omega))\,d\mu(\omega), \tag{7}$$

а $r = \nu(t,\omega)$ – обратная функция к мере шара $B(r,\omega)$ радиуса r с центром в точке ω

Для доказательства теоремы достаточно показать, что существует такая функция $\gamma(t)$, что

$$\chi(t) = -\frac{\ln t}{h}. \tag{8}$$

Действительно, подставляя (8) в (7), (6) и применяя формулу [2, 4.253.1], получим

$$\mathsf{E}r_n^{(k)}(\gamma) = \frac{H_n - H_{k-1}}{h},$$

где H_n – гармонические числа (5).

Следовательно,

$$\mathsf{E}\eta_n^{(k)}(\gamma) = \frac{1}{h}.$$

Для нахождения решения уравнения (8) введем вспомогательные функции

$$\chi_{n,j}(t) = \int_\Omega \gamma(j2^{-n} + 2^{-n}\nu(t,\omega))\,d\mu(\omega), \quad 0 \le j < 2^n, \quad n = 0,1,2,\dots \tag{9}$$

и будем считать, что $\chi_{0,0}(t) = \chi(t)$.

Для метрики (1) и меры Бернулли с параметрами $p_0 = q^2$, $p_1 = q$ функция $r = \nu(t,\omega)$ удовлетворяет рекуррентному уравнению

$$\begin{gathered}
\nu(t,0\omega) = \begin{cases} \frac{1}{2}\nu\left(\frac{t}{q^2},\omega\right), & 0 \le t \le q^2; \\ \frac{1}{2} + \frac{1}{2}\nu\left(\frac{t-q^2}{q},\omega\right) & q^2 \le t \le 1; \end{cases} \\
\nu(t,1\omega) = \begin{cases} \frac{1}{2}\nu\left(\frac{t}{q},\omega\right), & 0 \le t \le q; \\ \frac{1}{2} + \frac{1}{2}\nu\left(\frac{t-q}{q^2},\omega\right) & q \le t \le 1. \end{cases}
\end{gathered} \tag{10}$$

Подставляя (10) в (9), получим

$$\chi_{n,j}(t) = \begin{cases} q^2\chi_{n+1,2j}\left(\dfrac{t}{q^2}\right) + q\chi_{n+1,2j}\left(\dfrac{t}{q}\right), & 0 \le t \le q^2; \\ q^2\chi_{n+1,2j+1}\left(\dfrac{t-q^2}{q}\right) + q\chi_{n+1,2j}\left(\dfrac{t}{q}\right), & q^2 \le t \le q; \\ q^2\chi_{n+1,2j+1}\left(\dfrac{t-q^2}{q}\right) + q\chi_{n+1,2j+1}\left(\dfrac{t-q}{q^2}\right), & q \le t \le 1. \end{cases} \tag{11}$$

Найдем функцию $\chi_{n+1,2j+1}(t)$ через $\chi_{n,j}(t)$.

Лемма 1. *Пусть для некоторого a, $0 < a < 1$, функция $\chi_{n+1,2j+1}(t)$ ограничена при $t > a$, тогда*

$$\chi_{n+1,2j+1}(t) = \sum_{k=0}^{\infty}(-1)^k q^{k-1}\chi_{n,j}\left(1 - q^{k+2}(1-t)\right), \quad 0 \le t \le 1. \tag{12}$$

Доказательство. Положим

$$\overline{\chi}_{n,j}(t) = \chi_{n,j}(1-t),$$

тогда третье уравнение в (11) можно переписать как

$$\overline{\chi}_{n,j}(t) = q^2\overline{\chi}_{n+1,2j+1}\left(\frac{t}{q}\right) + q\overline{\chi}_{n+1,2j+1}\left(\frac{t}{q^2}\right), \quad 0 \le t \le q^2.$$

Сделав замену переменной, перепишем это уравнение в следующем виде:

$$\overline{\chi}_{n+1,2j+1}(t) = q^{-1}\overline{\chi}_{n,j}(q^2 t) - q\overline{\chi}_{n+1,2j+1}(qt), \quad 0 \le t \le 1. \tag{13}$$

Покажем, что $\forall m \ge 1$ справедливо равенство

$$\overline{\chi}_{n+1,2j+1}(t) = (-1)^m q^m \overline{\chi}_{n+1,2j+1}(q^m t) + \sum_{k=0}^{m-1}(-1)^k q^{k-1}\overline{\chi}_{n,j}(q^{k+2}t), \quad 0 \le t \le 1. \tag{14}$$

Проведем индукцию по m. При $m = 1$ равенство (14) совпадает с (13).

Предположим, что (14) справедливо при заданном m и покажем, что оно справедливо и для $m+1$.

