МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ

Е.Г. Лебедько

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Учебное пособие

Санкт-Петербург

2010

Лебедько Е.Г. Математические основы передачи информации, ч. 5: учеб. пособие для вузов. СПб: СПбГУИТМО, 2010. 93 с.

В пятой части настоящего учебного пособия излагаются основы теории информации. Все теоретические положения иллюстрируются примерами.

Учебное пособие предназначено для студентов, обучающихся по направлениям подготовки бакалавров и магистров 200200 «Оптотехника».

Рекомендовано Учебно-методическим объединением вузов Российской Федерации по образованию в области приборостроения и оптотехники для студентов высших учебных заведений, обучающихся по направлению подготовки 200200 – Оптотехника и специальности 200203 – Оптико-электронные приборы и системы, протокол №430 от 12.03.10

В 2009 году Университет стал победителем многоэтапного конкурса, в результате которого определены 12 ведущих университетов России, которым присвоена категория «Национальный исследовательский университет». Министерством образования и науки Российской Федерации была утверждена Программа развития государственного образовательного учреждения высшего профессионального образования «Санкт-Петербургский государственный университет информационных технологий, механики и оптики» на 2009–2018 годы.

Содержание

Часть 5. Основы теории информации

Понятие «*информация*» можно истолковать как некоторую совокупность сведений о тех или иных событиях, явлениях или фактах. Информация имеет количественную и качественную стороны. При этом количественная и качественная меры информации не должны противоречить нашим интуитивным представлениям и охватывать то общее, что присуще всему многообразию различной информации.

Первая попытка определения количественной меры была предпринята Р.В.Хартли в 1928 году. [1]. Однако это определение оказалось недостаточно универсальным. Основные соотношения, определяющие количественную меру информации, и теоремы были сформулированы К. Шенноном и опубликованы в 1948 – 1949 годах. Математический аппарат современной теории информации был разработан советским академиком А.Н.Колмогоровым к 1947 году [2]. В трудах В.А.Котельникова, А.Н.Колмогорова, А.А. Харкевича, Р.Л. Стратоновича, Р.Л. Добродушина, А.Файнстейна, Р.Фано и др. теория информации получила свое дальнейшее развитие. Важный вклад в определение качественной меры информации внес Р.Л. Стратонович, связав шенноновскую теорию информации с теорией решающих функций А.Вальда.

В теории информации рассматриваются:

- общая количественная мера информации, не зависящая от природы объектов;
- общие законы передачи, обработки и хранения информации;
- ценность и достоверность информации;
- общие закономерности, определяющие влияние тех или иных факторов на преобразование, скорость передачи, потерю и ценность информации.

5.1 Количественная мера информации

5.1.1 Количественная мера информации дискретных сообщений

Материальная форма воплощения информации называется *сообщением*. Сообщения могут быть представлены в виде показаний приборов, печатного текста, картины, произносимой речи, наблюдаемого изображения и т.д. Сообщения можно разделить на дискретные и непрерывные как по состояниям, так и по времени.

Информация добывается в результате опыта. Таким опытом является сообщение. До опыта можно только предполагать о состоянии элементов сообщения, т.е. до опыта имеет место большая или меньшая неопределенность в интересующей ситуации. После опыта ситуация становится более определенной и существующая до опыта неопределенность уменьшится.

Уменьшение неопределенности в результате опыта может быть принято за наиболее общую меру количества получаемой информации.

Меру количества информации, получаемой в результате того или иного опыта, можно было бы установить как отношение равновозможных ответов до опыта $\left(K\right)$ и после опыта $\left(K_0\right)$. Удобно использовать понятие о вероятности появления состояний. Тогда при равновозможных состояниях априорная (до опыта) вероятность появления состояния x_i равна $P\left(x_i\right) = 1/K$, а апостериорная (после опыта) вероятность $P_0 = 1/K_0$. В этом случае количество информации, которое получаем о состоянии x_i, будет равно $\dfrac{K}{K_0} = \dfrac{P_0\left(x_i\right)}{P\left(x_i\right)}$.

Однако выбор этого отношения за меру количества информации связан с рядом неудобств. Например, при заведомо известном состоянии x_i, сообщение о котором не несет никакой информации, это соотношение будет равно единице. Это соотношение неудобно и тем, что при сложении количества информации нескольких независимых источников не выполняется условие аддитивности.

Более удобной является логарифмическая мера количества информации, которая определяется соотношением

$$I\left(x_i\right) = k\log_a \frac{P_0\left(x_i\right)}{P\left(x_i\right)}. \tag{5.1}$$

Выбор коэффициента k и основания логарифма a не имеет принципиального значения. Обычно выбирают $k = 1$, а основание логарифма $a = 2$. В этом случае единица количества информации называется двоичной (или бит).

Рассмотрим некоторое конечное множество X состояний $x_1, x_2, ..., x_m$. При этом состояния независимы и несовместимы, а априорные вероятности их соответственно равны $P\left(x_1\right), P\left(x_2\right), ..., P\left(x_m\right)$ и выполняется условие

$P\left(x_1\right) + P\left(x_2\right) + \cdots + P\left(x_m\right) = 1$.

Последнее означает, что хоть одно из состояний будет иметь место.

Множество с известным распределением вероятностей его состояний называют *ансамблем.*

Рассматриваемый ансамбль состояний можно представить схемой

$$X = \begin{bmatrix} x_1 & x_2 & \dots & x_m \\ P(x_1) & P(x_2) & \dots & P(x_m) \end{bmatrix}. \tag{5.2}$$

Ансамбль можно рассматривать как модель сообщения. Достоверное сообщение о том, что из всех состояний X появляется x_i, несет в себе количество информации, равное

$$\log\left(1 \Big/ P(x_i) \right) = -\log P(x_i). \tag{5.3}$$

Как видно, чем меньше априорная вероятность появления данного состояния, тем большее количество информации содержится в этом сообщении.

Важнейшей характеристикой в теории информации является среднее количество информации, приходящееся на один элемент сообщения, являющейся мерой неопределенности, и определяется зависимостью

$$H(X) = m_1\left\{ I(x_i) \right\} = -\sum_{i=1}^{m} P(x_i) \log P(x_i). \tag{5.4}$$

Величина $H(X)$ носит название энтропии ансамбля X.

Приписывая определенное значение энтропии каждому состоянию x_i и, учитывая, что x_i - случайная величина, $-\log P(x_i) = H(x_i)$ называют *случайной энтропией* [3],

Таким образом, количество информации о состоянии x_i согласно формуле (5.1) будет равно

$$I(x_i) = \log P_0(x_i) - \log P(x_i) = H(x_i) - H_0(x_i), \tag{5.5}$$

где $H(x_i)$ и $H_0(x_i)$ - соответственно априорная и апостериорная случайные энтропии.

Среднее количество информации в ансамбле будет определяться соотношением

$$\begin{gathered} I(X) = H(X) - H_0(X) = \\ = \sum_{i=1}^{m} P_0(x_i) \log P_0(x_i) - \sum_{i=1}^{m} P(x_i) \log P(x_i) \end{gathered} \tag{5.6}$$

Как видно из формул (5.5) и (5.6) количество информации соответствует величине снятой неопределенности в результате проведения опыта.

В приведенных и последующих формулах знаком $\log$ обозначается логарифм с основанием два.

5.1.2 Основные свойства энтропии

1.Энтропия есть величина вещественная, ограниченная и неотрицательная: $H(X) \geq 0$

Рассмотрим одно слагаемое суммы в формуле (5.4) –

$$-P(x_i)\log P(x_i). \qquad (5.7)$$

Так как вероятности $P(x_i)$ являются неотрицательными величинами, заключенными в промежутке $\left[0 \leq P(x_i) \leq 1\right]$, то при изменении $P(x_i)$ от нуля до единицы положительное вещественное слагаемое (5.7) возрастает от нуля, достигает максимума, а затем снова уменьшается до нуля.

Устремим $P(x_i)$ к нулю и найдем предел величины (5.7)

$$\lim_{P(x_i)\to 0}\left[-P(x_i)\log P(x_i)\right] = \lim_{P(x_i)\to 0}\left[\frac{\log\left(1/P(x_i)\right)}{1/P(x_i)}\right].$$

Обозначим $1/P(x_i) = \alpha$ и определим неопределенность по правилу Лопиталя:

$$\lim_{\alpha\to\infty}\frac{\log\alpha}{\alpha} = \lim_{\alpha\to\infty}\frac{\left(1/\alpha\right)\log e}{1} = 0.$$

Таким образом, величина (1.7) обращается в нуль при $P(x_i) = 0$.

При $P(x_i) = 1$:

$$\log P(x_i) = 0 \text{ и } P(x_i)\log P(x_i) = 0.$$

Максимальное значение слагаемого (5.7) определяется из условия

$$\frac{d}{dP(x_i)}\left[-P(x_i)\log P(x_i)\right] = -\log P(x_i) - \log e = 0.$$

Отсюда $P(x_i) = 1/e$.

Следовательно, максимальное значение слагаемого (5.7) равно 0,531.

Так как слагаемое (5.7) вещественно, неотрицательно и ограничено, то при любом конечном m энтропия $H(X)$ является величиной вещественной, неотрицательной и ограниченной.

2. Энтропия заведомо известных сообщений минимальна и равна нулю.

Так как сообщение заведомо известно, то вероятность появления того или иного состояния равна либо нулю, либо единице. Следовательно, как видно из доказательства предыдущего свойства, все слагаемые (5.7) обращаются в ноль и энтропия равна нулю.

3. Энтропия дискретных сообщений максимальна при равновероятном распределении состояний элементов.

Задаваясь дополнительным условием

$$\sum_{i=1}^{m} P(x_i) = 1, \tag{5.8}$$

найдем значение вероятности $P(x_i)$, обеспечивающей максимум энтропии $H(X)$, решая вариационную задачу методом неопределенных множителей Лагранжа.

С учетом (5.4) и (5.8) составим функцию

$$f = -P(x_i)\log P(x_i) - \lambda P(x_i)$$

и решим уравнение

$$\frac{df}{dP(x_i)} = -\log P(x_i) - \log e - \lambda = 0.$$

Отсюда $\log P(x_i) = -\log e - \lambda$, $(i = 1, 2, ..., m)$.

Как видно $\log P(x_i)$ и, следовательно, $P(x_i)$ не зависят от номера i, что может быть только в том случае, когда все $P(x_i)$ равны между собой.

Следовательно, энтропия максимальна при равновероятных состояниях дискретных сообщений:

$$P(x_1) = P(x_2) = \cdots = P(x_m) = \frac{1}{m}.$$

Величина же максимальной энтропии будет равна

$$H(X)=H_{\max}(X)=-\sum_{i=1}^{m}\frac{1}{m}\log\frac{1}{m}=\log m.$$

Таким образом, для равновероятного распределения состояний энтропия вычисляется по формуле

$$H(X)=\log m=H(x_i) \tag{5.9}$$

и равна случайной энтропии.

4. Энтропия бинарных сообщений может изменяться от нуля до единицы.

Бинарные сообщения характеризуются двумя возможными состояниями и их энтропия равна

$$H(X)=-\sum_{i=1}^{2}P(x_i)\log P(x_i)=-P\log P-(1-P)\log(1-P)$$

Из формулы видно, что максимальное значение энтропии $\left[H(X)=1\right]$ имеет место при $P=0,5$.

5. Энтропия объединения нескольких независимых ансамблей равна сумме энтропий этих ансамблей – свойство аддитивности энтропий

Под объединением двух ансамблей X и Y понимают объединенный ансамбль, который характеризуется совместными вероятностями $P(x_i,y_j)$. Энтропия такого ансамбля по аналогии с формулой (5.4) равна

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

Для независимых ансамблей $P(x_i,y_j)=P(x_i)P(y_j)$, следовательно

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

$$=-\sum_{i=1}^{m}P(x_i)\log P(x_i)-\sum_{j=1}^{n}P(y_j)\log P(y_j)=H(X)+H(Y)$$

Это соотношение обобщается на несколько независимых ансамблей:

$$H(X_1,X_2,...,X_k)=\sum_{l=1}^{k}H(X_l). \tag{5.10}$$

5.1.3 Энтропия при непрерывном состоянии элементов

Выражение при дискретном состоянии элементов можно в определенном смысле обобщить на случай непрерывных сообщений.

Путь задана плотность вероятности $W(x)$ состояний элементов непрерывного сообщения. Характер случайного процесса существенно не изменится, если непрерывные состояния заменить дискретными $x_1, x_2, \ldots, x_m$, отстоящими друг от друга на интервал Δx и имеющими вероятности $P(x_i) = W(x_i)\Delta x$. Тогда в соответствии с формулой (5.4) можно записать

$$H_{ä}(X) = -\sum_i P(y_j)\log P(y_j) = -\sum_i \left[W(x_i)\Delta x\right]\log\left[W(x_i)\Delta x\right] =$$

$$= -\sum_i \left[W(x_i)\log W(x_i)\right]\Delta x - \log\Delta x\left[\sum_i W(x_i)\Delta x\right]$$

Переходя к пределу при $\Delta x \to 0$, получим, что энтропия непрерывного сообщения равна

$$H(X) = \lim_{\Delta x\to 0} H_{ä}(X) = -\int_{-\infty}^{\infty} W(x)\log W(x)dx - \lim_{\Delta x\to 0}\log\Delta x, \qquad (5.11)$$

так как $\int_{-\infty}^{\infty} W(x)dx = 1$.

Величина $H^*(X) = -\int_{-\infty}^{\infty} W(x)\log W(x)dx$ называется *приведенной или дифференциальной энтропией* непрерывного сообщения и имеет конечное значение. Величина $H^*(X)$ не зависит от Δx и определяется только плотностью вероятностей непрерывного состояния x. Второй член правой части соотношения (5.11) зависит от выбранного интервала Δx и именно он обращает значение $H(X)$ для непрерывного сообщения в бесконечность. Однако среднее количество информации при непрерывных сообщениях будет определяться только разностью априорной $\left[H^*(X)\right]$ и апостериорной $\left[H_0^*(X)\right]$ приведенными энтропиями:

$$I(X) = H(X) - H_0(X) = H^*(X) - H_0^*(X).$$

Свойства приведенной энтропии

1. *Приведенная энтропия зависит от масштаба переменной x и даже может менять знак.*

Например, изменив масштаб в k раз, что соответствует переходу к переменной $x_1 = kx$, и, учтя, что $W\left(x_1\right) = \frac{1}{k}W\left(x\right)$, получим

$$H^*\left(X_1\right) = -\int_{-\infty}^{\infty} W\left(x_1\right)\log W\left(x_1\right)dx_1 = -\int_{-\infty}^{\infty} \frac{W\left(x\right)}{k}\log\frac{W\left(x\right)}{k}kdx =$$

$$= -\int_{-\infty}^{\infty} W\left(x\right)\log W\left(x\right)dx + \log k\int_{-\infty}^{\infty} W\left(x\right)dx = H^*\left(X\right) + \log k.$$

Следует заметить, что разность приведенных энтропий двух распределений при одном и том же масштабе переменной x не зависит от выбора масштаба.

2. *Приведенная энтропия не зависит от математического ожидания состояния непрерывного сообщения.*

$$H^*\left(X - m_1\right) = -\int_{-\infty}^{\infty} W\left(x - m_1\right)\log W\left(x - m_1\right)d\left(x - m_1\right) =$$

$$= -\int_{-\infty}^{\infty} W\left(x\right)\log W\left(x\right)dx = H^*\left(X\right).$$

3. *Приведенная энтропия аддитивна.*

Т.е. для независимых сообщений имеем:

$$H^*\left(X,Y\right) = H^*\left(X\right) + H^*\left(Y\right).$$

5.1.4 Определение плотности вероятностей состояний непрерывных сообщений, обладающих максимальной энтропией

Определим вид функций плотности вероятностей состояний $W\left(x\right)$, обеспечивающих максимальную приведенную энтропию для двух важных случаев:

1. Дисперсия состояний сообщения – величина заданная:
2. На дисперсию состояний сообщения не накладываются никакие ограничения.

В обеих случаях, естественно, выполняется условие

$$\int_{-\infty}^{\infty} W(x)\,dx = 1.$$

Для решения этих вариационных задач воспользуемся методом неопределенных множителей Лагранжа, в соответствие с которым функция $f(x)$, обеспечивающая экстремальное значение функционала

$$J = \int_a^b F\left[x, f(x)\right] dx \tag{5.12}$$

при дополнительных условиях

$$\begin{cases} \int_a^b \varphi_1\left[x, f(x)\right] dx = k_1 = const \\ \int_a^b \varphi_2\left[x, f(x)\right] dx = k_2 = const \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \int_a^b \varphi_n\left[x, f(x)\right] dx = k_n = const \end{cases}, \tag{5.13}$$

находится решением уравнения

$$\frac{dF}{df(x)} + \lambda_1 \frac{d\varphi_1}{df(x)} + \lambda_2 \frac{d\varphi_2}{df(x)} + \cdots + \lambda_n \frac{d\varphi_n}{df(x)} = 0. \tag{5.14}$$

Здесь $\lambda_1, \lambda_2, \dots, \lambda_n$ - неопределенные множители Лагранжа, которые определяются подстановкой $f(x)$ в равенства (5.13).

1. Для первой задачи требуется найти такую функцию $W(x)$, для которой $H^*(X) = -\int_{-\infty}^{\infty} W(x) \log W(x)\,dx$ - максимальна при условиях:

$$\text{а) } \int_{-\infty}^{\infty} x^2 W(x)\,dx = \sigma^2, \quad \text{б) } \int_{-\infty}^{\infty} W(x)\,dx = 1. \tag{5.15}$$

Функции F, φ_1 и φ_2 в данном случае имеют вид $F = W(x)\log W(x)$, $\varphi_1 = x^2 W(x)$, $\varphi_2 = W(x)$ и уравнение (5.14) принимает вид

$$\log W(x) + \log e + \lambda_1 x^2 + \lambda_2 = 0.$$

Из последнего следует, что

$$\log W(x) = -\lambda_1 x^2 - \lambda_2 - \log e.$$

Учитывая, что переход от одной системы логарифмов к другой осуществляется по формуле $\log_b c = \log_b a \log_a c$, получим

$$\frac{\ln W(x)}{\ln 2} = -\lambda_1 x^2 - \lambda_2 - \frac{1}{\ln 2},$$

или

$$\ln W(x) = -\lambda_1^* x^2 - \lambda_2^* - 1, \tag{5.16}$$

где $\lambda_1^* = \lambda_1 \ln 2$, $\lambda_2^* = \lambda_2 \ln 2$.

Из соотношения (5.16) определяем $W(x)$:

$$W(x) = \exp\left[-\left(\lambda_2^* + 1\right)\right]\exp\left(-\lambda_1^* x^2\right).$$

Используя свойство (5.15 б), получим

$$\int_{-\infty}^{\infty} W(x)dx = \int_{-\infty}^{\infty} \exp\left[-\left(\lambda_2^* + 1\right)\right]\exp\left(-\lambda_1^* x^2\right)dx =$$

$$= 2\exp\left[-\left(\lambda_2^* + 1\right)\right]\int_0^{\infty} \exp\left(-\lambda_1^* x^2\right)dx = 1$$

Отсюда

$$\exp\left[-\left(\lambda_2^* + 1\right)\right] = \frac{1}{2\int_0^{\infty} \exp\left(-\lambda_1^* x^2\right)dx} = \sqrt{\frac{\lambda_1^*}{\pi}}$$

С учетом этого выражения плотность вероятностей $W(x)$ можно теперь записать в виде

$$W(X) = \sqrt{\frac{\lambda_1^*}{\pi}}\exp\left(-\lambda_1^* x^2\right). \tag{5.17}$$

Подставим, полученное значение $W(x)$ в выражение для условия (5.15а) и получим

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 W(x)dx = 2\sqrt{\frac{\lambda_1^*}{\pi}}\int_0^{\infty} x^2 \exp\left(-\lambda_1^* x^2\right)dx = \frac{1}{2\lambda_1^*}.$$

Следовательно $\lambda_1^* = \dfrac{1}{2\sigma^2}$.

Решение приведенных интегралов проводилось по формуле [4]

$$\int\limits_0^{\infty} x^m e^{-bx^n}\,dx = \frac{1}{n}\cdot\frac{\tilde{A}\left(\dfrac{m+1}{n}\right)}{b^{\frac{m+1}{n}}}, \qquad (5.18)$$

где $\tilde{A}(z)$ - гамма-функция.

Подставляя значение λ_1^* в (5.17), получим

$$W(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}.$$

Эта функция представляет собой одномерную плотность вероятностей нормального случайного процесса с нулевым средним значением.

