

Оптимизация, системный анализ и исследование операций

© 2023 г. А.А. ГАЛЯЕВ, д-р техн. наук (galaev@ipu.ru),

П.В. ЛЫСЕНКО, канд. физ.-мат. наук (pavellysen@ipu.ru),

Л.М. БЕРЛИН (berlin.lm@phystech.edu)

(Институт проблем управления им. В.А. Трапезникова РАН, Москва)

СТАТИСТИЧЕСКАЯ СЛОЖНОСТЬ КАК КРИТЕРИЙ ЗАДАЧИ ОБНАРУЖЕНИЯ ПОЛЕЗНОГО СИГНАЛА¹

Рассматриваются три варианта функции статистической сложности, которая используется в качестве критерия в задаче обнаружения полезного сигнала в сигнально-шумовой смеси. Получены вероятностные распределения, максимизирующие рассматриваемые варианты статистической сложности, и сделаны выводы об эффективности использования того или иного варианта для задач обнаружения. На примере синтезированных сигналов продемонстрировано сравнение рассмотренных информационных характеристик и проиллюстрированы аналитические результаты. Предложен способ выбора порога информационного критерия, превышение которого позволяет сделать вывод о появлении полезного сигнала в сигнально-шумовой смеси. Выбор порога априорно зависит от максимальных значений, полученных аналитически. В результате сложность на основе полной вариации меры показала наилучшую способность обнаружения полезного сигнала.

Ключевые слова: статистическая сложность, обнаружение сигнала, информационная дивергенция.

DOI: 10.31857/S0005231023070073, **EDN:** FDVSTA

1. Введение

Понятие информационной энтропии было впервые введено в статье Клода Шеннона [1] в 1948 г. Эта работа положила начало новой области науки, получившей название теории информации [2]. Развитие теории информации дало возможность аналитического и практического исследования многих прикладных областей науки и техники. Были введены в обиход и получили свою интерпретацию такие термины, как энтропии Гиббса и фон Неймана, расстояние Кульбака–Лейблера, дивергенция Йенсена–Шеннона, информационные дивергенции и некоторые другие, которые стали использоваться в качестве критериев различных оптимизационных задач распознавания [3], классификации [4], фильтрации и отвечать за существо и качество принимаемых на их основе решений.

¹ Работа выполнена при финансовой поддержке Российского научного фонда (грант № 23-19-00134).

К концу прошлого века различные информационные критерии, в основном информационной энтропии Шеннона [1], начали активно применяться в задачах цифровой обработки сигналов, в частности в задаче обнаружения полезного сигнала в шумовой обстановке [5]. Появилось понятие спектральной энтропии [6], связанной с Фурье-спектром рассматриваемого сигнала, что оказалось особенно актуально в анализе акустических сигналов [7]. Кроме того, энтропийный подход успешно применяется в анализе временных рядов в медицинской сфере, например ЭКГ или ЭЭГ [8]. Позднее была предложена функция статистической сложности как развитие понятия энтропии [9–11]. Однако в статьях авторов не приводится аналитического исследования свойств предложенной ими функции, что оказывается особенно важным при решении задачи различения двух гипотез.

Нужно отметить, что существует несколько классических способов решения задачи обнаружения. Первый из них основан на решении задачи оптимальной фильтрации и требует знания свойств сигнала: периодичности, полосы частот и др. [12]. Второй способ опирается на лемму Неймана–Пирсона, решает задачу различения двух гипотез, определяет факт превышения оптимального порога при заданной вероятности ложной тревоги и требует оценки статистических свойств выборочных распределений шума и смеси сигнала и шума [13]. Третий способ эквивалентен решению задачи о разладке при изменении неизвестных статистических характеристик распределений шума и смеси сигнала и шума. Схожую постановку имеет задача поиска аномалий [14]. Все эти способы демонстрируют качественную и надежную работу при превышении сигнала над шумом, но при малых отношениях сигнал/шум или в тяжелой помеховой обстановке зачастую дают неправильный ответ.

Данная статья посвящена проблеме обнаружения полезного сигнала в сигнально-шумовой смеси и сочетает в себе все три ранее перечисленных способа решения задачи обнаружения. Предлагается использовать вариант леммы Неймана–Пирсона для задачи различения двух гипотез [13], который справедлив, в том числе когда вероятность ошибки различения близка к единице и зависит от полной вариации меры двух распределений нулевой и первой гипотез. На основании аналитического выражения функции этой ошибки формализуется критерий задачи обнаружения сигнала как один из вариантов статистической сложности [15], который учитывает детерминированную природу сигнала, примешанного к шуму. Особенностью статистической сложности является то, что она мультипликативна и состоит из двух множителей, один из которых обращается в нуль на детерминированных синусоидальных сигналах одной частоты, а в физике — на объектах заданной структуры, например кристаллах [9], а другой обращается в нуль на равномерных функциях распределений [10], соответствующих, например, белому шуму. Затем введенный критерий сравнивается с уже известными двумя вариантами статистической сложности на основе квадрата евклидова расстояния и дивергенции Йенсена–Шеннона, устанавливаются их свойства и проводится оптимизация как функции многих переменных на множестве дискретных распределе-

ний. В результате удастся выделить семейства оптимальных распределений и вычислить максимумы функций статистических сложностей.

Статья имеет следующую структуру. В разделе 1 приведен литературный обзор и освещено актуальное состояние исследований по теме статьи. Раздел 2 посвящен связи рассматриваемых в дальнейшем информационных критериев с классическим критерием задачи обнаружения сигналов. В разделе 3 исследуются свойства трех типов статистической сложности. В разделе 4 аналитические результаты предыдущего раздела подкрепляются численным моделированием на синтезированных сигналах. Раздел 5 объединяет полученные в статье результаты и перечисляет планы на будущее.

2. Лемма Неймана–Пирсона и статистическая сложность

Задача обнаружения сигнала $s(n)$ традиционно сводится к задаче различения двух гипотез

$$\begin{cases} \Gamma_0 : x(n) = w(n), \\ \Gamma_1 : x(n) = s(n) + w(n), \quad n = 1, \dots, N. \end{cases}$$

Гипотеза Γ_0 связана с принятием решения о приеме только шума, а гипотеза Γ_1 — о приеме смеси полезного сигнала и шума, где последовательности $\{x(n)\}$, $n = 1, \dots, N$ — временной ряд из принятых данных, $\{s(n)\}$ — полезный сигнал, $\{w(n)\}$ — аддитивный случайный шум, N — длина временного ряда данных.

Случайные величины временного ряда $(x(1), \dots, x(n), \dots, x(N))$ принимают значения $(x_1, \dots, x_n, \dots, x_N) \in \mathbb{R}^N$.

Для получения аналитического выражения для оценки вероятности ошибки различения гипотез можно применить одну из разновидностей леммы Неймана–Пирсона [13, 16].

Лемма 1 [Неймана–Пирсона]. Пусть имеется произвольная, называемая решающим правилом или тестом, измеримая функция многих переменных $(x_1, \dots, x_N) \in \mathbb{R}^N$, такая что

$$d(x_1, \dots, x_N) = \begin{cases} 1, & \text{верна гипотеза } \Gamma_0, \\ 0, & \text{верна гипотеза } \Gamma_1, \end{cases}$$

по которой находятся

$$\begin{aligned} \alpha(d) &= \text{Вероятность (принять } \Gamma_0 \mid \text{верна } \Gamma_1), \\ \beta(d) &= \text{Вероятность (принять } \Gamma_1 \mid \text{верна } \Gamma_0). \end{aligned}$$

Тогда решающее правило d^ оптимально, если*

$$(1) \quad \alpha(d^*) + \beta(d^*) = \inf_d [\alpha(d) + \beta(d)] = \mathcal{E}r(N; \Gamma_0, \Gamma_1) - \text{функция ошибок,}$$

где инфимум берется по всевозможным тестам.

Здесь $\alpha(\cdot)$ – вероятность ложной тревоги, а $\beta(\cdot)$ – вероятность пропуска полезного сигнала.

Для функции ошибок справедлива точная формула следующего вида:

$$(2) \quad \mathcal{E}r(N; \Gamma_0, \Gamma_1) = 1 - \frac{1}{2} \|P_0^{(N)} - P_1^{(N)}\| = 1 - TV(P_0, P_1),$$

где $P_0^{(N)}$ – многомерная функция распределения статистики наблюдений по гипотезе Γ_0 , $P_1^{(N)}$ – многомерная функция распределения статистики наблюдений по гипотезе Γ_1 , а $TV(P_0, P_1)$ – полная вариация меры со знаком, $\|Q\| = 2 \sup_A |Q(A)|$. Таким образом, если множества-носители мер P_0, P_1 не пересекаются, то возможно безошибочное различение гипотез. Если же меры $P_0^{(N)}$ и $P_1^{(N)}$ близки, то $\|P_0^{(N)} - P_1^{(N)}\| \approx 0$, и тогда $\mathcal{E}r(N; \Gamma_0, \Gamma_1) \approx 1$.

