Министерство образования и науки РФ
Нижнекамский химико-технологический институт (филиал)
Федерального государственного бюджетного
образовательного учреждения высшего профессионального
образования
«Казанский национальный исследовательский
технологический университет»

ИНФОРМАЦИОННЫЕ ХАРАКТЕРИСТИКИ ИСТОЧНИКОВ СООБЩЕНИЙ

Практикум по курсу «Теория информации»

Нижнекамск
2015

УДК 681.3

Печатается по решению редакционно-издательского совета Нижнекамского химико-технологического института (филиал) ФГБОУ ВПО «КНИТУ».

Рецензенты:
Галеев Э.Р., кандидат технических наук,
Саримов Н.Н., кандидат физико-математических наук.

Лежнева Н.В.
Информационные характеристики источников сообщений: Практикум по курсу «Теория информации» / Н.В. Лежнева. – Нижнекамск: Нижнекамский химико-технологический института (филиал) ФГБОУ ВПО «КНИТУ», 2015. - 29 с.

Приведены основные теоретические сведения по разделу курса «Теория информация»: информационные характеристики источников дискретных и непрерывных сообщений. Применение теории проиллюстрировано решением конкретных практических задач, а также приведены практические задания для самостоятельного решения.

Предназначены для студентов, обучающихся по направлениям: «Информатика и вычислительная техника», «Управление в технических системах», «Автоматизация технологических процессов и производств».

УДК 681.3

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Любая информационная система предназначена для передачи информации с заданными свойствами от различных источников к различным получателям. Важнейшей задачей системы, особенно системы реального времени, является обеспечение максимальной скорости передачи информации при высоком качестве функционирования и экономичности. Под качеством функционирования при этом понимают минимизацию потерь информации в системе. Для повышения качества функционирования системы передачи информации необходимо согласование статистических свойств источников сообщений и каналов связи.

Для того, чтобы установить потенциальные возможности системы по передаче информации вводят информационные характеристики сообщений и каналов связи (энтропия, количество информации при равновероятных и неравновероятных, взаимонезависимых и взаимозависимых сообщениях, скорость передачи информации, пропускная способность каналов связи и.т.п.). Поэтому возникает задача оценки количества информации в передаваемом сообщении, а также других информационных характеристик сообщения. Знание информационных характеристик позволяет определить пути повышения эффективности системы передачи информации.

Для определения энтропии, количества информации и других информационных характеристик используются вероятностные характеристики сообщения, которые не связаны с конкретным содержанием сообщения, а отражают степень их неопределенности.

I. ИНФОРМАЦИОННЫЕ ХАРАКТЕРИСТИКИ ДИСКРЕТНЫХ ИСТОЧНИКОВ СООБЩЕНИЙ

Совокупность сведений об окружающем мире, являющиеся объектом хранения, передачи и преобразования, называются *информацией* (в узко практическом смысле). Информация поступает в систему в форме сообщений. *Сообщение* – это совокупность знаков или первичных сигналов, содержащих информацию, т.е. сообщение – это информация, представленная в какой-либо форме [1-3]. Различают дискретные и непрерывные сообщения. Дискретные сообщения формируются в результате последовательной выдачи источником отдельных элементов - знаков. Множество различных знаков называется *алфавитом источника сообщений*, а число знаков – *объемом алфавита*. Непрерывные сообщения не разделимы на элементы, они описываются функциями времени, принимающими непрерывное множество значений. В зависимости от формы создаваемых сообщений различают источники *дискретных* и *непрерывных сообщений*.

1.1. Методические указания по теоретической части

Если источник сообщений в каждый момент времени может случайным образом принять одно из конечного множества возможных состояний, то такой источник сообщений называется *дискретным*. Одни состояния выбираются источником чаще, другие реже, поэтому он характеризуется дискретным ансамблем X, т.е. полной совокупностью состояний с вероятностями их появления: $\sum_{i=1}^{n} p(x_i) = 1$, где $p(x_i)$ - вероятность состояния x_i, n – объем алфавита источника сообщений.

В каждом сообщении содержится для его получателя определенная информация – совокупность сведений о состоянии дискретного источника сообщений. Мера неопределенности выбора дискретным источником сообщений состояния из ансамбля X называется *энтропией дискретного источника* или *энтропией конечного ансамбля*:

$$H(X) = -\sum_{i=1}^{n} p(x_i) log_2\ p(x_i).$$

Неопределенность, приходящаяся на символ первичного алфавита, составленного из равновероятных и взаимнонезависимых символов, определяется по формуле:

$$H(X) = log_2 n.$$

Свойства энтропии дискретного источника сообщений:

1) энтропия является вещественной, ограниченной и неотрицательной величиной;

2) энтропия минимальна и равна нулю, если состояние источника полностью определено;

3) энтропия максимальна, если все состояния источника равновероятны;

4) энтропия объединения нескольких статистически независимых источников сообщений u, v, …z равна сумме энтропий исходных источников (свойство аддитивности):

$$H(U, V, \ldots, Z) = H(U) + H(V) + \ldots + H(Z).$$

Количество информации, содержащейся в дискретном сообщении, измеряется величиной исчезнувшей (снятой) при получении сообщения неопределенности. Количество информации равно

$$I = k \cdot H \text{ бит},$$

где H- средняя энтропия, приходящаяся на одно сообщение, k- общее число сообщений.