Выводы:

1. *Если задана дисперсия состояний непрерывного сообщения, то такое сообщение обладает наибольшей информативностью при нормальном законе распределения состояний сообщения.*
2. *При заданной средней мощности помехи наиболее опасной является та, которая имеет нормальный закон распределения.*

Определим величину приведенной энтропии при нормальном распределении состояний непрерывных сообщений. При

$$W(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m_1)^2}{2\sigma^2}}$$ (m_1- математическое ожидание состояний сообщения) получим

$$H^*(X) = -\int\limits_{-\infty}^{\infty} W(x)\log W(x)\,dx = -\int\limits_{-\infty}^{\infty} W(x)\log\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m_1)}{2\sigma^2}}\,dx =$$

$$= \frac{1}{\ln 2}\int\limits_{-\infty}^{\infty} W(x)\ln\sqrt{2\pi\sigma^2}\,dx + \frac{1}{\ln 2}\int\limits_{-\infty}^{\infty} W(x)\frac{(x-m_1)^2}{2\sigma^2}\,dx =$$ Так

$$= \frac{1}{\ln 2}\left[\ln\sqrt{2\pi\sigma^2}\int\limits_{-\infty}^{\infty} W(x)\,dx + \frac{1}{2\sigma^2}\int\limits_{-\infty}^{\infty}(x-m_1)^2 W(x)\,dx\right]$$

как $\int\limits_{-\infty}^{\infty} W(x)\,dx = 1$, а $\int\limits_{-\infty}^{\infty}(x-m_1)^2 W(x)\,dx = \sigma^2$,

То получим $H^*(X) = \dfrac{1}{\ln 2}\left(\ln\sqrt{2\pi\sigma^2} + \dfrac{1}{2}\right).$

Имея в виду, что $\ln\sqrt{e} = 1/2$, окончательно получим выражение для приведенной энтропии

$$H^*(X) = \frac{1}{\ln 2}\left(\ln\sqrt{2\pi\sigma^2} + \ln\sqrt{e}\right) = \log\left(\sigma\sqrt{2\pi e}\right). \quad (5.19)$$

Таким образом, приведенная энтропия при гауссовой статистике состояний сообщений прямо пропорциональна логарифму среднеквадратического значения состояний.

Как и следовало ожидать, приведенная энтропия не зависит от математического ожидания.

2. Рассмотрим вторую задачу, при которой не накладываются какие-либо ограничения на дисперсию состояний. В этом случае используется только одно дополнительное условие (5.15б) и функции F и φ_1 имеют вид

$$F = W(x)\log W(x), \quad \varphi_1 = W(x).$$

Уравнение (5.14) в этом случае принимает вид

$$\log W(x) + \log e + \lambda_1 = 0.$$

После элементарных преобразований получим

$$W(x) = e^{-\left(\lambda_1^* + 1\right)}, \quad (5.20)$$

где $\lambda_1^* = \lambda_1 \ln 2$.

Так как функция плотности вероятностей (5.20) не зависит от x, то она является постоянной во всем интервале существования случайной величины, т.е. величина x распределена равновероятно.

Если состояния сообщения существуют в интервале $[a,b]$, то искомая функция распределения равна

$$W(x) = \begin{cases} \dfrac{1}{b-a} & \text{при} \quad a < x < b \\ 0 & \quad\quad\; x < a, x > b \end{cases}.$$

Выводы

1. *Если на дисперсию состояний не накладываются ограничения, то непрерывное сообщение обладает наибольшей информативностью при равновероятном распределении состояний.*
2. *При неизвестной мощности помехи наиболее опасной будет помеха с равновероятным распределением.*

Определим величину приведенной энтропии при равновероятном распределении состояний.

$$H^*(X) = -\int\limits_{-\infty}^{\infty} W(x)\log W(x)\,dx = -\int\limits_a^b \frac{1}{b-a}\log\frac{1}{b-a}\,dx =$$
$$= \log(b-a)$$

Отсюда следует, что величина приведенной энтропии при равновероятном распределении состояний прямо пропорциональна логарифму величины интервала изменения состояний.

5.1.5 Условная энтропия статистически зависимых сообщений

На практике приходится сталкиваться с двумя и более источниками, дающими зависимые друг от друга сообщения.

Положим, что имеются два элемента сообщения, первый из которых характеризуется ансамблем X

$$X = \begin{bmatrix} \dfrac{x_1}{P(x_1)}, \dfrac{x_2}{P(x_2)}, \ldots, \dfrac{x_m}{P(x_m)} \end{bmatrix},$$

а второй – ансамблем Y

$$Y = \begin{bmatrix} \dfrac{y_1}{P(y_1)}, \dfrac{y_2}{P(y_2)}, \ldots, \dfrac{y_n}{P(y_n)} \end{bmatrix}.$$

Взаимная статистическая связь между состояниями элементов можно характеризовать условными вероятностями $P(y_i|x_k)$, т.е. вероятностями появления состояния y_i при условии, что состояние x_k уже известно.

Схема условного ансамбля в этом случае принимает вид

$$Y|X = \begin{bmatrix} & x_1 & x_2 & \ldots & x_m \\ y_1 & P(y_1|x_1) & P(y_1|x_2) & \ldots & P(y_1|x_m) \\ y_2 & P(y_2|x_1) & P(y_2|x_2) & \ldots & P(y_2|x_m) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ y_n & P(y_n|x_1) & P(y_n|x_2) & \ldots & P(y_n|x_m) \end{bmatrix}. \qquad (5.21)$$

По аналогии с (5.3) достоверное сообщение о том, что из всех состояний y появляется y_i при условии, что имело место состояние x_k, несет в себе количество информации, равное [$-\log P(y_i|x_k)$], которое можно рассматривать как *случайную условную энтропию*.

При фиксированном состоянии x_k совокупность условных вероятностей (5.21) определяет частную условную энтропию

$$H\left(Y\middle|x_k\right)=-\sum_{i=1}^{n}P\left(y_i\middle|x_k\right)\log P\left(y_i\middle|x_k\right), \tag{5.22}$$

которая характеризует среднее количество информации, несущее элементом Y после того, как стало известно состояние x_k элемента X.
При сильной статистической связи X и Y частная условная энтропия будет малой и, наоборот, при слабой статистической связи – большой.

Если частную условную энтропию (5.22) усреднить по всем состояниям x_k с учетом вероятности появления каждого из состояний $P\left(x_k\right)$, то получим выражение условной энтропии сообщения Y относительно элемента сообщения X:

$$\begin{aligned}H\left(Y\middle|X\right)&=\sum_{k=1}^{m}P\left(x_k\right)H\left(Y\middle|x_k\right)=\\&=-\sum_{i=1}^{n}\sum_{k=1}^{m}P\left(x_k\right)P\left(y_i\middle|x_k\right)\log P\left(y_i\middle|x_k\right)\end{aligned}. \tag{5.23}$$

Так как в соответствии с теоремой умножения вероятностей имеем

$$P\left(x_k,y_i\right)=P\left(x_k\right)P\left(y_i\middle|x_k\right)=P\left(y_i\right)P\left(x_k\middle|y_i\right),$$

То выражение (5.23) можно переписать в виде

$$H\left(Y\middle|X\right)=-\sum_{i=1}^{n}\sum_{k=1}^{m}P\left(x_k,y_i\right)\log P\left(y_i\middle|x_k\right). \tag{5.24}$$

Основной смысл условной энтропии состоит в том, что она показывает, какую энтропию дает сообщение элемента Y, если уже известна энтропия сообщения X.

Свойства условной энтропии.

1. *Если элементы сообщения X и Y статистически независимы, то условная энтропия $H\left(Y\middle|X\right)$ равна энтропии $H\left(Y\right)$,* т.е.

$$H\left(Y\middle|X\right)=H\left(Y\right).$$

Действительно, если элементы X и Y статистически независимы, то $P\left(y_i\middle|x_k\right)=P\left(y_i\right)$ и соотношение (5.23) приобретает вид

$$H\left(Y|X\right)=-\sum_{i=1}^{n}\sum_{k=1}^{m}P\left(x_k\right)P\left(y_i\right)\log P\left(y_i\right)=$$

$$=-\sum_{k=1}^{m}P\left(x_k\right)\sum_{i=1}^{n}P\left(y_i\right)\log P\left(y_i\right).$$

Так как $\sum_{k=1}^{m}P\left(x_k\right)=1$, то

$$H\left(Y|X\right)=-\sum_{i=1}^{n}P\left(y_i\right)\log P\left(y_i\right)=H\left(Y\right).$$

2. *Если элементы сообщения X и Y являются статистически жестко связанными, то условная энтропия элемента Y относительно элемента X равна нулю*

$$H\left(Y|X\right)=0.$$

Это означает, что сообщение элемента Y не содержит никакой информации сверх той, которая содержится в сообщении элемента X.

Действительно, при полной статистической зависимости условные вероятности $P\left(y_i|x_k\right)$ либо раны нулю, либо – единице. Следовательно, энтропия будет равна нулю.

Для непрерывных сообщений частную условную приведенную энтропию можно представить интегралом

$$H^*\left(Y|x\right)=-\int_{-\infty}^{\infty}W\left(y|x\right)\log W\left(y|x\right)dy, \qquad (5.25)$$

а условную приведенную энтропию – соотношениями:

$$H^*\left(Y|X\right)=-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}W\left(x\right)W\left(y|x\right)\log W\left(y|x\right)dxdy$$

$$H^*\left(Y|X\right)=-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}W\left(x,y\right)\log W\left(y|x\right)dxdy. \qquad (5.26)$$

5.1.6. Энтропия объединения статистически зависимых сообщений

Положим, что известны ансамбли двух зависимых элементов сообщений X и Y. Зависимость состояний этих элементов можно задать совместными вероятностями $P\left(x_k, y_i\right)$ посредством схемы объединенного ансамбля вида:

$$X \times Y = \begin{bmatrix} & x_1 & x_2 & \dots & x_m \\ y_1 & P\left(x_1, y_1\right) & P\left(x_2, y_1\right) & \dots & P\left(x_m, y_1\right) \\ y_2 & P\left(x_1, y_2\right) & P\left(x_2, y_2\right) & \dots & P\left(x_m, y_2\right) \\ \dots & \dots & \dots & \dots & \dots \\ y_n & P\left(x_1, y_n\right) & P\left(x_2, y_n\right) & \dots & P\left(x_m, y_n\right) \end{bmatrix}$$

Объединенный ансамбль описывает более сложную систему, состоящую из двух подсистем X и Y, в которой возможны $n \times m$ различных состояний с заданным распределением вероятностей $P\left(x, y\right)$.

Достоверное сообщение о том, что элемент X из всех возможных состояний принимает значение x_k, а второй элемент Y при этом принимает состояние y_i, содержит в себе количество информации, равное $H\left(x_k, y_i\right) = -\log P\left(x_k, y_i\right)$, которое можно назвать *случайной энтропией объединения*.

Энтропия объединения будет равна

$$H\left(X, Y\right) = -\sum_{i=1}^{n} \sum_{k=1}^{m} P\left(x_k, y_i\right) \log P\left(x_k, y_i\right). \tag{5.27}$$

С учетом того, что $P\left(x_k, y_i\right) = P\left(x_k\right) P\left(y_i \middle| x_k\right)$ соотношение (5.27) можно представит в следующем виде

$$H\left(X, Y\right) = -\sum_{i=1}^{n} \sum_{k=1}^{m} P\left(x_k, y_i\right) \log\left[P\left(x_k\right) P\left(y_i \middle| x_k\right)\right] =$$

$$= -\sum_{i=1}^{n} \sum_{k=1}^{m} P\left(x_k, y_i\right) \log P\left(x_k\right) - \sum_{i=1}^{n} \sum_{k=1}^{m} P\left(x_k, y_i\right) \log P\left(y_i \middle| x_k\right) =$$

$$= -\sum_{i=1}^{n} \sum_{k=1}^{m} P\left(x_k\right) P\left(y_i \middle| x_k\right) \log P\left(x_k\right) - \sum_{i=1}^{n} \sum_{k=1}^{m} P\left(x_k, y_i\right) \log P\left(y_i \middle| x_k\right)$$

Выполняя суммирование по i в первом слагаемом представленного соотношения, получим

$$H\left(X, Y\right) = H\left(X\right) + H\left(Y \middle| X\right), \tag{5.28}$$

так как $\sum_{i=1}^{n} P\left(y_i \middle| x_k\right) = 1, \quad -\sum_{k=1}^{m} P\left(x_k\right) \log P\left(x_k\right) = H\left(X\right)$

и

$$-\sum_{i=1}^{n}\sum_{k=1}^{m}P(x_k,y_i)\log P(y_i|x_k)=H(Y|X).$$

Таким образом, энтропия объединения, определяемая зависимостью (5.28), равна сумме двух энтропий: энтропии $H(X)$ элемента сообщения X и условной энтропии $H(Y|X)$. Из равенства (5.28) более полно раскрывается смысл условной энтропии: она представляет собой то добавочное среднее количество информации, которое вносит сообщение элемента Y, если среднее количество информации сообщения элемента X уже известно.

Свойства энтропии объединения.

1.Очевидно, что

$$H(X,Y)=H(Y,X)=H(X)+H(Y|X)=H(Y)+H(X|Y) \qquad (5.29)$$

2. Для статистически независимых сообщений энтропия объединения равна сумме энтропий сообщений

$$H(X,Y)=H(X)+H(Y). \qquad (5.30)$$

3. При полной статистической зависимости элементов X и Y энтропия объединения равна

$$H(X,Y)=H(X)=H(Y), \qquad (5.31)$$

т.е. при полной статистической зависимости энтропии элементов X и Y равны между собой.

При вычислении энтропии иногда полезно использовать некоторые соотношения теории вероятностей:

$$\sum_k P(x_k)=1$$

$$\sum_i P(y_i)=1$$

$$\sum_k P(y_i|x_k)P(x_k)=\sum_k P(x_k,y_i)=P(y_i) \qquad (5.32)$$

$$\sum_i P(x_k|y_i)P(y_i)=\sum_i P(y_i,x_k)=P(x_k)$$

$$\sum_i\sum_k P(x_k)P(y_i|x_k)=\sum_i\sum_k P(y_i)P(x_k|y_i)=1$$

Для непрерывных сообщений приведенная энтропия объединения будет определяться соотношением

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

$$= -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(x,y)\log W(x)\,dxdy -$$

$$-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(x,y)\log W(y|x)\,dxdy = \qquad (5.33)$$

$$= -\int_{-\infty}^{\infty} W(x)\log W(x)dx - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(x,y)\log W(y|x)\,dxdy,$$

так как $\int_{-\infty}^{\infty} W(x,y)\,dy = W(x)$.

В качестве иллюстрации рассмотрим следующий пример.

Пример. Ансамбль состояний X и Y объединен и представлен в виде

$$X \times Y = \begin{bmatrix} & x_1 & x_2 & x_3 \\ y_1 & 0{,}10 & 0{,}20 & 0{,}30 \\ y_2 & 0{,}25 & 0 & 0{,}15 \end{bmatrix}.$$

Определить: а) энтропию ансамблей X и Y,

б) энтропию объединенного ансамбля $X \times Y$,

в) условные энтропии ансамблей.

Решение. Используя формулы (5.32), определим вероятности состояний в ансамблях X и Y

$$P(x_1) = P(x_1, y_1) + P(x_1, y_2) = 0{,}1 + 0{,}25 = 0{,}35.$$

Аналогично получим

$$P(x_2) = 0{,}2;\ P(x_3) = 0{,}45;\ P(y_1) = 0{,}6;\ P(y_2) = 0{,}4.$$

Зная вероятности состояний по формуле (5.4), получим

$$H(X) = -\sum_{k=1}^{3} P(x_k)\log P(x_k) = -0,35\log 0,35 -$$
$$-0,2\log 0,2 - 0,45\log 0,45 = 1,512\textit{äâ.åä.};$$
$$H(Y) = -\sum_{i=1}^{2} P(y_i)\log P(y_i) = 0,971\textit{äâ.åä.}$$

Для определения энтропии объединенного ансамбля воспользуемся выражением (5.27)

$$H(X,Y) = -\sum_{i=1}^{2}\sum_{k=1}^{3} P(x_k, y_i)\log P(x_k, y_i) =$$
$$= -0.1\log 0,1 - 0,25\log 0,25 - 0,2\log 0,2 -$$
$$-0,3\log 0,3 - 0,15\log 0,15 = 2,228\textit{äâ.åä.}$$

Условные энтропии вычислим по формулам (5.28) и (5.29)

$$H(X|Y) = H(X,Y) - H(Y) = 2,228 - 0,971 = 1,257\textit{äâ.åä.}$$

и

$$H(Y|X) = H(X,Y) - H(X) = 2,228 - 1,512 = 0,716\textit{äâ.åä.}$$

5.1.7. Количество информации при неполной достоверности результатов сообщения

Как указывалось выше, информация о том или ином состоянии добывается в результате опыта. Однако не всегда можно с полной достоверностью утверждать, какое именно состояние имело место. Определим среднее количество информации для такой ситуации. Начнем с простейшей задачи.

Допустим, что интересующие нас состояния составляют ансамбль X, а результаты сообщений, на основе которых мы выносим суждение о x_k, составляют ансамбль Y. Обозначим через $P(x_k|y_i)$ вероятность того, что при известном нам состоянии y_i имело место состояние x_k.

Например, если

$$P(x_k|y_i) = \begin{cases} 1 \\ 0 \end{cases} \text{при} \begin{array}{l} k = i \\ k \neq i \end{array},$$

то в результате опыта или сообщения ситуация полностью определена и можно с полной достоверностью утверждать, какое состояние x_k имело место. Так как неопределенность появления состояния x до опыта равна $H\left(X\right)$, а после опыта неопределенность отсутствует, то в этом случае среднее количество информации $I\left(Y,X\right)$, содержащееся в Y относительно X, будет равно

$$I\left(Y,X\right)=H\left(X\right). \tag{5.34}$$

Пример. Колода карт состоит из 32 карт от семерки до туза. Вытаскивается из колоды любая карта. Необходимо определить какая карта вытащена, задавая вопросы, на которые даются ответы «да» или «нет». Определить максимальное число вопросов, которое гарантирует определение вытащенной карты.

Решение. Так как любая карта может быть вытащена с равной вероятностью, то согласно (5.9) энтропия равна

$H\left(X\right)=\log 32=5$*äâ.åä*.

Такое же число двоичных единиц должно содержать сообщение об этом событии, т.е.

$$I\left(Y,X\right)=H\left(X\right).$$

Ответ «да» или «нет» содержит одну двоичную единицу информации. Следовательно, достаточно задать пять вопросов, ответы на которые должны позволить определить вытащенную карту.

Допустим, что вытащен валет треф. Тогда:

Вопрос 1: Масть красная?
Ответ: Нет.
Вывод: Вытащена черная масть.

Вопрос 2: Масть пика?
Ответ: Нет.
Вывод: Вытащена трефа.

Вопрос 3: Картинка?
Ответ: Да.
Вывод: Вытащена трефовая картинка.

Вопрос 4: Старше дамы?
Ответ: Нет.
Вывод: Вытащена трефовая дама или валет.

Вопрос 5: Валет?
Ответ: Да.
Вывод: Вытащен трефовый валет

В общем случае, когда $0<P\left(x_k\middle|y_i\right)<1$, количество информации можно определить следующим образом. Положим, что передаваемое количество информации равно $H\left(X\right)$, а принимаемое $H\left(Y\right)$. Тогда после приема остается неопределенность, которая в среднем может характеризоваться условной энтропией $H(X|Y)$. Иными словами, $H(X|Y)$, есть то количество информации, которое может дать полное знание X, когда известно количество информации, даваемое Y. Следовательно, $H(X|Y)$ можно рассматривать как то количество информации, которого недостает для полного знания энтропии объединения, когда известна энтропия $H(Y)$. Поэтому $H(X|Y)$ можно назвать *потерей информации*.

Если из количества информации $H(X)$ вычесть потерю информации $H(X|Y)$, то получим количество информации $I(Y,X)$, которое содержится в принятой совокупности сообщений Y относительно переданной X:

$$I(Y,X) = H(X) - H(X|Y). \quad (5.35)$$

Если элементы сообщений X и Y статистически зависимы, то (5.35) переходит к соотношению (5.34), т.е.

$$I(Y,X) = H(X).$$

Если элементы сообщения X и Y статистически независимы, то $I(Y,X) = 0$, так как $H(X|Y) = H(X)$. Это означает, что состояния элемента Y не содержат никакой информации о X.

Необходимо заметить, что

$$I(Y,X) = I(X,Y), \quad (5.36)$$

так как

$$H(X,Y) = H(Y,X) = H(X) + H(Y|X) = H(Y) + H(X|Y) \quad (5.37)$$

Соотношение (5.36) указывает на то, что количество информации, которое содержится в Y относительно X, равно количеству информации, которое содержится в X относительно Y. Поэтому $I(Y,X)$ и $I(X,Y)$ называются полной взаимной информацией.