Подставляя (13) в первое слагаемое (14), получим

$$\begin{aligned} \overline{\chi}_{n+1,2j+1}(t) = (-1)^m q^{m-1}\overline{\chi}_{n,j}(q^{m+2}t) - (-1)^m q^{m+1}\overline{\chi}_{n+1,2j+1}(q^{m+1}t) \\ + \sum_{k=0}^{m-1}(-1)^k q^{k-1}\overline{\chi}_{n,j}(q^{k+2}t) = \\ = (-1)^{m+1}q^{m+1}\overline{\chi}_{n+1,2j+1}(q^{m+1}t) + \sum_{k=0}^{m}(-1)^k q^{k-1}\overline{\chi}_{n,j}(q^{k+2}t). \end{aligned}$$

По условию леммы функция $\overline{\chi}_{n+1,2j+1}(t)$ ограничена при $t < 1-a$, тогда при $m \to \infty$ получаем (12). $\square$

При известной функции $\chi_{n+1,2j+1}(t)$ нахождение функции $\chi_{n+1,2j}(t)$ легко проводится по формулам (9), которые перепишем в следующем виде:

$$\begin{aligned} &\chi_{n+1,2j}(t) = q^{-1}\chi_{n,j}(qt) - q\chi_{n+1,2j+1}(t-q), \quad q \le t \le 1; \\ &\chi_{n+1,2j}(t) = q^{-1}\chi_{n,j}(qt) - q\chi_{n+1,2j}(\tfrac{t}{q}), \qquad\quad 0 \le t \le q. \end{aligned} \tag{15}$$

Покажем, что функция $\chi_{n+1,2j+1}(t)$ является аналитической, функция $\chi_{n+1,2j}(t)$ – кусочно-аналитической.

Обозначим через $\chi^{(m)}_{n,j}(t)$ аналитическую функцию, которая совпадает с $\chi_{n,j}(t)$ на интервале $(q^{m+1} < t < q^m)$, $m = 0, 1, 2, \ldots$. Отметим, что $\chi^{(m)}_{n+1,2j+1}(t) = \chi_{n+1,2j+1}(t)$.

В этих обозначениях формулы (12), (15) переписываются в следующем виде:

$$\begin{aligned} &\chi_{n+1,2j+1}(t) = \textstyle\sum_{k=0}^{\infty}(-1)^k q^{k-1}\chi^{(0)}_{n,j}\left(1 - q^{k+2}(1-t)\right), && 0 \le t \le 1; \\ &\chi^{(0)}_{n+1,2j}(t) = q^{-1}\chi^{(1)}_{n,j}(qt) - q\chi_{n+1,2j+1}(t-q), && q \le t \le 1; \\ &\chi^{(m)}_{n+1,2j}(t) = q^{-1}\chi^{(m+1)}_{n,j}(qt) - q\chi^{(m-1)}_{n+1,2j}(\tfrac{t}{q}), && q^{m+1} \le t \le q^m, \quad m \ge 1. \end{aligned} \tag{16}$$

Лемма 2. *Пусть функция* $\chi_{0,0}(t)$ *аналитична на интервале* $(0, 2)$, *тогда функция* $\chi_{n+1,2j+1}(t)$ *аналитична на интервале* $(0, 2q^{-1})$, *функция* $\chi^{(m)}_{n+1,2j}(t)$ *аналитична на интервале* $(0, 2q^{m-1})$, $j = 0, 1, \ldots 2^n - 1$, $n = 0, 1, \ldots$, $m = 0, 1, \ldots$

Доказательство. Проведем индукцию по n. По условию и предположению индукции функция $\chi^{(0)}_{n,j}(t)$ аналитична на $(0, 2)$.

Отсюда и из первого уравнения (16) следует, что $\chi_{n+1,2j+1}(t)$ аналитична на интервале $|1 - t| < q^{-2} = 2 + q$ и, следовательно, на интервале $(0, 2q^{-1})$.

По условию и предположению индукции функция $\chi^{(1)}_{n,j}(t)$ аналитична на $(0, 2)$. Следовательно, функция $\chi^{(1)}_{n,j}(qt)$ аналитична на $(0, 2q^{-1})$. Функция $\chi_{n+1,2j+1}(t - q)$ аналитична на интервале $|1 + q - t| < 2 + q$. Отсюда и из второго уравнения (16) следует, что $\chi^{(0)}_{n+1,2j}(t)$ аналитична на интервале $(0, 2q^{-1})$.

По условию и предположению индукции функция $\chi^{(m+1)}_{n,j}(t)$ аналитична на $(0, 2q^{m-1})$. По доказанному функция $\chi^{(m-1)}_{n+1,2j}(\frac{t}{q})$ аналитична на интервале $(0, 2q^{m-1})$. Отсюда и из третьего уравнения (16) следует, что $\chi^{(m)}_{n+1,2j}(t)$ аналитична на интервале $(0, 2q^{m-1})$. □

Непрерывную функцию $f(t)$, которая на каждом интервале $(q^{m+1} < t < q^m)$, $m = 0, 1, 2, \ldots$, является многочленом степени k, будем называть **квазимногочленом степени** k.

Следствие 1. *Пусть функция* $\chi_{n,j}(t)$ *– квазимногочлен степени* k, *тогда* $\chi_{n+1,2j+1}(t)$ *– многочлен степени* k, $\chi_{n+1,2j}(t)$ *– квазимногочлен степени* k.

