

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО  
ОБРАЗОВАНИЯ «НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ «ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Московский институт электроники и математики им. А.Н. Тихонова

Ефремов Виктор Васильевич, группа БИТ 203

Отчет  
по домашней работе 2

по дисциплине "Информатика"  
Тема: "Математические основы вычислительной техники"

Номер варианта: 6  
Дата сдачи отчета: 11.12.2020

Москва 2020

# Задание 1

Канал связи задан двумя распределениями:  $p(y/x)$  и  $p(x)$ .

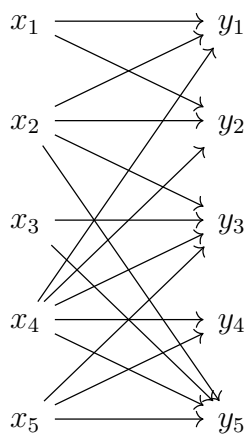
	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
$p(x)$	$\frac{4}{19}$	$\frac{6}{19}$	$\frac{4}{19}$	$\frac{2}{19}$	$\frac{3}{19}$

Таблица 1:  $p(x)$

	$y_1$	$y_2$	$y_3$	$y_4$	$y_5$
$x_1$	$\frac{8}{9}$	$\frac{1}{9}$	0	0	0
$x_2$	$\frac{3}{18}$	$\frac{5}{18}$	$\frac{4}{18}$	0	$\frac{6}{18}$
$x_3$	0	0	$\frac{2}{8}$	0	$\frac{6}{8}$
$x_4$	$\frac{9}{29}$	$\frac{9}{29}$	$\frac{1}{29}$	$\frac{6}{29}$	$\frac{4}{29}$
$x_5$	0	0	$\frac{1}{16}$	$\frac{8}{16}$	$\frac{7}{16}$

Таблица 2:  $p(y/x)$

Схематично нарисуем модель канала. Стрелки показывают что можно получить на приемнике при отправке определенного сообщения. Например, отправив  $x_1$  можно получить либо  $y_1$ , либо  $y_2$ .



Для того чтобы найти взаимную информацию вспомним несколько формул.

$$I(x, y) = H(y) - H(y/x)$$

$$H(y) = - \sum_j p(y_j) \cdot \log_2 p(y_j)$$

$$H(y/x) = - \sum_i \sum_j p(x_i, y_j) \cdot \log_2 p(y_j/x_i)$$

$$p(y_j) = \sum_i p(x_i, y_j)$$

$$p(x, y) = p(x) \cdot p(y/x)$$

Для начала посчитаем  $p(x, y)$ . Для этого нужно просто умножить каждый элемент матрицы  $p(y/x)$  на соответствующую вероятность  $p(x)$ . Например  $p(x_1, y_1) = p(x_1) \cdot p(y_1/x_1)$ ,  $p(x_1, y_2) = p(x_1) \cdot p(y_2/x_1)$ .

	$y_1$	$y_2$	$y_3$	$y_4$	$y_5$
$x_1$	$\frac{32}{171}$	$\frac{4}{171}$	0	0	0
$x_2$	$\frac{18}{342}$	$\frac{30}{342}$	$\frac{24}{342}$	0	$\frac{36}{342}$
$x_3$	0	0	$\frac{8}{152}$	0	$\frac{24}{152}$
$x_4$	$\frac{18}{551}$	$\frac{18}{551}$	$\frac{2}{551}$	$\frac{12}{551}$	$\frac{8}{551}$
$x_5$	0	0	$\frac{3}{304}$	$\frac{24}{304}$	$\frac{21}{304}$

Таблица 3:  $p(x, y)$

Найдем  $p(y_j)$ . Для этого нужно просто просуммировать элементы каждого столбца матрицы  $p(x, y)$ .

$$p(y_1) = \frac{32}{171} + \frac{18}{342} + 0 + \frac{18}{551} + 0 = \frac{1351}{4959}$$

$$p(y_2) = \frac{4}{171} + \frac{30}{342} + 0 + \frac{18}{551} + 0 = \frac{713}{4959}$$

$$p(y_3) = 0 + \frac{24}{342} + \frac{8}{152} + \frac{2}{551} + \frac{3}{304} = \frac{3605}{26448}$$

$$p(y_4) = 0 + 0 + 0 + \frac{12}{551} + \frac{24}{304} = \frac{111}{1102}$$

$$p(y_5) = 0 + \frac{36}{342} + \frac{24}{152} + \frac{8}{551} + \frac{21}{304} = \frac{3057}{8816}$$