Так как согласно формуле (5.37) условная энтропия может быть выражена в виде

$$H(X|Y) = H(X,Y) - H(Y),$$

то выражение для полной взаимной информации принимает вид

$$I(Y,X) = H(X) + H(Y) - H(X,Y). \quad (5.38)$$

Раскроим содержание этой формулы.

$$I(T,X) = -\sum_{k=1}^{m} P(x_k)\log P(x_k) - \sum_{i=1}^{n} P(y_i)\log P(y_i) +$$
$$-\sum_{k=1}^{m}\sum_{i=1}^{n} P(x_k, y_i)\log P(x_k, y_i)$$

Умножим первую сумму этого соотношения на $\sum_{i=1}^{n} P(y_i|x_k)$, а вторую на $\sum_{k=1}^{m} P(x_k|y_i)$, получим

$$I(Y,X) = -\sum_{k=1}^{m}\sum_{i=1}^{n} P(x_k)P(y_i|x_k)\log P(x_k) -$$
$$-\sum_{k=1}^{m}\sum_{i=1}^{n} P(y_i)P(x_k|y_i)\log P(y_i) + \sum_{k=1}^{m}\sum_{i=1}^{n} P(x_k, y_i)\log P(x_k, y_i) =$$
$$= -\sum_{k=1}^{m}\sum_{i=1}^{n} P(x_k, y_i)\log P(x_k) - \sum_{k=1}^{m}\sum_{i=1}^{n} P(x_k, y_i)\log P(y_i) +$$
$$+\sum_{k=1}^{m}\sum_{i=1}^{n} P(x_k, y_i)\log P(x_k, y_i). \qquad (5.39)$$

Принимая во внимание, что

$$-\log P(x_k) - \log P(y_i) + \log P(x_k, y_i) = \log\frac{P(x_k, y_i)}{P(x_k)P(y_i)},$$

выражение (5.39) принимает вид

$$I(Y,X) = \sum_{k=1}^{m}\sum_{i=1}^{n} P(x_k, y_i)\log\frac{P(x_k, y_i)}{P(x_k)P(y_i)}. \qquad (5.40)$$

Формула (5.40) определяет количество информации, содержащиеся в Y относительно X при известных вероятностях $P(x_k)$, $P(y_i)$ и $P(x_k, y_i)$.

В качестве иллюстрации рассмотрим следующий простой пример.

Пример. В результате наличия помехи при передачи бинарного сообщения вероятность правильного приема состояний (x_1 и x_2) уменьшается до 0,9.

Определить принимаемое количество информации относительно передаваемого.

Решение. Так как передается бинарное сообщение, то $P(x_1)=P(x_2)=0,5$.

Условные вероятности равны

$$P(y_1|x_1)=P(y_2|x_2)=0,9$$
$$P(y_2|x_1)=P(y_1|x_2)=0,1 \, .$$

Используя соотношение (5.32) определим $P(y_1)$ и $P(y_2)$

$$P(y_1)=P(y_1|x_1)P(x_1)+P(y_1|x_2)P(x_2)=0,9\cdot 0,5+0,1\cdot 0,5=0,5$$

$P(y_2)=1-P(y_1)=0,5$.

Воспользуемся формулой (5.40).

$$I(Y,X)=\sum_{k=1}^{m}\sum_{i=1}^{n}P(x_k,y_i)\log\frac{P(x_k,y_i)}{P(x_k)P(y_i)}=$$
$$=\sum_{k=1}^{m}\sum_{i=1}^{n}P(x_k)P(y_i|x_k)\log\frac{P(y_i|x_k)}{P(y_i)} \, ,$$

так как $P(x_k,y_i)=P(x_k)P(y_i|x_k)$.

Следовательно:

$$I(Y,X)=\sum_{k=1}^{2}\sum_{i=1}^{2}P(x_k)P(y_i|x_k)\log\frac{P(y_i|x_k)}{P(y_i)}=$$
$$=2\cdot 0,5\cdot 0,9\cdot\log\frac{0,9}{0,5}+2\cdot 0,5\cdot 0,1\cdot\log\frac{0,1}{0,5}=0,532 \text{äâ.åä.}$$

Если бы помехи не было, то $P(y_2|x_1)=P(y_1|x_2)=0$.

В этом случае $I(Y,X)=H(X)=1$дв.ед., так как энтропия бинарного сообщения равна одной двоичной единице.

Для непрерывных сообщений среднее количество информации, содержащееся в Y относительно X будет определяться выражением

$$I(Y,X)=H(X)-H(X|Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}W(x,y)\log\frac{W(x|y)}{W(x)}dxdy=$$
$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}W(x,y)\log\frac{W(x,y)}{W(x)W(y)}, \qquad (5.41)$$

где:

$$H(X) = -\int_{-\infty}^{\infty} W(x) \log W(x) dx - \lim_{\Delta x \to 0} \log \Delta x,$$

$$H(X|Y) = -\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(x, y) \log W(x|y) dxdy - \lim_{\Delta x \to 0} \log \Delta x.$$

Как энтропия $H(X)$, так и условная энтропия $H(X|Y)$ обращаются в бесконечность из-за члена $\lim_{\Delta x \to 0} \log \Delta x$. Однако величина $I(Y, X)$ конечна, так как эти члены одинаковы, но с различным знаком в формуле (5.41).

Так как величина условной энтропии $H(X|Y)$ может быть сколь угодно малой (чрезвычайно малая потеря информации), то $I(Y, X) = H^*(X)$, т.е. среднее количество информации будет определяться приведенной энтропией $H^*(X)$ непрерывного сообщения.

5.1.8. Частное количество информации

Прежде всего, рассмотрим частное количество информации, содержащееся в y_i относительно X.

Преобразуем (5.40) к следующему виду

$$I(Y, X) = \sum_{k=1}^{m} \sum_{i=1}^{n} P(x_k, y_i) \log \frac{P(x_k, y_i)}{P(x_k) P(y_i)} = \\ = \sum_{i=1}^{n} P(y_i) \sum_{k=1}^{m} P(x_k|y_i) \log \frac{P(x_k|y_i)}{P(x_k)} \tag{5.42}$$

Это количество информации можно рассматривать как среднее количество информации для всех состояний Y, полученное путем усреднения частных количеств информации, содержащихся в отдельных состояниях y_i, усредненных с весами, равными вероятностям $P(y_i)$.

Естественно, что величина

$$I(y_i, X) = \sum_{k=1}^{m} P(x_k|y_i) \log \frac{P(x_k|y_i)}{P(x_k)} \tag{5.43}$$

представляет собой частное количество информации, содержащееся в y_i относительно совокупности состояний X.

Полное среднее количество информации, содержащееся в Y относительно X, можно теперь записать в виде

$$I(Y,X)=\sum_{i=1}^{n}P(y_i)I(y_i,X). \tag{5.44}$$

Основное свойство частного количества информации состоит в том, что частное количество информации, содержащееся в отдельном состоянии y_i, всегда больше либо равно нулю, т.е. величина не отрицательная:

$$I(y_i,X)\geq 0.$$

Из этого свойства вытекают два важных следствия:

1. Так как частное количество информации $I(y_i,X)$ неотрицательно, то и полное среднее количество информации больше или равно нулю $\left[I(Y,X)\geq 0\right]$.

2. Так как $I(Y,X)\geq 0$, то $H(X,Y)\leq H(X)$.

Таким образом, потеря информации $H(X,Y)$ не превосходит энтропии исходного сообщения.

Во многих практически важных случаях возникает задача определения частного количества информации, содержащегося в y_i относительно x_k, т.е. определения $I(y_i,x_k)$. Например, по принятому одиночному импульсу, величина которого характеризуется y_i, определяется величина x_k передаваемого импульсного сигнала в условиях воздействия помехи, которая разрушает информацию об истинной величине передаваемого импульса. Наиболее достоверной в этом случае следует считать ту величину x_k, относительно которой содержится наибольшее количество информации. Это частное количество информации, содержащееся в y_i относительно x_k, равно

$$I(y_i,x_k)=\log\frac{P(x_k|y_i)}{P(x_k)}. \tag{5.45}$$

Это частное количество информации (5.45) называют также *случайной информацией связи.*

Рассмотрим основные свойства частного количества информации, содержащегося y_i относительно x_k.

1.Частное количество информации может иметь как положительные, так и отрицательные значения.

Представим формулу (5.45) в виде

$$I\left(y_i, x_k\right) = \log P\left(x_k \middle| y_i\right) - \log P\left(x_k\right).$$

Из этого выражения следует:

при $P\left(x_k \middle| y_i\right) > P\left(x_k\right) \to I\left(y_i, x_k\right) > 0$.

Если состояние y_i увеличивает вероятность передачи состояния x_k по сравнению с вероятностью передачи того же состояния до приема, то частное количество информации – положительно.

При $P\left(x_k \middle| y_i\right) < P\left(x_k\right) \to I\left(y_i, x_k\right) < 0$.

В этом случае знание состояния y_i уменьшает вероятность передачи состояния x_k по сравнению с априорной.

При $P\left(x_k \middle| y_i\right) = P\left(x_k\right) \to I\left(y_i, x_k\right) = 0$.

Частное количество информации равно нулю, когда значение принятого сообщения y_i не меняет вероятности передачи x_k.

2.Частное количество информации бесконечно велико и отрицательно, если $P\left(x_k \middle| y_i\right) = 0$.

Так как $\lim\limits_{z \to 0} \log z = -\infty$, то

$$P\left(x_k \middle| y_i\right) = \lim_{P\left(x_k | y_i\right) \to 0} \log \frac{P\left(x_k \middle| y_i\right)}{P\left(x_k\right)} = -\infty.$$

3. Частное количество информации, содержащееся в состоянии y_i, относительно x_k, равно частному количеству информации, содержащейся в состоянии x_k относительно y_i

$$I\left(y_i, x_k\right) = I\left(x_k, y_i\right).$$

Особое значение приобрела разность частных количеств информации.

Предположим, что передаваемый элемент характеризуется ансамблем X

$$X = \begin{bmatrix} x_1 & x_2 & \ldots & x_m \\ P\left(x_1\right) & P\left(x_2\right) & \ldots & P\left(x_m\right) \end{bmatrix}.$$

Принято одно состояние y_i. Вследствие воздействия помех однозначного соответствия между передаваемыми и принимаемыми состояниями нет. Поэтому наиболее вероятно переданным состоянием следует считать то, относительно которого в принятом состоянии содержится наибольшее количество информации.

Частные количества информации, содержащиеся в y_i относительно взятых x_k и x_r, соответственно равны

$$I(y_i, x_k) = \log \frac{P(x_k | y_i)}{P(x_k)} \text{ и } I(y_i, x_r) = \log \frac{P(x_r | y_i)}{P(x_r)}.$$

Разность частных количеств информаций будет равна

$$I(y_i, x_k) - I(y_i, x_r) = \log \frac{P(x_k | y_i) P(x_r)}{P(x_r | y_i) P(x_k)}. \tag{5.46}$$

Если величина разности (5.46) является положительной, то наиболее вероятно, что было передано состояние x_k, если – отрицательной, то наиболее вероятно, что было передано состояние x_r.

Формула (5.46) может быть представлена также в другом виде. Так как

$$\mathrm{P}(\mathrm{x_k}, y_i) = P(x_k) P(y_i | x_k) = P(y_i) P(x_k | y_i),$$

то

$$\frac{P(x_k | y_i)}{P(x_k)} = \frac{P(y_i | x_k)}{P(y_i)}. \tag{5.47}$$

Аналогично получим

$$\frac{P(x_r | y_i)}{P(x_r)} = \frac{P(y_i | x_r)}{P(y_i)}. \tag{5.48}$$

Подставляя (5.47) и (5.48) в (5.46), получим

$$I(y_i, x_k) - I(y_i, x_r) = \log \frac{P(y_i | x_k)}{P(y_i | x_r)}. \tag{5.49}$$

Разность частных количеств информации, представленная в виде (5.49), наиболее удобна для решения ряда важных практических задач.

Для непрерывных сообщений частное количество информации, которое содержится в y относительно X, определяется соотношением

$$I(y, X) = \int_{-\infty}^{\infty} W(x|y) \log \frac{W(x|y)}{W(x)} dx. \tag{5.50}$$

Следовательно, полное среднее количество информации Y относительно X можно представить в виде

$$I(Y, X) = \int_{-\infty}^{\infty} W(y) I(y, X) dy,$$

а частное количество информации, содержащееся в одиночном состоянии y относительно одиночного состояния x из совокупности X, (случайную информацию связи) определять формулой

$$I(y,x)=\log\frac{W(x|y)}{W(x)} \tag{5.51}$$

Разность частных количеств информаций для непрерывных сообщений можно представить соотношением

$$I(y,x_k)-I(y,x_r)=\log\frac{W(y|x_k)}{W(y|x_r)}. \tag{5.52}$$

Проиллюстрируем вычисление частного количества информации y_i относительно x_k на следующих примерах.

Пример. По двоичному симметричному каналу связи с помехами передается сообщение с ансамблем состояний

$$X=\begin{bmatrix} x_1 & x_2 \\ P(x_1)=\frac{3}{4} & P(x_2)=\frac{1}{4}\end{bmatrix}.$$

Из-за наличия помех вероятность правильного приема каждого из состояний (x_1 и x_2) уменьшается до $7/8$. Определить частное количество информации $I(x_1,y_2)$.

Решение. По условию задачи

$$P(x_1)=\frac{3}{4},P(x_2)=\frac{1}{4},P(y_1|x_1)=P(y_2|x_2)=\frac{7}{8},$$

$$P(y_1|x_2)=P(y_2|x_1)=\frac{1}{8}$$

Частное количество информации определим по формуле

$$I(x_2,y_2)=\log\frac{P(x_2|y_2)}{P(x_2)}=\log\frac{P(y_2|x_2)}{P(y_2)},$$

так как $P(x_2,y_2)=P(x_2)P(y_2|x_2)=P(y_2)P(x_2|y_2)$.

Величина $P(y_2|x_2)$ известна по условию и равна $7/8$. $P(y_2)$ определим согласно формуле (5.32).

$$P(y_2)=P(x_2)P(y_2|x_2)+P(x_1)P(y_2|x_2)=\frac{1}{4}\cdot\frac{7}{8}+\frac{3}{4}\cdot\frac{1}{8}=\frac{10}{32}$$

Следовательно, $I\left(x_2,y_2\right)$ будет равно

$$I\left(x_2,y_2\right)=\log\frac{7}{8}\cdot\frac{32}{10}=1,485\text{äâ.åä.}$$

Пример. На выходе фотоприемника имеется колебание $y\left(t\right)=x\left(t\right)+n\left(t\right)$, где $x\left(t\right)$- полезный сигнал, $n\left(t\right)$- помеха – независимые нормальные случайные процессы с нулевым средним значением и дисперсиями, равными σ_x^2 и σ_n^2 соответственно. Определить частное количество информации $I\left(x,y\right)$, которое содержится в принятом колебании $y\left(t\right)$ о сигнале $x\left(t\right)$.

Решение. По условию задачи

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

$$W\left(y|x\right)=\frac{1}{\sigma_n\sqrt{2\pi}}e^{-\frac{\left(y-x\right)^2}{2\sigma_n^2}}=\frac{1}{\sigma_n\sqrt{2\pi}}e^{-\frac{n^2}{2\sigma_n^2}}=W\left(n\right)$$

Так как $x\left(t\right)$ и $n\left(t\right)$- независимые, то $\sigma_y^2=\sigma_x^2+\sigma_n^2$, причем

$$W\left(x\right)=\frac{1}{\sigma_y\sqrt{2\pi}}e^{-\frac{y^2}{2\sigma_y^2}}.$$

По формуле (5.51) находим

$$I\left(x,y\right)=\log\frac{W\left(y|x\right)}{W\left(y\right)}=\frac{1}{\ln 2}\ln\frac{\dfrac{1}{\sigma_n\sqrt{2\pi}}e^{-\frac{n^2}{2\sigma_n^2}}}{\dfrac{1}{\sigma_y\sqrt{2\pi}}e^{-\frac{y^2}{2\sigma_y^2}}}=$$

$$=\frac{1}{\ln 2}\ln\frac{\sigma_y}{\sigma_n}e^{\frac{y^2}{2\sigma_y^2}-\frac{n^2}{2\sigma_n^2}}=\frac{1}{\ln 2}\ln\sqrt{1+\frac{\sigma_x^2}{\sigma_n^2}}+$$

$$+\frac{1}{\ln 2}\left[-\frac{n^2}{2\sigma_n^2}+\frac{y^2}{2\left(\sigma_x^2+\sigma_n^2\right)}\right]$$

5.1.9. Понятие об ε - энтропии

Допустим, что стоит задача о воспроизведении функции X по случайной функции Y. Будем искать количество информации, содержащееся в случайной функции Y относительно случайной функции X, при заданных требованиях к точности воспроизведения X.

При выборе критерия воспроизведения от него можно требовать, чтобы плотность вероятностей совместного распределения $W(x,y)$ принадлежала к некоторому классу функций.

В качестве такого критерия можно рассматривать средний риск, который характеризует качество принятия решений:

$$\begin{aligned} R &= \int\limits_{\Omega(\vec{x})}\int\limits_{\Omega(\vec{y})} \ddot{I}\,(\vec{x},\vec{y})W(\vec{x},\vec{y})d\Omega(\vec{x})d\Omega(\vec{y}) = \\ &= \int\limits_{\Omega(\vec{x})}\int\limits_{\Omega(\vec{y})} \ddot{I}\,(\vec{x},\vec{y})W(\vec{x})W(\vec{y}|\vec{x})d\Omega(\vec{x})d\Omega(\vec{y}), \end{aligned} \tag{5.53}$$

где: $\ddot{I}\,(\vec{x},\vec{y})$ - функция потерь, $\Omega(\vec{x})$ и $\Omega(\vec{y})$ соответственно пространства функций X и Y.

В этом случае требования к точности воспроизведения можно задать в виде

$$R \le \varepsilon, \tag{5.54}$$

где ε - приемлемо малая величина.

При известных функции потерь $\ddot{I}\,(\vec{x},\vec{y})$ и априорной плотности вероятностей $W(\vec{x})$ можно добиться выполнения условия (5.54), варьируя условной плотностью вероятностей $W(\vec{y}|\vec{x})$. При этом, возможно, условию (5.54) будет удовлетворять не одна, а несколько функций $W(\vec{y}|\vec{x})$ и, следовательно, функций $W(\vec{x},\vec{y})$.

Оптимальной функцией $W(\vec{y}|\vec{x})$ будет та, при которой полное среднее количество информации $I(Y,X)$, содержащееся в Y относительно X, будет минимальным, так как выбор такой функции позволяет при получении минимального количества информации выполнить заданные требования к точности воспроизведения.

Это *наименьшее значение $I(Y,X)$, при котором удовлетворяются заданные требования к точности воспроизведения, А.Н. Колмогоров назвал ε - энтропией* [5].

Следовательно, ε - энтропия равна

$$H_\varepsilon(X)=\underset{R\le\varepsilon}{Inf}\left[I(Y,X)\right]. \qquad (5.55)$$

Таким образом, ε- энтропия равна минимальному среднему количеству информации, которое должно содержать сообщение о случайной функции X при условии выполнения заданных требований к точности ее воспроизведения.

5.1.10. Избыточность сообщений

При одинаковом количестве различных состояний среднее количество информации, приходящееся на одно сообщение, или энтропия может быть различным в зависимости от статистических характеристик состояний. Сообщения, для которых энтропия максимальна, являются оптимальными в смысле наибольшего количества информации.

Мерой количественной оценки того, насколько данное сообщение отличается от соответствующего ему оптимального, служит *коэффициент сжатия*, численно равный отношению энтропии данного сообщения к энтропии ему оптимального

$$\mu=\frac{H(X)}{H_{\max}(X)}. \qquad (5.56)$$

Для того чтобы оценить, какая часть данного сообщения является избыточной по сравнению с соответствующим ему оптимальным сообщением, используется величина, называемая *коэффициентом избыточности*:

$$r=\frac{H_{\max}(X)-H(X)}{H_{\max}(X)}=1-\mu\ . \qquad (5.57)$$

Например, в русском алфавите 31 буква, если не различать «е» и «ё», а также твердый и мягкий знаки. Если к этим буквам алфавита добавит пробел между словами, то получится 32 состояния. Если бы все буквы появлялись с равной вероятностью, то энтропия такого языка была бы оптимальной и равна

$$H_{\max}(X)=\log 32=5\,\text{дв.ед.}$$

В действительности вероятность появления в тексте букв различны: так вероятность появления буквы «о» равна 0,09, буквы «ф» - 0,002. Кроме того, между буквами имеют место значительные корреляции. Проведенные исследования показали, что избыточность наиболее распространенных европейских языков превышает 50% [6 , 7].

Следует отметить, что если неоптимальные и оптимальные сообщения содержат одинаковое количество информации, то число элементов n неоптимального сообщения всегда будет больше числа элементов n_0, соответствующего ему оптимального сообщения.

Таким образом, избыточность приводит к увеличению времени передачи информации.

Для уменьшения избыточности можно осуществлять функциональное преобразование состояний, отыскивая такую функциональную зависимость, при которой плотности вероятностей будут приближаться к нормальному или равновероятному распределению.

Следует также заметить, что при передачи сообщений в условиях воздействия помех избыточность может быть использована для повышения помехоустойчивости сообщений.

