Про информацию

1 Про информацию

Информация всегда содержится в сообщении.

2 Как измерить информацию

2.1 Первое предположение:

Кол-во инфы в сообщение примерно пропорционально длине сообщения. К сожалению это не всегда так, нужно просто послушать пьяные бредни алкашей и всё станет понятно - Слов много, а смысла нет.

2.2 Второе предположение

Кол-во информации в сообщении зависит от вероятности события. Чем меньше вероятность возникновения события, тем больше информации несет сообщение.

2.3 Определение (Кол-во информации)

Пусть **кол-во информации** в сообщении "Произошло A" $J(A) = J(\text{произошло } A)$ - это функция от вероятности события A.

- $J(A) = J(P(A)) \Rightarrow J : [0, 1] \to \mathbb{R}^+ \cup \{0\}$
- J - это непрерывная функция (т.е если мы чуть-чуть изменим вероятность события A, то кол-во информации тоже изменится на чуть-чуть)
- J - монотонно убывающая функция
- Если $P(A) = 1 \Rightarrow J(A) = 0$
- Если A и B - независимые, то $J(A \cdot B) = J(A) + J(B)$
- Если $P(A) = \frac{1}{2}$, то $J(A) = 1$

3 Теорема об информации

1. $\forall p, q \in [0,1] : J(p \cdot q) = J(p) + J(q)$

2. $J(p)$ - непрерывная

3. $J\left(\frac{1}{2}\right) = 1$

Тогда $J(p) = -\log(p)$

3.1 Доказательство

Для 1-го свойства проведем индукцию по n и получим свойство 1'):
1') $J(p^n) = n \cdot J(p)$

- сделаем замену $p^n = q \to n = q^{\frac{1}{n}}$
  1") $\frac{1}{n} \cdot J(p) = J\left(q^{\frac{1}{n}}\right)$

- Получили свойство 1") для чисел степеней обратных натуральным
  $J\left(p^{\frac{m}{n}}\right) = \frac{1}{n} \cdot J(p^m) = \frac{m}{n} \cdot J(p)$

- Получили аналог свойства 1) для всех натуральных чисел
  $J\left(\left(\frac{1}{2}\right)^{\frac{m}{n}}\right) = \frac{m}{n} \cdot J\left(\frac{1}{2}\right) = \frac{m}{n}$

Теперь покажем, что $J(p)$ это именно $-\log(p)$

- т.к $-\log(p)$ - это действительное число, то можно построить 2 последовательности рациональных чисел, которые будут стремиться к нему сверху и снизу
  $\left\{\frac{m_i}{n_i}\right\}_{i=0}^{\infty} \nearrow -\log(p)$
  $\left\{\frac{u_n}{v_i}\right\}_{i=0}^{\infty} \searrow -\log(p)$
  Тогда $\forall i : \frac{m_i}{n_i} < \log(p) < \frac{u_n}{v_i} \Rightarrow$
  $-\frac{m_i}{n_i} > \log(p) > -\frac{u_n}{v_i} \Rightarrow$
  $2^{-\frac{m_i}{n_i}} > p > 2^{-\frac{u_n}{v_i}} \Rightarrow$

- т.к $J(p)$ - убываюящая, то меняем знак неравенства
  $J\left(2^{-\frac{m_i}{n_i}}\right) < J(p) < J\left(2^{-\frac{u_n}{v_i}}\right) \Rightarrow$
  $\frac{m_i}{n_i} < J(p) < \frac{u_n}{v_i} \Rightarrow$

- Делаем предельный переход по i
  $-\log(p) = \lim\limits_{i\to\infty} \frac{m_i}{n_i} \leq J(p) \leq \lim\limits_{i\to\infty} \frac{u_i}{v_i} = -\log(p)$
  ■

3.2 Определение (Стохастический вектор)

$(x_1, x_2, \cdots, x_n) = \vec{x}$ - **стохастический вектор**, если верно:

