УДК 519.214.8,519.722

ВВЕДЕНИЕ ИНФОРМАЦИОННОЙ ФУНКЦИИ КУЛЬБАКА – ЛЕЙБЛЕРА С ПОМОЩЬЮ РАЗБИЕНИЙ ВЕРОЯТНОСТНОГО ПРОСТРАНСТВА

Э. Э. СОКОЛ [1)]

[1)]*Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь*

Изучается проблема эквивалентности двух определений информационной функции Кульбака – Лейблера. Обычно ее вводят с помощью интеграла от логарифма плотности одной вероятностной меры по отношению к другой. При этом в последнее время активно исследуется понятие *t*-энтропии динамической системы, которое вводится с помощью измеримых разбиений фазового пространства и является обобщением информационной функции. Решается вопрос о том, в каких ситуациях указанные определения эквивалентны, а в каких нет. В частности, эквивалентность имеет место, если обе меры конечны.

Ключевые слова: информационная функция Кульбака – Лейблера; *t*-энтропия.

INTRODUCTION OF THE KULLBACK – LEIBLER INFORMATION FUNCTION BY MEANS OF PARTITIONS OF THE PROBABILITY SPACE

E. E. SOKAL [a]

[a]*Belarusian State University, 4 Niezaliežnasci Avenue, Minsk 220030, Belarus*

In the paper we study the equivalence problem of two definitions of the Kullback – Leibler information function. It is commonly defined by integration of the logarithm of the density of one probability measure with respect to another. On the other hand, recently a concept of *t*-entropy of a dynamical system (that is a generalization of the information function) is actively explored, and this concept is defined by means of measurable partitions of the phase space. In the paper we investigate in which situations the two definitions are equivalent and in which ones they are not. In particular, the equivalence holds if both measures are finite.

Key words: Kullback – Leibler information function; *t*-entropy.

Информационная функция Кульбака – Лейблера в теории вероятностей и математической статистике рассматривается как мера уклонения некоторой вероятностной меры ν от другой вероятностной меры μ. Она используется для оценки вероятностей больших уклонений в последовательности независимых испытаний. Если говорить точнее, она определяет логарифмическую асимптотику вероятности (по отношению к распределению μ) множества тех исходов, которые порождают эмпирические меры, близкие к ν [1].

Образец цитирования:
Сокол Э. Э. Введение информационной функции Кульбака – Лейблера с помощью разбиений вероятностного пространства // Журн. Белорус. гос. ун-та. Математика. Информатика. 2018. № 1. С. 59–67.

For citation:
Sokal E. E. Introduction of the Kullback – Leibler information function by means of partitions of the probability space. *J. Belarus. State Univ. Math. Inform.* 2018. No. 1. P. 59–67 (in Russ.).

Автор:
Эдвард Эдуардович Сокол – аспирант кафедры функционального анализа и аналитической экономики механико-математического факультета. Научный руководитель – доктор физико-математических наук, профессор В. И. Бахтин.

Author:
Edvard E. Sokal, postgraduate student at the department of functional analysis and analytical economics, faculty of mechanics and mathematics.
edward.e.sokol@gmail.com

Информационная функция Кульбака – Лейблера от пары аргументов ν и μ, являющихся вероятностными мерами на одном и том же измеримом пространстве, определяется при помощи формулы

$$\rho(\nu, \mu) = \int \varphi \ln \varphi \, d\mu, \;\; \varphi = \frac{d\nu}{d\mu}, \tag{1}$$

если ν абсолютно непрерывна относительно μ. Если же ν не абсолютно непрерывна относительно μ, значение $\rho(\nu, \mu)$ полагают равным $+\infty$. В формуле (1) при $\varphi = 0$ значение выражения $\varphi \ln \varphi$ считают равным нулю. При этом в работах [2–4] был рассмотрен аналог информационной функции для динамических систем, который называется t-энтропией. При определении t-энтропии используется не интеграл, а конечные разбиения фазового пространства. Кроме того, в работе [5] информационная функция $\rho(\nu, \mu)$ вводится для конечно-аддитивных мер ν, что аналогично определению t-энтропии и также базируется на конечных измеримых разбиениях пространства. Таким образом, в случае счетно-аддитивной меры ν имеются два разных определения информационной функции: с помощью формулы (1) и с помощью измеримых разбиений пространства. Цель настоящей работы – исследование их эквивалентности. В теореме 2, приведенной ниже, эта эквивалентность доказывается в случае конечной меры μ. В теореме 3, представленной далее, она устанавливается для σ-конечной меры μ и абсолютно непрерывной меры ν, удовлетворяющей условию

$$\int\limits_{\{\varphi < 1\}} \varphi \ln \varphi \, d\mu > -\infty, \;\; \varphi = \frac{d\nu}{d\mu}. \tag{2}$$

