

# Теория информации.

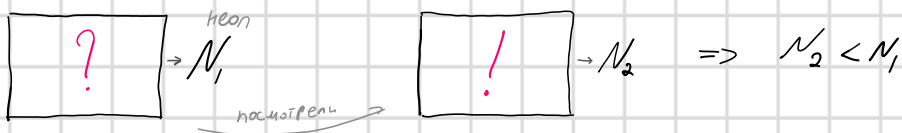
Пр:

1) 38 карточками 3 бита

Если 1000, то 10 бит

3 бита  $\Rightarrow 2^3$  сообщений  $< 1000$

## Неопределенность



$$\Delta N = N_2 - N_1$$

$$\text{информация} = -\Delta N$$

## Случайный источник

$$\Omega = \{\omega_1, \dots, \omega_n\}; p_1, \dots, p_n$$

$$C.V. - \mathbb{Q}, \mathbb{Q} \dots$$

$$P(\mathbb{Q}_i = a) = p_a, a \in \{1, \dots, h\}$$

Энтропия:  $H(p_1, \dots, p_n) : RS \xrightarrow{\text{случ. источник}} \mathbb{R}^+$

$$1) f(1) = e$$

$$2) f(a+b) = f(a)f(b)$$

$$3) f \in C(\mathbb{R})$$

$$f(x) = e^x \leftarrow \text{единственное решение}$$

$$H(p_1, \dots, p_n) = - \sum_{i=1}^n p_i \cdot \log p_i$$

$$\exists h(n) = H\left(\frac{1}{n}, \dots, \frac{1}{n}\right)$$

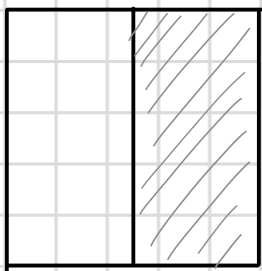
Свойства:

1) монотонность

$$h(n+1) > h(n)$$

2) Аддитивность

$$\Omega = \left\{ (1,1), (1,2), \dots, (1,m_1), \right. \\ \left. (2,1), (2,2), \dots, (2,m_2), \right. \\ \left. \vdots \right. \\ \left. (k,1), (k,2), \dots, (k,m_k) \right\}$$



$$p((i,j)) = q_{ij}$$

$$\sum q_{ij} = 1$$

$$p_i = \sum_{j=1}^{m_i} q_{ij}$$

$$p_1, \dots, p_k \rightarrow \sum p_i = 1$$

$$p((1,1) | (1,x)) = \frac{q_{11}}{p_1}$$

$$\Rightarrow p((i,j) | (i,x)) = \frac{q_{ij}}{p_i}$$

$$H(q_{11}, q_{12}, \dots, q_{km_k}) = H(p_1, p_2, \dots, p_k) + \sum_{i=1}^k p_i \cdot H\left(\frac{q_{i1}}{p_i}, \frac{q_{i2}}{p_i}, \dots, \frac{q_{im_i}}{p_i}\right)$$

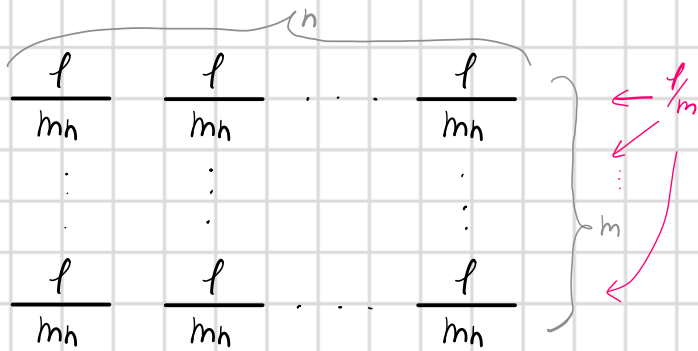
3) Непрерывность

$$H(p_1, \dots, p_n) - \text{непрерывна как ф-ция из } \mathbb{R}^n$$

Докажем это в 2 шага

$$1) h(m \cdot n) = h(m) + h(n) \quad // \quad h(n) = \log n$$





$$h(m \cdot n) = h(m) + \sum \frac{1}{m} \cdot H\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right) \ominus$$

$$q_{ij} = \frac{1}{m \cdot n}; p_i = \frac{1}{m}; \frac{q_{ij}}{p_i} = \frac{1}{n}$$

$$\ominus h(m) + h(n) \blacktriangleleft$$

**Lemma:**  $h(n^r) = r \cdot h(n), r \in \mathbb{N}$

$$\left. \begin{array}{l} x \geq 0 \\ \forall n: x \leq \frac{1}{n} \end{array} \right\} \Rightarrow x = 0$$

