Практика №2 (6.10.2023)

1. Неопределенность информации и энтропия.

1. [Che13 1.4] В урне r красных шаров, b черных и w белых. Случайный эксперимент X состоит в том, что выбирают $k \geq 2$ шаров и фиксируют их цвета. В каком случае величина $H(X)$ больше - при выборе с возвращением или без возвращения?

2. [Che13 1.23] Пусть случайные величины X и Y независимы и имеют распределения:
$X \sim \begin{array}{|c|c|c|c|} \hline 0 & 1 & 3 & 4 \\ \hline \frac{1}{8} & \frac{1}{8} & \frac{1}{4} & \frac{1}{2} \\ \hline \end{array} \quad Y \sim \begin{array}{|c|c|} \hline -2 & 2 \\ \hline \frac{3}{8} & \frac{5}{8} \\ \hline \end{array}$
при этом $Z = X + Y$. Вычислить энтропию $H(Z|Y)$.

3. [Che13 1.23] Пусть X и Y - зависимые случайные величины, $H(X) = 8$ бит, $H(Y) = 12$ бит. Какие значения может принимать $H(Y|X)$, если $H(X|Y)$ изменяется в максимально возможных пределах?

4. Пусть X - случайная величина, имеющее распределение Бернулли, т.е. $P(X = 1) = p$, $P(X = 0) = 1 - p$. Энтропия такой случайной величины выражается формулой

$$H(X) = -p \log_2 p - (1 - p) log_2(1 - p)$$

- Вычислить $H(1/4)$ по определению энтропии (не используя вычислительную формулу), если известно значение $\log_2 3 \approx 1.584$.

- Вычислить ожидаемое значение энтропии $H(p)$, если p равномерно распределено на отрезке $p \in [0, 1]$.

5. Пусть X и Y - случайные величины, принимающие значения из множеств $\{x1, \ldots, x_r\}$, $\{y_1, \ldots, y_s\}$ соответственно. Пусть $Z = X + Y$.

- Показать, что $H(Z|X) = H(Y|X)$. Доказать, что если X и Y независимы, то $H(Y) \leq H(Z)$ и $H(X) \leq H(Z)$. Следовательно, сложение независимых случайных величин увеличивает неопределенность.

- Привести пример (зависимых случайных величин), при котором $H(X) > H(Z)$ и $H(Y) > H(Z)$.
- В каком случае $H(Z) = H(X) + H(Y)$?

Домашнее задание

1. Известно, что женщин старше 80 лет в три раза больше, чем мужчин в той же возрастной группе. Сколько информации (в битах) несет сообщение о том, что человек старше 80 лет - мужчина?

2. Пусть имеется n дискретных случайных величин $X_1, \ldots, X_n$, энтропия случайной величины X_i равна $H(X_i)$, при этом $max\{H(X_i)\} = H(X_L)$. Каковы верхние и нижние оценки совместной энтропии $H(X_1, \ldots, X_n)$? При каких условиях эти границы достигаются?