МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ (ТУСУР)

Кафедра радиотехнических систем (РТС)

УТВЕРЖДАЮ
Заведующий кафедрой РТС
_____________Г.С.Шарыгин
_____________ 2012г

ТЕОРИЯ ИНФОРМАЦИИ

Учебно - методическое пособие для проведения практических занятий и самостоятельной работы студентов по дисциплине «ТЕОРИЯ ИНФОРМАЦИИ» для студентов, обучающихся по специальностям 090103 (075300) «Организация и технология защиты информации», 090104 (075400) «Комплексная защита объектов информатизации» и 090106 (075600) «Информационная безопасность телекоммуникационных систем».

РАЗРАБОТЧИК:
Профессор кафедры РТС, д.т.н.
___________Ю.П. Акулиничев
________________2012г.

2012

Рекомендовано к изданию кафедрой радиотехнических систем Томского государственного университета систем управления и радиоэлектроники

Ю.П. Акулиничев. Теория информации: Учебно - методическое пособие для проведения практических занятий и самостоятельной работы студентов - Томск: Томский государственный университет систем управления и радиоэлектроники, 2012.- 170с.

Учебно- методическое пособие содержит методические указания по для проведению практических занятий и самостоятельной работы студентов. В каждом разделе приводятся краткие теоретические сведения и расчетные формулы, даны решения нескольких типовых примеров и подобраны задачи для самостоятельного решения.

Материал пособия соответствует программе курса "«ТЕОРИЯ ИНФОРМАЦИИ» для студентов, обучающихся по специальностям 090103 (075300) "Организация и технология защиты информации ", 090104 (075400) " Комплексная защита объектов информатизации " и 090106 (075600) "Информационная безопасность телекоммуникационных систем"., а также будет полезно для студентов, обучающихся по направлениям "Телекоммуникации"и"Радиотехника".

© Ю.П. Акулиничев
© Томский гос. ун-т систем управления и радиоэлектроники, 2012.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Активное изучение любой математической дисциплины невозможно без решения задач, и теория информации не является исключением.

Каждый из разделов учебно- методического пособия содержит основные теоретические сведения, излагаемые в конспективной форме, и подробные решения ряда типичных примеров.

Учебно- методическое пособие охватывает важнейшие направления теории информации. Главный принцип подбора материала – это показать, как результаты теории информации позволяют находить предельные, потенциальные характеристики системы передачи информации. Именно по этой причине в пособии нашли отражение основные элементы теории информации и теории кодирования.

Материал пособия соответствует программе курса “ «ТЕОРИЯ ИНФОРМАЦИИ» ” для студентов, обучающихся по специальностям 090103 (075300) Организация и технология защиты информации ”, 090104 (075400) “ Комплексная защита объектов информатизации ” и 090106 (075600) “Информационная безопасность телекоммуникационных систем”.

Пособие будет полезно для студентов, обучающихся по направлениям “Телекоммуникации”и“Радиотехника”.

1 СИГНАЛЫ И ПОМЕХИ В СИСТЕМАХ ПЕРЕДАЧИ ИНФОРМАЦИИ

Сигнал. Принято считать, что передача информации происходит тогда, когда некоторый объект (получатель информации) приобретает знания, сведения, данные о состоянии другого объекта (источника информации).

Сигналами называются состояния объектов в цепи, связывающей источник с получателем, при условии, что изменение состояния источника приводит к изменению состояний остальных объектов в этой цепи по правилам, хотя бы частично известным получателю. Такие правила называются кодом.

Следует подчеркнуть, что всегда можно указать пару крайних элементов в этой цепи сигналов: переданный сигнал (сообщение), то есть состояние источника информации, и принятый сигнал, то есть состояние объекта, доступного для непосредственного наблюдения получателем.

Состояние любого объекта можно математически описать при помощи набора чисел (параметров). Поскольку значения этих параметров обычно изменяются со временем, то математической моделью сигнала служит функция времени (и, возможно, других аргументов, например пространственных координат). Интересоваться значением передаваемого сигнала $x(t)$ и наблюдать с этой целью принимаемый сигнал $y(t)$ имеет смысл только в том случае, когда значения этих функций заранее неизвестны получателю.

Наиболее распространенной математической моделью сигнала, удовлетворяющей этому требованию и допускающей строгое количественное описание, является случайная функция. При этом конкретной реализации сигнала, возникшей в результате опыта, ставится в соответствие конкретная реализация этой случайной функции. Множество всех возможных реализаций сигнала с указанием вероятностной меры, соответствующей каждой из реализаций, называется ансамблем.

Случайная функция называется дискретной по данному параметру, если множество возможных значений этого параметра конечно или счетно. В противном случае, когда параметр может принимать любое значение из некоторой непрерывной области, функция называется непрерывной по данному параметру. Переход от непрерывного сигнала к дискретному называется квантованием (дискретизацией), а обратный переход - интерполяцией (восстановлением). Случайная функция, дискретная по времени, называется случайной последовательностью.

Типичная процедура квантования по времени заключается в том, что берут отсчеты исходной непрерывной функции в моменты времени $t_1, t_2, ..., t_n$, отстоящие один от другого на величину Δt, называемую шагом квантования по времени (рис. 1.1, б). Полученная в результате квантования случайная последовательность – система непрерывных случайных величин $X^{(1)}, X^{(2)}, ..., X^{(n)}$ в соответствии с теоремой В.А. Котельникова полностью определяет исходную случайную функцию $X(t)$, если $n \to \infty$, а шаг квантования выбран из условия $\Delta t \le 1/(2F_B)$, где F_B - верхняя граничная частота в спектре сигнала $X(t)$.

Квантование по уровню, т.е. переход от непрерывных величин $X^{(k)}$ к дискретным $X_j^{(k)}$, иллюстрируется рис. 1.1в.

В дальнейшем будем рассматривать следующие модели сообщений и сигналов:

а) непрерывная или дискретная случайная величина;

б) непрерывная или дискретная случайная последовательность, т.е. система n непрерывных или дискретных случайных величин;

в) непрерывная случайная функция, т.е. непрерывная функция времени, значения которой в любые моменты времени являются непрерывными случайными величинами.

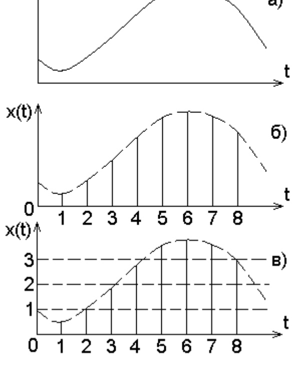


Рис. 1.1

Модель сигнала – случайная величина. Простейшая модель сигнала - случайная величина X .

Полное математическое описание дискретной случайной величины X дает ее закон распределения, т.е. таблица, в которой перечислены все возможные значения этой случайной величины $x_1, x_2, ...$ и соответствующие им вероятности $p(x_1), p(x_2), ...$. Дискретная случайная величина X называется m-ичным символом (буквой, цифрой), если множество ее возможных значений $x_1, x_2, ..., x_m$ конечно. Это множество называется алфавитом, а число m - основанием кода. Термин «символ x_j» применяется для указания конкретного j-го элемента алфавита.

Полное математическое описание непрерывной случайной величины X содержит ее функция распределения $F(x)$ либо плотность вероятности $W(x)$, если последняя существует.

Типичная процедура квантования непрерывной величины X , т.е. переход к дискретной величине X', заключается в следующем. Область возможных значений случайной величины X делят на m интервалов с длинами $\Delta x_1, \Delta x_2, ..., \Delta x_m$ и задают множество $x_1, x_2, ..., x_m$ возможных значений дискретной случайной величины X'. Обычно, хотя это не обязательно, в качестве x_j выбирают середину j-го интервала. Далее считают, что дискретная величина X' приняла значение x_j, если реализация непрерывной случайной величины X попала внутрь j-го интервала (рис. 1.2). Вероятность этого события равна

$$p(x_j) = \int_{x_j - \Delta x_j/2}^{x_j + \Delta x_j/2} W(x)dx. \qquad (1.1а)$$

Если длины интервалов настолько малы, что относительное изменение плотности $W(x)$ внутри интервала незначительно, то приближенно имеем

$$p(x_j) \approx W(x_j)\Delta x_j\,. \qquad (1.1б)$$

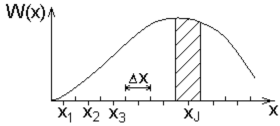


Рис. 1.2

Квантование называется равномерным, если длины всех интервалов равны между собой, тогда величина $\Delta x = \Delta x_j$ называется шагом квантования по уровню.

Кроме закона распределения, для описания случайной величины нередко используют числовые характеристики, являющиеся математическими ожиданиями некоторых функций $\varphi(x)$ от случайной величины X . Такие характеристики вычисляют по формуле

$$\text{M}\ \varphi(X)\ = \sum_{j=1}^{m} \varphi(x_j)\,p(x_j) \ \text{- для дискретных}\ X\ ,$$
$$\text{M}\ \varphi(X)\ = \int_{-\infty}^{\infty} \varphi(x)W(x)dx \ \text{- для непрерывных}\ X\ . \qquad (1.2)$$

Простейшей моделью пары «сообщение - сигнал» является система двух случайных величин XY .

Если эти случайные величины дискретны, то их полной характеристикой служит набор ms совместных вероятностей $p(x_j, y_k)$, где m – основание кода сообщения, s – основание кода сигнала.

Система непрерывных случайных величин XY полностью характеризуется функцией распределения $F(x, y)$ или совместной плотностью вероятности $W(x, y)$ при условии существования последней.

Для каждой из случайных величин, входящих в систему, можно вычислить безусловные и условные вероятности (плотности вероятности):

$$p(x_j) = \sum_{k=1}^{s} p(x_j, y_k)\,, \quad W(x) = \int_{-\infty}^{\infty} W(x, y)dy\,,$$
$$p(y_k) = \sum_{j=1}^{m} p(x_j, y_k)\,, \quad W(y) = \int_{-\infty}^{\infty} W(x, y)dx\,, \qquad (1.3)$$

$$p(x_j / y_k) = \frac{p(x_j, y_k)}{p(y_k)}, \quad W(x/y) = \frac{W(x,y)}{W(y)},$$
$$p(y_k / x_j) = \frac{p(x_j, y_k)}{p(x_j)}, \quad W(y/x) = \frac{W(x,y)}{W(x)}, \tag{1.4}$$

причем по формулам умножения вероятностей имеем

$$p(x_j, y_k) = p(x_j) p(y_k / x_j) = p(y_k) p(x_j / y_k),$$
$$W(x,y) = W(x) W(y/x) = W(y) W(x/y). \tag{1.5}$$

Случайные величины X и Y статистически независимы, если для всех значений x и y выполняется условие

$$p(x_j / y_k) = p(x_j), \quad W(x/y) = W(x). \tag{1.6}$$

Числовые характеристики системы вычисляются по формуле

$$\text{M}\ \varphi(X,Y) = \sum_{j=1}^{m} \sum_{k=1}^{s} \varphi(x_j, y_k) p(x_j, y_k),$$
$$\text{M}\ \varphi(X,Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi(x,y) W(x,y) dx dy. \tag{1.7}$$

Условные числовые характеристики, например математическое ожидание $\varphi(x,y)$ при условии, что случайная величина Y приняла конкретное значение y_k, вычисляем по формуле

$$\text{M}\ \varphi(X, y_k) = \sum_{j=1}^{m} \varphi(x_j, y_k) p(x_j / y_k),$$
$$\text{M}\ \varphi(X, y_k) = \int_{-\infty}^{\infty} \varphi(x, y_k) W(x / y_k) dx. \tag{1.8}$$

Модель сигнала – случайная последовательность. Более сложной моделью сообщения и сигнала являются случайные последовательности.

Дискретная по времени и по уровню случайная функция $X(t)$ полностью определяется последовательностью n случайных чисел – отсчетов $X^{(1)}, \ldots, X^{(n)}$, взятых в моменты времени

$t_1, t_2, ..., t_n$. Совокупность символов $X^{(1)}, ..., X^{(n)}$ образует кодовое слово, а число n символов в слове называется его длиной.

Аналогично для сигнала $Y(t)$ имеем последовательность отсчетов $Y^{(1)}, ..., Y^{(z)}$, образующую кодовое слово длины z.

Основание кода сигнала s также может отличаться от основания кода сообщения m.

Последовательность $X = (X^{(1)}, ..., X^{(n)})$ имеет $N_x = m^n$ возможных значений, а для последовательности $Y = (Y^{(1)}, ..., Y^{(z)})$ это число равно $N_y = s^z$. Поэтому для полного описания системы XY необходимо перечислить $N = N_x N_y = m^n s^z$ возможных реализаций этой системы и указать их вероятности $p(x_j^{(1)}, ..., x_k^{(n)}, y_r^{(1)}, ..., y_q^{(z)})$.

Из них могут быть получены безусловные и условные характеристики $X(t)$ и $Y(t)$ по формулам, аналогичным (1.3) - (1.8), например:

$$p(x_j^{(1)}, ..., x_k^{(n)}) = \sum_{r=1}^{s} \cdots \sum_{q=1}^{s} p(x_j^{(1)}, ..., x_k^{(n)}, y_r^{(1)}, ..., y_q^{(z)}),$$

$$p(x_j^{(1)}, ..., x_k^{(n)} / y_r^{(1)}, ..., y_q^{(z)}) = \frac{p(x_j^{(1)}, ..., x_k^{(n)}, y_r^{(1)}, ..., y_q^{(z)})}{p(y_r^{(1)}, ..., y_q^{(z)})}$$

и так далее.

Две последовательности бесконечной длины

$$X = ..., X^{(-1)}, X^{(0)}, X^{(1)}, ... \quad \text{и} \quad Y = ..., Y^{(-1)}, Y^{(0)}, Y^{(1)}, ...$$

называются стационарными и стационарно связанными, если для любых n и i

$$p(x_j^{(i)}, ..., x_k^{(i+n)}, y_r^{(i)}, ..., y_q^{(i+n)}) = p(x_j^{(1)}, ..., x_k^{(n)}, y_r^{(1)}, ..., y_q^{(n)}),$$

т.е. все характеристики таких последовательностей не зависят от начала отсчета.

Модель сигнала – непрерывная случайная функция. Осуществляя квантование по времени, переходим от непрерывной случайной функции $X(t)$ к системе непрерывных случайных величин $X^{(1)}, X^{(2)}, ..., X^{(n)}$, а от непрерывной случайной функции $Y(t)$ – к системе $Y^{(1)}, Y^{(2)}, ..., Y^{(n)}$. Полной статистической характеристикой сообщения и сигнала является $2n$-мерная совместная плотность вероятности $W(x^{(1)}, ..., x^{(n)}, y^{(1)}, ..., y^{(n)})$. Для вычисления других характеристик необходимо использовать обобщение формул (1.3) – (1.8).

Помехи. Помехой называется всякое стороннее воздействие, мешающее правильному приему сигнала. Если принимаемый сигнал $Y(t)$ связан с передаваемым сигналом $X(t)$ соотношением

$$Y(t) = X(t) + Q(t), \tag{1.9}$$

то помеха $Q(t)$ называется аддитивной, а в случае

$$Y(t) = X(t)Q(t) \text{ -} \tag{1.10}$$

мультипликативной.

В качестве помехи имеет смысл рассматривать только такую функцию $Q(t)$, значение которой заранее не известно получателю, так как в противном случае помеха может быть скомпенсирована в месте приема. Поэтому для описания помехи используют те же модели, что и для описания сигнала. Принципиальное отличие сигнала от помехи сводится к тому, что последняя не имеет информативных параметров.

РЕШЕНИЕ ТИПОВЫХ ПРИМЕРОВ

Пример 1.1. Случайная величина X – число бросаний монеты до первого выпадания герба. Найти:

а) ряд распределения случайной величины X,

б) математическое ожидание X,

в) математическое ожидание двоичного логарифма вероятности X.

Решение. Возможные значения случайной величины X равны 1, 2, 3,... . Для осуществления события $x = n$ необходимо, чтобы в первых n-1 бросаниях выпадали решки, а в n-м бросании выпал орел, поэтому $p(x = n) = (1/2)^n$ по формуле умножения вероятностей независимых событий. Ряд распределения дан в табл. 1.1.

Таблица 1.1

x_j	1	2	3	4	…
$p(x_j)$	1/2	1/4	1/8	1/16	…

Математическое ожидание числа бросаний вычислим по формуле (1.2), положив $\varphi(x) = x, \quad m = \infty$,

$$M\ X = \sum_{j=1}^{\infty} x_j p(x_j) = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + ... =$$

$$\frac{1}{2} \cdot (1 + 2 \cdot \frac{1}{2} + 3 \cdot \frac{1}{4} + ...) = \frac{1}{2} \cdot (1 + \frac{1}{2} + \frac{1}{4} + ...)^2 = \frac{1}{2} \cdot \left(\frac{1}{1 - 0,5} \right)^2 = 2.$$

Математическое ожидание двоичного логарифма вероятности X также вычисляем по формуле (1.2), положив $\varphi(x) = \log_2 p(x)$, т. е.

$$M\ \log_2 p(X) = \sum_{j=1}^{\infty} \log_2 p(x_j) p(x_j) =$$

$$= \left(\log_2 \frac{1}{2} \right) \cdot \frac{1}{2} + \left(\log_2 \frac{1}{4} \right) \cdot \frac{1}{4} + ... = -1 \cdot \frac{1}{2} - 2 \cdot \frac{1}{4} - 3 \cdot \frac{1}{8} - ... = -2.$$

Пример 1.2. По двоичному каналу связи с помехами (рис. 1.3) передаются сообщения x_1 и x_2 с априорными вероятностями $p(x_1) = 0,4$ и $p(x_2) = 0,6$. Влияние помех описывается переходными вероятностями:

$$p(y_1 / x_1) = 0,75, \qquad p(y_2 / x_1) = 0,25,$$

$$p(y_1 / x_2) = 0,5, \qquad p(y_2 / x_2) = 0,5.$$

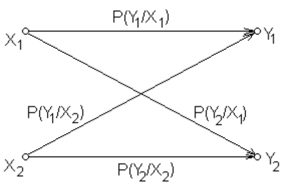


Рис. 1.3

Найти: а) безусловные вероятности сигналов на выходе канала;

б) наиболее вероятное значение X , если $y = y_1$;

в) наиболее вероятное значение X , если $y = y_2$.

Решение. Совместные вероятности

сообщения X и сигнала Y вычисляем по формуле умножения вероятностей (1.5):

$$p(x_1, y_1) = p(x_1)p(y_1 / x_1) =$$
$$0,4 \cdot 0,75 = 0,3,$$
$$p(x_1, y_2) = 0,4 \cdot 0,25 = 0,1, \qquad p(x_2, y_1) = 0,6 \cdot 0,5 = 0,3,$$
$$p(x_2, y_2) = 0,6 \cdot 0,5 = 0,3.$$

Безусловные вероятности сигналов на выходе канала вычислим по формуле полной вероятности (1.3):

$$p(y_1) = p(x_1, y_1) + p(x_2, y_1) = 0,3 + 0,3 = 0,6,$$
$$p(y_2) = p(x_1, y_2) + p(x_2, y_2) = 0,1 + 0,3 = 0,4 = 1 - p(y_1).$$

Условные вероятности сообщений на входе находим по формуле Байеса (1.4):

$$p(x_1 / y_1) = \frac{p(x_1, y_1)}{p(y_1)} = \frac{0,3}{0,6} = 0,5,$$
$$p(x_2 / y_1) = \frac{0,3}{0,6} = 0,5 = 1 - p(x_1 / y_1),$$
$$p(x_1 / y_2) = \frac{p(x_1, y_2)}{p(y_2)} = \frac{0,1}{0,4} = 0,25,$$
$$p(x_2 / y_2) = \frac{0,3}{0,4} = 0,75 = 1 - p(x_1 / y_2).$$

Сравнив $p(x_1 / y_2)$ и $p(x_2 / y_2)$, видим, что если принят сигнал y_2, то более вероятно, что было передано сообщение x_2. Сигнал y_1 мог быть с одинаковой вероятностью вызван сообщениями x_1 и x_2.

Пример 1.3. Сигнал $Y(t)$ на выходе непрерывного канала связи выражается через входной сигнал $X(t)$ соотношением $Y(t) = X(t) + Z(t)$, где $Z(t)$ – аддитивный нормальный стационарный белый шум с односторонней спектральной плотностью $N_0 = 10^{-18}$ Вт/Гц, ограниченный полосой от 0 до $F_B = 10$ МГц. Суммарная мощность составляющих в спектре сигнала $X(t)$, лежащих вне указанной полосы, пренебрежимо мала.

Осуществить квантование по времени сигнала $Y(t)$ на интервале от 0 до $T = 10^{-4}$ секунды. Для конкретной реализации входного сигнала (в вольтах)

$$x(t) = \frac{10^{-6}}{1 + 10^{5} \cdot (t - 5 \cdot 10^{-5})^{2}}$$

найти для квантованного сигнала:

а) вектор условных математических ожиданий;

б) условную корреляционную матрицу;

в) условную плотность вероятности квантованного сигнала на выходе.

Решение. Чтобы осуществить квантование непрерывного по времени сигнала $Y(t)$, необходимо взять его отсчеты $Y^{(1)}, Y^{(2)}, ..., Y^{(n)}$ в моменты времени $t_k = k\Delta t$, $k = 1, 2, ..., n$, где $T/\Delta t \le n < T/\Delta t + 1$.

Верхняя граничная частота суммы сигнала с шумом равна F_B, поэтому шаг квантования определяется в соответствии с теоремой Котельникова

$$\Delta t = \frac{1}{2F_B} = \frac{1}{2 \cdot 10^{7}} \text{ с} = 0,05 \text{ мкс.}$$

Требуемое число отсчетов равно $n = 10^{-4} \cdot 2 \cdot 10^{7} = 2000$.

Каждый отсчет сигнала $Y^{(k)} = Y(t_k)$ является суммой двух величин

$$Y^{(k)} = X^{(k)} + Z^{(k)},$$

где $X^{(k)} = X(t_k)$ - отсчет сообщения;

$Z^{(k)} = Z(t_k)$ - отсчет шума.

Вектор условных математических ожиданий сигнала состоит из следующих элементов

$$m^{(k)} = \mathrm{M}\left[Y^{(k)} / x^{(k)}\right] = x^{(k)} + \mathrm{M}\left[Z^{(k)}\right] = x^{(k)} =$$

$$\frac{10^{-6}}{1 + 10^{5} \cdot (5 \cdot 10^{-8} k - 5 \cdot 10^{-5})^{2}}$$

и определяется только передаваемым сообщением, так как математическое ожидание белого шума $Z(t)$ равно нулю.

Условная корреляционная матрица **B** сигнала $Y(t)$ при фиксированном $x(t)$ состоит из следующих элементов

$$B^{(kj)} = \mathrm{M}\left[Y^{(k)} - m^{(k)}\right]\cdot\left[Y^{(j)} - m^{(j)}\right] = \mathrm{M}\left[Z^{(k)}Z^{(j)}\right]$$

и равна корреляционной матрице отсчетов шума. Элементы этой матрицы есть отсчеты корреляционной функции шума

$$B^{(kj)} = B(t_k, t_j).$$

Шум стационарен, поэтому его корреляционная функция зависит от разности аргументов $\tau = t_1 - t_2$ и может быть найдена по теореме Винера – Хинчина

$$B(\tau) = \int_0^\infty G(f)\cos 2\pi f\tau \cdot df ,$$

где $G(f)$ – спектр плотности мощности шума.

По условию задачи, он равномерен в полосе $0\ldots F_B$,

$$G(f) = \begin{cases} N_0, & 0 \le f \le F_B, \\ 0, & f > F_B. \end{cases}$$

Находим выражение для корреляционной функции

$$B(\tau) = N_0 \int_0^{F_B} \cos 2\pi f\tau \cdot df = N_0 \frac{\sin 2\pi F_B \tau}{2\pi\tau}.$$

Поскольку $\Delta t = 1/(2F_B)$, то

$$B(\tau) = N_0 F_B \frac{\sin \pi\tau/\Delta t}{\pi\tau/\Delta t}.$$

Отсюда видно, что $B(\tau) = 0$ при $\tau = \Delta t; 2\Delta t; 3\Delta t,...,$ т.е. отсчеты $Z^{(1)},...,Z^{(n)}$, взятые с шагом квантования Δt, некоррелированы. Таким образом, в корреляционной матрице отсчетов сигнала не равны нулю будут только элементы, стоящие на главной диагонали,

$$B^{(kk)} = B(0) = N_0 F_B = 10^{-11},$$

численно равные дисперсии этих отсчетов (вольт^2).

Условная плотность вероятности квантованного сигнала есть совместная плотность вероятности системы n некоррелированных (следовательно, и независимых) нормальных случайных величин

$$W\left[y^{(1)},...,y^{(n)}/x(t)\right]=\prod_{k=1}^{n}W\left[y^{(k)}/x^{(k)}\right]=$$

$$\prod_{k=1}^{n}\frac{1}{\sqrt{2\pi B^{(kk)}}}\cdot\exp\left[-\frac{1}{2B^{(kk)}}(y^{(k)}-x^{(k)})^2\right]=$$

$$\frac{1}{(2\pi N_0F_B)^{n/2}}\exp\left[-\frac{1}{2N_0F_B}\sum_{k=1}^{n}(y^{(k)}-x^{(k)})^2\right].$$

Пример 1.4. Доказать, что для любой положительной случайной величины X (имеющей только положительные возможные значения) при $a>1$ справедливо неравенство Иенсена

$$\mathrm{M}\ \log_a X\ \le\log_a \mathrm{M}\ X\ .$$

Доказать, что для любой системы случайных величин $Q,L,...,Z$ и любой функции φ, таких, что $\varphi(Q,L,...,Z)>0$ при всех возможных значениях системы, справедливо аналогичное неравенство

$$\mathrm{M}\ \log_a\varphi(Q,L,...,Z)\ \le\log_a \mathrm{M}\ \varphi(Q,L,...,Z)\ .$$

Найти необходимые и достаточные условия, при которых неравенства обращаются в равенства.

Решение. Сначала убедимся, что непрерывная функция $y=\log_a x$ является строго выпуклой вверх, т.е. ее вторая производная отрицательна при любых $x>0$.

Действительно,

$$y'=\frac{1}{x\ln a},\quad y''=-\frac{1}{\ln a}\cdot\frac{1}{x^2}<0\quad \text{при } a>1.$$

Следовательно, график функции $y=\log_a x$ лежит ниже касательной, проведенной в любой точке $x_0>0$ (рис. 1.4):

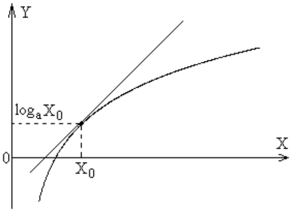


Рис. 1.4

$$y = \log_a x \le \log_a x_0 + \frac{1}{x_0 \ln a}(x - x_0),$$

причем знак равенства выполняется только в точке касания $x = x_0$.

Предположим, что X – положительная случайная величина, тогда полученное неравенство справедливо для любого из ее возможных значений и, следовательно, при усреднении обеих частей знак неравенства сохранится:

$$\mathrm{M}\ \ \log_a X\ \ \le \log_a x_0 + \frac{1}{x_0 \ln a}\ \ \mathrm{M}\ \ X\ \ - x_0\ \ .$$

Выбрав абсциссу точки касания $x_0 = \mathrm{M}\ \ X$, получим окончательно

$$\mathrm{M}\ \ \log_a X\ \ \le \log_a \mathrm{M}\ \ X\ \ .$$

Это неравенство обращается в равенство тогда и только тогда, когда все возможные значения случайной величины $X = x_0 = \mathrm{M}\ \ X$, т.е. если величина X не случайна.

Пусть случайная величина X получена в результате функционального преобразования системы случайных величин $X = \varphi(Q, L, ..., Z) > 0$, тогда в силу доказанного неравенства имеем

$$\mathrm{M}\ \ \log_a \varphi(Q, L, ..., Z)\ \ \le \log_a \mathrm{M}\ \ \varphi(Q, L, ..., Z)\ \ .$$

Это неравенство обращается в равенство тогда и только тогда, когда величина $X = \varphi(Q, L, ..., Z)$ не случайна.

ЗАДАЧИ

1.1 В городе Н в 50% случаев бывает ясная погода, в 30% - переменная и в 20% - дождливая. Сообщения о погоде передаются при помощи кодовых слов «ясно», «переменно», «дождь». Найти среднее количество букв в одном кодовом слове (математическое ожидание длины кодового слова).

1.2 В условиях задачи 1.1 сообщения о погоде передаются по линии связи при помощи импульсов амплитуды 5; 3 и 1 В соответственно. Определить среднюю мощность в импульсе, если входное сопротивление линии равно 100 Ом.

1.3. В условиях задачи 1.2 на выходе линии связи установлен идеальный ограничитель сверху на уровне 0,5 В. Имеется ли статистическая зависимость сигналов на входе линии и выходе ограничителя?

1.4 Сообщение на входе линии связи может с одинаковой вероятностью принимать одно из двух значений: $x_1 = -1$ В или $x_2 = +1$ В. На выходе линии установлен вольтметр, ошибка измерения которого распределена нормально с математическим ожиданием, равным нулю, и среднеквадратическим отклонением 0,5 В. Показание вольтметра в некоторый момент времени равно 0,75 В. Найти условные вероятности сообщений x_1 и x_2.

1.5 Температура в камере может принять с одинаковой вероятностью любое значение в интервале от 0 до 40°. Как выбрать шаг квантования этой величины, чтобы ошибка квантования никогда не превышала 2°? Построить ряд распределения для квантованных значений температуры.

1.6 В условиях задачи 1.5 найти средний квадрат ошибки квантования.

1.7 Напряжение в сети в момент измерения - случайная величина, имеющая нормальное распределение с параметрами $m = 220$ В, $\sigma = 5$ В. На интервале 190 - 240 В осуществить квантование этой величины с шагом $\Delta u = 5$ В и построить ряд распределения.

1.8 Сколько отсчетов по теореме Котельникова необходимо для передачи сигнала длительностью 10 мин с выхода микрофона, если спектр звуковых частот полностью заключен в полосе 20 - 20 000 Гц?

1.9 Можно ли без искажений восстановить телевизионный видеосигнал, если по линии связи передавались его отсчеты с шагом $\Delta t = 0,2$ мкс?

1.10 Система случайных величин XY имеет совместную плотность вероятности

$$W(x,y)=\begin{cases}0{,}125, & 0\le x\le 2,\quad 1\le y\le 5,\\ 0, & \text{вне этой области.}\end{cases}$$

а) Вычислить четвертый центральный момент величины X.
б) Являются ли эти случайные величины независимыми?

1.11 Билеты лотереи, среди которых только один выигрышный, занумерованы числами от 0 до 999. Указать среднее количество цифр, необходимых для передачи номера выигравшего билета.

1.12 По каналу связи передаются сообщения x_1 или x_2 с вероятностями $p(x_1)=0,7$, $p(x_2)=0,3$. Вследствие влияния помех сигнал на выходе может принимать одно из трех значений y_1, y_2 или y_3 с вероятностями:

$p(y_1/x_1)=0,4$, $p(y_2/x_1)=0,4$, $p(y_3/x_1)=0,2$,
$p(y_1/x_2)=0,2$, $p(y_2/x_2)=0,2$, $p(y_3/x_2)=0,6$.

Найти вероятность ошибочных решений, если выходным сигналам ставить в соответствие следующие решения:
$y_1\to x_1$, $y_2\to x_1$, $y_3\to x_2$.

1.13 Даны математические ожидания двух нормальных случайных величин: $m_x=26$, $m_y=-12$ и их корреляционная матрица

$$\left\|K_{jk}\right\|=\begin{Vmatrix}196 & -91\\ -91 & 169\end{Vmatrix}.$$

Определить плотность вероятности системы XY, считая совместное распределение также нормальным.

1.14 Плотность вероятности системы трех случайных величин равна

$$W(x,y,z)=\frac{\sqrt{3}}{16\pi^{\frac{3}{2}}}\exp\left\{-\frac{1}{8}\left[2x^2+4y^2-2y(z+5)+(z+5)^2\right]\right\}.$$

Найти совместную плотность вероятности величин X и Y.

1.15 Система случайных величин имеет нормальное распределение

$$W(x,y)=\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}}\exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{x^2}{\sigma_x^2}-\frac{2rxy}{\sigma_x\sigma_y}+\frac{y^2}{\sigma_y^2}\right]\right\}.$$

Определить: а) условную плотность вероятности $W(x/y)$,

б) условное математическое ожидание $M\quad X/y\quad$,

в) условную дисперсию $D\quad X/y\quad$.

1.16 Сообщение X на входе линии связи есть случайная величина, имеющая нормальное распределение с $m_x=0$ В, $\sigma_x=10$ В. Сигнал на выходе линии $Y=X+Z$, где помеха Z - случайная величина, независимая от X и имеющая нормальное распределение с параметрами $m_z=0$ В, $\sigma_z=1$ В. Указать наиболее вероятное значение переданного сообщения x, если сигнал на выходе линии равен 7 В?

1.17 Случайная величина X может принять одно из m значений $x_1,x_2,...,x_m$. Какими должны быть вероятности $p(x_j)$, чтобы величина $\mathrm{M}\left[\ln\frac{1}{p(X)}\right]$ приняла наибольшее значение? Чему равно это значение?

1.18 В двоичном симметричном канале с независимыми ошибками найти вероятность того, что в 1000-разрядной кодовой комбинации возникнет не более двух ошибок, если битовая вероятность ошибки равна:

а) 10^{-2}; а) 10^{-5}.

1.19 В двоичном симметричном канале с независимыми ошибками битовая вероятность ошибки равна 10^{-2}. Оценить вероятность того, что в 1000-разрядной кодовой комбинации возникнет от 70 до 90 ошибок.

Указание. На основании центральной предельной теоремы теории вероятностей использовать аппроксимацию огибающей биномиального распределения колокольной кривой.

1.20 Найти максимальное значение модуля ошибки квантования по уровню и ее среднеквадратическое значение, если шаг квантования равен 0,05 В.

2 ИНФОРМАЦИОННЫЕ ХАРАКТЕРИСТИКИ ЦИФРОВЫХ СИГНАЛОВ

2.1 Собственная информация. Взаимная информация

Описание дискретного канала. Цифровымым называется такой канал, сигналы на входе и выходе которого являются последовательностями дискретных случайных величин (символов).

Для полного описания канала на интервале времени, соответствующем передаче одного символа, необходимо задать ансамбли символов на входе X и выходе Y и условные вероятности переходов $p(y_k / x_j)$. В дальнейшем будем обозначать:

$X = \begin{Bmatrix} x_1, x_2, ..., x_j, ..., x_n \\ p(x_1), ..., p(x_j), ..., p(x_n) \end{Bmatrix}$ - ансамбль сообщений на входе,

$Y = \begin{Bmatrix} y_1, y_2, ..., y_k, ..., y_m \\ p(y_1), ..., p(y_k), ..., p(y_m) \end{Bmatrix}$ - ансамбль сигналов на выходе.

Собственная информация. Поскольку появление символа сообщения x_j на входе дискретного канала есть событие случайное, то имеет место неопределенность исхода. В результате опыта неопределенность уменьшается или даже исчезает полностью, и при этом получается некоторое количество информации.

Тогда собственная информация символа x_j (количество информации, доставляемое самим символом x_j или любым другим, однозначно с ним связанным) определяется как

$$I(x_j) = -\log p(x_j) = \log \frac{1}{p(x_j)}, \qquad (2.1.1)$$

т.е. информация в каком-либо событии измеряется логарифмом величины, обратной вероятности его появления.

Выбор основания логарифма $\log_a p(x_j)$ определяет единицу количества информации. Если $a=2$, то единица информации называется двоичной (бит), при $a=e$ – натуральной (нат), а при $a=10$ – десятичной. Двоичная единица количества информации, например, есть собственная информация символа, обладающего двумя равновозможными состояниями. Переход от одной системы логарифмов к другой равносилен простому изменению единицы измерения информации. Этот переход осуществляется по формуле

$$\log_b k = \log_b a \cdot \log_a k .$$

Отсюда $1\,\text{нат} = \log_2 e\ \text{бит} = 1{,}4427\ \text{бит}$,

$$1\,\text{дес. ед.} = \log_2 10\ \text{бит} = 3{,}3219\ \text{бит} .$$

Условная собственная информация. В общем случае сообщения X и сигналы Y на входе и выходе дискретного канала зависимы. Пусть $p(x_j / y_k)$ - условная вероятность того, что реализовалось состояние x_j ансамбля X при условии, что ансамбль Y принял состояние y_k. Тогда, по аналогии с собственной, информация, содержащаяся в символе сообщения x_j при условии, что сигнал принял значение y_k, определяется как

$$I(x_j / y_k) = -\log p(x_j / y_k) \tag{2.1.2}$$

и называется условной собственной информацией.

Взаимная информация. Обратимся снова к ансамблям X и Y. Пусть ансамбли зависимы. В результате опыта (приема символа сигнала y_k) апостериорная вероятность $p(x_j / y_k)$ появления символа x_j изменяется по сравнению с априорной $p(x_j)$. Тогда количество информации относительно символа сообщения x_j, доставляемое символом сигнала y_k, можно определить как логарифм отношения апостериорной вероятности к априорной

$$I(x_j ; y_k) = \log \frac{p(x_j / y_k)}{p(x_j)} . \tag{2.1.3}$$

Это и есть взаимная информация.