Приведенные в настоящей работе примеры показывают, что при нарушении условия (2), а также в случае бесконечной меры μ и не абсолютно непрерывной меры ν указанные два определения неэквивалентны.

Случай конечной меры μ

Пусть X – произвольное множество, $\mathfrak{A}$ – σ-алгебра его подмножеств. Введем следующие обозначения: $M(X)$ – множество конечных мер на $(X, \mathfrak{A})$, $M_1(X)$ – множество вероятностных мер на $(X, \mathfrak{A})$, $M_\sigma(X)$ – множество σ-конечных мер на $(X, \mathfrak{A})$.

Очевидно, справедлива цепочка включений:

$$M_1(X) \subset M(X) \subset M_\sigma(X).$$

Возьмем меры $\nu \in M_1(X)$ и $\mu \in M(X)$. В этом случае информационная функция $\rho(\nu, \mu)$ определяется с помощью (1), где интеграл берется по множеству X.

Пусть $G = \{X_1, X_2, \dots\}$ – произвольное дискретное, т. е. конечное либо счетное, разбиение X на измеримые подмножества X_i. Определим информационную функцию $\rho'(\nu, \mu)$ с помощью формулы

$$\rho'(\nu, \mu) = \sup_G \sum_{X_i \in G} \nu(X_i) \ln \frac{\nu(X_i)}{\mu(X_i)}. \tag{3}$$

Если $\nu(X_i) = 0$, то тогда соответствующее слагаемое в (3) считаем равным нулю (в том числе и при равенстве $\mu(X_i) = 0$). Если же $\nu(X_i) > 0$ и $\mu(X_i) = 0$, то соответствующее слагаемое и вся сумма в формуле (3) считаются равными $+\infty$.

Теорема 1. *Для конечных мер $\mu \in M(X)$ в определении (3) достаточно использовать лишь конечные разбиения G. От этого значение правой части (3) не изменится.*

Доказательство. Для любого $\varepsilon > 0$ существует такое конечное или счетное разбиение $G = \{X_1, X_2, \dots\}$, что

$$\sum_{X_i \in G} \nu(X_i) \ln \frac{\nu(X_i)}{\mu(X_i)} > \rho'(\nu, \mu) - \varepsilon. \tag{4}$$

Покажем, что разбиение G в (4) можно выбрать конечным. Предположим, что G есть счетное разбиение. Рассмотрим множества $Y_n = \bigcup_{i > n} X_i$. Очевидно, $\bigcap_n Y_n = \varnothing$. Поэтому $\nu(Y_n) \to 0$ и

$$\nu(Y_n)\ln\frac{\nu(Y_n)}{\mu(Y_n)} \geq \nu(Y_n)\ln\frac{\nu(Y_n)}{\mu(X)} \to 0. \tag{5}$$

Рассмотрим конечное разбиение $G_n = \{X_1, X_2, \dots, X_n, Y_n\}$. Из (5) следует, что

$$\sup_n\left(\sum_{i=1}^{n}\nu(X_i)\ln\frac{\nu(X_i)}{\mu(X_i)} + \nu(Y_n)\ln\frac{\nu(Y_n)}{\mu(Y_n)}\right) \geq \sum_{X_i\in G}\nu(X_i)\ln\frac{\nu(X_i)}{\mu(X_i)}.$$

Отсюда видно, что если в неравенстве (4) заменить G на конечное разбиение G_n с достаточно большим номером n, то оно останется в силе. В итоге для любого $\varepsilon > 0$ имеем конечное разбиение G, удовлетворяющее (4). В силу произвольности ε отсюда следует, что супремум в (3) не зависит от того, какие разбиения G рассматриваются – конечные или счетные.