5.2. Передача информации

5.2.1. Дискретные каналы без помех

Каналы передачи сообщений или информационные каналы преобразуют последовательности входных сообщений одного пространства в выходные сообщения другого пространства.

Информационные каналы классифицируются в соответствии с характеристиками входного и выходного пространств и распределением вероятностей, управляющим преобразованием одного пространства в другое. Канал называется дискретным, если входное и выходное пространства дискретны.

Функциональная схема дискретного информационного канала (дискретного канала передачи сообщений) без помех представлена на рис. 5.1

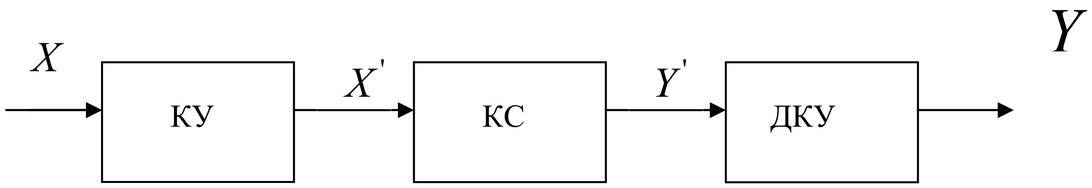

Рис.5.1
Функциональная схема дискретного информационного канала без помех

На рис. 5.1 приведены следующие обозначения:

X - сообщение от источника; $X^{'}$ - сигнал на входе канала связи;

$Y^{'}$ - сигнал на выходе канала связи; Y - сообщение для получателя;

КУ – кодирующее устройство; КС – канал связи;

ДКУ – декодирующее устройство.

Сообщение на вход информационного канала подается от источника сообщений. Источником сообщения является объект, состояния которого определяются физическим процессом, протекающим в пространстве или во времени по случайному закону.

Источник дискретных сообщений формирует некоторую последовательность элементов (символов), порядок следования которых случаен и характеризуется совокупностью вероятностей. При этом появляющиеся элементы могут иметь вероятностные связи с предшествующими элементами.

Источник дискретных сообщений называют эргодическим, если статистические свойства сообщения можно определить по одной достаточно длинной реализации. Для таких источников вероятностные связи распространяются на конечное число предшествующих элементов.

Любая достаточно длинная эргодическая последовательность будет *типичной*. Это означает, что частота появления любого элемента в такой последовательности с вероятностью сколь угодно близкой к единице равна вероятности появления этого элемента, а частота появления i элемента после j мало отличается от условной вероятности появления i элемента, если перед ним появился j элемент, и т.д. Таким образом, достаточно длинная последовательность, создаваемая эргодическим источником, с вероятностью сколь угодно близкой к единице характеризует вероятности появления отдельных элементов и вероятностные связи между ними.

Входное дискретное сообщение X преобразуется кодирующим устройством в сигнал. Сигналы передаются по каналу связи и на его конце восстанавливаются в сообщение для потребителя. Во многих дискретных информационных каналах используются специальные кодирующие и декодирующие устройства в цепях электрических сигналов.

С информационной точки зрения физическая реализация кодированных сигналов на передающей $X^{'}$ и на проемной $Y^{'}$ сторонах, а также выходных сообщений Y, значения не имеет. Важным является лишь установление соответствия между $X^{'}$ и X, между $Y^{'}$ и $X^{'}$ и, наконец, между Y и $Y^{'}$.

При отсутствии шумов можно принять $Y^{'} = X^{'}$. Способ же кодирования должен быть таким, чтобы по полученным кодированным сигналам можно было однозначно восстановить переданное сообщение. Это накладывает некоторые ограничения на допустимые комбинации элементов кодов.

Например, если закодировать элементы $x_1 \to 1$, $x_2 \to 0$, а $x_3 \to 01$, то при получении сигнала 01 мы не знаем было ли сообщение

x_3 или два сообщения x_2x_1. Следовательно, комбинацию x_2x_1 передавать нельзя. Могут накладываться и другие ограничения.

Совокупность запретов, обусловленных способом кодирования и построения аппаратуры, относятся к фиксированным ограничениям, накладываемым на информационный канал.

5.2.2. Пропускная способность дискретного канала без помех

Обозначим через X_T сообщение источника за время T, а соответствующие этому сообщению сигналы на входе и выходе канала связи и принимаемое сообщение через $X_T^{'}$, $Y_T^{'}$ и Y_T соответственно. Очевидно, что полное среднее количество информации $I\left(Y_T, X_T\right)$, содержащееся в Y_T относительно X_T, зависит от статистических характеристик состояний и от интервала времени T.

Рассмотрим предел

$$\lim_{T \to \infty} \frac{I\left(Y_T, X_T\right)}{T}. \tag{5.58}$$

При передаче сообщений эргодического источника при $T \to \infty$ с вероятностью, сколь угодно близкой к единице, последовательность сообщения X_T будет типичной. Например, простейшей типичной последовательностью является совокупность возможных состояний сообщения, состоящей из двух элементов a_1 и a_2, вероятности появления которых соответственно равны P_1 и $P_2 = 1 - P_1$. При этом элементы независимы, т.е. вероятность появления очередного элемента не зависит от значений предшествующих элементов. Вероятность того, что в последовательности из n элементов будет r элементов a_1 и $n-r$ элементов a_2 определяется биномиальным законом

$$P_{r,n-r} = C_n^r P_1^r \left(1 - P_1\right)^{n-r},$$

где $C_n^r = \dfrac{n!}{r!\left(n-r\right)!}$ - число различных последовательностей, содержащих r элементов a_1 и $n-r$ элементов a_2.

При увеличении числа элементов n значения r и $n-r$ в каждой реализации последовательности будет стремиться к математическим ожиданиям $P_1 n$ и $\left(1-P_1\right)n$ и именно такие последовательности будут типичными. При $n \to \infty$ вероятность появления типичной последовательности с соотношением элементов, строго соответствующим их математическим ожиданиям, стремится к единице, а все типичные последовательности с элементами a_1 и элементами a_2 имеют одинаковую вероятность появления, равную

$$P_1^{P_1 n}\left(1-P_1\right)^{\left(1-P_1\right)n}.$$

Аналогичный результат получается и для совокупности состояний из m независимых элементов $a_1, a_2, ..., a_m$ с вероятностями их появления $P_1, P_2, ..., P_m$. В этом случае число различных типичных последовательностей длительности n равно

$$\frac{n!}{\left(P_1 n\right)!\left(P_2 n\right)!\cdots\left(P_m n\right)!}$$

и все они имеют одинаковую вероятность появления

$$P_{\text{Òèï}} = P_1^{P_1 n} P_2^{P_2 n} \cdots P_m^{P_m n}.$$

Приведенные результаты можно обобщить и на случай корреляций между элементами.

Так как последовательность сообщений источника X будет типичной, то и последовательность выходных сообщений Y также будет типичной [8], а поэтому следует ожидать, что предел (5.58) может являться некоторой характеристикой работы информационного канала, указывающей на полное среднее количество информации, получаемое на выходе канала за единицу времени.

$$V\left(Y,X\right) = \lim_{T\to\infty} \frac{I\left(Y_T, X_T\right)}{T}. \qquad (5.59)$$

$V\left(Y,X\right)$ - определяет скорость передачи информации на выходе информационного канала.

В общем случае скорость передачи информации зависит от статистических характеристик источника сообщения, метода кодирования и свойств канала. Например, при одном и том же способе кодирования длительность элементов может быть различной и зависит от полосы пропускания канала. Следовательно, различной будет и скорость передачи информации.

Пропускной способностью информационного канала называется максимальное значение скорости передачи информации при заданных ограничениях

$$C = \underset{\substack{\beta_1 \in B_1 \\ \dots \\ \beta_k \in B_k}}{Sup}\left[V\left(Y, X\right)\right], \tag{5.60}$$

где $\beta_i \in B_i$ указывает на то, параметр канала β_i удовлетворяет заданному ограничению (принадлежит некоторой области B_i).

К ограничениям можно отнести, например, длительность передаваемых элементов, используемый код, методы декодирования и т.п.

Способ определения пропускной способности информационного канала зависит от совокупности ограничений. Для полностью определенного информационного канала пропускная способность определяется статистическими характеристиками источника сообщений.

Следует отметить, что пропускная способность есть характеристика собственно информационного канала и не зависит от производительности источника (средней скорости поступления информации от источника), которая равна

$$V\left(X\right) = \lim_{T \to \infty} \frac{H\left(X_T\right)}{T}. \tag{5.61}$$

Аналогично информационному каналу пропускная способность канала связи определяется зависимостью

$$C_c = \underset{\substack{\alpha_1 \in A_1 \\ \dots \\ \alpha_i \in A_i}}{Sup}\left[V\left(Y_T^{'}, X_T^{'}\right)\right], \tag{5.62}$$

где $V\left(Y^{'}, X^{'}\right) = \lim\limits_{T \to \infty} \dfrac{I\left(Y_T^{'}, X_T^{'}\right)}{T}$ - скорость передачи информации по каналу связи, $X_T^{'}$ и $Y_T^{'}$ - сообщения, длительностью T, на входе и выходе канала связи.

Так как канал связи является частью информационного канала, то имеется возможность использования оптимальных сигналов, при которых производительность кодирующего устройства будет наибольшей

$$V\left(X^{'}\right) = \lim_{T \to \infty} \frac{I\left(X_T^{'}, X_T\right)}{T}.$$

Реальный информационный канал обладает наличием дополнительных ограничений по способу кодирования, виду модуляции и структуре канала, что приводит к недоиспользованию пропускной способности канала. Поэтому

$C_c \geq C$.

Обратного неравенства быть не может.

На основании приведенных общих соотношений, учитывая, что при отсутствии помех можно считать $Y_T^{'} = X_T^{'}$ и $Y_T = X_T$, получим формулы для определения пропускных способностей дискретного информационного канала без помех и дискретного канала связи соответственно в виде

$$C = Sup\left[\lim_{T\to\infty}\frac{H\left(X_T\right)}{T}\right], \tag{5.63}$$

$$C_c = Sup\left[\lim_{T\to\infty}\frac{H\left(X_T^{'}\right)}{T}\right], \tag{5.64}$$

так как

$$I\left(Y_T, X_T\right) = I\left(X_T, X_T\right) = H\left(X_T\right)$$

и

$$I\left(Y_T^{'}, X_T^{'}\right) = I\left(X_T^{'}, X_T^{'}\right) = H\left(X_T^{'}\right)$$

Число возможных последовательностей сообщения длительностью T будем считать равными n_T. Тогда исходя из свойства энтропии дискретных сообщений, согласно которому энтропия будет максимальна, если все последовательности равновероятны, соотношения (5.63) и (5.64) соответственно примут вид

$$C = \lim_{T\to\infty}\frac{\log n_T}{T}, \tag{5.65}$$

$$C_c = \lim_{T\to\infty}\frac{\log n_{Tc}}{T}. \tag{5.66}$$

В формуле (5.66) n_{Tc}- число всех возможных последовательностей кодированных сигналов длительностью T.

В качестве иллюстрации рассмотрим канал, в котором используется элементы с α различными состояниями. Длительность элементов одинакова и равна τ. Других ограничений нет. Для определения пропускной способности канала связи составим последовательность из m элементов. Длительность такой последовательности равна $T = m\tau$. При $T \to \infty$ число элементов в одной последовательности $m \to \infty$. Всего можно образовать α^m последовательностей длиной в m элементов.

Следовательно, $n_{Tc} = \alpha^m$. Таким образом, пропускная способность канала связи C_c согласно (5.66) равна

$$C_c = \lim_{m\to\infty} \frac{\log \alpha^m}{m\tau} = \frac{\log \alpha}{\tau}. \qquad (5.67)$$

При использовании двоичного кода ($\alpha = 2$ - два состояния) получим $C_c = \frac{1}{\tau} \ \frac{\text{äâ.åä}}{\text{ñåê}}$.

Единица пропускной способности дискретного канала $\frac{\text{äâ.åä}}{\text{ñåê}}$ носит название áî ä.

Пример. Сообщение состоит из четырех состояний x_1, x_2, x_3, x_4 с вероятностями $P(x_1) = P(x_2) = P(x_3) = P(x_4) = 0{,}25$ и используется двоичный код. Сообщение кодируется по схеме:

$$x_1 \to 00,\ x_2 \to 01,\ x_3 \to 10,\ x_4 \to 11.$$

Длительность элементов двоичного кода одинакова и равна 2 мкс. Определить пропускную способность такого канала связи.

Решение. Так как передаваемые состояния равновероятны, то в длинной последовательности сообщений число нулей и единиц будет одинаково. Следовательно, вероятности передачи 0 и 1 равны. Отсюда следует, что

$$C_c = \frac{1}{\tau} = 5 \cdot 10^5 \textit{ бод.}$$

Рассмотрим производительность кодирующего устройства.

Символы кода имеют одинаковую длительность, равную τ. Длительность сигнала на выходе кодирующего устройства равна $T = m\tau$ (последовательность из m элементов).

Энтропия такого сигнала будет равна

$$H\left(X_T^{'}\right) = mH\left(X^{'}\right).$$

Тогда производительность кодирующего устройства определится зависимостью

$$V\left(X^{'}\right) = \lim_{T\to\infty} \frac{I\left(X_T^{'}, X\right)}{T} = \lim_{m\to\infty} \frac{mH\left(X^{'}\right)}{m\tau} = \frac{H\left(X^{'}\right)}{\tau}.$$

Естественно, что при такой производительности кодирующего устройства пропускная способность информационного канала будет определяться величиной

$$C = \frac{H_{\max}\left(X^{'}\right)}{\tau} \le C_c ,$$

где $H_{\max}\left(X^{'}\right)$ - максимально возможное значение энтропии кодированного сигнала с учетом накладываемых ограничений.

Пример. Сообщение состоит из x_1, x_2, x_3, x_4 с вероятностями: $P\left(x_1\right) = 0,5$, $P\left(x_{1=2}\right) = 0,25$, $P\left(x_3\right) = P\left(x_4\right) = 0,125$. Длительность элементов кодирующего сигнала одинакова и равна $\tau = 2$ мкс. Сообщение кодируется по схеме

$$x_1 \to 00,\ x_2 \to 01,\ x_3 \to 10,\ x_4 \to 11.$$

Определить производительность кодирующего устройства.

Решение. Каждая двоичная комбинация, соответствующая состояниям x_1, x_2, x_3, x_4 будет передаваться с указанными вероятностями. При этом, длительность одной двоичной комбинации равна 2τ. Следовательно, получим

$$V\left(X^{'}\right) = \frac{H\left(X^{'}\right)}{2\tau} = -\frac{\sum_{i=1}^{4} P\left(x_i\right)\log P\left(x_i\right)}{2\tau} = 0,437 \cdot 10^6 \, \frac{\text{äâ.åä}}{\text{ñåê}},$$

$$V\left(X^{'}\right) < C_c .$$

Из приведенного примера видно, что отклонение от оптимальных статистических характеристик приводит к тому, что фактическая скорость передачи информации меньше пропускной способности канала связи.

Рассмотрим также передачу сообщения, когда состояниям элементов $\alpha_1, \alpha_2, ..., \alpha_k$ соответствуют длительности $\tau_1, \tau_2, ..., \tau_k$. Анализ такого случая показывает, что на достаточно большом интервале T число возможных последовательностей равно [9]

$$n_{Tc} = A r_1^{\frac{T}{\tau_m}} ,$$

где: A - некоторая постоянная величина,

τ_m - интервал, равный минимальной длительности элемента,

r_1 - находится как наибольший действительный положительный корень уравнения

$$1 - r^{\frac{\tau_1}{\tau_m}} - r^{\frac{\tau_2}{\tau_m}} - \cdots - r^{\frac{\tau_k}{\tau_m}} = 0 .$$

Следовательно, пропускная способность канала связи будет равна

$$C_c = \frac{\log r_1}{\tau_m}.$$

5.2.3. Основная теорема Шеннона для дискретного канала без помех

Фактическая скорость передачи информации зависит от согласованности статистических характеристик источника сообщений со свойствами информационного канала. Согласование характеристик источника сообщений со свойствами информационного канала может быть осуществлена посредством выбора способа кодирования и декодирования сообщений. Кодирование сообщения, при котором достигается наилучшее использование пропускной способности канала связи, т.е. наибольшая скорость передачи информации, называется эффективным.

Ответ на вопрос о том, в какой мере скорость передачи информации может быть приближена к пропускной способности информационного канала, дает основная теорема Шеннона для дискретного канала без помех.

Теорема Шеннона: *если производительность источника сообщений меньше пропускной способности канала связи без помех, то всегда можно найти такой способ кодирования, который обеспечит передачу информации со средней скоростью, сколь угодно близкой к пропускной способности канала связи.*

Передавать информацию со средней скоростью большей пропускной способности канала связи невозможно.

Таким образом:

1.) Если $V(X) < C_c$, то всегда можно найти такой способ кодирования, при котором

$$V(Y,X) = C_c - \varepsilon,$$

где ε - сколь угодно малая величина.

2.) Передать информацию при $V(Y,X) > C_c$ невозможно.

Доказательство теоремы.

Число типичных последовательностей из m элементов достаточно большой длительности $m\tau = T$, создаваемых источником сообщений X, можно выразить через энтропию источника $H(X)$ в виде

$$n_T = 2^{mH(X)}. \tag{5.68}$$

Так как при $m \to \infty$ источник с вероятностью, близкой к единице, формирует лишь типичные последовательности, имеющие равную вероятность $1/n_T$, то энтропия такого источника равна $\log n_T$, а суммарное количество информации, содержащееся в m элементах источника с энтропией $H(X)$ при $m \to \infty$ равна $mH(X)$. Отсюда

$$mH(X) = \log n_T.$$

Следовательно, приходим к соотношению (5.68).
Учитывая, что производительность источника

$$V(X) = \frac{H(X)}{\tau},$$

выражение (5.68) принимает вид

$$n_T = 2^{TV(X)}.$$

Если эти последовательности с достаточно большой протяженностью T будем кодировать цифровыми комбинациями X оптимальным способом, то число различных кодовых комбинаций X' будет

$$n_{Tc} = 2^{TC_c}.$$

Теперь условие $V(X < c_c)$ запишем в виде $C_c = V(X) + \varepsilon$, где ε - сколь угодно малая величина.

Тогда

$$\frac{n_{Tc}}{n_T} = 2^{[TC_c - TV(X)]} = 2^{T\varepsilon}.$$

Если принять

$$\varepsilon > \frac{\log e}{Tn_T}, \text{ то}$$

$$\frac{n_{Tc}}{n_T} > e^{1/n_T} = 1 + \frac{1}{n_T} + \frac{1}{2!n_T^2} + \cdots$$

или

$$n_{Tc} > n_T + 1.$$

Таким образом, при выполнении условия $V(X) < C_c$ число различных кодовых комбинаций X', по крайней мере, на одну больше числа типичных последовательностей источника. Эту избыточную кодовую комбинацию поставим в соответствие всем истинным последовательностям, предопределив их недостоверную передачу. Так как при $T \to \infty$ и $m \to \infty$ вероятность появления нетипичной последовательности стремится к нулю, а величина ε, определяющая требуемое превышение пропускной способности канала над производительностью источника, бесконечно малая, то первую часть теоремы можно считать доказанной.

Для второй части теоремы при $V(X) > C_c$ получим неравенство

$n_T > n_{Tc} + 1$.

Таким образом, даже при оптимальном кодировании, обеспечивающем предельную скорость передачи информации по каналу уже невозможно закодировать и передать все типичные последовательности n_T.

Оптимальное кодирование для дискретного канала без помех сводится к предельному укрупнению всей совокупности возможных состояний канала. При этом одновременно устраняется корреляция между элементами укрупненной совокупности возможных состояний и благодаря типичных последовательностей обеспечивается равная вероятность появления элементов. В результате устраняется избыточность сообщения, передаваемого по каналу.

В качестве иллюстрации основной теоремы Шеннона для дискретного канала без шумов рассмотрим пример, в котором используется способ кодирования Шеннона – Фано.

Сущность этого способа кодирования заключается в том, что все состояния сообщения выписываются в порядке убывания вероятностей их появления. Затем производится последовательное деление на подгруппы ❶ и ❷ из условия возможного равенства сумм вероятностей появления состояний в подгруппах ❶ и ❷. Номер подгруппы, в которую попадает данное состояние при каждом делении, определяет символ (элемент) на соответствующей позиции записи кода.