Основные свойства взаимной информации.

1) Взаимная информация может быть отрицательной, положительной и равной нулю в зависимости от соотношения между априорной и апостериорной вероятностями

$$-\infty < I(x_j;y_k) < \infty. \qquad (2.1.4)$$

2) Взаимная информация не превышает собственную

$$I(x_j;y_k) \le I(x_j), \quad I(x_j;y_k) \le I(y_k). \qquad (2.1.5)$$

При данной вероятности $p(x_j)$ взаимная информация $I(x_j;y_k)$ достигает максимума, когда принятый символ y_k однозначно определяет переданный символ x_j. При этом

$$p(x_j / y_k) = \begin{bmatrix} 0, & j \ne k, \\ 1, & j = k, \end{bmatrix}$$

и максимальное значение взаимной информации

$$I(x_j;y_j) = -\log p(x_j),$$

равно собственной информации, определяемой только априорной вероятностью символа x_j.

3) Свойство симметрии

$$I(x_j;y_k) = I(y_k;x_j), \qquad (2.1.6)$$

т.е. информация, содержащаяся в y_k относительно x_j, равна информации, содержащейся в x_j относительно y_k. В силу этого свойства информацию $I(x_j;y_k)$ называют взаимной информацией между x_j и y_k.

4) Свойство аддитивности количества информации

$$I(x_j,z_i;y_k,q_l) = I(x_j;y_k) + I(z_i;q_l) \qquad (2.1.7)$$

Если пара XY независима от пары ZQ, то информация, содержащаяся в паре $y_k q_i$ относительно пары $x_j z_i$, равна сумме информации, содержащейся в y_k относительно x_j, и информации, содержащейся в q_l относительно z_i.

РЕШЕНИЕ ТИПОВЫХ ПРИМЕРОВ

Пример 2.1.1. Определить собственную информацию, содержащуюся в изображении, при условии, что оно разлагается на 500 строк по 500 элементов в каждой строке. Яркость каждого элемента передается одним из восьми квантованных уровней. Различные градации яркости равновероятны, а яркости разных элементов статистически независимы.

Решение. Обозначим случайной величиной X яркость одного элемента изображения. По условию задачи все 8 градаций яркости одинаково вероятны, т.е. $p(x_j) = 1/n$, где $n = 8$ и, следовательно, собственная информация одного элемента для любого *j* по формуле (2.1.1)

$$I(x_j) = \log_2 n .$$

Изображение содержит $N = 500 \cdot 500 = 2,5 \cdot 10^5$ элементов.

Так как яркости элементов независимы, то по свойству аддитивности информации

$$I(\text{изображения}) = NI(x_j) = N \log_2 n = 2,5 \cdot 10^5 \cdot 3 = 7,5 \cdot 10^5 \text{ бит.}$$

Пример 2.1.2. На экране индикатора РЛС, представляющего поле с 10 вертикальными и 10 горизонтальными полосами, появляется изображение объекта в виде яркостной отметки. Все положения объекта равновероятны.

Определить количество информации, содержащееся в сообщениях:

а) объект находится в 46-м квадрате экрана;

б) объект находится в 5-й горизонтальной строке экрана;

в) объект находится в 6-м вертикальном столбце и 3-й горизонтальной строке экрана.

Решение. а) Пусть x_{46} - сообщение о том, что объект находится в 46-м квадрате экрана.

Собственная информация в этом сообщении по формуле (2.1.1) равна $I(x_{46}) = -\log_2 p(x_{46})$. Безусловная вероятность сообщения – объект находится в 46-квадрате экрана – равна $p(x_{46}) = m/n$, где n – общее число возможных исходов (квадратов поля), m – число исходов, благоприятствующих событию x_{46}.

По условию задачи $n = 100$ квадратов, а $m = 1$. Тогда $p(x_{46}) = 0,01$ и $I(x_{46}) = -\log 0,01 = \log_2 100 = 6,6439$ бит.

б) Вероятность события y_5 – объект находится в 5-й горизонтальной строке экрана – по аналогии с рассмотренным случаем а) определится $p(y_5) = m/n = 10/100 = 0,1$ и собственная информация

$$I(y_5) = -\log_2 0,1 = \log_2 10 = 3,3219 \text{ бит.}$$

в) Вероятность события $z_{6,3}$ – объект находится в 6-м вертикальном столбце и 3-й горизонтальной строке – равна $p(z_{6,3}) = m/n = 0,01$, следовательно,

$$I(z_{6,3}) = -\log_2 0,01 = 6,6439 \text{ бит.}$$

Пример 2.1.3. Рассматривается ансамбль сообщений, приведенный в табл. 2.1.1.

Таблица 2.1.1

x_j	x_0	x_1	x_2	x_3	x_4	x_5	x_6
$p(x_j)$	1/2	1/4	1/8	1/32	1/32	1/32	1/32
Кодовое слово	001	010	100	011	101	110	111

Сообщение x_4 поступает в кодер. Вычислить дополнительную информацию об этом сообщении, доставляемую каждым последующим символом на выходе кодера.

Решение. На вход кодера поступает одно из сообщений $x_0, ..., x_6$, а кодер порождает соответствующие таблице двоичные символы. Так, сообщению x_4 соответствует на выходе кодовое слово 101. Символы на выходе кодера появляются последовательно, т.е. первый символ 1, второй 0 и третий 1. Первый символ кодового слова содержит некоторую информацию относительно того, какое сообщение поступает на вход кодера. Так, первый символ 1 показывает, что на входе могли быть сообщения x_2, x_4, x_5 или x_6. Второй символ 0 сужает выбор – теперь на входе возможно одно из двух сообщений: x_2 или x_4. И, наконец, последний, третий символ 1 однозначно определяет переданное сообщение.

По формуле (2.1.3) взаимная информация, содержащаяся в первом кодовом символе 1 относительно сообщения x_4, равна

$$I(x_4;1) = \log \frac{p(x_4/1)}{p(x_4)}.$$

Обратная вероятность $p(x_4/1)$ может быть найдена по формуле Байеса (1.4)

$$p(x_4/1) = \frac{p(x_4)p(1/x_4)}{\sum_{j=1}^{6} p(1/x_j)p(x_j)},$$

где $p(1/x_j) = \begin{cases} 1, & j = 2,4,5,6, \\ 0, & j = 0,1,3, \end{cases}$

т.е. условная вероятность $p(1/x_j) = 0$ для гипотез, при которых первый кодовый символ есть 0, и $p(1/x_j) = 1$ для гипотез, при которых первый кодовый символ 1. В знаменателе формулы Байеса таким образом учитываются те гипотезы, при которых возможно появление 1 на первом месте.

Итак,

$$p(x_4/1) = \frac{1/32}{1/8 + 3 \cdot 1/32} = \frac{1}{7},$$

$$p(x_4/10) = \frac{1/32}{1/8 + 1/32} = \frac{1}{5}, \quad p(x_4/101) = \frac{1/32}{1/32} = 1,$$

а взаимная информация, содержащаяся в первом кодовом символе 1 относительно сообщения x_4, равна

$$I(x_4;1) = \log \frac{p(x_4/1)}{p(x_4)} = \log_2 \frac{1/7}{1/32} = \log_2 32 - \log_2 7 = 2{,}1926 \text{ бит}.$$

Информация, содержащаяся во втором кодовом символе 0 при условии, что первый кодовый символ был 1, есть

$$I(x_4;0/1)=\log\frac{p(x_4/10)}{p(x_4/1)}=\log_2\frac{1/5}{1/7}=0,4854\text{ бит.}$$

Информация, содержащаяся в третьем кодовом символе 1 при условии, что ему предшествовали 10, есть

$$I(x_4;1/10)=\log_2\frac{p(x_4/101)}{p(x_4/10)}=\log_2\frac{1}{1/5}=2,3219\text{ бит.}$$

Так как сообщения x_J и кодовые слова однозначно связаны, то

$$I(x_4)=I(x_4;1)+I(x_4;0/1)+I(x_4;1/10).$$

Действительно, $I(x_4)=-\log_2 p(x_4)=\log_2 32=5$ бит, и это совпадает с приведенной суммой.

Пример 2.1.4. По дискретному каналу передаются сообщения x_1 или x_2. Вследствие действия шумов на выходе появляется один из сигналов y_1, y_2, y_3. Вероятности совместного появления $p(x_j;y_k)$ заданы табл. 2.1.2

Таблица 2.1.2

x_j	y_k		
	y_1	y_2	y_3
x_1	1/4	1/16	1/8
x_2	1/8	3/16	1/4

Вычислить взаимные информации $I(x_2;y_2)$, $I(x_1;y_3)$.

Решение. Дискретный канал с шумом удобно изображать в виде графа (рис. 2.1.2).

Определим взаимную информацию по формуле (2.1.3)

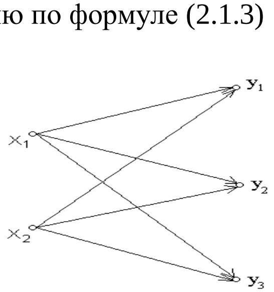


Рис. 2.1

$$I(x_1;y_3)=\log\frac{p(x_1/y_3)}{p(x_1)}$$

или в силу свойства симметрии

$$I(x_1;y_3)=I(y_3;x_1)=\log\frac{p(y_3/x_1)}{p(y_3)}$$

.

Условные и безусловные вероятности найдем, воспользовавшись таблицей. По формуле (1.3)

$$p(x_1) = p(x_1, y_1) + p(x_1, y_2) + p(x_1, y_3) = \frac{1}{4} + \frac{1}{16} + \frac{1}{8} = \frac{7}{16};$$

$$p(x_2) = \frac{1}{8} + \frac{3}{16} + \frac{1}{4} = \frac{9}{16}; \; p(y_1) = \frac{1}{4} + \frac{1}{8} = \frac{3}{8};$$

$$p(y_2) = \frac{1}{16} + \frac{3}{16} = \frac{1}{4}; \; p(y_3) = \frac{1}{8} + \frac{1}{4} = \frac{3}{8}.$$

Используя формулы (1.4), найдем условные вероятности:

$$p(y_2 / x_2) = \frac{p(x_2, y_2)}{p(x_2)} = \frac{3/16}{9/16} = \frac{1}{3},$$

$$p(x_1 / y_3) = \frac{p(x_1, y_3)}{p(y_3)} = \frac{1/8}{3/8} = \frac{1}{3}.$$

Тогда количество взаимной информации по формуле (2.1.3)

$$I(x_2; y_2) = \log_2 \frac{p(y_2 / x_2)}{p(y_2)} = \log_2 \frac{1/3}{1/4} = 0,415 \text{ бит.}$$

$$I(x_1; y_3) = \log_2 \frac{p(x_1 / y_3)}{p(x_1)} = \log_2 \frac{1/3}{7/16} = -0,3923 \text{ бит.}$$

Мы получили $I(x_1; y_3) < 0$, так как $p(x_1 / y_3) < p(x_1)$.

ЗАДАЧИ

2.1.1. На шахматной доске произвольным образом расставлены фигуры. Априори все положения фигур на доске одинаково вероятны. Определить собственную информацию, получаемую от сообщения, что конкретная фигура находится в одной из угловых клеток доски.

2.1.2. Сколько информации содержится в сообщении о том, что сумма очков на двух подброшенных игральных костях есть четное число?

2.1.3. Сколько информации содержится в сообщении том, что сумма очков на двух подброшенных игральных костях равна 7?

2.1.4. Брошены одновременно две игральные кости. Определить количество информации, содержащееся в сообщении о том, что произведение чисел выпавших очков четно.

2.1.5. В цехе работают 7 мужчин и 3 женщины. По табельным номерам наудачу отобраны 3 человека. Найти количество информации, содержащееся в сообщении «все отобранные люди – мужчины».

2.1.6. Из шести букв разрезной азбуки составлено слово «машина». Ребенок, не умеющий читать, рассыпал эти буквы и затем собрал в произвольном порядке. Какое количество информации будет содержаться в утверждении, что у него снова получилось слово «машина»?

2.1.7. Два стрелка, для которых вероятности попадания в мишень равны соответственно 0,6 и 0,7, производят по одному выстрелу. В результате оказалось, что мишень поражена. Какое количество информации содержится в этом сообщении?

2.1.8. Урна содержит 6 черных и 10 белых шаров. Случайно, без возвращения, из урны вынимают 3 шара, и результат опыта передается по системе связи. Пусть шары выбраны в следующей последовательности: черный, черный, белый.

а) Какое количество информации надо передать, если интересоваться только количеством шаров того и другого цвета?

б) Какое количество информации надо передать, если представляет интерес также и порядок, в котором выбраны шары?

2.1.9. В некотором городке четверть женщин — блондинки, половина — брюнетки и четверть — шатенки. Блондинки всегда приходят на свидание вовремя, брюнетки – подбрасывают монету и в зависимости от результата приходят вовремя или опаздывают, а шатенки всегда опаздывают.

1. Определить взаимную информацию между высказыванием «женщина пришла на свидание вовремя» относительно каждого из следующих предположений:

а) она блондинка;

б) брюнетка;

в) шатенка.

2. Сколько информации содержится в высказывании «женщина пришла вовремя на 3 свидания подряд» относительно предположения, что она брюнетка?

2.1.10. При фототелеграфной передаче изображения кадр состоит из $2{,}5 \cdot 10^6$ элементов. Для хорошего воспроизведения необходимы 12 градаций (уровней) яркости. Предполагается, что, все уровни яркости встречаются с одинаковой вероятностью. Элементы изображения независимы. Какое количество информации надо передать по каналу связи, если передача продолжается 5 мин?

2.1.11. По дискретному каналу передается одно из сообщений x_1, x_2, x_3. Вследствие действия шумов на выходе канала появляется сигнал y_1 или y_2. Вероятности совместного появления заданы табл. 2.1.3

Таблица 2.1.3

x_j	y_k	
	y_1	y_2
x_1	0,4	0,1
x_2	0,2	0,15
x_3	0,1	0,05

Вычислить взаимные информации $I(x_1; y_2)$, $I(x_3; y_1)$, $I(x_2; y_2)$.

2.1.12. По двоичному каналу с шумом передаются сообщения x_1, x_2, x_3 с вероятностями 0,2; 0,3; 0,5. На выходе канала проявляются сигналы y_1, y_2, y_3. Вероятности искажения в канале (условные вероятности переходов):

$$p(y_1/x_1)=3/4; \quad p(y_1/x_2)=1/8; \quad p(y_1/x_3)=1/8;$$

$$p(y_2/x_1)=1/8; \quad p(y_2/x_2)=3/4; \quad p(y_2/x_3)=1/8;$$

$$p(y_3/x_1)=1/8; \quad p(y_3/x_2)=1/8; \quad p(y_3/x_3)=3/4.$$

Найти взаимные информации I $(x_1;y_3)$, I$(x_3;y_1)$.

2.1.13. По двоичному каналу с помехами передаются равновероятные и статистически независимые сообщения x_1 и x_2. В результате действия помех они преобразуются в сигналы y_1, y_2, y_3. Условные вероятности переходов $p(y_k/x_j)$ заданы табл. 2.1.4. Вычислить взаимные информации $I(x_1;y_3)$ и $I(x_2;y_2)$.

Таблица 2.1.4

x_j	y_k		
	y_1	y_2	y_3
x_1	/8	/8	/8

$_2$	/8	/8	/8

2.1.14. Рассматривается ансамбль сообщений X, приведенный в табл. 2.1.5.

Таблица 2.1.5

x_j	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$p(x_j)$	1/4	1/4	1/8	1/8	1/16	1/16	1/16	1/16
Кодовое слово	000	001	010	011	100	101	110	111

Сообщение x_2 поступает в кодер. Вычислить дополнительную информацию об этом сообщении, доставляемую каждым последующим символом на выходе кодера.

2.1.15. Сообщения источника $x_1,...,x_4$ для согласования с каналом кодируются в соответствии с табл. 2.1.6.

Таблица 2.1.6

Сообщения x_j	x_1	x_2	x_3	x_4
$p(x_j)$	1/2	1/4	1/8	1/8
Кодовое слово	000	011	101	100

Пусть на вход кодера поступает сообщение x_3. Вычислить дополнительную информацию об этом сообщении, которую содержит каждый последующий символ на выходе кодера.

2.1.16. Среди студенток некоторого института 25% всех девушек – блондинки, а 75% всех блондинок имеют голубые глаза; всего же голубые глаза имеет половина всех девушек. Пусть мы знаем, что некоторая студентка имеет голубые глаза. Сколько дополнительной информации будет содержаться в сообщении о том, что эта девушка – блондинка?

2.2 Средняя собственная информация (энтропия)

Энтропия. Дискретный источник удобнее характеризовать количеством собственной информации, содержащимся в среднем в одном символе ансамбля X.

Это среднее количество собственной информации есть

$$I(X) = \mathrm{M}\left[-\log p(X)\right] = \sum_{j=1}^{n} I(x_j)\, p(x_j) = \qquad (2.2.1)$$

$$-\sum_{j=1}^{n} p(x_j)\log\, p(x_j) = H(X)$$

и названо энтропией (по аналогии с понятием энтропии в термодинамике).

Свойства энтропии.

1) Энтропия неотрицательна

$$H(X) \geq 0. \tag{2.2.2}$$

Знак равенства имеет место, когда X – неслучайна, т.е. $p(x_j) = 1$, а $p(x_i) = 0$ для $i \neq j$. При этом неопределенность относительно ансамбля X отсутствует. Таким образом, энтропия есть мера неопределенности случайного ансамбля.

2) Величина энтропии удовлетворяет неравенству

$$H(X) \leq \log n. \tag{2.2.3}$$

Знак равенства имеет место при равновероятности символов ансамбля X, т.е. при $p(x_j) = 1/n$.

3) Свойство аддитивности энтропии.

В последовательности i независимых символов энтропия равна сумме энтропий, содержащихся в отдельных символах

$$H\left[X^{(1)},\ldots,X^{(i)}\right] = H(X^{(1)}) + \ldots + H(X^{(i)}). \tag{2.2.4}$$

Вычисление энтропии по формуле (2.2.1) можно упростить, введя функцию $\eta(p) = -\log\, p$, тогда формула примет вид

$$H(X) = \sum_{j=1}^{n} \eta(p_j). \tag{2.2.5}$$

Значения функции $\eta(p)$ приведены в Приложении 1.

Условная энтропия. Пусть имеются два статистически зависимых конечных ансамбля символов X и Y. Пары символов $x_j y_k$ с вероятностями $p(x_j, y_k)$ можно рассматривать как элементарные символы объединенного ансамбля XY с энтропией

$$H(XY) = -\sum_{j=1}^{n}\ \sum_{k=1}^{m} p(x_j, y_k)\log\, p(x_j, y_k). \tag{2.2.6}$$

Появление символа x_j вызовет появление символа y_k с условной вероятностью

$$p(y_k / x_j) = \frac{p(x_{j,} y_k)}{p(x_j)}.$$

При этом условная энтропия ансамбля Y в предположении, что выбран символ x_j, будет

$$H(Y / x_j) = \mathrm{M}\left[-\log\, p(Y / x_j)\right] =$$

$$\sum_{k=1}^{m} p(y_k / x_j) \log\, p(y_k / x_j). \tag{2.2.7}$$

Здесь каждому x_j соответствует свое значение энтропии $H(Y / x_j)$, т.е. $H(Y / x_j)$ - случайная величина.

Тогда средняя условная энтропия случайной величины Y, вычисленная при условии, что известно значение другой случайной величины X, равна

$$H(Y / X) = \mathrm{M}\ -\log\, p(Y / X)\ = \sum_{j=1}^{n} p(x_j) H\ (Y / x_j) =$$

$$-\sum_{j=1}^{n} p(x_j) \sum_{k=1}^{m} p(y_k / x_j) \log\, p(y_k / x_j). \tag{2.2.8}$$

Энтропия объединенного ансамбля $H(XY)$ удовлетворяет следующим соотношениям:

а) $H(XY) = H(X) + H(Y / X) = H(Y) + H(X / Y),$ (2.2.9)

если X и Y зависимы;

в) $H(XY) = H(X) + H(Y),$ (2.2.10)

если X и Y независимы.

Для объединенного ансамбля XY условная энтропия удовлетворяет неравенствам:

$$H(Y / X) \le H(Y), \quad H(X / Y) \le H(X). \tag{2.2.11}$$

Избыточность. Считают, что имеется избыточность, если количество информации, содержащейся в сигнале (энтропия сигнала), меньше того количества, которое этот сигнал мог бы содержать по своей физической природе. Введем количественную меру избыточности. Пусть сигнал длиной в n символов (отсчетов) содержит количество информации H. Пусть далее наибольшее количество информации, которое в принципе может содержаться в данном сигнале с учетом наложенных на него ограничений (заданное основание кода, заданная средняя мощность сигнала и т.п.), равно H_{max}. Тогда количественной мерой избыточности является величина

$$R = 1 - H/H_{max} \,. \qquad (2.2.12)$$

Причины появления избыточности – это статистическая связь между символами (отсчетами) сигнала и неэкстремальность распределения вероятностей отдельных символов (отсчетов). Введение избыточности приводит к удлинению сигнала, но зато повышает его информационную устойчивость при воздействии помех.

РЕШЕНИЕ ТИПОВЫХ ПРИМЕРОВ

Пример 2.2.1. На измерительной станции имеются два прибора. Первый имеет шкалу, содержащую 100 делений, его показания могут меняться через каждые 0,05 с. Шкала второго прибора имеет 10 делений, и его показания могут меняться каждые 0,01 с.

Какова наибольшая средняя информация, поставляемая двумя приборами в 1 с?

Решение. 1-й прибор. Энтропия одного значения (отсчета) по формуле (2.2.3) $H_1(X) = \log m_1 = \log 100$.

Число отсчетов в 1 секунду равно $n_1 = 1/0,05 = 20$.

2-й прибор. Энтропия одного значения $H_2(X) = \log 10$, а число отсчетов в 1 секунду равно $n_2 = 100$.

Энтропия двух приборов в 1 с по формуле (2.2.4.) равна

$$H_\Sigma(X) = n_1 H_1(X) + n_2 H_2(X) = 20 \cdot 6,64 + 100 \cdot 3,32 \approx 465 \text{ бит/с}.$$

Пример 2.2.2. Производится стрельба по двум мишеням, по одной сделано 2 выстрела, по второй – 3. Вероятности попадания при одном выстреле соответственно равны 1/2 и 1/3. Исход стрельбы (число попаданий) по какой мишени является более определенным?

Решение. Исход стрельбы определяется числом попаданий в мишень, которое подчинено биномиальному закону распределения $p(X = m) = C_n^{\,m} p^m (1-p)^{n-m}$,

Таблица 2.2.1

m	0	1	2
$p(x=m)$	1/4	1/2	1/4

Таблица 2.2.2

m	0	1	2	3
$p(x=m)$	8/27	4/9	2/9	1/27

где $C_n^{\ m} = \dfrac{n!}{m!(n-m)!}$.

Составляем ряд распределения для числа попаданий в первую мишень при n=2 и p=1/2 (табл. 2.2.1) и вторую мишень при n=3 и p=1/3 (табл. 2.2.2).

Мерой неопределенности исхода стрельбы служит энтропия числа попаданий. Энтропия числа попаданий при стрельбе по первой мишени

$$H_1(X) = -\sum_{j=1}^{n_1} p(x_j)\log\, p(x_j) = -\left[2\cdot\frac{1}{4}\log_2\frac{1}{4} + \frac{1}{2}\log_2\frac{1}{2}\right] = 1{,}5 \text{ бит.}$$

Аналогично для второй мишени имеем

$$H_2(X) = -\left[\frac{8}{27}\log_2\frac{8}{27} + \frac{4}{9}\log_2\frac{4}{9} + \frac{2}{9}\log_2\frac{2}{9} + \frac{1}{27}\log_2\frac{1}{27}\right] = 1{,}7 \text{ бит,}$$

т.е. исход стрельбы по второй мишени обладает большей неопределенностью.

Пример 2.2.3. Источник сообщений вырабатывает ансамбль символов

$$X = \begin{Bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ 0{,}4 & 0{,}2 & 0{,}15 & 0{,}1 & 0{,}1 & 0{,}05 \end{Bmatrix}.$$

Символы в последовательности независимы.

Вычислить энтропию источника и определить избыточность.

Решение. Энтропия источника для случая неравновероятных и независимых сообщений определяется формулой (2.2.1)

$$H(X) = -\sum_{j=1}^{m} p(x_j)\log\, p(x_j) = -[0{,}4\cdot\log 0{,}4 + 0{,}2\cdot\log 0{,}2 +$$

$$0{,}15\cdot\log 0{,}15 + 2\cdot 0{,}1\cdot\log 0{,}1 + 0{,}05\cdot\log 0{,}05] = 2{,}2842 \text{ бит.}$$

Избыточность за счет неоптимальности (неравновероятности) распределения сообщений в источнике определяется формулой (1.2.12) $R = 1 - H/H_{max}$, где $H_{max} = \log m$ по формуле (2.2.3).

Отсюда $R = 1 - 2,2842/\log 6 = 0,1164.$

Пример 2.2.4. Алфавит источника состоит из трех букв: x_1, x_2, x_3.

Определить энтропию на 1 букву текста $X^{(1)}$, $X^{(2)}$ для следующих случаев:

а) буквы алфавита неравновероятны: $p(x_1) = 0,5$, $p(x_2) = p(x_3) = 0,25$, а символы в последовательности на выходе источника статистически зависимы. Условные вероятности $p(x_j^{(2)} / x_i^{(1)})$ заданы в табл. 2.2.3;

Таблица 2.2.3

i – индекс предыдущей буквы	j – индекс последующей буквы		
	1	2	3
1	0,4	0,2	0,4
2	0	0,6	0,4
3	0,3	0	0,7

б) вероятности букв те же, что и в п. а), но символы независимы;

в) символы в последовательности независимы, вероятности букв одинаковы.

Вычислить избыточность источников для случаев а) и б).

Решение. а) В случае неравновероятных и зависимых сообщений энтропия текста по формуле (2.2.9)

$$H(X^{(1)}X^{(2)}) = H(X^{(1)}) + H(X^{(2)} / X^{(1)}),$$

где $$H(X^{(1)}) = -\sum_{j=1}^{m} p(x_j) \log\, p(x_j) = 1,5 \text{ бит},$$

а условная энтропия по формуле (2.2.8) равна

$$H(X^{(2)} / X^{(1)}) = -\sum_{i=1}^{3} \sum_{j=1}^{3} p(x_i^{(1)}) p(x_j^{(2)} / x_i^{(1)}) \log p(x_j^{(2)} / x_i^{(1)}) =$$

$$-\{p(x_1^{(1)})\big[p(x_1^{(2)} / x_1^{(1)}) \log\, p(x_1^{(2)} / x_1^{(1)}) + p(x_2^{(2)} / x_1^{(1)}) \log\, p(x_2^{(2)} / x_1^{(1)}) +$$

$$p(x_3^{(2)} / x_1^{(1)}) \log\, p(x_3^{(2)} / x_1^{(1)})\big] + p(x_2^{(1)})\big[p(x_1^{(2)} / x_2^{(1)}) \log\, p(x_1^{(2)} / x_2^{(1)}) +$$

$$p(x_2^{(2)} / x_2^{(1)}) \log\, p(x_2^{(2)} / x_2^{(1)}) + p(x_3^{(2)} / x_2^{(1)}) \log\, p(x_3^{(2)} / x_2^{(1)})\big] +$$

$$+p(x_3^{(1)})\left[p(x_1^{(2)}/x_3^{(1)})\log p(x_1^{(2)}/x_3^{(1)})+p(x_2^{(2)}/x_3^{(1)})\log p(x_2^{(2)}/x_3^{(1)})+\right.$$
$$\left.p(x_3^{(2)}/x_3^{(1)})\log p(x_3^{(2)}/x_3^{(1)})\right]\Big\}=-\Big\{\frac{1}{2}\big[0,4\cdot\log_2 0,4+0,2\cdot\log_2 0,2+$$
$$0,4\cdot\log_2 0,4\big]+\frac{1}{4}\big[0+0,6\cdot\log_2 0,6+0,4\cdot\log_2 0,4\big]+\frac{1}{4}\big[0,3\cdot\log_2 0,3+$$
$$0,7\cdot\log_2 0,7\big]\Big\}\approx 1,224 \text{ бит.}$$

Энтропия на один символ

$$H_1(X^{(1)},X^{(2)})=\left[H(X^{(1)})+H(X^{(2)}/X^{(1)})\right]\big/2=1,362 \text{ бит/симв}.$$

б) При неравновероятных, но зависимых сообщениях энтропия вычисляется по формуле (2.2.1)

$$H(X)=-\sum_{j=1}^{n}p(x_j)\log p(x_j)=-\left[\frac{1}{2}\log_2\frac{1}{2}+2\cdot\frac{1}{4}\log_2\frac{1}{4}\right]=1,5\,\frac{\text{бит}}{\text{симв}}.$$

Избыточность, обусловленная статистической зависимостью

$$R_1=1-H_1(X)/H(X)=1-1,362/1,5=0,092.$$

в) В случае равновероятных и независимых сообщений энтропия по формуле (2.2.3)

$$H_{max}(X)=\log_2 m=\log_2 3=1,585 \text{ бит.}$$

Избыточность, обусловленная неоптимальностью распределения

$$R_2=1-H(X)/H_{max}(X)=1-1,5/1,585=0,054.$$

Полная избыточность (за счет неоптимальности распределения и наличия статистических взаимосвязей)

$$R=1-H_1(X)/H_{max}(X)=0,141.$$

ЗАДАЧИ

2.2.1. Алфавит русского языка состоит из 32 букв (если не различать е и ё, ь и ъ), включая промежуток между буквами. Вычислить энтропию однобуквенного текста, считая вероятности появления любой из букв в заданном месте текста одинаковыми.

2.2.2. Источник вырабатывает ансамбль сообщений

$$X = \begin{Bmatrix} x_1 & x_2 & x_3 & x_4 \\ 0,2 & 0,3 & 0,4 & 0,1 \end{Bmatrix}$$

Символы в последовательности независимы. Вычислить энтропию источника и определить избыточность.

2.2.3. Найти число значений m равномерно распределенной случайной величины Y, при котором ее энтропия будет равна энтропии случайной величины X, заданной табл. 2.2.4

Таблица 2.2.4

x_j	x_1	x_2	x_3	x_4	x_5	x_6
$p(x_j)$	1/2	1/4	1/8	1/16	1/32	1/6

2.2.4. Производится стрельба по мишени. Сделано 2 выстрела. Вероятность попадания при одном выстреле S. Найти неопределенность исхода стрельбы (числа попаданий) по мишени.

2.2.5. Определить энтропию случайной величины X, распределенной по биномиальному закону:

а) в общем случае,

б) при p=q= $\frac{1}{2}$ и m=3.

2.2.6. В двух урнах имеется по 15 шаров: в первой урне 5 красных, 7 белых и 3 черных; во второй соответственно 4, 4,7. Из каждой урны вынимается по одному шару. Сравнить неопределенности исхода опытов для двух урн.

2.2.7. Алфавит источника сообщений состоит из двух букв x_1 и x_2 с вероятностями 0,6 и 0,4. В последовательности на выходе источника символы статистически зависимы. Условные вероятности переходов $p(x_j^{(2)}/x_i^{(1)})$ заданы табл. 2.2.5

Таблица 2.2.5

i- индекс предыдущей буквы	j- индекс последующей буквы	
	1	2
1	0,2	0,8
2	0,7	0,3

Определить энтропию на один символ текста. $X^{(1)}X^{(2)}$. Вычислить избыточность источника.

2.2.8. Выполнить задачу 2.2.7, если условные вероятности переходов $p(x_j^{(2)}/x_i^{(1)})$ заданы табл. 2.2.6

Таблица 2.2.6

i – индекс предыдущей буквы	j – индекс последующей буквы			
	1	2	3	4
1	0,0	0,2	0,4	0,4
2	0,2	0,2	0,3	0,3
3	0,25	0,0	0,25	0,5
4	0,2	0,4	0,4	0,0

Безусловные вероятности букв x_1 …. x_4 равны соответственно 0,5; 0,25; 0,125; 0,125.

2.2.9. Символы азбуки Морзе могут появиться в сообщении с вероятностями: для точки — 0,51, для тире — 0,31, для промежутка между буквами — 0,12, между словами — 0,06. Определить среднее количество информации в сообщении из 500 символов данного алфавита, считая, что связь между последовательными символами отсутствует.

2.2.10. Система радиозонда измеряет давление. Барометр имеет 10 отметок шкалы, и его отсчеты могут изменяться до любого допустимого значения за 0,01 с. Связь между отсчетами отсутствует. Найти энтропию источника за 1 с, если показания барометра будут появляться со следующими вероятностями (табл. 2.2.7).

Таблица 2.2.7

Отметка шкалы	0	1	2	3	4	5
Вероятн.	0,05	0,05	0,05	0,05	0,1	0,2

2.2.11. Измеряемое напряжение лежит в пределах $0 \div 6$ В. Телеметрический датчик регистрирует приращение напряжения $\Delta u = 0{,}01\,B$. Найти наибольшее среднее количество информации, получаемое за 10 независимых отсчетов.

2.2.12*. В двоичной системе связи для передачи символов 0 и 1 используются импульсы длительности 2 и 4 мкс соответственно. Определить наибольшее среднее количество информации, которое можно передать в такой системе в 1 с.

2.2.13. На радиорелейной линии для передачи символов 0 и 1 используются радиоимпульсы мощностью 1 и 10 Вт соответственно. Для нормальной работы передатчика необходимо, чтобы отдаваемая им средняя мощность не превышала 3,4 Вт. Найти наибольшее количество информации, которое в среднем может переносить один импульс.

2.2.14. Показать, что энтропия H(X) ансамбля X с конечным множеством символов x_i, $i=\overline{1,m}$ достигает максимума, когда все символы равновероятны, и это максимальное значение равно

$$H_{\max}(X)=\log m.$$

2.2.15. Узлу связи А было передано 3 радиограммы, узлу связи В — 4 радиограммы. Вероятности неискаженного приема одной радиограммы на узлах соответственно равны $^2/_3$ и s. Считая, что качество связи характеризуется числом m правильно принятых радиограмм, подчиняющимся биномиальному закону распределения, оценить степень неопределенности качества связи для каждого из узлов.

2.2.16. Вероятность того, что проводная линия выдержит испытание, уменьшается с увеличением ее длины L по экспоненциальному закону

$p=\exp(-L/L_0)$, где $L_0=200$ *км.*

При какой длине линии исход испытания обладает наибольшей неопределенностью?

2.2.17. Передатчик, описанный в задаче 2.2.13, передает символы 1 и 0 с вероятностями 0,1 и 0,9 соответственно. Найти избыточность при ограничении на среднюю мощность передатчика.

2.2.18. Доказать, что любое преобразование вероятностей двух элементов ансамбля, которое делает эти вероятности более близкими друг к другу, увеличивает энтропию ансамбля.

2.2.19. Из многолетних наблюдений за погодой известно, что в некотором пункте А вероятность того, что 15 июня будет дождь, равна 0,4, а вероятность того, что осадков не выпадет, равна 0,6. Вероятность того, что будет дождь 15 ноября, равна 0,2, а вероятность того, что выпадет снег, равна 0,45. В какой из названных дней погоду в п. А следует считать более неопределенной, если:

а) интересоваться вопросом о наличии и характере осадков,

б) интересоваться вопросом только о наличии осадков.

2.3 Средняя взаимная информация

Определения. В изучении проблем связи, кроме рассмотренных выше величин, важную роль играет среднее значение взаимной информации между элементами различных ансамблей.

Рассмотрим условное среднее значение взаимной информации для объединенного ансамбля XY. Пусть сигнал принял значение y_k. Тогда информация, содержащаяся в реализации y_k принятого сигнала относительно ансамбля передаваемых сообщений X,

$$I(X;y_k) = \mathrm{M}\left[\log\frac{p(X/y_k)}{p(X)}\right] = \sum_{j=1}^{n} p(x_j/y_k) I(x_j;y_k) = \sum_{j=1}^{n} p(x_j/y_k)\log\frac{p(x_j/y_k)}{p(x_j)} - \tag{2.3.1}$$

есть средняя взаимная информация между ансамблем X и реализацией y_k.

Аналогично информация, содержащаяся в ансамбле принятых сигналов Y относительно реализации переданного сообщения x_j, определяется как

$$I(x_j;Y) = \mathrm{M}\left[\log\frac{p(Y/x_j)}{p(Y)}\right] = \sum_{k=1}^{m} p(y_k/x_j) I(x_j;y_k) = \sum_{k=1}^{m} p(y_k/x_j)\log\frac{p(y_k/x_j)}{p(y_k)}. \tag{2.3.2}$$

Это средняя взаимная информация между ансамблем Y и реализацией x_j.

Средняя взаимная информация между ансамблем принимаемых сигналов *Y* и ансамблем передаваемых сообщений *X*

$$I(X;Y)=\mathrm{M}\left[\log\frac{p(X/Y)}{p(X)}\right]=\sum_{k=1}^{k}p(y_k)I(X;y_k)=$$
$$\sum_{j=1}^{n}\sum_{k=1}^{m}p(x_j,y_k)\log\frac{p(x_j/y_k)}{p(x_j)} \qquad (2.3.3)$$

есть то количество информации, которое содержится в среднем в ансамбле принимаемых символов *Y* относительно ансамбля передаваемых символов *X*.