Теорема 2. *Для* $\nu \in M_1(X)$ *и* $\mu \in M(X)$ *выполняется равенство*

$$\rho(\nu, \mu) = \rho'(\nu, \mu).$$

Доказательство. Если мера $\nu \in M_1(X)$ не является абсолютно непрерывной относительно $\mu \in M(X)$, то по определению $\rho(\nu, \mu) = +\infty$. Кроме того, в этом случае существует такое множество $A \in \mathfrak{A}$, что $\mu(A) = 0$ и $\nu(A) > 0$. Рассмотрим разбиение $G = \{A, X\setminus A\}$. Для него формула (3) дает равенство $\rho'(\nu, \mu) = +\infty$, т. е. в случае меры $\nu \in M_1(X)$, которая не абсолютно непрерывна относительно $\mu \in M(X)$, равенство ρ и ρ' выполняется.

Пусть теперь ν абсолютно непрерывна относительно μ. Тогда по теореме Радона – Никодима она имеет плотность $\varphi = \frac{d\nu}{d\mu}$. Определим множество $Y = \{x \in X \mid \varphi(x) = 0\}$. Очевидно, $\nu(Y) = 0$. Рассмотрим любое дискретное измеримое разбиение G пространства X. Для него справедлива следующая выкладка:

$$-\int_X \varphi\ln\varphi\, d\mu = -\int_{X\setminus Y}\ln\frac{d\nu}{d\mu}\,d\nu = \int_{X\setminus Y}\ln\frac{d\mu}{d\nu}\,d\nu =$$

$$= \sum_{X_i\in G:\,\nu(X_i)>0}\nu(X_i)\int_{X_i\setminus Y}\frac{\ln\left(\frac{d\mu}{d\nu}\right)}{\nu(X_i)}\,d\nu \leq \sum_{X_i\in G:\,\nu(X_i)>0}\nu(X_i)\ln\left(\int_{X_i\setminus Y}\frac{\frac{d\mu}{d\nu}}{\nu(X_i)}\,d\nu\right) =$$

$$= \sum_{X_i\in G:\,\nu(X_i)>0}\nu(X_i)\ln\frac{\mu(X_i\setminus Y)}{\nu(X_i)} \leq -\sum_{X_i\in G}\nu(X_i)\ln\frac{\nu(X_i)}{\mu(X_i)}, \tag{6}$$

где в средней строке для каждого слагаемого используется неравенство Йенсена, т. е. для любого дискретного измеримого разбиения G верно неравенство

$$\rho(\nu, \mu) = \int_X \ln\frac{d\nu}{d\mu}\,d\nu \geq \sum_{X_i\in G}\nu(X_i)\ln\frac{\nu(X_i)}{\mu(X_i)}.$$

Отсюда вытекает, что $\rho(\nu, \mu) \geq \rho'(\nu, \mu)$.

Фиксируем произвольное $\varepsilon > 0$. Рассмотрим множества:

$$X_i = \{x \in X \mid (i-1)\varepsilon \leq \ln\varphi < i\varepsilon\},\ i \in \mathbb{Z}.$$

Очевидно, для них справедливы неравенства:

$$e^{(i-1)\varepsilon}\mu(X_i) \leq \nu(X_i) = \int_{X_i}\varphi\,d\mu \leq e^{i\varepsilon}\mu(X_i).$$

Из них следует, что

$$(i-1)\varepsilon \leq \ln\frac{\nu(X_i)}{\mu(X_i)} \leq i\varepsilon, \text{ если } \mu(X_i) > 0. \tag{7}$$

Также для $x \in X_i$ выполняется

$$(i-1)\varepsilon \le \ln \varphi(x) \le i\varepsilon. \tag{8}$$

Из (7) и (8) следует

$$-\varepsilon\nu(X_i) + \nu(X_i)\ln\frac{\nu(X_i)}{\mu(X_i)} \le \int_{X_i} \ln\varphi\, d\nu \le \nu(X_i)\ln\frac{\nu(X_i)}{\mu(X_i)} + \varepsilon\nu(X_i). \tag{9}$$