При равновероятных и взаимонезависимых символах первичного алфавита количество информации в k сообщениях алфавита объема m:

$$I = k\log_2 m \text{ бит},$$

а при неравновероятных символах:

$$I = -k\sum_{i=1}^{m} p(x_i)\log_2 p(x_i) \text{ бит/символ.}$$

Объем информации находится по формуле:

Q = k·l,

где 1 - средняя длина кодовых слов вторичного алфавита. Объем информации для равномерных кодов

$$I = k\log_2 n,$$

где n- длина кода.

В том случае, когда источники сообщений статистически зависимы, возникает понятие условной энтропии.

Условная энтропия ансамбля Y относительно ансамбля X:

$$H(X/Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j)\log_2 p(x_i/y_j),$$

$$H(X/Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(y_j)p(x_i/y_j)\log_2 p(x_i/y_j),$$

где n – объем алфавита источника x, m – объем алфавита источника y, $p(x_i/y_j)$- условная вероятность реализации состояния x_i ансамбля X при условии, что реализовалось состояние y_j ансамбля Y; $p(y_j)$- вероятность реализации состояния y_j ансамбля Y; $p(x_i, y_j)$ - вероятность совместной реализации взаимозависимых состояний x_i и y_j.

Аналогично, условная энтропия ансамбля Y по отношению к ансамблю X:

$$H(Y/X) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j)\log_2 p(y_j/x_i),$$

$$H(X/Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i)p(y_j/x_i)\log_2 p(y_j/x_i),$$

где $p(y_j/x_i)$ - условная вероятность реализации состояния y_j ансамбля Y при условии, что реализовалось состояние x_i ансамбля X; $p(x_i)$ - вероятность реализации состояния x_i ансамбля X.

Условная энтропия в теории информации используется для определения информационных потерь при передаче информации по каналу связи с шумами. Потери информации при передаче k символов по данному каналу связи: $\Delta I = kH(X/Y)$.

Для вычисления энтропии совместного появления статистически зависимых сообщений используется понятие энтропии объединения. Энтропия объединения в случае статистической взаимосвязи между символами кодируемого алфавита:

$$H(X,Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j)\log_2 p(x_i, y_j).$$

Энтропия объединения и условная энтропия связаны между собой следующими соотношениями:

H(X,Y) = H(X) + H(Y/X) = H(Y) + H(X/Y).

Количество переданной или принятой информации при наличии помех в канале связи определяется по формулам:

I(X,Y) = H(X) - H(X/Y) = H(Y) - H(Y/X)= H(X) + H(Y) - H(X,Y)=

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j)\log_2 \frac{p(x_i, y_j)}{p(x_i)p(y_j)},$$

где условные энтропии H(X/Y), H(Y/X) характеризуют потери информации в канале связи.

При отсутствии или низком уровне помех в канале связи условная энтропия H(X/Y)=0 и, следовательно, количество информации, содержащееся в принятом сообщении ансамбля, равно энтропии передаваемых сообщений ансамбля, т.е. *I(X,Y)=H(X)*.

1.2. Методические указания по практической части [4, 5]

При решении практических заданий по данной теме удобно использовать таблицу значений функции $-p\log_2 p$, приведенной в приложении.

Задание 1. Определить количество информации в одном сообщении, если известно:

а) максимально возможное количество сообщений N;

б) количество качественных признаков m, из которых составлены сообщения, а также количество символов n в каждом сообщении.

Решение: а) Если сообщения равновероятны, то количество информации в одном сообщении: $H = \log_2 N$;

б) Количество информации в одном сообщении: $I = n\log_2 m,$ а в N сообщениях: $I = N\,n\log_2 m.$ Например, сообщения передаются в двоичном коде, по десять символов в каждом, т.е. m=2, n=10, всего передается десять сообщений (N=10), тогда количество информации в десяти сообщениях $I = 10\cdot 10\log_2 2 = 100$ бит.

Задание 2. Определить объем и количество информации в сообщении "Теория информации", переданном 7-элементным телеграфным кодом.

Решение: Число принятых символов, включая пробел, k = 17. Объем информации Q = 17·7 = 119 бит. Количество информации для равновероятного алфавита $I = k \cdot H = 17 \cdot \log_2 32 = 17 \cdot 5 = 85$ бит.

Задание 3. Источник формирует трехбуквенные сообщения из алфавита A, B, C. Определить возможные сообщения источника и их количество, а также количество информации, приходящейся на одно сообщение.

Решение: Количество возможных сообщений источника $N = 3^3 = 27$. Возможные сообщения источника:

1) AAA;	10) BAA;	19) CAA;
2) AAB;	11) BAB;	20) CAB;
3) AAC;	12) BAC;	21) CAC;
4) ABA;	13) BBA;	22) CBA;
5) ABB;	14) BBB;	23) CBB;
6) ABC;	15) BBC;	24) CBC;
7) ACA;	16) BCA;	25) CCA;
8) ACB;	17) BCB;	26) CCB;
9) ACC;	18) BCC;	27) CCC.

Количество информации, приходящееся на одно сообщение, $I = \log_2 N = \log_2 27 \approx 4.75$ бит.

Задание 4. Определить максимальную энтропию следующих систем:

а) из двух элементов, каждый из которых может находиться в двух состояниях;

б) из трех элементов, каждый из которых может находиться в трех состояниях;

в) из четырех элементов, каждый из которых может находиться в трех состояниях.

Задание 5. Какое количество информации приходится на букву русского алфавита?

