Неравенство Йенсена и утверждение из теории информации

Пусть случайная величина K произвольно распределена на множестве $\mathcal{K}$.

Воспоминания о выпуклостях

Пусть функция φ выпукла вниз на отрезке $[a,\ b]$, то есть
$a \le x < y \le b:\ \forall\ z \in [x,\ y]:\ \varphi(z) \le \varphi(x) + \frac{\varphi(y)-\varphi(x)}{y-x}(z-x)$.

Неравенство Йенсена

Если функция $f(x)$ выпукла на $[a,b]$, то для любого стохастического вектора

$$\vec{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_k), \quad \sum_{i=1}^{k} \alpha_i = 1, \quad \alpha_i \ge 0,$$

и любых $x_1, x_2, \dots, x_k \in [a,b]$ верно неравенство:

$$f\left(\sum_{i=1}^{k} \alpha_i x_i\right) \le \sum_{i=1}^{k} \alpha_i f(x_i).$$

Доказательство

База индукции для $k=1$:

$$f(x_1) = f(x_1).$$

База индукции для $k=2$:
Нужно показать, что для любого $\alpha_1, \alpha_2 \ge 0$, $\alpha_1 + \alpha_2 = 1$, и $x_1, x_2 \in [a,b]$,

$$f(\alpha_1 x_1 + \alpha_2 x_2) \le \alpha_1 f(x_1) + \alpha_2 f(x_2).$$

Пусть $x_1 < x_2$, тогда

$$x = \alpha_1 x_1 + \alpha_2 x_2,$$

и очевидно, что $x_1 < x < x_2$.
Обозначим $\alpha_1 = \alpha$, тогда $\alpha_2 = 1 - \alpha$.
Так как f выпукла, то:

$$f(x) \le f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1).$$

Подставляя $x - x_1 = (1-\alpha)(x_2 - x_1)$, получаем

$$f(x) \le f(x_1) + (f(x_2) - f(x_1))(1-\alpha) = \alpha f(x_1) + (1-\alpha) f(x_2).$$

Индукционный переход:
Пусть неравенство доказано для k, нужно доказать для $k+1$:

$$f\left(\sum_{i=1}^{k+1} \alpha_i x_i\right) \le \sum_{i=1}^{k+1} \alpha_i f(x_i).$$

Обозначим:

$$y_i = x_i, \quad i = 1, \dots, k-1,$$

$$y_k = \frac{\alpha_k x_k + \alpha_{k+1} x_{k+1}}{\alpha_k + \alpha_{k+1}},$$

и

$$\beta_i = \alpha_i, \quad i = 1, \dots, k-1, \quad \beta_k = \alpha_k + \alpha_{k+1}.$$

Тогда $\vec{\beta} = (\beta_1, \dots, \beta_k)$ — стохастический вектор.

По предположению индукции,

$$f\left(\sum_{i=1}^{k} \beta_i y_i\right) \leq \sum_{i=1}^{k} \beta_i f(y_i).$$

По базе индукции для $k = 2$ и выпуклости f,

$$f(y_k) = f\left(\frac{\alpha_k x_k + \alpha_{k+1} x_{k+1}}{\alpha_k + \alpha_{k+1}}\right) \leq \frac{\alpha_k}{\alpha_k + \alpha_{k+1}} f(x_k) + \frac{\alpha_{k+1}}{\alpha_k + \alpha_{k+1}} f(x_{k+1}).$$

Подставляя обратно, получаем:

$$\sum_{i=1}^{k} \beta_i f(y_i) = \sum_{i=1}^{k-1} \alpha_i f(x_i) + (\alpha_k + \alpha_{k+1}) f(y_k) \leq \sum_{i=1}^{k+1} \alpha_i f(x_i).$$

Знак ■ — доказательство завершено.

Утверждение 3

$$\log P \geq -I(K \leftrightarrow C).$$

Доказательство

Пусть $c \in \mathcal{C}$ — допустимый код, если

$$D(c, k) \in \mathcal{M},$$

где k — ключ, используемый A и B.

Вероятность того, что c допустим:

$$0 \leq P(c \text{ допустим}) = \sum_{k \in \mathcal{K}} P(c \text{ допустим} \mid K = k) P(K = k).$$

Поскольку при известном ключе $P(c \text{ допустим} \mid K = k)$ либо 0, либо 1, определим индикатор

$$\delta(c, k) = \begin{cases} 1, & D(c, k) \in \mathcal{M}, \\ 0, & \text{иначе.} \end{cases}$$

Тогда

$$P(c \text{ допустим}) = \sum_{k \in \mathcal{K}} P(K = k) \delta(c, k).$$

Определим вероятностное распределение

$$Q_c(k) = \frac{P(K = k) \delta(c, k)}{P(c \text{ допустим})}.$$

Вектор $(Q_c(k))_{k \in \mathcal{K}}$ — стохастический.