Заметим, что в случае $\mu(X_i) = 0$ автоматически $\nu(X_i) = 0$ и неравенства (9) остаются в силе. Они также остаются в силе, если в них заменить X_i на множество $X_{-\infty} = \{x \in X \,|\, \varphi(x) = 0\}$. Рассмотрим разбиение $G = \{X_{-\infty}\} \bigcup \{X_i \,|\, i \in \mathbb{Z}\}$ множества X. Суммируя неравенства (9) по элементам этого разбиения, получаем

$$-\varepsilon + \sum_{X_i \in G} \nu(X_i)\ln\frac{\nu(X_i)}{\mu(X_i)} \le \int_X \ln\varphi\, d\nu \le \sum_{X_i \in G} \nu(X_i)\ln\frac{\nu(X_i)}{\mu(X_i)} + \varepsilon. \tag{10}$$

В силу произвольности ε из правого неравенства (10) вытекает, что $\rho(\nu, \mu) \le \rho'(\nu, \mu)$. Тем самым теорема доказана.

Случай σ-конечной меры μ

Пусть теперь $\nu \in M_1(X)$ и $\mu \in M_\sigma(X)$. В работе [6] было доказано, что в этой ситуации информационную функцию $\rho(\nu, \mu)$ можно определить следующим образом:

$$\rho(\nu, \mu) = \begin{cases} +\infty, \text{ если } \nexists\ \varphi(x) = \dfrac{d\nu(x)}{d\mu(x)}, \\ \displaystyle\int_X \varphi \ln\varphi\, d\mu, \text{ если } \exists\ \varphi(x) = \dfrac{d\nu(x)}{d\mu(x)} \text{ и } \int_{\{\varphi<1\}} \varphi\ln\varphi\, d\mu > -\infty, \\ -\infty, \text{ если } \exists\ \varphi(x) = \dfrac{d\nu(x)}{d\mu(x)} \text{ и } \displaystyle\int_{\{\varphi<1\}} \varphi\ln\varphi\, d\mu = -\infty. \end{cases}$$

Функцию $\rho'(\nu, \mu)$ по-прежнему будем определять формулой (3). Далее, нашей целью является исследование вопроса о том, когда и в каком смысле выполняется равенство $\rho(\nu, \mu) = \rho'(\nu, \mu)$.

Прежде всего заметим, что в случае бесконечной меры μ счетные разбиения G в формуле (3) нельзя заменить на конечные. Действительно, пусть $X = \mathbb{N}$. Рассмотрим такую меру $\mu \in M_\sigma(\mathbb{N})$, что $\mu(i) = 1$, и такую меру $\nu \in M_1(\mathbb{N})$, что $\nu(i) = \dfrac{1}{2^i}$ для всех $i \in \mathbb{N}$. Тогда

$$\rho'(\nu, \mu) \ge \sum_{i=1}^{\infty} \nu(i)\ln\frac{\nu(i)}{\mu(i)} = \sum_{i=1}^{\infty} \frac{1}{2^i}\ln\frac{1}{2^i} = -\sum_{i=1}^{\infty}\frac{i\ln 2}{2^i} > -\infty.$$

С другой стороны, для любого конечного разбиения $G = \{X_1, X_2, \ldots, X_n\}$ множества $\mathbb{N}$ найдется такое i, при котором $\mu(X_i) = +\infty$, и поэтому

$$\sup_G \sum_{X_i \in G} \nu(X_i)\ln\frac{\nu(X_i)}{\mu(X_i)} = -\infty,$$

если супремум берется по конечным разбиениям G.

В связи с этим далее будем считать, что супремум в (3) берется по всем конечным и счетным разбиениям G.

Теорема 3. *Пусть мера* $\nu \in M_1(X)$ *абсолютно непрерывна относительно меры* $\mu \in M_\sigma(X)$ *и плотность* $\varphi = \dfrac{d\nu}{d\mu}$ *удовлетворяет условию*

$$\int\limits_{\{\varphi<1\}} \varphi \ln \varphi \, d\mu > -\infty.$$

Тогда выполняется равенство

$$\rho(\nu, \mu) = \rho'(\nu, \mu).$$

Доказательство этой теоремы дословно воспроизводит доказательство теоремы 2 в случае абсолютно непрерывной меры ν. Поэтому мы не будем его повторять.