Для задачи обнаружения детерминированного полезного сигнала, например при малом отношении сигнал/шум, интересен случай $\|P_0^{(N)} - P_1^{(N)}\| = 2TV(P_0, P_1) \approx 0$ и возможность разумной оценки этой величины. Поэтому когда вероятность суммарной ошибки различения двух гипотез близка к единице, появляется возможность использовать аналитическое выражение $TV(P_0, P_1)$ для конструирования критерия в задаче обнаружения полезного сигнала в смеси. Но сначала обратимся к уже известным критериям и установим их свойства.

Чаще всего для удобства математического исследования полезный сигнал и шум моделируются случайными гауссовыми процессами с разными параметрами, в этом случае задача нахождения момента появления сигнала $s(n)$ в принимаемой последовательности отсчетов называется задачей о разладке [13].

Здесь и далее рассматриваются дискретные распределения вероятностей $p = (p_1, \dots, p_i, \dots, p_N)$, которые по определению обладают следующими свойствами:

$$(3) \quad \forall p_i \in [0, 1], \quad \sum_{i=1}^N p_i = 1.$$

Для формализации критериев, которые учитывают не только случайную составляющую сигналов, но и детерминированную, введем понятия функции дисбаланса D и статистической сложности C распределения. Простейшим примером дисбаланса является евклидово расстояние в пространстве дискретных вероятностных распределений [10].

Определение 1. Дисбаланс D_{SQ} имеет смысл дисперсии распределения относительно равномерного, а именно

$$(4) \quad D_{SQ}(p) = \sum_{i=1}^N \left(p_i - \frac{1}{N} \right)^2 = \sum_{i=1}^N p_i^2 - \frac{1}{N}.$$

Определение 2. Статистическая сложность, определяемая через выражение дисбаланса по определению 1, равна

$$(5) \quad C_{SQ}(p) = H(p) \cdot D_{SQ}(p),$$

где

$$(6) \quad H(p) = \frac{1}{\log N} \left(- \sum_i^N p_i \log p_i \right)$$

— нормированная энтропия Шеннона [1].

При вычислении суммы (6) принимается, что $\frac{0}{\log 0} = 0$ по непрерывности, и это предположение справедливо для всех последующих уравнений.

Из определения 1 следует, что дисбаланс вида (4) и сложность вида (5) удобно применять при оценке и сравнении сигналов, имеющих спектральное распределение, близкое к равномерному. В общем случае вместо равномерного распределения $q_i = 1/N$ при $i = 1, \dots, N$ в формулу (4) может входить произвольное дискретное распределение.

Выше введена формула (4), предложенная в [10] для вычисления дисбаланса относительно равномерного распределения, однако в большинстве работ в качестве такой величины используется дивергенция $JSD(p||q)$ Йенсена–Шеннона [17], так что

Определение 3. Дисбаланс Йенсена–Шеннона равен

$$(7) \quad D_{JSD}(p) = JSD(p||q),$$

где $q = (1/N, \dots, 1/N)$ — равномерное распределение.

Определение 4. Статистическая сложность, определяемая через выражение дисбаланса согласно определению 3, выражается как

$$(8) \quad C_{JSD}(p) = H(p) \cdot D_{JSD}(p).$$

Замечание 1. Выше отмечено, что $\sqrt{D_{SQ}}$ является евклидовой метрикой на пространстве дискретных распределений. В то же время $\sqrt{D_{JSD}}$ также является метрикой и пропорционален метрике Фишера.

Поскольку в лемме 1 Неймана–Пирсона получено, что функция ошибок различения двух гипотез зависит от полной вариации $TV(p, q)$, то введем еще одно понятие дисбаланса.

Определение 5. Дисбаланс по полной вариации меры равен

$$(9) \quad D_{TV}(p) = TV^2(p, q),$$

где $q = (1/N, \dots, 1/N)$.

Определение 6. Статистическая сложность, определяемая через выражение дисбаланса согласно определению 5, равна

$$(10) \quad C_{TV}(p) = H(p) \cdot D_{TV}(p).$$

Представленные выше функции информационных дивергенций, определяющие разные варианты функции дисбаланса, могут быть объединены общим понятием *f-дивергенции* [18]:

$$(11) \quad D_f(p||q) = \sum_{x \in \mathbb{R}^N} q(x) f\left(\frac{p(x)}{q(x)}\right).$$

Выбор функции f порождает целое семейство различных дивергенций:

- Дивергенция Кульбака–Лейблера $D_{KL}(p, q)$ получается из (11) выбором $f(x) = x \log(x)$, $x > 0$.
- Дивергенция Йенсена–Шеннона получается из (11) выбором

$$(12) \quad f(x) = x \log \frac{2x}{x+1} + \log \frac{2}{x+1}, \quad x > 0.$$

- Полная вариация получается при $f(x) = \frac{1}{2}|1 - x|$:

$$(13) \quad TV(p, q) = \frac{1}{2} \sum_{x \in \mathbb{R}^N} |p(x) - q(x)|;$$

TV также является метрикой на пространстве распределений вероятностей. Полная вариация связана с дивергенцией Йенсена–Шеннона следующим соотношением:

$$(14) \quad JSD(p||q) \leq TV(p, q).$$

Из соотношения (14) следует, что дивергенция Йенсена–Шеннона оценивается сверху полной вариацией.

Далее исследуем возможность применения каждого варианта статистической сложности как критерия для индикации появления сигнала, но сначала установим их свойства.

3. Оптимизация статистической сложности

3.1. Оптимизация C_{SQ}

Поставим задачу максимизации функции статистической сложности на множестве дискретных распределений $p = (p_1, \dots, p_N)$

$$(15) \quad C_{SQ}(p) = \frac{1}{\log N} \left(- \sum_{i=1}^N p_i \log p_i \right) \cdot \left(\sum_{i=1}^N \left(p_i - \frac{1}{N} \right)^2 \right) \rightarrow \max_p$$

при условии нормировки

$$(16) \quad \sum_{i=1}^N p_i = 1.$$

Для формулировки леммы о максимальном значении статистической сложности понадобится вспомогательный результат.

Лемма 2. Пусть $0 < x \leq y \leq z \leq 1$, тогда $f(x, y, z) = x^y y^{-x} z^x x^{-z} y^z z^{-y} \geq 1$, причем равенство возможно только при либо $x = y$, либо $y = z$.

Доказательство. Введем новую функцию $g(x, y, z) = \ln f(x, y, z)$,

$$g(x, y, z) = y \ln x - x \ln y + x \ln z - z \ln x + z \ln y - y \ln z.$$

Тогда требуется доказать, что $g(x, y, z) \geq 0$ при $0 < x \leq y \leq z \leq 1$.

По теореме Куна–Таккера решение задачи условной оптимизации функции трех переменных либо находится во внутренней точке ограничения, либо на его границе. Необходимые условия безусловного экстремума функции $g(x, y, z)$ принимают вид

$$(17) \quad \begin{aligned} \frac{\partial g}{\partial x} &= \ln z - \ln y + \frac{y - z}{x} = 0, \\ \frac{\partial g}{\partial y} &= \ln x - \ln z + \frac{z - x}{y} = 0, \\ \frac{\partial g}{\partial z} &= \ln y - \ln x + \frac{x - y}{z} = 0. \end{aligned}$$

Сложим все уравнения последней системы и в результате получим уравнение

$$\frac{y - z}{x} + \frac{z - x}{y} + \frac{x - y}{z} = 0,$$

которое можно переписать в виде

$$\frac{(y - z)(x - y)(z - x)}{xyz} = 0.$$

Это означает, что при выполнении одного из равенств либо $x = y$, либо $y = z$ функция $g(x, y, z)$, возможно, имеет минимум. Пусть $x = y$, тогда третье уравнение из (17) выполняется тождественно, а первое и второе уравнения идентичны и могут быть записаны в виде

$$\ln \eta = \eta - 1, \quad \eta = \frac{z}{y}.$$

Последнее уравнение имеет только один корень $\eta = 1$, т.е. $y = z$.

Вычислим вторые производные и запишем матрицу Гессе:

$$(18) \quad G(x, y, z) = \begin{pmatrix} \frac{z - y}{x^2} & \frac{1}{x} - \frac{1}{y} & \frac{1}{z} - \frac{1}{x} \\ \frac{1}{x} - \frac{1}{y} & \frac{x - z}{y^2} & \frac{1}{y} - \frac{1}{z} \\ \frac{1}{z} - \frac{1}{x} & \frac{1}{y} - \frac{1}{z} & \frac{y - x}{z^2} \end{pmatrix}.$$

Миноры матрицы Гессе равны

$$\begin{aligned}
 M_1(x, y, z) &= \frac{z - y}{x^2}, \\
 (19) \quad M_2(x, y, z) &= -\frac{(x - y)^2 + (z - x)^2 + (y - z)^2}{2x^2y^2}, \\
 M_3(x, y, z) &= 0.
 \end{aligned}$$

Матрица Гессе не является знакоопределенной, более того, ее детерминант равен нулю. Поэтому рассмотрим малую окрестность экстремальной точки.

