

ЧЕБЫШЕВСКИЙ СБОРНИК

Том 25. Выпуск 4.

УДК 517.956

DOI 10.22405/2226-8383-2024-25-4-27-41

Об информационном профиле трех случайных величин с двумя исходами

А. Аллеманд

Аллеманд Аллан Олегович — аспирант, Московский государственный университет им. М. В. Ломоносова (Москва).
e-mail: allan.allemand@gmail.com

Аннотация

В этой статье мы рассматриваем взаимную информацию для пары случайных величин и находим третью переменную (условие), которая максимизирует условную взаимную информацию трех из них.

Ключевые слова: взаимная информация, условная взаимная информация, энтропия, информационное неравенство.

Библиография: 5 названий.

Для цитирования:

Аллеманд, А. Об информационном профиле трех случайных величин с двумя исходами // Чебышевский сборник, 2024, т. 25, вып. 4, с. 27–41.

CHEBYSHEVSKII SBORNIK

Vol. 25. No. 4.

UDC 517.956

DOI 10.22405/2226-8383-2024-25-4-27-41

On the information profile of three random variables with two outcomes

A. Allemand

Allemand Allan Olegovich — postgraduate student, Lomonosov Moscow State University (Moscow).

e-mail: allan.allemand@gmail.com

Abstract

In this paper we consider mutual information for a pair of random variables and find a third variable (condition) that maximise conditional mutual information of three of them.

Keywords: mutual information, conditional mutual information, entropy, information inequality.

Bibliography: 5 titles.

For citation:

Allemand, A. 2024, “On the information profile of three random variables with two outcomes”, *Chebyshevskii sbornik*, vol. 25, no. 4, pp. 27–41.

1. Введение

Пусть заданы две дискретные случайные величины A и B вместе с их совместным распределением AB . Их совместную информацию можно определить как

$$I(A, B) = H(A) + H(B) - H(A, B)$$

где H - энтропия, которая, как известно, задается формулой

$$H(X) = - \sum_i p_i \log p_i$$

где p_i - распределение X , логарифм берется натуральный. При этом мы $p \log p$ по непрерывности доопределяем в точке $p = 0$ нулевым значением. Если же работать с условными распределениями, то можно определить условную совместную информацию двух случайных величин A и B при условии третьей случайной величины C :

$$I(A, B|C) = H(A, C) + H(B, C) - H(A, B, C) - H(C)$$

В силу простых неравенств на энтропию получается, что такая информация всегда неотрицательна. Этот минимум достигается, например, при $C = A$. Встает вопрос о нахождении максимума данной функции. Есть несколько методов решения подобных экстремальных задач, в том числе основанных на том, что $I(A, B|C)$ есть по сути разность двух вогнутых функций в силу вогнутости функции энтропии. Тем не менее, все эти методы являются численными, а также не гарантируют нахождения глобального максимума информации. В настоящей работе аналитическим путем решений уравнений на производную для внутренностей сначала всего параллелепипеда возможных значений, а потом - граней и ребер, находится максимальное значение (достигаемое в одной из вершин параллелепипеда) при конкретном распределении AB , где для удобства A, B, C принимают по два значения, но данный метод применим и для любых других распределений.

Нахождение максимального значения интересно тем, что $I(A, B)$ есть максимальная длина случайного секретного ключа, который могут совместно выработать владельцы случайных величин A и B , общаясь по открытому каналу. Аналогично максимальное значение $I(A, B|C)$ равно максимальной длине случайного секретного ключа, который могут совместно выработать владельцы случайных величин A и B , если дополнительно Центр, знающий и A , и B опубликует по открытому каналу некоторую третью случайную величину.

2. Внутренность параллелепипеда

Распишем условную совместную информацию через попарные совместные распределения:

$$\begin{aligned} I(A, B|C) &= H(A, C) + H(B, C) - H(A, B, C) - H(C) \\ &= - \sum_{i,k} p_{ik} \log(p_{ik}) - \sum_{j,k} p_{jk} \log(p_{jk}) + \sum_{i,j,k} p_{ijk} \log(p_{ijk}) + \sum_k p_k \log(p_k) \end{aligned}$$

Здесь у значений A индекс i , у значений B индекс j , у значений C индекс k , то есть

$$p_{ijk} = P(A = A_i, B = B_j, C = C_k)$$

В этой работе мы будем использовать следующее распределение:

$$AB = \begin{pmatrix} 0.2 & 0.2 \\ 0.0 & 0.6 \end{pmatrix}$$

Обозначим $x = p_{111}, y = p_{121}, z = p_{221}$.