Задание 6. Алфавит источника сообщений состоит из букв x, y, z, w. Определить количество информации на символ сообщения, если:

а) символы появляются в сообщении равновероятно;

б) вероятности появления букв равны 0.2, 0.3, 0.4, 0.1 соответственно.

Решение: Количество информации, приходящееся на один символ алфавита, – это энтропия данного алфавита. Следовательно,

а) $H(X) = \log_2 4 = 2$ бит/символ;

б)

$$H(X) = -\sum_{i=1}^{4} p_i \log_2 p_i = -0.2\log_2 0.2 - 0.3\log_2 0.3 - 0.4\log_2 0.4 -$$
$$-0.1\log_2 0.1 \approx 1.85 \text{ бит/символ.}$$

Задание 7. Найти максимально возможную энтропию системы, состоящей из десяти элементов, каждый из которых может быть в одном из восьми равновероятном состоянии.

Задание 8. Определить энтропию дискретного источника сообщений, характеризующегося следующими вероятностями появления символов на его выходе:

а) $X = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ 0.3 & 0.1 & 0.25 & 0.15 & 0.2 \end{pmatrix}$;

б) $X = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ 0.4 & 0.1 & 0.3 & 0.2 \end{pmatrix}$.

Задание 9. Вероятность появления сигнала на выходе канала связи равна p, а вероятность непоявления соответственно q = 1-

p. При каком p неопределенность появления или непоявления сигнала максимальна?

Задание 10. Определить энтропию, содержащуюся в изображении, при условии, что последнее разлагается на 625 строк по 840 элементов в каждой строке. Яркость каждого элемента передается восемью квантованными уровнями, а также яркости разных элементов некоррелированны.

Задание 11. В двух урнах имеется по 15 шаров, причем в первой урне - 5 красных, 7 белых и 3 черных, а во второй - 4, 4 и 7 соответственно. Из каждой урны вынимается по одному шару. Определить, в случае с какой из урн исход опыта является более определенным.

Задание 12. По заданным энтропиям ансамблей X и Y, т.е. H(X) и H(Y) и условной энтропии ансамбля X по отношению к ансамблю Y (H(X/Y)) определить среднюю условную энтропию H(X/Y) ансамбля Y по отношению к ансамблю X.

Задание 13. Сигнал формируется в виде двоичного кода с вероятностями появления символов 1 и 0, равными соответственно 0.6 и 0.4. Появление любого из символов взаимосвязано условными вероятностями

$$P(A/B)=\begin{pmatrix}0.1 & 0.9\\ 0.9 & 0.1\end{pmatrix}.$$

Определить условную энтропию.

Задание 14. Сообщения передаются по каналу связи в двоичном коде. Определить энтропию сообщений, если

а) вероятности появления 0 и 1 равны 0.8 и 0.2 соответственно, а влиянием помех в канале связи можно пренебречь, т.е. условные вероятности трансформации 0 в 1 или 1 в 0 равны нулю;

б) вероятности появления 0 и 1 равны, а значения условных вероятностей, обусловленных действиями помех в канале связи, следующие: p(0/0)= 0.75, p(0/1)= 0.25, p(1/0)= 0.3, p(1/1)= 0.7.

Задание 15. Влияние помех в канале связи описывается следующей матрицей:

$$P(B/A)=\begin{pmatrix} 0.01 & 0.98 & 0.01 \\ 0.75 & 0.1 & 0.15 \\ 0.3 & 0.2 & 0.5 \end{pmatrix},$$

вероятности появления символов на выходе источника сообщений равны $p(a_1)=0.2,\ p(a_2)=0.1,\ \ p(a_3)=0.7.$

Определить энтропию приемника сообщений.

Решение: Энтропия приемника сообщений:

$H(B)=-\sum_{j=1}^{3} p(b_j)\log_2 p(b_j),$ где $p(b_j)$ - вероятность приема символа b_j, равная $p(b_j)=\sum_{i=1}^{3} p(a_i)\log_2 p(b_j/a_i),$ где $p(b_j/a_i)$ – условная вероятность приема символа b_j при условии передачи символа a_i.

Следовательно,

$$p(b_1)= p(a_1)p(b_1/a_1)+ p(a_2)p(b_1/a_2)+ p(a_3)p(b_1/a_3)=0.2\cdot 0.01+$$
$$+0.1\cdot 0.75+0.7\cdot 0.3=0.002+0.075+0.21=0.287;$$

$$p(b_2)= p(a_1)p(b_2/a_1)+ p(a_2)p(b_2/a_2)+ p(a_3)p(b_2/a_3)=0.2\cdot 0.98+$$
$$+0.1\cdot 0.1+0.7\cdot 0.2=0.196+0.01+0.14=0.346;$$

$$p(b_3)= p(a_1)p(b_3/a_1)+ p(a_2)p(b_3/a_2)+ p(a_3)p(b_3/a_3)=0.2\cdot 0.01+$$
$$+0.1\cdot 0.15+0.7\cdot 0.5=0.002+0.015+0.35=0.367;$$

$$H(B)=-0.287\log_2 0.287-0.346\log_2 0.346-0.267\log_2 0.267\approx 1.577$$

бит/символ.

Задание 16. Две взаимозависимые системы X и Y объединены в одну, матрица вероятностей состояний которой имеет вид:

$$P(X,Y)=\begin{pmatrix}0.3 & 0 & 0\\ 0.2 & 0.3 & 0.1\\ 0 & 0.1 & 0\end{pmatrix}.$$

Определить полные условные энтропии H(X/Y), H(Y/X).