В малой окрестности $x = y = z$ при условии, что $\delta x \leq \delta y \leq \delta z$, вариация δg функции $g(x, y, z)$ записывается в виде

$$\begin{aligned}
 \delta g &= (x + \delta y) \ln(x + \delta x) - (x + \delta x) \ln(x + \delta y) + (x + \delta x) \ln(x + \delta z) - \\
 &\quad - (x + \delta z) \ln(x + \delta x) + (x + \delta z) \ln(x + \delta y) - (x + \delta y) \ln(x + \delta z) = \\
 &= (\delta z - \delta x)(\delta y - \delta x)(\delta z - \delta y) + o(((\delta x)^2 + (\delta y)^2 + (\delta z)^2)^{3/2}) \geq 0,
 \end{aligned}$$

где ненулевыми оказываются члены при кубах вариаций независимых переменных, а сама вариация функции $g(x, y, z)$ в силу условий леммы оказывается положительной. В случае, когда, например, $\delta y = \delta x$, имеем $g(x, y, z) \equiv 0$, а $f(x, y, z) \equiv 1$. Поэтому экстремум функции является ее нестрогим минимумом. ■

Лемма 3. Максимум статистической сложности (15) достигается на распределении вида

$$(20) \quad \begin{cases} p_i = \frac{1 - p_{\max}}{N - 1}, & i = \overline{1, N} \setminus k, \\ p_k = p_{\max}, \end{cases}$$

где $p_{\max} = \text{const}$, т.е. при появлении отдельной дискреты в произвольном разряде k над равномерным распределением.

Доказательство. Без потери общности положим $k = N$. Из уравнения (16) выразим одну переменную p_N из набора p_i через все остальные:

$$(21) \quad p_N = 1 - \sum_{i=1}^{N-1} p_i.$$

Перепишем уравнение (15) в виде

$$\begin{aligned}
 (22) \quad C_{SQ} &= -\frac{1}{\log N} \left(\sum_{i=1}^{N-1} p_i \log p_i + p_N \log p_N \right) \times \\
 &\quad \times \left(\sum_{i=1}^{N-1} \left(p_i - \frac{1}{N} \right)^2 + \left(p_N - \frac{1}{N} \right)^2 \right).
 \end{aligned}$$

Необходимым условием экстремума функции во внутренней точке области (симплекса 3) является равенство нулю всех частных производных по p_i :

$$(23) \quad \frac{\partial C_{SQ}}{\partial p_i} = 0, \quad i = 1, \dots, N-1.$$

Подстановка функции (22) в (23) дает (с учетом того, что $\frac{\partial p_N}{\partial p_i} = -1$):

$$(24) \quad \begin{aligned} \frac{\partial C_{SQ}}{\partial p_i} = & -\frac{1}{\log N} (\log p_i - \log p_N) \cdot \left(\sum_{i=1}^{N-1} \left(p_i - \frac{1}{N} \right)^2 + \left(p_N - \frac{1}{N} \right)^2 \right) - \\ & - \frac{2}{\log N} \left(\sum_{i=1}^{N-1} p_i \log p_i + p_N \log p_N \right) \cdot (p_i - p_N) = 0, \quad i = 1, \dots, N-1. \end{aligned}$$

В более удобном виде уравнения имеют вид

$$(25) \quad \frac{\partial C_{SQ}}{\partial p_i} = \frac{1}{\log N} (-\log p_i + \log p_N) \cdot D + 2H \cdot (p_i - p_N) = 0, \\ i = 1, \dots, N-1.$$

Запишем разность любых двух уравнений из системы выше для индексов i и j :

$$(26) \quad \frac{\partial C_{SQ}}{\partial p_i} - \frac{\partial C_{SQ}}{\partial p_j} = \frac{1}{\log N} (-\log p_i + \log p_j) \cdot D + 2H \cdot (p_i - p_j) = 0.$$

Учитывая, что величины D и H положительны, из уравнений (25) и (26) можно составить следующие равенства при условии, что рассматриваемые вероятности p_j , $j = 1, \dots, N-1$ не равны p_N :

$$(27) \quad \frac{\log p_i - \log p_j}{\log p_N - \log p_j} - \frac{p_i - p_j}{p_N - p_j} = 0,$$

$$(28) \quad (p_N - p_j) \log p_i + (p_i - p_N) \log p_j + (p_j - p_i) \log p_N = 0,$$

$$(29) \quad p_i^{p_N - p_j} \cdot p_j^{p_i - p_N} \cdot p_N^{p_j - p_i} = 1.$$

Применяя лемму 2, делаем вывод, что последнее уравнение может быть выполнено при равенстве $p_i = p_j$.

Таким образом, получено, что каждая из вероятностей p_i может принимать одно из двух различных значений, которые задают распределение вида

$$(30) \quad \begin{cases} p_i = \frac{1 - p_{\max}}{K}, & \forall i = 1, \dots, K, \\ p_i = p_N = \frac{p_{\max}}{N - K}, & \forall i = K + 1, \dots, N. \end{cases}$$

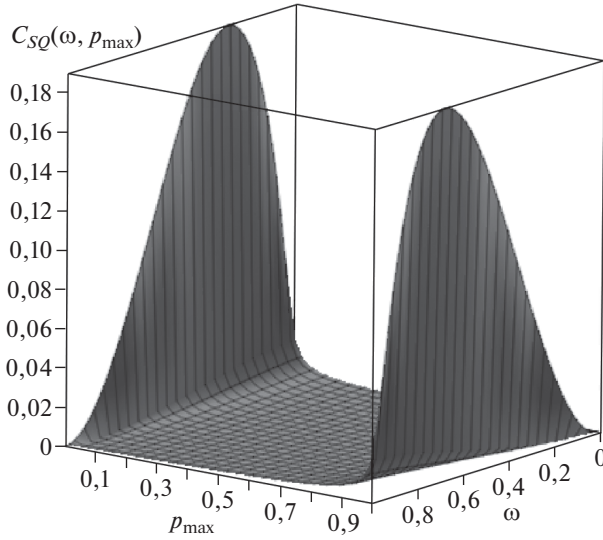


Рис. 1. Поверхности уровня статистической сложности $C_{SQ}(\omega, p_{\max})$.

Теперь нужно показать, что максимуму сложности соответствуют значения $K = 1$ и $K = N - 1$. Для этого вычислим значение дисбаланса (4) на распределении (30), которое обозначим через $D^{(K)}(\omega, p_{\max})$:

$$(31) \quad D^{(K)}(\omega, p_{\max}) = \frac{1}{N} \frac{(p_{\max} + \omega - 1)^2}{\omega(1 - \omega)}, \quad \omega = \frac{K}{N}.$$

В свою очередь, энтропия равна

$$(32) \quad H^{(K)}(\omega, p_{\max}) = 1 - \frac{1}{\log N} \left((1 - p_{\max}) \log \frac{1 - p_{\max}}{\omega} + p_{\max} \log \frac{p_{\max}}{1 - \omega} \right).$$

Максимум $C_{SQ}(\omega, p_{\max})$ при $N \leq 100$ был исследован численно, и он достигался при $K = 1$. Из выражения для $D^{(K)}(\omega, p_{\max})$ (31) видно, что при $N \geq 101$ и при изменении с $K = 1$ на $K = 2$ или с $K = N - 1$ на $K = N - 2$ его значение меняется почти в два раза, в то время как энтропия (32) меняется незначительно. Таким образом, распределение вероятностей (30), доставляющее функции сложности экстремальное значение при $K = 1$ или $K = N - 1$, имеет вид (20). ■

Для наглядности на рис. 1 показан график $C_{SQ} = C_{SQ}(\omega, p_{\max})$ при $N = 1024$, на котором ω меняется непрерывно (хотя K меняется дискретно).

Следствие 1. Подставим значения p_i и $p_N = p_{\max}$ из (20) в (22) и рассмотрим сложность C_{SQ} как функцию от $p_{\max}$. При достаточно больших значениях N она примет вид

$$C_{SQ} \approx (1 - p_{\max}) \cdot p_{\max}^2.$$

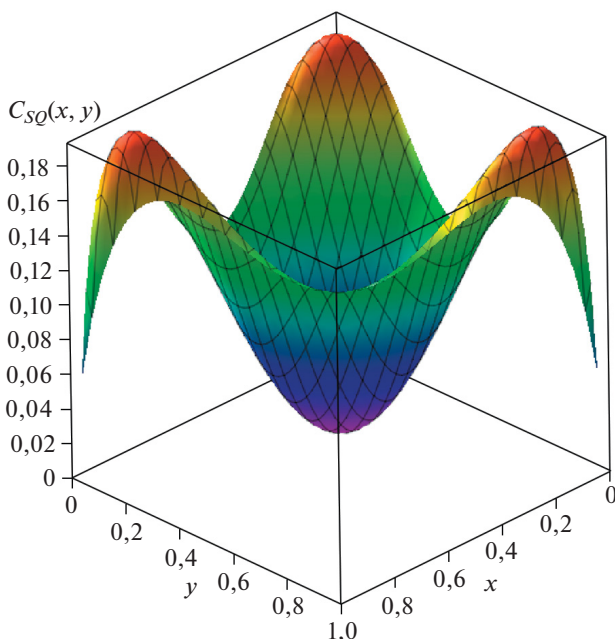


Рис. 2. Поверхности уровня статистической сложности $C_{SQ}(x, y)$ для $p = \{p_1 = x, p_2 = y, p_3 = 1 - x - y\}$.