Вероятность $P(C = c)$ можно переписать как

$$P(C = c) = \sum_{k \in \mathcal{K}} P(C = c \mid K = k) P(K = k) = \sum_{k \in \mathcal{K}} P(C = c \mid K = k) P(K = k) \delta(c, k),$$

так как домножение на $\delta(c, k)$ не изменяет сумму.

Подставляя $Q_c(k)$, получаем

$$P(C = c) = P(c \text{ допустим}) \sum_{k \in \mathcal{K}} P(C = c \mid K = k) Q_c(k).$$

Рассмотрим выражение

$$P(C = c) \log P(C = c) = P(C = c) \log P(c \text{ допустим}) + P(C = c) \log \left(\sum_{k \in \mathcal{K}} P(C = c \mid K = k) Q_c(k)\right).$$

Распишем второй член подробнее:

$$= P(C=c)\log P(c \text{ допустим}) + P(c \text{ допустим}) \sum_{k\in\mathcal{K}} P(C=c \mid K=k) Q_c(k) \log\left(\sum_{k\in\mathcal{K}} P(C=c \mid K=k) Q_c(k)\right).$$

Функция $t \mapsto t\log t$ выпукла вниз, поэтому применяем неравенство Йенсена:

$$P(C=c)\log P(C=c) \le P(C=c)\log P(c \text{ допустим}) + P(c \text{ допустим}) \sum_{k\in\mathcal{K}} Q_c(k) P(C=c \mid K=k) \log P(C=c \mid K=k).$$

Подставляя $Q_c(k)$, получаем:

$$P(C=c)\log P(C=c) \le P(C=c)\log P(c \text{ допустим}) + \sum_{k\in\mathcal{K}} \delta(c,k) P(K=k) P(C=c \mid K=k) \log P(C=c \mid K=k).$$

Так как

$$P(K=k)P(C=c \mid K=k) = P(K=k, C=c),$$

а $\delta(c,k)$ можно опустить, поскольку если оно равно 0, то $P(C=c \mid K=k)=0$,
то

$$P(C=c)\log P(C=c) \le P(C=c)\log P(c \text{ допустим}) + \sum_{k\in\mathcal{K}} P(K=k, C=c) \log P(C=c \mid K=k).$$

Просуммируем по всем $c \in \mathcal{C}$:

$$\sum_{c\in\mathcal{C}} P(C=c)\log P(C=c) \le \sum_{c\in\mathcal{C}} P(C=c)\log P(c \text{ допустим}) + \sum_{c\in\mathcal{C}}\sum_{k\in\mathcal{K}} P(K=k, C=c) \log P(C=c \mid K=k).$$

Используя обозначения энтропий:

$$\sum_{c\in\mathcal{C}} P(C=c)\log P(C=c) = -H(C),$$

$$\sum_{c\in\mathcal{C}}\sum_{k\in\mathcal{K}} P(K=k, C=c) \log P(C=c \mid K=k) = -H(C \mid K),$$

получаем

$$-H(C) \le \sum_{c\in\mathcal{C}} P(C=c)\log P(c \text{ допустим}) - H(C \mid K).$$

Оценим $\log P(c \text{ допустим})$:

$$\log P(c \text{ допустим}) \le \max_{c\in\mathcal{C}} \log P(c \text{ допустим}),$$

а так как

$$\sum_{c\in\mathcal{C}} P(C=c) = 1,$$

получаем

$$-H(C) \le \max_{c\in\mathcal{C}} \log P(c \text{ допустим}) - H(C \mid K).$$

Обозначим

$$P = \max_{c\in\mathcal{C}} P(c \text{ допустим}).$$

Тогда

$$H(C \mid K) - H(C) \le \log P.$$

Применяя цепное правило и теорему о взаимной информации, получаем

$$I(K \leftrightarrow C) = H(C) - H(C \mid K),$$

откуда следует

$$\log P \ge -I(K \leftrightarrow C).$$

■

Смысл утверждения

Чтобы уменьшить вероятность имитационной ошибки P, необходимо увеличить взаимную информацию

$$I(K \leftrightarrow C).$$

Эта величина отражает, в какой степени ключ используется для защиты от атаки имитации.

Определение

Шифр обладает совершенной имитационной стойкостью, если

$$\log P = -I(K \leftrightarrow C) \quad \Rightarrow \quad P = \left(\frac{1}{2}\right)^{I(K \leftrightarrow C)}.$$