

ИНФОРМАЦИЯ И ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ. ИЗМЕРЕНИЕ И КОДИРОВАНИЕ ИНФОРМАЦИИ

6. Измерение информации

6.1. Меры информации

Универсальная схема передачи информации



Аспекты исследования сообщений

Синтаксический - рассматриваются
внутренние свойства сообщений

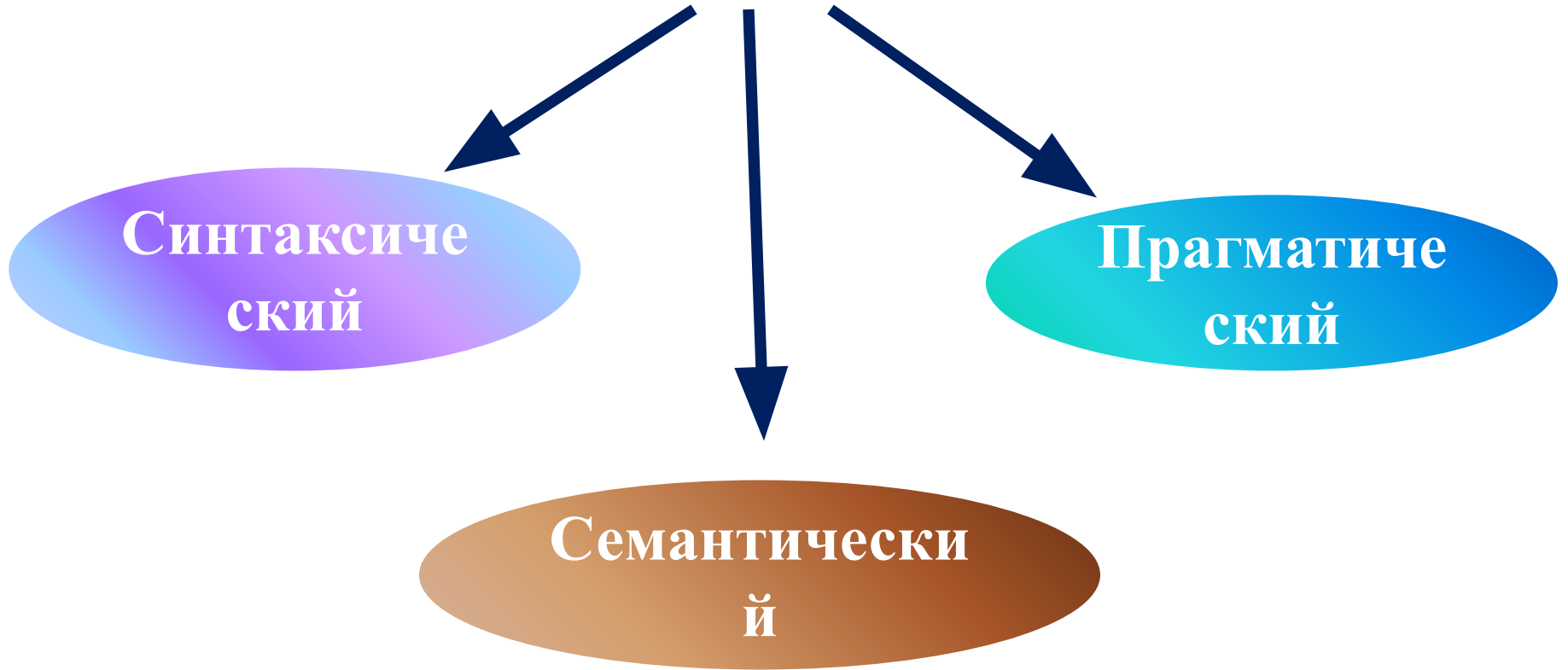
Семантический - анализируется
смысловое содержание сообщения, его
отношение к источнику информации

Прагматический - рассматривается
потребительское содержание
сообщения, его отношение к
получателю

Универсальная схема передачи информации



Уровни изучения передачи информации



Синтаксический уровень

Идея:

- Технические проблемы совершенствования методов передачи сообщений и их материальных носителей – сигналов
- Проблемы доставки получателю сообщений
- Полностью абстрагируются от смыслового содержания сообщений и их целевого предназначения

Учитывают:

- Тип носителя
- Способ представления информации
- Скорость передачи и обработки
- Размеры кодов представления информации и т.