Ясно, что $p_{112} = 0.2 - x, p_{122} = 0.2 - y, p_{222} = 0.6 - z, p_{211} = 0, p_{212} = 0$.

Далее будем работать, используя эти три независимые переменные. Выразим через них все совместные распределения:

$$AC = \begin{pmatrix} x + y & 0.4 - x - y \\ z & 0.6 - z \end{pmatrix}$$

$$BC = \begin{pmatrix} x & 0.2 - x \\ y + z & 0.8 - y - z \end{pmatrix}$$

$$ABC = \begin{pmatrix} x & 0.2 - x \\ y & 0.2 - y \\ z & 0.6 - z \end{pmatrix}$$

$$C = (x + y + z \quad 1 - x - y - z)$$

Для нахождения максимума будем действовать следующим образом: сначала рассмотрим функцию для точек, находящихся во внутренности параллелепипеда. Возьмем производные по всем трем переменным и приравняем их к нулю. Решив уравнения, мы поймем, что точек экстремума во внутренности нет. Повторим данную операцию для внутренностей граней и ребер параллелепипеда, чем докажем, что максимум достигается в одной из его вершин. Нам в таком случае останется только посчитать значения во всех восьми вершинах и выбрать среди них наибольшее.

Итак, используя полученные значения, рассмотрим систему из приравненных к нулю производных по всем трем переменным

$$\begin{cases} I'_x = -\log(x + y) + \log(0.4 - x - y) - \log(x) + \log(0.2 - x) \\ \quad + \log(x) - \log(0.2 - x) + \log(x + y + z) - \log(1 - x - y - z) = 0 \\ I'_y = -\log(x + y) + \log(0.4 - x - y) - \log(y + z) + \log(0.8 - y - z) \\ \quad + \log(y) - \log(0.2 - y) + \log(x + y + z) - \log(1 - x - y - z) = 0 \\ I'_z = -\log(z) + \log(0.6 - z) - \log(y + z) + \log(0.8 - y - z) \\ \quad + \log(z) - \log(0.6 - z) + \log(x + y + z) - \log(1 - x - y - z) = 0 \end{cases}$$

Сократим одинаковые слагаемые и возьмем от полученного экспоненту:

$$\begin{cases} (x + y)(1 - x - y - z) = (x + y + z)(0.4 - x - y) \\ (0.4 - x - y)(0.8 - y - z)(x + y + z)y = (x + y)(y + z)(0.2 - y)(1 - x - y - z) \\ (y + z)(1 - x - y - z) = (x + y + z)(0.8 - y - z) \end{cases}$$

С помощью первого и третьего уравнений можно выразить y, z через x :

$$\begin{cases} (x + y) - (x + y)(x + y + z) = (x + y + z)(0.4 - x - y) \\ (y + z) - (y + z)(x + y + z) = (x + y + z)(0.8 - y - z) \end{cases}$$

$$\begin{cases} x + y = 0.4(x + y + z) \\ y + z = 0.8(x + y + z) \end{cases}$$

$$2(x + y) = y + z$$

Из чего, путем подстановки в одно из исходных уравнений, получаем

$$\begin{cases} y = x \\ z = 3x \end{cases}$$

По теореме о производной сложной функции нескольких переменных, производная нашей функции по x при подстановке вместо y и z выражений выше будет суммой полученных ранее производных с коэффициентами 1, 1, 3:

$$\begin{aligned} I'_x = & -\log(2x) + \log(0.4 - 2x) - \log(x) + \log(0.2 - x) \\ & + \log(x) - \log(0.2 - x) + \log(5x) - \log(1 - 5x) \\ & - \log(2x) + \log(0.4 - 2x) - \log(4x) + \log(0.8 - 4x) \\ & + \log(x) - \log(0.2 - x) + \log(5x) - \log(1 - 5x) \\ & - 3\log(3x) + 3\log(0.6 - 3x) - 3\log(4x) + 3\log(0.8 - 4x) \\ & + 3\log(3x) - 3\log(0.6 - 3x) + 3\log(5x) - 3\log(1 - 5x) = 0 \end{aligned}$$