Основные свойства средней взаимной информации.

1) Средняя взаимная информация симметрична

$$I(X;Y)=I(Y;X). \qquad (2.3.4)$$

2) Средняя взаимная информация не превышает собственную

$$I(X;Y)\le\begin{cases}H(X)\\H(Y).\end{cases} \qquad (2.3.5)$$

3) Средняя взаимная информация всегда неотрицательна

$$I(X;Y)\ge 0. \qquad (2.3.6)$$

4) Следующее свойство устанавливает соотношение между средней взаимной информацией и энтропиями, относящимися к объединенному ансамблю. Величина

$$I(X;Y)=H(X)-H(X/Y) \qquad (2.3.7)$$

– среднее количество информации о сообщении, содержащееся в принятом сигнале – равна среднему количеству информации, требуемому для определения сообщения *X*, минус среднее количество информации, которое все еще потребуется для определения *X* после приема сигнала *Y*. Тогда энтропию $H(X)$ мы понимаем как среднее количество переданной информации, а условную энтропию $H(X/Y)$ – как среднее количество информации, потерянное вследствие влияния шума («ненадежность»).

В другом варианте

$$I(X;Y)=H(Y)-H(Y/X) \qquad (2.3.8)$$

среднее количество информации есть разность между средним количеством информации, необходимым для определения принятого сигнала, и средним количеством информации, необходимым для определения того же сигнала, когда известно переданное сообщение. Тогда $H(Y/X)$ можно трактовать как среднее количество информации, необходимое для определения помехи в канале, т. е. это есть энтропия шума в канале.

При отсутствии в канале помех

$$I(X;Y) = H(X), \tag{2.3.9}$$

т.е. принимаемый сигнал Y доставляет получателю всю информацию, содержащуюся в переданном сигнале.

В этом случае y_k и x_j связаны однозначно

$$p(x_j / y_i) = \begin{cases} 1 & \text{при } i = j, \\ 0 & \text{при } i \neq j, \end{cases} \tag{2.3.10}$$

и условная энтропия $H(X/Y)$=0.

При значительном уровне помех прием y_k не дает ответа относительно переданного x_j, следовательно, X и Y можно приближенно считать статистически независимыми, тогда $p(x_j / y_k) = p(x_j)$ – прием символа y_k никак не определяет переданного символа x_j – и среднее количество информации

$$I(X;Y) = H(X) - H(X/Y) = 0.$$

Очевидно, формулы (2.3.3), (2.3.7) и (2.3.8) дают тождественные результаты, и выбор той или иной формулы при решении конкретной задачи производится из соображений удобства математических выкладок.

РЕШЕНИЕ ТИПОВЫХ ПРИМЕРОВ

Пример 2.3.1. Вычислить для конкретного канала, заданного в примере 2.1.4, средние количества информации.

$$I(X;y_1), \qquad I(x_1;Y), \qquad I(X;Y).$$

Решение. а) Средняя взаимная информация в реализации сигнала на выходе y_1 относительно случайной величины X на входе канала определяется формулой (2.3.1)

$$I(X;y_1)=\sum_{j=1}^{n}p(x_j/y_1)\log\frac{p(x_j/y_1)}{p(x_j)}=$$

$$p(x_1/y_1)\log\frac{p(x_1/y_1)}{p(x_1)}+p(x_2/y_1)\log\frac{p(x_2/y_1)}{p(x_2)}.$$

Из примера 2.1.4 известны вероятности:

$$p(x_1)=7/16,\quad p(x_2)=9/16,$$

$$p(y_1)=3/8,\quad p(y_2)=1/4,\quad p(y_3)=3/8.$$

Определим условные вероятности:

$$p(x_1/y_1)=\frac{p(x_1,y_1)}{p(y_1)}=\frac{1/4}{3/8}=\frac{2}{3},$$

$$p(x_2/y_1)=\frac{p(x_2,y_1)}{p(y_1)}=\frac{1/8}{3/8}=\frac{1}{3}.$$

Средняя информация

$$I(X;y_1)=\frac{2}{3}\log_2\frac{2/3}{3/8}+\frac{1}{3}\log_2\frac{1/3}{3/8}=0,4967\text{ бит.}$$

б) Средняя взаимная информация в выходной величине Y относительно реализации случайной величины на входе x_1 определяется формулой (2.3.2)

$$I(x_1;Y)=\sum_{k=1}^{m}p(y_k/x_1)\log\frac{p(y_k/x_1)}{p(y_k)}=p(y_1/x_1)\log\frac{p(y_1/x_1)}{p(y_1)}+$$

$$p(y_2/x_1)\log\frac{p(y_2/x_1)}{p(y_2)}+p(y_3/x_1)\log\frac{p(y_3/x_1)}{p(y_3)}.$$

Определяем условные вероятности

$$p(y_1/x_1)=\frac{p(x_1,y_1)}{p(x_1)}=\frac{1/4}{7/16}=\frac{4}{7},$$

$$p(y_2/x_1)=\frac{p(x_1,y_2)}{p(x_1)}=\frac{1/16}{7/16}=\frac{1}{7},$$

$$p(y_3/x_1)=\frac{p(x_1,y_3)}{p(x_1)}=\frac{1/8}{7/16}=\frac{2}{7}.$$

Условная средняя взаимная информация равна

$$I(x_1;Y)=\frac{4}{7}\log_2\frac{4/7}{3/8}+\frac{1}{7}\log_2\frac{1/7}{1/4}+\frac{2}{7}\log_2\frac{2/7}{3/8}=0{,}1198\text{ бит.}$$

в) Средняя взаимная информация между случайной величиной Y на выходе канала и случайной величиной X на входе определяется формулой (2.3.3)

$$I(X;Y)=\sum_{j=1}^{n}\sum_{k=1}^{m}p(x_j)p(y_k/x_j)\log\frac{p(y_k/x_j)}{p(y_k)}.$$

Воспользовавшись результатами вычислений $I(x_1;Y)$, получим для средней взаимной информации

$$I(X;Y)=\mathrm{M}\left[I(x_j;Y)\right]=\sum_{j=1}^{n}p(x_j)I(x_j;Y)=$$

$$p(x_1)\Big[p(y_1/x_1)\log\frac{p(y_1/x_1)}{p(y_1)}+$$

$$p(y_2/x_1)\log\frac{p(y_2/x_1)}{p(y_2)}+p(y_3/x_1)\log\frac{p(y_3/x_1)}{p(y_3)}\Big]+$$

$$+p(x_2)\Big[p(y_1/x_2)\log\frac{p(y_1/x_2)}{p(y_1)}+$$

$$p(y_2/x_2)\log\frac{p(y_2/x_2)}{p(y_2)}+p(y_3/x_2)\log\frac{p(y_3/x_2)}{p(y_3)}\Big]=$$

$$\frac{7}{16}\left[\frac{4}{7}\log_2\frac{4/7}{3/8}+\frac{1}{7}\log_2\frac{1/7}{1/4}+\frac{2}{7}\log_2\frac{2/7}{3/8}\right]+$$

$$\frac{9}{16}\left[\frac{2}{9}\log_2\frac{2/9}{3/8}+\frac{1}{3}\log_2\frac{1/3}{1/4}+\frac{4}{9}\log_2\frac{4/9}{3/8}\right]=0{,}0972\text{ бит.}$$

Пример 2.3.2. На вход приемника двоичных сигналов поступают посылки (1) и паузы (0). Априорные вероятности $p(0)=p(1)=1/2$. Из-за помех посылка, появившаяся на входе приемника, регистрируется в решающем устройстве правильно (как посылка) с вероятностью $\alpha=0{,}8$, ошибочно (как пауза) – с вероятностью $1-\alpha$. При поступлении паузы на вход приемника она принимается правильно (как пауза) с вероятностью $\beta = 0,6$ и ошибочно (как посылка) – с вероятностью $1-\beta$. Определить среднюю информацию о входном сигнале, содержащуюся в наблюдаемом сигнале на выходе приемника.

Решение. Обозначим X – ансамбль сигналов на входе, Y – ансамбль сигналов на выходе. Ищем среднюю взаимную информацию по формуле (2.3.8)

$$I(X;Y) = H(Y) - H(Y/X),$$

где $H(Y) = -\sum_{k=1}^{2} p(y_k)\log p(y_k)$ – энтропия на выходе,

$$H(Y/X) = -\sum_{j=1}^{2}\sum_{k=1}^{2} p(x_j)\ p(y_k/x_j)\log p(y_k/x_j)$$ – средняя

условная энтропия шума в канале, определяемая условными вероятностями искажений $p(y_k/x_j)$. По условию задачи они имеют следующие значения

$$p(1/1) = \alpha = 0,8, \qquad p(0/1) = 1-\alpha = 0,2,$$

$$p(0/0) = \beta = 0,6, \qquad p(1/0) = 1-\beta = 0,4.$$

Для вычисления вероятностей появления сигналов на выходе $p(Y=1)$ и $p(Y=0)$ воспользуемся формулой полной вероятности

$$p(y=1) = p(x=1)\cdot p(y=1/x=1) + p(x=0)\cdot p(y=1/x=0) =$$

$$0,5\cdot 0,8 + 0,5\cdot 0,4 = 0,6.$$

$$p(y=0) = p(x=0)p(y=0/x=0) + p(x=1)\cdot p(y=0/x=1) =$$

$$0,5\cdot 0,6 + 0,5\cdot 0,2 = 0,4.$$

Тогда средняя взаимная информация

$$I(X;Y) = H(Y) - H(Y/X) =$$

$$-p(y=0)\cdot\log p(y=0) - p(y=1)\cdot\log p(y=1) +$$

$$p(x=0)[\,p({}^{y=0}/_{x=0})\log p({}^{y=0}/_{x=0})+$$

$$p({}^{y=1}/_{x=0})\log p({}^{y=1}/_{x=0})]+$$

$$p(x=1)[\,p({}^{y=0}/_{x=1})\log p({}^{y=0}/_{x=1})+$$

$$p({}^{y=1}/_{x=1})\log p({}^{y=1}/_{x=1})]=$$

$$-\ \ 0,4\cdot\log_2 0,4+0,6\log_2 0,6\ \ +$$

$$\{0,5\ \ 0,6\cdot\log_2 0,6+0,4\cdot\log_2 0,4\ \ +$$

$$0,5\ \ 0,2\cdot\log_2 0,2+0,8\cdot\log_2 0,8\ \}=0,1245\text{ бит.}$$

Пример 2.3.3. Эргодический источник имеет алфавит, состоящий из 8 букв. Средние частоты повторения букв одинаковы. При передаче по каналу с шумом в среднем половина всех букв принимается правильно, в другой половине случаев имеют место ошибки, при этом любая буква переходит в любую другую с одинаковой вероятностью. Какова средняя информация в принятой букве относительно переданной?

Решение. Обозначим

$$X=\begin{Bmatrix} x_1 & x_2 & x_3 & \ldots & x_8 \\ p(x_1) & p(x_2) & p(x_3) & \ldots & p(x_8) \end{Bmatrix}-$$

ансамбль переданных букв,

$$Y=\begin{Bmatrix} y_1 & y_2 & y_3 & \ldots & y_8 \\ p(y_1) & p(y_2) & p(y_3) & \ldots & p(y_8) \end{Bmatrix}-$$

ансамбль принятых букв.

Так как для эргодической последовательности средние по времени частоты повторения букв совпадают с вероятностями, то по условию задачи вероятности появления букв на входе канала $p(x_j)=1/n=1/8$.

Ищем условные вероятности $p(y_k/x_j)$. Поскольку половина всех букв принимается правильно, то $p(y_k/x_j)=1/2$ при $k=j$.

Другая половина случаев есть ошибочный прием, причем по условию задачи все возможные ошибки равновероятны. Число возможных переходов (ошибок) равно 7. Тогда вероятность ошибки $p(y_k / x_j) = 0,5 \cdot 1/7 = 1/14$ при $k \neq j$.

Вероятности появления букв на выходе найдем по (1.3)

$$p(y_k) = \sum_{j=1}^{8} p(x_j)\, p(y_k / x_j) = \frac{1}{8} \cdot \frac{1}{2} + 7 \cdot \frac{1}{8} \cdot \frac{1}{14} = \frac{1}{8} \text{ для любого } \kappa.$$

Этот же результат следует непосредственно из того факта, что рассматриваемый канал – симметричный (набор вероятностей ошибок одинаков для любого *X*), тогда при равномерном распределении на входе распределение на выходе также равномерно.

Среднюю взаимную информацию находим по формуле (2.3.3)

$$I(X;Y) = \sum_{j=1}^{8} \sum_{k=1}^{8} p(x_j)\, p(y_k / x_j) \log \frac{p(y_k / x_j)}{p(y_k)} =$$

$$p(x_1)\Big[p(y_1 / x_1) \log \frac{p(y_1 / x_1)}{p(y_1)} + p(y_2 / x_1) \log \frac{p(y_2 / x_1)}{p(y_2)} + \ldots +$$

$$p(y_8 / x_1) \log \frac{p(y_8 / x_1)}{p(y_8)} \Big] + p(x_2)\ \cdot\ + \ldots + p(x_8)\ \cdot\ .$$

Выражения в квадратных скобках $\cdot$ численно равны, поэтому

$$I(X;Y) = \ \cdot\ \ p(x_1) + p(x_2) + \ldots + p(x_8) \ = \ \cdot\ =$$

$$\left[\frac{1}{2} \log \frac{1/2}{1/8} + 7 \cdot \frac{1}{14} \log \frac{1/14}{1/8} \right] = 0,5963 \text{ бит.}$$

ЗАДАЧИ

2.3.1. В эргодической последовательности, составленной из. букв алфавита А, В, С, Д, средние частоты повторения всех букв одинаковы, а связи между ними отсутствуют. Пусть половина всех букв принимается неверно, причем в результате ошибок поровну возникают все остальные буквы. Какова средняя информация в принятой букве относительно переданной буквы?

2.3.2. Алфавит состоит из 8 согласных и 8 гласных букв. Все буквы алфавита равновероятны и взаимные связи отсутствуют. Согласные принимаются правильно всегда, гласные – в половине случаев, в другой половине случаев имеют место ошибки, в результате которых гласные заменяются другими гласными, при этом в случае ошибки каждая гласная переходит в любую другую гласную с одинаковой вероятностью. Какова в среднем информация в принятой букве относительно переданной?

2.3.3. Источник создает последовательность из алфавита в 16 равновероятных и статистически независимых букв. При передаче по каналу с шумом буквы искажаются так, что четверть всех букв принимается неправильно, причем все ошибки одинаково вероятны. Определить среднюю информацию в принятой букве относительно переданной.

2.3.4. Сообщения источника $\boldsymbol{x_1}$ и $\boldsymbol{x_2}$ для передачи по каналу связи кодируются 0 и 1 соответственно. Кодовые символы равновероятны, $\boldsymbol{p}(0)$ =$\boldsymbol{p}(1)$. Шумы в канале вызывают: ошибки так, что в среднем 1 символ из 100 принимается неверно, причем ошибкам одинаково подвержены как нули, так и единицы. Найти информацию, содержащуюся в среднем в одном символе на входе приемника.

2.3.5. Вычислить среднюю взаимную информацию для канала:

а) заданного в задаче 2.1.12; б) заданного в задаче 2.1.13.

2.3.6. Для передачи по дискретному каналу показания телеизмерительного прибора кодируются 0 и 1. Вследствие действия помех на выходе канала появляются символы 0, -1,1. Вероятности совместного появления символов на входе $\boldsymbol{x_j}$ и выходе канала $\boldsymbol{y_к}$ приведены в табл. 2.3.1.

Таблица 2.3.1

	$y_к$		
j	1	y 2	3
		1	

$_1$	/4	/16	/8
$_2$	/8	3 /16	/4

Сколько информации содержится в среднем в принятом символе относительно переданного?

2.3.7. На вход приемника телеметрической системы поступает импульс с вероятностью $\boldsymbol{\alpha}$. При наличии импульса на входе вследствие действия помех решающее устройство дает ответ «да» с вероятностью $\boldsymbol{\beta}$ и «нет» с вероятностью 1- $\boldsymbol{\beta}$; при отсутствии импульса на входе – ответ «да» с вероятностью $\boldsymbol{\gamma}$ и «нет» с вероятностью 1– $\boldsymbol{\gamma}$. Определить среднее количество информации о переданном сигнале по наблюдению принятого сигнала: а) в общем виде,
б) при $\boldsymbol{\alpha} = 0{,}7$, $\boldsymbol{\beta} = 0{,}8$, $\boldsymbol{\gamma} = 0{,}4$.

2.3.8. Известно, что из 100 изготовленных деталей в среднем 10 деталей имеют дефекты. Для выявления брака используется метод, дающий всегда отрицательный ответ, если деталь изготовлена с браком. Если брак отсутствует, то деталь признается годной лишь в 80% случаев. Какое количество информации о качестве детали можно получить в среднем по результату такого метода отбраковки?

2.3.9. Для некоторого пункта вероятность того, что 15 июня будет идти дождь, равна 0,5, а вероятность дождя 15 октября равна 0,8. Используемый метод прогноза погоды на 15 июня оказывается правильным, в 3/5 тех случаев, когда предсказывается дождь, и в 4/5 случаев, когда предсказывается отсутствие осадков. Этот же метод прогноза погоды на 15 октября оказывается правильным в 9/10 случаев, когда предсказывается дождь, и в половине случаев, когда предсказывается отсутствие дождя. В какой из указанных дней прогноз дает в среднем больше информации о реальной погоде?

2.3.10. По дискретному каналу связи передаются сообщения «да» и «нет», приему которых соответствует появление зеленого или красного светового сигнала. Из-за воздействия шума приему зеленого сигнала в 4 случаях из 5 соответствует передача сигнала «да» и в одном случае – сигнала «нет», приему красного сигнала в 3 случаях из 5 соответствует передача сигнала «да» и в 2 – сигнала «нет». Определить среднее количество информации о переданном сигнале, содержащееся в принятом сообщении, если прием зеленого и красного сигналов одинаково вероятен.

2.3.11. В условиях задачи 2.3.10 определить количество информации, содержащееся в сообщении, если приему красного или зеленого сигнала соответствует один и тот же процент передачи сигналов «да» или «нет».

2.3.12. По дискретному каналу с помехами передаются равновероятные и статистически независимые телеграфные посылки (+) и паузы (–). В результате действия помех на выходе приемника могут регистрироваться символы 0, +, –. Условные вероятности заданы табл. 2.3.2.

Таблица 2.3.2

j	y_k		
	y_1	y_2	y_3
1	1/32	61/64	1/64
2	1/32	1/64	61/64

Какова в среднем взаимная информация в принятом символе относительно переданного?

2.3.13. Радиостанция противника может работать на частотах f_1 и f_2 (события x_1 и x_2) в режимах АТ и ЧТ (события y_1 и y_2). Совместные плотности вероятностей заданы табл. 2.3.3.

Таблица 2.3.3

y_j	x_k	
	x_1	x_2
y_1	0.7	0.05
y_2	0.1	0.15

Вычислить среднее количество информации относительно режима работы станции, если стала известна ее частота.

2.3.14. Студент может получить зачет с вероятностью 0,3, не проработав весь материал, и с вероятностью 0,9, проработав весь материал курса. Какое количество информации о подготовленности студента к зачету можно получить по данным о результатах сдачи зачета? В среднем 90% студентов готовы к сдаче зачета.

2.3.15. Имеется последовательность, состоящая из ***n*** двоичных символов $X^{(1)},...,X^{(n)}$. Возможно появление только тех реализаций, которые содержат четное число единиц, и все такие реализации равновероятны:
а) вычислить вероятность появления реализации, состоящей из одних нулей;
б) найти средние взаимные информации
$I(X^{(2)};X^{(1)})$; $I(X^{(3)};X^{(1)},X^{(2)})$; $I(X^{(n)};X^{(1)},...,X^{(n-1)})$;

в) проверить полученные результаты при ***n*** = 3.

2.3.16. Запись погоды в некотором городе дается в приводимой ниже табл. 2.3.4, числа указывают относительную частоту соответствующих событий.

Таблица 2.3.4

Предсказание	На самом деле	
	Дождь	Нет дождя
Дождь	1/8	3/16
Нет дождя	1/16	10/16

Умный студент заметил, что составитель прогноза прав только в 12/16 случаев, но мог бы быть правым в 13/16 случаев, предсказывая всегда отсутствие дождя. Студент объяснил ситуацию и написал заявление о приеме на работу в бюро прогнозов, но начальник бюро прогнозов, который, разумеется, был специалистом по теории информации, отклонил его заявление. Почему?

2.3.17. Двоичный стирающий канал является одним из наиболее простых типов канала с шумом. В нем переданные символы могут быть «стертыми», но никогда не могут быть приняты ошибочно. Найти среднее количество информации, переносимое одним символом в таком канале, если вероятность стирания равна 0,1 и не зависит от переданного символа.

2.3.18. Докажите, что никакое преобразование реализаций случайной величины ***Y*** не может увеличить количество информации, содержащейся в ней относительно другой случайной величины ***X***.

2.4 Информационные характеристики случайных последовательностей

Энтропия. Пусть имеется отрезок случайной функции $X(t)$, дискретной по времени и амплитуде, и он полностью определяется последовательностью отсчетов $X^{(1)},\ldots,X^{(n)}$, взятых в моменты времени $t_1,\ldots,t_n$. Тогда энтропия отрезка этой случайной функции длиною в n отсчетов (m-ичных символов) равна

$$H_n\left[X\;\;t\;\;\right]=$$

$$-\sum_{j=1}^{m}\cdots\sum_{k=1}^{m}p\;\;x_j^{1},\ldots,x_k^{n}\;\;\log p\;\;x_j^{1},\ldots,x_k^{n}\;\;. \qquad (2.4.1)$$

Для стационарных случайных последовательностей используют более удобную характеристику, называемую энтропией случайной последовательности

$$H\left[X\;\;t\;\;\right]=\lim_{n\to\infty}\frac{H_n\left[X\;\;t\;\;\right]}{n}\;\;\frac{\text{бит}}{\text{симв}}. \qquad (2.4.2)$$

Эта величина имеет смысл средней энтропии, приходящейся на один отсчёт (символ) в реализациях бесконечно большой длины.

Энтропия случайной последовательности удовлетворяет неравенствам (2.4.3) и (2.4.4).

$$H\left[X\;\;t\;\;\right]\le H_1\left[X\;\;t\;\;\right], \qquad (2.4.3)$$

где $H_1\left[X\quad t\quad\right]$ – энтропия одного из отсчетов.

Энтропия случайной последовательности достигает наибольшего значения (знак равенства в (2.4.3)) только тогда, когда отдельные отсчеты функции *X(t)* (т.е. случайные величины $X^{(1)},\ldots,X^{(n)}$) статистически независимы (см. формулу (2.2.4)).

$$H_1\left[X\quad t\quad\right]\le H_{1\max}, \tag{2.4.4}$$

где $H_{1\max}$ - максимальное значение энтропии одного символа. Энтропия случайной последовательности максимальна, когда все возможные значения символов равновероятны (см. формулу (2.2.3)). В итоге получаем

$$H\left[X\quad t\quad\right]\le \log m. \tag{2.4.5}$$

Можно найти и среднюю условную энтропию стационарной случайной последовательности *X(t)*. Она вычисляется при условии, что известно значение другой дискретной случайной последовательности *Y(t)*, также стационарной и стационарно связанной с *X(t)*:

$$H\left[X\quad t\ /Y\quad t\quad\right]=\lim_{n\to\infty}\frac{H_n\left[X\quad t\ /Y\quad t\quad\right]}{n}\quad\frac{\text{бит}}{\text{симв}}, \tag{2.4.6}$$

где

$$H_n\left[X\quad t\ /Y\quad t\quad\right]=$$

$$-\sum_{j=1}^{m}\cdots\sum_{k=1}^{m}\sum_{r=1}^{m}\ldots\sum_{q=1}^{m}p\quad x_j^{1},\ldots,x_k^{n},y_r^{1},\ldots,y_q^{n}\quad\cdot \tag{2.4.7}$$

$$\log p\quad x_j^{1},\ldots,x_k^{n},y_r^{1},\ldots,y_q^{n}\quad .$$

Взаимная информация. Среднее количество информации, содержащейся в отрезке последовательности *Y(t)* относительно отрезка *X(t)*, равно

$$I_n\left[X\quad t\ ;Y\quad t\quad\right]=H_n\left[X\quad t\quad\right]-H_n\left[X\quad t\ /Y\quad t\quad\right],$$

а средняя информация, приходящаяся на один отсчет (символ)

$$I\left[X\quad t\ ;Y\quad t\quad\right]=H\left[X\quad t\quad\right]-H\left[X\quad t\ /Y\quad t\quad\right]. \tag{2.4.8}$$

Возможен и другой способ вычисления средней энтропии (условной и безусловной) дискретной стационарной случайной последовательности, приводящий к тем же результатам.

Энтропия последовательности *X(t)* равна

$$H\left[X\quad t\ \right]=\lim_{n\to\infty}H\left[X\ ^{n}\big/X\ ^{1}\ ,\ldots,X\ ^{n-1}\ \right] \tag{2.4.9}$$

и характеризует среднее количество дополнительной информации, доставляемой одним символом в реализациях бесконечной длины.

Средняя условная энтропия последовательности $X(t)$ вычисляется по формуле

$$H\left[X\quad t\ \big/Y\quad t\ \right]= \tag{2.4.10}$$

$$\lim_{n\to\infty}H\left[X\ ^{n}\big/X\ ^{1}\ ,\ldots,X\ ^{n-1}\ ,Y\ ^{1}\ ,\ldots,Y\ ^{n}\ \right].$$

РЕШЕНИЕ ТИПОВЫХ ПРИМЕРОВ

Пример 2.4.1. Сообщение *X* есть стационарная последовательность независимых символов, имеющих ряд распределения

x_j	x_1	x_2	x_3
$p(x_j)$	0,2	0,1	0,7

Сигнал *Y* является последовательностью двоичных символов, связанных с сообщением следующим неудачным образом: $x_1\to 00$, $x_2\to 00$, $x_3\to 1$.

Определить: а) средние безусловную и условную энтропии, приходящиеся на 1 символ сообщения;

б) средние безусловную и условную энтропии, приходящиеся на 1 символ сигнала;

в) среднюю взаимную информацию в расчете на 1 символ сообщения.

Решение. Символы в последовательности на выходе источника независимы, поэтому из (2.4.3)

$$H\left[X\quad t\ \right]=H_1\left[X\quad t\ \right]=H\quad X\quad =$$

$$-0,2\cdot\log 0,2-0,1\cdot\log 0,1-0,7\cdot\log 0,7=1,1568\ \text{бит/симв.}$$

Обозначив y_1=00, y_2=1, вычислим условные и безусловные вероятности сигнала:

$$p\quad y_1/x_1\quad =p\quad y_1/x_2\quad =1,\qquad p\quad y_1/x_3\quad =0,$$

$$p\quad y_2/x_1\quad =p\quad y_2/x_2\quad =0,\qquad p\quad y_2/x_3\quad =1,$$

$$p(y_1) = p(x_1)\,p(y_1/x_1) + p(x_2)\,p(y_1/x_2) + p(x_3)\,p(y_1/x_3) = 0,2 + 0,1 = 0,3,$$

$$p(y_2) = 0,7.$$

Энтропия случайной величины Y равна

$$H(Y) = -0,3\cdot\log 0,3 - 0,7\cdot\log 0,7 = 0,8813\text{ бит},$$

а условная энтропия $H(Y/X)$=0, так как сигнал однозначно определяется сообщением.

Взаимная информация

$$I(X;Y) = H(Y) - H(Y/X) = 0,8813\text{ бит}.$$

Условная энтропия сообщения

$$H(X/Y) = H(X) - I(X;Y) = 1,1568 - 0,8813 = 0,2755\text{ бит}.$$

Итак, получаем, что условная энтропия сообщения равна 0,2755 бит/симв, а взаимная информация в расчете на 1 символ сообщения равна 0,8813 бит/симв.

Среднее количество символов сигнала, приходящихся на 1 символ сообщения, равно

$$L = \sum_{j=1}^{3} p(x_j)\,l_j = 0,2\cdot 2 + 0,1\cdot 2 + 0,7\cdot 1 = 1,3\text{ симв},$$

где l_j – длина кодового слова, соответствующего x_j.

Энтропия последовательности $Y(t)$ равна

$$H\left[Y(t)\right] = H(Y)/L = 0,8813/1,3 = 0,6779\text{ бит/симв},$$

а условная энтропия равна нулю.

Пример 2.4.2. Вычислить энтропию однородной неразложимой марковской последовательности, заданной матрицей вероятностей переходов $p(X^{k}/X^{k-1})$ от символа X^{k-1} к символу X^{k} (табл. 2.4.1).

Таблица 2.4.1

$X^{(k)}$	$X^{(k-1)}$		
	x_1	x_2	x_3
x_1	0,2	0,3	0,6
x_2	0,4	0,5	0,1
x_3	0,4	0,2	0,3

Решение. Однородность последовательности проявляется в том, что при $k\to\infty$ безусловная вероятность p_j того, что k-й символ примет значение x_j, не зависит от k и от значения первого символа и связана с вероятностями переходов $p_{rj} = p\left(x_j^{\,k} / x_r^{\,k-1}\right)$

уравнением $p_j = \sum_{r=1}^{m} p_r p_{rj}$.

Записываем это уравнение для каждого j и получаем систему уравнений

$$p_1 = p_1 p_{11} + p_2 p_{21} + p_3 p_{31},$$

$$p_2 = p_1 p_{12} + p_2 p_{22} + p_3 p_{32},$$

$$p_3 = p_1 p_{13} + p_2 p_{23} + p_3 p_{33}.$$

Подставляем численные значения вероятностей переходов и получим

$$-0{,}8 p_1 + 0{,}3 p_2 + 0{,}6 p_3 = 0,$$

$$0{,}4 p_1 - 0{,}5 p_2 + 0{,}1 p_3 = 0,$$

$$0{,}4 p_1 + 0{,}2 p_2 - 0{,}7 p_3 = 0.$$

Определитель этой системы равен нулю, поэтому вероятности p_j определяется с точностью до постоянного множителя. Выбираем его так, чтобы выполнялось условие нормировки $p_1 + p_2 + p_3 = 1$, и получим

$$p_1 = 0{,}3548, \quad p_2 = 0{,}3441, \quad p_3 = 0{,}3011.$$

Энтропия последовательности $H[X(t)]$ по смыслу есть средняя дополнительная (условная) информация, приносимая одним символом последовательности, поэтому для марковской последовательности ее можно вычислить по формуле (2.4.9)

$$H\left[X(t)\right] = H\left(x^{k} / X^{k-1}\right) = \sum_{j=1}^{m} p_j H\left(X^{k} / X^{k-1}\right) =$$

$$= \sum_{j=1}^{m} p_j \left(\sum_{r=1}^{m} p_{jr} \log p_{jr} \right) = 0{,}3548 \cdot 1{,}5219 + 0{,}3441 \cdot 1{,}4855 +$$

$$0{,}3011 \cdot 1{,}2955 = 1{,}4412 \text{ бит/симв.}$$

ЗАДАЧИ

2.4.1. Периодически, один раз в 5 с, бросают игральную кость, и результат каждого бросания передают при помощи трехразрядного двоичного числа. Найти:
а) энтропию шестеричной последовательности на входе,
б) энтропию двоичной последовательности на выходе,
в) среднюю информацию, содержащуюся в отрезке выходного сигнала длительностью 1 мин.

2.4.2. Известно, что энтропия, приходящаяся на 1 букву русского текста, составляет приблизительно 1,2 бит. Каково минимальное среднее количество десятичных символов, необходимых для передачи информации, содержащейся в телеграмме из 100 букв?

2.4.3. Периодически проводятся тиражи лотереи. В каждом тираже участвуют 1024 билета, но выигрывает из них один. Результат каждого тиража передается при помощи 10-разрядного двоичного числа. Из-за наличия помех возможны искажения символов с вероятностью 0,001. Ошибки в отдельных символах независимы. Найти:
а) энтропию исходной 1024-ичной последовательности,
б) энтропию двоичной последовательности на входе,
в) энтропию двоичной последовательности на выходе,
г) энтропию выходной 1024-ичной последовательности,
д) среднее количество передаваемой информации в расчете на двоичный символ входной последовательности.

2.4.4. В условиях задачи 2.4.3 для повышения помехоустойчивости каждый двоичный символ передается трехкратно. В месте приема из каждых трех символов формируется один, причем решение выносится простым «большинством голосов». Найти:
а) энтропию двоичной последовательности на выходе,
б) среднее количество информации в расчете на передаваемый двоичный символ.

2.4.5. В некотором районе состояние погоды зависит только от того, что было накануне. При ежедневной регистрации погоды различают лишь два состояния: x_1 – ясно и x_2 – переменно. Вероятности переходов равны

$$p\ \ x_1/x_1\ \ = 0.9 \qquad p\ \ x_2/x_1\ \ = 0.1$$
$$p\ \ x_1/x_2\ \ = 0.7 \qquad p\ \ x_2/x_2\ \ = 0.3$$

Вычислить энтропию последовательности сообщений о состоянии погоды.

2.4.6. Найти энтропию последовательности на выходе троичного марковского источника, если вероятности всех переходов равны 1/3.

2.4.7. Буквы в последовательности на выходе источника выбираются из алфавита A, B, C и образуют марковскую последовательность. Совместные вероятности всевозможных пар букв заданы в табл. 2.4.2.

Таблица 2.4.2

(k)	$X^{(k+1)}$		
	A	B	C
	0,2	0,05	0,15
	0,15	0,05	0,1
	0,05	0,2	0,05

Вычислить энтропию последовательности.

2.4.8. Троичная марковская последовательность задана матрицей вероятности переходов ***P*** от $X^{(k)}$ (строка) к $X^{(k+1)}$ (столбец)

$$P = \begin{Vmatrix} 0 & 3/4 & 1/4 \\ 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \end{Vmatrix}$$

Вычислить:

а) энтропию одного символа,
б) энтропию последовательности и избыточность.

2.4.9. На кольцевой трассе расположено 10 станций. По трассе в направлении часовой стрелки курсирует автобус, делая остановки на каждой станции. Исключение составляет только станция № 1, с которой с вероятностью 0,8 автобус может вернуться «а предыдущую станцию. С каждой станция-водитель посылает сообщение о направлении дальнейшего движения. Найти среднюю информацию, содержащуюся в одном сообщении, и избыточность.

2.4.10. На водохранилище имеются 3 контрольные отметки. Сброс воды из водохранилища производится периодически один раз в год, а в промежутках между этими моментами оно заполняется. Поступление воды поднимает уровень на одну отметку с вероятностью 0,7 и с вероятностью 0,3 сохраняет его неизменным. Величины поступлений в различные годы независимы. Если уровень воды достигает отметки № 3, то после сброса он понижается до отметки № 1; в остальных случаях сброс не производится. Найти среднее количество информации, содержащейся в сообщениях об уровне воды.

2.4.11. На входе канала связи имеется последовательность двоичных независимых равновероятных символов. Ошибки в канале приводят к изменению значений некоторых символов на обратные, причем вероятность ошибки в k–м символе зависит лишь от наличия ошибки в предыдущем, (k–1)-м символе, а именно, она равна 0,2 при наличии ошибки в (k–1)-м символе и равна 0,05 при отсутствии ее. Найти среднюю величину передаваемой информации в расчете на символ.

3 ИНФОРМАЦИОННЫЕ ХАРАКТЕРИСТИКИ АНАЛОГОВЫХ СИГНАЛОВ

3.1 Взаимная информация

Определение непрерывного канала. Непрерывным (аналоговым) называется канал, в котором сигналы на входе $x(t)$ и выходе $y(t)$ являются реализациями непрерывных случайных функций.

В дальнейшем в качестве входного и выходного сигналов будут рассматриваться только такие функции, спектр которых заключен в ограниченной полосе частот от 0 до F_B. Такие функции полностью определяются последовательностью отсчетов, взятых с шагом $\Delta t = 1/2F_B$.

Для полного описания канала в фиксированный момент времени t_1 достаточно задать совместную плотность вероятностей $W(x, y)$ системы двух отсчетов, т.е. непрерывных случайных величин $X=X(t_1)$ и $Y=Y(t_1)$. Часто для расчетов более удобными являются следующие характеристики:

$W(x)$ – одномерная плотность вероятности сообщения,

$W(y)$ – одномерная плотность вероятности сигнала,

$W(x/y)$ – условная плотность вероятности сообщения x при известном y,

$W(y/x)$ – условная плотность вероятности сигнала y при заданном x.

Соотношения, связывающие эти плотности вероятности, см. в разд. 1, формулы (1.3) – (1.5).

Обобщим понятие взаимной информации на непрерывные сигналы.

Взаимная информация. По аналогии с дискретным случаем взаимной информацией называют величину

$$I(x;y) = \log\frac{W(x,y)}{W(x)W(y)} = \log\frac{W(x/y)}{W(x)} = \log\frac{W(y/x)}{W(y)}. \qquad (3.1.1)$$

Свойства взаимной информации $I(x;y)$ для непрерывных случайных величин x и y аналогичны свойствам взаимной информации $I(x_j;y_k)$ для дискретных случайных величин, за исключением свойства 2, связанного с понятием собственной информации, неприменимым к непрерывной случайной величине (см. далее в разд. 3.2).