Если, например, сообщение состоит из состояний x_1, x_2, x_3, x_4 с вероятностями $P(x_1) = 0{,}5$, $P(x_2) = 0{,}25$, $P(x_3) = P(x_4) = 0{,}125$ и с длительностью каждого $\tau = 2$ мкс, то построение кода для этого сообщения определяется согласно табл. 5.1

Таблица 5.1

<table>
<tr><th rowspan="3">Состояние</th><th rowspan="3">Вероятность</th><th colspan="3" rowspan="2">Номера деления на подгруппы</th><th colspan="3">Символы кода</th><th rowspan="3">Длительность сигнала</th></tr>
<tr><th colspan="3">Позиции</th></tr>
<tr><th>1</th><th>2</th><th>3</th><th>1</th><th>2</th><th>3</th></tr>
<tr><td>x_1</td><td>0,5</td><td>❶</td><td></td><td></td><td>0</td><td></td><td></td><td>τ</td></tr>
<tr><td>x_2</td><td>0,25</td><td rowspan="3">❷</td><td>❶</td><td></td><td>1</td><td>0</td><td></td><td>2τ</td></tr>
<tr><td>x_3</td><td>0,125</td><td rowspan="2">❷</td><td>❶</td><td>1</td><td>1</td><td>0</td><td>3τ</td></tr>
<tr><td>x_4</td><td>0,125</td><td>❷</td><td>1</td><td>1</td><td>1</td><td>3τ</td></tr>
</table>

Подгруппе ❶ соответствует символ кода «0», а подгруппе ❷ - «1».

Как следует из табл. 5.1 состоянию x_1 соответствует кодовый сигнал «0», x_2 - «10», x_3 - «110», x_4 - «111».

Из таблицы видно, что полученный код является неравномерным, так как сигналы разных состояний могут иметь различное число символов, а, следовательно, и различную длительность. Средняя длительность сигнала, затрачиваемая на передачу одного состояния, в данном примере определяется выражением:

$$\tau_c = \tau P\left(x_1\right) + 2\tau P\left(x_2\right) + 3\tau P\left(x_3\right) + 3\tau P\left(x_4\right) = 3{,}5 \cdot 10^{-6}\, c\,.$$

Скорость передачи информации

$$V\left(Y,X\right) = \frac{H\left(X\right)}{\tau_c} = \frac{-\sum_{i=1}^{4} P\left(x_i\right) \log P\left(x_i\right)}{\tau_c} = \frac{1{,}75}{3{,}5} = 0{,}5 \cdot 10^6 \ \frac{\ddot{a}\hat{a}.\mathring{a}\ddot{a}}{\tilde{n}\mathring{a}\hat{e}}\,.$$

Пропускная способность такого канала

$$\tilde{N}_{\tilde{n}} = 0{,}5 \cdot 10^6 \ \frac{\ddot{a}\hat{a}.\mathring{a}\ddot{a}}{\tilde{n}\mathring{a}\hat{e}}\,.$$

Как видно из приведенного примера, используя код Шеннона – Фано, удалось полностью согласовать статистические характеристики источника со свойствами канала. Именно этим примером К.Шеннон проиллюстрировал свою основную теорему [10]. Однако это не всегда удается.

Следует обратить внимание на то, что приближение скорости передачи информации к пропускной способности канала связи при кодировании и декодировании связано с задержкой передачи сообщения. Величина этой задержки достигает двойного значения длительности сообщения.

Однако теорема Шеннона играет важную роль, как предельная теорема, которая послужила толчком к развитию прикладной теории кодирования.

Отметим также, что надежность отождествления передаваемых сообщений в дискретных информационных каналах без помех определяется только операцией кодирования, так как нарушение соответствия передаваемого и принимаемого сообщения в каналах связи исключается.

5.2.4. Дискретные каналы с помехами

Функциональная схема дискретного информационного канала с помехами представлена на Рис.5.2 и отличается от схемы, приведенной на Рис.5.1 наличием источника помех (ИП).

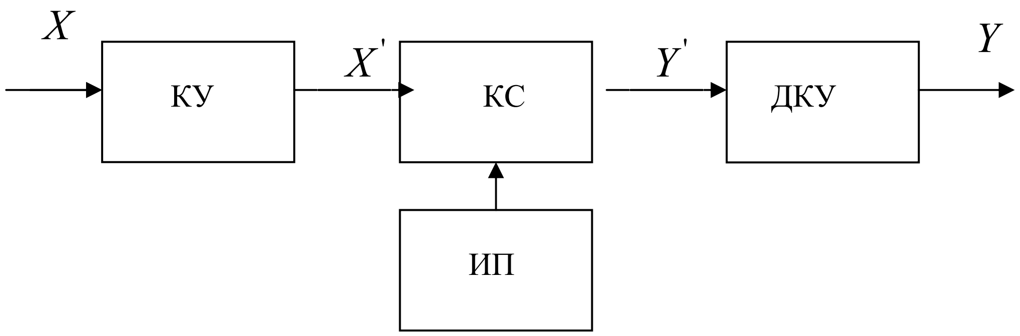

Рис.5.2

Функциональная схема дискретного информационного канала с помехами.

К помехам относятся естественные и искусственные помехи. А также собственные шумы канала связи, например, для открытых оптических систем связи естественной помехой является фоновая засветка, собственными шумами канала связи – темновые шумы фотоприемника, тепловые шумы нагрузки и шумы входных каскадов электронной схемы. Для радиотехнических систем легко создаются искусственные помехи.

Воздействие помех приводит к искажению передаваемых информационных параметров (состояний) сигналов. В результате чего нарушается взаимно – однозначное соответствие между параметрами сигналов на выходе и входе канала связи. При передаче параметра $x_i^{'}$ параметры сигнала на выходе канала связи $y_i^{'}$ могут принимать различные значения из–за воздействия помехи. В общем случае параметр $y_i^{'}$ может принимать множество различных значений и принятие решения в этом случае основывается на выбранном статистическом критерии качества.

Основные соотношения для дискретного канала с помехой в теории информации выведены из условия равномерно ограниченной ошибки при принятии решения о передаваемом параметре $x_i^{'}$. В этом случае вероятность правильного принятия решения $P\left(y_i^{'}\middle|x_i^{'}\right)$, т.е. условная вероятность того, что при передачи параметра $x_i^{'}$ будет принято решение $y_i^{'}$ справедливо соотношение [8]

$$P\left(y_i^{'}\middle|x_i^{'}\right) \geq P_{\text{ïð}},$$

где $P_{\text{ïð}}$- вероятность правильного принятия решения для всех параметров (состояний) сообщения.

5.2.5. Пропускная способность канала с помехами

Соотношения (5.59), (5.60), (5.62), (5.64), приведенные в разделе 5.2.2, являются общими и справедливы для дискретных каналов с помехой. Разница между каналами без помех и с помехами заключается только в способе вычисления полного среднего количества информации, содержащегося в последовательности выходных сигналов $Y_T^{'}$ относительно входных $X_T^{'}$.

Для вычисления $I\left(Y_T^{'}, X_T^{'}\right)$ воспользуемся формулами (5.35) и (5.36), в соответствии с которыми получим

$$I\left(Y_T^{'}, X_T^{'}\right) = H\left(Y_T^{'}\right) - H\left(Y_T^{'}\middle|X_T^{'}\right) = H\left(X_T^{'}\right) - H\left(X_T^{'}\middle|Y_T^{'}\right), \qquad (5.69)$$

где $X_T^{'}$ и $Y_T^{'}$ - сигналы длительностью T на входе и выходе канала связи.

Будем, при этом, исходить из предположения, что помеха в канале связи имеет эргодический характер, а, следовательно, при длительной передачи сигнала образует типичную последовательность. Будем также рассматривать только информационные каналы без памяти или постоянные каналы [11], у которых на каждый передаваемый элемент помеха воздействует независимо от того, какие элементы передавались ранее, т.е. помеха не вызывает дополнительных корреляционных связей между элементами.

Скорость передачи информации по каналу связи в этом случае в соответствии с (5.59) и с учетом (5.69) можно представить в виде

$$V\left(Y^{'}, X^{'}\right) = \lim_{T\to\infty} \frac{I\left(Y_T^{'}, X_T^{'}\right)}{T} =$$

$$= \lim_{T\to\infty} \frac{H\left(Y_T^{'}\right) - H\left(Y_T^{'}\middle|X_T^{'}\right)}{T} = \lim_{T\to\infty} \frac{H\left(X_T^{'}\right) - H\left(X_T^{'}\middle|Y_T^{'}\right)}{T}. \qquad (5.70)$$

Для последовательностей длительностью T, содержащей m элементов, имеем:

$$H\left(Y_T^{'}\right)=mH\left(Y^{'}\right),\quad H\left(X_T^{'}\right)=mH\left(X^{'}\right),$$

$$H\left(Y_T^{'}\middle|X_T^{'}\right)=mH\left(Y^{'}\middle|X^{'}\right),\ H\left(X_T^{'}\middle|Y_T^{'}\right)=mH\left(X^{'}\middle|Y^{'}\right).$$

Учитывая, что $T=m\tau_c$, где τ_c - средняя длительность элемента, из (5.70) получим:

$$V\left(Y^{'},X^{'}\right)=\lim_{T\to\infty}\frac{H\left(Y_T^{'}\right)-H\left(Y_T^{'}\middle|X_T^{'}\right)}{T}=$$

$$=\lim_{T\to\infty}\frac{m\left[H\left(Y^{'}\right)-H\left(Y^{'}\middle|X^{'}\right)\right]}{m\tau_c}=\frac{H\left(Y^{'}\right)}{\tau_c}-\frac{H\left(Y^{'}\middle|X^{'}\right)}{\tau_c}=\quad. \qquad (5.71)$$

$$=\frac{H\left(X^{'}\right)}{\tau_c}-\frac{H\left(X^{'}\middle|Y^{'}\right)}{\tau_c}=V\left(X^{'}\right)-V\left(X^{'}\middle|Y^{'}\right)$$

Или

$$V\left(Y^{'},X^{/}\right)=V\left(Y^{'}\right)-V\left(Y^{'}\middle|X^{'}\right), \qquad (5.72)$$

где: $V\left(X^{'}\right)$ - производительность кодирующего устройства;

$V\left(Y^{'}\right)$ - поток информации на выходе канала связи;

$V\left(X^{'}\middle|Y^{'}\right)=\frac{H\left(X^{'}\middle|Y^{'}\right)}{\tau_c}$ - апостериорная энтропия за единицу времени, которая носит название *ненадежность канала*.

Из соотношения (5.64) следует, что пропускная способность канала связи с помехой будет определяться зависимостями:

$$C_c=Sup\left[V\left(X^{'}\right)-V\left(X^{'}\middle|Y^{'}\right)\right], \qquad (5.73)$$

$$C_c=Sup\left[V\left(Y^{'}\right)-V\left(Y^{'}\middle|X^{'}\right)\right]. \qquad (5.74)$$

Зависимости (5.73) и (5.74) равноправны и дают одно и тоже значение C_c, а использование каждого из них определяется удобством проводимого анализа.

Как видно из соотношения (5.73) пропускная способность канала с помехами определяется как предельное количество информации, которое достигает выхода канала в единицу времени, если на вход в единицу времени поступает максимальное количество информации, которое только может быть воспринято каналом при отсутствии помех.

Пропускная способность дискретного канала с помехами, как видно из (5.73) зависит от ненадежности канала $V\left(X^{'}\middle|Y^{'}\right)$, характеризующей средние потери информации, приходящейся на один элемент, из-за действия помехи. При этом ненадежность канала согласно (5.23) может быть представлена в виде

$$V\left(X^{'}\middle|Y^{'}\right)=-\frac{1}{\tau}\sum_{i=1}^{n}\sum_{k=1}^{m}P\left(y_i^{'}\right)P\left(x_k^{'}\middle|y_i^{'}\right)\log P\left(x_k^{'}\middle|y_i^{'}\right).$$

Учитывая, что

$$P\left(y_i^{'}\right)P\left(x_k^{'}\middle|y_i^{'}\right)=P\left(x_i^{'}\right)P\left(y_k^{'}\middle|x_i^{'}\right) \quad \text{и} \quad P\left(y_i^{'}\right)=\sum_{k=1}^{m}P\left(x_k^{'}\right)P\left(y_i^{'}\middle|x_k^{'}\right),$$

получим

$$V\left(X^{'}\middle|Y^{'}\right)=-\frac{1}{\tau}\sum_{i=1}^{n}\sum_{k=1}^{m}P\left(x_k^{'}\right)P\left(y_i^{'}\middle|x_k^{'}\right)\log\frac{P\left(x_k^{'}\right)P\left(y_i^{'}\middle|x_k^{'}\right)}{\sum_{k=1}^{m}P\left(x_k^{'}\right)P\left(y_i^{'}\middle|x_k^{'}\right)}. \qquad (5.75)$$

Из приведенного выражения (5.75) видно, что ненадежность канала зависит как от статистических характеристик входного сообщения $X^{'}$, так и от вероятностных характеристик искажения элементов сообщения из-за помехи. Эти вероятностные характеристики могут быть заданы матрицей условных вероятностей:

$$\begin{vmatrix} P\left(y_1^{'}\middle|x_1^{'}\right) & P\left(y_1^{'}\middle|x_2^{'}\right) & \ldots & P\left(y_1^{'}\middle|x_m^{'}\right) \\ P\left(y_2^{'}\middle|x_1^{'}\right) & P\left(y_2^{'}\middle|x_2^{'}\right) & \ldots & P\left(y_2^{'}\middle|x_m^{'}\right) \\ \ldots & \ldots & \ldots & \ldots \\ P\left(y_n^{'}\middle|x_1^{'}\right) & P\left(y_n^{'}\middle|x_2^{'}\right) & \ldots & P\left(y_n^{'}\middle|x_m^{'}\right) \end{vmatrix}.$$

При отсутствии помех ненадежность канала равна нулю

$$V\left(X^{'}\middle|Y^{'}\right)=0.$$

В этом случае пропускная способность канала будет равна

$$C_c=Sup V\left(X^{'}\right)=Sup\frac{H\left(X^{'}\right)}{\tau_c},$$

т.е. пропускной способности канала без помех.

В том случае, когда помеха настолько велика, что сообщения на входе $X^{'}$ и выходе $Y^{'}$ канала связи становятся статистически независимыми, тогда $H\left(X^{'}\middle|Y^{'}\right)=H\left(X^{'}\right)$, а значит и $V\left(X^{'}\middle|Y^{'}\right)=V\left(X^{'}\right)$. Следовательно, пропускная способность такого канала будет минимальной и равна нулю, т.е. $C_c=0$.

Таким образом, в зависимости от уровня помехи пропускная способность дискретного канала с помехами может меняться с учетом (5.67) в пределах

$$0\le C_c\le\frac{1}{\tau}\log\alpha . \tag{5.76}$$

Пример. По бинарному каналу связи без фиксированных ограничений передаются два элемента $x_1^{'}$ и $x_2^{'}$ с одинаковой длительностью τ и вероятностями $P\left(x_1^{'}\right)=P\left(x_2^{'}\right)=0,5$. На приемной стороне решающее устройство разделяет принимаемые элементы в области $Y_1^{'}$ и $Y_2^{'}$ так, что, если элемент попадает в область $Y_1^{'}$, то принимается решение о том, что передан элемент $x_1^{'}$, а если - в область $Y_2^{'}$, то – элемент $x_2^{'}$. Матрица условных вероятностей $P\left(Y^{'}\middle|X^{'}\right)$ имеет вид

$$\begin{vmatrix} P\left(y_1^{'}\middle|x_1^{'}\right)=P_{11} & P\left(y_1^{'}\middle|x_2^{'}\right)=P_{12} \\ P\left(y_2^{'}\middle|x_1^{'}\right)=P_{21} & P\left(y_2^{'}\middle|x_2^{'}\right)=P_{22} \end{vmatrix},$$

$P\left(y_1^{'}\middle|x_1^{'}\right)=P_{11}$ и $P\left(y_2^{'}\middle|x_2^{'}\right)=P_{22}$- вероятности правильного принятия решения,
$P\left(y_2^{'}\middle|x_1^{'}\right)=P_{21}$ и $P\left(y_1^{'}\middle|x_2^{'}\right)=P_{12}$- вероятности ошибочных решений.

Естественно, что $P\left(y_1^{'}\middle|x_1^{'}\right)=1-P_{21}$ и $P\left(y_2^{'}\middle|x_2^{'}\right)=1-P_{12}$.

Определить пропускную способность канала связи.

Решение. В соответствии с формулами (5.73) и (5.75) запишем

$$C_c = Sup\left[V\left(X^{'}\right) - V\left(X^{'}\middle|Y^{'}\right)\right] = \frac{H\left(X^{'}\right)}{\tau} - \frac{H\left(X^{'}\middle|Y^{'}\right)}{\tau} =$$

$$= \frac{1}{\tau}\log 2 + \frac{1}{\tau}\sum_{i=1}^{n}\sum_{k=1}^{m} P\left(x_k^{'}\right)P\left(y_i^{'}\middle|x_k^{'}\right)\log\frac{P\left(x_k^{'}\right)P\left(y_i^{'}\middle|x_k^{'}\right)}{\sum\limits_{k=1}^{m} P\left(x_k^{'}\right)P\left(y_i^{'}\middle|x_k^{'}\right)} =$$

$$= \frac{1}{\tau}\left[1 + \sum_{i=1}^{n}\sum_{k=1}^{m} P\left(x_k^{'}\right)P\left(y_i^{'}\middle|x_k^{'}\right)\log\frac{P\left(x_k^{'}\right)P\left(y_i^{'}\middle|x_k^{'}\right)}{\sum\limits_{k=1}^{m} P\left(x_k^{'}\right)P\left(y_i^{'}\middle|x_k^{'}\right)}\right] =$$

$$= 0,5\left[2 + P_{11}\log\frac{P_{11}}{P_{11}+P_{12}} + P_{12}\log\frac{P_{12}}{P_{11}+P_{12}} + \right.$$

$$\left. + P_{21}\log\frac{P_{21}}{P_{21}+P_{22}} + P_{22}\log\frac{P_{22}}{P_{21}+P_{22}}\right] =$$

$$= \frac{0,5}{\tau}\left[2 + \left(1-P_{21}\right)\frac{1-P_{21}}{1-P_{21}+P_{12}} + P_{12}\frac{P_{12}}{1-P_{21}+P_{12}} + \right.$$

$$\left. + P_{21}\frac{1-P_{21}}{1-P_{12}+P_{21}} + \left(1-P_{12}\right)\frac{1-P_{12}}{1-P_{12}+P_{21}}\right]$$

Проанализируем полученное выражение в зависимости от вероятностей ошибочных решений.

А.) Вероятности ошибочных решений одинаковы при передаче элементов $x_1^{'}$ и $x_2^{'}$, т.е. $P_{12} = P_{21} = P_{\text{îø}}$.

Результаты расчетов пропускной способности канала связи приведены в виде графика функции $C_c P_{\text{îø}} \tau$ на Рис 5.3 ($C_c P_{\text{îø}} \tau$ - относительная пропускная способность канала связи).

Из графика видно, что при отсутствии помехи $P_{\text{îø}} = 0$ пропускная способность достигает максимального значения, равного значению C_c канала без помех. С ростом $P_{\text{îø}}$ пропускная способность канала падает и достигает минимального значения, равного нулю, при $P_{\text{îø}} = 0,5$. И это естественно, так как при передаче каждого из элементов $x_1^{'}$ и $x_2^{'}$ на выходе канала с равной вероятностью может быть принято решение $y_1^{'}$ и $y_2^{'}$. С дальнейшим ростом вероятности ошибочных решений пропускная способность канала возрастает. Эта парадоксальная ситуация объясняется тем, что изменяется правило распознавания элементов на обратное, т.е. при $P_{\text{ош}} > 0,5$ следует считать решение $y_1^{'}$ соответствует передаче элемента $x_2^{'}$, а решение $y_2^{'}$ - передаче $x_1^{'}$.

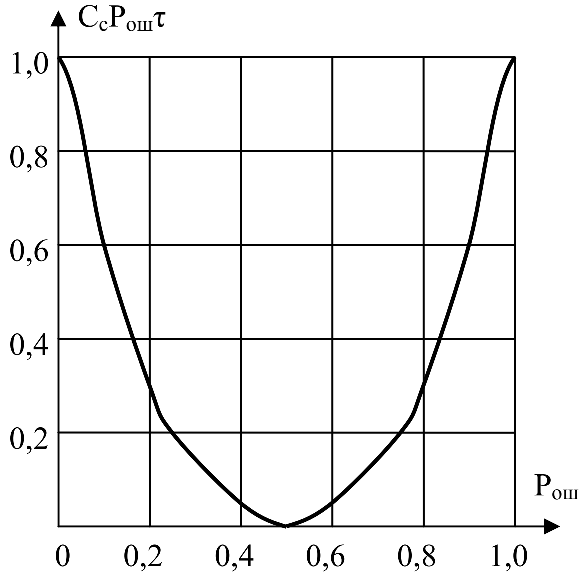

Рис 5.3

Относительная пропускная способность при одинаковых вероятностях ошибочных решений решений

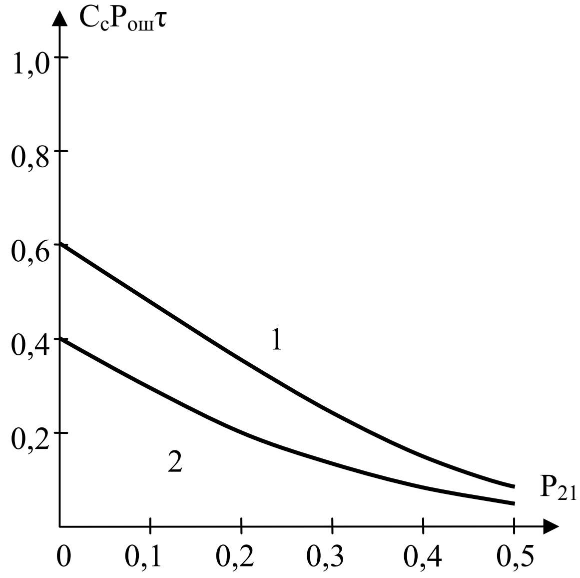

Рис 5.4

Относительная пропускная способность при различных вероятностях ошибочных

Б.) Вероятности ошибочных решений различны при передаче элементов $x_1^{'}$ и $x_2^{'}$, т.е. $P_{12} \neq P_{21}$.