Решение: Безусловные вероятности систем X и Y:

$p(x_i)=\sum_{j=1}^{3} p(x_i,y_j)$, $\quad p(y_j)=\sum_{i=1}^{3} p(x_i,y_j)$, т.е. безусловные вероятности находятся суммированием совместных вероятностей по строкам и столбцам матрицы P(X,Y):

$$\begin{array}{rcc} & & p(x_i)\\ P(X,Y)= & \begin{pmatrix}0.3 & 0 & 0\\ 0.2 & 0.3 & 0.1\\ 0 & 0.1 & 0\end{pmatrix} & \begin{matrix}0.3\\ 0.6\\ 0.1\end{matrix}\\ p(y_j) & \begin{matrix}0.5 & 0.4 & 0.1\end{matrix} & \end{array}$$

Условные вероятности рассчитываем по формулам: $p(x_i/y_j)=p(x_i,y_j)/p(y_j)$, $\;p(y_j/x_i)=p(x_i,y_j)/p(x_i)$.

Следовательно,

$p(x_1/y_1)=0.3/0.5=0.6$, $p(x_2/y_1)=0.2/0.5=0.4$,

$p(x_3/y_1)=0/0.5=0$, $p(x_1/y_2)=0/0.4=0$,

$p(x_2/y_2)=0.3/0.4=0.75$, $p(x_3/y_2)=0.1/0.4=0.25$,

$p(x_1/y_3)=p(x_3/y_3)=0/0.1=0$, $\;p(x_2/y_3)=0.1/0.1=1$.

Таким образом, $P(X/Y)=\begin{pmatrix}0.6 & 0 & 0\\ 0.4 & 0.75 & 1\\ 0 & 0.25 & 0\end{pmatrix}.$

Полная условная энтропия:

$$H(X/Y) = -\sum_{i=1}^{3}\sum_{j=1}^{3} p(y_j)p(x_i/y_j)\log_2 p(x_i/y_j) = -0.5(0.6\log_2 0.6 +$$

$$0.4\log_2 0.4) - 0.4(0.75\log_2 0.75 + 0.25\log_2 0.25) - 0.1\cdot 1\cdot \log_2 1 \approx 0.809$$
бит.

Аналогично,

$p(y_1/x_1) = 0.3/0.3 = 1,\ p(y_2/x_1) = p(y_3/x_1) = 0/0.3 = 0,$

$p(y_1/x_2) = 0.2/0.6 \approx 0.333,\ p(y_2/x_2) = 0.3/0.6 = 0.5,$

$p(y_3/x_2) = 0.1/0.6 \approx 0.167,\ p(y_1/x_3) = p(y_3/x_3) = 0/0.1 = 0,$

$p(x_1/y_3) = p(x_3/y_3) = 0/0.1 = 0,\ \ p(y_2/x_3) = 0.1/0.1 = 1.$

Следовательно, $P(X/Y) = \begin{pmatrix} 1 & 0 & 0 \\ 0.333 & 0.5 & 0.167 \\ 0 & 1 & 0 \end{pmatrix}.$

Полная условная энтропия:

$$H(Y/X) = -\sum_{i=1}^{3}\sum_{j=1}^{3} p(x_i)p(y_j/x_i)\log_2 p(y_j/x_i) = -0.3\cdot 1\cdot \log_2 1 -$$

$$-0.6(0.333\log_2 0.333 + 0.5\log_2 0.5 + 0.167\log_2 0.167) - 0.1\cdot 1\cdot \log_2 1 \approx$$

≈ 0.876 бит.

Задание 17. Взаимодействие двух систем X и Y описывается следующей матрицей:

$$P(X,Y) = \begin{pmatrix} 0.4 & 0.1 & 0 \\ 0. & 0.2 & 0.1 \\ 0 & 0 & 0.2 \end{pmatrix}.$$

Определить безусловные энтропии системы X и системы Y.

Задание 18. Канал связи описывается следующей канальной матрицей:

$$P(X,Y)=\begin{pmatrix}0.1 & 0.1 & 0\\ 0 & 0.2 & 0.1\\ 0 & 0.1 & 0.4\end{pmatrix}.$$

Определить энтропии источника и приемника сообщений, а также энтропию объединения *H(X,Y)*.

Задание 19. Определить полные условные энтропии двух взаимозависимых систем X и Y, если матрица вероятностей объединенной системы, следующая:

$$P(X,Y)=\begin{pmatrix}0.2 & 0 & 0\\ 0.1 & 0.2 & 0\\ 0 & 0.1 & 0.4\end{pmatrix}.$$

Задание 20. Определить все возможные информационные характеристики канала связи, в котором взаимосвязь источника с приемником описывается следующей матрицей:

$$P(X,Y)=\begin{pmatrix}0.2 & 0.1 & 0\\ 0.2 & 0 & 0.2\\ 0.1 & 0.1 & 0.1\end{pmatrix}.$$

Задание 21. Канал связи описывается следующей канальной матрицей:

$$P(Y/X)=\begin{pmatrix}0.98 & 0.01 & 0.01\\ 0.1 & 0.75 & 0.15\\ 0.2 & 0.3 & 0.5\end{pmatrix}.$$

Определить:

а) количество информации, которое переносится одним символом сообщения, если вероятности появления символов источника следующие: $p(x_1)=0.7,\ \ p(x_2)=0.2,\ p(x_3)=0.1;$

б) информационные потери при передаче сообщений из ста символов алфавита x_1, x_2, x_3;

в) количество принятой информации.