Откуда следует, что эта функция принимает максимальное значение $C_{SQ}^* \approx 4/27$ при $p_{\max} = 2/3$.

Следствие 2. Минимальное значение $C_{SQ} = 0$ достигается на равномерном распределении $p = (1/N, \dots, 1/N)$.

Действие леммы 3 для случая, когда дискретное распределение $p = \{p_1, p_2, p_3\}$ состоит только из трех отсчетов, продемонстрировано на рис. 2. Сложность зависит от двух переменных, поскольку одна из вероятностей может быть выражена через другие. Здесь C_{SQ} имеет три одинаковых ярко выраженных максимума и три одинаковых локальных минимума, относящихся к ситуациям $p_1 = p_2$, $p_2 = p_3$, $p_1 = p_3$ при выполнении необходимых условий экстремума, а также глобальный минимум при $p_1 = p_2 = p_3$.

В табл. 1 показано изменение оптимальных параметров $C_{SQ}(w, p_{\max})$ при росте N .

Таблица 1. Оптимальные параметры $C_{SQ}(\omega, p_{\max})$ при различных значениях N

N	$C_{SQ}(\omega^*, p_{\max}^*)$	$p_{\max}^*$	ω^*	$N - K^*$
3	0,1932	0,8315	0,6666	1
256	0,1994	0,7044	0,9960	1
512	0,1942	0,7008	0,9980	1
1024	0,1898	0,6979	0,9990	1
2048	0,1861	0,6955	0,9995	1

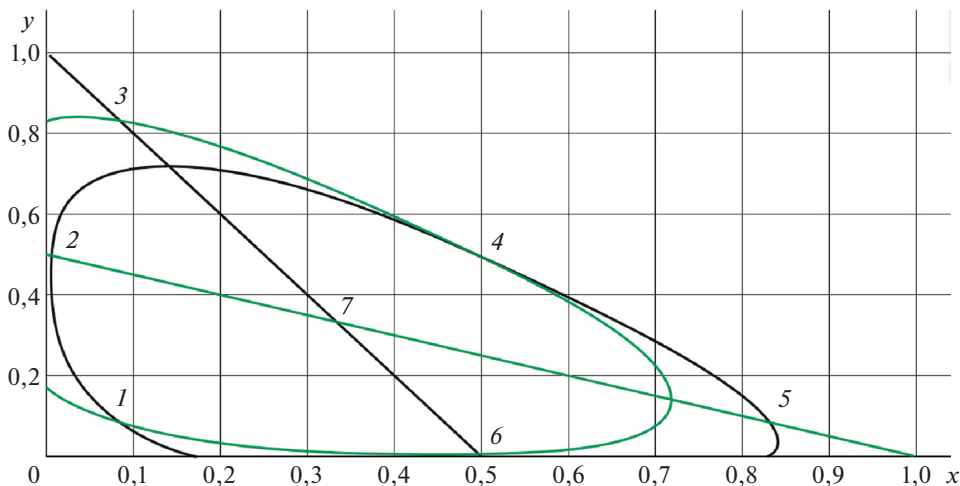


Рис. 3. Кривые, соответствующие необходимым условиям экстремума C_{SQ} для $p = \{p_1 = x, p_2 = y, p_3 = 1 - x - y\}$.

Необходимые условия экстремума C_{SQ} для дискретного распределения $p = \{p_1 = x, p_2 = y, p_3 = 1 - x - y\}$ выписываются согласно (25) следующим образом:

$$(33) \quad \begin{cases} (-\log x + \log(1 - x - y)) \left(\left(x - \frac{1}{3}\right)^2 + \left(y - \frac{1}{3}\right)^2 + \left(1 - x - y - \frac{1}{3}\right)^2 \right) - \\ - 2(x \log x + y \log y + (1 - x - y) \log(1 - x - y))(-1 + y) = 0, \\ (-\log y + \log(1 - x - y)) \left(\left(x - \frac{1}{3}\right)^2 + \left(y - \frac{1}{3}\right)^2 + \left(1 - x - y - \frac{1}{3}\right)^2 \right) - \\ - 2(x \log x + y \log y + (1 - x - y) \log(1 - x - y))(-1 + x) = 0. \end{cases}$$

Неявные уравнения системы (33) описывают кривые, представленные на рис. 3.

Черная и зеленые кривые отвечают первому и второму неявным уравнениям системы (33) соответственно. На рис. 3 отмечены семь точек экстремума, для которых подсчитано значение статистической сложности. Все полученные данные приведены в табл. 2.

Таблица 2. Экстремальные точки статистической сложности при $N = 3$

p	1	2	3	4	5	6	7
p_1	0,08425	0,006	0,08425	0,497	0,8315	0,497	0, (3)
p_2	0,08425	0,497	0,8315	0,497	0,08425	0,006	0, (3)
p_3	0,8315	0,497	0,08425	0,006	0,08425	0,497	0, (3)
C_{SQ}	0,1932	0,1062	0,1932	0,1062	0,1932	0,1062	0

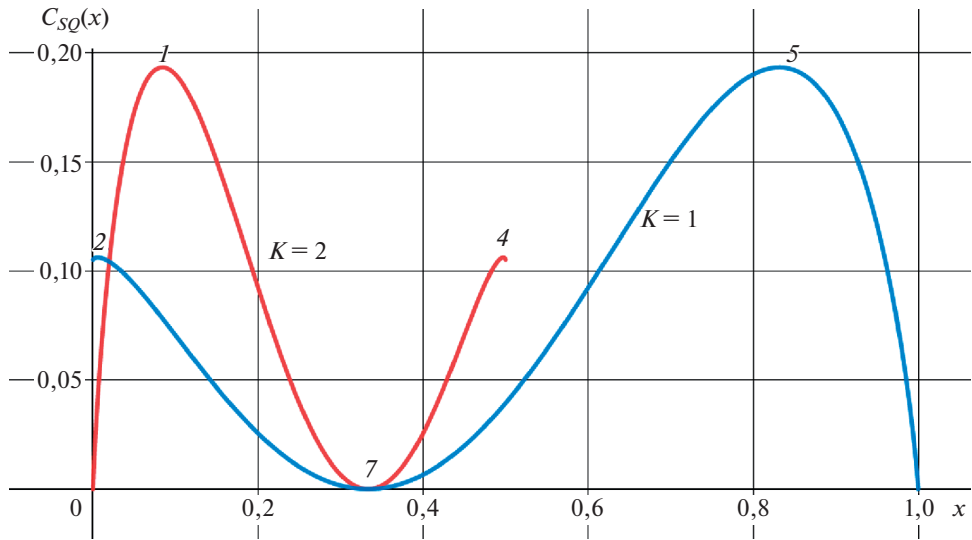


Рис. 4. Статистическая сложность C_{SQ} для $p = \{p_1 = x, p_2 = x, p_3 = 1 - 2x\}$.

Первая, третья и пятая точки максимума соответствуют одному и тому же значению максимума функции. Вторая, четвертая и шестая точки минимума соответствуют одному и тому же значению минимума функции. Стоит отметить, что экстремальные точки отвечают случаю $K = N - 1 = 2$ за исключением глобального минимума, где все вероятности равны между собой, и тем самым описывают три локальных минимума, один глобальный минимум и три равных между собой максимума статистической сложности на рис. 2. Отдельно можно построить статистическую сложность в случае $p = \{p_1 = x, p_2 = x, p_3 = 1 - 2x\}$.

На рис. 4 отмечены точки экстремума в соответствии с табл. 2, которые охватывают все случаи $p = \{p_1 = x, p_2 = y, p_3 = 1 - x - y\}$.

3.2. Оптимизация C_{JSD}

Применим аналогичный подход к статистической сложности, использующей дивергенцию Йенсена–Шеннона в качестве дисбаланса. Сложность в этом случае выражается как

$$(34) \quad C_{JSD}(p) = H(p) \cdot JSD(p||q), \quad q_j = 1/N, \quad j = 1, \dots, N.$$

Раскрыв $JSD(p||q)$, получаем

$$(35) \quad C_{JSD}(p) = H(p) \cdot \left(H(m) - \frac{1}{2}(H(p) + H(q)) \right) \cdot \log N, \quad m = \frac{p+q}{2}.$$

Лемму, подобную лемме 3, доказать не удастся, однако можно выписать необходимые условия экстремума для статистической сложности вида (35).