Семантический уровень

Идея:

- Проблемы связаны с формализацией и учетом смысла передаваемой информации.
- Проблемы этого уровня чрезвычайно сложны, так как смысловое содержание информации больше зависит от получателя, чем от семантики сообщения, представленного на каком-либо языке.

На данном уровне:

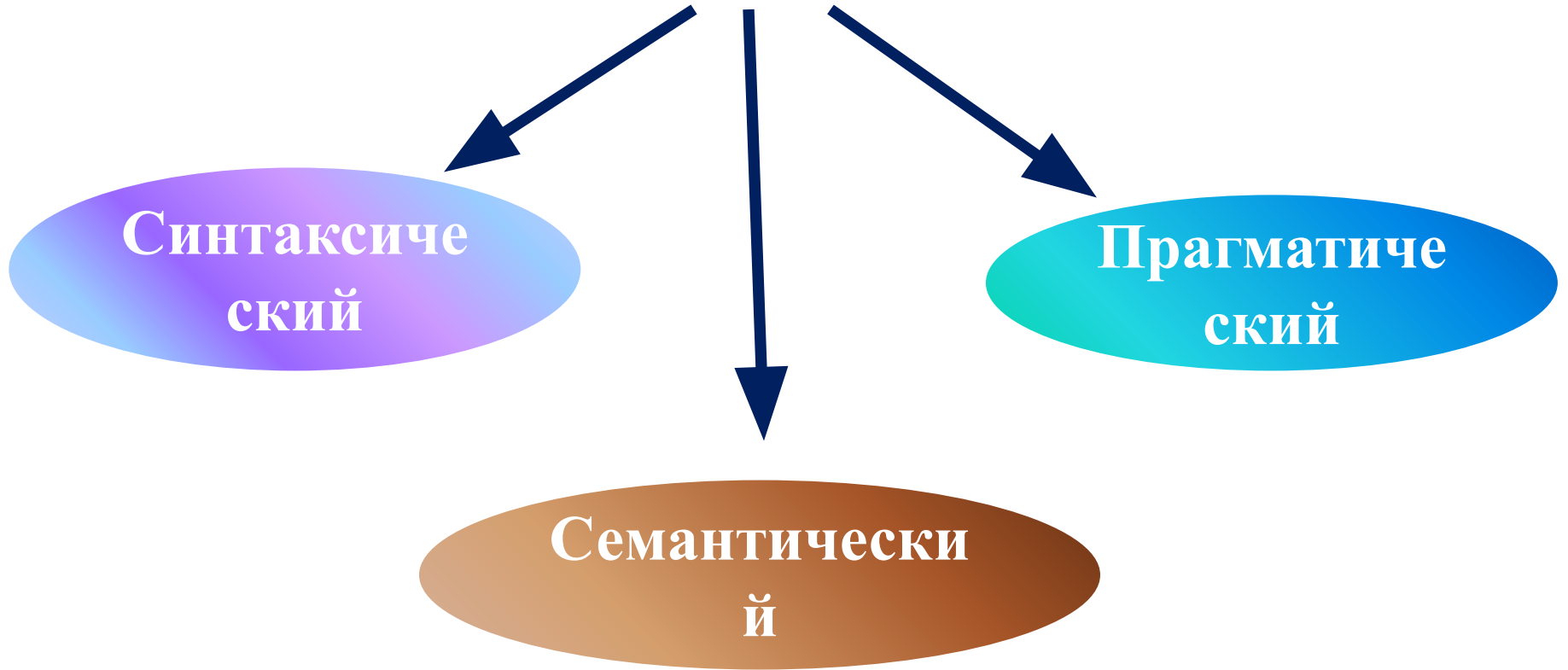
- Анализируется сведения, которые отражает информация
- Выявляется смысл информации
- Выявляется содержание информации
- Обеспечивается обмен информацией

Прагматический уровень

Идея:

- ▣ проблемы этого уровня связаны с определением ценности и полезности информации для потребителя
- ▣ интересуют последствия от получения и использования данной информации потребителем.