Примеры

Далее, исследуем случаи, которые теоремы 2 и 3 не охватывают: 1) случай не абсолютно непрерывной меры $\nu \in M_1(X)$ по отношению к бесконечной мере $\mu \in M_\sigma(X)$; 2) случай абсолютно непрерывной меры $\nu \in M_1(X)$ относительно бесконечной меры $\mu \in M_\sigma(X)$ с плотностью $\varphi = \dfrac{d\nu}{d\mu}$, удовлетворяющей условию

$$\int\limits_{\{\varphi<1\}} \varphi \ln \varphi \, d\mu = -\infty. \tag{11}$$

Сначала установим справедливость следующего утверждения.

Лемма 4. *Пусть заданы конечные меры* $\nu \in M(X)$, $\mu \in M(X)$ *и дискретное измеримое разбиение* $\{X_i\}$ *пространства X. Тогда выполняется неравенство*

$$\nu(X) \ln \frac{\nu(X)}{\mu(X)} \le \sum_i \nu(X_i) \ln \frac{\nu(X_i)}{\mu(X_i)}. \tag{12}$$

Доказательство. Справедливость (12) вытекает из вогнутости логарифмической функции и следующей выкладки, аналогичной выкладке (6):

$$-\sum_i \nu(X_i) \ln \frac{\nu(X_i)}{\mu(X_i)} = \sum_{\nu(X_i)>0} \nu(X_i) \ln \frac{\mu(X_i)}{\nu(X_i)} =$$

$$= \nu(X) \sum_{\nu(X_i)>0} \frac{\nu(X_i)}{\nu(X)} \ln \frac{\mu(X_i)}{\nu(X_i)} \le \nu(X) \ln \left(\sum_{\nu(X_i)>0} \frac{\nu(X_i)}{\nu(X)} \frac{\mu(X_i)}{\nu(X_i)} \right) \le$$

$$\le \nu(X) \ln \frac{\mu(X)}{\nu(X)} = -\nu(X) \ln \frac{\nu(X)}{\mu(X)}.$$

1. Пусть мера $\nu \in M_1(X)$ не абсолютно непрерывна по отношению к бесконечной мере $\mu \in M_\sigma(X)$. Тогда по определению $\rho(\nu, \mu) = +\infty$. С другой стороны, выясним, какой результат может давать формула (3).

В качестве примера рассмотрим пространство $X = \mathbb{N}$, меру μ вида

$$\mu(i) = \begin{cases} 0, & i = 1, \\ 1, & i > 1 \end{cases}$$

и меру ν вида

$$\nu(i) = \frac{C}{(i+1)\ln^2(i+1)}, \quad i = 1, 2, \ldots,$$

где нормировочная константа $C > 0$ определяется условием

$$\sum_{i=1}^{+\infty} \frac{C}{(i+1)\ln^2(i+1)} = 1.$$

Рассмотрим произвольное разбиение $G = \{X_1, X_2, \ldots\}$ множества $\mathbb{N}$. Не ограничивая общности, можно считать, что $1 \in X_1$. При этом возможны три случая:

а) $X_1 = \{1\}$;

б) X_1 конечно и содержит элементы, отличные от единицы;

в) X_1 бесконечно.

В случае а) получаем $\nu(X_1)\ln\dfrac{\nu(X_1)}{\mu(X_1)} = +\infty.$ Применяя лемму 4 к множествам X_i и их разбиениям на отдельные элементы, получаем

$$\sum_{i=2}^{+\infty}\nu(X_i)\ln\frac{\nu(X_i)}{\mu(X_i)} \le \sum_{i=2}^{+\infty}\sum_{j\in X_i}\nu(j)\ln\frac{\nu(j)}{\mu(j)} = \sum_{j=2}^{+\infty}\nu(j)\ln\frac{\nu(j)}{\mu(j)} =$$
$$= \sum_{j=2}^{+\infty}\frac{C}{(j+1)\ln^2(j+1)}\ln\frac{C}{(j+1)\ln^2(j+1)} = \sum_{j=3}^{+\infty}\frac{C\ln C}{j\ln^2 j} - \sum_{j=3}^{+\infty}\frac{C}{j\ln j} - \sum_{j=3}^{+\infty}\frac{2C\ln(\ln j)}{j\ln^2 j} = -\infty.$$

В этой ситуации сумма в правой части (3) обладает тем свойством, что в ней одно слагаемое равно $+\infty$, а сумма остальных слагаемых обращается в $-\infty$, и, следовательно, не существует разумного способа приписать определенное значение правой части (3).