Следуя той же логике, что и в лемме 3, перепишем последнее уравнение в переменных p_i и p_N :

$$(36) \quad C_{JSD}(p) = - \left(\sum_{i=1}^{N-1} p_i \log p_i + p_N \log p_N \right) \times \\ \times \left(H(m) - \frac{1}{2} \left(-\frac{1}{\log N} \left(\sum_{i=1}^{N-1} p_i \log p_i + p_N \log p_N \right) + 1 \right) \right),$$

где

$$(37) \quad H(m) = -\frac{1}{\log N} \left(\sum_{i=1}^{N-1} \frac{p_i + \frac{1}{N}}{2} \log \frac{p_i + \frac{1}{N}}{2} + \frac{p_N + \frac{1}{N}}{2} \log \frac{p_N + \frac{1}{N}}{2} \right).$$

Тогда с учетом (21)

$$(38) \quad \frac{\partial H(m)}{\partial p_i} = -\frac{1}{\log N} \left(\frac{1}{2} \log \frac{p_i + \frac{1}{N}}{2} - \frac{1}{2} \log \frac{p_N + \frac{1}{N}}{2} \right), \quad i = 1, \dots, N-1.$$

Объединяя все частные производные, получаем необходимые условия экстремума в виде

$$(39) \quad \frac{\partial C_{JSD}(p)}{\partial p_i} = \frac{H(p)}{2} \cdot \left(- \left(\log \frac{p_i + \frac{1}{N}}{2} - \log \frac{p_N + \frac{1}{N}}{2} \right) + (\log p_i - \log p_N) \right) - \\ - (\log p_i - \log p_N) \cdot JSD(p||q) = 0, \quad i = 1, \dots, N-1.$$

После упрощения остаются уравнения

$$(40) \quad H(p) \cdot \left(\log \frac{p_i + \frac{1}{N}}{2} - \log \frac{p_N + \frac{1}{N}}{2} \right) + \\ + (\log p_i - \log p_N) \cdot (2JSD(p||q) - H(p)) = 0, \quad i = 1, \dots, N-1.$$

Тогда разность уравнений (40) для индексов i, j будет иметь вид

$$(41) \quad (\log p_i - \log p_j) \cdot (2JSD(p||q) - H(p)) + \\ + H(p) \cdot \left(\log \frac{p_i + \frac{1}{N}}{2} - \log \frac{p_j + \frac{1}{N}}{2} \right) = 0.$$

Замечание 2. Из вида системы уравнений (41) следует, что система выполняется, если $p_i = p_j$, что является одним из необходимых условий экстремума функции (35). Вследствие нелинейности системы, состоящей из уравнений (41), у нее могут быть и другие корни.

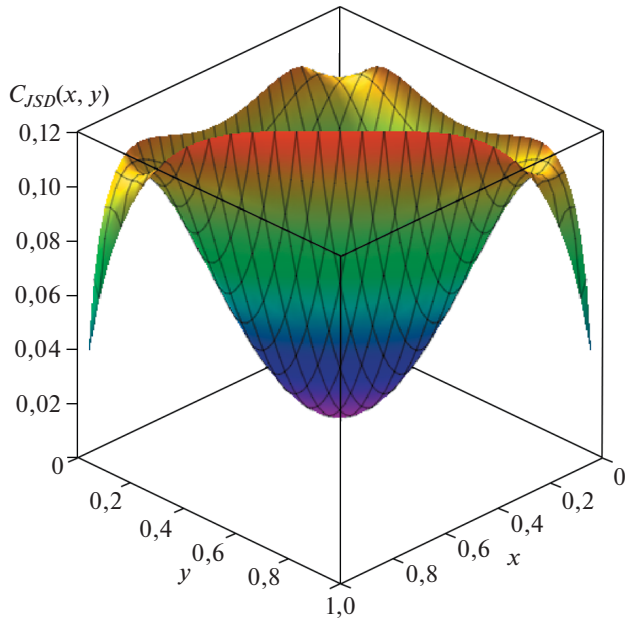


Рис. 5. Поверхности уровня статистической сложности C_{JSD} для $p = \{p_1 = x, p_2 = y, p_3 = 1 - x - y\}$.

Для иллюстрации этого факта на рис. 5 приведен график поверхности уровня статистической сложности вида (35), когда дискретное распределение $p = \{p_1, p_2, p_3\}$ состоит только из трех отсчетов.

На рис. 5 видно, что в точках, удовлетворяющих $p_1 = p_2$, $p_2 = p_3$ и $p_1 = p_3$, при выполнении необходимых условий экстремума находятся седловые точки поверхности.

Ранее было установлено, что распределение (30) доставляет экстремум C_{SQ} при $K = N - 1$. Далее будет показано, что оно же доставляет экстремум сложности, основанной на полной вариации меры $TV(p, q)$. Поэтому предлагается найти максимум C_{JSD} на этом распределении и провести сравнение полученных оптимальных параметров распределения при фиксированных N . Выпишем сложность в явном виде и получим

$$(42) \quad C_{JSD}^{(K)} = H^{(K)} \cdot \left(H^{(K)}(m) - \frac{1}{2}(H^{(K)} + 1) \right) \cdot \log N,$$

где $H^{(K)}$ соответствует (32), а $H^{(K)}(m)$ дается следующей формулой:

$$(43) \quad H^{(K)}(m) = 1 - \frac{1}{\log N} \left(\frac{(1 - p_{\max} + \omega)}{2} \log \frac{(1 - p_{\max} + \omega)}{2\omega} + \frac{(1 + p_{\max} - \omega)}{2} \log \frac{(1 + p_{\max} - \omega)}{2 - 2\omega} \right).$$

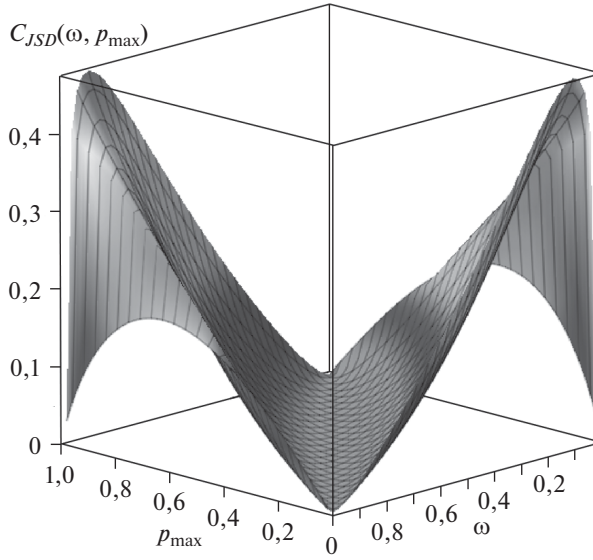


Рис. 6. Поверхности уровня статистической сложности $C_{JSD}(\omega, p_{\max})$.

В табл. 3 показано изменение оптимальных параметров C_{JSD} при росте N .

Для наглядности на рис. 6 показан график $C_{JSD} = C_{JSD}(\omega, p_{\max})$ при $N = 1024$, на котором ω меняется непрерывно (хотя K меняется дискретно).

Результаты, приведенные в табл. 3, демонстрируют, что для выбранного класса распределений (30) оптимальным является набор из N дискрет, где K равны между собой, а остальные нулевые. Стоит обратить внимание, что на полученном распределении C_{JSD} не равна нулю, в том числе благодаря слагаемому $H^{(K)}(m)$, которое отвечает уже “сдвинутому” распределению, состоящему из K элементов, равных $\frac{1}{K} + \frac{1}{N}$, и $N - K$ отсчетов по $\frac{1}{2N}$ каждый.

Таблица 3. Оптимальные параметры $C_{JSD}(\omega, p_{\max})$ при различных значениях N

N	$C_{JSD}(\omega^*, p_{\max}^*)$	$p_{\max}^*$	ω^*	$N - K^*$
3	0,1266	1	0,4083	1 или 2
256	0,4482	1	0,8703	33
512	0,4790	1	0,8897	56
1024	0,5065	1	0,9051	97
2048	0,5312	1	0,9171	170

3.3. Оптимизация C_{TV}

Приступим к анализу статистической сложности на основе полной вариации

$$(44) \quad C_{TV}(p) = -\frac{1}{4 \log N} \left(\sum_{i=1}^N p_i \log p_i \right) \cdot \left(\sum_{i=1}^N \left| p_i - \frac{1}{N} \right| \right)^2.$$

Предложение 1. Согласно выражению для функции ошибок (2) из леммы 1 Неймана–Пирсона и определению (44) в качестве критерия предлагается использовать C_{TV} для решения задачи различения двух гипотез и индикации появления детерминированной составляющей полезного сигнала в шуме.

Справедлива следующая лемма.

Лемма 4. Максимум статистической сложности (44) достигается на семействе распределений (30).