$$h(n) = \log_2 n \cdot h(2)$$

$$\triangleleft 2^k \leq n^r < 2^{k+1}, r \in \mathbb{N}$$

$$\frac{k}{r} \leq \log_2 n < \frac{k+1}{r}$$

$$h(2^k) \leq h(n^r) \leq h(2^{k+1})$$

$$\exists h(2) = \alpha$$

$$\frac{h(2) \cdot k}{r} \leq h(n) \leq \frac{(k+1)h(2)}{r}$$

$$\frac{\alpha k}{r} \leq h(n) \leq \frac{\alpha (k+1)}{r}$$

$$\frac{\alpha k}{r} \leq \alpha \log_2 n < \frac{\alpha (k+1)}{r}$$



$$\forall r: r \in \mathbb{N}, |\alpha \log_2 n - h(n)| \leq \frac{\alpha}{r}$$

$$\Rightarrow \alpha \log_2 n = h(n) \blacktriangleleft$$

$$p_i \in \mathbb{Q} \quad p_i = \frac{q_i}{b}, \sum q_i = b$$

$$H(p_1, \dots, p_n) = ?$$

$$\frac{1}{b}, \frac{1}{b}, \dots, \frac{1}{b} \leftarrow q_1 \text{ wtyk}$$

$$\frac{1}{b}, \frac{1}{b}, \dots, \frac{1}{b} \leftarrow q_2 \text{ wtyk}$$

$$\dots \leftarrow q_n \text{ wtyk}$$

$$h(b) = H(p_1, \dots, p_n) + \sum p_i \cdot h(q_i)$$

$$\alpha = h(2)$$

$$\sum p_i \cdot \log_2 b \cdot \alpha = H(p_1, \dots, p_n) + \sum p_i \cdot \log_2 q_i \cdot \alpha$$

$$\log_2 q_i - \log_2 b = \log_2 \frac{q_i}{b} = \log_2 p_i$$

$$\rightarrow H(p_1, \dots, p_n) = -\sum p_i \cdot \log p_i \quad \checkmark$$

$$L = h(2) \quad ?$$

$$h(2) = 1$$

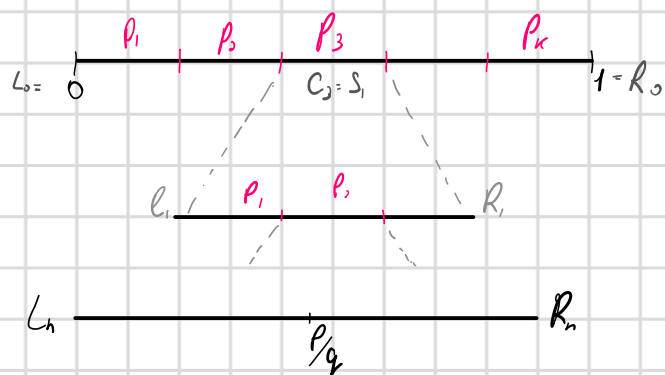
Опр.:  $\mathcal{H}$  - Энтропия честной монеты

$$[h(1000)] = [g \dots] = 10$$

Арифметическое кодирование

$$\exists S \in \Sigma^* \quad |S| = h \quad |\Sigma| = *$$

$$f_i - \text{кон-во } c_i \text{ в } S; \quad p_i = \frac{f_i}{n}$$



$$-\log_2 (R_h - L_h) \leftarrow \log_2 q \text{ нгдифер}$$

$$1 \cdot p_{s_1} \cdot p_{s_2} \cdot \dots \cdot p_{s_n} = \prod_{i=1}^k p_i^{f_i}$$

Сколько раз встречается  $p_i$

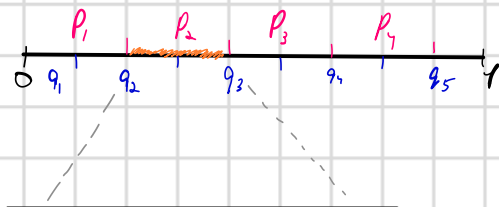
$$-\log_2 \left( \prod_{i=1}^k p_i^{f_i} \right) = -n \sum p_i \cdot \log_2 p_i$$

$$|a| \leq n \cdot H(p_1, \dots, p_k)$$

Эмпирические энтропийные источники.

$$\exists p_1, \dots, p_n$$

$$q_1, \dots, q_m$$



нат. описание к-ва изоб:  $\approx \frac{H(q_1, \dots, q_k)}{H(p_1, \dots, p_n)}$