Сокращаем одинаковые слагаемые:

$$\begin{aligned} I'_x = & -2\log(2x) + 2\log(0.4 - 2x) \\ & + \log(x) - \log(0.2 - x) - 4\log(4x) + 4\log(0.8 - 4x) \\ & + 5\log(5x) - 5\log(1 - 5x) = 0 \end{aligned}$$

Берем экспоненту:

$$\begin{aligned} (2x)^2(0.2 - x)(4x)^4(1 - 5x)^5 = \\ x(0.4 - 2x)^2(0.8 - 4x)^4(5x)^5 \end{aligned}$$

Упрощаем:

$$1024x^6(0.2 - x)(1 - 5x)^5 = 3125x^6(0.4 - 2x)^2(0.8 - 4x)^4$$

Выносим коэффициенты, чтобы получить одинаковые скобки с обеих сторон:

$$1024x^6(1 - 5x)^6 = 5 * 0.4^2 * 0.8^4 * 3125x^6(1 - 5x)^6$$

Перемножаем числа:

$$1024x^6(1 - 5x)^6 = 1024x^6(1 - 5x)^6$$

Получили тождество. Таким образом, производная всегда равна нулю, функция здесь константа и ее значение всюду равно значению в точках $(0, 0, 0)$ и $(0.2, 0.2, 0.6)$, которые мы разберем в конце.

3. Грани

Теперь будем доказывать, что экстремумов не найдется во внутренности ни одной из 6 граней параллелепипеда. Для этого будем фиксировать два крайних значения по каждой переменной и приравнять производные по остальным переменным к нулю:

3.1. z=0

$$AC = \begin{pmatrix} x+y & 0.4-x-y \\ 0 & 0.6 \end{pmatrix}$$

$$BC = \begin{pmatrix} x & 0.2-x \\ y & 0.8-y \end{pmatrix}$$

$$ABC = \begin{pmatrix} x & 0.2-x \\ y & 0.2-y \\ 0 & 0.6 \end{pmatrix}$$

$$C = \begin{pmatrix} x+y & 1-x-y \end{pmatrix}$$

В данном случае, уравнение на производную по x совсем не имеет решений, поэтому достаточно рассмотреть только его:

$$\begin{aligned} I'_x &= -\log(x+y) + \log(0.4-x-y) - \log(x) + \log(0.2-x) \\ &\quad + \log(x) - \log(0.2-x) + \log(x+y) - \log(1-x-y) = 0 \end{aligned}$$

после сокращения одинаковых слагаемых и взятие экспоненты имеем уравнение

$$0.4 - x - y = 1 - x - y$$

которое не имеет решений.

3.2. z=0.6

$$AC = \begin{pmatrix} x+y & 0.4-x-y \\ 0.6 & 0 \end{pmatrix}$$

$$BC = \begin{pmatrix} x & 0.2-x \\ y+0.6 & 0.2-y \end{pmatrix}$$

$$ABC = \begin{pmatrix} x & 0.2-x \\ y & 0.2-y \\ 0.6 & 0 \end{pmatrix}$$

$$C = \begin{pmatrix} x+y+0.6 & 0.4-x-y \end{pmatrix}$$

В данном случае, уравнение на производную по x также совсем не имеет решений, поэтому достаточно рассмотреть только его:

$$\begin{aligned} I'_x &= -\log(x+y) + \log(0.4-x-y) - \log(x) + \log(0.2-x) \\ &\quad + \log(x) - \log(0.2-x) + \log(x+y+0.6) - \log(0.4-x-y) = 0 \end{aligned}$$

после сокращения одинаковых слагаемых и взятие экспоненты имеем уравнение

$$x+y = x+y+0.6$$

3.3. $y=0$

$$AC = \begin{pmatrix} x & 0.4 - x \\ z & 0.6 - z \end{pmatrix}$$

$$BC = \begin{pmatrix} x & 0.2 - x \\ z & 0.8 - z \end{pmatrix}$$

$$ABC = \begin{pmatrix} x & 0.2 - x \\ 0 & 0.2 \\ z & 0.6 - z \end{pmatrix}$$

$$C = (x + z \quad 1 - x - z)$$

Система уравнений на производную:

$$\begin{cases} I'_x = -\log(x) + \log(0.4 - x) - \log(x) + \log(0.2 - x) \\ + \log(x) - \log(0.2 - x) + \log(x + z) - \log(1 - x - z) = 0 \\ I'_z = -\log(z) + \log(0.6 - z) - \log(z) + \log(0.8 - z) \\ + \log(z) - \log(0.6 - z) + \log(x + z) - \log(1 - x - z) = 0 \end{cases}$$