Уровни изучения передачи информации



- **Меры информации**

- **синтаксические**

- **объем**

- **Данных**

- **V**

- **Количество информации**

- **I**

- **семантические**

■ Меры информации

- **Количество информации**

- *I*

- **Прагматические**
- **меры**

- **Синтаксические**
- **меры**

- **Объем данных**

- *V*

ИНФОРМАЦИЯ И ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ. ИЗМЕРЕНИЕ И КОДИРОВАНИЕ ИНФОРМАЦИИ

6. Измерение информации

6.2. Объем данных.

Универсальная схема передачи информации



Объем данных

Объем данных измеряется количеством символов (минимальных неделимых единиц данных) в сообщении

Алфавит

При использовании современных компьютеров данные кодируются группами двоичных разрядов – *битов*

Для кодирования символов могут быть использованы кодовые группы различной длины. Кодовая группа минимальной длины, достаточная для кодирования одного символа, носит название *байт*

Байтовый алфавит

Принято считать, что длина байта – 8 символов.

Понятие байта как *единицы измерения объема*

■ Меры информации

■ синтаксические

- объем
- Данных
- V

- Количество информации
- I

■ семантические

■ прагматические

■ Меры информации

■ синтаксические

- объем
- Данных
- V

- Количество информации
- I

■ семантические

■ прагматические

ИНФОРМАЦИЯ И ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ. ИЗМЕРЕНИЕ И КОДИРОВАНИЕ ИНФОРМАЦИИ

6. Измерение информации

6.3. Основные сведения теории информации

Теория информации

наука, изучающая
количественные
закономерности, связанные с
получением, передачей,
обработкой и хранением
информации, структуру и
закономерности протекания
информационных процессов.

Задачи теории информации

- 1. Отыскание наиболее экономных методов кодирования, позволяющих передать заданную информацию с помощью минимального количества символов**
- 2. Определение пропускной способности канала, чтобы он передавал всю поступающую информацию без помех и искажений**
- 3. Определение объема запоминающих устройств, поиск способов ввода в эти устройства данных и вывода их для**

**Для решения перечисленных
задач нужно, прежде всего,
научиться измерять
количественный объем
передаваемой или хранимой
информации, пропускную
способность каналов связи и их
чувствительность к помехам
(искажениям)**

Ученые, внесшие вклад в развитие теории информации

Р. Фишер

Х. Найквист

Р. Хартли

К. Шеннон

Основные положения количественной оценки информации

Информация об объекте (источнике информации) передается в виде сообщений, имеющих различную синтаксическую форму

Следовательно, сообщение—это форма представления информации

Можно пренебречь содержательной стороной информации, выразить её количество числом, и, следовательно, сравнить количество информации, содержащейся в различных группах данных

Основные положения количественной оценки информации

**в теории информации используются
синтаксические меры информации,
оперирующие с обезличенной
информацией и не учитывающие ее
семантического и прагматического
аспектов**

Основные положения количественной оценки информации

За отправную точку принимается то, что информация, содержащаяся в сообщении, полностью или частично устраняет априорную неопределенность объекта, процесса, явления

В качестве объекта, о котором передается информация, в теории информации рассматривают некоторую физическую систему, которая может оказаться в том или ином состоянии, т.е. систему, которой заведомо присуща некоторая степень неопределенности

Основные положения количественной оценки информации

**Описание любого события или
объекта формально можно
рассматривать как указание на то, в
каком из возможных состояний
находится описываемый объект**

**Тогда протекание событий во
времени есть не что иное, как
смена возможных состояний
системы**

Основные положения количественной оценки информации

**Следовательно, сообщение имеет
смысл только тогда, когда состояние
объекта заранее неизвестно.**

Основные положения количественной оценки информации

**Таким образом, информацию,
содержащуюся в сообщении, можно
рассматривать с точки зрения её
новизны то есть, уменьшения
неопределённости наших знаний об
объекте.**