Сделаем проверку. Сумма всех  $p(y_i)$  должна быть равна единице, так как это полная группа событий.

$$\frac{1351}{4959} + \frac{713}{4959} + \frac{3605}{26448} + \frac{111}{1102} + \frac{3057}{8816} = 1$$

	$y_1$	$y_2$	$y_3$	$y_4$	$y_5$
$p(y)$	$\frac{1351}{4959}$	$\frac{713}{4959}$	$\frac{3605}{26448}$	$\frac{111}{1102}$	$\frac{3057}{8816}$

Таблица 4:  $p(y)$

Посчитаем энтропию  $H(y)$ .

$$H(y) = - \left( \frac{1351}{4959} \cdot \log_2 \frac{1351}{4959} + \frac{713}{4959} \cdot \log_2 \frac{713}{4959} + \frac{3605}{26448} \cdot \log_2 \frac{3605}{26448} + \right. \\ \left. \frac{111}{1102} \cdot \log_2 \frac{111}{1102} + \frac{3057}{8816} \cdot \log_2 \frac{3057}{8816} \right) \approx 2.1686847$$

Посчитаем энтропию  $H(y/x)$ .

$$H(y/x) = - \left( \frac{32}{171} \cdot \log_2 \frac{8}{9} + \frac{4}{171} \cdot \log_2 \frac{1}{9} + 0 + 0 + 0 + \right. \\ \frac{18}{342} \cdot \log_2 \frac{3}{18} + \frac{30}{342} \cdot \log_2 \frac{5}{18} + \frac{24}{342} \cdot \log_2 \frac{4}{18} + 0 + \frac{36}{342} \cdot \log_2 \frac{6}{18} + \\ 0 + 0 + \frac{8}{152} \cdot \log_2 \frac{2}{8} + 0 + \frac{24}{152} \cdot \log_2 \frac{6}{8} + \\ \frac{18}{551} \cdot \log_2 \frac{9}{29} + \frac{18}{551} \cdot \log_2 \frac{9}{29} + \frac{2}{551} \cdot \log_2 \frac{1}{29} + \frac{12}{551} \cdot \log_2 \frac{6}{29} + \frac{8}{551} \cdot \log_2 \frac{4}{29} + \\ \left. 0 + 0 + \frac{3}{304} \cdot \log_2 \frac{1}{16} + \frac{24}{304} \cdot \log_2 \frac{8}{16} + \frac{21}{304} \cdot \log_2 \frac{7}{16} \right) \approx 1.3137435$$

Взяв разность найденных энтропий получим взаимную информацию.

$$I(x, y) \approx 2.1686847 - 1.3137435 \approx 0.8549412 \text{ бит}$$

Посмотрим что будет при другом распределении вероятностей над входным алфавитом. Чтобы отличать новые распределения от старых будем обозначать их  $\tilde{x}$  и  $\tilde{y}$ . Пусть все сообщения источника равновероятны, т.е. все  $p(\tilde{x}_i)$  равны.

	$\tilde{x}_1$	$\tilde{x}_2$	$\tilde{x}_3$	$\tilde{x}_4$	$\tilde{x}_5$
$p(\tilde{x})$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

Таблица 5:  $p(\tilde{x})$

Аналогично предыдущему посчитаем совместную информацию  $I(\tilde{x}, \tilde{y})$ .  
Посчитаем  $p(\tilde{x}, \tilde{y})$ .

	$\tilde{y}_1$	$\tilde{y}_2$	$\tilde{y}_3$	$\tilde{y}_4$	$\tilde{y}_5$
$\tilde{x}_1$	$\frac{8}{45}$	$\frac{1}{45}$	0	0	0
$\tilde{x}_2$	$\frac{3}{90}$	$\frac{5}{90}$	$\frac{4}{90}$	0	$\frac{6}{90}$
$\tilde{x}_3$	0	0	$\frac{2}{40}$	0	$\frac{6}{40}$
$\tilde{x}_4$	$\frac{9}{145}$	$\frac{9}{145}$	$\frac{1}{145}$	$\frac{6}{145}$	$\frac{4}{145}$
$\tilde{x}_5$	0	0	$\frac{1}{80}$	$\frac{8}{80}$	$\frac{7}{80}$

Таблица 6:  $p(\tilde{x}, \tilde{y})$

Посчитаем  $p(\tilde{y})$ .