В случае б) выражение $\nu(X_1)\ln\dfrac{\nu(X_1)}{\mu(X_1)}$ конечно и, используя лемму 4, аналогично предыдущей выкладке получаем

$$\sum_{i=2}^{+\infty}\nu(X_i)\ln\frac{\nu(X_i)}{\mu(X_i)} \le \sum_{j\in\mathbb{N}\setminus X_1}\nu(j)\ln\frac{\nu(j)}{\mu(j)} = -\infty.$$

В этой ситуации правая часть (3) обращается в $-\infty$.

В случае в) имеем $\nu(X_1)\ln\dfrac{\nu(X_1)}{\mu(X_1)} = -\infty$ и, поскольку $\nu(j) < \mu(j)$ для всех $j \ge 2$,

$$\sum_{i=2}^{+\infty}\nu(X_i)\ln\frac{\nu(X_i)}{\mu(X_i)} \le 0.$$

Значит, в этой ситуации, так же как и в случае б), правая часть (3) принимает значение $-\infty$.

Если же необходимо иметь равенство $\rho(\nu,\mu) = \rho'(\nu,\mu)$, то неизбежно приходим к необходимости считать значение правой части (3) равным $+\infty$ в случае а), что выглядит неестественно. Другой возможный способ добиться равенства $\rho(\nu,\mu) = \rho'(\nu,\mu)$ в случае не абсолютно непрерывной меры ν – просто положить $\rho'(\nu,\mu) = +\infty$ без какой-либо связи с равенством (3).

2. Теперь рассмотрим случай, когда мера $\nu\in M_1(X)$ абсолютно непрерывна относительно бесконечной меры $\mu\in M_\sigma(X)$ с плотностью $\varphi = \dfrac{d\nu}{d\mu}$, удовлетворяющей условию (11). Тогда по определению $\rho(\nu,\mu) = -\infty$. Выясним, какой результат в этой ситуации может давать формула (3).

В качестве примера рассмотрим множество $X = \mathbb{Z}$, меру μ вида

$$\mu(i) = \begin{cases} 1, & i \ge 1, \\ 2^{i-1}, & i < 1 \end{cases} \tag{13}$$

и меру ν вида

$$\nu(i) = \begin{cases} \dfrac{A}{(i+1)\ln^2(i+1)}, & i \ge 1, \\ \dfrac{B}{(i-1)^2}, & i < 1, \end{cases} \tag{14}$$

где нормировочные константы $A > 0$ и $B > 0$ определяются условиями:

$$\sum_{i=1}^{\infty}\frac{A}{(i+1)\ln^2(i+1)} = \frac{1}{2},\quad \sum_{i=-\infty}^{0}\frac{B}{(i-1)^2} = \frac{1}{2}.$$

Очевидно, ν абсолютно непрерывна относительно меры μ с плотностью

$$\varphi(i)=\begin{cases}\dfrac{A}{(i+1)\ln^2(i+1)}, & i\geq 1,\\ \dfrac{B}{2^{i-1}(i-1)^2}, & i<1.\end{cases}$$

Проверим, что выполняется условие (11). Действительно, нетрудно видеть, что

$$\sum_{i=1}^{+\infty}\varphi(i)\ln(\varphi(i))\mu(i)=\sum_{i=1}^{+\infty}\frac{A}{(i+1)\ln^2(i+1)}\ln\frac{A}{(i+1)\ln^2(i+1)}=-\infty. \tag{15}$$

Кроме того, по построению $\sum_{i=1}^{+\infty}\varphi(i)=\frac{1}{2},$ поэтому $\varphi(i)<1$ при $i\geq 1$.

Также легко видеть, что

$$\sum_{i=-\infty}^{0}\varphi(i)\ln(\varphi(i))\mu(i)=\sum_{i=-\infty}^{0}\frac{B}{(i-1)^2}\ln\frac{B}{2^{i-1}(i-1)^2}=\sum_{i=0}^{+\infty}\frac{B}{(i+1)^2}\ln\frac{B\,2^{i+1}}{(i+1)^2}=\sum_{i=1}^{+\infty}\frac{B}{i^2}\ln\frac{B\,2^i}{i^2}=+\infty. \tag{16}$$

Кроме того, $\varphi(i)\to+\infty$ при $i\to-\infty$. Это означает, что существует такое i', что $\varphi(i)\geq 1$ для всех $i\leq i'$.

