

---

# НЕОПРЕДЕЛЁННОСТЬ ЭНТРОПИЯ

---

- 
- Основоположенник теории информации *Клод Шеннон* определил *информацию*, как *снятую неопределенность*.
  - Задача, которая решается в ходе снятия неопределенности – уменьшение количества рассматриваемых вариантов (уменьшение разнообразия), и в итоге выбор одного соответствующего ситуации варианта из числа возможных. В этом управляющая роль информации
-

- Ситуация *максимальной неопределенности* предполагает наличие нескольких *равновероятных* альтернатив (вариантов), т.е. ни один из вариантов не является более предпочтительным. Причем, *чем больше равновероятных вариантов* наблюдается, тем больше неопределенность, тем сложнее сделать однозначный выбор и *тем больше информации требуется* для этого получить. Для  $N$  вариантов эта ситуация описывается следующим распределением вероятностей:  $\{1/N, 1/N, \dots, 1/N\}$ .
- *Минимальная неопределенность* равна  $0$ , т.е. эта ситуация *полной определенности*, означающая что выбор сделан, и вся необходимая информация получена. Распределение вероятностей для ситуации полной определенности выглядит так:  $\{1, 0, \dots, 0\}$ .

- Величина, характеризующая количество неопределенности в теории информации обозначается символом  $H$  и имеет название *энтропия*, точнее *информационная энтропия*.
- *Энтропия ( $H$ ) – мера неопределенности, выраженная в битах. Так же энтропию можно рассматривать как меру равномерности распределения случайной величины.*
- *Количество информации  $I$  и энтропия  $H$  характеризуют одну и ту же ситуацию, но с качественно противоположенных сторон.  $I$  – это количество информации, которое требуется для снятия неопределенности  $H$ . По определению Леона Бриллюэна информация есть отрицательная энтропия (негэнтропия).*

# Формула Шеннона

- В общем случае, энтропия  $H$  и количество получаемой в результате снятия неопределенности информации  $I$  зависят от исходного количества рассматриваемых вариантов  $N$  и вероятностей реализации каждого из них  $P: \{p_0, p_1, \dots, p_{N-1}\}$ , т.е.  $H=F(N, P)$ . Расчет энтропии в этом случае производится по формуле Шеннона, предложенной им в 1948 году.
- В частном случае, когда все варианты равновероятны, остается зависимость только от количества рассматриваемых вариантов, т.е.  $H=F(N)$ . В этом случае формула Шеннона совпадает с формулой Хартли, которая впервые была предложена американским инженером Ральфом Хартли в 1928 году.

$$H = -\sum_{i=0}^{N-1} p_i \log_2(p_i) = \sum_{i=0}^{N-1} p_i \log_2\left(\frac{1}{p_i}\right)$$

- Знак минус в формуле (1) не означает, что энтропия – отрицательная величина.
- Энтропия интерпретируется как частное количество информации, получаемое в случае реализации  $i$ -ого варианта.
- Энтропия в формуле Шеннона является средней характеристикой – математическим ожиданием распределения случайной величины  $\{I_0, I_1, \dots, I_{N-1}\}$ .

# Формула Хартли

- формула Хартли – частный случай формулы Шеннона для равновероятных альтернатив.
- Из нее следует, что чем больше количество альтернатив ( $N$ ), тем больше неопределенность ( $H$ ). Эти величины связаны в формуле через двоичный логарифм и приводит количество вариантов к единицам измерения информации – битам.

$$H = \log_2(N)$$

- Для решения обратных задач, когда известна неопределенность ( $H$ ) или полученное в результате ее снятия количество информации ( $I$ ), используют обратную формулу Хартли:  $N = 2^H$