	$\tilde{y}_1$	$\tilde{y}_2$	$\tilde{y}_3$	$\tilde{y}_4$	$\tilde{y}_5$
$p(\tilde{y})$	$\frac{713}{2610}$	$\frac{73}{522}$	$\frac{2377}{20880}$	$\frac{41}{290}$	$\frac{2309}{6960}$

Таблица 7:  $p(\tilde{y})$

Посчитаем  $H(\tilde{y})$ .

$$H(\tilde{y}) = -\left( \frac{713}{2610} \cdot \log_2 \frac{713}{2610} + \frac{73}{522} \cdot \log_2 \frac{73}{522} + \frac{2377}{20880} \cdot \log_2 \frac{2377}{20880} + \frac{41}{290} \cdot \log_2 \frac{41}{290} + \frac{2309}{6960} \cdot \log_2 \frac{2309}{6960} \right) \approx 2.1923048$$

Посчитаем  $H(\tilde{y}/\tilde{x})$ .

$$\begin{aligned}
H(\tilde{y}/\tilde{x}) = & - \left( \frac{8}{45} \cdot \log_2 \frac{8}{9} + \frac{1}{45} \cdot \log_2 \frac{1}{9} + 0 + 0 + 0 + \right. \\
& \frac{3}{90} \cdot \log_2 \frac{3}{18} + \frac{5}{90} \cdot \log_2 \frac{5}{18} + \frac{4}{90} \cdot \log_2 \frac{4}{18} + 0 + \frac{6}{90} \cdot \log_2 \frac{6}{18} + \\
& 0 + 0 + \frac{2}{40} \cdot \log_2 \frac{2}{8} + 0 + \frac{6}{40} \cdot \log_2 \frac{6}{8} + \\
& \frac{9}{145} \cdot \log_2 \frac{9}{29} + \frac{9}{145} \cdot \log_2 \frac{9}{29} + \frac{1}{145} \cdot \log_2 \frac{1}{29} + \frac{6}{145} \cdot \log_2 \frac{6}{29} + \frac{4}{145} \cdot \log_2 \frac{4}{29} + \\
& \left. 0 + 0 + \frac{1}{80} \cdot \log_2 \frac{1}{16} + \frac{8}{80} \cdot \log_2 \frac{8}{16} + \frac{7}{80} \cdot \log_2 \frac{7}{16} \right) \approx 1.324153
\end{aligned}$$

Посчитаем  $I(\tilde{x}, \tilde{y})$ .

$$I(\tilde{x}, \tilde{y}) \approx 2.1923048 - 1.324153 \approx 0.8681518 \text{ бит}$$

Видно, что  $I(\tilde{x}, \tilde{y}) > I(x, y)$ . Поэтому пропускная способность  $C$  канала из задачи не меньше  $I(\tilde{x}, \tilde{y}) \approx 0.8681518$ . Но может быть и больше, т.к.  $C = \max_{p(x)} I(x, y)$  и возможно найдется распределение источника дающее ещё большую совместную информацию.

## Задание 2

Используем следующую таблицу встречаемости букв (взято в руссокой википедии).

Буква	Частота, %
а	8.01
б	1.59
в	4.54
г	1.70
д	2.98
е	8.45
ё	0.04
ж	0.94
з	1.65
и	7.35
й	1.21
к	3.49
л	4.40
м	3.21
н	6.70
о	10.97
п	2.81
р	4.73
с	5.47
т	6.26
у	2.62
ф	0.26
х	0.97
ц	0.48
ч	1.44
ш	0.73
щ	0.36
ъ	0.04
ы	1.90
ь	1.74
э	0.32
ю	0.64
я	2.01

Таблица 8: Встречаемость букв руссокого алфавита

Информация, которая содержится в последовательности, суть сумма информации каждого символа, которые равны  $-\log_2 p$ . ФИО - ефремов виктор васьевич. Для удобства счета разобьем буквы по их количеству в сообщении. в - 4; е - 3; и, о, р - 2; а, к, л, м, с, т, ф, ч, ь.

Поэтому в моем ФИО содержится

$$I = -(4 \cdot \log_2 0.0454 + 3 \cdot \log_2 0.0845 + 2 \cdot (\log_2 0.0735 + \log_2 0.01097 + \log_2 0.0473) + \\ 1 \cdot (\log_2 0.0801 + \log_2 0.0349 + \log_2 0.0440 + \log_2 0.0321 + \log_2 0.0547 + \\ \log_2 0.0626 + \log_2 0.0026 + \log_2 0.0144 + \log_2 0.0174)) \approx 104.586 \text{ бит}$$

Если закодировать все то же ФИО равномерным двоичным кодом, то на каждый символ уйдет по 6 бит, т.к.  $2^5 < 33 \leq 2^6$  и всего информации будет  $6 \cdot 23 = 138$  бит.