Доказательство. С учетом симметричности функции (44) и симплекса (3) без ограничения общности найдется целое $K \in \{1, \dots, N-1\}$, при котором максимум в данной лемме достигается на части симплекса (3), определяемой ограничениями $p_i \leq 1/N$ для $i = 1, \dots, K$ и $p_i \geq 1/N$ для $i = K+1, \dots, N$. Перепишем уравнение (44) в виде

$$(45) \quad C_{TV} = -\frac{1}{4 \log N} \left(\sum_{i=1}^{N-1} p_i \log p_i + p_N \log p_N \right) \times \\ \times \left(\sum_{i=1}^K \left(-p_i + \frac{1}{N} \right) + \sum_{i=K+1}^N \left(p_i - \frac{1}{N} \right) \right)^2.$$

Тогда для $i = 1, \dots, K$ необходимые условия экстремума принимают вид

$$(46) \quad \frac{\partial C_{TV}}{\partial p_i} = -\frac{1}{\log N} (\log p_i - \log p_N) \cdot D_{TV} - 2H(p) \sqrt{D_{TV}} = 0, \\ i = 1, \dots, K,$$

а для $i = K+1, \dots, N$ справедливо

$$(47) \quad \frac{\partial C_{TV}}{\partial p_i} = -\frac{1}{\log N} (\log p_i - \log p_N) \cdot D_{TV} = 0, \quad i = K+1, \dots, N.$$

Составим разность двух уравнений из (46) для индексов i и j . Откуда следует, что если $D_{TV} \neq 0$, то $p_i = p_j$ при $i = 1, \dots, K$. В то время как из (47) следует, что $p_i = p_N$ при $i = K+1, \dots, N$. Снова получили, что семейство распределений (30) доставляет максимум функции сложности, теперь уже C_{TV} . ■

Далее требуется определить оптимальные значения K и $p_{\max}$. Для этого вычислим значение дисбаланса $D_{TV}^{(K)}$ на распределении (30):

$$(48) \quad D_{TV}^{(K)} = (p_{\max} + \omega - 1)^2, \quad \omega = \frac{K}{N}.$$

В свою очередь, энтропия равна

$$(49) \quad H^{(K)} = 1 - \frac{1}{\log N} \left((1 - p_{\max}) \log \frac{1 - p_{\max}}{\omega} + p_{\max} \log \frac{p_{\max}}{1 - \omega} \right).$$

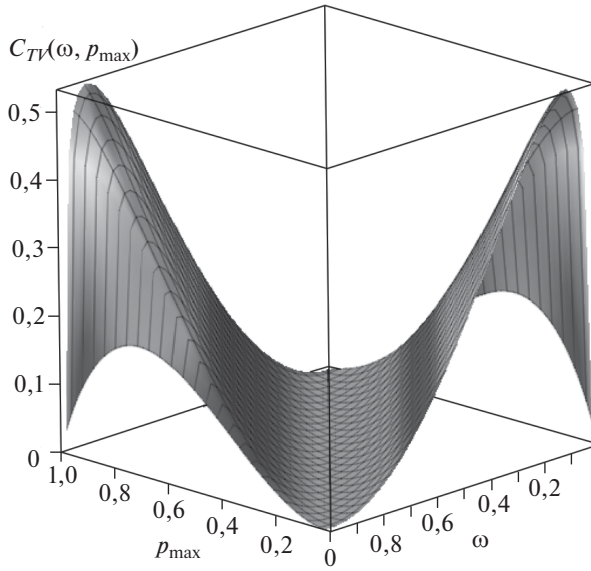


Рис. 7. Поверхности уровня статистической сложности $C_{TV}(\omega, p_{\max})$.

Для наглядности на рис. 7 представлен график $C_{TV} = C_{TV}(\omega, p_{\max})$ при $N = 1024$, на котором ω меняется непрерывно (хотя K меняется дискретно).

Составим необходимые условия (46) и (47) экстремума статистической сложности C_{TV} , выписанной через переменные $p_{\max}$, ω .

$$(50) \quad \begin{cases} f_1^N(p_{\max}, \omega) := 2(p_{\max} + \omega - 1) \left(\left(1 - \frac{(1-p_{\max}) \log \frac{1-p_{\max}}{\omega} + p_{\max} \log \frac{p_{\max}}{1-\omega}}{\log N} \right) - \right. \\ \left. - \frac{(p_{\max} + \omega - 1)}{2 \log N} \left(\log \frac{p_{\max}}{1-\omega} - \log \frac{1-p_{\max}}{\omega} \right) \right) = 0, \\ f_2^N(p_{\max}, \omega) := 2(p_{\max} + \omega - 1) \left(\left(1 - \frac{(1-p_{\max}) \log \frac{1-p_{\max}}{\omega} + p_{\max} \log \frac{p_{\max}}{1-\omega}}{\log N} \right) - \right. \\ \left. - \frac{(p_{\max} + \omega - 1)}{2 \log N} \left(\frac{p_{\max}}{1-\omega} - \frac{1-p_{\max}}{\omega} \right) \right) = 0. \end{cases}$$

Пересечения кривых, соответствующих неявным уравнениям (50), отвечают точкам экстремума C_{TV} . Составим разность двух необходимых условий экстремума

$$(51) \quad \begin{aligned} f_3^N(p_{\max}, \omega) &:= f_1^N(p_{\max}, \omega) - f_2^N(p_{\max}, \omega) = \\ &= \frac{(p_{\max} + \omega - 1)^2}{\log N} \left(-\log \frac{p_{\max}}{1-\omega} + \log \frac{1-p_{\max}}{\omega} + \frac{p_{\max}}{1-\omega} - \frac{1-p_{\max}}{\omega} \right) = 0. \end{aligned}$$

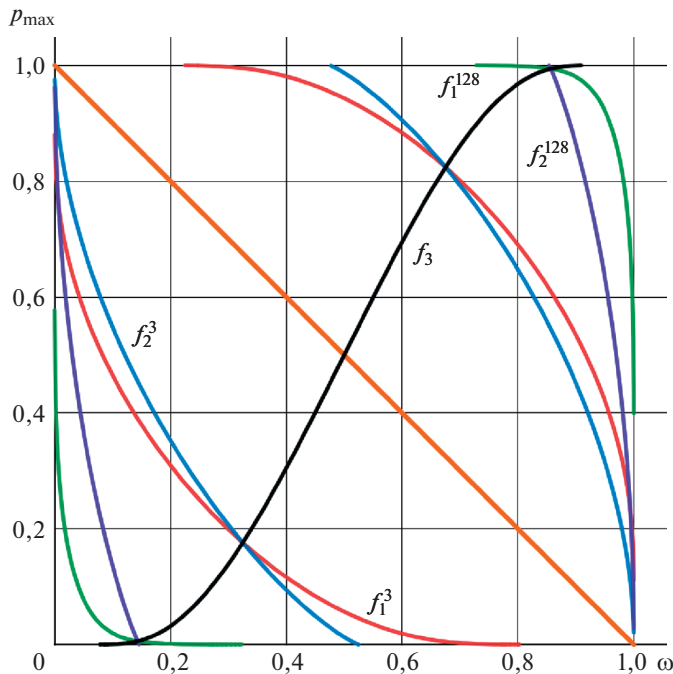


Рис. 8. Кривые необходимых условий экстремума C_{TV} для $N = 3$ и $N = 128$. Уравнения $f_1^N(p_{\max}, \omega)$ и $f_2^N(p_{\max}, \omega)$ имеют общее решение $(p_{\max} + \omega - 1 = 0)$ (оранжевая кривая).

Построим неявные кривые уравнений $f_1^N(p_{\max}, \omega)$, $f_2^N(p_{\max}, \omega)$, $f_3^N(p_{\max}, \omega)$ для некоторых значений N . Для удобства индекс N у $f_3^N(p_{\max}, \omega)$ может быть опущен, так как неявная кривая уравнения (51) не зависит от N .

Согласно рис. 8 статистическая сложность имеет, помимо точек минимума $(p_{\max} + \omega - 1 = 0)$, где $C_{TV} = 0$, еще и две точки максимума для каждого значения N : $(p_{\max}^*, \omega^*)$ и $(1 - p_{\max}^*, 1 - \omega^*)$, которые лежат на кривой $f_3(p_{\max}, \omega)$.

В табл. 4 приведены оптимальные значения параметров формул (48) и (49), доставляющие максимум статистической сложности C_{TV} .

Таблица 4. Оптимальные параметры $C_{TV}(\omega, p_{\max})$ при различных значениях N

N	$C_{TV}(\omega^*, p_{\max}^*)$	$p_{\max}^*$	ω^*	$N - K^*$
3	0,1289	0,8241	0,6751	1 или 2
256	0,4789	0,9976	0,8752	32
512	0,5120	0,9991	0,8901	56
1024	0,5410	0,9997	0,9022	100
2048	0,5667	0,9999	0,9122	180

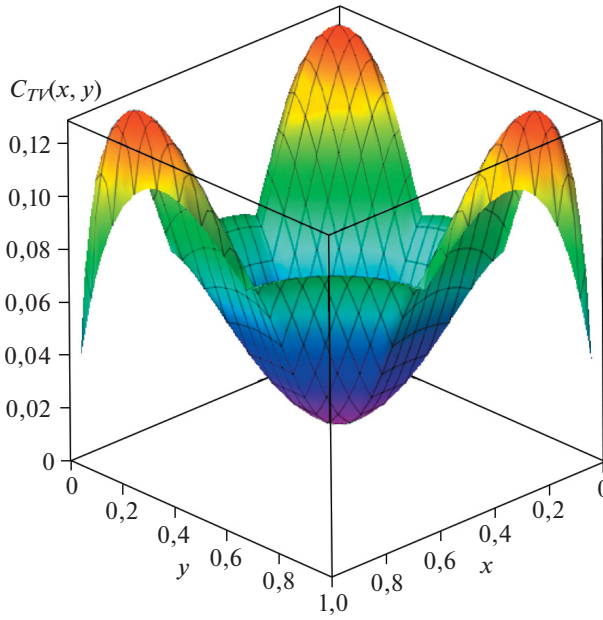


Рис. 9. Поверхности уровня статистической сложности $C_{TV}(x, y)$ для $p = \{p_1 = x, p_2 = y, p_3 = 1 - x - y\}$.