после сокращения одинаковых слагаемых и взятие экспоненты имеем уравнение

$$\begin{cases} x(1 - x - z) = (0.4 - x)(x + z) & \begin{cases} x - x^2 - xz = 0.4x + 0.4z - x^2 - xz \\ z - xz - z^2 = 0.8x + 0.8z - zx - z^2 \end{cases} \\ z(1 - x - z) = (0.8 - z)(x + z) \end{cases}$$

$$\begin{cases} 0.6x = 0.4z \\ 0.8x = 0.2z \end{cases}$$

Система совместна только в точке $x = z = 0$, которая не лежит во внутреннейности грани.

3.4. $y=0.2$

$$AC = \begin{pmatrix} x + 0.2 & 0.2 - x \\ z & 0.6 - z \end{pmatrix}$$

$$BC = \begin{pmatrix} x & 0.2 - x \\ z + 0.2 & 0.6 - z \end{pmatrix}$$

$$ABC = \begin{pmatrix} x & 0.2 - x \\ 0.2 & 0 \\ z & 0.6 - z \end{pmatrix}$$

$$C = (x + z + 0.2 \quad 0.8 - x - z)$$

Система уравнений на производную:

$$\begin{cases} I'_x = -\log(x + 0.2) + \log(0.2 - x) - \log(x) + \log(0.2 - x) \\ + \log(x) - \log(0.2 - x) + \log(x + z + 0.2) - \log(0.8 - x - z) = 0 \\ I'_z = -\log(z) + \log(0.6 - z) - \log(z + 0.2) + \log(0.6 - z) \\ + \log(z) - \log(0.6 - z) + \log(x + z + 0.2) - \log(0.8 - x - z) = 0 \end{cases}$$

после сокращения одинаковых слагаемых и взятие экспоненты имеем уравнение

$$\begin{cases} (x + 0.2)(0.8 - x - z) = (0.2 - x)(x + z + 0.2) \\ (z + 0.2)(0.8 - x - z) = (0.6 - z)(x + z + 0.2) \end{cases}$$

$$\begin{cases} 0.8x - x^2 - xz + 0.16 - 0.2x - 0.2z = 0.2x + 0.2z + 0.04 - x^2 - xz - 0.2x \\ 0.8z - xz - z^2 + 0.16 - 0.2x - 0.2z = 0.6x + 0.6z + 0.12 - zx - z^2 + 0.2z \end{cases}$$

$$\begin{cases} 0.6x + 0.12 = 0.4z \\ -0.8x + 0.04 = 0.2z \end{cases}$$

$$\begin{cases} 0.6x + 0.12 = 0.4z \\ -1.6x + 0.08 = 0.4z \end{cases}$$

Вычитывая второе уравнение из первого получаем

$$2.2x + 0.04 = 0$$

Что не имеет решений, так как x не может быть отрицательным.

3.5. $x=0$

$$AC = \begin{pmatrix} y & 0.4 - y \\ z & 0.6 - z \end{pmatrix}$$

$$BC = \begin{pmatrix} 0 & 0.2 \\ y + z & 0.8 - y - z \end{pmatrix}$$

$$ABC = \begin{pmatrix} 0 & 0.2 \\ y & 0.2 - y \\ z & 0.6 - z \end{pmatrix}$$

$$C = \begin{pmatrix} y + z & 1 - y - z \end{pmatrix}$$

При данном распределении уравнение на производную по z не будет иметь решений, чем и можно будет ограничиться:

$$\begin{aligned} I'_z &= -\log(z) + \log(0.6 - z) - \log(y + z) + \log(0.8 - y - z) \\ &\quad + \log(z) - \log(0.6 - z) + \log(y + z) - \log(1 - y - z) \end{aligned}$$

после сокращения и взятие экспоненты получаем уравнение

$$0.8 - y - z = 1 - y - z$$

которое не имеет решений.