Таким образом, в этом примере выполняется (11) и по определению имеем $\rho(\nu,\mu)=-\infty$. А из (15) и (16) следует, что ряд

$$\sum_{i=-\infty}^{+\infty}\nu(i)\ln\frac{\nu(i)}{\mu(i)} \tag{17}$$

расходится в обе стороны.

Зафиксируем произвольное разбиение $G=\{X_1, X_2, \ldots\}$ множества $\mathbb{Z}$. Убедимся, что для заданных формулами (13) и (14) мер у суммы

$$\sum_{X_i\in G}\nu(X_i)\ln\frac{\nu(X_i)}{\mu(X_i)} \tag{18}$$

можно выделить частичную сумму, равную $-\infty$. Если существует такое $X_i\in G$, что $\mu(X_i)=+\infty$, то соответствующее слагаемое в (18) обращается в $-\infty$, и тогда выделяем частичную сумму, состоящую из этого одного слагаемого. Поэтому далее можем предполагать, что $\mu(X_i)<+\infty$ для всех $X_i\in G$.

Определим G_+ как совокупность элементов разбиения G, имеющих непустое пересечение с $\mathbb{N}$. Для удобства введем для элементов G_+ новое обозначение и будем считать, что $G_+=\{Y_1, Y_2, \ldots\}$. Достаточно проверить, что выполняется

$$\sum_{Y_i\in G_+}\nu(Y_i)\ln\frac{\nu(Y_i)}{\mu(Y_i)}=-\infty. \tag{19}$$

В силу леммы 4 сумма в (19) может лишь увеличиться при разбиении каждого Y_i на более мелкие подмножества. Поэтому при доказательстве (19) достаточно считать, что $Y_i\cap\mathbb{N}=\{i\}$. Тогда для $i\geq 1$ будем иметь $\nu(i)\leq\nu(Y_i)\leq 1,\ \mu(Y_i)\geq\mu(i)=1,$

$$\sum_{i=1}^{+\infty}\nu(i)\ln\frac{\nu(i)}{\mu(i)}=\sum_{i=1}^{+\infty}\nu(i)\ln\nu(i)=-\infty, \tag{20}$$

$$\sum_{i=1}^{+\infty}\nu(Y_i)\ln\frac{\nu(Y_i)}{\mu(Y_i)}\leq\sum_{i=1}^{+\infty}\nu(Y_i)\ln\nu(Y_i)\leq\sum_{i=1}^{+\infty}\nu(i)\ln\nu(Y_i). \tag{21}$$

Для того чтобы оценить правую часть (21) сверху, для произвольного фиксированного натурального n рассмотрим задачу на условный экстремум

$$\begin{cases}\sum_{i=1}^{n}\nu(i)\ln y_i\to\operatorname{extr},\\ \sum_{i=1}^{n}y_i=C_n,\ C_n=\sum_{i=1}^{n}\nu(Y_i).\end{cases}$$

Решая ее методом множителей Лагранжа, получаем, что целевая функция достигает локального максимума при

$$y_i = \frac{C_n \nu(i)}{D_n}, \; i = 1, 2, \ldots,$$

где $D_n = \sum_{i=1}^{n} \nu(i).$

Значение этого максимума равно

$$D_n \ln C_n + \sum_{i=1}^{n} \nu(i) \ln \nu(i) - D_n \ln D_n.$$

А поскольку целевая функция, что очевидно, вогнута и рассматривается на выпуклом множестве, то этот максимум является глобальным, т. е. выполняется

$$\sum_{i=1}^{n} \nu(i) \ln \nu(Y_i) \le D_n \ln C_n + \sum_{i=1}^{n} \nu(i) \ln \nu(i) - D_n \ln D_n. \tag{22}$$

Переходя к пределу при $n \to +\infty$ в (22), с учетом оценки (20) имеем

$$\sum_{i=1}^{+\infty} \nu(i) \ln \nu(Y_i) \le -\infty. \tag{23}$$

Таким образом, принимая во внимание оценки (21) и (23), убеждаемся в справедливости (19). А поскольку ряд (17) расходится в обе стороны, то для произвольного разбиения $G = \{X_1, X_2, \ldots\}$ множества $\mathbb{Z}$ для суммы (18) возникает ситуация, аналогичная ситуации из предыдущего примера:

а) либо возможно выделить две частичные суммы, одна из которых будет равна $-\infty$, а другая $+\infty$;

б) либо возможно выделить только одну частичную сумму, равную $-\infty$, а сумма остальных слагаемых будет конечным числом.