Лемма 3. *Существует функция* $\gamma(t)$, *для которой* $\chi(t) = 1 - t$, *т.е.*

$$1 - t = \int_{\Omega} \gamma(\nu(t, \omega))\, d\mu(\omega).$$

Более того, существуют такие константы a_n, $\sum_{k=1}^{\infty} a_k = 1$, *что*

$$\gamma(2^{-n} + 2^{-n}t) = a_n\gamma(t) + \sum_{k=1}^{n-1} a_k, \quad n = 1, 2, \ldots \tag{17}$$

Эти константы задаются через производящую функцию $a(x) = \sum_{k=1}^{\infty} a_k x^{k-1}$ *по формуле*

$$a(x + x^2) = \frac{q}{(1-x)(1+q^2+q^2x)}. \tag{18}$$

Доказательство. Индукцией по n покажем, что можно выбрать числа $a_1, a_2, \ldots$ так, что

$$\chi(t) = 1 - t, \quad q^{n+1} \le t \le q^n.$$

При $n = 0$ из (17) имеем

$$\chi_{1,1}(t) = a_1\chi(t).$$

Из (18) имеем

$$a_1 = \frac{q}{1+q^2}$$

Поэтому третье уравнение в (11) при $n = j = 1$ переписывается в следующем виде:

$$\chi(t) = \frac{q^3}{1+q^2}\chi\left(\frac{t-q^2}{q}\right) + \frac{q^2}{1+q^2}\chi\left(\frac{t-q}{q^2}\right). \tag{19}$$

Поскольку отображения $S_0(t) = q^2 + qt$, $S_1(t) = q + q^2t$ являются сжатиями, то по известной теореме Хатчинсона [4] это уравнение с условием нормировки $\chi(0) = 1$ имеет единственное решение. Нетрудно проверить, что

$$\chi(t) = 1 - t$$

является решением уравнения (19).

Итак, из третьего уравнения (11) имеем

$$\chi(t) = 1 - t, \quad q \le t \le 1.$$

Введем обозначения. Положим $\chi_{0,0}(t) = 1 - t$, тогда из уравнений (16) находятся функции $\chi_{n,j}^{(m)}(t)$. По следствию 1 имеем, что из того, что $\chi_{0,0}(t) = 1 - t$, следует, что функции $\chi_{n,j}^{(m)}(t)$ являются линейными, поэтому положим

$$\chi_{n,j}^{(m)}(t) = B_{n,j}^{(m)} - A_{n,j}^{(m)}t. \tag{20}$$

Подставляя в первое уравнение (16), получим

$$\begin{aligned} A_{n+1,2j+1} &= \tfrac{q}{1+q^2}A_{n,j}^{(0)}, \\ B_{n+1,2j+1} &= B_{n,j}^{(0)} - \tfrac{1-q+q^2}{1+q^2}A_{n,j}^{(0)}. \end{aligned} \tag{21}$$

Подставляя во второе уравнение (16), получим

$$A^{(0)}_{n+1,2j} = A^{(1)}_{n,j} - qA_{n+1,2j+1},$$
$$B^{(0)}_{n+1,2j} = q^{-1}B^{(1)}_{n,j} - qB_{n+1,2j+1} - q^2 A_{n+1,2j+1}. \tag{22}$$

Подставляя в третье уравнение (16), получим

$$A^{(m)}_{n+1,2j} = A^{(m+1)}_{n,j} - A^{(m-1)}_{n+1,2j},$$
$$B^{(m)}_{n+1,2j} = q^{-1}B^{(m+1)}_{n,j} - qB^{(m-1)}_{n+1,2j}. \tag{23}$$

Найдем величины $A^{(m)}_{n,0}$. Через

$$\alpha(x,y) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} A^{(m)}_{n,0} x^m y^n \tag{24}$$

обозначим производящую функцию этих величин.

Умножая уравнения (22)– (23) на $x^{m+1}y^{n+1}$ и суммируя, получим

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} A^{(m)}_{n+1,0} x^{m+1} y^{n+1} =$$
$$= x\alpha(x,y) - x\alpha(x,0) =$$
$$= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} A^{(m+1)}_{n+1,0} x^{m+1} y^{n+1} - \sum_{n=0}^{\infty}\sum_{m=1}^{\infty} A^{(m-1)}_{n+1,0} x^{m+1} y^{n+1} - \frac{q^2}{1+q^2}\sum_{n=0}^{\infty} A^{(0)}_{n,0} x^1 y^{n+1} =$$
$$= y\alpha(x,y) - y\alpha(0,y) - x^2\alpha(x,y) + x^2\alpha(x,0) - \frac{q^2}{1+q^2}xy\alpha(0,y).$$

Подставляя $\alpha(x,0) = \frac{1}{1-x}$, получим

$$(y - x - x^2)\alpha(x,y) = \left(y + \frac{q^2}{1+q^2}xy\right)\alpha(0,y) - \frac{x+x^2}{1-x}. \tag{25}$$

Поскольку функция $\alpha(x,y)$ аналитична в области $|x| < 1$, $|y| < 1$, то на кривой $y = x + x^2$ правая и левая части уравнения (25) обращаются в 0. Поэтому

$$\alpha(0, x + x^2) = \frac{1+q^2}{(1-x)(1+q^2+q^2x)}. \tag{26}$$

Следовательно, для производящей функции $a(x)$ справедливо уравнение (18).