3.6. $x=0.2$

$$AC = \begin{pmatrix} y + 0.2 & 0.2 - y \\ z & 0.6 - z \end{pmatrix}$$

$$BC = \begin{pmatrix} 0.2 & 0 \\ y + z & 0.8 - y - z \end{pmatrix}$$

$$ABC = \begin{pmatrix} 0.2 & 0 \\ y & 0.2 - y \\ z & 0.6 - z \end{pmatrix}$$

$$C = (y + z + 0.2 \quad 0.8 - y - z)$$

При данном распределении уравнение на производную по z не будет иметь решений, чем и можно будет ограничиться:

$$\begin{aligned} I'_z &= -\log(z) + \log(0.6 - z) - \log(y + z) + \log(0.8 - y - z) \\ &+ \log(z) - \log(0.6 - z) + \log(y + z + 0.2) - \log(0.8 - y - z) \end{aligned}$$

после сокращения и взятие экспоненты получаем уравнение

$$y + z = 0.2 + y + z$$

которое не имеет решений. Таким образом, экстремума во внутренностях граней параллелепипеда нет.

4. Ребра

В этой части мы проверим, нет ли экстремумов во внутренностях 12 ребер параллелепипеда. Для этого будем фиксировать значения пары переменных в крайних значениях и решать уравнение на приравненную к нулю производную по третьей переменной.

4.1. $x=0, y=0$

$$AC = \begin{pmatrix} 0 & 0.4 \\ z & 0.6 - z \end{pmatrix}$$

$$BC = \begin{pmatrix} 0 & 0.2 \\ z & 0.8 - z \end{pmatrix}$$

$$ABC = \begin{pmatrix} 0 & 0.2 \\ 0 & 0.2 \\ z & 0.6 - z \end{pmatrix}$$

$$C = (z \quad 1 - z)$$

Получаем уравнение

$$\begin{aligned} I'_z &= -\log(z) + \log(0.6 - z) - \log(z) + \log(0.8 - z) \\ &+ \log(z) - \log(0.6 - z) + \log(z) - \log(1 - z) = 0 \end{aligned}$$

Сокращая одинаковые слагаемые и используя экспоненты получаем:

$$0.8 - z = 1 - z$$

которое не имеет решений

4.2. $x=0, y=0.2$

$$AC = \begin{pmatrix} 0.2 & 0.2 \\ z & 0.6 - z \end{pmatrix}$$

$$BC = \begin{pmatrix} 0 & 0.2 \\ z + 0.2 & 0.6 - z \end{pmatrix}$$

$$ABC = \begin{pmatrix} 0 & 0.2 \\ 0.2 & 0.2 \\ z & 0.6 - z \end{pmatrix}$$

$$C = (z + 0.2 \quad 0.8 - z)$$

Получаем уравнение

$$I'_z = -\log(z) + \log(0.6 - z) - \log(z + 0.2) + \log(0.6 - z) + \log(z) - \log(0.6 - z) + \log(z + 0.2) - \log(0.8 - z) = 0$$

Сокращая одинаковые слагаемые и используя экспоненты получаем:

$$0.6 - z = 0.8 - z$$

которое не имеет решений

4.3. $x=0.2, y=0$

$$AC = \begin{pmatrix} 0.2 & 0.2 \\ z & 0.6 - z \end{pmatrix}$$

$$BC = \begin{pmatrix} 0.2 & 0 \\ z & 0.8 - z \end{pmatrix}$$

$$ABC = \begin{pmatrix} 0.2 & 0 \\ 0 & 0.2 \\ z & 0.6 - z \end{pmatrix}$$

$$C = (0.2 + z \quad 0.8 - z)$$

Получаем уравнение

$$I'_z = -\log(z) + \log(0.6 - z) - \log(z) + \log(0.8 - z) + \log(z) - \log(0.6 - z) + \log(z + 0.2) - \log(0.8 - z) = 0$$