Средняя взаимная информация. Рассмотрим среднее значение взаимной информации для объединенного ансамбля XY. Здесь возможны следующие случаи. Средняя информация, содержащаяся в реализации y принятого сигнала относительно ансамбля передаваемых сообщений X, есть

$$I(X;y)=\mathrm{M}\left[\log\frac{W(X/y)}{W(X)}\right]=\int_{-\infty}^{\infty}W(x/y)\,I(x;y)\,dx=$$
$$\int_{-\infty}^{\infty}W(x/y)\log\frac{W(x/y)}{W(x)}dx. \qquad (3.1.2)$$

Иначе, это – средняя взаимная информация между ансамблем X и реализацией y.

Средняя информация, содержащаяся в реализации передаваемого сообщения x относительно ансамбля Y, равна

$$I(Y;x)=\mathrm{M}\left[\log\frac{W(Y/x)}{W(Y)}\right]=\int_{-\infty}^{\infty}W(y/x)\,I(x;y)\,dy=$$
$$\int_{-\infty}^{\infty}W(y/x)\log\frac{W(y/x)}{W(y)}dy \qquad (3.1.3)$$

Это – средняя взаимная информация между ансамблем $\boldsymbol{Y}$ и реализацией $\boldsymbol{x}$.

Средняя информация о непрерывной случайной величине X, заключенная в непрерывной случайной величине Y,

$$I(X;Y)=\mathrm{M}\left[\log\frac{W(X/Y)}{W(X)}\right]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}W(x,y)\log\frac{W(x/y)}{W(x)}dxdy \qquad (3.1.4)$$

есть средняя взаимная информация между ансамблями X и Y.

Средняя взаимная информация неотрицательна, обладает свойством симметрии и связана с относительными энтропиями (см. разд. 3.2) непрерывных случайных величин X и Y следующим образом:

$$I(X;Y)=H(X)-H(X/Y)=H(Y)-H(Y/X). \qquad (3.1.5)$$

РЕШЕНИЕ ТИПОВЫХ ПРИМЕРОВ

Пример 3.1.1. Вольтметром измеряется напряжение в электрической сети. Ошибка измерения не зависит от истинного значения напряжения и распределена нормально с нулевым математическим ожиданием и среднеквадратическим отклонением, равным 2 В. Истинное значение напряжения в сети также имеет нормальное распределение с математическим ожиданием 220 В и СКО, равным 10 В.

Найти:

а) зависимость величины получаемой информации от показаний прибора,

б) среднюю величину получаемой информации.

Решение. Введем обозначения:

X – напряжение в сети,

V – ошибка измерения,

Y = X + V – показание прибора.

Из условия задачи записываем плотности вероятностей:

$$W(x) = \frac{1}{\sqrt{2\pi}\sigma_x}\exp\left[-\frac{(x-m_x)^2}{2\sigma_x^2}\right],$$

$$W(y/x) = \frac{1}{\sqrt{2\pi}\sigma_v}\exp\left[-\frac{(y-x)^2}{2\sigma_v^2}\right].$$

Безусловную плотность вероятности величины *Y* можно найти по общей формуле (1.3), но проще поступить следующим образом. Показание прибора *Y* есть сумма двух независимых нормальных случайных величин и, следовательно, *Y* также имеет нормальное распределение с математическим ожиданием

$$m_y = m_x + m_v = m_x = 220 \text{ В}$$

и дисперсией

$$\sigma_y^2 = \sigma_x^2 + \sigma_v^2 = 104 \text{ В}^2.$$

Итак,

$$W(y) = \frac{1}{\sqrt{2\pi}\sigma_y}\exp\left[-\frac{(y-m_x)^2}{2\sigma_y^2}\right].$$

Условную плотность вероятности *x* находим по формуле (1.4)

$$W(x/y) = \frac{W(y/x)\,W(x)}{W(y)} =$$

$$\frac{\sqrt{2\pi}\sigma_y}{\sqrt{2\pi}\sigma_v\sqrt{2\pi}\sigma_x}\exp\left\{-\frac{1}{2}\left[\frac{(y-x)^2}{\sigma_v^2}+\frac{(x-m_x)^2}{\sigma_x^2}-\frac{(y-m_x)^2}{\sigma_y^2}\right]\right\}.$$

В выражении, стоящем под знаком экспоненты, осуществляем возведение в квадрат и затем группируем члены, содержащие x^2 и x. В итоге убеждаемся, что условное распределение величины x также нормально

$$W(x/y)=\frac{1}{\sqrt{2\pi}\delta}\exp\left\{-\frac{\left[x-S(y)\right]^2}{2\delta^2}\right\},$$

где $S(y)=m_x+\dfrac{\sigma_x^2}{\sigma_y^2}(y-m_x)$ – условное математическое ожидание X,

$\delta^2=\dfrac{\sigma_x^2\sigma_v^2}{\sigma_y^2}$ - условная дисперсия X.

Для ответа на вопрос п. а) следует по формуле (3.1.2) вычислить величину средней взаимной информации между ансамблем X и реализацией y

$$I(X;y)=\mathrm{M}\left[\ln\frac{W(X/y)}{W(X)}\right]=$$

$$\mathrm{M}\left\{\ln\frac{\sigma_x}{\delta}-\frac{\left[X-S(y)\right]^2}{2\delta^2}+\frac{(X-m_x)^2}{2\sigma_x^2}\right\}.$$

Находим квадраты разностей и выносим за знак математического ожидания слагаемые и множители, не зависящие от x. Далее учитываем, что вычисляются условные математические ожидания при конкретном значении y, поэтому

$$\mathrm{M}(X)=S(y),\quad \mathrm{M}\left[X^2\right]=S^2(y)+\delta^2.$$

В итоге получаем

$$I(X;y)=\frac{1}{2}\ln\frac{\sigma_y^2}{\sigma_v^2}+\frac{1}{2}\frac{\sigma_x^2}{\sigma_y^2}\left[\left(\frac{y-m_x}{\sigma_y}\right)^2-1\right]=$$

$$\frac{1}{2}\ln\frac{104}{4}+\frac{100}{2\cdot 104}\left[\frac{(y-220)^2}{104}-1\right]=$$

$$1{,}629+0{,}4801\left[\frac{(y-220)^2}{104}-1\right] \text{ нат.}$$

Таким образом, искомая зависимость есть параболическая функция разности $(y-m_x)$, причем наименьшее количество информации, равное 1,1482 нат, доставляет наиболее вероятное показание прибора $y=m_x=220\ B$.

Для ответа на вопрос п. б) необходимо найти среднюю величину взаимной информации между ансамблями *X* и *Y*. Вычисляем безусловное математическое ожидание величины *I*(*X*; *y*). При этом учитываем, что $\mathrm{M}\left[(Y-m_x)^2\right]=\sigma_y^2$, и получаем

$$I(X;Y)=\frac{1}{2}\ln\frac{\sigma_y^2}{\sigma_v^2}=\frac{1}{2}\ln\left(1+\frac{\sigma_x^2}{\sigma_v^2}\right)=1{,}629 \text{ нат.}$$

Обратите внимание, что в среднем количество получаемой информации в данном примере зависит только от отношения «сигнал/ошибка» σ_x/σ_v.

Пример 3.1.2. Для контроля исправности блока периодически измеряют значение напряжения в контрольной точке схемы. Если блок исправен, то это напряжение равно 1 В, в противном случае – 0 В. Ошибка вольтметра распределена равномерно с нулевым математическим ожиданием, но ширина этого распределения зависит от величины измеряемого напряжения: она равна 2 В, если напряжение на входе 1 В, и 1 В – в противном случае. В среднем в 90% времени блок исправен.

Найти количество информации, доставляемой прибором при одном измерении.

Решение. Введем обозначения:

X – измеряемое напряжение, Y – показание вольтметра.

Из условия задачи следует, что величина X – двоичная, причем $x_1 = 1$; $x_2 = 0$; $p(x_1) = 0{,}9$; $p(x_2) = 0{,}1$.

Сигнал Y есть непрерывная случайная величина, и для нее заданы условные плотности вероятности:

$$W(y/x_1) = \begin{cases} 0{,}5, & 0 \le y \le 2, \\ 0, & y < 0,\ y > 2, \end{cases}$$

$$W(y/x_2) = \begin{cases} 1, & -0{,}5 \le y \le 0{,}5, \\ 0, & y < -0{,}5,\ y > 0{,}5. \end{cases}$$

Рассматриваемый канал является дискретно-непрерывным, так как входной сигнал дискретен, а выходной непрерывен. Требуется найти среднюю взаимную информацию $I(X;Y)$.

Найдем безусловную плотность вероятности выходного сигнала:

$$W(y) = p(x_1)W(y/x_1) + p(x_2)W(y/x_2) =$$

$$\begin{cases} 0, & y < -0{,}5, \\ 0{,}1, & -0{,}5 \le y < 0, \\ 0{,}55, & 0 \le y \le 0{,}5, \\ 0{,}45, & 0{,}5 < y \le 2, \\ 0, & 2 < y. \end{cases}$$

Далее вычисляем величину средней взаимной информации, учитывая соотношение (3.1.3)

$$I(X;Y) = p(x_1)I(Y;x_1) + p(x_2)I(Y;x_2) =$$

$$p(x_1)\int_0^2 W(y/x_1)\log\frac{W(y/x_1)}{W(y)}dy +$$

$$p(x_2)\int_{-0.5}^{0.5} W(y/x_2)\log\frac{W(y/x_2)}{W(y)}dy =$$

$$0,9\left(\int_{0}^{0,5}0,5\cdot\log\frac{0,5}{0,55}dy+\int_{0,5}^{2}0,5\cdot\log\frac{0,5}{0,45}dy\right)+$$

$$0,1\left(\int_{-0,5}^{0}1\cdot\log\frac{1}{0,1}dy+\int_{0}^{0,5}1\cdot\log\frac{1}{0,55}dy\right)=0,2809\text{ бит.}$$

ЗАДАЧИ

3.1.1. Автобусы ходят ровно через 10 мин. Пассажир случайно вышел на остановку и узнал, что в последние 7,5 мин, автобуса не было. Сколько информации он получил?

3.1.2. Погрешность фазометра распределена нормально со с.к.о. 3°. Найти количество информации, получаемой при измерении значения начальной фазы радиосигнала, если она может с одинаковой вероятностью принять любое значение.

3.1.3. Некто обладает развитым чувством времени и может без часов угадать текущее значение времени. Указанное им значение всегда бывает меньше истинного и распределено по экспоненциальному закону со средним 10 мин. Человек посмотрел, на часы, о которых известно, что они всегда отстают, причем ошибка их показаний распределена также экспоненциально со средним 30 с. Сколько информации он при: этом получил?

3.1.4. Найти величину взаимной информации между исходной и квантованной величинами в задаче 1.5.

3.1.5. Вычислить взаимную информацию между случайными величинами:
а) из задачи 1.10; б) из задачи 1.13.

3.1.6. Случайные величины X и Y имеют совместное нормальное распределение. Найти зависимость средней взаимной информации от коэффициента корреляции.

3.1.7. Вес изделия из большой партии есть случайная величина, распределенная по нормальному закону со с.к.о., равным 10 кг. Из партии наугад берут изделие и 100 раз взвешивают на стрелочных весах со среднеквадратичной погрешностью 100 г. Результаты всех взвешиваний суммируют.

Вычислить количество информации, содержащейся в суммарном результате.

3.1.8. Сигнал на выходе радиолокационного приемника является суммой отраженного от цели сигнала и независимого внутреннего нормального шума, имеющего нулевое среднее к с.к.о. 0,1 В. Амплитуда отраженного от цели сигнала равна 5 В.

Найти величину получаемой информации, если вероятность появления цели в зоне действия радиолокатора равна 0,01?

3.1.9. Сигнал; имеющий функцию распределения

$$F(x) = \begin{cases} 0, & x < 0, \\ 0.01x^2, & 0 \le x \le 10, \\ 1, & 10 < x. \end{cases}$$

подвергается равномерному квантованию с шагом, разным 2. Сколько информации содержится в квантованном сигнале?

3.1.10. Сигнал X, имеющий нормальное распределение, одновременно передается по двум параллельным каналам. В каждом канале действует аддитивный нормальный шум Q с нулевым средним и с.к.о. $\sigma_q = k\,\sigma_x$. Шумы в каналах независимы между собой и от входного сигнала:

а) вычислить среднюю взаимную информацию между входным сигналом X и парой выходных сигналов YZ;

б) привести пример преобразования двух выходных сигналов в один G = f(Z, Y), не разрушающего информации, т. е. такого, что

I(X; G) = I(X; YZ).

3.1.11. * Сообщение $X = (X^1, \ldots, X^n)$ и сигнал $Y = (Y^1, \ldots, Y^n)$ в совокупности образуют систему 2n нормальных случайных величин с единичными с.к.о. Найти среднюю взаимную информацию между подсистемами X и Y, если заданы корреляционные матрицы R, B и Q, имеющие элементы:

$$R_{ik} = M\left[X^i X^k\right], \quad B_{ik} = M\left[Y^i Y^k\right],$$

$$Q_{ik} = M\left[X^i Y^k\right], \qquad i, k = 1, \ldots, n.$$

3.2 Относительная (дифференциальная) энтропия

Определения. Из формулы (2.2.3) видно, что даже в дискретном случае при бесконечном числе состояний энтропия не имеет верхней грани. Тогда неопределенность реализации одного из бесконечного множества состояний (т.е. непрерывной случайной величины) может быть сколь угодно велика – непрерывные случайные величины не допускают введения конечной абсолютной меры неопределенности. Поэтому вводится относительная мера неопределенности – относительная энтропия

$$H_{д}(X) = \mathrm{M}\left[-\log W(X)\right] = -\int_{-\infty}^{\infty} W(x)\log W(x)\,dx. \qquad (3.2.1)$$

Благодаря связи с дифференциальным законом распределения вероятностей ее называют еще дифференциальной энтропией (в дальнейшем изложении индекс д и слово «дифференциальная» будем опускать, но помнить, что если речь идет о непрерывной случайной величине, то ее энтропия $H(X)$ есть величина относительная).

Аналогично для относительной энтропии непрерывной случайной величины Y имеем

$$H(Y) = -\int_{-\infty}^{\infty} W(y)\log W(y)\,dy. \qquad (3.2.2)$$

Для средней условной энтропии имеем

$$H(X/Y) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(x,y)\log W(x/y)\,dxdy, \qquad (3.2.3)$$

$$H(Y/X) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(x,y)\log W(y/x)\,dxdy. \qquad (3.2.4)$$

Свойства. Энтропия непрерывных величин обладает свойствами, во многом аналогичными свойствам энтропии дискретных случайных величин.

1) Относительная энтропия может принимать любые значения (сравнить с (2.2.2))

$$-\infty < H(X) < \infty. \qquad (3.2.5)$$

2) Свойство аддитивности.

Энтропия объединения двух непрерывных случайных величин

$$H(XY)=H(X)+H(Y/X)=H(Y)+H(X/Y), \qquad (3.2.6)$$

если X и Y статистически связаны между собой. Здесь обозначено

$$H(XY)=-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}W(x,y)\log W(x,y)\,dxdy. \qquad (3.2.7)$$

Энтропия объединения двух независимых непрерывных случайных величин равна сумме их энтропий

$$H(XY)=H(X)+H(Y). \qquad (3.2.8)$$

3) Относительная условная энтропия не превышает безусловной

$$H(X/Y)\le H(X), \qquad H(Y/X)\le H(Y). \qquad (3.2.9)$$

4) Свойство экстремальности энтропии.

Среди всех непрерывных случайных величин, удовлетворяющих условиям:

$$\text{а) } \int_{-\infty}^{\infty}\varphi(x)W(x)\,dx=\theta, \qquad (3.2.10)$$

$$\text{б) } \int_{-\infty}^{\infty}W(x)\,dx=1 \text{ (условие нормировки),}$$

наибольшей энтропией обладает случайная величина X, имеющая плотность вероятности

$$W(x)=\exp\left[\lambda_1\varphi(x)+\lambda_2\right]. \qquad (3.2.11)$$

Коэффициенты λ_1 и λ_2 выбираются так, чтобы функция $W(x)$ удовлетворяла условиям а) и б).

РЕШЕНИЕ ТИПОВЫХ ПРИМЕРОВ

Пример 3.2.1. Положительная непрерывная случайная величина X распределена по экспоненциальному закону с математическим ожиданием m_x=3. Вычислить значение дифференциальной энтропии величины X.

Решение. Заданный закон распределения имеет вид

$$W(x) = \frac{1}{m_x}\exp\left(-\frac{x}{m_x}\right), x \geq 0.$$

Дифференциальная энтропия непрерывной случайной величины X определяется формулой (3.2.1)

$$H(X) = \mathrm{M}\left[-\log W(x)\right] = \mathrm{M}\left[-\log\frac{1}{m_x}\exp\left(-\frac{x}{m_x}\right)\right] =$$

$$\log m_x + \frac{\log e}{m_x}\mathrm{M}(X) = \log m_x + \frac{\log e}{m_x}m_x = \log(em_x).$$

Подставляя $m_x = 3$, получим

$$H(X) = \log(2,7183\cdot 3) = 3,0277 \text{ бит}.$$

Пример 3.2.2. Среди всех случайных величин, которые могут принимать значения из интервала (a, $в$) и имеют следующие числовые характеристики:

$$\int_a^b \psi_1(x)W(x)\,dx = \theta_1, \quad \int_a^b \psi_2(x)W(x)\,dx = \theta_2, \ldots,$$

$$\int_a^b \psi_k(x)W(x)\,dx = \theta_k,$$

найти случайную величину X, обладающую наибольшей энтропией.

Решение. Поставленная задача состоит в определении плотности вероятности $W(x)$, доставляющей наибольшее значение функционалу энтропии

$$H(X) = -\int_a^b W(x)\ln W(x)\,dx,$$

при указанных k условиях и условии нормировки

$$(k+1) \qquad \int_a^b W(x)\,dx = 1.$$

Составляем вспомогательный функционал

$$F(X) = H(X) + \sum_{i=1}^{k+1}\lambda_i\theta_i,$$

где λ_j – произвольные постоянные коэффициенты (неопределенные множители Лагранжа). Так как сумма $\sum_{i=1}^{k+1}\lambda_i\theta_i$ есть по условию задачи величина постоянная, не зависящая от конкретного вида функции $W(x)$, то плотность $W(x)$, обеспечивающая максимум функционалу $F(X)$, обеспечит также максимальное значение функционалу $H(X)$.

Вспомогательный функционал с учетом выражений (1) - (k+1) принимает вид

$$F(X)=\int_a^b W(x)\left[-\ln W(x)+\sum_{i=1}^{k+1}\lambda_i\psi_i(x)\right]dx,$$

где $\psi_{k+1}=1$.

Воспользуемся очевидным соотношением

$$\lambda_i\psi_i(x)=\ln\left(\exp\left[\lambda_i\psi_i(x)\right]\right)$$

и получим

$$F(X)=\int_a^b W(x)\cdot\ln\frac{\exp\left[\sum_{i=1}^{k+1}\lambda_i\psi_i(x)\right]}{W(x)}dx=$$

$$\text{M}\left\{\ln\frac{\exp\left[\sum_{i=1}^{k+1}\lambda_i\psi_i(X)\right]}{W(X)}\right\}.$$

Далее пользуемся неравенством Иенсена (см. пример 1.4) и получаем

$$F(x)\le\ln\text{M}\left\{\frac{\exp\left[\sum_{i=1}^{k+1}\lambda_i\psi_i(X)\right]}{W(X)}\right\}=$$

$$\ln\int_a^b W(x)\frac{\exp\left[\sum_{i=1}^{k+1}\lambda_i\psi_i(x)\right]}{W(x)}dx=\ln\int_a^b\exp\left[\sum_{i=1}^{k+1}\lambda_i\psi_i(x)\right]dx.$$

Правая часть этого неравенства есть постоянная, не зависящая от вида распределения $W(x)$, поэтому наибольшее значение $F(x)$ будет достигнуто для функции $W(x)$, обращающей неравенство в равенство. Как следует из примера 1.4, неравенство Иенсена обращается в равенство только в том случае, когда

$$\frac{\exp\left[\sum_{i=1}^{k+1}\lambda_i\psi_i(x)\right]}{W(x)}=C,$$

где C – некоторая постоянная.

Полагая C=1, получим

$$W(x)=\exp\left[\sum_{i=1}^{k+1}\lambda_i\psi_i(x)\right].$$

Подставляя эту функцию в выражения (1) – (k+1) и решая систему k+1 уравнений, можно найти конкретные значения коэффициентов $\lambda_1,\lambda_2,\ldots,\lambda_{k+1}$.

Пример 3.2.3. Среди всех случайных величин, для которых задано математическое ожидание m_x, а $W(x)$=0 при x<0, найти случайную величину X, обладающую наибольшей энтропией.

Решение. Задача состоит в нахождении плотности вероятности $W(x)$, доставляющей максимум функционалу энтропии

$$H(X)=-\int_{-\infty}^{\infty}W(x)\ln W(x)\,dx$$

при дополнительном условии

$$1)\qquad \int_0^{\infty}xW(x)=m_x \tag{3.2.12}$$

и условии нормировки

$$2)\qquad \int_0^{\infty}W(x)\,dx=1.$$

Имеем решение задачи в общем виде (формула (3.2.11))

$$W(x)=\exp\left[\lambda_1\varphi(x)+\lambda_2\right].$$

Сравнивая выражения (3.2.10) и (3.2:12), убеждаемся, что $\varphi(x)=x$, и решение для данного случая примет вид

$$W(x)=\exp(\lambda_1 x+\lambda_2).$$

Значения коэффициентов λ_1 и λ_2 ищем, подставляя $W(x)$ в условия (3.2.12). Получим систему уравнений с двумя неизвестными:

$$1)\quad \int_0^\infty xe^{\lambda_1 x+\lambda_2}dx=m_x, \qquad 2)\quad \int_0^\infty e^{\lambda_1 x+\lambda_2}dx=1.$$

Решаем эту систему уравнений. Из уравнения 2) находим

$$e^{\lambda_2}=\frac{1}{\int_0^\infty e^{\lambda_1 x}dx}.$$

Воспользуемся табличным интегралом вида $\int_0^\infty e^{cx}dx=-\frac{1}{c}$, где $c<0$. Тогда $\int_0^\infty e^{\lambda_1 x}dx=-\frac{1}{\lambda_1}$ и $e^{\lambda_2}=-\lambda_1$.

Подставляя $e^{\lambda_2}=-\lambda_1$ в уравнение 1) системы

$$\int_0^\infty xe^{\lambda_1 x}(-\lambda_1)dx=m_x$$

и воспользовавшись табличным интегралом вида

$$\int_0^\infty xe^{cx}dx=-\frac{1}{c^2},$$

получим

$$(-\lambda_1)\int_0^\infty xe^{\lambda_1 x}dx=(-\lambda_1)\cdot\frac{1}{(-\lambda_1)^2}=m_x.$$

Отсюда $\lambda_1 = -\frac{1}{m_x}$ и $e^{\lambda_2} = -\lambda_1 = \frac{1}{m_x}$.

Искомый закон распределения имеет вид

$$W(x) = e^{\lambda_1 x} e^{\lambda_2} = \frac{1}{m_x} \exp\left(-\frac{x}{m_x}\right), \quad x \geq 0.$$

Это – одностороннее экспоненциальное распределение

ЗАДАЧИ

3.2.1. По каналу связи передаётся частотно-модулированный сигнал, частота которого f изменяется

а) с равной вероятностью в пределах от f_1= 700 МГц до f_2 = 900 МГц,

б) по экспоненциальному закону

$$W(x) = \begin{cases} \frac{1}{m} e^{-\frac{f-900}{m}}, f > 900 МГц \\ 0, f < 900 МГц \end{cases}$$

со средним значением m=200 МГц. Вычислить относительную энтропию для обоих случаев.

3.2.2. Два стрелка делают по одному выстрелу в мишень. Координаты X и Y точек попадания для первого стрелка распределены равномерно в прямоугольнике: $0 \leq x \leq 4$; $0 \leq y \leq 8$, а для второго — имеют совместное нормальное распределение с параметрами: m_x = m_y = 0; σ_x = 1; σ_y = 8; r_{xy} = 0,7. Для которого стрелка исход стрельбы является, более определенным и насколько?

3.2.3. Сигнал X с выхода детектора радиолокационного приемника должен быть подан на цифровой измеритель и с этой целью подвергается равномерному квантованию с шагом Δ x. Сигнал распределен по рэлеевскому закону:

$$W(x) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right), x > 0$$

Найти приближенное выражение для вычисления величины взаимной информации между этими сигналами при Δ x<<a.

3.2.4. Пусть $W_1(x)$ и $W_2(x)$ есть плотности вероятностей. непрерывных случайных величин. Доказать, что

$$\int_{-\infty}^{\infty} W_1(x)\log W_1(x)dx \geq \int_{-\infty}^{\infty} W_2(x)\log W_2(x)dx$$

Найти необходимое и достаточное условие, при котором. Неравенство обращается в равенство.

3.2.5. Осуществляется прием на две антенны, разнесенные в пространстве так, что сигналы в ветвях статистически независимы, имеют нормальное распределение вероятностей с ну левыми средними, дисперсиями σ_1^2 = 16 В^2 и σ_2^2 = 25 В^2 со ответственно.

Вычислить энтропию колебания на выходе суммирующего устройства.

3.2.6. Среди всех законов распределения непрерывной случайной величины X, имеющей ограниченные пределы $-\dfrac{a}{2} < x < \dfrac{a}{2}$, найти закон распределения с максимальной энтропией.

3.2.7. Вычислить значение энтропии для случайной величи-ны X, имеющей нормальный закон распределения с дисперсией σ_x^2

3.2.8. Найти максимально возможную энтропию сигналов с ограниченной пиковой мощностью, т. е. при условии $|x_{\max}| = +\sqrt{P_{пик}}$.

3.2 9. Среди всех законов распределения непрерывной случайной величины X, имеющей заданную дисперсию D_x = $=\sigma_x^2$ = const, найти закон распределения с максимальной энтропией.

3.2.10. Вычислить значение энтропия для случайной величии-ны, найденной в задаче **3.2.9.,** $\sigma_x^2 = 14B^2$

3.2.11*. Среди множества законов распределения системы n случайных величин, имеющих заданную корреляционную матрицу R, найти закон, обеспечивающий максимальное значение энтропии этой системы.

3.2.12. Определить, насколько энтропия сигнала с нормальным распределением значений отсчетов превышает энтропию сигнала с равномерным распределением отсчётов, при одинаковой величине средней мощности.

3.2.13. Во сколько раз средняя мощность сигнала с равномерным, распределением значений отсчетов должна быть больше мощности сигнала с нормальным распределением отсчетов при условии, что оба сигнала имеют одинаковые энтропии: $H_р(X) = H_н(X)$

3.2.14. Доказать, что из всех положительных случайных величин, имеющих заданную энтропию, наименьшим средним значением обладает экспоненциально распределенная слу-чайная величина.

3.2.15. Из всех случайных величин, имеющих заданную относительную энтропию H(X)=1,5 бит, найти такую случайную величину, которая обладает минимальной дисперсией σ_x

3.2.16 *. На вход безынерционного квадратичного детектора с характеристикой Y = βX^2 поступает сигнал, представляющий собой непрерывную случайную величину с нормальным законом распределения с нулевым средним. Вычислить энтропию случайной величины Y на выходе детектора.

3.2.17. Две системы (X_1,…, X_n) и (Y_1,…, Y_n) случайных величин связаны линейными соотношениями:

$$Y_k = \sum_{j=1}^{n} a_{kj} X_j ; k = (1,2,..., n)$$

Определить разность энтропий

$$H(Y_1,..., Y_n) - H(X_1,..., X_n)$$

а) в общем случае,

б) при n=3 и матрице преобразований

$$\left\| a_{kj} \right\| = \begin{Vmatrix} 3 & 2 & -1 \\ 1 & 4 & 2 \\ 0 & -3 & 5 \end{Vmatrix}.$$

3.2.18. Определить среднюю условную энтропию H(X/Y) случайной величины X и H(X/Y) случайной величины Y для системы XY, подчиняющейся нормальному закону распределения.

3.2.19. Доказать, что средняя условная энтропия случайной величины X H(X/Y) не превосходит безусловной, т. е. $\mathrm{H(X/Y)} \le \mathrm{H(X)}$.

Какому случаю соответствует знак равенства?

3.3 Информационные характеристики непрерывных случайных функций

Имеются два различных способа определения информационных характеристик непрерывных стационарных случайных функций.

Первый способ предусматривает квантование во времени исходной непрерывной случайной функции $X(t)$ и переход к последовательности непрерывных случайных величин-отсчётов $\ldots,\ \mathrm{X}^{(-1)}, \mathrm{X}^{(0)}, \mathrm{X}^{(1)}, \ldots$, взятых через интервалы $\Delta t = 1\big/\, 2\Delta F$, где ΔF - ширина полосы частот на положительной полуоси, занимаемых спектром плотности мощности этой функции. Для полученной непрерывной последовательности вводятся характеристики, аналогичные тем, которые были введены для дискретных последовательностей (формулы (2.4.1) – (2.4.3), (2.4.6) – (2.4.10)), но используется понятие относительной энтропии.

Определённая таким образом энтропия непрерывной случайной величины функции $X(t)$ равна средней относительной энтропии, приходящейся на один отсчёт. Эта величина также удовлетворяет неравенствам (2.4.3) и (2.4.4), причём в качестве $H_{1\max}$ используется максимальное значение относительной энтропии непрерывной случайной величины при заданных ограничениях. Например, если задана средняя мощность σ^2 случайной функции $X(t)$, оба неравенства совместно дают

$$H\left[X(t)\right] \le \frac{1}{2}\log(2\pi e\sigma^2), \tag{3.3.1}$$

причём максимальной относительной энтропией на один отсчёт обладает нормальная случайная функция с нулевым математическим ожиданием, дисперсией (средней мощностью) σ^2 и равномерным в полосе ΔF энергетическим спектром (белый шум).

«Энтропийной мощностью» случайной функции $X(t)$, имеющей ширину спектра ΔF и энтропию H нат/отсчёт, называется средняя мощность белого шума с такой же шириной спектра и тем же значением энтропии на отсчёт

$$\overline{N} = \frac{1}{2\pi}\exp(2H - 1). \tag{3.3.2}$$

Относительная энтропия на один отсчёт нормальной случайной функции, имеющей спектр плотности мощности $G(f)$, вычисляется по формуле

$$H\ X(t)\ = \frac{1}{2\Delta F}\int\limits_{\Delta F}\ln\ 2\pi e\cdot\Delta F\cdot G(f)\ df, \tag{3.3.3}$$

а энтропийная мощность равна

$$\overline{N} = \Delta F\exp\left[-\frac{1}{\Delta F}\int\limits_{\Delta F}\ln G(f)df\right]. \tag{3.3.4}$$

Второй способ введения энтропийных характеристик случайной функции опирается на понятие точности воспроизведения реализации этой функции. Пусть $x(t)$ есть реализация непрерывной случайной функции $X(t)$, которую необходимо передать, а $z(t)$ – реализация другой случайной функции $Z(t)$, которая в действительности принимается. Обозначим как $\rho(X,Z)$ некоторую заранее заданную количественную меру различия этих двух случайных функций. Тогда ε-энтропией $H_\varepsilon^\rho\left[X(t)\right]$ случайной функции $X(t)$ называется минимальное среднее количество взаимной информации в единицу времени между $X(t)$ и $Z(t)$, необходимое для того, чтобы $\rho(X,Z)\le\varepsilon$.

Часто в качестве меры отличия используют среднеквадратическую ошибку воспроизведения стационарной случайной функции $X(t)$ при помощи стационарной и стационарно связанной с ней случайной функции $Z(t)$

$$\rho_0(X,Z)=\sqrt{\mathrm{M}\left[(X-Z)^2\right]}. \tag{3.3.5}$$

Если $Z(t)$ – это цифровой сигнал, то ε-энтропия численно равна минимальному среднему количеству двоичных символов в единицу времени, необходимых для восстановления функции $X(t)$ со среднеквадратической ошибкой, не превышающей ε.

Для нормальной стационарной случайной функции $X(t)$, имеющей спектр плотности мощности $G(f)$, ε-энтропия при среднеквадратическом критерии точности (3.3.5) вычисляется по формуле

$$H_\varepsilon^0\ X(t)\ =\int_{\Delta f}\log\frac{G(f)}{\lambda}df\ , \tag{3.3.6}$$

где Δf - полоса частот, в которой $G(f)\geq\lambda$. Коэффициент λ выбирается таким образом, чтобы площадь фигуры, ограниченной снизу осью f, а сверху – прямой $G=\lambda$ (в области Δf) либо кривой $G(f)$ (вне области Δf), была равна ε^2. Эта фигура заштрихована на рис. 3.3.1.

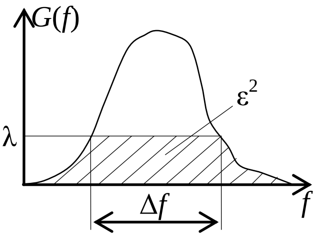


Рис. 3.1

Понятие ε - энтропии может быть применено и к последовательности дискретных случайных величин.

При вычислении ε-энтропии случайной величины X, когда расстояние между X и Z задано в виде математического ожидания некоторой функции

$$\rho(X,Z)=\mathrm{M}\ \ \varphi(V) \tag{3.3.7}$$

их разности $V=X-Z$, справедливо соотношение

$$H(X/Z)=H(V/Z)\leq H(V). \tag{3.3.8}$$

Оно показывает, что средняя условная энтропия ошибки при заданном ограничении $\rho\leq\varepsilon$ достигает максимального значения, когда X и Z независимы. Это требование, однако, не является достаточным.

РЕШЕНИЕ ТИПОВЫХ ПРИМЕРОВ

Пример 3.3.1. Нормальный стационарный узкополосный случайный процесс, имеющий корреляционную функцию

$$B(\tau) = \frac{1}{1 + a^2\tau^2}\cos 2\pi f_0 \tau\ ,$$

пропускают через идеальный фильтр с полосой, заключённой в интервале частот от $f_1 = f_0 - F$ до $f_2 = f_0 + F$. Найти относительную энтропию в единицу времени для процесса на выходе фильтра.

Решение. Односторонний спектр плотности мощности исходного процесса в соответствии с теоремой Винера-Хинчина равен преобразованию Фурье от корреляционной функции

$$G(f) = 2\int_{-\infty}^{\infty} B(\tau)\cos 2\pi f\tau d\tau = 2\int_{-\infty}^{\infty} \frac{1}{1 + a^2\tau^2}\cos 2\pi (f - f_0)\tau d\tau.$$

Этот интеграл – табличный:

$$G(f) = \frac{\pi}{a}\left\{\exp\left[-\left|\frac{2\pi}{a}(f - f_0)\right|\right] + \exp\left[-\left|\frac{2\pi}{a}(f + f_0)\right|\right]\right\}, \quad f \ge 0.$$

Поскольку процесс узкополосен, то справедливо соотношение $f_0 >> a$, поэтому в полученном выражении можно пренебречь вторым слагаемым и получим

$$G(f) \approx \frac{\pi}{a}\left\{\exp\left[-\left|\frac{2\pi}{a}(f - f_0)\right|\right]\right., \quad f \ge 0.$$

Спектр плотности мощности процесса на выходе идеального полосового фильтра равен

$$G_{вых}(f) = \frac{\pi}{a}\left\{\exp\left[-\left|\frac{2\pi}{a}(f - f_0)\right|\right]\right., \quad f_1 \le f \le f_2.$$

Величину относительной энтропии на один отсчёт вычисляем по формуле (3.3.3)

$$H\ \ X(t)\ =\frac{1}{2(f_2-f_1)}\int\limits_{f_1}^{f_2}\ln\left\{\frac{\pi}{a}\exp\left[-\left|\frac{2\pi}{a}(f-f_0)\right|\right]\right\}df+$$

$$+\frac{1}{2}\ln\ \ 2\pi e(f_2-f_1)\ =\frac{1}{2}\ln\left[\frac{2\pi^2 e}{a}(f_2-f_1)\right]+\int\limits_{f_1}^{f_2}\left[-\left|\frac{2\pi}{a}(f-f_0)\right|\right]df.$$

Далее учитываем, что $f_2-f_0=f_0-f_1=F$ и, осуществив интегрирование, получим

$$H[X(t)]=\frac{1}{2}\ln\frac{4\pi^2 eF}{a}-\frac{2\pi F^2}{a}\quad\frac{\text{нат}}{\text{отсчёт}}.$$

В одну секунду берется $n=2\Delta F=4F$ отсчётов, поэтому относительная энтропия в единицу времени равна

$$H[X(t)]=2F\ln\frac{4\pi^2 eF}{a}-\frac{8\pi F^2}{a}\quad\frac{\text{нат}}{\text{с}}.$$

Пример 3.3.2. Уровень воды в водоёме $U(t)$ меняется со временем так, что он полностью определяется отсчётами, взятыми через полгода. Отдельные отсчёты независимы и имеют одностороннее экспоненциальное распределение с математическим ожиданием, равным 4 м. При учёте запаса воды допустимы только положительные ошибки $v=u\text{-}z$, причём среднее значение ошибки не должно превышать 0,25 м.