Основные положения количественной оценки информации

1. **что значит большая или меньшая степень неопределенности системы?**
2. **чем можно ее измерить?**

ИНФОРМАЦИЯ И ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ. ИЗМЕРЕНИЕ И КОДИРОВАНИЕ ИНФОРМАЦИИ

6. Измерение информации

6.4. Структурный подход Хартли

Основные положения количественной оценки информации

1. **что значит большая или меньшая степень неопределенности системы?**
2. **чем можно ее измерить?**

Структурный подход Хартли

Хартли предположил, что любая система может с одинаковой степенью вероятности находиться в любом из своих возможных состояний, т.е. все состояния системы являются равновероятными. Поэтому в рамках этого подхода степень неопределенности определяется числом возможных состояний системы

Структурный подход Хартли (примеры)

Пусть имеются пять различных объектов: партнер по игре в крестики–нолики; поезд; вещество; студент; колода карт. О каждом из них нас интересует следующая информация:

1. В какой из четырех углов сделает свой первый ход партнер по игре в крестики–нолики?
2. В какое время суток (утром, днем, вечером или ночью) прибывает поезд?
3. В каком состоянии (твердое, жидкое, газообразное или плазма) находится некоторое вещество?
4. Какой из четырех оценок (неудовлетворительно, удовлетворительно, хорошо или отлично) оценил преподаватель знания студента?
5. Какой из четырех возможных мастей оказалась карта, наугад вынутая из колоды?

Структурный подход Хартли

Хартли предложил структурный подход к определению т.н. емкости системы, передающей и накапливающей информацию, и количества информации в сообщении. Этот подход называют структурным, т.к. он учитывает структуру источника информации, т.е. количество его возможных состояний.

Структурный подход Хартли

Если количество возможных сообщений в множестве возможных сообщений равно n , то, согласно Хартли, информация I_1 , $I_1 = \log_a n$ приходящаяся на одно сообщение, определяется логарифмом общего числа возможных сообщений n :

$$I_1 = \log_a n$$

Свойства меры Хартли

При увеличении число возможных состояний источника информации возрастает степень его неопределенности, и, следовательно, увеличивается количество информации в сообщении;

Если источник информации может находиться только в одном состоянии (передать только одно сообщение), то, следовательно, его состояние полностью определено, и количество информации в передаваемом сообщении равно нулю;

Свойства меры Хартли

Пусть сообщения исходят из двух независимых источников, которые могут передавать n_1 и n_2 возможных сообщений соответственно. Каждое сообщение от первого источника может быть объединено с сообщением от второго источника, образуя новое сложное сообщение. Общее число таких сообщений $n = n_1 \cdot n_2$. Следовательно, количество информации, приходящееся на одно сложное сообщение, определяется формулой:

$$I_1 = \log_a n = \log_a (n_1 \cdot n_2) = \log_a n_1 + \log_a n_2$$

Свойства меры Хартли

логарифмическая мера информации обладает свойством *аддитивности*.

В самом деле, количество информации на сообщение в этом случае оказывается равным сумме количеств информации, которые были бы получены на одно сообщение из двух источников, взятых порознь

ИНФОРМАЦИЯ И ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ. ИЗМЕРЕНИЕ И КОДИРОВАНИЕ ИНФОРМАЦИИ

6. Измерение информации

6.5. Математический аппарат, используемый в теории информации

Основные понятия теории вероятностей

Многим явлениям присуща случайная природа, и их надо изучать с точки зрения закономерностей, присущих им как случайным явлениям

Опыт показывает, что если нельзя конкретно судить об исходе отдельных явлений, то наблюдая в совокупности массы однородных случайных явлений, мы обнаруживаем в них вполне определенные закономерности (устойчивости), свойственные именно массовым случайным явлениям

Основные понятия теории вероятностей

Наука, изучающая специфические закономерности случайных явлений и процессов, получила название теории вероятностей

Понятие события

Под событием будем понимать всякий факт (результат какого-либо опыта, наблюдения, эксперимента и т.п.), который в результате опыта может произойти (наступить, осуществиться) или не произойти.