Сокращая одинаковые слагаемые и используя экспоненты получаем:

$$0.2 + z = z$$

которое не имеет решений

4.4. $x=0.2, y=0.2$

$$AC = \begin{pmatrix} 0.4 & 0 \\ z & 0.6 - z \end{pmatrix}$$

$$BC = \begin{pmatrix} 0.2 & 0 \\ z + 0.2 & 0.6 - z \end{pmatrix}$$

$$ABC = \begin{pmatrix} 0.2 & 0 \\ 0.2 & 0 \\ z & 0.6 - z \end{pmatrix}$$

$$C = (0.4 + z \quad 0.6 - z)$$

Получаем уравнение

$$\begin{aligned} I'_z &= -\log(z) + \log(0.6 - z) - \log(z + 0.2) + \log(0.6 - z) \\ &+ \log(z) - \log(0.6 - z) + \log(z + 0.4) - \log(0.6 - z) = 0 \end{aligned}$$

Сокращая одинаковые слагаемые и используя экспоненты получаем:

$$z + 0.2 = z + 0.4$$

которое не имеет решений.

4.5. $x=0, z=0$

$$AC = \begin{pmatrix} y & 0.4 - y \\ 0 & 0.6 \end{pmatrix}$$

$$BC = \begin{pmatrix} 0 & 0.2 \\ y & 0.8 - y \end{pmatrix}$$

$$ABC = \begin{pmatrix} 0 & 0.2 \\ y & 0.2 - y \\ 0 & 0.6 \end{pmatrix}$$

$$C = (y \quad 1 - y)$$

Получаем уравнение:

$$\begin{aligned} I'_y &= -\log(y) + \log(0.4 - y) - \log(y) + \log(0.8 - y) \\ &+ \log(y) - \log(0.2 - y) + \log(y) - \log(1 - y) = 0 \end{aligned}$$

$$\begin{aligned} (0.4 - y)(0.8 - y) &= (0.2 - y)(1 - y) \\ 0.32 - 0.4y - 0.8y + y^2 &= 0.2 - 0.2y - y + y^2 \\ 0.32 &= 0.2 \end{aligned}$$

которое не имеет решений

4.6. $x=0, z=0.6$

$$AC = \begin{pmatrix} y & 0.4 - y \\ 0.6 & 0 \end{pmatrix}$$

$$BC = \begin{pmatrix} 0 & 0.2 \\ y + 0.6 & 0.2 - y \end{pmatrix}$$

$$ABC = \begin{pmatrix} 0 & 0.2 \\ y & 0.2 - y \\ 0.6 & 0 \end{pmatrix}$$

$$C = (y + 0.6 \quad 0.4 - y)$$

Уравнение будет:

$$\begin{aligned} I'_y &= -\log(y) + \log(0.4 - y) - \log(y + 0.6) + \log(0.2 - y) \\ &+ \log(y) - \log(0.2 - y) + \log(y + 0.6) - \log(0.4 - y) = 0 \end{aligned}$$

Получаем, что производная тождественно равна нулю. Это означает, что функция на этом отрезке есть константа, значение которой равно значению в вершине $(0, 0, 0.6)$, которую мы рассмотрим в последней части.

4.7. $x=0.2, z=0$

$$AC = \begin{pmatrix} 0.2 + y & 0.2 - y \\ 0 & 0.6 \end{pmatrix}$$

$$BC = \begin{pmatrix} 0 & 0.2 \\ y & 0.8 - y \end{pmatrix}$$

$$ABC = \begin{pmatrix} 0.2 & 0 \\ y & 0.2 - y \\ 0 & 0.6 \end{pmatrix}$$

$$C = (0.2 + y \quad 0.8 - y)$$

Уравнение будет:

$$\begin{aligned} I'_y &= -\log(0.2 + y) + \log(0.2 - y) - \log(y) + \log(0.8 - y) \\ &+ \log(y) - \log(0.2 - y) + \log(0.2 + y) - \log(0.8 - y) = 0 \end{aligned}$$

Получаем, что производная тождественно равна нулю. Это означает, что функция на этом отрезке есть константа, значение которой равно значению в вершине $(0.2, 0, 0)$, которую мы рассмотрим в последней части.