Найти ε-энтропию функции $U(t)$.

Решение. Задача заключается в вычислении минимального среднего количества информации, которое необходимо передавать в год для воспроизведения функции $U(t)$ при помощи другой функции с заданной точностью. За год уровень воды измеряется два раза, и эти отсчёты независимы, поэтому можно вычислить ε -энтропию одного отсчёта и, пользуясь свойством аддитивности ε -энтропии, удвоить полученный результат.

Средняя взаимная информация между U и Z равна

$$I(U;Z)=H(U)-H(U/Z).$$

Относительная энтропия экспоненциально распределённой положительной случайной величины U равна (см. пример (3.2.1))

$$H(U)=\log_2(em_u)=\log_2(4e)\ \text{бит}.$$

В условии задачи расстояние между X и Z задано в виде математического ожидания их разности, поэтому средняя условная энтропия $H(U/Z)$ достигает максимального значения, когда V и Z независимы (см. неравенство (3.3.8)). Далее необходимо в классе распределений $W(v)$ положительной случайной величины V, удовлетворяющих заданному в условии задачи ограничению

$$1) \qquad \mathrm{M}[V] = \int_0^\infty vW(v)dv \le 0,25$$

и условию нормировки

$$2) \qquad \int_0^\infty W(v)dv = 1,$$

найти такое распределение, которое обладает наибольшей энтропией.

Как следует из примера 3.2.3, случайная величина V при этих ограничениях также должна иметь одностороннее экспоненциальное распределение

$$W(v) = \frac{1}{0,25}\exp(-\frac{v}{0,25}), \quad v > 0.$$

Тогда $H(U/Z)_{\max} = H(V)_{\max} = \log_2(0{,}25e)$ бит,

и ε -энтропия на один отсчёт равна

$$H_\varepsilon^\rho(U) = I(U;Z)_{\min} = \log_2(4e) - \log_2(0,25e) = 4 \text{ бит/отсчёт}.$$

Для ε -энтропии случайной функции $U(t)$ окончательно получим $H_\varepsilon^\rho[U(t)] = 8 \text{ бит/год}$.

ЗАДАЧИ

3.3.1. Радиосигнал подаётся на усилитель, имеющий коэффициент передачи K(f) = 1000. Насколько изменится энтропия на 1 отсчёт выходного сигнала по сравнению с входным.

3.3.2. Сигнал на выходе микрофона стационарен, имеет среднюю мощность 10^{-6} Вт, а его энергетический спектр заключён в полосе частот 100 – 5000 Гц. Найти наибольшее возможное значение относительной энтропии такого сигнала:

а) на один отсчёт,

б) в 1 мин.

3.3.3. Вычислить «энтропийную мощность» сигнала U(t) из примера 3.3.2., полагая, что этот сигнал имеет ограниченный спектр.

3.3.4. Радиосигнал, спектр которого заключён в полосе частот 100 – 300 МГц, подаётся на усилитель промежуточной частоты, имеющий амплитудно-частотную характеристику колокольной формы

$$K(f) = 10^5 \exp[-\frac{(f - f_0)^2}{(\Delta f)^2}]$$

с центральной частотой $f_0 = 200$ МГц и полосой пропуска-ния $\Delta f = 20$ МГц.

Вычислить, насколько изменится энтропия выходного сигнала по сравнению с входным.

3.3.5. Отсчёты сигнала U(t) из примера 3.3.2 подвергаются равномерному квантованию с шагом 2 м. Найти величину взаимной информации в единицу времени между исходным и квантованным сигналами.

3.3.6*. Вычислить ε - энтропию нормальной случайной ве-ли-чины X, имеющей параметры $m = 5$ и $\sigma = 8$, если допустима среднеквадратическая ошибка представления этой величины равная 1. Найти условную плотность вероятности W(z/x) для воспроизводящей величины Z.

3.3.7. Вычислить ε - энтропию равномерно распределённой в интервале 0 – 2 случайной величины, если при воспроизведе-

нии допустима ошибка, не превышающая по модулю 0,01. Объяснить полученный результат и указать способ воспроизведения с требуемой точностью.

3.3.8. Получить формулу относительной энтропии на один отсчёт для системы n нормальных случайных величин (отсчётов), имеющих корелляционную матрицу R.

3.3.9. Вероятности значений 0 и 1 двоичного символа X равны. Вычислить минимальное среднее количество информации, необходимое для того, чтобы полная вероятность ошибки при воспроизведении значений этого символа не пре-вышала 0,11.

3.3.10. При передаче значений символа X из задачи 3.3.9 на выходе линии связи имеется информация, равная в среднем 0,55 бит/символ. Укажите можно ли при наличии такого количества информации воспроизводить значения символа X так, чтобы полная вероятность ошибки не превышала 0,02?

3.3.11. Доказать неравенство 3.3.8.

4 КОДИРОВАНИЕ. ПРОПУСКНАЯ СПОСОБНОСТЬ КАНАЛА.

4.1 Основные определения. Пропускная способность канала

Модель системы передачи информации. Системой передачи информации (СПИ) называется совокупность технических средств, используемых для передачи информации в пространстве (связь, телекоммуникации). В сущности, те же закономерности характерны и для системы передачи информации во времени (хранение информации). Модель системы передачи информации представлена на рис. 4.1.

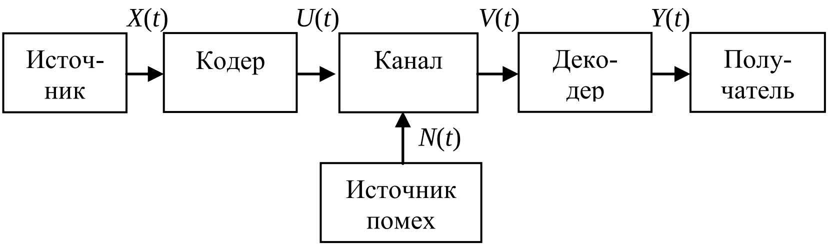


Рис. 4.1

Источником называется устройство или человек, генерирующие сообщение, подлежащее передаче, а *получателем* – устройство или человек, получающие информацию. *Каналом* называется совокупность физических объектов, используемых для переноса сигнала (кабель, проводная линия, лента магнитофона, среда, в которой распространяются волны, и т.п.). *Кодером* называется устройство, преобразующее сигнал на выходе источника в сигнал, пригодный для передачи по данному каналу. *Декодером* называется устройство, преобразующее сигнал на выходе канала к виду, пригодному для использования потребителем. Воздействие помех (шумов) на сигнал в процессе его распространения в канале отражено на схеме введением *источника помех*.

Источник информации. Источник информации называется дискретным или непрерывным в зависимости от вида создаваемых им сообщений. Источник называется стационарным, если сообщение на его выходе есть стационарная случайная функция (последовательность). Дискретный источник называется источником без памяти, если символы в последовательности на его выходе независимы.

Канал. Канал называется дискретным, непрерывным, дискретно-непрерывным или непрерывно-дискретным в зависимости от вида сигналов на его входе и выходе. Если входной $A(t)$ и выходной $B(t)$ сигналы связаны взаимно-однозначно, то такой канал называется каналом без шума. В канале с шумом возможны случайные ошибки при преобразовании входного сигнала в выходной.

Дискретный канал с шумом называется каналом без памяти, если ошибки в отдельных символах выходной последовательности статистически независимы. Канал называется стационарным (постоянным), если условные вероятности перехода от $A(t)$ к $B(t)$ не зависят от начала отсчёта.

Кодер и декодер. Кодирующее устройство выполняет следующие операции:

а) согласование источника с каналом (перевод реальных сообщений в электрические сигналы, модуляция непрерывных сигналов, квантование непрерывных сообщений, представление s-ичного дискретного сообщения в m-ичном коде и т.п.);

б) экономное представление информации с минимальной избыточностью либо, наоборот, разумное введение избыточности в сигнал, передаваемый по каналу, с целью повышения его помехоустойчивости.

Функции декодера в значительной степени обратны функциям кодера. Кроме того, введение при кодировании избыточности в сигнал $A(t)$ часто представляет возможность обнаруживать и исправлять ошибки, возникающие в канале из-за влияния помех. Эта операция также выполняется декодирующим устройством.

Скорость передачи информации. Одной из важных характеристик источника является скорость создания информации. Для стационарного дискретного источника она равна энтропии последовательности символов на его выходе, вычисленной по формуле (2.4.2). Скорость создания информации непрерывным источником, вообще говоря, бесконечна (см. разд. 3.2). Поэтому в качестве скорости создания информации стационарным непрерывным источником принимают значение ε - энтропии непрерывной функции на его выходе, вычисленной при заданной точности воспроизведения этой функции (см. разд. 3.3).

Важнейшей характеристикой СПИ в целом является скорость передачи информации. Скоростью передачи информации называется средняя величина взаимной информации (в единицу времени или на отсчёт) между сигналом $X(t)$ на выходе источника и сигналом $Y(t)$, поступающим к потребителю. Эта величина вычисляется по формуле (2.4.8), но для непрерывных источников в формуле используются значения относительной энтропии (условной и безусловной).

Условная энтропия в правой части выражения (2.4.8) иногда называется ненадёжностью передачи информации от дискретного источника. Она численно равна разности между скоростью создания информации и скоростью её передачи, поэтому ненадёжность, по существу, есть скорость потерь информации в канале (и, возможно, в кодере и декодере). Ненадёжность равна нулю только в случае безошибочной передачи информации.

Пропускная способность канала. Скорость передачи информации зависит в значительной степени от скорости её создания, способов кодирования и декодирования. Наибольшая возможная в данном канале скорость передачи информации называется его пропускной способностью C. Пропускная способность канала, по определению, есть скорость передачи информации при использовании «наилучших» для данного канала источника, кодера и декодера, поэтому она характеризует только канал.

Пропускная способность дискретного (цифрового) канала без помех

$$C = \log m \quad \text{бит/символ}, \qquad (4.1.1)$$

где m – основание кода сигнала, используемого в канале (см. 4.2.2). Скорость передачи информации в дискретном канале без шумов равна его пропускной способности, когда символы в канале независимы, а все m букв алфавита равновероятны (используются одинаково часто).

Пропускная способность дискретного канала с шумом без памяти

$$C = \max_{p(A)} \sum_{j=1}^{m} \sum_{k=1}^{m} p(a_j) p(b_k / a_j) \log \frac{p(b_k / a_j)}{p(a_j)} \quad \text{бит/симв.} \qquad (4.1.2)$$

При заданных вероятностях перехода $p(b_k / a_j)$ задача вычисления пропускной способности сводится к отысканию «наилучшего» распределения вероятностей $p(a_1),..., p(a_m)$ на входе канала. Скорость передачи информации в таком канале равна его пропускной способности, когда символы в канале независимы, а распределение вероятностей m входных букв равно найденному.

Пропускная способность непрерывного канала с ограниченной полосой частот Δf Гц, в котором действует аддитивный белый гауссовский шум (1.9) со спектральной плотностью N_0 Вт/Гц, вычисляется по формуле Шеннона-Таллера

$$C = \Delta f \log(1 + \frac{P_c}{P_ш}) \quad \text{бит/с}, \qquad (4.1.3)$$

где $P_c = \mathrm{M}[B^2(t)]$ – средняя мощность стационарного полезного сигнала $B(t)$ на выходе этого канала,

$P_ш = N_0 \Delta f$ – средняя мощность шума, попадающего в полосу частот Δf , на выходе этого канала.

Скорость передачи информации в непрерывном канале с постоянными параметрами равна его пропускной способности, когда входной сигнал $A(t)$ также является стационарным гауссовским случайным процессом с нулевым математическим ожиданием и спектром плотности мощности, равномерным в полосе частот Δf .

РЕШЕНИЕ ТИПОВЫХ ПРИМЕРОВ

$$\mathbf{P}=\begin{Vmatrix}1 & 0 & 0\\ 0 & 1-p_e & p_e\\ 0 & p_e & 1-p_e\end{Vmatrix}$$

Пример 4.1.1. Несимметричный троичный канал без памяти задан матрицей **P** условных вероятностей переходов.

Стационарный источник без памяти генерирует символы с частотой $F=1/\tau$. Априорные вероятности символов на входе равны $p(a_1)=p,\quad p(a_2)=p(a_3)=q$.

Нарисовать граф канала. Вычислить скорость создания информации, скорость передачи информации, ненадёжность.

Решение. Канал можно изобразить в виде графа (рис.4.2), используя матрицу P в общем виде

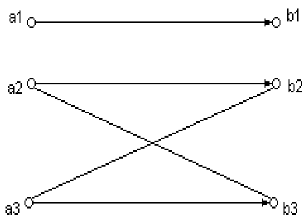


Рис. 4.2

$$\mathbf{P}=\begin{Vmatrix}p(b_1/a_1) & p(b_2/a_1) & p(b_3/a_1)\\ p(b_1/a_2) & p(b_2/a_2) & p(b_3/a_2)\\ p(b_1/a_3) & p(b_2/a_3) & p(b_3/a_3)\end{Vmatrix}.$$

Источник и канал стационарны и не имеют памяти, поэтому скорость создания информации равна $H(A)$, скорость передачи информации равна $I(A;B)$ и ненадёжность равна $H(A/B)$.

Для нахождения условной энтропии $H(A/B)$ надо знать вероятности $p(a_j/b_k)$. Их можно найти по формулам (1.3) и (1.4). Так как символ a_1 не искажается, то $p(b_1)=p(a_1)=p$. Поскольку символы a_2 и a_3 искажаются симметрично, а $p(a_2)=p(a_3)$, то вероятности $p(b_2)=p(b_3)$, и тогда обратные вероятности

$$p(a_2/b_2)=p(a_3/b_3)=1-p_e,\quad p(a_3/b_2)=p(a_2/b_3)=p_e.$$

Энтропия сигнала на входе

$$H(A) = -\sum_{i=1}^{3} p(a_i)\log p(a_i) = -p\log p - 2q\log q \quad \frac{\text{бит}}{\text{симв}}.$$

Скорость передачи равна средней взаимной информации

$$I(A;B) = -\sum_{j=1}^{3} p(b_j)\log p(b_j) +$$

$$p(a_1)[p(b_1/a_1)\log p(b_1/a_1) + p(b_2/a_1)\log p(b_2/a_1) +$$
$$p(b_3/a_1)\log p(b_3/a_1)] + p(a_2)[p(b_1/a_2)\log p(b_1/a_2) +$$
$$p(b_2/a_2)\log p(b_2/a_2) + p(b_3/a_2)\log p(b_3/a_2)] +$$
$$p(a_3)[p(b_1/a_3)\log p(b_1/a_3) + p(b_2/a_3)\log p(b_2/a_3) +$$
$$p(b_3/a_3)\log p(b_3/a_3)]\} =$$
$$-p\log p - 2q\log q + 0 + q[0 + (1-p_e)\log(1-p_e) +$$
$$p_e\log p_e] + q[0 + (1-p_e)\log(1-p_e) + p_e\log p_e] =$$
$$\{-p\log p - 2q\log q + 2q[(1-p_e)\log(1-p_e) + p_e\log p_e]\}.$$

Обозначив $\beta = -(1-p_e)\log(1-p_e) - p_e\log p_e$,

получим

$$I(A;B) = -p\log p - 2q\log q - 2q\beta \quad \text{бит/символ.}$$

Ненадёжность равна разности между скоростью создания и скоростью передачи информации

$$H(A|B) = H(A) - I(A;B) = 2q\beta \quad \text{бит/символ.}$$

Пример 4.1.2. Для канала из примера 4.1.1 вычислить пропускную способность в следующих случаях:

а) при произвольном шуме,

б) при отсутствии шума,

в) при очень большом уровне шума.

Вероятности входных символов a_2 и a_3 должны быть равными.

Решение. а) Пропускная способность есть верхняя грань выражения (4.1.2) при дополнительном условии нормировки

$$\sum_{j=1}^{3} p(a_j) = p + 2q = 1.$$

Для определения верхней грани $I(A;B)$ применим метод неопределённых множителей Лагранжа.

Образуем вспомогательную функцию

$$F = I(A;B) + \lambda(p+2q) =$$
$$[-p\log p - 2q\log q - 2q\beta + \lambda(p+2q)].$$

Составляем систему уравнений:

$$\frac{dF}{dp} = 0 \text{ и } \frac{dF}{dq} = 0.$$

Система уравнений примет вид

$$1)\ \frac{dF}{dp} = -\log_2 p + \lambda - \frac{1}{\ln 2} = 0,$$

$$2)\ \frac{dF}{dq} = -2\log_2 q - 2\beta + 2\lambda - \frac{2}{\ln 2} = 0.$$

Из системы уравнений находим p и q:

$$\log p = \lambda - \frac{1}{\ln 2} \quad \text{и} \quad p = 2^{\lambda - \frac{1}{\ln 2}},$$

$$\log q = -\beta + \lambda - \frac{1}{\ln 2} \quad \text{и} \quad q = 2^{\lambda - \frac{1}{\ln 2}} \cdot 2^{-\beta} = \frac{p}{\alpha},$$

где обозначено $\alpha = 2^{\beta}$.

Множитель λ найдём из условия нормировки, которому должны удовлетворять искомые вероятности p и q:

$$p + 2q = 2^{\lambda - \frac{1}{\ln 2}} + 2 \cdot 2^{\lambda - \frac{1}{\ln 2}} \cdot \frac{1}{\alpha} = 1,$$

$$2^{\lambda - \frac{1}{\ln 2}}\left[1 + \frac{2}{\alpha}\right] = 1,$$

откуда

$$p = 2^{\lambda - \frac{1}{\ln 2}} = \frac{1}{1 + 2/\alpha} = \frac{\alpha}{2+\alpha} \quad \text{и} \quad q = \frac{p}{\alpha} = \frac{1}{2+\alpha}.$$

Тогда пропускная способность канала

$$C = \frac{1}{\tau}[-p\log p - 2q\log q - 2q\beta] =$$

$$\frac{1}{\tau}\left[-\frac{\alpha}{2+\alpha}\log\frac{\alpha}{2+\alpha}-2\frac{1}{2+\alpha}\log\frac{1}{2+\alpha}-2\beta\frac{1}{2+\alpha}\right].$$

Произведя простейшие преобразования

$$C=\frac{1}{\tau}\left[-\frac{\alpha}{2+\alpha}\log\alpha+\frac{\alpha}{2+\alpha}\log(2+\alpha)+\right.$$

$$\left.\frac{\alpha}{2+\alpha}\log(2+\alpha)-\frac{2\beta}{2+\alpha}\right]=\frac{1}{\tau}\ \log(2+\alpha)-$$

$$\left.\frac{\alpha}{2+\alpha}\log\alpha-\frac{2\beta}{2+\alpha}\right]$$

и осуществив потенцирование, получим окончательно

$$C=\frac{1}{\tau}\log_2\left[(2+\alpha)\cdot 2^{-\frac{2\beta}{2+\alpha}}\cdot\alpha^{-\frac{\alpha}{2+\alpha}}\right].$$

б) Рассмотрим частный случай, когда помехи в канале отсутствуют, следовательно, $p_e=0$. При этом

$$\beta=-p_e\log p_e-(1-p_e)\log(1-p_e)=0,\quad \alpha=2^{\beta}=1.$$

Априорные вероятности

$$p=\frac{\alpha}{2+\alpha}=\frac{1}{3},\qquad q=\frac{1}{2+\alpha}=\frac{1}{3}.$$

Пропускная способность максимальна и равна

$$C_{\max}=\frac{1}{\tau}\log_2 3,$$

что совпадает с формулой (4.1.1).

в) Помехи значительны, так что $p_e=0{,}5$. При этом

$$\beta=-p_e\log_2 p_e-(1-p_e)\log_2(1-p_e)=1,$$

$$\alpha=2,\quad p=\frac{\alpha}{2+\alpha}=\frac{1}{2},\quad q=\frac{1}{2+\alpha}=\frac{1}{4},$$

и пропускная способность равна

$$C=\frac{1}{\tau}\left[\log_2 4-\frac{1}{2}\log_2 2-\frac{1}{2}\right]=\frac{1}{\tau}\log_2 2\quad\frac{\text{бит}}{\text{с}}.$$

Фактически канал стал двоичным, так как символы a_2 и a_3 неразличимы, и результат соответствует пропускной способности двоичного канала без помех.

Пример 4.1.3. Изображение с энтропией $5 \cdot 10^7$ бит передаётся по линии связи. Определить чувствительность приёмного устройства, если на входе действует белый гауссовский шум со спектральной плотностью $N_0 = 2 \cdot 10^{-17}$ Вт/кГц. Полоса приёмного устройства 100 кГц, время передачи T=1 час.

Решение. Поскольку необходимо определить чувствительность, т.е. минимально возможное значение мощности сигнала на входе приёмного устройства, то радиолиния должна быть идеальной, т.е. скорость передачи информации должна быть равна пропускной способности канала

$$V = \frac{H(A)}{T} = C,$$

где C определяется по формуле (4.1.3)

$$C = \Delta f \log\left(1 + \frac{P_c}{P_ш}\right) \quad \frac{\text{бит}}{\text{с}}.$$

Скорость создания информации

$H_t(A) = 5 \cdot 10^7 / 3600 = 1,4 \cdot 10^4$ бит/с.

Так как $\log\left(1 + \frac{P_c}{P_ш}\right) = \frac{C}{\Delta f} = \frac{H_t(A)}{\Delta f}$,

то $\frac{P_c}{P_ш} = 2^{\frac{H_t(A)}{\Delta f}} - 1 = 2^{\frac{1,4 \cdot 10^4}{10^5}} - 1 = 0,1019$, и окончательно

$$P_c = 0,1019 P_ш = 0,1019 \cdot \Delta f N_0 =$$
$$0,1019 \cdot 100 \cdot 2 \cdot 10^{-17} \approx 2 \cdot 10^{-16} \quad \text{Вт}.$$

ЗАДАЧИ

4.1.1. В информационном канале используется алфавит с четырьмя различными символами. Длительности всех символов одинаковы и равны $\tau = 1$ мкс. Определить пропускную способность канала при отсутствии шумов.

4.1.2. Вычислить «ненадёжность» для канала, заданного матрицей

$$P = \begin{Vmatrix} 0,1 & 0,2 & 0,3 & 0,4 \\ 0,4 & 0,3 & 0,2 & 0,1 \\ 0,3 & 0,4 & 0,1 & 0,2 \end{Vmatrix},$$

если вероятности входных сигналов равны $p(x_1) = 0$, $p(x_2) = p(x_3) = 0{,}25$.

4.1.3. Вычислить пропускную способность дискретного m – ого симметричного канала связи, если условные вероятности переходов в канале

$$p(b_k / a_j) = \begin{cases} 1 - p, & k = j \\ \dfrac{p}{m-1}, & k \neq j \end{cases},$$

а частота следования символов в канале равна F_k.

4.1.4. Вычислить пропускную способность двоичного симмет-рического постоянного канала, заданного матрицей

$$P = \begin{Vmatrix} 1-p & p \\ p & 1-p \end{Vmatrix},$$

где p – вероятность искажения в канале.

Длительность символа сигнала τ. Построить зависимость $\dfrac{c}{c_m} = f(p)$, где C_m - пропускная способность канала без шумов.

4.1.5. Дискретный канал задан матрицей

$$P = \begin{Vmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{Vmatrix}$$

Длительность символа сигнала 1 мс.

Вычислить пропускную способность канала. Сравнить с пропускной способностью при отсутствии шумов.

4.1.6. Сигнал на входе дискретного канала без помех может принимать одно из следующих значений: 0; 1; 2; 3; 4; 5. Вычислить пропускную способность канала, полагая, что среднее значение входного не превышает единицы.

4.1.7. Сообщения источника с производительностью 850 $бит/с$ поступают на вход двоичного симметричного канала с вероятностью искажения p = 0,05. Длительность символов сигнала в канале $\tau = 1$мс. Достаточна ли пропускная способность канала для передачи всей информации, поступающей от источника?

4.1.8. Источник сообщений с производительностью 300 $бит/с$ подключён к двум симметричным каналам. Первый канал имеет вероятность искажения $p_1 = 0{,}01$ и длительность символов сигнала в канале $\tau_1 = 10$мс, второй - $p_2 = 0{,}02$ и $\tau_2 = 5$мс соответственно. Достаточна ли пропускная способность обоих каналов для передачи всей информации, поставляемой источником в 1 с?

4.1.9. Два двоичных симметричных канала с вероятностью искажения p = 0,05 соединено последовательно. Как изменится пропускная способность нового канала по сравнению с одним каналом?

4.1.10. Решить задачу 4.1.9 для параллельного соединения двух каналов, то есть при условии, что один и тот же символ передаётся по обоим каналам, а два принятых символа не обязательно совпадают.

4.1.11. Дискретный канал задан матрицей

$$P = \frac{1}{2}\begin{Vmatrix} 1-p & 1-p & p & p \\ p & p & 1-p & 1-p \end{Vmatrix}.$$

Показать, что он эквивалентен двоичному симметричному каналу.

4.1.12. Дискретный канал задан матрицей

$$P = \begin{Vmatrix} 1-p & p & 0 & 0 \\ p & 1-p & 0 & 0 \\ 0 & 0 & 1-p & p \\ 0 & 0 & p & 1-p \end{Vmatrix}$$

Нарисовать граф канала. Вычислить пропускную способность канала.

4.1.13. Вычислить пропускную способность троичного канала связи. Частота следования символов в канале F_u. Условные вероятности переходов имеют следующие значения:

$$p(b_1/a_1) = p(b_2/a_2) = p(b_3/a_3) = 1 - p_1 - p_2$$

$$p(b_2/a_1) = p(b_3/a_2) = p(b_1/a_2) = p_1$$

$$p(b_1/a_2) = p(b_3/a_1) = p(b_2/a_3) = p_2$$

4.1.14. Вычислить пропускную способность дискретного канала, заданного матрицей условных вероятностей переходов

$$P = \begin{Vmatrix} 1-p & 0 & 0 \\ 0 & 1-p & 0 \\ 0 & 0 & 1-p \end{Vmatrix}$$

4.1.15. По каналу связи передаются две буквы $p(a_1) = p(a_2) = 0{,}5$. Одна кодируется 0, другая – 1. При отсутствии помех по канала может передаваться 1000 двоичных знаков. Под действием помех в среднем 1 буква из 100 принимается неверно. Найти фактическое количество информации, передаваемой в среднем за 1 с.

4.1.16. Определить требуемую полосу пропускания канала передачи телевизионного чёрно-белого изображения для принятого в СССР стандарта: в 1 с передаётся 25 кадров, в кадре 625 строк по 800 элементов в строке, для передачи которой программы достаточно передавать 8 градаций освещённости каждого элемента изображения, причём любая градация возникает с одинаковой вероятностью. Элементы изображения статистически независимы. Изображение может принимать наиболее хаотичный вид – вид «белого шума». Отношение сигнал/шум на входе приёмника $q = \dfrac{P_c}{P_ш} = 50$.

4.1.17. Определить максимально возможную скорость передачи информации по телеметрическому каналу связи, если полоса канала 1 кГц, а относительная среднеквадратическая ошибка в канале $\delta_m = 1\%$

4.1.18. Источник фототелеграфного изображения с производительностью 200 бит/с подключен ко входу канала связи. Определить минимальную полосу канала, если для требуемого качества передачи достаточно иметь $P_c / P_ш = 20$.

4.1.19. Вычислить пропускную способность стандартного телефонного канала с полосой (0,3 – 3,4) кГц, если шум в канале белый гауссов, а для обеспечения требуемого качества приёма необходимо иметь $P_c / P_ш = 20$ дБ. Как изменится это отношение при той же производительности источника, если сузить полосу канала до 0,8 кГц?

4.1.20. Показать, чему равна предельная пропускная способность (формула (4.1.3)) при использовании широкополосных методов модуляции $(\Delta f \to \infty)$. Построить график зависимости

$C/C_\infty = \varphi(\Delta f)$.

4.1.21. Из формулы (4.1.3) вытекает возможность обмена полосы канала Δf на отношение $P_c / P_ш$ при постоянной пропускной способности. Показать, каким соотношениям подчиняется эта операция обмена для случаев:

а) $P_c / P_ш \ll 1$,

б) $P_c / P_ш \gg 1$.

4.1.22. Оценить степень использования теоретической пропускной способности телеграфного канала связи при амплитудной манипуляции $\mu = \dfrac{B}{C}$, где $B = \dfrac{1}{\tau}$ - скорость передачи в бит/с, С – пропускная способность канала в бит/с, τ - длительность посылки.

Современные стартстопные буквопечатающие аппараты работают со скоростью 100 слов/мин. Средняя длительность слова, включая и интервал между словами, составляет 6 букв. Каждая буква передаётся 1 стартовой, 5 рабочими и 1 стоповой посылкой, причём стоповая посылка на 50% превышает рабочую. При автоматизированной работе буквы передаются непрерывно. Для правильного воспроизведения пря-

4.2 Кодирование в дискретных каналах без шума

Общие принципы кодирования. Дискретным m-ичным каналом без шумов называется канал, в котором сигналы на входе $A(t)$ и выходе $B(t)$ являются последовательностями дискретных случайных величин-символов с основанием кода m, причем входной и выходной сигналы связаны взаимно однозначно (см. рис. 4.1.1)..

Пропускная способность такого канала равна наибольшему возможному значению энтропии сигнала $A(t)$ на его входе

$$C = H\ A(t)\ _{\max}, \quad \text{бит/симв.}$$

Из (2.4.5) имеем

$$C = \log m, \text{ бит/симв} \quad \text{или} \quad C = F_k \log m, \text{ бит/с}, \tag{4.2.1}$$

где F_k – частота следования символов в канале, симв/с.

Пусть $\mathrm{X}^{(1)}$, $\mathrm{X}^{(2)}$, ... – последовательность символов s-ичного источника сообщений при многократном независимом выборе. Тогда скорость создания информации равна

$$H = H(X)F_u, \tag{4.2.2}$$

где F_u – частота следования символов источника,

$H(X)$ – энтропия алфавита источника.

Функции кодирующего устройства заключаются в том, чтобы каждой из букв s-ичното алфавита источника поставить в соответствие кодовое слово, состоящее из букв m-ичного алфавита, используемого для передачи по каналу.

Кодирующее устройство должно удовлетворять двум основным требованиям:

1) обеспечивать безошибочную передачу информации, т.е. взаимно однозначное соответствие между $X(t)$ и $A(t)$ (см. рис. 4.1.1);

2) осуществлять кодирование наиболее экономным образом (с минимальной избыточностью).

Для выполнения первого требования необходимо, во-первых, чтобы различным буквам алфавита источника соответствовали различные кодовые слова и, во-вторых, была предусмотрена возможность разделения кодовых слов при их последовательной передаче.

Для осуществления возможности разделения кодовых слов применяется одна из следующих мер:

1) использование специальной буквы;

2) применение кодовых слов одинаковой длины без разделительных букв – равномерное кодирование;

3) кодовая таблица составляется так, чтобы никакое кодовое слово не являлось началом другого, более длинного.

Требование экономичности заключается в том, что необходимо добиваться при кодировании минимальной средней длины кодового слова

$$L = \sum_{k=1}^{s} l_k p(x_k)\,, \tag{4.2.3}$$

где l_k – длина k-го кодового слова с учетом разделительной буквы, если таковая используется.

Средняя длина кодового слова ограничена снизу, так как для безошибочной передачи информации необходимо выполнение соотношения $H<C$. Это обстоятельство сформулировано в теореме Шеннона о кодировании в дискретных каналах без шумов.

Теорема Шеннона. При кодировании множества сигналов с энтропией $H(X)$ при условии отсутствия шумов средняя длина кодового слова не может быть меньше чем $\dfrac{H(X)}{\log m}$, где m – основание кода. Если вероятности сигналов не являются целочисленными отрицательными степенями числа m, то точное достижение указанной нижней границы невозможно, но при кодировании достаточно длинными блоками к ней можно сколь угодно приблизиться.

Существует несколько способов, позволяющих получать коды с малой избыточностью, причем все они обладают следующими свойствами.

1) Более вероятным буквам источника соответствуют более короткие кодовые слова (принцип статистического кодирования).

2) Никакое кодовое слово не является началом другого, более длинного.

3) Все буквы алфавита, используемого для передачи по каналу, приблизительно равновероятны.

4) Символы в последовательности на выходе кодера практически независимы.

Наилучшим среди этих кодов является код, предложенный Д.А. Хафманом.

Код Хафмана. Кодирование алфавита, состоящего из s букв, по методу Хафмана осуществляется следующим образом.

1) Буквы алфавита источника располагаем в порядке убывания их вероятностей.

2) Выбираем целое число m_0, такое, что

$$2 \le m_0 \le m \quad \text{и} \quad (s - m_0)/(m-1) = a\,,$$

где a – целое положительное число.

При кодировании в двоичном канале $m_0 = m = 2$.

Производим первое сжатие алфавита, т.е. группируем вместе m_0 букв алфавита источника, имеющих наименьшие вероятности, и обозначаем новой буквой. Вычисляем общую вероятность такого сгруппированного подмножества букв.

3) Буквы нового алфавита, полученного в результате первого сжатия, снова располагаем в порядке убывания вероятностей.

4) Производим второе сжатие этого алфавита, т.е. снова группируем вместе m букв, имеющих наименьшие вероятности, и обозначаем новой буквой. Вычисляем общую вероятность сгруппированного подмножества букв.

5) Буквы нового алфавита, полученного на 4-м шаге, располагаем в порядке убывания вероятностей.

6) Осуществляем последовательные сжатия алфавита путем повторения операций 4 и 5, пока в новом алфавите не останется единственная буква.

7) Проводим линии, которые соединяют буквы, образующие последовательные вспомогательные алфавиты. Получается дерево, в котором отдельные сообщения являются концевыми узлами. Соответствующие им кодовые слова получаем, если приписать различные буквы m-ичного алфавита ветвям, исходящим из каждого промежуточного узла.

Код Хафмана имеет среднюю длину кодового слова, не большую чем любой другой код.

Код Шеннона-Фано. Кодирование осуществляется следующим образом.

1) Буквы алфавита источника располагаем в порядке убывания вероятностей.

2) Делим алфавит источника на m групп так, чтобы общие вероятности букв в каждой из групп были примерно равны. В качестве первого символа кодового слова выбираем 0,1, ..., m-1 в зависимости от того, в которой из групп оказалась кодируемая буква.

3) Каждую из групп снова разбиваем на m примерно равновероятных подгрупп и в качестве второго символа кодового слова выбираем 0,1, …, m-1 в зависимости от того, в которой из подгрупп оказалась кодируемая буква.

4) Такое разбиение на все более мелкие группы проводится до тех пор, пока в каждой из групп не окажется по одной букве.

Блоковое кодирование. Выше были рассмотрены способы побуквенного кодирования, когда каждой букве на выходе источника приписывалось свое кодовое слово. Такой способ кодирования позволяет достигнуть минимальной средней длины кодового слова, равной $H(X)/\log m$, только в том случае, когда вероятности букв являются отрицательными целочисленными степенями числа m. Если вероятности не удовлетворяют этому требованию, избыточность при побуквенном кодировании может оказаться недопустимо большой. В таком случае применяют блоковое кодирование.

При кодировании блоками по k букв прежде всего строят вспомогательный алфавит, состоящий из $N=s^k$ букв, так что каждой букве этого алфавита соответствует своя комбинация (блок) из k букв алфавита источника. Вероятность буквы $z_{ij\ldots q}$ вспомогательного алфавита, соответствующей комбинации $x_i x_j \ldots x_q$, в простейшем случае вычисляется по формуле

$$p(z_{ij\ldots q}) = p(x_i)\,p(x_j)\ldots p(x_q)\,. \qquad (4.2.4)$$

Далее буквы вспомогательного алфавита располагаются в порядке убывания вероятностей, и осуществляется кодирование методом Шеннона-Фэно или Хафмана.

Таким образом, кодирующее устройство при блоковом кодировании разбивает последовательность букв на выходе источника на блоки длиной в k букв каждый и генерирует последовательность кодовых слов, соответствующих каждому из блоков.

Главный недостаток обоих кодов заключаются в том, что для их применения нужно знать вероятности появления букв (или их комбинаций при кодировании блоками).

Код Лемпела–Зива свободен от этого недостатка. Здесь кодовая таблица, изначально почти пустая, заполняется одновременно в пунктах передачи и приема в процессе кодирования (декодирования), причем в эту таблицу вносятся лишь такие все более длинные отрезки передаваемого сообщения, которые еще не встречались ранее. Каждому отрезку в таблице присваивается n-разрядный номер. При внесении очередной записи (строки) в таблицу передается блок, содержащий:

1) номер отрезка, уже имеющегося в таблице;

2) символ, следующий в передаваемом сообщении за этим отрезком.

РЕШЕНИЕ ТИПОВЫХ ПРИМЕРОВ

Пример 4.2.1. Закодировать сообщения источника, приведенные в табл 4.2.1, двоичным кодом Хафмана. Оценить эффективность полученного кода.