Дополнительно случай $N = 3$ показан на рис. 9, где приведен график поверхности уровня C_{TV} , когда дискретное распределение $p = \{p_1, p_2, p_3\}$ состоит из трех отсчетов.

4. Моделирование и сравнение статистических сложностей

Проведем анализ оптимальных параметров, доставляющих максимум различным типам статистической сложности, и сравним значения в табл. 1, 3 и 4. Основной интерес представляют максимальные значения сложностей и оптимальные значения K . Максимальные значения $C_{TV}(\omega^*, p_{\max}^*) \in [0, 1]$, $C_{JSD}(\omega^*, p_{\max}^*) \in [0, 1]$ близки друг другу и растут с увеличением N . Оптимальные значения K для этих двух видов сложностей также близки.

Для демонстрации аналитических результатов, полученных в предыдущих разделах, показано применение трех вариантов статистической сложности в задаче индикации полезного сигнала в шумовой смеси для синтезированных сигналов. Применен алгоритм из [15], основанный на вычислении дискретных распределений p по спектральному представлению временных рядов.

Синтезированный сигнал длительностью 10 секунд представляет собой сумму конечного числа косинусоидальных колебаний в смеси с белым шумом:

$$(52) \quad x(t) = I(t) \sum_{i=1}^K A_i \cos(2\pi f_i t + \Delta\phi_i) + w(t), \quad t \in [0, 10],$$

где A_i , f_i , $\Delta\phi_i$ – амплитуды, частоты и случайные фазы гармонических колебаний соответственно, $w(t)$ – белый шум, а $I(t)$ – индикаторная функция наличия полезного сигнала в сигнально-шумовой смеси.

$I(t)$ выбрана таким образом, что гармонические сигналы присутствуют в середине итоговой последовательности $x(t)$.

$$(53) \quad I(t) = \begin{cases} 0, & t \in [0, 3), \\ 1, & t \in [3, 7], \\ 0, & t \in (7, 10]. \end{cases}$$

Работа алгоритма осуществляется следующим образом:

1. Синтезированный с частотой дискретизации f_s сигнал делится на короткие окна, содержащие по $N = 2048$ отсчетов.

2. Далее вычисляется спектр для каждого окна с использованием алгоритма БПФ.

3. На основе спектра путем его нормирования вычисляются дискретные плотности p_i , $i = 1, \dots, N$.

4. Для полученного набора p_i вычисляются информационные характеристики $C_{SQ}(p)$, $C_{JSD}(p)$, $C_{TV}(p)$.

5. Полученная последовательность значений информационных характеристик выводится вместе с сигналом по временной оси.

Стоит отметить, что параметры f_s и N подобраны так, чтобы исключить эффект растекания спектра, т.е. получить в спектре четкие дискреты, соответствующие K гармоническим функциям из формулы (52). Отношения амплитуд сигнала и дисперсии шума выбрано близким к единице.

Порог γ для решающего правила предлагается выбрать как 25% от максимального значения критерия для выбранного N из табл. 1, 3, 4:

$$(54) \quad \begin{aligned} \gamma_{CQ} &= 0,25 \cdot 0,1861 = 0,0465; \\ \gamma_{JSD} &= 0,25 \cdot 0,5312 = 0,1328; \\ \gamma_{TV} &= 0,25 \cdot 0,5667 = 0,1417. \end{aligned}$$

Удобство выбора такого порога заключается в том, что он не зависит от конкретной реализации шума и основан на аналитически полученных максимальных значениях функций статистической сложности.

На всех графиках синим цветом обозначена амплитуда исходного сигнала, а красным — статистическая сложность, которая рассчитана по алгоритму, описанному выше. Горизонтальная ось отвечает за время в секундах, а вертикальные — за величину амплитуды сигнала (слева) и критерия (справа). Черной пунктирной линией очерчено значение выбранного порога γ .

В первом эксперименте количество синусоидальных сигналов и соответственно дискрет в спектре равно $K = 3$ при $N = 2048$. На рис. 10 показаны зависимости статистических сложностей от времени для синтезированного сигнала.

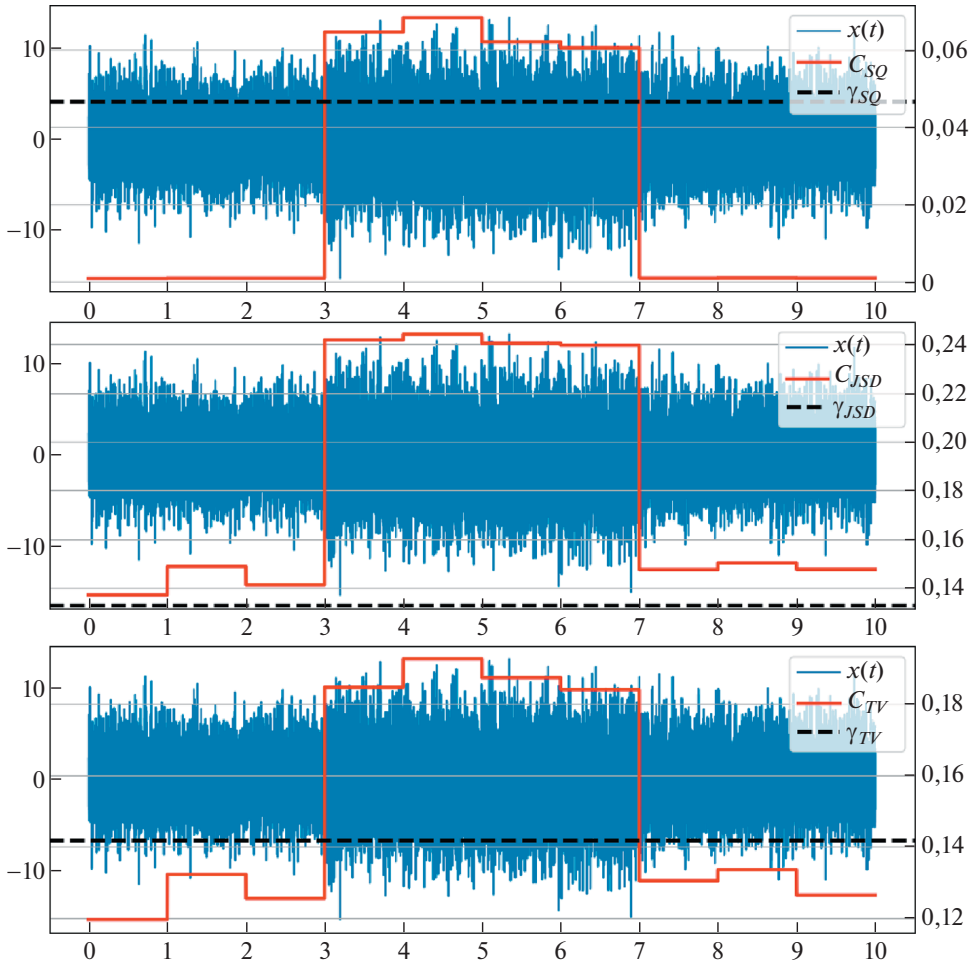


Рис. 10. Три дискреты, $K = 3$.

Как видно, значения C_{SQ} и C_{TV} превышают выбранный порог на интервале присутствия сигнала, что позволяет с уверенностью сделать вывод о его появлении. Что касается C_{JSD} , то априорный выбор порога оказался неудачен в силу того, что истинное значение ее максимума неизвестно, как показано в подразделе 3.2. Если изменить значение порога в большую сторону на 20%, то детектирование на основе C_{JSD} будет таким же успешным, как и на основе C_{TV} .

Во втором эксперименте количество дискрет в спектре $K = 30$. В этом случае C_{SQ} перестает показывать удовлетворительный результат в смысле превышения выбранного порога, поскольку функция C_{SQ} сильно деградирует с ростом K , однако все еще позволяет провести индикацию, как видно на рис. 11. Функция сложности C_{TV} все еще уверенно превышает порог, как и в первом эксперименте, а C_{JSD} превышает порог на всем сигнале, как и в первом эксперименте.

