Энтропия и пересчет

На семинаре мы аксиоматически ввели «меру» $S : [0,1] \to \mathbb{R}^+$ удивления. Было доказано, что если S удовлетворяет следующим условиям:

- $S(1) = 0$;
- Если $p < q$, то $S(p) > S(q)$;
- S непрерывна на $[0,1]$;
- $S(pq) = S(p) + S(q)$;
- $S(1/2) = 1$;

то, с необходимостью, $S(p) = -\log_2 p$. В дальнейшем основание логарифма будем опускать.

Для последующего изложения введем некоторые обозначения. Пусть $\boldsymbol{X}$ и $\boldsymbol{Y}$ — случайные элементы, заданные на одном вероятностном пространстве и принимающие лишь конечное число значений. Положим

$$p(\boldsymbol{x}) = \mathbf{P}\left(\boldsymbol{X} = \boldsymbol{x}\right), \quad p(\boldsymbol{x} \mid E) = \mathbf{P}\left(\boldsymbol{X} = \boldsymbol{x} \mid E\right), \quad p(\boldsymbol{x} \mid \boldsymbol{y}) = \mathbf{P}\left(\boldsymbol{X} = \boldsymbol{x} \mid \boldsymbol{Y} = \boldsymbol{y}\right).$$

Определение 1. *Энтропия $H(\boldsymbol{X})$ случайного элемента $\boldsymbol{X}$ задается формулой*

$$H(\boldsymbol{X}) = \sum_{\boldsymbol{x}} -p(\boldsymbol{x}) \log p(\boldsymbol{x}),$$

где сумирование ведется по всем $\boldsymbol{x}$, лежащим в области значений $\boldsymbol{X}$.

Свойство 1. *Для любого случайного элемента $\boldsymbol{X}$, принимающего конечное число значений,*

$$H(\boldsymbol{X}) \leq \log |\operatorname{range}(\boldsymbol{X})|,$$

где $\operatorname{range}(\boldsymbol{X})$ *— множество значений, которые $\boldsymbol{X}$ принимает с положительной вероятностью, при этом, равенство достигается на равномерном распределении.*

Определение 2. *Условная энтропия $H(\boldsymbol{X} \mid E)$ случайного элемента $\boldsymbol{X}$ относительно события E задается формулой*

$$H(\boldsymbol{X} \mid E) = \sum_{\boldsymbol{x}} -p(\boldsymbol{x} \mid E) \log p(\boldsymbol{x} \mid E),$$

а относительно случайного элемента $\boldsymbol{Y}$ формулой

$$H(\boldsymbol{X} \mid \boldsymbol{Y}) = E_{\boldsymbol{Y}}(H(\boldsymbol{X} \mid \{\boldsymbol{Y} = \boldsymbol{y}\})) = \sum_{\boldsymbol{y}} p(\boldsymbol{y}) \sum_{\boldsymbol{x}} p(\boldsymbol{x} \mid \boldsymbol{y}) \log p(\boldsymbol{x} \mid \boldsymbol{y}).$$

Если случайный элемент $\boldsymbol{X}$ принимает значения в пространстве $\mathbb{R}^n$, мы будем писать $H(X_1, \ldots, X_n)$ для обозначения его энтропии.

Свойство 2. *(Цепное правило) Для случайных величин $X_1, \ldots, X_n$, принимающих конечное число значений,*

$$H\left(X_1, \ldots, X_n\right)=\sum_{i=1}^n H\left(X_i \mid X_1, \ldots, X_{i-1}\right).$$

Свойство 3. *(Отбрасывание условия) Для случайных величин X, Y, Z, принимающих конечное число значений,*

$$H(X \mid Y, Z) \leq H(X \mid Y) \leq H(X).$$

Свойство 4. *(Субаддитивность) Для случайных величин $X_1, \ldots, X_n$, принимающих конечное число значений,*

$$H\left(X_1, \ldots, X_n\right) \leq \sum_{i=1}^n H\left(X_i\right).$$

Лемма 1. *(лемма Shearer's) Пусть $\mathcal{F}$ — некоторое семейство подмножеств (возможно с повторениями) множества $\{1, \ldots, n\}$ такое, что каждое число $i \in \{1, \ldots, n\}$ входит, по крайней мере, в t элементов семейства $\mathcal{F}$. Для случайного вектора $(X_1, \ldots, X_n)$, принимающего конечное число значений,*

$$H\left(X_1, \ldots, X_n\right) \leq \frac{1}{t} \sum_{F \in \mathcal{F}} H\left(\boldsymbol{X}_{\boldsymbol{F}}\right)$$

где $\boldsymbol{X}_{\boldsymbol{F}}$ — вектор $(X_i, i \in F)$

В качестве примера применения энтропии к задачам пересчета была разобрана

Теорема 1. *Для любого n и для фиксированного $\alpha = k/n \leq 1/2$,*

$$\sum_{i \leq \alpha n}\binom{n}{i} \leq 2^{H(\alpha) n}.$$