Таблица 4.2.1

u_k	u_1	u_2	u_3	u_4	u_5	u_6
$p(u_k)$	0,1	0,2	0,25	0,05	0,25	0,15

Решение. В соответствии с алгоритмом построения кода Хафмана делаем последовательно следующие шаги:

1) располагаем сообщения источника в порядке убывания вероятностей;

2) образуем вспомогательный алфавит, объединяя наиболее маловероятные буквы u_1 и u_4 (m_0=2), тогда вероятность новой буквы равна $p_1=p(u_1)+p(u_4)$=0,1+0,05=0,15. Оставляем эту букву на месте, так как $p_1=p(u_6)$;

3) объединяем первую вспомогательную букву и букву u_6, тогда вероятность второй вспомогательной буквы равна $p_2=p_1+p(u_6)$=0,15+0,15=0,3; перемещаем ее вверх в соответствии с этой вероятностью;

4) объединение продолжаем до тех пор, пока в ансамбле не останется единственное сообщение с вероятностью единица.

Построение кода Хафмана приведено на рис. 4.3.

Комбинация	u_j
10	u_3
01	u_5
00	u_2
111	u_6
1101	u_1
1100	u_4

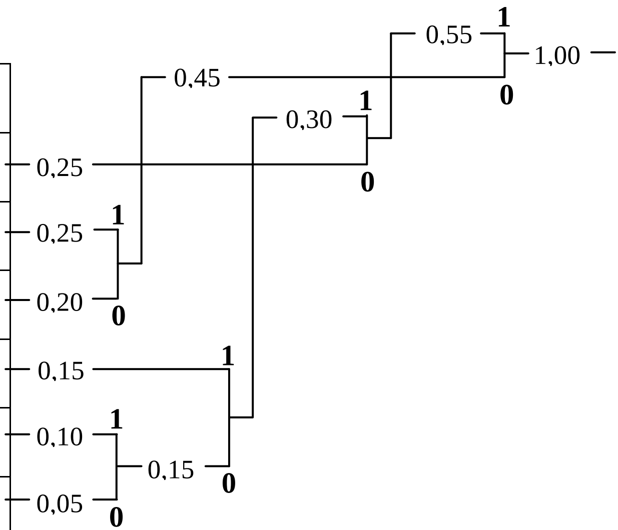


Рис. 4.3

Сообщения источника являются теперь концевыми узлами кодового дерева. Приписав концевым узлам значения символов

1 и 0, записываем кодовые обозначения, пользуясь следующим правилом: чтобы получить кодовое слово, соответствующее сообщению u_4, проследим переход u_4 в группировку с наибольшей вероятностью, кодовые символы записываем справа налево (от младшего разряда к старшему), получим 1100.

Для сообщения u_1 – 1101 и т.д. (см. рис. 4.3).

Оценим эффективность полученного кода.

Энтропия источника сообщений:

$$H(U) = -\sum_{k=1}^{6} p(u_k)\log p(u_k) =$$

$$2\eta(0,25)+\eta(0,2)+\eta(0,15)+\eta(0,1)+\eta(0,05) = 2,4232 \text{ бит}$$

на одну букву на выходе источника.

Средняя длина кодового слова (формула (4.2.3))

$$L = \sum_{k=1}^{6} l_k p(u_k) = 2\cdot 0,25 + 2\cdot 0,25 +$$

$$2\cdot 0,2 + 3\cdot 0,15 + 4\cdot 0,1 + 4\cdot 0,05 = 2,45 \text{ дв. симв/букву.}$$

Для оценки эффективности кода используем коэффициент эффективности $\gamma = H(U)/\ L\log m$.

Для оптимального двоичного кода $H(U) = L$ и $\gamma = 1$.

Полученный нами код имеет $\gamma = 2,4232/2,45 = 0,9891$, избыточность R=0,0109, т.е. код близок к оптимальному.

Пример 4.2.2. Сообщение источника X составляется из статистически независимых букв, извлекаемых из алфавита A, B, C с вероятностями 0,7; 0,2; 0,1. Произвести двоичное кодирование по методу Шеннона-Фано отдельных букв и двухбуквенных блоков. Сравнить коды по их эффективности.

Решение. Производим побуквенное кодирование методом Шеннона-Фано.

1) Располагаем буквы алфавита источника в порядке убывания вероятностей.

2) Делим алфавит источника на две (m=2) примерно равновероятные группы. Всем сообщениям верхней группы (буква A) приписываем в качестве первого кодового символа 1, всем сообщениям нижней группы приписываем символ 0.

3) Производим второе разбиение на две группы (буквы B и

С) и снова букве в верхней группе (В) приписываем символ 1, а в нижней (С) в качестве второго символа кодового слова приписываем 0. Так как в каждой группе оказалось по одной букве, кодирование заканчиваем. Результат приведен в табл. 4.2.2.

Таблица 4.2.2

x_j	$p(x_j)$	Разбиения	Кодовое слово
A	0,7	———	1
B	0,2	—	01
C	0,1		00

Оценим эффективность полученного кода. Энтропия источника

$$H(X) = -\sum_{k=1}^{s} p(x_k)\log_2 p(x_k) =$$

$$\eta(0,7)+\eta(0,2)+\eta(0,1)=1,1568 \text{ бит/букву.}$$

Средняя длина кодового слова

$$L=\sum_{k=1}^{s} l_k\, p(x_k) = 0,7\cdot 1+0,2\cdot 2+0,1\cdot 2 = 1,3 \text{ бит/букву.}$$

Видим, что $L>H(X)$, и коэффициент эффективности $\gamma_1 = 1,1568/1,3 = 0,8898,$ а избыточность R_1=0,1102.

Покажем, что кодирование блоками по 2 буквы (k=2) увеличивает эффективность кода. Строим вспомогательный алфавит из $N=3^2$ блоков. Вероятности блоков находим по формуле (4.2.4), считая буквы исходного алфавита независимыми. Располагаем блоки в порядке убывания вероятностей и осуществляем кодирование методом Шеннона-Фано. Все полученные двухбуквенные блоки, вероятности их и соответствующие кодовые обозначения сведены в табл. 4.2.3.

При блоковом кодировании средняя длина кодового слова на одну букву

$$L_2 = L/2 = 0,5(1\cdot 0,49+3\cdot 2\cdot 0,14+4\cdot 2\cdot 0,07+4\cdot 0,04+$$
$$5\cdot 0,02+6\cdot 0,02+6\cdot 0,01) = 1,165 \text{ бит/букву.}$$

При этом коэффициент эффективности

$$\gamma_2 = H(X)/L_2 = 1,1568/1,165 = 0,9955.$$

Избыточность при двухбуквенном кодировании R_2=0,0045.

Получили $\gamma_2 > \gamma_1$, $R_2 \ll R_1$, что и требовалось показать.

Таблица 4.2.3

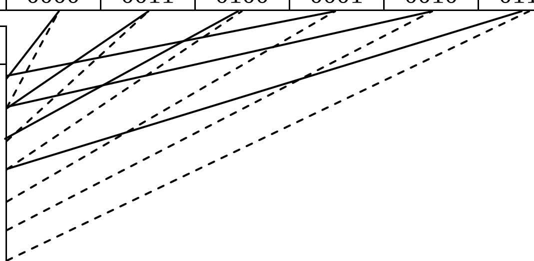
Двухбук-венные блоки	Вероят-ности	Разбиения	Кодовые слова
AA	0,49		1
AB	0,14		0 1 1
BA	0,14		0 1 0
AC	0,07		0 0 1 1
CA	00,7		0 0 1 0
BB	0,04		0 0 0 1
BC	0,02		0 0 0 0 1
CB	0,02		0 0 0 0 0 1
CC	0,01		0 0 0 0 0 0

Пример 4.2.3. Закодировать двоичным кодом Лемпела–Зива двоичное сообщение 0010101000101. Номера комбинаций в кодовой таблице состоят из трех битов.

Решение.

На рис. 4.4 показаны двоичные последовательности на входе и выходе кодера, причем ради наглядности они разбиты на части, соответствующие отдельным шагам. Сплошные линии указывают, какие номера извлекаются из кодовой таблицы и подаются на выход кодера. Штриховые линии показывают путь внесения очередных записей в кодовую таблицу (первая запись

Вход	0	01	010	1	00	0101
Выход	0000	0011	0100	0001	0010	0111



Отрезок	Номер
-	000
0	001
01	010
010	011
1	100
00	101
0101	110

Рис. 4.4 – Пример кодирования кодом Лемпела–Зива

“пробел” и его номер 000 внесены в таблицу заранее). Убедитесь, что по передаваемой последовательности в пункте приема можно в том же порядке заполнять кодовую таблицу и декодировать сигнал.

ЗАДАЧИ

4.2.1. Убедиться в том, что при кодировании двоичным кодом Шеннона-Фано сообщений источника, заданного табл. 4.2.4, может быть достигнута длина кодового слова, удовлетворяющая условию $L_{min} = H(X)/\log m$.

Сравнить полученный код с равномерным кодом.

Таблица 4.2.4

x_j	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
$p(x_j)$	1/2	1/4	1/8	1/16	1/32	1/64	1/128	1/128

4.2.2. Источник статистически независимых сообщений задан табл. 4.2.5.

Таблица 4.2.5

x_j	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_7	x_7
$p(x_j)$	1/4	1/4	1/8	1/8	1/16	1/16	1/16	1/32	1/64	1/64

Закодировать сообщения источника двоичным кодом так, чтобы средняя длина кодового слова была минимальна. Оценить эффективность полученного кода.

4.2.3. Ансамбль сообщений задан табл. 4.2.6.

Таблица 4.2.6

x_j	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
$p(x_j)$	0,2	0,25	0,15	0,1	0,05	0,05	0,14	0,06

Сообщения источника статистически независимы. Передача производится по двоичному каналу. Длительности символов кодированного сигнала одинаковы $\tau_0=\tau_1=\tau$. Определить скорость передачи информации при использовании равномерного кода и кода Шеннона-Фано. Сравнить коды по их эффективности.

4.2.4. Закодировать статистически независимые сообщения источника (табл. 4.2.7) двоичным кодом Хафмана. Найти вероятности появления нулей и единиц в полученной последовательности.

Таблица 4.2.7

x_j	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$p(x_j)$	0,25	0,23	0,15	0,12	0,1	0,08	0,07

4.2.5. Закодировать двоичным кодом Шеннона-Фано алфавит источника, состоящий из четырех букв (А, В, С, Д) с вероятностями 0,28; 0,14; 0,48 и 0,1 соответственно. Оценить эффективность полученного кода.

4.2.6. Сколько вопросов в среднем надо задать, чтобы отгадать задуманное собеседником целое положительное число, не превосходящее 10, если спрашиваемый на все вопросы отвечает лишь «да» или «нет»?

4.2.7. Известно, что жители некоторого города А всегда говорят правду, а жители соседнего города Б всегда обманывают. Наблюдатель Н знает, что он находится в одном из этих двух городов, но не знает, в каком именно. Путем опроса встречного ему требуется определить: а) в каком городе он находится, б) в каком городе живет его собеседник (в каждом пункте можно с одинаковой вероятностью встретить жителей обоих городов), в) то и другое вместе.

Каково наименьшее число вопросов, которые должен задать Н, если на все вопросы Н встречный отвечает лишь «да» или «нет»?

4.2.8. Сообщение на выходе источника без памяти состоит из букв, принимающих значение А и В с вероятностями 0,7 и 0,3. Произвести кодирование по методу Шеннона-Фано отдельных букв, двух- и трехбуквенных блоков. Сравнить коды по их эффективности.

4.2.9. Имеются три города (А, Б и В), причем жители А всех случаях говорят правду, жители Б – только неправду, а жители В через раз отвечают на вопросы верно и неверно. Наблюдатель Н хочет выяснить, в каком городе он находится и в каком городе живет встреченный им человек. Сколько вопросов ему потребуется задать этому встречному, если на все вопросы он отвечает лишь «да» или «нет»?

4.2.10. Некто задумал два различных целых положительных числа, не превосходящих четырех. Сколько в среднем надо задать ему вопросов для того, чтобы определить сумму этих чисел, если на каждый вопрос спрашиваемый отвечает лишь «да» или «нет»?

4.2.11. Построить оптимальное множество кодовых слов для ансамбля X, заданного табл. 4.2.8.

Таблица 4.2.8

x_j	x_1	x_2	x_3	x_4	x_5
$p(x_j)$	1/2	1/4	1/8	1/16	1/32

Показать, что если сообщения статистически независимы, то в получающейся последовательности кодовых слов символы 0 и 1 независимы и появляются с равными вероятностями.

4.2.12. Сообщения гипотетического алфавита, заданного табл. 4.2.9, для передачи по двоичному каналу преобразуются кодером в последовательности кодовых символов 0 и 1 со средней длиной кодового слава $L_{ср} = 3{,}19 \frac{дв.симв}{симв.ист}$

Таблица 4,2.9

a_j	a_1	a_2	a_3	a_4	a_5
$p(a_j)$	0.22	0.16	0.14	0.12	0.1

a_6	a_7	a_8	a_9	a_{10}	a_{11}	а$_{12}$
0.07	0.06	0.05	0.04	0.02	0.01	0.01

Является ли полученный код оптимальным в статистическом смысле?

4.2.13. Радиотехническое устройство состоит из 5 блоков (А, Б, В, Г, Д). Блок А в среднем выходит из строя 1 раз в 100 дней, блок Б – 1 раз в 25 дней, В – 1 раз в 5 дней, Г– один раз в 4 дня и блок Д – 1 раз в 2 дня. Контрольный прибор позволяет за одно измерение проверить работоспособность в целом любой комбинации блоков.

Как нужно проводить контроль, чтобы затратить на поиски неисправного блока в среднем минимальное количество проверок? Найти это среднее значение.

4.2.14. Имеются 12 монет одного достоинства; 11 из них имеют одинаковый вес, а одна – фальшивая – отличается по весу от остальных (причем не известно, легче она или тяжелее настоящих). Каково наименьшее число взвешиваний на чашечных весах без гирь, которое позволяет обнаружить фальшивую монету и выяснить, легче она, чем остальные монеты, или тяжелее?

4.2.15. На кабельной линии длиной 20 км имеются 5 контрольных отводов, расположенных на расстоянии 1, 3, 5, 9 и 14 км от ее начала. Вероятность пробоя кабеля в любой точке линии одинакова. При осуществлении одной проверки можно установить исправность отрезка кабеля между любыми контрольными точками.

Как нужно проводить измерения, чтобы найти участок с поврежденным кабелем и произвести в среднем минимальное количество проверок?

4.2.16. Все буквы алфавита равновероятны, то есть избыточность отсутствует. При каком условии удастся закодировать этот алфавит двоичным кодом без избыточности?

4.2.17. Закодировать кодом Лемпела–Зива двоичную последовательность 0 0 0 0 1 0 0 0 0 0… и определить количество передаваемых двоичных символов, если используются трехразрядные номера комбинаций.

4.3 Кодирование в дискретном канале с шумом

Теоремы Шеннона о кодировании в дискретном канале с шумом. Дискретным каналом с шумом называется канал, в котором сигналы $A(t)$ и $B(t)$ (см. рис. 4.1.1) являются последовательностями символов, однозначная связь между входными и выходными сигналами отсутствует, и поэтому передача символов в канале сопровождается случайными ошибками.

Ответ на вопрос о возможности безошибочной передачи информации дается теоремами Шеннона о кодировании в дискретных каналах с шумом.

Первая и вторая теоремы. Если скорость создания информации H источником на входе шумящего канала без памяти с пропускной способностью C меньше пропускной способности, то существует такой код, при котором вероятность ошибок на приемном конце сколь угодно мала, а скорость передачи информации сколь угодно близка к скорости ее создания.

Обратная теорема. Если $H > C$, то никакой код не сможет сделать вероятность ошибок сколь угодно малой и достигнуть ненадежности, меньшей чем H-C.

Основной способ повышения помехоустойчивости системы передачи информации – это разумное введение избыточности в передаваемый сигнал $A(t)$.

Корректирующие коды. Идея рационального введения избыточности в передаваемый сигнал для борьбы с ошибками в каналах с шумами реализуется применением корректирующих кодов. Так называются коды, которые позволяют обнаруживать и даже исправлять ошибки, возникающие в канале. Наиболее часто применяют двоичные равномерные корректирующие коды.

Пусть $\mathbf{a_1}$, $\mathbf{a_2}$, ... , $\mathbf{a_s}$ – двоичные кодовые слова длиной в n символов каждое, соответствующие s буквам алфавита источника. Каждое кодовое слово можно формально рассматривать как вектор-строку, соответствующий некоторой точке в n-мерном пространстве.

Расстояние Хэмминга d_{jk} между двумя кодовыми словами $\mathbf{a_j}$ и $\mathbf{a_k}$ равно числу символов, в которых эти слова отличаются одно от другого.

Удобно вычислять расстояние между двумя кодовыми словами, пользуясь операцией суммирования по модулю 2. Правила сложения сведены в таблице 4.3.1, где $\oplus$ – знак операции суммирования по модулю 2.

Таблица 4.3.1

$0 \oplus 0=0$	$0 \oplus 1=1$
$1 \oplus 0=1$	$1 \oplus 1=1$

Сложение и вычитание по модулю 2 эквивалентны.

Расстояние между двумя кодовыми словами можно теперь определить как количество единиц в кодовом слове, полученном при суммировании этих слов по модулю 2.

Наименьшее значение расстояния между словами в выбранной системе кодирования называется кодовым расстоянием $d_{код}$.

Считают, что в канале произошла q-кратная ошибка, если кодовое слово на выходе канала отличается от кодового слова на его входе ровно в q символах (т.е. расстояние между этими словами-векторами равно q). Таким образом, вектор на выходе канала равен сумме по модулю 2 входного вектора и вектора ошибки.

Код позволяет обнаруживать любую ошибку кратности $q_о$, если его кодовое расстояние удовлетворяет условию

$$d_{код} \geq q_о + 1. \tag{4.3.1}$$

Код позволяет исправлять любую ошибку кратности $q_и$, если выполняется условие

$$d_{код} \geq 2q_и + 1. \tag{4.3.2}$$

Линейные блочные коды. Среди всех корректирующих кодов наибольшее применение нашли линейные блочные коды. Это – двоичный корректирующий (n,k)-код, удовлетворяющий следующим требованиям:

1) Все кодовые слова содержат по n символов, из которых k символов – информационные, а остальные n-k символов – проверочные.

2) Сумма по модулю 2 любых двух кодовых слов снова дает кодовое слово, принадлежащее этому же коду.

Линейный блочный код код задается производящей матрицей **G**, имеющей k строк и n столбцов. Строки производящей матрицы должны быть линейно независимыми, т.е. сумма любых строк (по 2, по 3 и т.д. в любых комбинациях) не должна давать нулевое слово (состоящее из одних нулей).

Следующим этапом в построении линейного блочного кода является определение проверочной матрицы **H**, имеющей r=n-k строк и n столбцов. Любой вектор-строка проверочной матрицы **H** ортогонален любому вектору-строке производящей матрицы **G**, следовательно,

$$\mathbf{G}\mathbf{H}^{\mathbf{T}}=\mathbf{0}, \tag{4.3.3}$$

где $\mathbf{H}^{\mathbf{T}}$ – транспонированная матрица **H**.

Заметим, что два вектора $\mathbf{u}=u_1,\ldots,u_n$ и $\mathbf{v}=v_1,\ldots,v_n$ являются, по определению, ортогональными, если

$$\mathbf{uv} = u_1v_1 \oplus u_2v_2 \oplus \ldots \oplus u_nv_n = 0 .$$

Работа кодера для (n,k)-кода заключается в том, что к поступающему на его вход k-разрядному вектору информационных символов $\mathbf{x}$ он должен добавить r проверочных символов. Правило образования кодового слова $\mathbf{a}$ определяется соотношением

$$\mathbf{a}=\mathbf{xG}. \tag{4.3.4}$$

Умножив обе части этого выражения справа на $\mathbf{H}^{\mathrm{T}}$ и учитывая (4.3.3), получим другое соотношение, определяющее функции кодера

$$\mathbf{a}\,\mathbf{H}^{\mathrm{T}} = \mathbf{0}, \tag{4.3.5}$$

т.е. он должен добавлять r проверочных символов таким образом, чтобы любой из получаемых в итоге кодовых векторов $\mathbf{a}$ удовлетворял соотношению (4.3.5).

Один из способов декодирования принятого вектора $\mathbf{b}$ предусматривает вычисление r-разрядного исправляющего вектора (синдрома) $\mathbf{c}$ по правилу

$$\mathbf{c}=\mathbf{bH}^{\mathrm{T}}. \tag{4.3.6}$$

Из формулы (4.3.5) следует, что значение синдрома определяется только вектором ошибки. Если $\mathbf{c} \neq \mathbf{0}$, делают заключение о наличии ошибки (обнаружение ошибки). Так как различным ошибкам кратности $q_{\text{и}}$, удовлетворяющей неравенству (4.3.2), соответствуют различные значения синдрома, то вычисленное значение синдрома $\mathbf{c}$ однозначно определяет положение символов, в которых произошли такие ошибки. Эти ошибки исправляются в декодере суммированием принятого вектора $\mathbf{b}$ с соответствующим вектором ошибок.

Основными элементами кодирующих и декодирующих устройств для систематических кодов являются сумматоры по модулю 2 и сдвигающие регистры. Сдвигающим регистром называют цепочку, состоящую из двоичных ячеек (в дальнейшем они изображаются прямоугольниками), меняющую свое состояние в дискретные моменты времени (шагами). На каждом шаге двоичный символ, имеющийся на входе ячейки, перемещается на ее выход.

Ниже приведены краткие данные наиболее распространенных систематических кодов.

Код Хэмминга имеет кодовое расстояние, равное трем, и полностью характеризуется числом проверочных символов r. Общее число символов в кодовом слове равно n=2^r-1. В качестве столбцов проверочной матрицы **H** выбираются всевозможные r-разрядные двоичные числа, исключая число нуль.

Код Рида-Малера полностью характеризуется двумя целыми положительными числами: $m \geq 3$ и $\delta < m$. Число δ называется порядком кода. Остальные параметры кода определяются из соотношений:

$$n = 2^m, \quad k = \sum_{i=1}^{\delta} C_m^i, \quad d_{код} = 2^{m-\delta}. \tag{4.3.7}$$

Первая строка производящей матрицы **G** состоит из единиц. Следующие m строк (базисные векторы первого порядка) можно получить, если записать n столбцов, состоящих из m-разрядных двоичных чисел. Следующие C_m^2 строк (базисные векторы второго порядка) получают, вычисляя поэлементные произведения различных пар базисных векторов первого порядка, и т.д. до получения базисных векторов δ-го порядка.

Циклический код полностью определяется первой строкой производящей матрицы **G**. Остальные строки получают в результате циклического сдвига первой строки на 1,2,...,k-1 элементов. Таким же циклическим свойством обладает и проверочная матрица **H**.

Более удобным является общепринятое описание кодовых комбинаций циклического кода при помощи полиномов. Кодовой комбинации $\mathbf{v}=(v_0,v_1,\ldots,v_{n-1})$ соответствует полином $v(x)= v_0x^0+ v_1x^1+,\ldots,v_{n-1}x^{n-1}$. Способы кодирования и декодирования для конкретного кода полностью определяются, если задать производящий полином $g(x)=1+g_1x+ g_2x^2+\ldots+ g_{r-1}x^{r-1}+x^r$.

Фундаментальное свойство циклического кода заключается в том, что полином, соответствующий любой разрешенной (передаваемой) кодовой комбинации, делится без остатка на производящий полином.

РЕШЕНИЕ ТИПОВЫХ, ПРИМЕРОВ

Пример 4.3.1. Способ кодирования задан кодовой таблицей

a$_1$ = 0000000
a$_2$ = 0110111
a$_3$ = 1011010
a$_4$ = 1101100.

а) составить матрицу расстояний d_{ij} $(i, j = 1, 2, 3, 4)$;

б) найти кодовое расстояние;

в) определить кратности обнаруживаемых и исправляемых ошибок;

г) определить избыточность кода, полагая буквы источника равновероятными.

Решение. а) Матрицу расстояний записываем в виде таблицы (табл. 4.3.2).

Таблица 4.3.2

j	i			
	1	2	3	4
1	0	5	4	4
2	5	0	5	5
3	4	5	0	4
4	4	5	4	0

б) По табл. 4.3.2 находим кодовое расстояние $d_{код} = \min d_{ij} = 4, \quad i \neq j.$

в) Кратность обнаруживаемых ошибок определяется по формуле (4.3.1), откуда $q_о \leq 3$.

Кратность исправляемых ошибок находим по формуле (4.3.2) $q_и \leq 1{,}5$.

Следовательно, приведенный код позволяет обнаруживать всевозможные однократные, двукратные и трехкратные ошибки и исправлять любые однократные ошибки.

г) Избыточность кода находим из следующих соображений. Для передачи равновероятных сигналов **a**$_1$, **a**$_2$, **a**$_3$, **a**$_4$ достаточно передавать кодовые слова 00, 10, 01 и 11 соответственно. Такой код не имеет избыточности, но не позволяет обнаруживать и, тем более, исправлять ошибки. Для обнаружения и исправления ошибок введены пять избыточных символов, т.е. количественно избыточность равна $R = (7-2)/7 = 71\%$.

Пример 4.3.2. Линейный блочный (5,2)-код задан производящей матрицей в систематической (канонической) форме

$$\mathbf{G} = \begin{Vmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{Vmatrix} = \begin{Vmatrix} 10110 \\ 01011 \end{Vmatrix}$$

Пусть принят вектор **b**=00110 и известно, что возможны только одиночные ошибки.

Произвести декодирование следующими методами:

а) по минимуму расстояния;

б) вычислением синдрома.

Решение. Если производящая матрица записана в каноническом виде, это значит, что она состоит из двух блоков: диагональной матрицы размера $k \times k$, состоящей из единиц, и прямоугольной матрицы размера $r \times k$. В этом случае первые k символов в любом кодовом слове являются информационными.

Проверочная матрица **H** также может быть записана в каноническом виде и состоит из двух блоков. Первый блок есть прямоугольная матрица размера $k \times r$, полученная транспонированием второго блока матрицы **G**. В качестве второго блока матрицы **H** записывают диагональную матрицу размера $r \times r$, состоящую из единиц.

Для заданного кода получаем

$$\mathbf{H} = \begin{Vmatrix} 10100 \\ 11010 \\ 01001 \end{Vmatrix}.$$

Убедимся, что полученная таким образом матрица **H** удовлетворяет соотношению (4.3.3). Действительно,

$$\mathbf{h_1 g_1} = (10100) \cdot (10110) = 0,$$
$$\mathbf{h_1 g_2} = (10100) \cdot (01011) = 0,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\mathbf{h_3 g_2} = (01001) \cdot (01011) = 0,$$

т.е. все 6 элементов матрицы $\mathbf{GH^T}$ равны нулю.

Построим кодовую таблицу, воспользовавшись правилом образования кодовых слов по формуле (4.3.4):

$\mathbf{a_1}$=00000, $\mathbf{a_2}$=10110, $\mathbf{a_3}$=01011, $\mathbf{a_4}$=11101=$\mathbf{a_2}$+$\mathbf{a_3}$.

Всего кодовая таблица содержит $2^k = 4$ вектора.

Из кодовой таблицы определяем величину кодового расстояния $d_{код} = 3$.

Следовательно, рассматриваемый код обнаруживает однократные и двукратные ошибки и исправляет однократные.

Декодируем принятый вектор **b** = 00110.

а) Метод декодирования по минимуму расстояния заклю-

Таблица 4.3.3

$\mathbf{a_i}$	00000	10110	01011	11101
d_{ib}	2	1	3	4

чается в том, что, вычислив расстояния вектора **b** относительно всех векторов $\mathbf{a_i}$ кодовой таблицы, отождествляем принятый вектор с тем, расстояние до которого минимально. Расстояния d_{ib} приведены в табл. 4.3.3.

По величине d_{ibmin}=1 решаем, что передавался вектор

$\mathbf{a}_2$= 10110,

следовательно, ошибка в первом символе кодового слова, а информационная последовательность имеет вид **x** = 10.

б) Метод декодирования с помощью вычисления синдрома включает следующие операции:

1) По формуле (4.3.6) заранее устанавливаем однозначную связь между векторами однократных ошибок **e** и соответствующими им синдромами. Все возможные векторы $\mathbf{c} = \mathbf{eH}^{\mathbf{T}}$ приведены в табл. 4.3.4.

Таблица 4.3.4

e	**c**
00001	001
00010	010
00100	100
01000	011
10000	110

2) Вычисляем синдром для принятого слова **b** по формуле (4.3.6) $\mathbf{c}=\mathbf{bH}^{\mathbf{T}}$:

c_1=(00110)(10100)=1, c_2=(00110)(11010)=1,

c_3=(00110)(01001)=0,

т.е. вектор **c** = 110. Синдром не равен нулю, следовательно, есть ошибка.

Вектору **c** = 110 в табл. 4.3.4 соответствует вектор ошибки в первом символе **e**=10000. Декодируем, суммируя принятый вектор с вектором ошибки $\breve{\mathbf{a}} = \mathbf{b} \oplus \mathbf{e} = 00110 \oplus 10000 = 10110.$

Итак, получили тот же результат: передавался вектор $\mathbf{a}_2$=10110, соответствующий информационной последовательности **x**=10.

Пример 4.3.3. Для кода, заданного в примере 4.3.2, составить схему кодирующего устройства.

Решение. Обозначим буквами $a_1,...,a_5$ символы на выходе кодера, причем a_1 и a_2 есть информационные символы, поступающие на его вход, а символы a_3, a_4 и a_5 образуются в кодере.

Из соотношений (4.3.4) или (4.3.5) получаем

$$a_3 = a_1, \qquad a_4 = a_1 + a_2, \qquad a_5 = a_2.$$

Один из вариантов схемы кодера, определяемой этими соотношениями, показан на рис. 4.4.

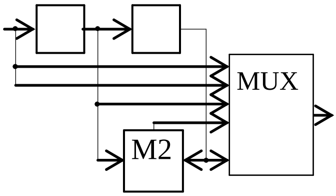


Рис. 4.4

Кодирующее устройство работает следующим образом. В течение первых двух шагов двухразрядный регистр сдвига заполняется информационными символами a_1 и a_2, а в течение следующих трех тактов его состояние сохраняется неизменным. Одновременно с заполнением регистра начинается переключение мультиплексора (переключателя) MUX. Таким образом формируются 5 символов кодового слова, удовлетворяющих требуемым соотношениям.

Пример 4.3.4. Построить код Хэмминга, имеющий параметры: $d_{код}$=3, r = 3.

Решение. Построение кода начинаем с проверочной матрицы. В качестве семи столбцов проверочной матрицы **H** выбираем всевозможные 3-разрядные двоичные числа, исключая число нуль. Проверочная матрица имеет вид

$$\mathbf{H} = \begin{Vmatrix} 0001111 \\ 0110011 \\ 1010101 \end{Vmatrix}.$$

Чтобы определить места проверочных и информационных символов, матрицу **H** представляем в каноническом виде

$\mathbf{H_0} = \|\mathbf{Q},\mathbf{I}\|$, где **I** – единичная матрица.

Для этого достаточно в матрице **H** выбрать столбцы, содержащие по одной единице, и перенести их в правую часть. Тогда получим

$$\mathbf{H_0} = \begin{Vmatrix} 0111000 \\ 1011010 \\ 1101001 \end{Vmatrix}.$$

Из формулы (4.3.5) получаем следующие соотношения для символов a_1, ..., a_7 кодового слова:

$$a_5 = a_2 \oplus a_3 \oplus a_4, \qquad a_6 = a_1 \oplus a_3 \oplus a_4, \qquad a_7 = a_1 \oplus a_2 \oplus a_4.$$

Пользуясь этими соотношениями, составляем кодовую таблицу

$$0000 \to 0000000, \quad 0001 \to 0001111, \ldots, 1111 \to 1111111.$$

Пример 4.3.5. Для кода Хэмминга, определяемого проверочной матрицей **H**, построенного в примере 4.3.4, декодировать принятый вектор 1011001.

Решение. Для принятого вектора **b**=1011001 по формуле (4.3.6) вычисляем синдром

$$c_1 = (1011001) \cdot (0001111) = 0, \quad c_2 = (1011001) \cdot (0110011) = 0,$$

$$c_3 = (1011001) \cdot (1010101) = 1, \quad \text{в итоге} \quad \mathbf{c} = 001.$$

Синдром **c**≠0, т.е. имеет место ошибка. В случае одиночной ошибки (и только лишь) вычисленное значение синдрома всегда совпадает с одним из столбцов матрицы **H**, причем номер столбца указывает номер искаженного символа. Видим, что ошибка произошла в первом символе и передавался вектор $\mathbf{a} = 0011001$.

ЗАДАЧИ

4.3.1. Составить кодовую таблицу, определить кодовое расстояние и вычислить минимальное значение избыточности 3-разрядного двоичного кода, удовлетворяющего требованиям:

а) код содержит максимальное количество кодовых слов;

б) код обнаруживает все однократные ошибки;

в) код исправляет все однократные ошибки.

Построить геометрические модели полученных кодов.

4.3.2. Дать правило построения простейшего кода, обнаруживающего все ошибки нечетной кратности.

4.3.3. Показать, что векторы

V_1 = 10111, V_2 = 01011,V_3 - = 01110

линейно независимы, а векторы

U_1=10001, U_2=10011, U_3=0110

ортогональны векторам V_I.

4.3.4. Найти все кодовые векторы для линейного блочного кода, заданного порождающей матрицей

$$G = \begin{Vmatrix} 11010 \\ 11001 \\ 01110 \end{Vmatrix}$$

4.3.5. Для кода, заданного порождающей матрицей

$$G = \begin{Vmatrix} 1001101 \\ 0101011 \\ 0010111 \end{Vmatrix}$$

построить проверочную матрицу Н.

4.3.6. Какими корректирующими свойствами обладает код заданный кодовой таблицей

$$a_1 = 000000$$
$$a_2 = 000111$$
$$a_3 = 111000$$
$$a_4 = 111111$$

Является ли этот код линейным блочным?

4.3.7. Кодирующее устройство осуществляет n-кратное повторение каждого из поступающих на его вход двоичных, символов источника:

а) убедиться в том, что такой код является систематическим и записать производящую и проверочную матрицы для n = 4;

б) каким должно быть число повторений n, чтобы вероятность искажения символов источника при декодировании непревышала 0, 0002, если вероятность искажения символа в двоичном стационарном симметричном канале без памяти равна 0,1?

4.3.8. Систематический код задан производящей матрицей

$$G = \begin{Vmatrix} 10011 \\ 01001 \\ 00110 \end{Vmatrix}$$

а) составить схему кодирующего устройства;

б) закодировать информационную последовательность, заданную в виде десятичного числа x = 7.

4.3.9. Систематический (n, k) код задан производящей матрицей

$$G = \begin{Vmatrix} 1000011 \\ 0100101 \\ 0010110 \\ 0001111 \end{Vmatrix}$$

Декодировать принятый вектор 0111011 методами минимального расстояния и вычисления синдрома.

4.3.10. Кодирующее устройство отображает последовательность информационных символов простого равномерного кода 00, 01, 10, 11 в последовательность кодовых векторов 00000, 01101, 10111, 11010, принадлежащих корректирующему коду:

а) показать, что полученный код является систематическим. Выразить каждый символ кодового слова в виде линейной комбинации информационных символов;

б) найти для этого кода производящую и проверочную матрицы;

в) привести таблицу декодирования (вычисления синдрома) для случая q = 1.

4.3.11. Построить код Хэмминга, заданный параметрами $d_{код} = 3$, r = 4, без приведения матрицы H к каноническому виду. Декодировать принимаемый вектор

b=110101010101010.

4.3.12. Код Рида—Малера задан параметрами m = 3 и δ =1. Закодировать десятичное число x = 11.

4.3.13. Построить код Рида–Малера с параметрами m = 4 и δ= 2 для информационной последовательности

x = 10101100111.

4.3.14. Составить схему кодирующего устройства для (8,4)-кода Рида–Малера.

4.3.15. Какому коду принадлежит проверочная матрица

$$H = \begin{Vmatrix} 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \end{Vmatrix}?$$

4.3.16. Представленное на рис. 4.3.2 устройство используется для кодирования символов при передаче по двоичному симметричному каналу.

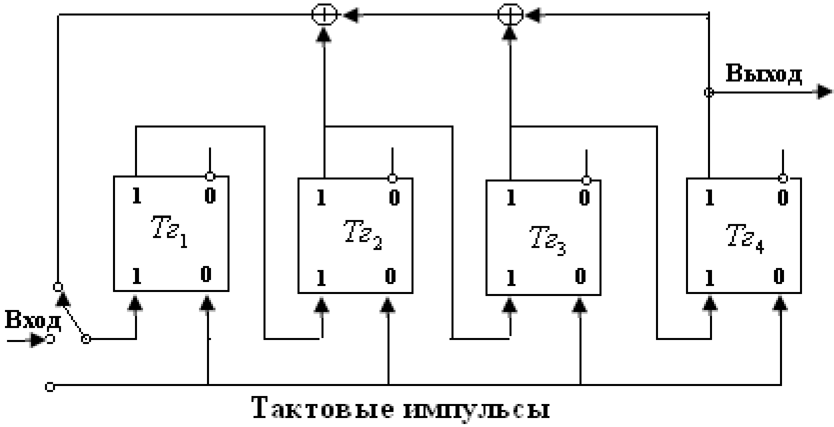


Рис. 4.3.2

В положении ключа 1 на вход подается информационная последовательность 1010, через 4 шага информационные символы занимают все ячейки регистра. Ключ **К** переводится положение 2, замыкая цепь обратной связи, и регистр совершает еще 7 шагов, причем при каждом шаге на выходе появляется очередной символ кодового вектора. Убедиться в том, что этот код принадлежит циклическому, и записать проверочную матрицу.