Найдем величины $B^{(m)}_{n,0}$. Через

$$\beta(x,y) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} B^{(m)}_{n,0} x^m y^n \tag{27}$$

обозначим производящую функцию этих величин.

Умножая уравнения (22)– (23) на $x^{m+1}y^{n+1}$ и суммируя, получим

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} B_{n+1,0}^{(m)} x^{m+1}y^{n+1} =$$
$$= x\beta(x,y) - x\beta(x,0) =$$
$$= q^{-1}\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} B_{n+1,0}^{(m+1)} x^{m+1}y^{n+1} - q\sum_{n=0}^{\infty}\sum_{m=1}^{\infty} B_{n+1,0}^{(m-1)} x^{m+1}y^{n+1} -$$
$$- q\sum_{n=0}^{\infty} B_{n,0}^{(0)} x^{1}y^{n+1} + \frac{q^3}{1+q^2}\sum_{n=0}^{\infty} A_{n,0}^{(0)} x^{1}y^{n+1} =$$
$$= q^{-1}y\beta(x,y) - q^{-1}y\beta(0,y) - qx^2\beta(x,y) + qx^2\beta(x,0) - qxy\beta(0,y) + \frac{q^3}{1+q^2}xy\alpha(0,y).$$

Подставляя $\beta(x,0) = \frac{1}{1-x}$, получим

$$(y - qx - q^2x^2)\beta(x,y) = \left(y + q^2xy\right)\beta(0,y) - \frac{q^4}{1+q^2}xy\alpha(0,y) - \frac{qx+q^2x^2}{1-x}. \tag{28}$$

Поскольку функция $\beta(x,y)$ аналитична в области $|x| < 1$, $|y| < 1$, то на кривой $y = qx + q^2x^2$ правая и левая части уравнения (28) обращаются в 0. Поэтому

$$(1 + q^2x)\beta(0, qx + q^2x^2) = \frac{1}{1-x} + \frac{q^4x}{1+q^2}\alpha(0, qx + q^2x^2). \tag{29}$$

Утверждение 1. *Функции $\alpha(0,y)$, $\beta(0,y)$ удовлетворяют равенству*

$$(1-y)\beta(0,y) = \left(1 - \frac{2q^2}{1+q^2}y\right)\alpha(0,y). \tag{30}$$

Доказательство. Умножив равенство (29) на $(1-x)$ и сделав замену переменной $t = qx$, получим

$$(1 - q^{-1}t)(1 + qt)\beta(0, t + t^2) = 1 + \frac{q^2t(q-t)}{1+q^2}\alpha(0, t + t^2).$$

Подставляя (26) и упрощая, получим

$$(1 - t - t^2)\beta(0, t + t^2) =$$
$$= \left((1-t)\left(1 + \frac{q^2}{1+q^2}t\right) + \frac{q^2t(q-t)}{1+q^2}\right)\frac{1+q^2}{(1-t)(1+q^2+q^2t)} =$$
$$= \left(1 - \frac{2q^2}{1+q^2}t - \frac{2q^2}{1+q^2}t^2\right)\alpha(0, t + t^2).$$

Подставляя $y = t + t^2$, получим (30). □

Приравнивая коэффициенты при y^n в (30), получим

Следствие 2. *Справедливы равенства*

$$B_{n,0}^{(0)} - B_{n-1,0}^{(0)} = A_{n,0}^{(0)} - \frac{2q^2}{1+q^2} A_{n-1,0}^{(0)}. \tag{31}$$

Утверждение 2. *Функции $\alpha(x, y)$, $\beta(x, y)$ удовлетворяют равенству*

$$(1-x)\beta(x,y) - (1-qx)\alpha(qx,y) = \beta(0,y) - \alpha(0,y). \tag{32}$$

Доказательство. Применяя (28) и (25), получим

$$\begin{aligned}(y - qx - q^2x^2)\left((1-x)\beta(x,y) - (1-qx)\alpha(qx,y) - \beta(0,y) + \alpha(0,y)\right) = \\ = \left(y(1-x)(1+q^2x) - y + qx + q^2x^2\right)\beta(0,y) - \\ - \left(y(1-qx)(\left(1 + \frac{q^3}{1+q^2}x\right) - y + qx + q^2x^2\right)\alpha(0,y) = \\ = (1-y)(qx+q^2x^2)\beta(0,y) - \left(1 - \frac{2q^2}{1+q^2}y\right)(qx+q^2x^2)\alpha(0,y).\end{aligned}$$

Подставляя (30), получим (32). □

Приравнивая коэффициенты при $x^{m+1}y^n$ в (32), получим

Следствие 3. *Справедливы равенства*

$$B_{n,0}^{(m+1)} - q^{m+1}A_{n,0}^{(m+1)} = B_{n,0}^{(m)} - q^{m+1}A_{n,0}^{(m)}.$$

Полученное равенство означает, что

$$\chi_{n,0}^{(m+1)}(q^{m+1}) = \chi_{n,0}^{(m)}(q^{m+1}),$$

т.е. что функция $\chi_{n,0}(t)$ непрерывна при $t > 0$.