4.8. $x=0.2, z=0.6$

$$AC = \begin{pmatrix} y + 0.2 & 0.2 - y \\ 0.6 & 0 \end{pmatrix}$$

$$BC = \begin{pmatrix} 0.2 & 0 \\ y + 0.6 & 0.2 - y \end{pmatrix}$$

$$ABC = \begin{pmatrix} 0.2 & 0 \\ y & 0.2 - y \\ 0.6 & 0 \end{pmatrix}$$

$$C = (0.8 + y \quad 0.2 - y)$$

Получаем уравнение:

$$\begin{aligned} I'_y &= -\log(y + 0.2) + \log(0.2 - y) - \log(y + 0.6) + \log(0.2 - y) \\ &+ \log(y) - \log(0.2 - y) + \log(y + 0.8) - \log(0.2 - y) = 0 \end{aligned}$$

Сокращаем одинаковые слагаемые и берем экспоненты, получаем:

$$\begin{aligned} (0.2 + y)(y + 0.6) &= y(y + 0.8) \\ 0.2y + 0.12 + y^2 + 0.6y &= y^2 + 0.8y \\ 0.12 &= 0 \end{aligned}$$

которое не имеет решений.

4.9. $y=0, z=0$

$$AC = \begin{pmatrix} x & 0.4 - x \\ 0 & 0.6 \end{pmatrix}$$

$$BC = \begin{pmatrix} x & 0.2 - x \\ 0 & 0.8 \end{pmatrix}$$

$$ABC = \begin{pmatrix} x & 0.2 - x \\ 0 & 0.2 \\ 0 & 0.6 \end{pmatrix}$$

$$C = (x \quad 1 - x)$$

Получаем уравнение:

$$\begin{aligned} I'_x &= -\log(x) + \log(0.4 - x) - \log(x) + \log(0.2 - x) \\ &+ \log(x) - \log(0.2 - x) + \log(x) - \log(1 - x) = 0 \end{aligned}$$

Которое путем сокращений и взятие экспоненты приводится к виду

$$0.4 - x = 1 - x$$

которое не имеет решений.

4.10. $y=0, z=0.6$

$$AC = \begin{pmatrix} x & 0.4 - x \\ 0.6 & 0 \end{pmatrix}$$

$$BC = \begin{pmatrix} x & 0.2 - x \\ 0.6 & 0.2 \end{pmatrix}$$

$$ABC = \begin{pmatrix} x & 0.2 - x \\ 0 & 0.2 \\ 0.6 & 0 \end{pmatrix}$$

$$C = (x + 0.6 \quad 0.4 - x)$$

Получаем уравнение:

$$I'_x = -\log(x) + \log(0.4 - x) - \log(x) + \log(0.2 - x) + \log(x) - \log(0.2 - x) + \log(x + 0.6) - \log(0.4 - x) = 0$$

Которое путем сокращений и взятие экспоненты приводится к виду

$$x = x + 0.6$$

которое не имеет решений.

4.11. $y=0.2, z=0$

$$AC = \begin{pmatrix} x + 0.2 & 0.2 - x \\ 0 & 0.6 \end{pmatrix}$$

$$BC = \begin{pmatrix} x & 0.2 - x \\ 0.2 & 0.6 \end{pmatrix}$$

$$ABC = \begin{pmatrix} x & 0.2 - x \\ 0.2 & 0 \\ 0 & 0.6 \end{pmatrix}$$

$$C = (x + 0.2 \quad 0.8 - x)$$

Получаем уравнение:

$$I'_x = -\log(x + 0.2) + \log(0.2 - x) - \log(x) + \log(0.2 - x) + \log(x) - \log(0.2 - x) + \log(x + 0.2) - \log(0.8 - x) = 0$$

Которое путем сокращений и взятие экспоненты приводится к виду

$$0.2 - x = 0.8 - x$$

которое не имеет решений.

4.12. $y=0.2, z=0.6$

$$AC = \begin{pmatrix} x + 0.2 & 0.2 - x \\ 0.6 & 0 \end{pmatrix}$$

$$BC = \begin{pmatrix} x & 0.2 - x \\ 0.8 & 0 \end{pmatrix}$$

$$ABC = \begin{pmatrix} x & 0.2 - x \\ 0.2 & 0 \\ 0.6 & 0 \end{pmatrix}$$

$$C = (x + 0.8 \quad 0.2 - x)$$

Получаем уравнение:

$$\begin{aligned} I'_x &= -\log(x + 0.2) + \log(0.2 - x) - \log(x) + \log(0.2 - x) \\ &+ \log(x) - \log(0.2 - x) + \log(x + 0.8) - \log(0.2 - x) = 0 \end{aligned}$$

Которое путем сокращений и взятие экспоненты приводится к виду

$$0.2 + x = 0.8 + x$$

которое не имеет решений.