Результаты расчетов представлены на Рис.5.4 в виде графиков функций $C_c P_{\text{ош}} \tau = C_c P_{21} \tau$ (кривая 1 при $P_{12} = 0,2$ и кривая 2 при $P_{12} = 0,4$).

Из графиков можно проследить тенденцию изменения пропускной способности канала при увеличении вероятностей ошибок принятия решений.

5.2.6. Основная теорема Шеннона для дискретного канала с помехами

К. Шеннон доказал, что пропускная способность дискретного канала с помехами определяет верхнюю границу скорости достоверной передачи информации.

Основную **теорему Шеннона** для дискретных каналов с помехами можно сформулировать в следующем виде:

1. *Если пропускная способность дискретного канала с помехой превышает производительность источника сообщений, то существует такой способ кодирования, при котором может быть обеспечена безошибочная передача всей информации.*
2. *Осуществить безошибочную передачу информации невозможно, если производительность источника сообщений превышает пропускную способность канала.*

Таким образом, если $C_c > V\left(X\right)$, то существует такой способ кодирования, при котором $P_{\text{îø}} < \eta$, где η- сколь угодно малая положительная величина.

Безошибочная передача невозможна при $C_c < V\left(X\right)$.

Доказательство теоремы.

Число типичных последовательностей из m элементов достаточно большой длительности $T = m\tau$, создаваемых источником сообщений X производительностью $V\left(X\right)$ можно представить (по аналогии с теоремой для канала без помех) в виде

$$n_T = 2^{TV(X)}.$$

Для кодирования этих последовательностей используются коды той же длительности T. Тогда общее число различных кодов можно представить как

$$n_{Tc} = 2^{TC_c}.$$

Запишем условие $C_c > V\left(X\right)$ в развернутом виде

$$Sup\left[V\left(X'\right) - V\left(X'\middle|Y'\right)\right] = \frac{1}{\tau}\log\alpha - V\left(X'\middle|Y'\right) > V\left(X\right). \quad (5.77)$$

Так как $V\left(X'\middle|Y'\right) > 0$, то неравенство (5.77) только усилится, если его записать в виде

$$\frac{1}{\tau}\log\alpha > V\left(X\right).$$

Тогда

$$n_{Tc} = 2^{m \log \alpha} >> 2^{TV(X)} = n_T.$$

Это означает, что число различных кодов n_{Tc}, которые могут быть использованы для кодирования, много больше типичных последовательностей n_T, подлежащих кодированию. Необходимо доказать, что среди всех способов кодирования имеется, по крайней мере. Один, обеспечивающий безошибочную передачу сообщения. Найдем вероятность правильного принятия решения, осредненную по всем возможным способам кодирования [12]. При ненадежности канала $V\left(X^{'}\middle|Y^{'}\right)$ число типичных входных кодовых последовательностей, которые могут трансформироваться в выходное сообщение длительности T, (по аналогии с теоремой раздела 5.2.3) можно записать в виде

$$n_{Tc} = 2^{TV\left(X^{'}\middle|Y^{'}\right)}.$$

Правильное принятие решения обеспечивается в этом случае, когда среди n_{T1} входных последовательностей, которые могли бы дать данную выходную последовательность, лишь одна была бы использована при кодировании, а остальные $\left[n_{T1} - 1\right]$ последовательности вообще не могли передаваться.

Определим по всем возможным способам кодирования вероятность P_n того. Что ни одна из последовательностей $\left[n_{T1} - 1\right]$ не использовалась при кодировании. Средняя для всех возможных способов кодирования последовательностей n_T вероятность использования конкретной кодовой комбинации из n_{Tc} возможных равна

$$P_T = \frac{n_T}{n_{Tc}},$$

так как усреднению по всем возможным способам кодирования соответствует равновероятный выбор кодовых комбинаций для кодирования каждой входной последовательности.

Средняя вероятность того, что конкретная кодовая комбинация не была использована при кодировании, равна

$$1 - P_T = 1 - \frac{n_T}{n_{Tc}},$$

а средняя вероятность того, что не использовались конкретные кодовые комбинации $\left[n_{T1} - 1\right]$, определяющая среднюю вероятность правильного решения, будет определяться соотношением

$$P_n = \left(1 - P_T\right)^{\left[n_{T1}-1\right]} = \left(1 - \frac{n_T}{n_{Tc}}\right)^{\left[n_{T1}-1\right]}. \qquad (5.78)$$

Так как $\left(1 - \dfrac{n_T}{n_{Tc}}\right) < 1$, то увеличивая показатель степени до значения n_{T1}, из (5.78) приходим к неравенству

$$P_n > \left(1 - \frac{n_T}{n_{Tc}}\right)^{n_{T1}}.$$

Разложим правую часть этого неравенства в ряд:

$$P_n > 1 - n_{T1}\frac{n_T}{n_{Tc}} + \frac{n_{T1}\left(n_{T1}-1\right)}{2}\left(\frac{n_T}{n_{Tc}}\right)^2 - \cdots . \qquad (5.79)$$

Так как $\dfrac{n_T}{n_{Tc}} << 1$, то правая часть неравенства (5.79) только уменьшится, если оставить лишь первый член разложения. Тогда получим

$$P_n > 1 - n_{T1}\frac{n_T}{n_{Tc}}.$$

Используя выражения для n_T, n_{Tc} и n_{T1}, приходим к следующей записи

$$P_{\hat{\imath}\o} = 1 - P_n < 2^{-T\left[\log\alpha - V\left(X'\middle|Y'\right) - V(X)\right]} = 2^{-T\left[C_c - V(X)\right]}. \qquad (5.80)$$

При $T \to \infty$ в соответствии с (5.80) $P_{\hat{\imath}\o} \to 0$. Таким образом, при любой заданной положительной величине $\eta > 0$ можно выбрать такое T, при котором обеспечивается $P_{\hat{\imath}\o} < \eta$.

Обратное утверждение вытекает из $V\left(X\right) > C_c$. В этом случае из неравенства следует, что не все типичные последовательности сообщений могут быть закодированы таким образом, чтобы они опознавались на выходе канала связи с приемлемо малой ошибкой.

Теорема Шеннона для канала с помехами не указывает на конкретный способ кодирования, а только доказывает теоретическую возможность существования такого кода.

Из формулы (5.80) вытекает одно важное практическое следствие о том, что вероятность правильного принятия решения будет тем больше, чем длительнее кодированная последовательность и больше запас по величине $\left[C_c - V\left(X\right)\right]$. Формулируя иначе, можно сказать, что чем больше избыточность входного сообщения

$$r = 1 - \frac{H\left(X\right)}{\log \alpha},$$

тем более достоверную передачу можно обеспечить.

Таким образом, для обеспечения достоверной передачи информации достаточно ввести во входное сообщение избыточность, превышающую ненадежность канала.

Следует также отметить, что удлинение кодируемых последовательностей вызывает не снижение пропускной способности, а лишь увеличение задержки в приеме информации.

В качестве иллюстрации рассмотрим простой пример эффективного кодирования для дискретного канала с помехой [10].

Пример эффективного кодирования. По дискретному каналу с помехой передаются группы из 7 элементов в двоичном коде (0,1). В результате воздействия помехи группа из семи элементов передается либо без ошибки, либо в ней оказывается ошибочным один элемент из семи. При этом вероятности правильных и ошибочных решений одинаковы и, следовательно, равны $1/8$. Необходимо составить код, обеспечивающий безошибочную передачу информации.

В данном случае пропускная способность канала согласно (5.75) будет равна $C_c = \frac{4}{7}\ \frac{\text{äâ.åä.}}{\text{ýëåì åí ò}}$.

Эффективный код, обеспечивающий безошибочную передачу с вычисленной пропускной способностью, будет следующим.

Положим, в группу входит семь элементов $x_1', x_2', ..., x_7'$, из них x_3', x_5', x_6', x_7' являются элементами сообщения, в которых заложена передаваемая информация. Остальные три элемента являются избыточными и выбираются из условия:

x_1' выбирается так, чтобы $\alpha = x_1' + x_3' + x_5' + x_7'$ было четным,

x_2' выбирается так, чтобы $\beta = x_2' + x_3' + x_6' + x_7'$ было четным,

x_4' выбирается так, чтобы $\gamma = x_4' + x_5' + x_6' + x_7'$ было четным.

Когда группа из семи элементов принята, то вычисляются α, β, γ. Если они окажутся четными, то это означает, что каждая из них соответствует нулю, а если – нечетными, то – единице. Двоичные цифры

α, β, γ дадут тогда индексы тех $x_i^{'}$, которые являются ошибочными. Если получится «0», то это означает отсутствие ошибок.

5.2.7. Непрерывные информационные каналы. Передача информации через линейные и нелинейные устройства

Непрерывные информационные каналы обеспечивают передачу непрерывных сообщений. Функциональная схема непрерывного информационного канала аналогична схеме Рис. 5.2 и может отличаться тем, что вместо кодирующих и декодирующих устройств могут использоваться преобразователи сообщения в сигнал и сигнала в сообщение. Для передачи сигналов используются различные методы модуляции, детектирования и каналы связи. Как в преобразователях, так и в каналах связи непрерывные сигналы проходят через линейные и нелинейные устройства.

В данном разделе ограничимся рассмотрением достаточно важного случая передачи непрерывной информации через линейные и нелинейные устройства при отсутствии помех. В этой ситуации количество информации можно характеризовать приведенной энтропией.

Как указывалось в разделе 5.1.3, приведенная энтропия непрерывных сообщений является относительной к координатной системе. Если изменить координаты, то приведенная энтропия в общем случае также изменится. Например, при переходе от координат $x_1, x_2, ..., x_n$ к координатам $y_1, y_2, ..., y_n$ новое значение приведенной энтропии будет выражаться в виде

$$\begin{aligned} H^*(Y) = -\int\limits_{-\infty}^{\infty} ..(n).. \int\limits_{-\infty}^{\infty} W(x_1, x_2, ..., x_n) \left| J\left(\frac{x}{y}\right) \right| \times \\ \times \log W(x_1, x_2, ..., x_n) \left| J\left(\frac{x}{y}\right) \right| dy_1 dy_2 \cdots dy_n \end{aligned}, \qquad (5.81)$$

где $\left| J\left(\frac{x}{y}\right) \right|$ - якобиан преобразования координат по абсолютной величине (так как плотность вероятностей – величина положительная).

Разложив логарифм и заменив переменные на $x_1, x_2, ..., x_n$ соотношение (5.81) примет вид

$$H^*(Y) = H^*(X) -$$
$$- \int_{-\infty}^{\infty} ..(n).. \int_{-\infty}^{\infty} W(x_1, x_2, ..., x_n) \log \left| J\left(\frac{x}{y}\right) \right| dx_1 dx_2 \cdots dx_n \quad (5.82)$$

Как видим, при переходе к новой системе координат значение приведенной энтропии будет равно значению приведенной энтропии в старой системе координат минус ожидаемый логарифм якобиана.

При линейном преобразовании $y_j = \sum_i \alpha_{ij} x_i$ якобиан представляет собой простой определитель $\left|\alpha_{ij}\right|^{-1}$ и в этом случае приведенная энтропия принимает вид

$$H^*(Y) = H^*(X) + \log\left|\alpha_{ij}\right| \quad (5.83)$$

При вращении координатной системы якобиан преобразования $J = 1$ и

$$H^*(Y) = H^*(X).$$

Рассмотрим, какова будет приведенная энтропия $H_2^*(X)$ непрерывного сообщения на выходе линейного устройства (линейного фильтра) с амплитудно – частотной характеристикой $K(\omega)$ (фазовые соотношения не учитываются, так как на вход поступает случайный процесс), если на входе сообщение имело приведенную энтропию $H_1^*(X)$ в полосе частот $\Delta\omega$. Действие линейной системы представляет собой линейное преобразование координат. Если частотные составляющие сообщения на входе рассматривать как первичные координаты, то новые частотные составляющие будут представлять собой старые, умноженные на некоторые коэффициенты. В этом случае матрица преобразования координат относительно новых координат является диагональной и якобиан преобразования равен

$$J = \prod_{i=1}^{n} \left|K(\omega_i)\right|^2 .$$

Подставляя этот якобиан в выражение (5.83) и осуществляя предельный переход, получим [10]

$$H_2^*(X) = H_1^*(X) + H_n^* = H_1^*(X) + \frac{1}{\Delta\omega} \int_{\Delta\omega} \log\left|K(\omega)\right|^2 d\omega . \quad (5.84)$$

Величина $H_n^* = \frac{1}{\Delta\omega}\int\limits_{\Delta\omega} \log\left|K(\omega)\right|^2 d\omega$ называется потерями приведенной энтропии в линейных устройствах.

Для непрерывных сообщений часто используют понятие энтропийной мощности. *Под энтропийной мощностью понимают мощность нормального случайного процесса с постоянным энергетическим спектром, ограниченного такой же полосой частот, что и рассматриваемое сообщение, и имеющего такую же приведенную энтропию.*[10].

Если приведенная энтропия сообщения $H^*(X)$, то энтропийная мощность равна

$$\sigma_{\acute{y}}^2 = \frac{1}{2\pi e}\exp\left\{2H^*(X)\right\}. \tag{5.85}$$

Изменение приведенной энтропии на выходе линейного устройства можно выразить через энтропийную мощность в следующем виде

$$\sigma_{2\acute{y}}^2 = \sigma_{1\acute{y}}^2 k_{\acute{y}\grave{\imath}} = \sigma_{1\acute{y}}^2 \exp\frac{1}{\Delta\omega}\int\limits_{\Delta\omega} \log\left|K(\omega)\right|^2 d\omega. \tag{5.86}$$

Здесь $\sigma_{1\acute{y}}^2$ и $\sigma_{2\acute{y}}^2$ энтропийные мощности на входе и выходе линейного устройства соответственно.

Величина $k_{\acute{y}\grave{\imath}} = \exp\frac{1}{\Delta\omega}\int\limits_{\Delta\omega} \log\left|K(\omega)\right|^2 d\omega$ называется *коэффициентом энтропийной мощности.*

В качестве иллюстрации в табл.5.2 приведены потери энтропии в линейных устройствах с некоторыми идеализированными передаточными функциями с единичной полосой пропускания.

Таблица 5.2

$K(\omega)$	H_n^*	$k_{\acute{y}\grave{\imath}}$
$1-\omega$	-1,45	0,136
$1-\omega^2$	-0,87	0,3
$1-\omega^3$	-0,7	0,384
$\sqrt{1-\omega^2}$	-0,45	0,54

Как видно из табл.5.2 при передаче сообщения через линейные устройства происходит потери приведенной энтропии. Приведенная энтропия сообщения на выходе линейного устройства уменьшается по отношению к приведенной энтропии на входе. И это естественно, так как уменьшается неопределенность из-за того, что более высокочастотные составляющие сигнала передаются с заведомо известными искажениями.

Пример. На вход линейного фильтра с передаточной функцией

$$K(\omega) = 1 - \omega^2$$

поступает, соответствующий передаваемому сообщению, сигнал гауссовой статистикой и средней мощностью $\sigma^2 = 1$. Определить приведенную энтропию сообщения на выходе этого фильтра.

Решение. Приведенная энтропия сообщения на входе фильтра равна

$$H_1^*(X) = \log \sigma \sqrt{2\pi e} = 4{,}12 \text{ } \textit{дв.уд.}$$

Энтропия на выходе фильтра согласно соотношению (5.84) и значений H_n^*, приведенных в таблице 5.2, будет равна

$$H_2^*(X) = H_1^*(X) + H_n^* = 4{,}12 - 0{,}87 = 3{,}25 \text{ } \textit{дв.ед}$$

Рассмотрим изменение приведенной энтропии при прохождении сигнала через нелинейные устройства.

Положим, что в безинерционном нелинейном устройстве осуществляется функциональное преобразование вида $y = f(x)$. На вход такого устройства поступает непрерывный информационный процесс с известной плотностью вероятностей $W_n(x_1, x_2, ..., x_n)$ и приведенной энтропией $H^*(X)$.

Плотность вероятностей функционально преобразованного процесса определяется зависимостью [13]

$$W_n(y_1, y_2, ..., y_n) = \sum_k W(x_{1k}, x_{2k}, ..., x_{nk}) \left| J\left(\frac{x_k}{y}\right) \right|, \qquad (5.87)$$

где: $x_{ik}(y_1, y_2, ..., y_n)$ - k-я ветвь обратного преобразования ($i = 1, 2, ..., n$),

$$J\left(\frac{x_k}{y}\right) = \begin{vmatrix} \dfrac{\partial x_{1k}}{\partial y_1} & \dots & \dfrac{\partial x_{1k}}{\partial y_n} \\ \dots & \dots & \dots \\ \dfrac{\partial x_{nk}}{\partial y_1} & \dots & \dfrac{\partial x_{nk}}{\partial y_n} \end{vmatrix}$$ - якобиан преобразования от случайных величин $x_1, x_2, ..., x_n$ к случайным величинам $y_1, y_2, ..., y_n$.

Следовательно, приведенная энтропия сообщения на выходе нелинейного устройства будет определяться соотношением

$$H_2^*\left(X\right)=H^*\left(Y\right)=$$

$$=-\int\limits_{-\infty}^{\infty}..(n)..\int\limits_{-\infty}^{\infty}W\left(y_1,y_2,...,y_n\right)\log W\left(y_1,y_2,...,y_n\right)dy_1dy_2\cdots dy_n=$$

$$=-\int\limits_{-\infty}^{\infty}..(n)..\int\limits_{-\infty}^{\infty}\sum_k W_n\left(x_{1k},x_{2k},...,x_{nk}\right)\left|J\left(\frac{x_k}{y}\right)\right|\times$$

$$\times\log\sum_k W_n\left(x_{1k},x_{2k},...,x_{nk}\right)\left|J\left(\frac{x_k}{y}\right)\right|dy_1dy_2\cdots dy_n \qquad (5.88)$$

Рассмотрим простой пример.

Пример. На вход нелинейного устройства с передаточной функцией $y=\left|\sqrt{x}\right|$ поступает сигнал, одномерная плотность вероятностей которого имеет вид

$$W\left(x\right)=\frac{1}{2\sigma\sqrt{2\pi x}}\exp\left(-\frac{x}{2\sigma^2}\right).$$

Следует определить плотность вероятностей сигнала и оценить его энтропию на выходе этого устройства.

Решение.

В данном случае $x=y^2$, $\left|\frac{\partial x}{\partial y}\right|=2y$, плотность вероятностей $W\left(y\right)$ получает вид $W\left(y\right)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{y^2}{2\sigma^2}\right)$.

Получили нормальный закон распределения. Следовательно, в данном случае приведенная энтропия выходного сигнала имеет максимальное значение, т.е. $H^*\left(Y\right)=H_2^*\left(X\right)>H_1^*\left(X\right)$ ($H_1^*\left(X\right)$ - приведенная энтропия сигнала на входе).

При прохождении сигнала через нелинейные устройства приведенная энтропия на выходе может как увеличиться, так и уменьшиться по отношению к входной приведенной энтропии. Однако если входной процесс имеет нормальное распределение, то в этом случае всегда имеет место уменьшение приведенной энтропии.

5.2.8. Теорема Котельникова

В непрерывных информационных каналах с помехой входные сигналы являются непрерывными функциями времени, а выходные будут их искаженными копиями. Реально непрерывных сигналов не существует, все они в той или иной мере ограничены во времени. Однако воспользоваться моделью непрерывного сигнала чрезвычайно удобно, так как в этом случае его можно рассматривать в ограниченной полосе частот Ω и аналитически представлять, используя теорему Котельникова.

Теорема Котельникова.

Если функция времени $s(t)$ имеет спектральную функцию $S(j\omega)$, ограниченную по полосе частот $\Omega = 2\pi F$, то такая функция будет полностью определяться значением ее ординат в точках, отстоящих друг от друга на интервалы времени Δt, равные

$$\Delta t = \frac{\pi}{\Omega} = \frac{1}{2F}. \qquad (5.89)$$

Функция времени $s(t)$ на интервале наблюдения T может быть представлена с помощью k отсчетов

$$k = \frac{\Omega T}{\pi} = 2FT. \qquad (5.90)$$

Доказательство теоремы.