Отметим, что коэффициенты $B_{n,j}^{(m)}$ обеспечивают непрерывность функции $\chi_{n,j}(t)$ и при $j > 0$, но далее будут нужны только коэффициенты $A_{n,j}^{(m)}$ при $j = 0, 1$, а для коэффициентов $B_{n,j}^{(m)}$ ограничимся уравнением

$$B_{n,1} = B_{n,0}^{(0)} - A_{n,0}^{(0)}, \tag{33}$$

которое вытекает при подстановке (21) в равенство (31).

Отметим, что это уравнение эквивалентно тому, что $\chi_{n,1}(0) = \chi_{n,0}^{(0)}(1)$.

Перейдем к доказательству леммы.

Из уравнений (21)–(23) имеем

$$\begin{aligned} &A_{n+1,1} = \tfrac{q}{1+q^2} A_{n,0}^{(0)}, \\ &A_{n+1,0}^{(0)} = A_{n,0}^{(1)} - qA_{n+1,1}, \\ &A_{n+1,0}^{(m)} = A_{n,0}^{(m+1)} - A_{n+1,0}^{(m-1)}, \quad m = 1, 2, \ldots \end{aligned} \tag{34}$$

С начальным условием $A^{(m)}_{0,0} = 0$.

Выберем числа $a_1, a_2, \dots$ так, что

$$a_n = A_{n,1}, \quad n = 1, 2, \dots . \tag{35}$$

Покажем, что при выборе чисел $a_1, a_2, \dots, a_n$ по формуле (35) для функции $\chi(t)$ выполняется условие

$$\chi(t) = 1 - t, \quad t \geq q^n. \tag{36}$$

Обозначим через $A'^{(m)}_{k,0}$ величины, полученные из решения уравнений (34), при заданных $A_{k,1} = a_k$, $k = 1, 2, \dots, n$.

Для удобства перепишем уравнения (34) в следующем виде:

$$\begin{aligned} A'^{(0)}_{k,0} &= \tfrac{1+q^2}{q} A_{k+1,1}, \\ A'^{(1)}_{k,0} &= A'^{(0)}_{k+1,0} + q A_{k+1,1}, \\ A'^{(m+1)}_{k,0} &= A'^{(m)}_{k+1,0} + A'^{(m-1)}_{k+1,0}, \quad m = 1, 2, \dots \end{aligned} \tag{37}$$

Покажем, что при $m = 0, 1 \dots, n-k$, $k = n, n-1, \dots, 0$, выполняется равенство.

$$A'^{(m)}_{k,0} = A^{(m)}_{k,0}.$$

При $k = n$ это равенство выполнено по построению. Поскольку в формулах (37) значения величин $A'^{(m)}_{k,0}$ находятся через $A'^{(m)}_{k+1,0}$ и $A_{k+1,1}$, то эти равенства выполняются и для меньших k.

Следовательно,

$$A'^{(m)}_{0,0} = A^{(m)}_{0,0} = 1, \quad m = 1, 2, \dots, n.$$

т.е. уравнение (36) выполнено.

Подставляя в (18) $x = q$, получим

$$a(1) = \sum_{k=1}^{\infty} a_k = 1.$$

Покажем, что выполнено равенство (17).

По построению

$$\chi_{n,1}(t) = B_{n,1} - A_{n,1} + A_{n,1}\chi_{0,0}(t).$$

Для того, чтобы это уравнение выполнялось, достаточно взять

$$\gamma(2^{-n} + 2^{-n}t) = B_{n,1} - A_{n,1} + A_{n,1}\gamma(t), \quad n = 1, 2, \dots .$$

Подставляя в уравнение (33) второе уравнение из (22) при $j = 0$, получим

$$B_{n+1,1} - B_{n,1} = A_{n+1,1}.$$

Подставляя обозначения (35), получим (17).

□

Лемма 4. *Существует функция $\gamma(t)$, удовлетворяющая условию (17) (для некоторых констант a_i), для которой функция*

$$f_k(t) = (1-t)^k - \int_\Omega \gamma(\nu(t,\omega))\, d\mu(\omega)$$

является квазимногочленом степени $k-1$.

Доказательство. Индукцией по n покажем, что можно выбрать числа $a_1, a_2, \ldots, a_n$ так, что $f_k(t)$ является квазимногочленом степени $k-1$ при $t \geq q^n$.

При $n = 0$ из (17) имеем

$$\chi_{1,1}(t) = a_1 \chi(t).$$

Поэтому третье уравнение в (11) при $n = j = 1$ переписывается в следующем виде:

$$\chi(t) = a_1 q^2 \chi\left(\frac{t-q^2}{q}\right) + a_1 q \chi\left(\frac{t-q}{q^2}\right). \tag{38}$$

Поскольку отображения $S_0(t) = q^2 + qt$, $S_1(t) = q + q^2 t$ являются сжатиями, то по известной теореме Хатчинсона [4] это уравнение с условием нормировки $\chi(0) = 1$ имеет единственное решение.