4.3.17. После кодирования линейным блочным кодом сигнал с выхода стационарного источника передается по стационарному двоичному симметричному каналу без памяти. Вероятность ошибки в канале равна 0,04. Найти минимальное отношение числа проверочных символов к числу информационных, необходимое для того, чтобы вероятность ошибки про декодирования была пренебрежимо мала.

4.3.18. Определить кодовое расстояние для кода, если комбинации в кодовой таблице – это сдвинутые во времени не более чем на период отрезки бесконечной М-последовательности, формируемой генератором, содержащим 3-разрядный двоичный регистр сдвига.

4.3.19. Построить ряд распределения для расстояния между двумя комбинациями кода Хэмминга (7,4), если одна из них (опорная) – это комбинация 0000000, а другая взята наугад. Проверить, изменится ли он, если в качестве опорной взять другую комбинацию.

4.3.20. Проводится каскадное кодирование, при этом первый этап – это кодирование кодом Хэмминга (7,4), а на втором этапе полученные семь символов кодируются кодом с проверкой на четность. Указать какой-то другой из известных кодов, эквивалентный полученному коду, то есть для которого основные параметры имеют те же значения.

4.3.21. Для передачи в двоичном симметричном канале с вероятностью ошибки $5{\cdot}10^{-3}$ применяется код Хэмминга (31,26). Определить битовую вероятность ошибки на выходе, если:

а) используется идеальный канал переспроса;

б) канал переспроса отсутствует.

4.3.22. Построить систему ортогональных двоичных сигналов для n = 8. Для той же длины комбинаций записать систему биортогональных сигналов. В обоих случаях найти кодовое расстояние.

4.3.23. Комбинация, закодированная кодом Хэмминга с проверочной матрицей $H = \begin{Vmatrix} 0\,0\,0\,1\,1\,1\,1 \\ 0\,1\,1\,0\,0\,1\,1 \\ 1\,0\,1\,0\,1\,0\,1 \end{Vmatrix}$, на входе приемника имеет вид 1 1 1 X 0 X 0, где буквой X обозначены стертые символы, а вероятность того, что остальные символы приняты верно, очень велика. Восстановить значения стертых символов.

4.3.24. Передача осуществляется в двоичном симметричном канале с независимыми ошибками, битовая вероятность ошибки p = 10^{-5}. Таков же канал переспроса, но p_o = 10^{-7}. Найти вероятность правильного декодирования кодовой комбинации, если в прямом канале применен простейший код с проверкой на четность (n = 100).

4.3.25. Передача осуществляется в двоичном симметричном канале с независимыми ошибками, битовая вероятность ошибки $p = 10^{-4}$. Таков же канал переспроса, но $p_o = 10^{-8}$. Найти вероятность повторной передачи кодовой комбинации, если в прямом канале применен линейный блочный (3,1)-код.

4.3.26. Передача осуществляется в двоичном симметричном канале с независимыми ошибками, битовая вероятность ошибки $p = 10^{-5}$. Таков же канал переспроса, но $p_o = 10^{-7}$. Найти вероятность правильного декодирования кодовой комбинации, если в прямом канале применен код (32,31).

4.3.27. Циклический код (7,4) задан проверочным полиномом $h(x) = 1 + x + x^2 + x^4$. Декодировать принятое кодовое слово 0101010 по минимуму расстояния и методом вычисления синдрома, сравнить результаты.

4.3.28. Найти вероятность правильного декодирования кодового слова в СПИ с каналом переспроса, если в прямом канале используется код Хэмминга с пятью проверочными символами. Вероятность ошибки в одном символе в прямом канале $p = 10^{-6}$, а в канале переспроса $p_o = 10^{-6}$, ошибки независимые.

4.3.29. Комбинации блочного (n,k)-кода заданы путем случайного выбора.

а) Найти вероятность того, что в кодовой таблице окажется более одной нулевой комбинации.

б) Найти математическое ожидание и СКО расстояния между комбинациями.

Решить задачу для случая $n = 100$, $k = 10$.

4.3.30. Систематический сверточный код, характеризуемый степенью кодирования k/n=1/3, задан полиномами $g_1(x) = 1$, $g_2(x) = 1 + x$, $g_3(x) = 1 + x^2$. Определить битовую вероятность ошибки при пороговом декодировании, если битовая вероятность ошибки в канале равна 10^{-3}.

4.3.31. Записать обе проверочные последовательности для сверточного кода из задачи 4.3.30., если информационная последовательность на входе кодера – периодическая и имеет вид …011110011110… .

4.3.32. Способ кодирования определен кодовой таблицей

Сообщение	да	нет

Кодовая комбинация	11111	00000

Является ли этот код систематическим и, если да, то записать для него производящую и проверочную матрицы.

4.3.33. Кодовую комбинацию, соответствующую десятичному числу 13, закодировать кодом Рида–Малера первого порядка так, чтобы в комбинации было не более 20 символов.

4.3.34. Кодовую таблицу, содержащую все возможные 5-разрядные двоичные комбинации, сократить (отобрать разрешенные к передаче кодовые слова) так, чтобы получившийся код позволял обнаруживать любые однократные и двукратные ошибки.

4.3.35. Кодовую таблицу, содержащую все возможные 15-разрядные двоичные комбинации, сократить (отобрать разрешенные к передаче кодовые слова) так, чтобы получившийся код позволял исправлять любые однократные и двукратные ошибки.

5 ДРУГИЕ МЕРЫ ИНФОРМАЦИИ

5.1 Информация по Кульбаку

Определение. Свойства. Пусть наблюдаемый сигнал **Y** описывается системой непрерывных случайных величин $Y^{(1)}, Y^{(1)}, ..., Y^{(n)}$, которые при гипотезе H_1 имеют совместную плотность вероятности $W_1(y^{(1)}, ..., y^{(n)})$, а при гипотезе H_2 – плотность $W_2(y^{(1)}, ..., y^{(n)})$. Тогда логарифм отношения правдоподобия

$$I(1:2;\mathbf{y}) = \log \frac{W_1(y^{(1)}, ..., y^{(n)})}{W_2(y^{(1)}, ..., y^{(n)})} \tag{5.1.1}$$

называется информацией для различения в пользу H_1 против H_2, содержащейся в выборке $y^{(1)}, ..., y^{(n)}$. Эта величина случайна, так как значение выборки неизвестно до опыта.

В качестве гипотез H_1 и H_2 могут выбираться любые предположения, в том числе такие:

H_1 – «переданное значение сообщения равно x_1»,

H_2 – «переданное значение сообщения равно x_2 ».

Математическое ожидание случайной величины (5.1.1) при условии, что справедливо предположение H_1,

$$I(1:2) = \mathrm{M}_{H_1}\left[\log \frac{W_1(y^{(1)},...,y^{(n)})}{W_2(y^{(1)},...,y^{(n)})}\right] = \tag{5.1.2}$$

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} W_1(y^{(1)},...,y^{(n)})\log\frac{W_1(y^{(1)},...,y^{(n)})}{W_2(y^{(1)},...,y^{(n)})}dy^{(1)},...,dy^{(n)}$$

называется средней информацией для различения в пользу H_1 против H_2.

Аналогично величина

$$I(2:1) = \mathrm{M}_{H_2}\left[\log \frac{W_2(y^{(1)},...,y^{(n)})}{W_1(y^{(1)},...,y^{(n)})}\right] = \tag{5.1.3}$$

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} W_2(y^{(1)},...,y^{(n)})\log\frac{W_2(y^{(1)},...,y^{(n)})}{W_1(y^{(1)},...,y^{(n)})}dy^{(1)},...,dy^{(n)}$$

называется средней информацией для различения в пользу H_2 против H_1.

Сумма (5.1.2) и (5.1.3)

$$J(1,2) = I(1:2) + I(2:1) \tag{5.1.4}$$

называется информационным расхождением между гипотезами H_1 и H_2.

Формулы (5.1.1) – (5.1.4) можно использовать и для вычисления информации, содержащейся в системе дискретных случайных величин **Y**. При этом следует вместо плотностей вероятностей подставить соответствующие вероятности, а интегрирование заменить суммированием по всем возможным значениям системы дискретных случайных величин.

Перечислим основные свойства:

1) Выпуклость.

$$I(1:2) \geq 0, \qquad I(2:1) \geq 0, \qquad J(1,2) \geq 0. \tag{5.1.5}$$

Неравенства обращаются в равенства тогда и только тогда, когда **Y** не зависит от H, т. е.

$$W_1(y^{(1)},...,y^{(n)}) = W_2(y^{(1)},...,y^{(n)}).$$

2) Аддитивность.

Если подсистемы случайных величин $Y^{(1)},...,Y^{(k)}$ и $Y^{(k+1)},...,Y^{(n)}$ независимы при каждой из гипотез, т. е.

$$W_i(y^{(1)},,..,y^{(k)},y^{(k+1)},...,y^{(n)}) = W_i(y^{(1)},...,y^{(k)})\cdot W_i(y^{(k+1)},...,y^{(n)}), \qquad \imath = 1,2, \tag{5.1.6}$$

то информация, содержащаяся в системе $Y^{(1)},...,Y^{(n)}$, равна сумме информаций, содержащихся в подсистемах.

Неравенство для ошибок первого и второго рода. Рассмотрим задачу проверки гипотез. Пространство всех возможных значений сигнала $\mathbf{Y} = (Y^{(1)},...,Y^{(n)})$ разбиваем на две непересекающиеся части: E_1 и E_2. Если выборка **y** попала в область E_1, то считаем, что справедлива гипотеза H_1, в противном случае принимаем гипотезу H_2. При использовании такого правила решения возможны ошибки двух видов.

1) Ошибка первого рода.

Если справедливо предположение H_1 (сигнал **Y** имеет плотность вероятности $W_1(y^{(1)},...,y^{(n)})$, но в результате опыта выборочное значение **y** попало в область E_2 и, следовательно, была принята гипотеза H_2, то возникает ошибка первого рода (неправильное отклонение гипотезы H_1).

2) Ошибка второго рода.

Если справедливо предположение H_2, но выборочное значение **y** попало в область E_1 и была принята гипотеза H_1, то возникает ошибка второго рода.

Вероятность ошибки первого рода

$$\alpha = \int\underset{E_2}{\cdots}\int W_1(y^{(1)},...,y^{(n)})dy^{(1)}...dy^{(n)} \tag{5.1.7}$$

есть вероятность попадания **y** в область E_2 при условии, что **y** имеет распределение $W_1(\mathbf{y})$.

Вероятность ошибки второго рода

$$\beta = \int\underset{E_1}{\cdots}\int W_2(y^{(1)},...,y^{(n)})dy^{(1)}...dy^{(n)} \tag{5.1.8}$$

есть вероятность попадания **y** в область E_1 при условии, что **y** имеет распределение $W_2(\mathbf{y})$.

Вероятности ошибок первого и второго рода связаны с информационными мерами Кульбака следующими неравенствами:

$$I(2:1) \geq \beta \log \frac{\beta}{1-\alpha} + (1-\beta) \log \frac{1-\beta}{\alpha}, \tag{5.1.9а}$$

$$I(1:2) \geq \alpha \log \frac{\alpha}{1-\beta} + (1-\alpha) \log \frac{1-\alpha}{\beta}. \tag{5.1.9б}$$

Неравенства обращаются в равенства тогда и только тогда, когда плотности вероятности $W_1(\mathbf{y})$ и $W_2(\mathbf{y})$ таковы, что

$$\frac{W_1(y^{(1)},...,y^{(n)})}{W_2(y^{(1)},...,y^{(n)})} = C_1 \tag{5.1.10а}$$

для всех $\mathbf{y}$ из области E_1 и

$$\frac{W_1(y^{(1)},...,y^{(n)})}{W_2(y^{(1)},...,y^{(n)})} = C_2 \tag{5.1.10б}$$

для всех $\mathbf{y}$ из области E_2, где C_1 и C_2 – некоторые постоянные.

Задав конкретные значения α и β, по (5.1.9) можно проверить, являются ли эти значения принципиально достижимыми при заданных вероятностных свойствах сигнала $\mathbf{Y}$. Таким образом, зная информационные характеристики сигнала, по (5.1.9) можно вычислить предельные, потенциальные значения вероятностей ошибок первого и второго рода.

Критерий Неймана-Пирсона. Найден оптимальный метод различения гипотез H_1 и H_2, который при заданной вероятности ошибок первого рода обеспечивает минимальную вероятность ошибки второго рода. Такое правило называется критерием Неймана-Пирсона и основано на сравнении логарифма отношения правдоподобия с порогом C, который выбирается таким образом, чтобы обеспечить требуемую величину α. Область E_1 состоит из тех значений $\mathbf{Y}$, для которых справедливо неравенство

$$I(1:2;\mathbf{y}) = \log \frac{W_1(y^{(1)},...,y^{(n)})}{W_2(y^{(1)},...,y^{(n)})} \geq C. \tag{5.1.11}$$

Таким образом, гипотеза H_1 принимается только в том случае, когда величина информации в пользу H_1 против H_2, содержащейся в выборке **y**, больше заданной величины *C*. В противном случае H_1 отвергается и принимается гипотеза H_2.

РЕШЕНИЕ ТИПОВЫХ ПРИМЕРОВ

Пример 5.1.1. Случайная величина Y имеет нормальное распределение с известным среднеквадратическим отклонением σ . Вычислить $I(1:2)$, $I(2:1)$, $J(1,2)$ для следующих гипотез о математическом ожидании этой величины:

H_1: $m=m_1$,

H_2: $m=m_2$.

Решение. Запишем плотности вероятности случайной величины Y , соответствующие каждой из гипотез:

$$W_1(y)=\frac{1}{\sqrt{2\pi}\sigma}\exp[-\frac{1}{2\sigma^2}(y-m_1)^2],$$

$$W_2(y)=\frac{1}{\sqrt{2\pi}\sigma}\exp[-\frac{1}{2\sigma^2}(y-m_2)^2].$$

По формуле (5.1.1) находим информацию для различения в пользу H_1 против H_2, содержащуюся в выборочном значении y (в этой задаче удобнее использовать натуральные единицы информации):

$$I(1:2;y)=\frac{1}{2\sigma^2}[(y-m_2)^2-(y-m_1)^2]=$$

$$\frac{1}{2\sigma^2}[2y(m_1-m_2)+m_2^2-m_1^2].$$

По формуле (5.1.2) находим среднюю информацию для различения в пользу H_1 против H_2

$$I(1:2)=\mathrm{M}_{\mathrm{H}_1}\left\{\frac{1}{2\sigma^2}[2Y(m_1-m_2)+m_2^2-m_1^2]\right\}=$$

$$\frac{1}{2\sigma^2}\ \ 2(m_1-m_2)\cdot\mathrm{M}_{\mathrm{H}_1}[Y]+m_2^2-m_1^2\ \ .$$

Далее учтем, что при гипотезе H_1 математическое ожидание $\mathrm{M}_{\mathrm{H}_1}(Y)=m_1$, и получим окончательно

$$I(1:2)=\frac{1}{2\sigma^2}[2(m_1-m_2)m_1+m_2{}^2-m_1{}^2]=\frac{(m_1-m_2)^2}{2\sigma^2}.$$

По формулам (5.1.3.) и (5.1.4) находим

$$I(2:1)=\mathrm{M}_{\mathrm{H}_2}\left\{\frac{1}{2\sigma^2}[2Y(m_2-m_1)+m_1{}^2-m_2{}^2]\right\}=\frac{(m_1-m_2)^2}{2\sigma^2},$$

$$J(1;2)=I(1:2)+I(2:1)=\frac{(m_1-m_2)^2}{\sigma^2}.$$

Таким образом, средняя информация для различения гипотез H_1 и H_2 в данной задаче пропорциональна квадрату расстояния между математическими ожиданиями сигнала Y и обратно пропорциональна его дисперсии.

Пример 5.1.2. Случайная величина Y имеет одностороннее экспоненциальное распределение

$$W(y)=\frac{1}{m}\exp(-\frac{y}{m}),\quad y\geq 0.$$

Математическое ожидание m этой случайной величины может принять одно из двух значений:

H_1: $m=m_1=1$,

H_2: $m=m_2=10$.

а) Найти правило различения гипотез H_1 и H_2, которое при вероятности ошибки первого рода α =0,01832 обеспечивает минимальную вероятность ошибки второго рода β.

б) Вычислить β, соответствующую найденному правилу.

в) Пользуясь неравенством (5.1.9), установить, возможно ли при проверке гипотез в данной задаче обеспечить $\alpha=\beta=0{,}1$.

Решение. а) Запишем плотности вероятности случайной величины Y, соответствующие каждой из гипотез:

$$W_1(y)=\frac{1}{m_1}\exp(-\frac{y}{m_1}),\quad W_2(y)=\frac{1}{m_2}\exp(-\frac{y}{m_2}),\quad y\geq 0.$$

Графики этих функций приведены на рис. 5.1.

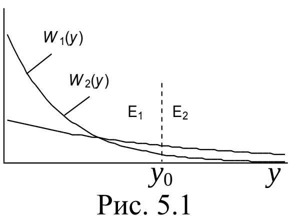


Рис. 5.1

По формуле (5.1.1) вычислим информацию, содержащуюся в выборочном значении y

$$I(1:2;y)=\ln\frac{m_2}{m_1}-y\frac{m_2-m_1}{m_1 m_2}.$$

В соответствии с критерием Неймана-Пирсона эту величину необходимо сравнить с порогом C и принять гипотезу H_1, если выполняется неравенство $I(1:2;y)\ge C$. Отсюда область E_1 принятия гипотезы H_1 состоит из всех значений y, удовлетворяющих неравенству

$$y\le\left(C-\ln\frac{m_2}{m_1}\right)m_1 m_2\Big/(m_2-m_1)=y_0.$$

Правая часть этого неравенства есть постоянная, которую мы обозначили y_0, и, следовательно, область E_1 состоит из всех значений сигнала, для которых $0\le y\le y_0$. Остальные значения $y>y_0$ относим к области E_2 (см. рис. 5.1).

Вероятность ошибки первого рода

$$\alpha=\int\limits_{y_0}^{\infty}W_1(y)dy=\int\limits_{y_0}^{\infty}\frac{1}{m_1}\exp(-\frac{y}{m_1})dy=-\exp(-\frac{y}{m_1})\Bigg|_{y_0}^{\infty}=\exp(-\frac{y_0}{m_1}).$$

Границу y_0 необходимо выбрать так, чтобы обеспечить заданное значение $\alpha=\exp(-y_0/m_1)=0,01832,$ отсюда

$$y_0=-m_1\ln 0,01832=4m_1=4.$$

Итак, правило проверки гипотез формулируем следующим образом. Если случайная величина Y приняла в результате опыта значение $0\le y\le 4,$ считаем, что математическое ожидание этой величины $m=m_1=1.$ В противном случае считаем, что $m=m_2=10.$

б) Вычислим вероятность ошибки второго рода

$$\beta=\int\limits_0^{y_0}W_2(y)dy=(1/m_2)\int\limits_0^{y_0}\exp(-y/m_2)\,dy=$$
$$1-\exp(-y_0/m_2)=1-\exp(-0,4)=0,3297.$$

в) Наконец, определим, возможно ли в принципе обеспечить $\alpha=\beta=0{,}1$. По формуле (5.1.2) вычисляем

$$I(1:2)=\int_0^\infty \frac{1}{m_1}\exp\left(-\frac{y}{m_1}\right)\left[\ln\frac{m_2}{m_1}-y\frac{m_2-m_1}{m_1m_2}\right]dy=$$

$$\ln\frac{m_2}{m_1}-m_1\frac{m_2-m_1}{m_1m_2}=\ln 10-0{,}9=1{,}4026.$$

Правая часть неравенства (5.1.9 а) равна

$$\beta\ln\frac{\beta}{1-\alpha}+(1-\beta)\ln\frac{1-\beta}{\alpha}=0{,}1\cdot\ln\frac{0{,}1}{0{,}9}+0{,}9\cdot\ln\frac{0{,}9}{0{,}1}=1{,}7578.$$

Требуемые значения $\alpha=\beta=0{,}1$ не удовлетворяют уже первому из неравенств (5.1.9), следовательно, не существует правило проверки гипотез, которое обеспечивало бы такие значения вероятностей ошибок первого и второго рода.

Пример 5.1.3. В условиях примера 1.3 вычислить для гипотез:

H_1 – сообщение на входе $x(t)=x_1(t)$;

H_2– сообщение на входе $x(t)=x_2(t)$,

где $x_1(t)$ и $x_2(t)$ – известные наблюдателю функции. Найти предельное выражение для $I(1:2)$ при условии, что верхняя граничная частота шума $F_в$ стремится к бесконечности, а полоса сигнала $x(t)$ остается неизменной.

Решение. Совместные плотности вероятности n отсчетов квантованного сигнала при каждой из гипотез равны (см. пример 1.3):

$$W_1\left[y^{(1)},\dots,y^{(n)}\right]=\prod_{k=1}^{n}\frac{1}{\sqrt{2\pi N_0F_в}}\cdot\exp\left[-\frac{1}{2N_0F_в}(y^{(k)}-x_1^{(k)})^2\right],$$

$$W_2\left[y^{(1)},\dots,y^{(n)}\right]=\prod_{k=1}^{n}\frac{1}{\sqrt{2\pi N_0F_в}}\cdot\exp\left[-\frac{1}{2N_0F_в}(y^{(k)}-x_2^{(k)})^2\right],$$

где $x_1^{(k)}=x_1(t_k)$, $x_2^{(k)}=x_2(t_k)$.

Случайные величины $Y^{(1)},\dots,Y^{(n)}$ независимы при каждой из гипотез, следовательно, для вычисления $I(1:2)$ можно воспользоваться свойством аддитивности информации

$$I(1:2)=\sum_{k=1}^{n} I_k(1:2),$$

где $I_k(1:2)=\int\limits_{-\infty}^{\infty} W\left(y^{(k)}/x_1^{(k)}\right)\ln\dfrac{W\left(y^{(k)}/x_1^{(k)}\right)}{W\left(y^{(k)}/x_2^{(k)}\right)}dy^{(k)}.$

В примере 5.1.1 было установлено, что

$$I_k(1:2)=\frac{(m_1-m_2)}{2\sigma^2}=\frac{\left[x_1^{(k)}-x_2^{(k)}\right]^2}{2N_0F_{\text{в}}},$$

и окончательно получаем

$$I(1:2)=\frac{1}{2N_0F_{\text{в}}}\sum_{k=1}^{n}\left[x_1^{(k)}-x_2^{(k)}\right]^2.$$

Вычислим предельное значение средней информации для различения в пользу H$_1$ против H$_2$ при $F_{\text{в}}\to\infty$ и $\Delta t\to 0$

$$I^*(1:2)=\lim_{\Delta t\to 0}\left\{\sum_{k=1}^{n}\frac{1}{2N_0F_{\text{в}}}\left[x_1^{(k)}-x_2^{(k)}\right]^2\right\}.$$

Внесем множитель $\Delta t=1/(2F_{\text{в}})$ под знак суммы и заметим, что ее пределом является определенный интеграл, вычисленный на интервале $(0,T)$

$$I^*(1:2)=\frac{1}{N_0}\lim_{\Delta t\to 0}\left\{\sum_{k=1}^{n}\left[x_1(t_k)-x_2(t_k)\right]^2\Delta t\right\}=$$
$$\frac{1}{N_0}\int\limits_0^T\left[x_1(t)-x_2(t)\right]^2dt.$$

Таким образом, величина средней информации для различения двух сообщений известной формы на фоне аддитивного белого шума зависит только от интенсивности шума и от квадрата расстояния между этими двумя сообщениями (сравните с результатами примера 5.1.1).

5.2 Информация по Фишеру

Определение. Свойства. Пусть сообщение X есть непрерывная случайная величина, а сигнал $\mathbf{y} = y^{(1)},..., y^{(n)}$ характеризуется условной плотностью вероятности $W(y^{(1)},..., y^{(n)} / x)$. Рассмотрим две гипотезы:

H_1 – передано сообщение x_1, то есть $W_1 = W(y^{(1)},..., y^{(n)} / x_1)$,

H_2 – передано сообщение x_2, то есть $W_2 = W(y^{(1)},..., y^{(n)} / x_2)$.

Среднюю информацию в пользу H_1, вычисленную по формуле (5.1.2)

$$I(1:2) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} W(y^{(1)},..., y^{(n)} / x_1) \cdot \log \frac{W(y^{(1)},..., y^{(n)} / x_1)}{W(y^{(1)},..., y^{(n)} / x_2)} dy^{(1)},..., dy^{(n)}, \tag{5.2.1}$$

можно формально рассматривать как функцию аргумента x_2, зависящую от параметра x_1. На рис.5.2 в качестве примера приприведен возможный вид этой функции при различных значениях x_1.

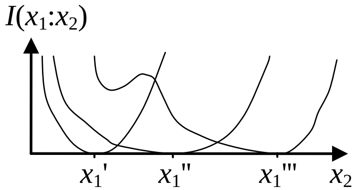


Рис. 5.2

Если существует вторая производная этой функции по x_2 в точке $x_2 = x_1 = x$,

$$I_F(x) = \frac{\partial^2 I(x_1 : x_2)}{\partial x_2^2} = -\int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} W(y^{[1]},..., y^{[n]} / x) \cdot \frac{\partial^2 \ln W(y^{[1]},..., y^{[n]} / x)}{\partial x^2} \cdot dy^{[1]}...dy^{[n]}. \tag{5.2.2}$$

то ее величина называется информацией по Фишеру о параметре x, содержащейся в сигнале **y.**

То же значение информации можно получить в результате вычисления по формуле

$$I_F(x) = \int\limits_{-\infty}^{+\infty} \ldots \int\limits_{-\infty}^{+\infty} W(y^{[1]}, \ldots, y^{[n]} / x) \cdot$$
$$\left[\frac{\partial \ln W(y^{[1]}, \ldots, y^{[n]} / x)}{\partial x} \right]^2 dy^{[1]} \ldots dy^{[n]}. \tag{5.2.3}$$

Выбор одной из двух формул определяется только удобством вычислений. Информация Фишера показывает, насколько быстро возрастает функция $I(x_1{:}x_2)$ при увеличении $\Delta x = |x_1 - x_2|$ и, следовательно, является мерой различимости близких значений x_1 и x_2.

Информация Фишера, как и информация Кульбака, обладает свойствами выпуклости и аддитивности. Единицы измерения информации Фишера – это единицы измерения величины x^{-2}.

Неравенство Рао-Крамера. Это – главный результат теории информации по Фишеру.

Рассмотрим задачу оценки непрерывного сообщения x. Правило оценивания сводится к тому, что каждому значению выборки $\mathbf{y} = y^{(1)}, \ldots, y^{(n)}$ ставится в соответствие некоторое значение оцениваемого параметра x. Таким образом, оценка $\hat{x}(y^{(1)}, \ldots, y^{(n)})$ является функцией выборочных данных и поэтому также является случайной величиной. Обычно вычисляют следующие числовые характеристики оценки:

1) Математическое ожидание при условии, что истинное значение сообщения равно x,

$$m(x) = \int\limits_{-\infty}^{+\infty} \ldots \int\limits_{-\infty}^{+\infty} \hat{x}(y^{[1]}, \ldots, y^{[n]}) \cdot W(y^{[1]}, \ldots, y^{[n]} / x) \cdot dy^{[1]} \ldots dy^{[n]}. \tag{5.2.4}$$

2) Смещение

$$b(x) = m(x) - x, \tag{5.2.5}$$

т.е. систематическую ошибку, сопутствующую выбранному правилу оценивания.

3) Дисперсию, вычисляемую также при условии, что истинное значение сообщения равно x,

$$D(x)=\int\limits_{-\infty}^{+\infty}\ldots\int\limits_{-\infty}^{+\infty}\left[\hat{x}(y^{[1]},\ldots,y^{[n]})-m(x)\right]^2\cdot \tag{5.2.6}$$
$$W(y^{[1]},\ldots,y^{[n]}/x)\cdot dy^{[1]}\ldots dy^{[n]}.$$

Дисперсия оценки является основной количественной мерой точности оценивания.

Пусть функция правдоподобия $W(y^{(1)},\ldots,y^{(n)}/x)$ дифференцируема по параметру *x*, информация Фишера (5.2.2) существует и не равна нулю для всех значений параметра в окрестности точки *x*, тогда дисперсия и смещение любой оценки связаны с информацией Фишера неравенством Рао-Крамера

$$D(x)\geq\frac{\left[1+\dfrac{db(x)}{dx}\right]^2}{I_F(x)}. \tag{5.2.7}$$

Для несмещенных оценок ($b(x)=0$) или для оценок с постоянным, не зависящим от *x* смещением ($b(x)=c$), числитель в (5.2.7) равен единице и тогда

$$D(x)\geq\frac{1}{I_F(x)}. \tag{5.2.8}$$

Таким образом, информация Фишера является количественной мерой предельной, потенциальной точности оценивания непрерывного сообщения *x*, так как дисперсия несмещенной оценки не может быть меньше величины, обратной информации Фишера.

Неравенства (5.2.7) и (5.2.8) обращаются в равенства тогда и только тогда, когда одновременно выполняются два условия:

1) Функция правдоподобия выборки может быть представлена в виде

$$W(y^{[1]},\ldots,y^{[n]}/x)=h\left(y^{[1]},\ldots,y^{[n]}\right)f\left[\hat{x}\left(y^{[1]},\ldots,y^{[n]}\right),x\right], \tag{5.2.9}$$

где $h\left(y^{[1]},\ldots,y^{[n]}\right)$ – некоторая функция выборки **y**, не зависящая от *x* и $\hat{x}$,

$f\left(\hat{x},x\right)$ – функция, зависящая только от *x* и $\hat{x}\left(y^{[1]},\ldots,y^{[n]}\right)$.

Оценка $\hat{x}\left(y^{[1]},...,y^{[n]}\right)$, удовлетворяющая условию (5.2.9), называется достаточной, поскольку она сохраняет всю информацию о x, содержащуюся в самой выборке.

2) Функция правдоподобия выборки такова, что для любого x выполняется соотношение

$$\frac{\partial \ln W(y^{[1]},...,y^{[n]}/x)}{\partial x} = k(x)\left[\hat{x}\left(y^{[1]},...,y^{[n]}\right) - m(x)\right], \qquad (5.2.10)$$

где $k(x)$ – некоторая функция x.

Оценка, удовлетворяющая этому уравнению, называется эффективной, а семейство распределений, задаваемых уравнением (5.2.10) при различных значениях x, называется экспоненциальным семейством. Легко убедиться, что эффективная оценка всегда достаточна, но обратное утверждение неверно.

Среди всех оценок с заданным смещением именно эффективные оценки обладают минимальной дисперсией. К сожалению, эффективная оценка существует далеко не во всех случаях, и тогда потенциальная точность оценивания сообщения недостижима.

Оценка максимального правдоподобия. Метод максимального правдоподобия широко используется на практике. В качестве оценки $\hat{x}\left(y^{[1]},...,y^{[n]}\right)$ выбирается такое значение x, при котором функция правдоподобия $W(y^{[1]},...,y^{[n]}/x)$ достигает наибольшего значения. Это значит, что в качестве оценки максимального правдоподобия выбирается решение уравнения правдоподобия

$$\frac{\partial W(y^{[1]},...,y^{[n]}/x)}{\partial x} = 0 \quad \text{или} \quad \frac{\partial \ln W(y^{[1]},...,y^{[n]}/x)}{\partial x} = 0. \qquad (5.2.11)$$

Доказано, что если эффективная оценка существует, то она может быть реализована методом максимального правдоподобия.

РЕШЕНИЕ ТИПОВЫХ ПРИМЕРОВ

Пример 5.2.1. Случайная величина Y имеет экспоненциальное распределение

$$W(y/x)=\frac{1}{x}\exp(-\frac{y}{x}),\qquad y\ge 0\,.$$

а) Найти максимально правдоподобную оценку математического ожидания m этой случайной величины.

б) Найти статистические характеристики (математическое ожидание и дисперсию) этой оценки.

в) Определить, является ли оценка достаточной.

г) Определить, является ли оценка эффективной.

д) Найти информацию Фишера и по неравенству Рао-Крамера проверить сделанное заключение об эффективности оценки.

Решение. Запишем уравнение правдоподобия

$$\frac{\partial \ln W(y/x)}{\partial x}=\frac{\partial}{\partial x}\left(-\ln x-\frac{y}{x}\right)=-\frac{1}{x}+y\frac{1}{x^2}=\frac{1}{x^2}(y-x)=0.$$

Отсюда $\hat{x}(y)=y$, т.е. максимально правдоподобная оценка математического ожидания равна наблюдаемому выборочному значению y.

По формуле (5.2.4) находим математическое ожидание оценки $m(x)=\int\limits_0^\infty y\frac{1}{x}\exp\left(-\frac{y}{x}\right)dy=x$.

Таким образом, оценка $\hat{x}(y)$ является несмещенной.

Дисперсию оценки вычисляем по формуле (5.2.5)

$$D(x)=\int\limits_0^\infty [y-x]^2\frac{1}{x}\exp\left(-\frac{y}{x}\right)dy=x^2\,.$$

Для того чтобы определить, является ли оценка достаточной, сравним функцию правдоподобия с выражением (5.2.9).

Видим, что $h(y)$= 1, а функция $f[\hat{x}(y),x]=\frac{1}{x}\exp(-\frac{y}{x})$ зависит только от x и $\hat{x}(y)=y$. Функция правдоподобия выборки может быть представлена в форме (5.2.9), и оценка $\hat{x}(y)=y$ является достаточной.

Полученное ранее выражение $\dfrac{\partial \ln W(y/x)}{\partial x} = \dfrac{1}{x^2}(y-x)$ сравним с выражением (5.2.10). Полагая $k(x) = 1/x^2$ и подставляя полученные нами выражения $\hat{x}(y) = y$ и $m(x)$=x, видим, что производная от логарифма функции правдоподобия может быть представлена в виде (5.2.10), и, следовательно, оценка максимального правдоподобия $\hat{x}(y) = y$ является эффективной.

Информацию Фишера находим по формуле (5.2.3)

$$I_F(x) = \int_0^\infty \frac{1}{x}\exp\left(-\frac{y}{x}\right)\left[\frac{1}{x^2}(y-x)\right]^2 dy = \frac{2}{x^2} - \frac{2}{x^2} + \frac{1}{x^2} = \frac{1}{x^2}.$$

Видим, что неравенство Рао-Крамера (5.2.8) обращается в равенство. Этого и следовало ожидать, так как мы уже показали, что оценка эффективна. Лучшей оценки, т.е. обладающей меньшей дисперсией при отсутствии систематической ошибки, не существует.

Пример 5.2.2. Система независимых случайных величин $Y^{(1)},...,Y^{(n)}$ распределена нормально с одинаковым среднеквадратическим отклонением σ и одинаковым, но неизвестным математическим ожиданием x.

а) Найти выражение для максимально правдоподобной оценки математического ожидания x.

б) Найти смещение и дисперсию этой оценки.

в) Определить потенциальную точность оценки величины математического ожидания x.

Решение. Функция правдоподобия выборки $y^{(1)},...,y^{(n)}$ равна

$$W(y^{(1)},...,y^{(n)}/x) = \prod_{k=1}^{n} \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2\sigma^2}(y^{(k)}-x)^2\right] =$$

$$\frac{1}{(2\pi\sigma^2)^{n/2}}\exp\left[-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(y^{(k)}-x)^2\right].$$

Запишем уравнение правдоподобия и преобразуем его к удобному виду

$$\frac{\partial \ln W(y^{(1)},....,y^{(n)}/x)}{\partial x}=\frac{\partial}{\partial x}\left[-\frac{n}{2}\ln(2\pi\sigma^2)-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(y^{(k)}-x)^2\right]=$$

$$\frac{1}{\sigma^2}\sum_{k=1}^{n}(y^{(k)}-x)=\frac{n}{\sigma^2}\left[\frac{1}{n}\sum_{k=1}^{n}y^{(k)}-x\right]=0.$$

Решением уравнения является значение

$$\hat{x}(y^{(1)},...,y^{(n)})=\frac{1}{n}\sum_{k=1}^{n}y^{(k)}.$$

Таким образом, максимально правдоподобная оценка математического ожидания x равна выборочному среднему.

По формуле (5.2.4) найдем математическое ожидание самой оценки

$$m(x)=\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}\left[\frac{1}{n}\sum_{k=1}^{n}y^{(k)}\right]\cdot$$

$$\prod_{\kappa=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}\exp[-\frac{1}{2\sigma^2}(y^{(k)}-x)^2]dy^{(1)}...dy^{(n)}=\frac{1}{n}\sum_{k=1}^{n}x=x,$$

так как математическое ожидание суммы случайных величин равно сумме математических ожиданий. Таким образом, смещение оценки равно нулю.