Любую вещественную функцию $s(t)$ можно представить в виде обратного преобразования Фурье

$$s(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(j\omega)e^{j\omega t}d\omega, \qquad (5.91)$$

где $S(j\omega)$ - спектральная функция $s(t)$.

Эта спектральная функция ограничена полосой частот Ω, т.е.

$S(j\omega) = 0$ при $|\omega| > \Omega$.

Следовательно, соотношение (5.91) можно переписать в виде

$$s(t) = \frac{1}{2\pi}\int_{-\Omega}^{\Omega} S(j\omega)e^{j\omega t}d\omega. \qquad (5.92)$$

Определим эту функцию только для дискретных моментов времени, равных

$$t_k = \frac{k\pi}{\Omega}, \ (k = 1, 2, 3, ...).$$

В этом случае соотношение (5.92) принимает вид

$$s\left(\frac{k\pi}{\Omega}\right)=\frac{1}{2\pi}\int_{-\Omega}^{\Omega}S(j\omega)\exp\left(j\omega\frac{k\pi}{\Omega}\right)d\omega. \tag{5.93}$$

Представим спектральную функцию $S(j\omega)$ на интервале ее существования $[-\Omega,\Omega]$ рядом Фурье

$$S(j\omega)=\frac{1}{2}\sum_{k=-\infty}^{\infty}\dot{D}_K e^{j\omega\frac{k\pi}{\Omega}}, \tag{5.94}$$

где $$\dot{D}_k=\frac{1}{\Omega}\int_{-\Omega}^{\Omega}S(j\omega)\exp\left(-j\omega\frac{k\pi}{\Omega}\right)d\omega. \tag{5.95}$$

Если сравним (5.95) с (5.93), то комплексную амплитуду гармоник $\dot{D}_k$ можно представить в следующем виде

$$\dot{D}_k=\frac{2\pi}{\Omega}s\left(-\frac{k\pi}{\Omega}\right).$$

Теперь ряд (5.94) примет вид

$$S(j\omega)=\frac{\pi}{\Omega}\sum_{k=-\infty}^{\infty}s\left(-\frac{k\pi}{\Omega}\right)e^{j\omega\frac{k\pi}{\Omega}}=\frac{\pi}{\Omega}\sum_{k=-\infty}^{\infty}s\left(\frac{k\pi}{\Omega}\right)e^{-j\omega\frac{k\pi}{\Omega}}. \tag{5.96}$$

Подставим значение $S(j\omega)$ из (5.96) в формулу (5.92) и получим

$$s(t)=\frac{1}{2\pi}\cdot\frac{\pi}{\Omega}\int_{-\Omega}^{\Omega}\left\{\sum_{k=-\infty}^{\infty}s\left(\frac{k\pi}{\Omega}\right)\exp\left(-j\omega\frac{k\pi}{\Omega}\right)\right\}e^{j\omega t}d\omega=$$

$$=\frac{1}{2\pi}\cdot\frac{\pi}{\Omega}\sum_{k=-\infty}^{\infty}s\left(\frac{k\pi}{\Omega}\right)\int_{-\Omega}^{\Omega}\exp\left[j\omega\left(t-\frac{k\pi}{\Omega}\right)\right]d\omega=$$

$$=\sum_{k=-\infty}^{\infty}s\left(\frac{k\pi}{\Omega}\right)\frac{\sin\Omega\left(t-\frac{k\pi}{\Omega}\right)}{\Omega\left(t-\frac{k\pi}{\Omega}\right)}=\sum_{k=-\infty}^{\infty}s\left(\frac{k\pi}{\Omega}\right)Sa(\Omega t-k\pi). \tag{5.97}$$

В формуле (5.97) величины $s\left(\frac{k\pi}{\Omega}\right)$ представляют значения непрерывной функции $s(t)$ в дискретные моменты времени $\frac{k\pi}{\Omega}$,

$Sa\left(\Omega t - k\pi\right)$ - называется функцией отсчетов, которая в моменты $t = \frac{k\pi}{\Omega}$ достигает своего максимума, равного единице, а в моменты $t = \frac{\left(k \pm \nu\right)\pi}{\Omega}$ $(\nu = 1, 2, 3, \ldots)$ обращается в ноль. (Функции отсчетов ортогональны на бесконечном интервале).

Таким образом, выражение (5.97) представляет аналитическую запись теоремы Котельникова и указывает на то, что любая функция с ограниченной спектральной функцией определяется своими значениями в дискретные моменты времени, отстоящими на интервал $\Delta t = \frac{\pi}{\Omega}$.

Теорема Котельникова известна также под названием теорема отсчетов.

Следует отметить, что если длительность сигнала ограничена по времени, то теорему Котельникова можно применить к амплитудно частотной характеристики сигнала:

$$S\left(\omega\right) = \sum_{k=-\infty}^{\infty} S\left(k\Omega_1\right) Sa\left(\frac{\pi\omega}{\Omega_1} - k\pi\right),$$

где $\Omega_1 = \frac{2\pi}{\tau}$, τ - длительность сигнала.

Условия теоремы Котельникова для реальных сигналов не удовлетворяются ввиду их временного ограничения. Однако практически всегда можно ограничить спектральную функцию реального сигнала достаточно большой частотой, при которой искажения сигнала будут минимальны.

5.2.9. Пропускная способность непрерывного информационного канала

В соответствии с теоремой Котельникова передаваемый непрерывный сигнал в интервале T можно представить $\frac{\Omega T}{\pi}$ числами, т.е. энергетический спектр передаваемого сигнала ограничен частотой Ω. Статистические свойства сигнала $x\left(t\right)$ и помехи $z\left(t\right)$ будем характеризовать плотностями вероятностей $W\left(x_1, x_2, \ldots, x_n\right)$ и $W\left(z_1, z_2, \ldots, z_n\right)$ соответственно.

По аналогии с дискретными каналами скорость передачи информации по непрерывному каналу с помехой и его пропускная способность будут определяться соответственно соотношениями

$$V\left(Y^{'},X^{'}\right)=\lim_{T\to\infty}\frac{H^{*}\left(Y^{'}\right)-H^{*}\left(Y^{'}\middle|X^{'}\right)}{T}, \tag{5.98}$$

$$C_c=Sup\left[V\left(Y^{'}\right)-V\left(Y^{'}\middle|X^{'}\right)\right]. \tag{5.99}$$

Будем исходить из условия аддитивности сигнала и помехи и отсутствия между ними корреляции. В этом случае процесс на выходе канала связи $y(t)=x(t)+z(t)$. Если $X^{'}$ и Z статистически независимы, то количество информации, содержащееся в $Y^{'}$ при уже известном $H\left(X^{'}\right)$, обусловлено только помехой. Тогда ненадежность канала будет равна

$$V\left(Y^{'}\middle|X^{'}\right)=V(Z)=\frac{H(Z)}{T},$$

где $H(Z)$ - энтропия помехи.

Следовательно, соотношение (5.99) примет вид

$$C_c=Sup\left[V\left(Y^{'}\right)-V(Z)\right]. \tag{5.100}$$

Определим пропускную способность непрерывного канала связи при воздействии наиболее опасной помехи, которой при известной средней мощности является помеха с нормальным законом распределения.

Естественно также, что максимальный поток информации на выходе канала связи $V\left(Y^{'}\right)$ имеет место при гауссовой статистике сигнала.

Будем рассматривать каждую из энтропий как энтропию объединения, т.е.

$$\begin{aligned}H\left(Y^{'}\right)&=H\left(y_1^{'},y_2^{'},...,y_n^{'}\right)\\H(Z)&=H\left(z_1,z_2,...,z_n\right)\end{aligned}.$$

Так как состояния $y_1^{'},y_2^{'},...,y_n^{'}$ и $z_1,z_2,...,z_n$ статистически независимы, то энтропия объединения будет равна сумме энтропий всех состояний, и для гауссовой статистики получим

$$H\left(Y^{'}\right)=\frac{\Omega T}{\pi}H\left(y^{'}\right)=\frac{\Omega T}{\pi}\left[\log\sqrt{2\pi e\sigma_y^2}-\log\Delta y\right], \tag{5.101}$$

$$H(Z)=\frac{\Omega T}{\pi}H(z)=\frac{\Omega T}{\pi}\left[\log\sqrt{2\pi e\sigma_z^2}-\log\Delta x\right], \tag{5.102}$$

где Δx - интервал квантования входного сигнала, а, следовательно, и помехи,

$\sigma_y^2=\sigma_x^2+\sigma_z^2$ - мощность сигнала на выходе канала связи,

σ_x^2 и σ_z^2 - соответственно мощности сигнала и помехи.

Если точность квантования сигналов на входе Δx и на выходе Δy канала связи одинаковы, то с учетом (5.101) и (5.102) пропускная способность непрерывного канала связи с помехой будет определяться зависимостью

$$c_c=\frac{\Omega}{\pi}\left[\log\sqrt{2\pi e\sigma_y^2}-\log\sqrt{2\pi e\sigma_z^2}\right]=\frac{\Omega}{2\pi}\log\left(1+\frac{\sigma_x^2}{\sigma_z^2}\right). \tag{5.103}$$

Формула (5.103) указывает на то, что максимальная скорость передачи информации по непрерывному каналу с помехой прямо пропорциональна полосе частот $\frac{\Omega}{2\pi}$ и логарифму суммы $1+\frac{\sigma_x^2}{\sigma_z^2}$.

При этом максимальное количество информации $I_{max}\left(Y',X'\right)$, содержащееся в Y' относительно X' будет равно

$$I_{max}\left(Y',X'\right)=\frac{\Omega T}{2\pi}\log\left(1+\frac{\sigma_x^2}{\sigma_z^2}\right). \tag{5.104}$$

Величину $\frac{\Omega T}{2\pi}\log\frac{\sigma_x^2}{\sigma_z^2}=FT\log\mu_p$ называют *объемом сигнала*.

Здесь $\mu_p=\frac{\sigma_x^2}{\sigma_z^2}$ - отношение сигнала к шуму по мощности.

Как видим из (5.104) одно и тоже количество информации можно передать, сохраняя постоянным объем сигнала, но варьируя шириной полосы пропускания $\frac{\Omega}{2\pi}$, длительностью сообщения T или отношением сигнала к шуму μ_p.

Если $\mu_p<<1$, то, учитывая, что при малом α в разложении

$$\log(1+\alpha) = \log e \cdot \ln(1+\alpha) = \log e \cdot \left(\alpha - \frac{\alpha^2}{2} + \frac{\alpha^3}{3} - \cdots\right)$$ можно ограничиться одним членом ряда, получим:

$$C_c = \frac{\Omega}{2\pi}\mu_p \log e = 1{,}443\frac{\Omega}{2\pi}\mu_p = 1{,}443F\mu_p.$$

Таким образом, при малом отношении мощности входного сигнала к мощности помехи, пропускная способность непрерывного канала прямо пропорциональна отношению сигнала к помехе по мощности. При $\mu_p \to 0$ и $C_c \to 0$.

Для $\mu_p >> 1$ имеем

$$C_c = \frac{\Omega}{2\pi}\log\mu_p = F\log\mu_p.$$

При большом отношении сигнала к помехе по мощности пропускная способность непрерывного канала пропорциональна логарифму от этого отношения.

Определенный интерес представляет зависимость пропускной способности непрерывного канала от полосы пропускания при постоянной средней мощности входного сигнала.

Запишем выражение для пропускной способности канала связи в следующем виде

$$C_c = \frac{n\Omega}{2\pi}\log\left[1 + \frac{\int_0^{n\Omega} G_x(\omega)\,d\omega}{n\Omega G_z}\right], \tag{5.105}$$

где G_z - энергетический спектр помехи (белый шум),

$G_x(\omega) = \begin{cases} \mu_p G_z & \text{при } 0 \le \omega \le \Omega \\ 0 & \omega > \Omega \end{cases}$ - энергетический спектр сигнала.

Тогда для $0 < n < 1$ соотношение (5.105) принимает вид

$$C_c = \frac{n\Omega}{2\pi}\log\left(1 + \mu_p\right).$$

Таким образом, при сокращении полосы пропускания ($n < 1$) имеем линейное уменьшение пропускной способности канала связи вплоть до нуля.

При $n > 1$ (расширение полосы пропускания) соотношение (5.105) получит вид

$$C_c = \frac{n\Omega}{2\pi}\log\left(1+\frac{\mu_p}{n}\right). \qquad (5.106)$$

При $n >> 1$ соотношение (5.106) приводится к виду

$$C_c = 1,443\frac{\Omega}{2\pi}\mu_p .$$

На рис.5.5 приведено семейство кривых, характеризующих изменения относительных величин пропускной способности непрерывного канала связи $\frac{C_c}{C_{c/n=1}}$ от ширины полосы пропускания при трех значениях отношения сигнала к помехе по мощности.

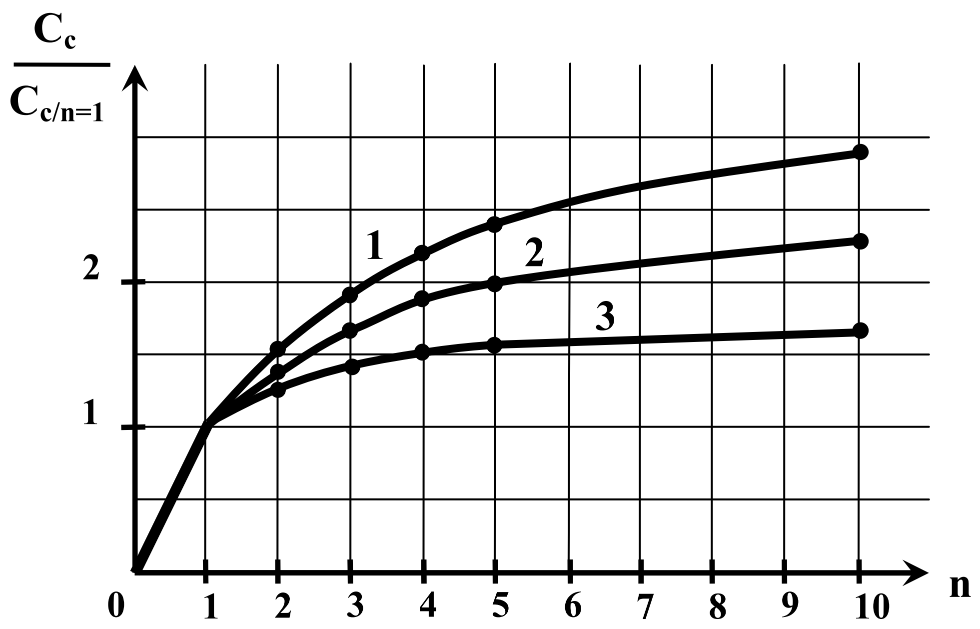

Рис.5.5

Зависимости относительной величины пропускной способности Непрерывного канала связи от ширины полосы пропускания при

1- $\mu_p = 10$, *2 –* $\mu_p = 5$, *3 -* $\mu_p = 2$

Из приведенных соотношений и кривых следует, что с увеличением полосы пропускания канала при заданной мощности передаваемого сигала, значение пропускной способности непрерывного канала стремится к некоторому пределу, величина которого прямо пропорциональна отношению сигнала к помехе по мощности.

5.2.10. Влияние энергетического спектра сигнала и помехи на скорость передачи непрерывной информации

В общем случае в канал связи поступают сигналы и помехи с неравномерными энергетическими спектрами. Положим, что полоса частот канала связи $\Omega = \omega_2 - \omega_1$, энергетические спектры сигнала и помехи соответственно равны $G_x(\omega)$ и $G_z(\omega)$. Тогда в элементарной полосе частот $\Delta\omega$ значение скорости передачи информации можно представить в виде [14]

$$V_{\Delta\omega}\left(Y', X'\right) = \frac{\Delta\omega}{2\pi}\log\frac{G_x(\omega) + G_z(\omega)}{G_z(\omega)}. \tag{5.107}$$

Просуммируем все составляющие (5.107) в полосе частот Ω и в пределе получим

$$V\left(Y', X'\right) = \frac{1}{2\pi}\int_{\omega_1}^{\omega_2}\log\frac{G_x(\omega) + G_z(\omega)}{G_z(\omega)}d\omega. \tag{5.108}$$

Определим энергетический спектр сигнала, при котором скорость передачи информации достигает максимального значения, если известен энергетический спектр помехи, а средняя мощность сигнала сохраняется неизменной.

Это вариационная задача на отыскание функции $G_x(\omega)$, обеспечивающая максимум функционалу (5.108) при дополнительном условии

$$\frac{1}{2\pi}\int_{\omega_1}^{\omega_2} G_x(\omega)d\omega = \sigma_x^2 = const. \tag{5.109}$$

При решении этой вариационной задачи воспользуемся методом неопределенных множителей Лагранжа и составим уравнение

$$\frac{dF}{dG_x(\omega)} + \lambda\frac{d\varphi}{dG_x(\omega)} = 0, \tag{5.110}$$

где $F = \log\dfrac{G_x(\omega) + G_z(\omega)}{G_z(\omega)}$ и $\varphi = G_x(\omega)$.

Решение уравнения (5.110) дает значение $G_x(\omega)$:

$$G_x(\omega) = -\frac{1}{\lambda} - G_z(\omega).$$

Подставляя полученное значение $G_x(\omega)$ в (5.109) получим

$$\lambda = -\frac{\Omega}{2\pi\left(\sigma_x^2 + \sigma_z^2\right)}.$$

Таким образом, искомая величина энергетического спектра сигнала будет равна

$$G_x(\omega) = \frac{2\pi\left(\sigma_x^2 + \sigma_z^2\right)}{\Omega} - G_z(\omega).$$

Скорость передачи непрерывной информации по каналу связи при этом будет равна

$$V\left(Y', X'\right) = \frac{\Omega}{2\pi}\log 2\pi\frac{\sigma_x^2 + \sigma_z^2}{\Omega} - \frac{1}{2\pi}\int_{\omega_1}^{\omega_2}\log G_z(\omega)\,d\omega.$$

Как видим, например, если энергетический спектр помехи постоянен в полосе пропускания канала связи, то максимальная скорость передачи информации будет иметь место также при постоянном энергетическом спектре сигнала. При убывающем с частотой энергетическом спектре помехи максимальная скорость передачи непрерывной информации по каналу связи, естественно, будет иметь место при возрастающем по частоте энергетическом спектре сигнала.

5.3. Ценность информации

5.3.1. Определение ценности информации

Известны несколько способов количественного определения ценности информации. Все они основаны на представлении о цели, достижению которой способствует полученная информация. Чем в большей мере информация помогает достижению цели, тем более ценной она считается.

Имеются несколько предложений по определению меры ценности информации. К ним относятся:

1-мера ценности, предложенная М. М. Бонгардом и А. А Харкевичем, которая может быть выражена через приращение вероятности достижения цели и определяется как

$$v = \log\frac{P}{p},$$

где

p - вероятность достижения цели до получения информации (априорная вероятность);

P - вероятность достижения цели после получения информации (апостериорная вероятность).

В системах передачи информации цель сводится к правильной передаче сообщений независимо от их конкретного содержания и формулируется относительно каждого состояния множества X. Положим, что целью является принятие решения в пользу x_i. Тогда относительно этой цели, видимо, ценность информации, содержащейся в принятом y_i равна

$\log\frac{P(x_i|y_i)}{P(x_i)}$. Здесь $P(x_i)$- является априорной вероятностью передачи состояния x_i, а $P(x_i|y_i)$- апостериорной вероятностью переданного состояния x_i после принятия состояния y_i.

При такой формулировке цели ценность информации совпадает с частным количеством информации, которое рассматривалось выше.

Так как частное количество информации может иметь как положительные, так и отрицательные значения, то в последнем случае ценность отрицательна и такая информация называется дезинформацией.

2 - мера ценности, предложенная В. И. Корогодиным, определяется величиной

$$v = \frac{P - p}{1 - p}$$

Она обладает фактически теми же свойствами, что ценность в формуле М. М.Бонгарда и А. А. Харкевича, но изменяется от 0 до 1.

3 – мера ценности, предложенная Стратоновичем Р.Л. [3], связывает шенноновскую информацию с теорией решающих функций А.Вальда и определяется уменьшением среднего риска, которое характеризует качество принимаемых решений, при получении информации.

Для технических целей представляется наиболее удобным понятие о мере ценности информации по Стратоновичу Р.Л..

Действительно, например, при определении размера детали, используя равномерную функцию потерь (Рис.5.6), информация об уменьшении погрешности измерения до определенных пределов казалось бы не

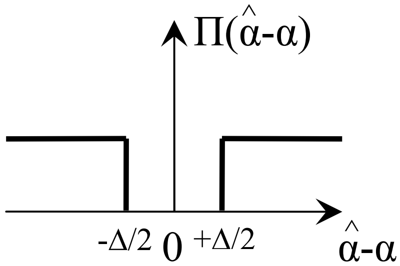

Рис.5.6 Равномерная функция потерь

имеет никакой ценности. Дальнейшее уменьшение погрешности измерения так же, казалось бы, не увеличивает ценности информации. Естественно, что для различных функций потерь величина ценности информации будет своя.