Нетрудно проверить, что

$$\chi(t) = (1-t)^k$$

является решением уравнения (38), если выбрать

$$a_1 = \frac{q^{2k-1}}{1+q^{k+1}}. \tag{39}$$

Итак, из третьего уравнения (11) имеем $f_k(t) = 0$ при $t \geq q$.

Введем обозначения. Положим $\chi_{0,0}(t) = (1-t)^k$, тогда из уравнений (16) находятся функции $\chi^{(m)}_{n,j}(t)$. По следствию 1 имеем, что из того, что $\chi_{0,0}(t) = (1-t)^k$ следует, что функции $\chi^{(m)}_{n,j}(t)$ являются многочленами степени k, поэтому положим

$$\chi^{(m)}_{n,j}(t) = A^{(m)}_{n,j}(1-t)^k + B^{(m)}_{n,j}(t), \tag{40}$$

где $B^{(m)}_{n,j}(t)$ – многочлен степени $k-1$.

Подставляя (40) в уравнение (16), получим

$$\begin{aligned} A_{n+1,2j+1} &= \tfrac{q^{2k-1}}{1+q^{k+1}} A^{(0)}_{n,j}, \\ A^{(0)}_{n+1,2j} &= q^{k-1} A^{(1)}_{n,j} - q A_{n+1,2j+1}, \\ A^{(m)}_{n+1,2j} &= q^{k-1} A^{(m+1)}_{n,j} - q^{1-k} A^{(m-1)}_{n+1,2j}. \end{aligned} \tag{41}$$

С начальным условием $A^{(m)}_{0,0} = 0$.

Выберем числа $a_1, a_2, \ldots$ так, что

$$a_n = A_{n,1}, \quad n = 1, 2, \ldots \tag{42}$$

Покажем, что при выборе чисел $a_1, a_2, \ldots, a_N$ по формуле (42) для функции $\chi(t)$ выполняется условие $\chi(t) - (1-t)^k$ – квазимногочлен степени $k-1$ при $t \geq q^N$.

Обозначим через $A'^{(m)}_{n,0}$ величины, полученные из решения уравнений (41), при заданных $A_{n,1} = a_n$, $n = 1, 2, \ldots, N$.

Для удобства перепишем уравнения (41) при $j = 0$ в следующем виде:

$$\begin{aligned} A'^{(0)}_{n,0} &= \tfrac{1+q^{k+1}}{q^{2k-1}} A_{n+1,1}, \\ A'^{(1)}_{n,0} &= q^{1-k} A'^{(0)}_{n+1,0} + q^{2-k} A_{n+1,1}, \\ A'^{(m+1)}_{n,0} &= q^{1-k} A'^{(m)}_{n+1,0} + q^{2-2k} A'^{(m-1)}_{n+1,0}, \quad m = 1, 2, \ldots \end{aligned} \tag{43}$$

Покажем, что при $m = 0, 1 \ldots, N-n$, $n = N, N-1, \ldots, 0$, выполняется равенство

$$A'^{(m)}_{n,0} = A^{(m)}_{n,0}.$$

При $n = N$ это равенство выполнено по построению. Поскольку в формулах (43) значения величин $A'^{(m)}_{n,0}$ находятся через $A'^{(m)}_{n+1,0}$ и $A_{n+1,1}$, то эти равенства выполняются и для меньших n.

Следовательно,

$$A'^{(m)}_{0,0} = A^{(m)}_{0,0} = 1, \quad m = 1, 2, \ldots, N,$$

т.е. $\chi(t) - (1-t)^k$ – квазимногочлен степени $k-1$.

Как и при доказательстве предыдущей леммы, можно найти производящую функцию $\alpha(x, y)$ (26) величин $A^{(m)}_{n,0}$.

Умножая уравнения (41) (при $j = 0$) на $x^{m+1} y^{n+1}$ и суммируя, получим

$$\begin{aligned} &\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} A^{(m)}_{n+1,0} x^{m+1} y^{n+1} = \\ &= x\alpha(x,y) - x\alpha(x,0) = \\ &= q^{k-1} y \sum_{m=1}^{\infty} \sum_{n=0}^{\infty} A^{(m)}_{n,0} x^m y^n - q^{1-k} x^2 \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} A^{(m)}_{n,0} x^m y^n - \frac{q^{2k}}{1+q^{k+1}} xy \sum_{n=0}^{\infty} A^{(0)}_{n,0} y^n = \\ &= q^{k-1} y\alpha(x,y) - q^{k-1} y\alpha(0,y) - q^{1-k} x^2 \alpha(x,y) + q^{1-k} x^2 \alpha(x,0) - \frac{q^{2k}}{1+q^{k+1}} xy\alpha(0,y). \end{aligned}$$