вероятность события

Чтобы сравнивать между собой события по степени их возможности, с каждым из событий связывают определенное число, которое называют вероятностью события

•

вероятность события

В качестве единицы измерения степени возможности (вероятности) события принята вероятность **достоверного** события, которому приписывают вероятность, равную единице.

Противоположным по отношению к достоверному событию является **невозможное** событие, которому приписывается вероятность, равная нулю.

вероятность события

Таким образом, вероятности возможных событий представляют собой числа в пределах от 0 до 1

Непосредственный подсчет вероятности

все исходы опыта образуют
полную группу

все исходы опыта попарно
несовместны

все исходы опыта обладают
симметрией и в силу этого
объективно

одинаково возможны

Непосредственный подсчет вероятности

Несколько событий в данном опыте образуют *полную группу событий*, если в результате опыта непременно должно появиться хотя бы одно из них

События называют *несовместными*, если появление одного из них исключает появление других событий в том же опыте

События называют *равновозможными*, если по условиям симметрии опыта есть основание считать, что ни одно из этих событий не является более объективно возможным, чем другое

Непосредственный подсчет вероятности

Несовместные равновозможные события, образующие полную группу, называют *случаями*

Случай называют *благоприятным* (благоприятствующим), некоторому событию, если появление этого случая влечет за собой появление соответствующего события

Непосредственный подсчет вероятности

$$P(A) = \frac{m}{n}$$

$P(A)$ — вероятность события A ;

n — общее число случаев,

m — число случаев, благоприятных
событию A .

определение статистической вероятности (частоты) события

Если проведена серия из n опытов, в каждом из которых могло появиться или не появиться событие A , то частотой появления события A в данной серии опытов называется отношение числа опытов m в которых появилось событие A к общему числу произведенных опытов.

$$P(A) = \frac{m}{n}$$

случайная величина

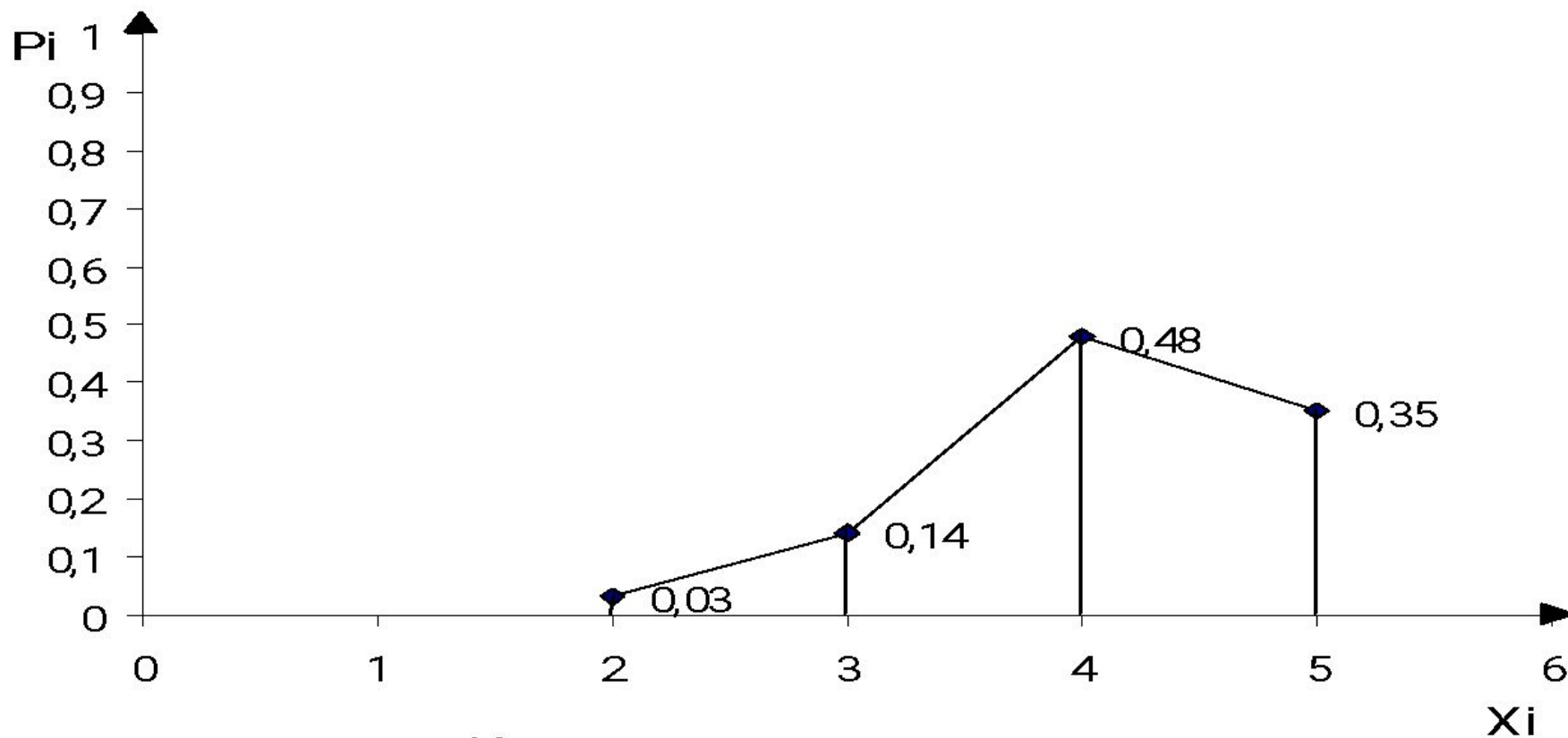
X
 x_1, x_2, \dots, x_n p_1, p_2, \dots, p_n