5. Вершины

Мы получили, что максимум достигается в одной из 6 вершин параллелепипеда (в частях 3.6 и 3.7 выявлено, что в двух парах вершин функция имеет одинаковое значение). Нам остается посчитать значения в этих вершинах и определить среди них самое большое.

5.1. $x=0, y=0, z=0$

После подстановки значений и сокращения одинаковых слагаемых получаем:

$$-0.4 \log 0.4 - 0.8 \log 0.8 + 0.2 \log 0.2 = 0.32$$

5.2. $x=0, y=0, z=0.6$

После подстановки значений все слагаемые взаимно сокращаются и дают нулевое значение. Такое же значение принимается в точке $(0, 0.2, 0.6)$.

5.3. $x=0, y=0.2, z=0$

После подстановки значений и сокращения одинаковых слагаемых получаем:

$$-0.2 \log 0.2 - 0.6 \log 0.6 + 0.8 \log 0.8 = 0.45$$

5.4. $x=0.2, y=0, z=0$

После подстановки значений все слагаемые взаимно сокращаются и дают нулевое значение. Такое же значение принимается в точке $(0.2, 0.2, 0)$.

5.5. $x=0.2, y=0, z=0.6$

После подстановки значений и сокращения одинаковых слагаемых получаем:

$$-0.2 \log 0.2 - 0.6 \log 0.6 + 0.8 \log 0.8 = 0.45$$

5.6. $x=0.2, y=0.2, z=0.6$

После подстановки значений и сокращения одинаковых слагаемых получаем:

$$-0.4 \log 0.4 - 0.8 \log 0.8 + 0.2 \log 0.2 = 0.32$$

Получается, что максимум достигается в точках $(0, 0.2, 0)$ и $(0.2, 0, 0.6)$ и равен 0.45.

6. Заключение

Мы рассмотрели условную совместную информацию двух дискретных случайных величин с заданным совместным распределением при условии третьей случайной величины. Была решена задачи максимизации такой информации среди всех дискретных случайных величин, принимающих два значения. Задача была решена путем поэтапной проверки на наличие экстремальных точек внутренностей параллелепипеда, заключающего все возможные значения набора вероятностей, а затем внутренностей граней и ребер. Было доказано, что функция достигает своего максимума лишь в вершинах параллелепипеда. После проверки значений во всех вершинах было выявлено максимальное значение информации и найдены все совместные распределения трех случайных величин, в которых оно достигается. Отметим, что метод, примененный в настоящей работе, может быть использован для нахождения точного глобального максимума для любых дискретных совместных распределений, в особенности, если функции принимают небольшое число значений. Тем не менее, остается открытым вопрос, для всех ли распределений функций, принимающих два значения, максимум будет достигаться в вершинах параллелепипеда.

СПИСОК ЦИТИРОВАННОЙ ЛИТЕРАТУРЫ

1. Верещагин, Н.К. Лекции по теории информации // 2023.
2. Верещагин, Н.К., Щепин, Е.В. Информация, кодирование и предсказание // Москва, МЦНМО .- 2012.
3. Яглом, А.М., Яглом, И.М. Вероятность и информация // Москва: Наука.- 1973.
4. Чисар, И., Кернер, Я. Теория информации // Москва: Мир.- 1985.
5. Cesa-Bianchi, N., Lugosi, G. Prediction, learning, and games // Cambridge University Press.- 2006.

REFERENCES

1. Vereshchagin, N.K. 2023, “Lectures on Information Theory”.
2. Vereshchagin, N.K. and Shchepin, E.V. 2012 , “Information, Coding and Prediction”, *Moscow: MCNMO*.
3. Yaglom, A.M. and Yaglom, I.M. 1973, “Probability and Information”, *Moscow: Nauka*.
4. Chisar, I. and Kerner, Y. 1985, “Information Theory”, *Moscow: Mir*.
5. Cesa-Bianchi, N. and Lugosi, G. 2006, “Prediction, Learning, and Games”, *Cambridge University Press*.

Получено: 23.06.2024

Принято в печать: 24.12.2024