Подставляя $\alpha(x, 0) = \frac{1}{1-x}$, получим

$$(q^{k-1} y - x - q^{1-k} x^2)\alpha(x,y) = \left(q^{k-1} y + \frac{q^{2k}}{1+q^{k+1}} xy \right) \alpha(0,y) - \frac{x + q^{1-k} x^2}{1-x}. \tag{44}$$

Поскольку функция $\alpha(x, y)$ аналитична в области $|x| < 1$, $|y| < 1$, то на кривой $q^{k-1} y - x - q^{1-k} x^2 = 0$ правая и левая части уравнения (44) обращаются в 0. Поэтому

$$\alpha(0, q^{1-k} x + q^{2-2k} x^2) = \frac{1 + q^{k+1}}{(1-x)(1 + q^{k+1} + q^{k+1} x)}.$$

Следовательно, производящая функция $a(x)$ задается как

$$a(q^{1-k}x + q^{2-2k}x^2) = \frac{q^{2k-1}}{(1-x)(1+q^{k+1}+q^{k+1}x)}.$$

$\square$

Перейдем к доказательству теоремы.

Обозначим через $\gamma_1(t)$ функцию, которая построена в лемме 3, т.е.

$$\int_\Omega \gamma_1(\nu(t,\omega))\, d\mu(\omega) = 1 - t.$$

По лемме 4 существует функция $\gamma_k(t)$, для которой

$$\int_\Omega \gamma_k(\nu(t,\omega))\, d\mu(\omega) = (1-t)^k.$$

Следовательно, для функции

$$\gamma(t) = \frac{1}{h}\sum_{k=1} \frac{1}{k}\gamma_k(t)$$

выполнено условие (8).

$\square$

Подчеркнем, что для функции $\gamma(t)$, найденной в теореме, оценка $\eta_n^{(k)}(\gamma)$ является несмещенной, а обобщение оценки Грассбергера $r_n^{(k)}(\gamma)/\ln n$ – можно тоже исправить до несмещенной оценки, исправив постоянный множитель – величину $\frac{r_n^{(k)}(\gamma)}{H_n - H_{k-1}}$.

В заключение приведем эскиз графика функции $\gamma_1(t)$, построенной в лемме 3.

Функция $\gamma_1(t)$ разрывна в каждой двоично-рациональной точке, поэтому будет приведен только эскиз графика, который получен соединением значений функции $\gamma_1(t)$ в точках $j2^{-12}$, $j = 0, 1, \ldots, 2^{-12}$. Эти значения находятся как

$$\gamma_1((2j+1)2^{-n}) = \chi_{n,2j+1}(0) = B_{n,2j+1}, \quad j = 0, 1, \ldots, 2^{n-1} - 1.$$

Величины $B_{n,2j+1}$ находим из рекуррентных уравнений (21)–(23).