Решение: Энтропия источника сообщений

$$H(X) = -\sum_{i=1}^{3} p(x_i)\log_2 p(x_i) = -0.7\log_2 0.7 - 0.2\log_2 0.2 - 0.1\log_2 0.1 \approx$$

$$\approx 0.3602 + 0.4644 + 0.3322 = 1.1568 \text{ бит /символ.}$$

Общая условная энтропия:

$$H(Y/X) = -\sum_{i=1}^{3}\sum_{j=1}^{3} p(x_i)p(y_j/x_i)\log_2 p(y_j/x_i) = -0.7(0.98\cdot\log_2 0.98 +$$

$$+2\cdot 0.01\log_2 0.01) - 0.2\,(0.1\log_2 0.1 + 0.75\log_2 0.75 + 0.15\log_2 0.15) -$$

$$-0.1(0.2\log_2 0.2 + 0.3\log_2 0.3 + 0.5\log_2 0.5) \approx 0.473 \text{ бит/символ.}$$

Потери информации в канале связи:

$$\Delta I = k\cdot H(Y/X) = 100\cdot 0.473 = 47.3 \text{ бит.}$$

Энтропия приемника: $H(Y) = -\sum_{j=1}^{3} p(y_j)\log_2 p(y_j)$.

Для определения энтропии приемника найдем вероятности появления символов на входе приемника:

$$p(y_1) = \sum_{i=1}^{3} p(x_i)p(y_1/x_i) = 0.7\cdot 0.98 + 0.2\cdot 0.1 + 0.1\cdot 0.2 = 0.726,$$

$$p(y_2) = \sum_{i=1}^{3} p(x_i)p(y_2/x_i) = 0.7\cdot 0.01 + 0.2\cdot 0.75 + 0.1\cdot 0.3 = 0.187,$$

$$p(y_3) = \sum_{i=1}^{3} p(x_i)p(y_3/x_i) = 0.7\cdot 0.01 + 0.2\cdot 0.15 + 0.1\cdot 0.5 = 0.087.$$

Следовательно,

$H(Y) = -0.726\log_2 0.726 - 0.187\log_2 0.187 - 0.087\log_2 0.087 \approx 1.094$ бит/символ.

Среднее количество полученной информации:

$$I = k\left[H(Y) - H(Y/X)\right] = kH(Y) - \Delta I = 100\cdot 1.094 - 47.3 = 62.1 \text{ бит.}$$

Задание 22. Определить информационные потери в канале связи, описываемом следующей канальной матрицей:

а) $P(A/B)=\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$,

б) $P(A/B)=\begin{pmatrix} 0.99 & 0.01 & 0 \\ 0.01 & 0.98 & 0 \\ 0 & 0.01 & 0 \end{pmatrix}$

при равновероятном появлении символов алфавита в сообщениях.

Задание 23. Определить информационные потери при передаче N символов по данному каналу связи, если он описывается следующей матрицей:

$$P(A,B)=\begin{pmatrix} 0.1 & 0.1 & 0 \\ 0 & 0.2 & 0.1 \\ 0 & 0.2 & 0.3 \end{pmatrix}.$$

Задание 24. Канал связи с помехами описывается матрицей:

$$P(X,Y)=\begin{pmatrix} 0.1 & 0.1 & 0 \\ 0 & 0.2 & 0.1 \\ 0 & 0.2 & 0.3 \end{pmatrix}.$$

Определить I(X, Y).

II. ИНФОРМАЦИОННЫЕ ХАРАКТЕРИСТИКИ НЕПРЕРЫВНЫХ ИСТОЧНИКОВ СООБЩЕНИЙ

2.1. Методические указания по теоретической части

Источник сообщений, множество возможных состояний которого составляет континуум [1-3], называется *непрерывным источником.* В большинстве случаев такие источники преобразуются в дискретные с помощью устройств дискретизации и квантования, но существуют системы, в которых информация передается и преобразуется непосредственно в форме непрерывных сигналов (например, телевидение, телефонная связь и.т.д.).

Для того, чтобы оценить потенциальные возможности передачи сообщений по непрерывным каналам связи, необходимо ввести количественные характеристики непрерывных сообщений и каналов. Формулы для информационных характеристик дискретных источников сообщений, введенные в п. 1.1, не могут непосредственно использоваться для непрерывных источников, т.к. в этом случае вероятность любого конкретного значения равна нулю. Поэтому в случае непрерывных источников используются вместо вероятностей значений плотности распределения вероятностей.

Энтропия непрерывного источника сообщений:

$$H(U) = -\int_{-\infty}^{\infty} p(u)\log_2 p(u)du - \lim_{\Delta u \to 0} \log_2 \Delta u, \qquad (1)$$

где p(u)- плотность распределения вероятностей значений непрерывной случайной величины U, Δu - шаг квантования непрерывной случайной величины U (точность измерения).

Формула (1) неудобна для практического использования, поэтому обычно на практике информационные свойства непрерывного источника характеризуются величиной

$$h(U) = -\int_{-\infty}^{\infty} p(u)\log_2 p(u)du,$$

называемой *дифференциальной энтропией непрерывного источника сообщений* (непрерывного распределения случайной величины U). Она характеризует среднюю неопределенность выбора случайной величины U с произвольным законом распределения по сравнению со средней неопределенность выбора случайной величины U', равномерно распределенной в единичном интервале.