Тогда в ситуации а) непонятно, какое значение приписать правой части (3), а в случае б) правая часть (3) обращается в $-\infty$.

Таким образом, в случае, когда мера $\nu \in M_1(X)$ абсолютно непрерывна относительно бесконечной меры $\mu \in M_\sigma(X)$ с плотностью $\varphi = \frac{d\nu}{d\mu}$, удовлетворяющей условию (11), если хотим иметь равенство $\rho(\nu, \mu) = \rho'(\nu, \mu)$, то приходим к необходимости считать значение правой части (3) равным $-\infty$ в случае а). Это вступает в смысловое противоречие с нашими действиями в предыдущем примере, где в таком же случае двусторонней расходимости ряда в правой части (3) приходилось полагать его значение равным $+\infty$. Другой возможный способ добиться равенства $\rho(\nu, \mu) = \rho'(\nu, \mu)$ в этой ситуации – просто положить $\rho'(\nu, \mu) = -\infty$ без какой-либо связи с равенством (3).

Библиографические ссылки

1. *Боровков А. А.* Математическая статистика: Оценка параметров. Проверка гипотез. М. : Наука, 1984.

2. *Bakhtin V. I.* On *t*-entropy and variational principle for the spectral radius of weighted shift operators // Ergod. Theory Dyn. Syst. 2010. Vol. 30, issue 5. P. 1331–1342. DOI: 10.1017/S0143385709000716.

3. *Antonevich A. B., Bakhtin V. I., Lebedev A. V.* On *t*-entropy and variational principle for the spectral radii of transfer and weighted shift operators // Ergod. Theory Dyn. Syst. 2011. Vol. 31, issue 4. P. 995–1042. DOI: 10.1017/S0143385710000210.

4. *Antonevich A. B., Bakhtin V. I., Lebedev A. V.* A road to the spectral radius of transfer operators // Contemp. Math. 2012. Vol. 567. P. 17–51. DOI: 10.1090/conm/567/11252.

5. *Бахтин В. И.* Спектральный потенциал, действие Кульбака и принцип больших уклонений для конечно-аддитивных мер // Тр. Ин-та математики НАН Беларуси. 2015. Т. 24, № 1. С. 1–13.

6. *Bakhtin V., Sokal E.* The Kullback – Leibler Information Function For Infinite Measures // Entropy. 2016. Vol. 18, issue 12. DOI:10.3390/e18120448.

References

1. Borovkov A. A. Matematicheskaya statistika: Otsenka parametrov. Proverka gipotez. M. : Nauka, 1984 (in Russ.).

2. Bakhtin V. I. On *t*-entropy and variational principle for the spectral radius of weighted shift operators. *Ergod. Theory Dyn. Syst.* 2010. Vol. 30, issue 5. P. 1331–1342. DOI: 10.1017/S0143385709000716.

3. Antonevich A. B., Bakhtin V. I., Lebedev A. V. On *t*-entropy and variational principle for the spectral radii of transfer and weighted shift operators. *Ergod. Theory Dyn. Syst.* 2011. Vol. 31, issue 4. P. 995–1042. DOI: 10.1017/S0143385710000210.

4. Antonevich A. B., Bakhtin V. I., Lebedev A. V. A road to the spectral radius of transfer operators. *Contemp. Math.* 2012. Vol. 567. P. 17–51. DOI: 10.1090/conm/567/11252.

5. Bakhtin V. I. [Spectral potential, Kullback action, and large deviation principle for finitely-additive measures]. *Proc. of the Inst. of Math. of the Natl. Acad. of Sci. of Belarus.* 2015. Vol. 24, No. 1. P. 1–13 (in Russ.).

6. Bakhtin V., Sokal E. The Kullback – Leibler Information Function For Infinite Measures. *Entropy*. 2016. Vol. 18, issue 12. DOI:10.3390/e18120448.

Статья поступила в редколлегию 27.10.2017.
Received by editorial board 27.10.2017.