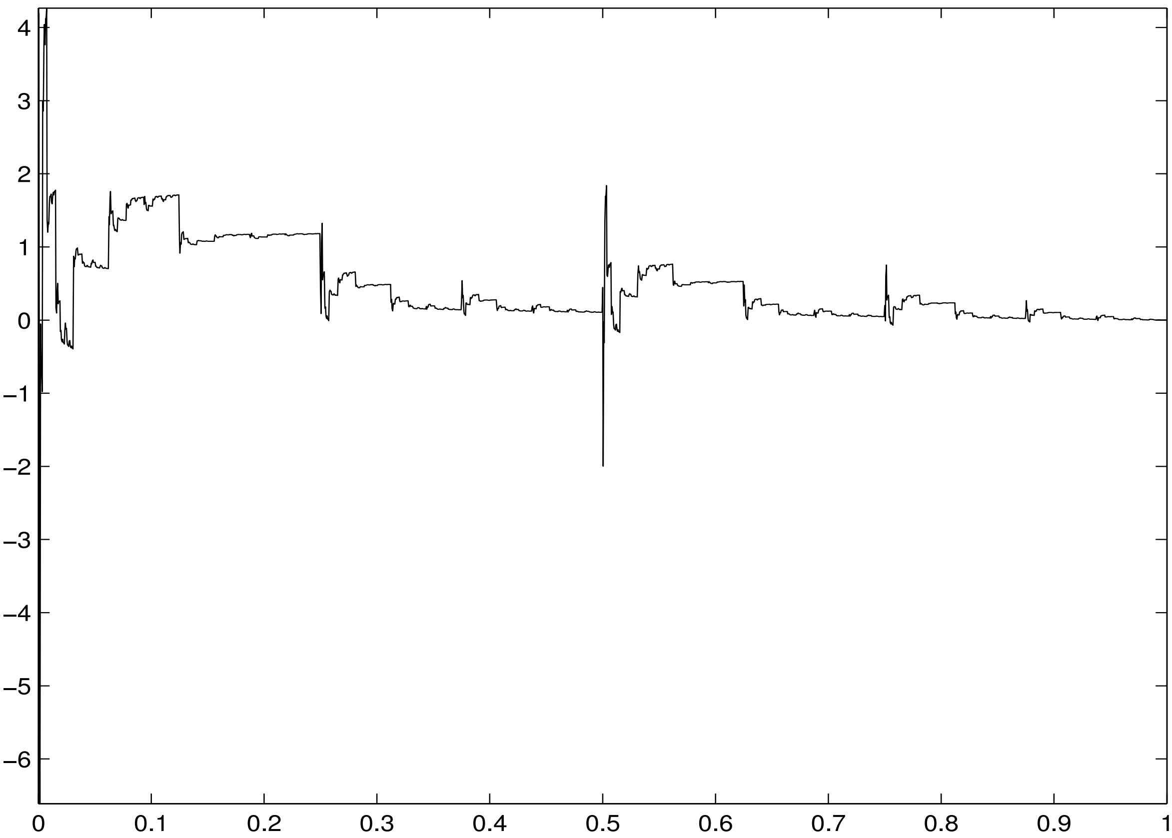

Рис. 1. График функции $\gamma_1(t)$
Fig. 1. The function $\gamma_1(t)$

Список литературы / References

[1] Falconer K. J., *Fractal geometry: Mathematical Foundation and Applications*, John Wiley & Sons, NY, USA, 1990.

[2] Gradshtein I. S., Ryzhik I. M., *Table of integrals, Series, and Products*, Fifth Edition, Academic Press, 1994.

[3] Grassberger P., "Estimating the information content of symbol sequences and efficient codes", *IEEE Trans. Inform. Theory*, **35** (1989), 669–675.

[4] Hutchinson J. E., "Fractals and sel-similarity", *Indiana Univ. Math. J.*, **30** (1981), 713–747.

[5] Timofeev E. A., "Selection of a Metric for the Nearest Neighbor Entropy Estimators", *Journal of Mathematical Sciences*, **203**:6 (2014), 892–906.

[6] Kaltchenko A., Timofeeva N., "Entropy Estimators with Almost Sure Convergence and an $O(n^{-1})$ Variance", *Advances in Mathematics of Communications*, **2**:1 (2008), 1–13.

[7] Kaltchenko A., Timofeeva N., "Rate of convergence of the nearest neighbor entropy estimator", *AEU – International Journal of Electronics and Communications*, **64**:1 (2010), 75–79.

[8] Тимофеева Н. Е., "Построение оценки энтропии для специальной метрики и произвольной функции", *Модел. и анализ информ. систем*, **20**:6 (2013), 174–178; [Timofeeva N. E., "Construction of Entropy Estimator with Special Metric and Arbitrary Function", *Modeling and Analysis of Information Systems*, **20**:6 (2013), 174–178, (in Russian)].

[9] Timofeev E. A., “Bias of a nonparametric entropy estimator for Markov measures”, *Journal of Mathematical Sciences*, **176**:2 (2011), 255–269.

[10] Timofeev E. A., “Statistical Estimation of measure invariants”, *St.Petersburg Math. J.*, **17**:3 (2006), 527–551.

Timofeev E. A., "Existence of an Unbiased Entropy Estimator for the Special Bernoulli Measure", *Modeling and Analysis of Information Systems*, **24**:5 (2017), 521–536.

DOI: 10.18255/1818-1015-2017-5-521-536

Abstract. Let $\Omega = \mathcal{A}^{\mathbb{N}}$ be a space of right-sided infinite sequences drawn from a finite alphabet $\mathcal{A} = \{0, 1\}$, $\mathbb{N} = \{1, 2, \dots\}$,

$$\rho(\boldsymbol{x}, \boldsymbol{y}) = \sum_{k=1}^{\infty} |x_k - y_k| 2^{-k}$$

a metric on $\Omega = \mathcal{A}^{\mathbb{N}}$, and μ is a probability measure on Ω. Let $\boldsymbol{\xi_0}, \boldsymbol{\xi_1}, \dots, \boldsymbol{\xi_n}$ be independent identically distributed points on Ω. We study the estimator $\eta_n^{(k)}(\gamma)$ of the reciprocal of the entropy $1/h$ that are defined as

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

where

$$r_n^{(k)}(\gamma) = \frac{1}{n+1} \sum_{j=0}^{n} \gamma\left(\min_{i:i\neq j}{}^{(k)} \rho(\boldsymbol{\xi_i}, \boldsymbol{\xi_j})\right),$$

$\min^{(k)}\{X_1, \dots, X_N\} = X_k$, if $X_1 \le X_2 \le \dots \le X_N$. The number k and the function $\gamma(t)$ are auxiliary parameters. The main result of this paper is

Theorem. Let μ be the Bernoulli measure with probabilities $p_0, p_1 > 0$, $p_0 + p_1 = 1$, $p_0 = p_1^2$. There exists a function $\gamma(t)$ such that

$$\mathsf{E}\eta_n^{(k)}(\gamma) = \frac{1}{h}.$$

Keywords: measure, metric, entropy, estimator, unbias, self-similar, Bernoulli measure

On the author:
Evgeniy A. Timofeev, orcid.org/0000-0002-0980-2507, ScD, professor
P.G. Demidov Yaroslavl State University,
14 Sovetskaya str., Yaroslavl 150003, Russia, e-mail: timofeevEA@gmail.com