Условная энтропия непрерывного источника сообщений:

$$H(U/V) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(u,v)\log_2 \frac{p(u,v)}{p(v)}dudv - \lim_{\Delta u \to 0} \log_2 \Delta u,$$

$$h(U/V) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(u,v)\log_2 p(u/v)dudv,$$

$$h(U/V) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(u,v)\log_2 \frac{p(u,v)}{p(v)}dudv,$$

где h(U/V)- *дифференциальная условная энтропия непрерывного источника*, $p(u,v)$- совместная плотность распределения вероятностей значений непрерывных случайных величин U и V. Дифференциальная условная энтропия характеризует неопределенность выбора непрерывной случайной величины U при условии, что известны результаты реализации значений другой статистически связанной с ней непрерывной случайной величины V, и по сравнению со средней неопределенностью выбора случайной величины U', равномерно распределенной в единичном интервале.

Аналогично,

$$H(V/U) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(u,v)\log_2 \frac{p(u,v)}{p(u)}dudv - \lim_{\Delta v \to 0} \log_2 \Delta v,$$

$$h(V/U) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(u,v)\log_2 p(v/u)dudv,$$

$$h(V/U) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(u,v)\log_2 \frac{p(u,v)}{p(u)}dudv.$$

Свойства дифференциальной энтропии:

1. Дифференциальная энтропия в отличие от энтропии дискретного источника сообщений является относительной мерой неопределенности. Ее значения зависят от масштаба случайной величины U, а, следовательно, и от выбора единицы ее измерения. Из относительности дифференциальной энтропии следует, что она может принимать положительные, отрицательные и нулевые значения.
2. Дифференциальная энтропия не зависит от конкретных значений случайной величины U.
3. Максимальной дифференциальной энтропией обладают следующие непрерывные распределения:

а) если единственным ограничением для случайной величины U является область ее возможных значений [a, b], то максимальной дифференциальной энтропией обладает равномерное распределение вероятностей в этой области, т.е. $p(u) = \frac{1}{b-a}, a \le u \le b,$ причем $h_{\max}(U) = \log_2(b-a);$

б) при отсутствии ограничения на область значений непрерывной случайной величины U, но ограниченной дисперсии максимальной дифференциальной энтропией обладает нормальное распределение величины U, т.е. $p(u) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{u^2}{2\sigma^2}}$, причем $h_{\max}(U) = \log_2 \sigma\sqrt{2\pi e},$

где σ^2 – дисперсия случайной величины U (заданное ограничение), σ - среднеквадратическое отклонение U;
4. Формулы для дифференциальной энтропии объединения статистически зависимых непрерывных источников аналогичны формулам для дискретных источников:

h(U,V)= h(U)+ h(V/U)= h(V)+ h(U/V),

где $h(U,V) = -\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} p(u,v)\log_2 p(u,v)dudv.$

Количество информации, получаемой от непрерывного источника, по каналу связи с помехами:

$$I(U,V)= h(U)- h(U/V)= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} p(u,v)\log_2 \frac{p(u,v)}{p(u)p(v)}dudv.$$

2.2. Методические указания по практической части [6]

Задание 1. Найти энтропию непрерывной системы X, все состояния которой равномерно распределены в интервале [a, b], т.е

$$p(x) = \begin{cases} \dfrac{1}{b-a}, a \le x \le b \\ 0, x < a \text{ и } x > b. \end{cases}$$

Значения a и b заданы по вариантам в таблице 1.

Таблица 1. Параметры распределения

№ варианта	a	b	m_x	σ_x^2	λ
1	0	1	0	0.1	0.001
2	1	2	1	0.1	0.002
3	0	2	0	0.2	0.003
4	2	3	1	0.2	0.004
5	0	3	0	0.3	0.005
6	1	3	1	0.3	0.006
7	2	3	2	0.1	0.007
8	0	4	2	0.2	0.008
9	1	4	2	0.3	0.009
10	2	4	3	0.1	0.01
11	3	4	3	0.2	0.011
12	0	5	3	0.3	0.012
13	1	5	4	0.1	0.013
14	2	5	4	0.2	0.014
15	3	5	4	0.3	0.015