x_i	x_1	x_2	\dots	x_n
p_i	p_1	p_2	\dots	p_n

$$p_1 + p_2 + \dots + p_n = 1$$

Многоугольник распределения

оценка	2	3	4	5
вероятность	0,03	0,4	0,48	0,35



Многоугольник распределения

характеристики положения

Определяют некоторое среднее ориентировочное значение, около которого группируются все возможные значения случайной величины

Характеристики положения характеризуют положение случайной величины на числовой оси

Таким образом, среднее значение случайной величины есть некоторое число, заменяющее ее при расчетах

математическое ожидание (среднее значение)

X_i	X_1	X_2	\square	X_n
p_i	p_1	p_2	\square	p_n

$$M(X) = \frac{p_1 x_1 + p_2 x_2 + \square p_n x_n}{p_1 + p_2 + \square p_n} = \frac{\sum_{i=1}^n p_i x_i}{\sum_{i=1}^n p_i}$$

математическое ожидание

$$M(X) = \frac{p_1 x_1 + p_2 x_2 + \dots + p_n x_n}{p_1 + p_2 + \dots + p_n} = \frac{\sum_{i=1}^n p_i x_i}{\sum_{i=1}^n p_i}$$

$$\sum_{i=1}^n p_i = 1$$

$$M(X) = \sum_{i=1}^n p_i x_i$$

математическое ожидание

$$M(X) = 0,03 \cdot 2 + 0,14 \cdot 3 + 0,48 \cdot 4 + 0,35 \cdot 5 = 4,15$$

Логарифм, свойства логарифма

$$y = \log_a x. \quad a > 0, a \neq 1, x > 0.$$

$$a^y = x, \text{ т.е. } a^{\log_a x} = x. \quad \log_a 1 = 0.$$

$$\log_a (x_1 \cdot x_2) = \log_a |x_1| + \log_a |x_2|$$

$$\log_a \left(\frac{x_1}{x_2} \right) = \log_a |x_1| - \log_a |x_2|$$

$$\log_a x^n = n \cdot \log_a |x| \quad \log_a x = \frac{1}{\log_b a} \cdot \log_b x$$

$$\frac{1}{\log_b a}$$