- $\vec{x} \in \mathbb{R}^n$
- $\forall i : x_i \geq 0$
- $\sum_{i=1}^{n} x_i = 1$

3.3 Определение (Энтропия)

Пусть $\mathcal{X}$ - дискретная случайная величина.
Тогда $J(\mathcal{X} = x_i) = -\log(P(\mathcal{X} = x_i))$

- $J(p)$ - это функция от случайной величины $\Rightarrow$ тоже случайная величина

Энтропия случайной величины $\mathcal{X}$ - Это мат.ожидание кол-ва информации

- $H(x) = M[J(x)]$
- $H(x) = \sum_{i=0}^{n} p_i \cdot J(p_i) = -\sum_{i=0}^{n} p_i \cdot \log(p_i)$
- $H(x) = H(p_1, p_2, \ldots, p_n) = H(\vec{p})$
- слагаемые вида $0 \cdot \log(0)$ можно не учитывать в подсчете энтропии, они дают 0

3.4 Свойства энтропии

1. $H(p_1, p_2, \ldots, p_n, 0) = H(p_1, p_2, \ldots, p_n)$
2. Пусть $\sigma \in S_n$. Тогда $H(p_{\sigma(1)}, p_{\sigma(2)}, \ldots, p_{\sigma(n)}) = H(\vec{p})$
   - $\sigma \in S_n$ - это перестановки длины n
   - Т.е как бы мы не переставляли аргументы функции $H(p)$, результат её не изменится
3. $0 \leq H(p_1, p_2, \ldots, p_n) \leq \log(n) = H\left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right)$ **правое равенство достигается при равномерном распределении**
4. $H(p_1, \ldots, p_n) = H(p_1, \cdots, p_{n-2}, p_{n-1}+p_n) + H(p_{n-1}+p_n) \cdot H\left(\frac{p_{n-1}}{p_{n-1}+p_n}, \frac{p_n}{p_{n-1}+p_n}\right)$
   - Для доказательства просто подставь все в формулы

3.5 Доказательство ($\leq$ 3-его свойства)

Сначала покажем, что $\ln(x) \leq x - 1$

- Графики имеют единственную точку пересечения в $(1, 0)$

Сравним производные:

- $(\ln(x))' = \frac{1}{x}$
- $(x-1)' = 1$
- Видим, что $(\ln(x))' = \frac{1}{x} < 1 = (x-1)' (\forall x > 1)$
- При $(0 < x < 1) : 1 < \frac{1}{x}$

$\angle H(p) - \log(n) =$
$-\sum_{i=1}^{n} p_i \cdot \log(p_i) - \log(n) =$
$* \sum_{i=1}^{n} p_i = 1$
$-\sum_{i=1}^{n} p_i \cdot \log(p_i) - \sum_{i=1}^{n} p_i \cdot \log(n) =$
$-\sum_{i=1}^{n} p_i \cdot (\log(p_i) + \log(n)) =$
$\sum_{i=1}^{n} p_i \cdot \log\left(\frac{1}{n \cdot p_i}\right) =$
$* \log(x) = \frac{\ln(x)}{\ln(2)} * \ln(x) = \frac{\log(x)}{\log(e)}$
$\log(e) \cdot \sum_{i=1}^{n} p_i \cdot \ln\left(\frac{1}{n \cdot p_i}\right) \leq$
$\log(e) \cdot \sum_{i=1}^{n} p_i \cdot \left(\frac{1}{n \cdot p_i} - 1\right) =$
$\log(e) \left(\sum_{i=1}^{n} \frac{1}{n} - \sum_{i=1}^{n} p_i\right) = 0$
$\Rightarrow H(\vec{p}) - \log(n) \leq 0$

Равенство будет при равномерном распределении. В этом случае $p_i = \frac{1}{n}$ и тогда $\log(e) \cdot \sum_{i=1}^{n} p_i \cdot \left(\frac{1}{n \cdot p_i} - 1\right) = 0$