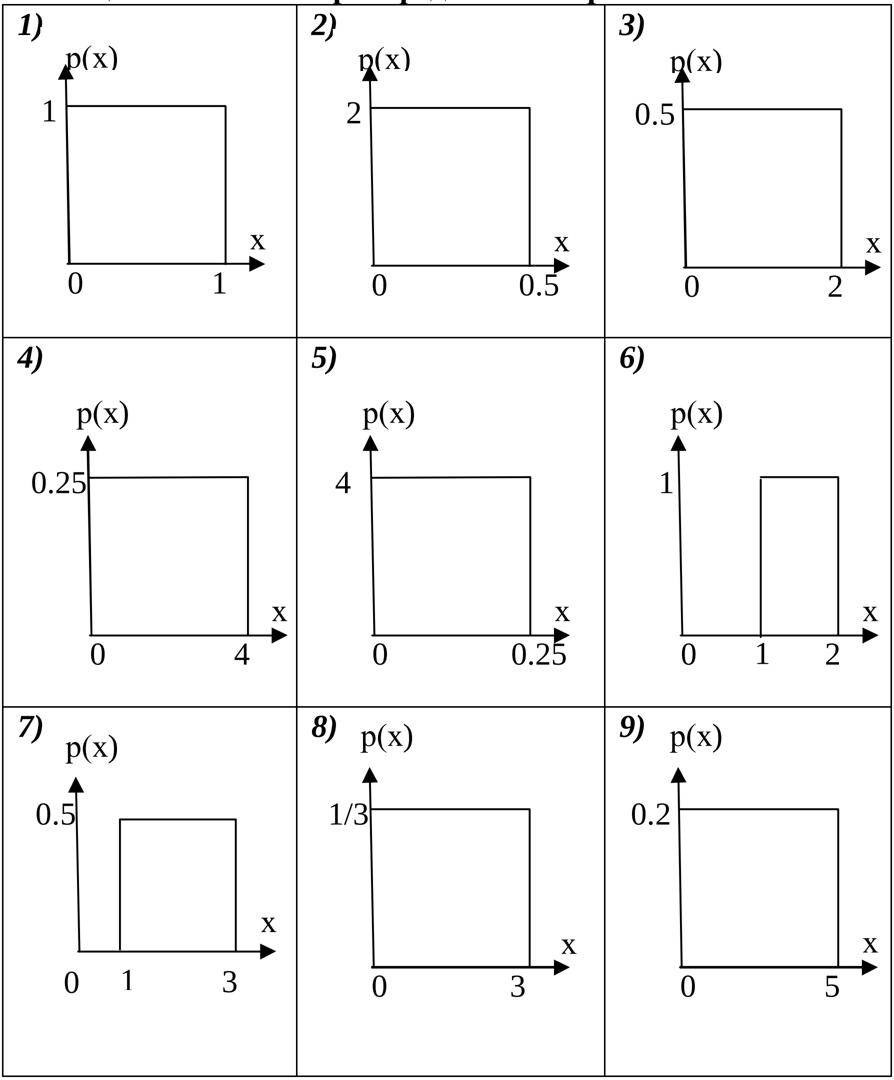
Задание 2. Найти дифференциальную энтропию системы X, состояния которой распределены по нормальному закону:

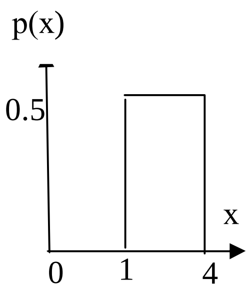
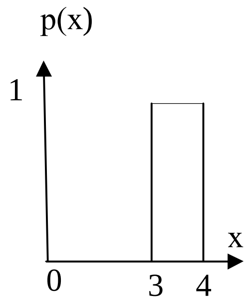
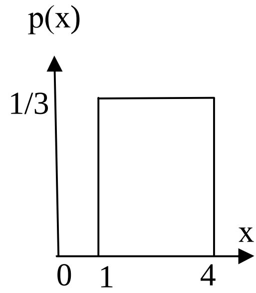
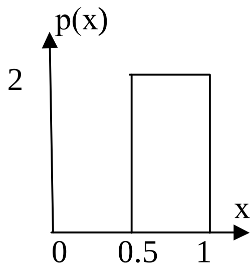
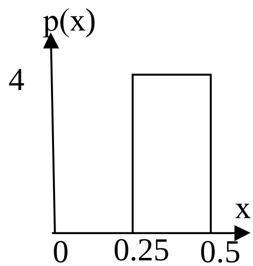
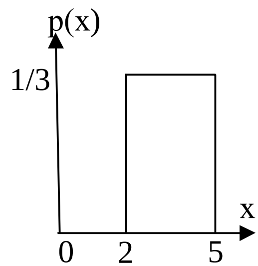
$$p(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}},$$

где m_x и σ_x^2 - параметры нормального распределения, приведенные по вариантам в таблице 1.

Задание 3. Плотность вероятности случайного процесса x(t) имеет вид: $p(x) = \lambda e^{-\lambda x}$, $x \geq 0$, где $\lambda > 0$ – параметр распределения, заданный в таблице 1. Найти дифференциальную энтропию процесса x(t).

Таблица 2. Плотности распределения вероятностей



10) 	***11)*** 	***12)*** 
13) 	***14)*** 	***15)*** 

Задание 4. Найти дифференциальную энтропию непрерывного источника, имеющего закон распределения плотности вероятности случайной величины X, приведенный в таблице 2.

Задание 5. По линии связи передаются непрерывные амплитудно-модулированные сигналы x(t), распределенные по нормальному закону со средним значением m_x В и дисперсией σ_x^2 B^2 (таблица 3). Определить энтропию сигнала при точности его измерения Δx В (таблица 3).

Задание 6. Найти энтропию непрерывной системы X, все состояния которой распределены по закону Лапласа:

$$p(x) = \frac{\alpha}{2} e^{-\alpha\left|x-\beta\right|},$$

где α>0, -∞<β<∞ - параметры распределения, заданные по варианту в таблице 3.

Таблица 3. Параметры распределения

№ варианта	m_x	σ_x^2	Δx	α	β	F_1	F_2	ΔF
1	0	2	0.1	1	0.1	1	2	0.01
2	0.5	2	0.03	2	0.01	1	4	0.02
3	0	1	0.01	3	0.015	2	4	0.03
4	0.5	1	0.02	4	0.02	10	20	0.01
5	0.2	0.2	0.01	5	0.03	5	10	1
6	0.3	0.1	0.02	1	0.04	10	15	1
7	0.3	0.2	0.03	2	0.03	2	10	2
8	1	0.5	0.02	3	0.01	2	12	5
9	2	0.3	0.02	4	0.05	10	30	5
10	2	0.3	0.05	5	0.02	15	30	3
11	1	0.2	0.03	1	0.02	10	40	3
12	4	0.5	0.04	2	0.01	20	50	4
13	4	0.4	0.04	3	0.03	25	50	5
14	5	0.3	0.04	4	0.01	30	50	2
15	5	0.5	0.05	5	0.04	20	40	4

Задание 7. Информация передается с помощью частотно-модулированных синусоидальных сигналов, рабочая частота которых изменяется с равной вероятностью от F_1 МГц до F_2 МГц (таблица 3). Определить энтропию сигнала, если точность измерения частоты составляет ΔF КГц (таблица 3).