Дисперсия оценки может быть найдена непосредственно по формуле (5.2.6) либо же следующим, более простым способом. Величина $\sum_{k=1}^{n}y^{(k)}$ является суммой n независимых случайных величин, причем дисперсия каждого из слагаемых равна σ^2. Тогда дисперсия этой суммы равна $n\sigma^2$, и тогда дисперсия оценки $D(x)=\sigma^2/n$.

Вычисление информации Фишера, содержащейся в выборке, удобно произвести с использованием свойства аддитивности

$$I_F(x)=\sum_{k=1}^{n}I_F(x)_k,$$

где

$$I_F(x)_k = \int_{-\infty}^{\infty} W(y^{(k)}/x)\left[\frac{\partial \ln W(y^{(k)}/x)}{\partial x}\right]^2 dy^{(k)} = \frac{1}{\sigma^2}.$$

Неравенство Рао-Крамера обращается в равенство, и, следовательно, найденная нами оценка достаточна и эффективна.

Пример 5.2.3. На интервале времени (0,*T*) принимается сигнал $Y(t) = s(t-\tau) + Z(t)$,

где $s(t-\tau)$ – сигнал известной формы, но с неизвестной задержкой τ мкс,

$Z(t)$ – белый нормальный шум со спектральной плотностью N_0.

Найти потенциальную точность оценки временной задержки τ.

Решение. По определению (5.2.2) информация Фишера численно равна

$$I_F(\tau) = \frac{\partial^2 I(\tau_1;\tau_2)}{\partial \tau_2^2} \quad \text{при } \tau_1 = \tau_1 = \tau,$$

где $I(\tau_1 : \tau_2)$ – средняя информация для различения в пользу гипотезы H_1 (передан сигнал $s(t-\tau_1)$) против гипотезы H_2 (передан сигнал $s(t-\tau_2)$). Эту величину мы вычисляли в примере 5.1.3, где получили

$$I(\tau_1 : \tau_2) = \frac{1}{N_0}\int_0^T [s(t-\tau_1) - s(t-\tau_2)]^2 dt .$$

Дифференцируем эту функцию дважды по τ_2:

$$\frac{\partial I(\tau_1 : \tau_2)}{\partial \tau_2} = \frac{1}{N_0}\int_0^T 2[s(t-\tau_1) - s(t-\tau_2)] \cdot \left[-\frac{\partial s(t-\tau_2)}{\partial \tau_2}\right] dt,$$

$$\frac{\partial^2 I(\tau_1 : \tau_2)}{\partial \tau_2} = \frac{2}{N_0}\int_0^T \left\{\left[\frac{\partial s(t-\tau_2)}{\partial \tau_2}\right]^2 - \; s(t-\tau_1) - s(t-\tau_2) \; \frac{\partial^2 s(t-\tau_2)}{\partial \tau_2^2}\right\} dt \cdot$$

Полагаем далее $\tau_1 = \tau_2 = \tau$ и получаем

$$I_F(\tau)=\frac{2}{N_0}\int_0^T\left[\frac{\partial s(t-\tau)}{\partial\tau}\right]^2 dt \quad \frac{1}{\text{мкс}^2}.$$

Применяя неравенство Рао-Крамера, получим для дисперсии несмещенной оценки параметра

$$D(\tau)\geq\frac{N_0}{2\int_0^T\left[\frac{\partial s(t-\tau)}{\partial\tau}\right]^2 dt} \quad \text{мкс}^2.$$

ЗАДАЧИ

5.2.1. Случайная величина распределена нормально с известным средним значением и неизвестной дисперсией. Найти потенциальную точность несмещенной оценки дисперсии.

5.2.2. Решить задачу 5.2.1 в предположении, что имеется n независимых реализаций случайной величины.

5.2.3. Случайная величина Y распределена равномерно на интервале известной длины, но с неизвестным средним значением x. Показать, почему в задаче оценки параметра x не может быть использовано неравенство Крамера-Рао.

5.2.4. Найти потенциальную точность оценки направления на источник при помощи двухантенного фазового пеленгатора, если разность фаз на входе фазометра равна

$$\varphi=\frac{2\pi d}{\lambda}\alpha,$$

где d – база пеленгатора,

λ -длина волны,

α — угол между нормалью к плоскости антенн и направлением на источник.

Ошибка измерения разности фаз распределена нормально с нулевым средним и среднеквадратическим отклонением 2°.

5.2.5. Найти потенциальную точность оценки допплеровского сдвига частоты од узкополосного сигнала известной формы

$S(t) = S_0(t)\cos(\omega_0 + \omega_\partial)t$.

Шум на входе измерителя – белый аддитивный нормальный со спектральной плотностью N_0.

5.2.6. Найти потенциальную точность оценки амплитуды A прямоугольного импульса длительности τ , принимаемого на фоне аддитивного белого нормального шума со спектральной плотностью N_0.

5.2.7. Найти потенциальную точность оценки начальной фазы φ прямоугольного радиоимпульса длительности τ

$$S(t) = A\cos(\omega_0 t + \varphi), \qquad t_1 \le t \le t_1 + \tau , \qquad \omega_0 \gg \frac{1}{\tau}$$

принимаемого на фоне аддитивного белого нормального шума со спектральной плотностью N_0.

5.2.8. Производится однократное бросание монеты неправильной формы:

а) найти максимально правдоподобную оценку вероятности p выпадения герба;

б) вычислить смещениё и дисперсию такой оценки, предполагая, что истинное значение p = 0,4;

в) определять, является ли оценка эффективной.

5.2.9. Решить задачу 5.2.8. для случая, когда производится n независимых бросаний монеты.

5.2.10. Производится однократное бросание шестигранной игральной кости неправильной формы. Пусть $p_1,\dots, p_2$ - вероятности выпадения соответствующего количества очков. Если выпало очков, считают, что $\hat{p}_j = 1$, а $\hat{p}_k = 0\,(k \ne j)$. Является ли такая оценка эффективной?

5.2.11 . Система n случайных величин имеет нормальное распределение, причем вектор-столбец средних значений S(x) и корреляционная матрица R(x) известным образом зависят от неизвестного параметра x . Вычислить информацию Фишера о параметре x .

5.2.12. Решить задачу 5.2.11 для случаев, когда от x зависит только S(x) или только R(x).

5.2.13. В условиях задачи 5.1.11 вычислить информацию

Фишера о параметре x, считая его непрерывным.

5.2.14. Случайная величина Y распределена равномерно на интервале длиной 10 м, но ее среднее значение x неизвестно. Является ли оценка $\hat{x}(y) = y$ оценкой максимального правдоподобия?

6 ИНДИВИДУАЛЬНЫЕ ЗАДАНИЯ №1 и №2

Каждое из заданий представить в виде отдельной брошюры. Все расчеты сопровождать подробными пояснениями вплоть до подстановки численных значений. После завершения всех вычислений по каждой из задач результаты округляются до двух знаков после десятичной точки и приводятся в виде таблицы в том же порядке, как они даны в задании. Последнее (дополнительное) значение в таблице ответов – это сумма S всех приведенных в ней значений (контрольная сумма).

Номер варианта N студенты выбирают по формуле

$$\mathrm{N} = (50 \cdot P)\mathrm{div}100,$$

где P – значение двух последних цифр пароля студента (00,…,99),

div100 – целочисленное деление на 100.

6.1 Индивидуальное задание №1

1) Вероятностное описание символа

Для дискретной случайной величины X, принимающей одно из трех значений x_j с вероятностями p_j, записать ряд распределения и функцию распределения, привести соответствующие графики и найти следующие числовые характеристики:
математическое ожидание и СКО, математическое ожидание модуля X, $\mathrm{M}[X^2]$, $\mathrm{M}[p(X)]$, $\mathrm{M}[(p(X))^{-1}]$, энтропию $\mathrm{M}[-\log_2 p(X)]$.

N	x_j			p_j		
0	-7	-3	3	0.41	0.40	0.19
1	-5	-2	8	0.34	0.15	0.51
2	-8	2	10	0.44	0.11	0.45
3	-8	2	4	0.20	0.27	0.53
4	-9	-1	9	0.40	0.21	0.39

5	-5	3	5	0.26	0.44	0.30
6	-8	2	9	0.15	0.45	0.40
7	-10	-2	6	0.31	0.35	0.34
8	-4	0	7	0.21	0.29	0.50
9	-3	2	4	0.31	0.32	0.37
10	-8	3	8	0.44	0.10	0.46
11	-6	2	4	0.42	0.48	0.10
12	-8	3	4	0.49	0.31	0.20
13	-7	-2	4	0.71	0.21	0.08
14	-7	2	5	0.15	0.19	0.66
15	-5	0	9	0.13	0.19	0.68
16	-6	2	5	0.43	0.32	0.25
17	-8	-3	6	0.14	0.42	0.44
18	-8	-2	4	0.39	0.31	0.30
19	-4	-2	7	0.29	0.34	0.37
20	-7	2	5	0.38	0.26	0.36
21	-5	3	9	0.18	0.50	0.32
22	-5	-2	3	0.34	0.31	0.35
23	-5	3	8	0.24	0.41	0.35
24	-6	1	7	0.08	0.47	0.45
25	-8	1	5	0.03	0.64	0.33
26	-7	-1	6	0.32	0.40	0.28
27	-7	-3	7	0.44	0.31	0.25
28	-7	-1	3	0.27	0.12	0.61
29	-4	1	7	0.21	0.72	0.07
30	-6	-1	10	0.07	0.09	0.84
31	-4	1	6	0.44	0.21	0.35
32	-7	2	9	0.89	0.02	0.09
33	-6	-1	9	0.29	0.41	0.30
34	-8	2	5	0.58	0.03	0.39
35	-4	-3	4	0.23	0.60	0.17
36	-8	-3	7	0.35	0.32	0.33
37	-10	-1	7	0.31	0.18	0.51
38	-7	1	6	0.31	0.56	0.13

39	-4	2	9	0.11	0.51	0.38
40	-4	3	10	0.45	0.21	0.34
41	-9	0	3	0.29	0.65	0.06
42	-7	-2	5	0.29	0.36	0.35
43	-5	1	9	0.20	0.42	0.38
44	-4	0	4	0.41	0.21	0.38
45	-8	-1	5	0.08	0.63	0.29
46	-5	-3	6	0.09	0.16	0.75
47	-9	-2	5	0.33	0.36	0.31
48	-8	-1	8	0.15	0.07	0.78
49	-4	-3	7	0.40	0.17	0.43

Форма таблицы ответов:

N=28

m_x	σ_x	M[|X|]	M[X^2]	M[$p(X)$]
2.25	6.32	5.28	12.84	1.33

M[$(p(X))^{-1}$]	M[$-\log_2 p(X)$]	S
-0.33	-1.00	18.35

2) Вероятностное описание двух символов

Два символа X и Y имеют возможные значения x_1, x_2 и y_1, y_2 соответственно. Задана матрица совместных вероятностей с элементами $p_{j,k}=p(x_j,y_k)$. Найти:
ряд распределения случайной величины X, повторить то же при каждом из условий $Y=y_1$ и $Y=y_2$, а также m_x, σ_x, M[$-\log_2 p(X,Y)$].

N	x_1	x_2	$p_{1,1}$ $p_{2,1}$ $p_{1,2}$ $p_{2,2}$
0	4	8	0.11 0.36 0.31
1	3	5	0.22
2	0	7	0.17 0.28 0.45
3	4	8	0.10
4	2	5	0.25 0.27 0.22
5	1	7	0.26
6	1	7	0.30 0.29 0.14
7	2	8	0.27

8	2	9	0.21 0.20 0.25
9	3	5	0.34
10	4	7	0.28 0.27 0.23
11	0	7	0.22
12	0	9	0.36 0.10 0.28
13	2	8	0.26
14	4	8	0.24 0.21 0.17
15	4	8	0.38
16	3	8	0.33 0.11 0.19
17	0	7	0.37
18	4	9	0.18 0.27 0.20
19	4	9	0.35
20	3	9	0.01 0.18 0.38
21	1	9	0.43
22	4	8	0.24 0.27 0.34
23	3	6	0.15
24	4	9	0.38 0.20 0.37
25	0	9	0.05
26	2	7	0.65 0.13 0.10
27	1	10	0.12
28	3	6	0.21 0.26 0.24
29	1	9	0.29
30	2	6	0.44 0.10 0.18
31	1	10	0.28
32	0	6	0.59 0.25 0.09
33	0	8	0.07
34	2	5	0.44 0.02 0.25
35	3	8	0.29
36	3	5	0.24 0.18 0.29
37	1	9	0.29
38	0	9	0.21 0.32 0.12
39	4	6	0.35
40	3	9	0.22 0.21 0.27
41	4	10	0.30

42	2	7	0.41 0.36 0.18 0.05
43	2	5	0.02 0.45 0.11 0.42
44	4	5	0.14 0.32 0.39 0.15
45	3	6	0.57 0.08 0.11 0.24
46	0	5	0.28 0.28 0.31 0.13
47	3	10	0.12 0.42 0.41 0.05
48	2	6	0.10 0.21 0.41 0.28
49	3	5	0.20 0.38 0.33 0.09
			0.11 0.36 0.27 0.26
			0.33 0.07 0.17 0.43
			0.07 0.47 0.26 0.20
			0.22 0.33 0.10 0.35
			0.20 0.29 0.19 0.32
			0.22 0.34 0.06 0.38
			0.21 0.14 0.09 0.56
			0.11 0.26 0.20 0.43
			0.10 0.25 0.34 0.31

		0.45 0.09 0.03 0.43 0.29 0.25 0.21 0.25 0.29 0.16 0.37 0.18 0.27 0.21 0.27 0.25 0.04 0.39 0.29 0.28 0.02 0.38 0.31 0.29 0.19 0.13 0.32 0.36 0.24 0.17 0.21 0.38 0.30 0.23 0.11 0.36 0.14 0.37 0.09 0.40 0.11 0.15 0.47 0.27 0.34 0.05 0.24 0.37

Форма таблицы ответов:

N=28

$p(x_1)$	$p(x_2)$	$p(x_1/y_1)$	$p(x_2/y_1)$	$p(x_1/y_2)$	$p(x_2/y_2)$
-2.25	6.32	5.28	12.84	1.33	-0.33

m_x	σ_x	$M[-\log_2 p(X,Y)]$	S
1.00	4.24	-25.14	18.35

3) Аналого-цифровое преобразование непрерывных сигналов

m-разрядный АЦП рассчитан на входные напряжения в интервале (U_{min}, U_{max}) и проводит квантование во времени с шагом Δt=1. Записать последовательность, состоящую из 5 двоичных комбинаций на выходе АЦП, если на вход поступает сигнал $U(t)=u_0+u_1t+u_2t^2$, для $0 \le t \le 4$. Найти среднеквадратическую величину ошибки квантования по уровню для данного сигнала σ и затем ее теоретическое значение $\sigma_o=\Delta u/(\sqrt{12})$, где Δu – шаг квантования по уровню.

Полученные двоичные комбинации представить в форме целых неотрицательных десятичных чисел $Z_0, Z_1, \ldots, Z_4$, например: 00011010=26.

N	m	U_{min}	U_{max}	u_0	u_1	u_2
0	7	-0.13	53.09	-0.1	-2.5	3.9
1	9	-8.65	6.40	-2.5	9.8	-2.7
2	9	-7.59	6.40	-5.7	7.9	-1.3
3	4	-112.08	-4.14	-4.2	-6.4	-3.4
4	6	-236.27	-9.36	-9.5	-4.0	-9.5
5	6	4.62	72.38	6.9	2.5	3.4
6	5	-0.67	92.88	-0.5	3.8	4.8
7	4	-141.36	1.02	1.0	-2.4	-6.1
8	7	-210.31	-8.27	-8.4	-3.4	-8.5
9	6	0.00	73.49	0.0	5.7	3.1
10	9	-36.74	2.13	-1.6	7.1	-3.4
11	4	-11.58	18.37	-8.7	0.3	1.6
12	7	5.55	117.24	8.3	-5.6	8.1
13	4	1.34	110.04	2.0	2.6	6.0
14	7	-134.84	1.52	1.5	3.5	-7.3
15	4	-157.07	-9.06	-9.2	-2.0	-6.3
16	9	-42.46	-6.89	-9.5	5.2	-2.7
17	4	-75.34	2.23	1.4	5.9	-5.1
18	8	-23.03	-3.25	-3.3	-1.1	-0.6
19	9	-123.53	0.41	0.4	-0.9	-5.6
20	8	-131.78	5.48	5.4	-1.7	-6.1
21	7	-192.61	-0.69	-0.7	-9.6	-6.6

22	5	6.62	53.09	9.9	7.0	0.9
23	7	2.47	71.56	3.7	1.9	3.7
24	9	4.88	122.32	7.3	-1.3	7.4
25	7	-9.72	113.18	-7.3	-3.1	8.2
26	9	-33.41	-1.08	-1.1	-9.6	0.9
27	4	-63.63	7.21	6.2	5.7	-4.8
28	5	-173.84	-4.92	-5.0	5.0	-9.1
29	9	1.81	24.36	9.2	-9.9	3.4
30	9	-102.49	-1.77	-1.8	-8.0	-2.7
31	7	4.21	75.52	6.4	-5.8	5.7
32	4	-2.13	4.87	4.8	-2.0	0.1
33	6	-5.32	124.25	-4.0	7.6	6.0
34	6	-2.13	151.86	-1.6	1.4	9.1
35	4	-76.01	0.10	0.1	-9.1	-1.3
36	4	-5.99	141.20	-4.5	1.5	8.6
37	8	-2.66	98.67	-2.0	2.8	5.5
38	5	-108.75	-3.25	-3.3	-7.6	-3.0
39	9	-3.99	11.57	-3.0	8.0	-1.1
40	5	-152.68	4.57	4.5	-2.6	-6.8
41	4	-9.85	74.10	-7.4	-1.5	5.4
42	5	-168.38	-0.10	-0.1	-7.6	-6.0
43	4	6.09	47.00	9.1	7.3	0.5
44	7	-236.54	-1.28	-1.3	-8.5	-8.9
45	6	-8.39	84.15	-6.3	-5.7	7.0
46	4	-153.08	1.02	1.0	6.2	-8.8
47	4	-8.79	86.28	-6.6	-6.7	7.4
48	6	-213.64	-4.83	-4.9	-7.7	-7.8
49	4	6.62	133.49	9.9	6.8	5.9

Форма таблицы ответов:

N=28

Z_0	Z_1	Z_2	Z_3	Z_4
25	32	28	184	133

σ	σ_o	S

0.33	1.05	218.35

6.2 Индивидуальное задание №2

4) Нормальные случайные величины

Система случайных величин $X,У$ имеет нормальное распределение $W(x,y)$, которое характеризуется вектором-строкой математических ожиданий $\mathbf{a}=(m_x,m_y)$ и ковариационной матрицей $\mathbf{R}$. Найти: σ_x, σ_y, коэффициент ковариации r, значение условного СКО $\sigma_x(y_o)$, величину средней взаимной информации $I=\mathbf{M}\left[\log_2\frac{W(X/Y)}{W(X)}\right]$, $x_{mp}(y_o)$ – наиболее вероятное значение x при заданном y_o.

N	m_x	m_y	R_{11}	R_{22} y_o	R_{12}	
0	-9.97	-7.05	4.42	6.14	0.43	
1	-6.13	-7.17		-6.86		
2	1.70	3.86	5.82	0.82	-1.91	-
3	-2.99	-1.47	6.83			
4	6.46	9.33	6.32	6.57	0.31	
5	-6.52	-6.93		4.65		
6	4.21	6.43	5.09	1.14	-1.91	
7	-3.92	-6.17		0.42		
8	-8.17	6.34	6.99	2.35	3.11	
9	-7.05	-6.89		9.93		
10	9.77	4.64	1.98	9.21	-2.96	-
11	-7.62	-4.41	3.01			
12	-9.82	3.64	1.87	6.66	2.99	
13	0.63	4.44		11.12		
14	2.04	-7.54	4.63	4.98	-4.26	-
15	-6.68	6.69	2.06			
16	-0.98	0.34	1.07	5.02	-0.74	
17	-8.86	-1.48		7.02		
18	5.67	8.99	1.03	5.14	-0.46	-
19	0.40	0.99	4.31			
20	7.52	-0.57	9.32	6.91	7.02	

21	9.12	6.94		9.60	
22	0.79	-0.88	8.96	6.10	0.50
23	-0.76	9.66		-1.55	
24	7.24	4.78	2.35	0.16	0.23
25	5.59	-6.08		3.84	
26	9.94	6.79	4.17	1.10	-1.01
27	2.23	0.02		5.17	
28	-4.68	-9.45	6.32	8.65	0.37
29	6.80	1.45		-2.86	
30	-2.48	0.63	4.57	7.50	-1.16
31	3.54	6.86		7.08	
32	-9.82	3.15	6.02	3.86	0.81
33	-4.48	6.84		2.66	
34	1.76	-7.80	8.56	5.57	2.49
35	6.75	-3.72		-0.55	
36	-0.30	-4.28	6.29	9.56	-0.98
37	4.87	-7.19		12.97	
38	-0.84	6.69	5.70	1.85	-3.14
39	4.89	2.00		2.63	
40	1.98	-4.95	1.92	1.40	0.33
41	4.70	-9.97		0.47	
42	1.45	6.12	5.60	9.52	1.10
43	-6.97	-5.79		9.39	
44	-1.50	1.06	2.50	0.38	-0.53 -
45	0.34	-7.72	0.15		
46	5.03	5.04	6.09	0.65	-1.93
47	-6.62	0.87		11.19	
48	-0.16	-1.27	5.89	1.40	-2.35
49	4.00	3.92		5.71	
			5.00	8.66	4.58
				-4.99	
			7.43	7.24	4.47
				9.88	
			6.24	0.25	0.47

					0.65
		8.06	7.11	0.45	-9.19
		5.80	2.25	-2.53	2.23
		9.12	1.77	1.79	3.01
		7.30	3.48	-4.63	9.58
		6.71	3.74	1.42	4.67
		3.22	8.04	-4.13	11.97
		3.13	5.31	3.58	-6.01
		1.17	8.00	-1.35	-1.72
		8.53	1.53	3.41	-2.44
		1.63	4.08	-1.05	-4.20
		0.89	1.45	-0.29	7.17
		6.45	0.76	-1.22	3.19
		5.50	5.78	-2.87	-3.18
		4.15	5.53	3.78	-9.46
		4.71	3.18	3.05	6.84
		1.61	8.10	-2.97	-2.06
		7.41	2.91	-2.99	

		3.41 8.28 8.96 4.66 -7.24 8.75 7.46 -4.58 5.19 3.07 3.68 -3.27 2.05 1.36 2.36 -1.36 - 0.83 7.87 2.37 -1.60 4.13

Форма таблицы ответов:

N=28

σ_x	σ_y	r	$\sigma_x(y_o)$
25	32	28	184

I	$x_{mp}(y_o)$	S
133	0.33	218.35

5) Корректирующие коды

Строки производящей матрицы линейного блочного (n,3)-кода – это три n-разрядные комбинаций (младший разряд – справа), которые в двоичной форме представляют десятичные числа g_0, g_1, g_2. Найти: кодовое расстояние $d_{код}$, максимальные кратности гарантированно обнаруживаемых $q_о$ и исправляемых $q_и$ ошибок. Закодировать двоичную комбинацию, соответствующую десятичному числу *in*, затем двоичную комбинацию на выходе кодера представить в форме десятичного числа *out*.

Примечание: верхняя строка производящей матрицы g_0 соответствует младшему разряду комбинации на входе кодера.

N	n	*in*	g_0	g_1	g_2
0	11	1	793	1261	1689
1	11	6	823	1528	1613
2	11	3	909	1342	1888
3	8	4	99	143	223
4	8	2	103	179	255
5	10	7	407	641	883
6	11	5	641	1398	1856
7	11	2	796	1252	2045
8	9	2	150	298	467
9	8	6	107	159	243
10	8	3	88	180	249
11	9	2	221	378	430
12	9	2	186	367	389
13	10	7	298	541	798
14	10	1	443	683	975
15	11	2	616	1222	1833
16	10	1	353	542	779
17	12	5	1025	2484	3544

18	11	3	798	1496	1974
19	8	1	111	144	242
20	12	5	1558	2769	3821
21	9	5	230	306	390
22	8	4	112	137	234
23	9	1	230	381	421
24	8	2	70	137	211
25	12	7	1857	2967	3347
26	10	1	479	728	778
27	11	5	742	1481	1815
28	12	7	2047	2936	3222
29	10	3	505	695	933
30	11	5	575	1216	2004
31	10	6	370	570	808
32	10	7	449	566	965
33	11	7	948	1299	1981
34	8	4	87	130	217
35	9	3	153	272	498
36	11	3	798	1240	1978
37	12	1	1722	3029	3879
38	11	6	581	1034	1919
39	12	1	1810	2694	3770
40	10	5	324	677	835
41	11	2	692	1267	1742
42	8	1	102	155	216
43	11	1	954	1382	1844
44	12	5	1411	3007	4059
45	11	7	720	1279	1983
46	11	4	813	1167	1983
47	9	2	238	280	497
48	11	5	542	1155	1595
49	11	1	537	1196	1799

Форма таблицы ответов:

N=28

$d_{код}$	$q_о$	$q_и$	*out*	S
5	3	8	184	218

6) Линейные блочные коды

Двоичные комбинации, соответствующие пяти десятичным числам (n, in, g_0, g_1, g_2) из задачи 5, считать строками проверочной матрицы **H** кода (n,n-5).

Определить: способен ли этот код обнаружить любую однократную ошибку (d=1, если способен, d=0 в противном случае);

способен ли этот код исправить любую однократную ошибку (c=1, если способен, c=0 в противном случае).

Форма таблицы ответов:

N=28

d	c	S
1	1	2

7) Неравенство Хэмминга для линейного блочного кода	N	*n*	*k*	*p*
Требуется построить линейный блочный (n,k)-код. Определить теоретический предел для этого кода – найти максимальную кратность исправляемых ошибок $q_и$. Определить вероятность ошибочного декодирования кодовой комбинации $P_{ош}$, если ошибки в отдельных символах в канале передачи происходят с вероятностью p, а ошибки в разных символах независимы. В ответе для величины $P_{ош}$ оставить 6 знаков после десятичной точки.	0	32	19	0.129
	1	38	11	0.085
	2	34	14	0.196
	3	48	13	0.181
	4	45	11	0.192
	5	44	25	0.056
	6	46	23	0.084
	7	48	21	0.072
	8	48	23	0.038
	9	23	13	0.148
	10	34	10	0.141
	11	20	9	0.021
	12	33	18	0.129
	13	24	11	0.072
	14	26	12	0.154
	15	43	18	0.184
	16	39	19	0.129

	17	43	8	0.161
	18	45	22	0.118
	19	34	12	0.091
	20	22	10	0.038
	21	30	6	0.025
	22	32	20	0.067
	23	28	13	0.061
	24	42	12	0.175
	25	21	14	0.029
	26	30	8	0.030
	27	24	5	0.035
	28	23	6	0.193
	29	25	7	0.178
	30	27	15	0.093
	31	30	6	0.139
	32	39	7	0.190
	33	47	18	0.020
	34	23	12	0.029
	35	36	13	0.096
	36	46	12	0.129
	37	46	11	0.095
	38	31	18	0.132
	39	40	22	0.021
	40	28	16	0.091
	41	49	20	0.077
	42	37	16	0.174
	43	23	7	0.056
	44	38	9	0.104
	45	30	8	0.098
	46	34	11	0.095
	47	48	11	0.124
	48	49	14	0.081
	49	35	8	0.191

Форма таблицы ответов:

N=28

$q_и$	$P_{ош}$	S
5	0.000124	5.000124

7. ВОПРОСЫ ДЛЯ САМОПРОВЕРКИ ПО КУРСУ

1. В чем разница понятий “информация” и “сигнал”?
2. Приведите примеры радиоэлектронных устройств, предназначенных не для передачи информации.
3. Назовите два основных признака того, что сигнал не несет информации.
4. Почему для математического описания сигналов используют вероятностные модели?
5. Может ли детерминированный сигнал переносить информацию?
6. Какие случайные события (величины) называются независимыми?
7. Что нужно задать для полного вероятностного описания: символа? последовательности символов? последовательности отсчетов сигнала? непрерывной случайной функции?
8. Из каких соображений выбирается шаг квантования непрерывного сигнала:
по времени? по напряжению?
9. Опишите этапы аналого-цифрового преобразования непрерывного сигнала.
10. Опишите этапы цифро-аналогового преобразования.
11. Изобразите обобщенную модель системы передачи информации. Опишите функции кодера и декодера.
12. Назовите способы манипуляции гармонической несущей. Чем обусловлен выбор того или иного способа?
13. Каковы недостатки многопозиционных методов манипуляции гармонической несущей?
14. В чем отличие аддитивной помехи от мультипликативной? Приведите примеры каналов связи с такими помехами.

15. Что такое собственная информация и энтропия дискретной случайной величины?
16. Дайте определение взаимной информации переданного и принятого символов.
Как влияет на ее величину интенсивность помех в канале связи?
17. От чего зависит пропускная способность непрерывного канала связи с аддитивным белым шумом?
18. Что такое избыточность сигнала? В каких случаях она полезна, а когда нет?
19. Когда полезно применять кодирование с малой избыточностью?
20. Какой смысл вкладывают в понятия: “кодирование источника”? “канальное кодирование”?
21. Каково значение минимально-возможной средней длины кодовой комбинации?
22. Всегда ли удается закодировать сигнал так, чтобы избыточность на выходе кодера была нулевой?
23. Когда полезно кодировать блоки букв, а не отдельные буквы?
24. Какой способ разделения кодовых комбинаций применяется в кодах, обладающих малой избыточностью?
25. В чем заключается главный недостаток кодов Хафмана и Шеннона-Фано?
26. Откуда берется кодовая таблица, используемая при кодировании кодом Лемпела-Зива?
27. Чем определяется корректирующая способность кода? Поясните на примере.
28. Какие коды называются корректирующими?
29. Что значит “обнаружить ошибки” при декодировании кодовой комбинации?
30. Что значит “исправить ошибки” при декодировании кодовой комбинации?
31. Каков характерный признак, позволяющий от-

личить кодовую таблицу линейного блочного кода от кодовых таблиц других кодов?

32. Что такое проверочная матрица линейного блочного кода? Как она используется при обнаружении ошибок в принятой комбинации?

33. Каков характерный признак, позволяющий отличить кодовую таблицу циклического кода от кодовых таблиц других кодов?

34. Чему равно количество комбинации в кодовой таблице линейного блочного кода?

35. Почему в проверочной матрице не может быть нулевых: столбцов? строк?

36. Какой смысл имеют строки проверочной матрицы?

37. По каким признакам можно определить, что проверочная матрица принадлежит коду, способному исправить любую одиночную ошибку?

38. Чем обусловлена популярность циклических кодов? Из каких логических элементов состоят кодер и декодер?

39. В чем заключается фундаментальное свойство комбинаций циклического кода?

40. Поясните суть декодирования по минимуму расстояния.

41. Почему декодирование по минимуму расстояния применяется редко?

8. СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

Основная

1. Акулиничев Ю.П. Теория электрической связи. Часть 1. – Томск, ТМЦДО. – 127 с.

2. Колесник В.Д., Полтырев Г.Ш. Курс теории информации. - М.: Наука, 1982. - 416 с.

3. Галлагер Р. Теория информации и надежная связь. - М.: Сов. радио, 1974.

4. Зюко А.Г, Кловский Д.Д., Коржик В.И., Назаров М.В. Теория электрической связи, под ред. Д.Д. Кловского. - М.: Радио и связь, 1997. - 432 с.

5. Зюко А.Г. и др. Теория передачи сигналов. - М.: Связь, 1980 (и последующие издания), 288 с.

6. Тарасенко Ф.П. Введение в курс теории информации. -Томск: ТГУ, 1963.

Дополнительная

7. Харкевич А..А. Борьба с помехами. - М.: Наука, 1965.

8. Хэмминг Р.В. Теория кодирования и теория информации. - М: Радио и связь, 1983. - 176 с.

9. Кульбак С. Теория информации и статистика. - М.: Наука, 1967.

10. Клюев Л.Л. Теория электрической связи.- Мн.: Дизайн ПРО, 1998.- 336 с.

11. Кловский Д.Д., Шилкин В.А. Теория электрической связи. Сб. задач и упражнений. Учебное пособие для втузов. - М.: Радио и связь, 1990. - 280 с.

12. Орлов В.А., Филиппов Л.И. Теория информации в упражнениях и задачах. Учебное пособие для втузов. М.: Высшая школа, 1976. - 136 с.

13. Цимбал В.П. Задачник по теории информации и кодированию. - Киев: Вища школа, 1976. – 276 с.

14. Лосев Ю.И., Плотников Н.Д. Основы теории передачи

данных. Сборник задач. - Киев: Вища школа, 1977. - 160 с.
15. Акулиничев Ю.П., Дроздова В.И. Сборник задач по теории информации. - Томск: ТГУ, 1976. - 146 с.

ПРИЛОЖЕНИЯ

Приложение 1

Таблица значений вспомогательной функции $\eta(p)=p\log_2 p$

p	$h(p)$	p	$h(p)$	p	$h(p)$	p	$h(p)$
0,00	0,0000						
0,01	0,0664	0,26	0,5053	0,51	0,4954	0,76	0,3009
0,02	0,1129	0,27	0,5100	0,52	0,4906	0,77	0,2903
0,03	0,1518	0,28	0,5142	0,53	0,4854	0,78	0,2796
0,04	0,1858	0,29	0,5179	0,54	0,4800	0,79	0,2687
0,05	0,2161	0,30	0,5211	0,55	0,4744	0,80	0,2575
0,06	0,2435	0,31	0,5238	0,56	0,4684	0,81	0,2462
0,07	0,2686	0,32	0,5260	0,57	0,4623	0,82	0,2348
0,08	0,2915	0,33	0,5278	0,58	0,4558	0,83	0,2231
0,09	0,3127	0,34	0,5292	0,59	0,4491	0,84	0,2113
0,10	0,3322	0,35	0,5301	0,60	0,4422	0,85	0,1993
0,11	0,3503	0,36	0,5306	0,61	0,4350	0,86	0,1871
0,12	0,3671	0,37	0,5307	0,62	0,4276	0,87	0,1748
0,13	0,3826	0,38	0,5305	0,63	0,4199	0,88	0,1623
0,14	0,3971	0,39	0,5298	0,64	0,4121	0,89	0,1496
0,15	0,4105	0,40	0,5288	0,65	0,4040	0,90	0,1368
0,16	0,4230	0,41	0,5274	0,66	0,3956	0,91	0,1238
0,17	0,4346	0,42	0,5256	0,67	0,3871	0,92	0,1107
0,18	0,4453	0,43	0,5236	0,68	0,3783	0,93	0,0974
0,19	0,4552	0,44	0,5211	0,69	0,3694	0,94	0,0839
0,20	0,4644	0,45	0,5184	0,70	0,3602	0,95	0,0703
0,21	0,4728	0,46	0,5153	0,71	0,3508	0,96	0,0565
0,22	0,4806	0,47	0,5120	0,72	0,3412	0,97	0,0426
0,23	0,4877	0,48	0,5083	0,73	0,3314	0,98	0,0286
0,24	0,4941	0,49	0,5043	0,74	0,3215	0,99	0,0144
0,25	0,5000	0,50	0,5000	0,75	0,3113	1,00	0,0000

Приложение 2

Значения двоичных логарифмов целых чисел от 1 до 100

n	$\log_2 n$	n	$\log_2 n$	n	$\log_2 n$	n	$\log_2 n$
1	0,00000	26	4,70044	51	5,67243	76	6,24793
2	1,00000	27	4,75489	52	5,70044	77	6,26679
3	1,58496	28	4,80735	53	5,72792	78	6,28540
4	2,00000	29	4,85798	54	5,75489	79	6,30378
5	2,32193	30	4,90689	55	5,78136	80	6,32193
6	2,58496	31	4,95420	56	5,80735	81	6,33985
7	2,80735	32	5,00000	57	5,83289	82	6,35755
8	3,00000	33	5,04439	58	5,85798	83	6,37504
9	3,16993	34	5,08746	59	5,88264	84	6,39232
10	3,32193	35	5,12928	60	5,90689	85	6,40939
11	3,45943	36	5,16993	61	5,93074	86	6,42626
12	3,58496	37	5,20945	62	5,95420	87	6,44294
13	3,70044	38	5,24793	63	5,97728	88	6,45943
14	3,80735	39	5,28540	64	6,00000	89	6,47573
15	3,90689	40	5,32193	65	6,02237	90	6,49185
16	4,00000	41	5,35755	66	6,04439	91	6,50779
17	4,08746	42	5,39232	67	6,06609	92	6,52356
18	4,16993	43	5,42626	68	6,08746	93	6,53916
19	4,24793	44	5,45943	69	6,10852	94	6,55459
20	4,32193	45	5,49185	70	6,12928	95	6,56986
21	4,39232	46	5,52356	71	6,14975	96	6,58496
22	4,45943	47	5,55459	72	6,16993	97	6,59991
23	4,52356	48	5,58496	73	6,18982	98	6,61471
24	4,58496	49	5,61471	74	6,20945	99	6,62936
25	4,64386	50	5,64386	75	6,22882	100	6,64386

$$\log_2 10^k = 3{,}32193k.$$