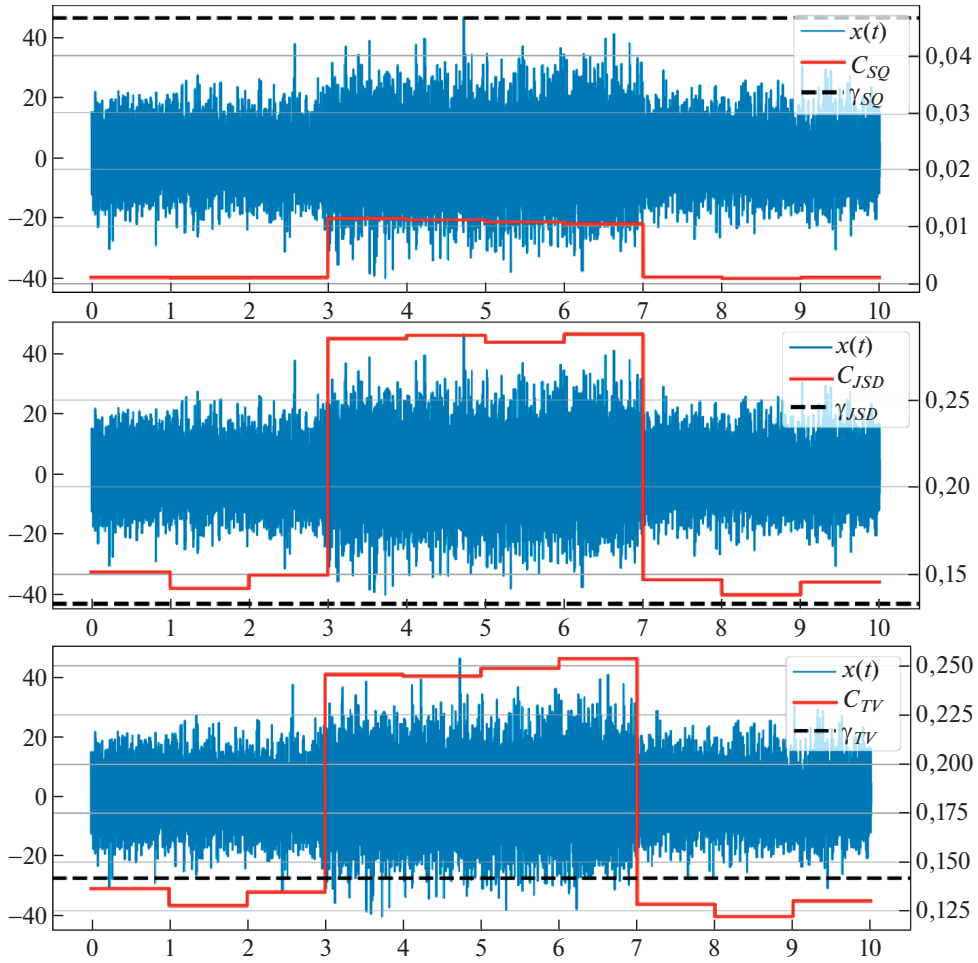


Рис. 11. Тридцать дискрет, $K = 30$.

Таким образом, можно сделать вывод о том, что C_{TV} наиболее удобна в практическом смысле, поскольку хорошо работает на сигналах с большим количеством дискрет в спектре и позволяет принимать решение о появлении полезного сигнала, основываясь на довольно простом правиле, связанном с выбором порога исходя из теоретически максимального значения для статистической сложности.

5. Заключение

В работе дано теоретическое обоснование использования статистической сложности в качестве критерия для решения задачи различения двух гипотез в условиях, когда вероятность ошибки их различения близка к единице. Рассмотрены три варианта вычисления статистической сложности для разных функций дисбаланса. Введены новые понятия дисбаланса и статисти-

стической сложности на основе полной вариации меры двух распределений. Проведено сравнение информационных критериев и найден класс дискретных распределений, на которых достигается максимум различных типов статистической сложности. Найдены значения максимумов при фиксированных значениях количества отсчетов распределения. Показано, что статистическая сложность C_{TV} на основе полной вариации непосредственно связана с задачей различения двух гипотез, а статистическая сложность C_{JSD} на основе энтропии Йенсена–Шеннона дает близкую оценку C_{TV} на выборочных распределениях. В свою очередь C_{SQ} наиболее перспективна для обнаружения отдельной дискретности над равномерным распределением. Предложен способ выбора порога для решающего правила обнаружения полезного сигнала, учитывающий максимальные полученные значения критериев, и показана результативность такого подхода на синтезированных сигналах.

Будущая работа будет посвящена исследованию информационных критериев на основе двух- и многомерных распределений, а также рассмотрению типовых акустических сигналов с реалистичным фоновым шумом.

СПИСОК ЛИТЕРАТУРЫ

1. *Shannon C.E.* A Mathematical Theory of Communication // Bell Syst. Tech. J. 1948. V. 27. P. 379–423.
2. *Gray R.M.* Entropy and Information Theory. New York: Springer, 2011.
<https://doi.org/10.1007/978-1-4419-7970-4>
3. *Holub A., Perona P., Burl M.C.* Entropy-based Active Learning for Object Recognition // Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE. 2008. P. 1–8.
<https://doi.org/10.1109/CVPRW.2008.4563068>
4. *Osisanwo F.Y., Akinsola J.E.T., Awodele O. et al.* Supervised Machine Learning Algorithms: Classification and Comparison // Int. J. Comput. Trends Technol. (IJCTT). 2017. V. 48. No. 3. P. 128–138.
<https://doi.org/10.14445/22312803/IJCTT-V48P126>
5. *Shen J., Hung J., Lee L.* Robust Entropy-based Endpoint Detection for Speech Recognition in Noisy Environments // Proc. 5th International Conference on Spoken Language Processing (ICSLP). 1998. <https://doi.org/10.21437/icslp.1998-527>
6. *Ribeiro M., Henriques T., Castro L., Souto A., Antunes L., Costa-Santos C., Teixeira A.* The Entropy Universe // Entropy. 2021. V. 222. No. 2. art. 222.
<https://doi.org/10.3390/e23020222>
7. *Ramirez J., Segura J.C., Benitez C. et al.* A New Kullback-Leibler VAD for Speech Recognition in Noise // IEEE Signal Proc. Lett. 2004. V. 11. No. 2. P. 266–269.
<https://doi.org/10.1109/LSP.2003.821762>
8. *Horie T., Burioka N., Amisaki T., Shimizu E.* Sample Entropy in Electrocardiogram During Atrial Fibrillation // Yonago Acta Medica. 2018. V. 61. No. 1. P. 49–57.
<https://doi.org/10.33160/yam.2018.03.007>
9. *Lamberti P.W., Martin M.T., Plastino A., Rosso O.A.* Intensive Entropic Non-Triviality Measure // Phys. A: Stat. Mech. Appl. 2004. V. 334. No. 1. P. 119–131.
<https://doi.org/10.1016/j.physa.2003.11.005>

10. *Lopez-Ruiz R.* Shannon Information, LMC Complexity and Renyi Entropies: A Straightforward Approach // *Biophys. Chem.* 2005. V. 115. No. 3. P. 215–218. <https://doi.org/10.1016/j.bpc.2004.12.035>
11. *Zunino L., Soriano M.C., Rosso O.A.* Distinguishing Chaotic and Stochastic Dynamics from Time Series by Using a Multiscale Symbolic Approach // *Phys. Rev. E. Stat. Nonlin. Soft. Matter Phys.* 2012. V. 86. No. 4. P. 1–5. <https://doi.org/10.1103/PhysRevE.86.046210>
12. *Ronald L.A., Duncan W.M.* Signal Analysis: Time, Frequency, Scale, and Structure. N.J.: IEEE Press, 2004.
13. *Ширяев А.Н.* Вероятностно-статистические методы в теории принятия решений. М.: МЦНМО: НМУ, 2020.
14. *Kishan G.M., Chilukuri K.M., HuaMing Huang.* Anomaly Detection Principles and Algorithms. Cham: Springer. 2017. <https://doi.org/10.1007/978-3-319-67526-8>
15. *Berlin L.M., Galyaev A.A., Lysenko P.V.* Comparison of Information Criteria for Detection of Useful Signals in Noisy Environments // *Sensors.* 2023. V. 23. No. 4. art. 2133. <https://doi.org/10.3390/s23042133>
16. *Johnson P., Moriarty J., Peskir G.* Detecting Changes in Real-Time Data: A User's Guide to Optimal Detection // *Philos. Trans. Royal Soc. A.* 2017. V. 375. P. 16. art. 2100. <https://doi.org/10.1098/rsta.2016.0298>
17. *Li Z., Li Y., Zhang K.A.* Feature Extraction Method of Ship-Radiated Noise Based on Fluctuation-Based Dispersion Entropy and Intrinsic Time-Scale Decomposition // *Entropy.* 2019. V. 21. No. 7. art. 693. <https://doi.org/10.3390/e21070693>
18. *Sason I.* On f-Divergences: Integral Representations, Local Behavior, and Inequalities // *Entropy.* 2018. V. 20. No. 5. art. 383. <https://doi.org/10.3390/e20050383>

Статья представлена к публикации членом редколлегии А.В. Назиным.

Поступила в редакцию 21.03.2023

После доработки 19.04.2023

Принята к публикации 11.05.2023