Задание 8. Вычислить дифференциальную энтропию системы, состояния которой распределены:

а) по нормальному закону с дисперсией σ_x^2,

б) по равномерному закону с той же дисперсией σ_x^2.

Задание 9. Состояние самолета характеризуются тремя случайными величинами: высотой H, модулем скорости V и углом θ, определяющим направление самолета. Высота самолета распределена по равномерному закону на интервале от h_1 до h_2; скорость – по нормальному закону с математическим ожиданием v_0 и средним квадратическим отклонением σ_V; угол- по равномерному закону на интервале от 0 до π. Величины H, V, θ независимы. Найти энтропию объединенной системы.

Задание 10. Определить энтропию непрерывного амплитудно-модулированного сигнала, который передается по линии связи, если его значения распределены:

а) по нормальному закону;

б) равномерно в области изменения сигнала.

При этом среднее значение сигнала, дисперсия и точность отсчетов на приемной стороне Δx приведены в таблице 3.

Задание 11. Измерительное устройство вырабатывает случайный сигнал x(t) с нормальной плотностью вероятности и корреляционной функцией вида $k_x(\tau) = \sigma_x^2 e^{-\alpha|\tau|}$. Определить энтропию сигнала и его избыточность, вызванную наличием корреляции при значении σ_x^2, заданном по варианту в таблице 3.

Задание 12. Ансамбль сигналов, проходящих через усилитель, имеет значения, ограниченное сверху величиной n, а снизу m. Определить:

а) максимальную энтропию $H_{\max}(X)$;

б) энтропию H(X) на единицу времени, если ширина полосы пропускания усилителя равна F.

Приложение

Таблица значений функции $-p\log_2 p$

p	$-p\log_2 p$	p	$-p\log_2 p$	p	$-p\log_2 p$
0.01	0.0664	0.34	0,5292	0.67	0,4954
0.02	0.1128	0.35	0,5301	0.68	0,4906
0.03	0.1518	0.36	0,5306	0.69	0,4854
0.04	0.1858	0.37	0,5307	0.7	0,4800
0.05	0.2161	0.38	0,5304	0.71	0,4744
0.06	0.2435	0.39	0,5298	0.72	0,4685
0.07	0.2696	0.4	0,5288	0.73	0,4623
0.08	0,2915	0.41	0,5274	0.74	0,3215
0.09	0,3126	0.42	0,5256	0.75	0,3113
0.1	0,3322	0.43	0,5236	0.76	0,3009
0.11	0,3503	0.44	0,5210	0.77	0,2903
0.12	0,3671	0.45	0,5184	0.78	0,2796
0.13	0,3826	0.46	0,5153	0.79	0,2687
0.14	0,3971	0.47	0,5120	0.8	0,2575
0.15	0,4105	0.48	0,5083	0.81	0,2462
0.16	0,4230	0.49	0,5043	0.82	0,2348
0.17	0,4346	0.5	0,5000	0.83	0,2231
0.18	0,4453	0.51	0,4954	0.84	0,2112
0.19	0,4552	0.52	0,4906	0.85	0,1992
0.2	0,4644	0.53	0,4854	0.86	0,1871

0.21	0,4728	0.54	0,4800	0.87	0,1748
0.22	0,4806	0.55	0,4744	0.88	0,1623
0.23	0,4877	0.56	0,4685	0.89	0,1496
0.24	0,4941	0.57	0,4623	0.9	0,1368
0.25	0,5278	0.58	0,4558	0.91	0,1238
0.26	0,5053	0.59	0,4491	0.92	0,1107
0.27	0,5100	0.6	0,4422	0.93	0,0974
0.28	0,5142	0.61	0,4350	0.94	0,0839
0.29	0,5179	0.62	0,4276	0.95	0,0703
0.3	0,5211	0.63	0,4199	0.96	0,0765
0.31	0,5238	0.64	0,4121	0.97	0,0426
0.32	0,5260	0.65	0,4040	0.98	0,0286
0.33	0,5278	0.66	0,3957	0.99	0,0144

Библиографический список

Основная литература

1. Литвинская О.С. Основы теории передачи и информации/ О.С. Литвинская, Н.И. Чернышев.-М: КноРус, 2010.
2. Вентцель Е.С. Теория случайных процессов и ее инженерные приложения/ Е.С. Вентцель, Л.А. Овчаров. – М.: Высш.школа, 2010, -448 с.

Дополнительная литература

3. Дмитриев В.И. Прикладная теория информации. –М.: Высш. Шк., 1989. -320 с.
4. Цымбал В.П. Задачник по теории информации. –Киев: Вища школа, 1976. -276 с.
5. Ризаев И.С., Эминов Ф.И. Элементы теории информации и кодирования: Практическое пособие. –Казань: КГТУ, 2001. -29 с.
6. Ризаев И.С. Сборник задач по курсу «Теория информации и

кодирования»: Учебное пособие. –Казань: Казанс. авиац. институт им. А.Н.Туполева, 1976. – 51 с.