Однако интуиция подсказывает, что ценность информации зависит не только от функции потерь, но и от количества информации, т.е. между количеством и качеством (ценностью) информации существует связь.

Для установления этой связи рассмотрим следующую вариационную задачу.

Количество информации можно представить как разницу между энтропиями информационных непрерывных параметров до опыта $H_1(\alpha)$ и после опыта $H_2(\alpha)$

$$I = H_1(\alpha) - H_2(\alpha). \qquad (5.111)$$

Определим плотность вероятностей параметра $W(\alpha)$, при которой для заданной приведенной энтропии параметра

$$H^*(\alpha) = -\int_{-\infty}^{\infty} W(\alpha) \ln W(\alpha) d\alpha = H^* \qquad (5.112)$$

(При определении энтропии основание логарифма не играет роли, так как всегда можно перейти от логарифма с одним основанием к логарифму с другим основанием по формуле $\log_b K = \log_b a \log_a K$. Для нашей задачи удобно использовать натуральный логарифм)

средний риск R будет минимальным

$$R = \int_{-\infty}^{\infty} \ddot{I}(\alpha) W(\alpha) d\alpha \to \min\Big|_{W(\alpha)}, \qquad (5.113)$$

где $\ddot{I}(\alpha)$ - функция потерь.

Кроме дополнительного условия (5.112) будем использовать также еще одно дополнительное условие в виде

$$\int_{-\infty}^{\infty} W(\alpha)\,d\alpha = 1. \tag{5.114}$$

Для решения этой вариационной задачи воспользуемся методом неопределенных множителей Лагранжа. В этом случае экстремум определяется решением уравнения

$$\frac{d\ddot{I}\ (\alpha)}{dW(\alpha)} + \lambda_1 \frac{dW(\alpha)\ln W(\alpha)}{dW(\alpha)} - \lambda_2 \frac{dW(\alpha)}{dW(\alpha)} = 0$$

или

$$\ddot{I}\ (\alpha) + \lambda_1 \ln W(\alpha) + \lambda_1 - \lambda_2 = 0. \tag{5.115}$$

Решение уравнения (5.115) дает

$$W(\alpha) = \exp\left(-1 + \frac{\lambda_2}{\lambda_1} - \frac{\ddot{I}\ (\alpha)}{\lambda_1}\right). \tag{5.116}$$

Используя дополнительное условие (5.114), получим

$$\int_{-\infty}^{\infty} W(\alpha)\,d\alpha = \exp\left(-1 + \frac{\lambda_2}{\lambda_1}\right)\int_{-\infty}^{\infty} e^{-\frac{\ddot{I}\ (\alpha)}{\lambda_1}}\,d\alpha = 1.$$

Отсюда

$$exp\left(-1 + \frac{\lambda_2}{\lambda_1}\right) = \frac{1}{\int_{-\infty}^{\infty} e^{-\frac{\ddot{I}\ (\alpha)}{\lambda_1}}\,d\alpha}. \tag{5.117}$$

Обозначим

$$\int_{-\infty}^{\infty} e^{-\frac{\ddot{I}\ (\alpha)}{\lambda_1}}\,d\alpha = \varphi \tag{5.118}$$

Назовем φ - статистическим коэффициентом.

Теперь плотность вероятностей $W(\alpha)$ можно представить в виде

$$W(\alpha) = \frac{\exp\left(-\frac{\ddot{I}\ (\alpha)}{\lambda_1}\right)}{\varphi}. \tag{5.119}$$

Случайная энтропия в этом случае будет равна

$$\ln W(\alpha) = -\frac{\ddot{I}\ (\alpha)}{\lambda_1} - \ln\varphi. \tag{5.120}$$

Правую и левую части соотношения (10) усредним по α и получим

$$\int_{-\infty}^{\infty} W(\alpha)\ln W(\alpha)d\alpha = -\frac{1}{\lambda_1}\int_{-\infty}^{\infty} \ddot{I}\ (\alpha)W(\alpha)d\alpha - \ln\varphi\int_{-\infty}^{\infty} W(\alpha)d\alpha$$

или

$$\int_{-\infty}^{\infty} \ddot{I}\ (\alpha)W(\alpha)d\alpha = \lambda_1\left[-\int_{-\infty}^{\infty} W(\alpha)\ln W(\alpha)d\alpha - \ln\varphi\right]. \qquad (5.121)$$

Учитывая (5.112) и (5.113) соотношение (5.121) можно переписать в виде

$$R = \lambda_1\left(H^* - \ln\varphi\right). \qquad (5.122)$$

При этом λ_1 определяется дополнительным условием (5.112), а, следовательно, при заданной функции потерь зависит от энтропии H^*

$$\lambda_1 = \lambda_1\left(H^*\right)\Big|_{\ddot{I}\ (\alpha)}. \qquad (5.123)$$

Следует отметить, что статистический коэффициент φ зависит от λ_1 и, следовательно, от H^*

$$\varphi = \varphi\left(H^*\right).$$

Таким образом, при заданной энтропии средний риск будет равен

$$R = \lambda_1\left(H^*\right)\Big|_{\ddot{I}\ (\alpha)}\left[H^* - \ln\varphi\left(H^*\right)\right] = R\left(H^*\right). \qquad (5.124)$$

Как видим, средний риск функционально связан с энтропией.

Так как средний риск характеризует качество принятия решений, то разность средних рисков до получения информации и после получения ее указывает на ту пользу, которую принесла полученная информация.

Таким образом, количественная мера ценности информации будет равна разности средних рисков до получения информации и после ее получения (т.е. среднего риска неопределенности до получения информации и среднего риска той неопределенности, которая остается после получения информации)

$$v = \left\{R\left(H_1^*\right) - R\left(H_1^* - H_2^*\right)\right\}_{\ddot{I}\ (\alpha)} = \left\{\lambda_1\left(H_1^*\right)\left[H_1^* - \ln\varphi\left(H_1^*\right)\right] -\right.$$

$$\left. -\lambda_1\left(H_1^* - H_2^*\right)\left[\left(H_1^* - H_2^*\right) - \ln\varphi\left(H_1^* - H_2^*\right)\right]\right\}_{\ddot{I}\ (\alpha)}. \qquad (5.125)$$

Естественно, что ценность информации будет зависеть от выбранного вида функции потерь $\ddot{I}\ (\alpha)$.

Например, если в качестве функции потерь взять квадратичную функцию

$$\ddot{I}\ (\alpha) = \alpha^2, \qquad (5.126)$$

то в соответствие с (5.119) имеем:

$$W(\alpha)=\frac{\exp\left(-\frac{\alpha^2}{\lambda_1}\right)}{\varphi}, \tag{5.127}$$

где $\varphi=\int\limits_{-\infty}^{\infty}\exp\left(-\frac{\alpha^2}{\lambda_1}\right)d\alpha$. (5.128)

В этом случае дополнительное условие (5.112) принимает вид

$$-\int\limits_{-\infty}^{\infty}W(\alpha)\ln W(\alpha)d\alpha=\frac{1}{\lambda_1\varphi}\int\limits_{-\infty}^{\infty}\alpha^2\exp\left(-\frac{\alpha^2}{\lambda_1}\right)d\alpha+\ln\varphi=H^*. \tag{5.129}$$

Для решения интегралов (5.128), (5.129) воспользуемся формулой (5.18), т.е.

$$\int\limits_{0}^{\infty}x^m e^{-bx^n}dx=\frac{1}{n}\cdot\frac{\tilde{A}\left(\frac{m+1}{n}\right)}{b^{\frac{m+1}{n}}},$$

где $\tilde{A}(z)$ - гамма-функция.

Согласно этой формуле интегралы будут раны;

$$\int\limits_{-\infty}^{\infty}\exp\left(-\frac{\alpha^2}{\lambda_1}\right)d\alpha=\sqrt{\pi\lambda_1}$$

и

$$\int\limits_{-\infty}^{\infty}\alpha^2\exp\left(-\frac{\alpha^2}{\lambda_1}\right)d\alpha=\frac{1}{2}\sqrt{\pi\lambda_1^3}.$$

Теперь соотношение (5.129) принимает вид

$$\frac{1}{2}+\ln\sqrt{\pi\lambda_1}=H^*. \tag{5.130}$$

Отсюда неопределенный множитель Лагранжа λ_1 будет равен

$$\lambda_1=\frac{1}{\pi}e^{2H^*-1} \tag{5.131}$$

Подставляя полученное значение λ_1 в формулу (5.125) получим значение ценности информации при квадратичной функции потерь

$$v=R\left(H_1^*\right)-R\left(H_1^*-H_2^*\right)=\frac{e^{2H_1^*-1}}{2\pi}\left(1-e^{-2H_2^*}\right). \tag{5.132}$$

На рис.5.7 представлен график, характеризующий ценность информации в зависимости от увеличения энтропии $H_2^*=H$, т.е. по мере уменьшения исходной неопределенности (увеличения количества информации)

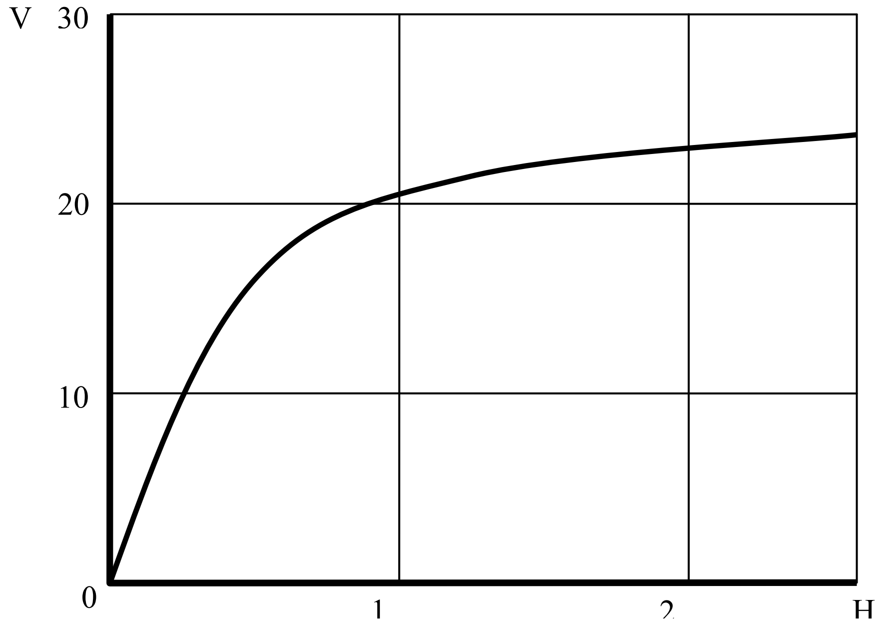

Рис.5.7

Зависимость ценности информации от энтропии H_2 при квадратичной функции потерь

Как видно из графика, приведенного на рис.5.7 для квадратичной функции потерь ценность информации возрастает с увеличением количества информации по закону $\left[1-e^{-2H^*}\right]$

Таким образом, качество информации функционально связано с количеством информации, и эта функциональная зависимость определяется используемой функцией потерь. Для квадратичной функции потерь по мере роста количества информации приращение качества уменьшается.

5.3.2. Ценность информации при неполной достоверности ее передачи

Положим, что передается состояние (параметр) X, решение о котором выносится по оценке Y. При этом будем считать, что среднее количество информации, содержащееся в Y относительно X, задано и равно I_0, т.е.

$$I(Y,X)=H(X)-H(X|Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}W(x,y)\ln\frac{W(x,y)}{W(x)W(y)}dxdy=I_0,$$

где $H\left(x\middle|Y\right)$ - потеря информации,

$W\left(x,y\right)$ - совместная плотность вероятностей состояния X и оценки Y,

$W\left(x\right)$ и $W\left(y\right)$ - плотности вероятностей состояния и оценки соответственно.

Аналогично предыдущему разделу 5.3.1 будем искать такую совместную плотность вероятностей $W\left(x,y\right)$, при которой для заданного $I\left(Y,X\right)=I_0$ средний риск R будет минимальным, т.е.

$$R=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\ddot{I}\left(x,y\right)W\left(x,y\right)dxdy\rightarrow\min\Big|_{W\left(x,y\right)} \qquad (5.133)$$

при следующих дополнительных условиях:

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}W\left(x,y\right)\ln\frac{W\left(x,y\right)}{W\left(x\right)W\left(y\right)}dxdy=$$

$$=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}W\left(x,y\right)\left[\ln W\left(x,y\right)-\ln W\left(x\right)-\ln\int\limits_{-\infty}^{\infty}W\left(x,y\right)dx\right]dxdy=I_0, \qquad (5.134)$$

$$\int\limits_{-\infty}^{\infty}W\left(x,y\right)dy=W\left(x\right). \qquad (5.135)$$

Для решения этой вариационной задачи воспользуемся методом неопределенных множителей Лагранжа, будем отыскивать экстремум из соотношения

$$\frac{d\ddot{I}\left(x,y\right)W\left(x,y\right)}{dW\left(x,y\right)}+\lambda\frac{W\left(x,y\right)[\ln]W\left(x,y\right)-\ln W\left(x\right)-\ln\int\limits_{-\infty}^{\infty}W\left(x,y\right)dx}{dW\left(x,y\right)}+$$

$$+\gamma\left(x\right)\frac{dW\left(x,y\right)}{dW\left(x,y\right)}=0, \qquad (5.136)$$

где λ и $\gamma\left(x\right)$ - неопределенные множители Лагранжа, определяемые дополнительными условиями (5.134) и (5.135).

Получим

$$\ddot{I}\left(x,y\right)+\lambda\left[\ln W\left(x,y\right)-\ln W\left(x\right)-\ln W\left(y\right)+1\right]+\gamma\left(x\right)=0$$

или

$$\ln W(x,y)=\ln\left[W(x)W(y)\right]-\frac{\ddot{I}\ (x,y)}{\lambda}-\frac{\gamma(x)}{\lambda}-1. \qquad (5.137)$$

Из (5.237) получим выражение для совместной плотности вероятностей $W(x,y)$

$$W(x,y)=W(x)W(y)\exp\left[-\frac{\ddot{I}\ (x,y)}{\lambda}-\frac{\gamma(x)}{\lambda}-1\right]. \qquad (5.138)$$

Воспользуемся дополнительным условием (5.135)

$$\int_{-\infty}^{\infty}W(x)W(y)\exp\left[-\frac{\ddot{I}\ (x,y)}{\lambda}-\frac{\gamma(x)}{\lambda}-1\right]dy=W(x)$$

Отсюда получим

$$\exp\left[-\frac{\gamma(x)}{\lambda}-1\right]\int_{-\infty}^{\infty}W(y)\exp\left[-\frac{\ddot{I}\ (x,y)}{\lambda}\right]dy=1. \qquad (5.139)$$

Обозначим

$$\int_{-\infty}^{\infty}W(y)\exp\left[-\frac{\ddot{I}\ (x,y)}{\lambda}\right]dy=\psi(x).$$

Теперь совместную плотность вероятностей $W(x,y)$ можно записать в следующем виде

$$W(x,y)=\frac{W(x)W(y)}{\psi(x)}\exp\left[-\frac{\ddot{I}\ (x,y)}{\lambda}\right]. \qquad (5.140)$$

Случайная информация связи в этом случае будет равна

$$I(y,x)=\ln\frac{W(x,y)}{W(x)W(y)}=\ln\frac{\exp\left[-\frac{\ddot{I}\ (x,u)}{\lambda}\right]}{\psi(x)}=-\frac{\ddot{I}\ (x,y)}{\lambda}-\ln\psi(x).$$

Усредняя обе части этого равенства по x и y, получим

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}I(y,x)W(x,y)\,dxdy=-\frac{1}{\lambda}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\ddot{I}\ (x,y)W(x,y)\,dxdy-$$

$$-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\ln\psi(x)W(x,y)\,dxdy\ . \qquad (5.141)$$

Функция $\psi(x)$ зависит от λ и, следовательно, от I, так как λ определяется дополнительным условием (5.134). Поэтому обозначим второй двойной интеграл в правой части соотношения (5.141) за $\varphi(I)$,т.е.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \ln\psi(x)W(x,y)\,dxdy = \varphi(I).$$

Так как в формуле (5.141)

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I(y,x)W(x,y)\,dxdy = I(Y,X) = I,$$

а

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \ddot{I}\ (x,y)W(x,y)\,dxdy = R,$$

то соотношение (5.41) можно переписать в виде

$$R = \lambda\left[-I - \varphi(I)\right]. \qquad (5.142)$$

Так как λ функционально зависит от I, то окончательно получим

$$R = \lambda(I)\left[-I - \varphi(I)\right]. \qquad (5.143)$$

Как видим функциональная связь среднего риска со средним количеством информации, содержащемся в Y относительно X, идентична рассмотренной в разделе 5.3.1 функциональной связи среднего риска и энтропии.

Таким образом, ценность информации при неполной достоверности ее передачи можно представить в виде

$$v = R\left(H^*\right) - R\left(H^* - I\right) = R\left(H^*\right) - R\left[H^*\left(X|Y\right)\right], \qquad (5.144)$$

где $H^*\left(X|Y\right) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(x,y)\ln W\left(x|y\right)dxdy$ - потеря информации.

Условная плотность вероятностей $W\left(x|y\right)$ для рассматриваемой экстремальной задачи равна

$$W\left(x|y\right) = \frac{W(x)}{\psi(x)}\exp\left[-\frac{\ddot{I}\ (x,y)}{\lambda}\right].$$

Из формулы (5.144) видно, что ценность информации при неполной достоверности ее передачи возрастает с уменьшением среднего риска потери информации, что интуитивно и следовало ожидать. Однако следует заметить, что функциональная зависимость ценности от условной энтропии достаточно сложная.

5.4 Приложение

5.4.1 Приведенная энтропия некоторых непрерывных законов распределения

Закон распре-де-ления	Аналитическое выражение плотности вероятностей $W(x)$	Приведенная энтропия $H^* = H^*(X)$
Равно-мер-ный	$W(x) = \begin{cases} 0 & -\infty < x < a \\ \dfrac{1}{b-a} \text{ при } & a < x < b \\ 0 & b < x < \infty \end{cases}$	$H^* = \log(b-a) =$ $= \log \sigma 2\sqrt{3}$
Симп-сона (треу-голь-ный)	$W(X) = \begin{cases} 0 & -\infty < x < a \\ \dfrac{4(x-a)}{(b-a)^2} & a < x < \dfrac{a+b}{2} \\ \dfrac{4(b-x)}{(b-a)^2} \text{ при } & \dfrac{a+b}{2} < x < b \\ 0 & b < x < \infty \end{cases}$	$H^* =$ $= \log \dfrac{(b-a)\sqrt{e}}{2} =$ $= \log \sigma \sqrt{6e}$
Закон $\sec h^2 x$	$W(X) = \dfrac{a}{2ch^2 ax} = \dfrac{a}{2} \sec h^2 x$	$H^* = \log \dfrac{2}{a} +$ $+2(\log e - \log 2) =$ $= \log \dfrac{e^2}{2a}$
Арк-синуса	$W(x) = \begin{cases} 0 & -\infty < x < -a \\ \dfrac{1}{\pi\sqrt{a^2 - x^2}} \text{ при } & -a < x < a \\ 0 & a < x < \infty \end{cases}$	$H^* = \log \pi +$ $+ \int_0^a \dfrac{\log(a^2 - x^2)dx}{\pi\sqrt{a^2 - x^2}}$
Коши	$W(x) = \dfrac{a}{\pi} \cdot \dfrac{1}{(x-\mu)^2 + a^2}$	$H^* = \log 4\pi a$

Экспо-нен-циаль-ный	$W(x)=\begin{cases} 0 & \\ ae^{ax} & \end{cases}$ при $\begin{matrix} -\infty < x < 0 \\ 0 < x < \infty \end{matrix}$	$H^{*}=\log\frac{e}{a}=$ $=\log\sigma e$
Лап-ласа	$W(x)=\frac{\lambda}{2}e^{-\lambda\lvert x-\mu\rvert}$	$H^{*}=\log\frac{2e}{a}=$ $=\log\sigma e\sqrt{2}$
Гипер-экспо-ненци-аль-ный	$W(x)=\begin{cases} 0 & -\infty < x < 0 \\ & \text{ï ðè} \\ \sum\limits_{n=1}^{N}a_n\lambda_n e^{-\lambda_n x} & 0 < x < \infty \end{cases}$	$H^{*}=-\sum\limits_{n=1}^{N}a_n\lambda_n\times$ $\times\int\limits_{0}^{\infty}e^{-\lambda_n x}\log\sum\limits_{n=1}^{N}a_n\times$ $\times\lambda_n e^{-\lambda_n x}dx$
Пока-затель-но-сте-пенной	$W(x)=\begin{cases} 0 & \\ \frac{x^m}{m!}e^{-x} & \end{cases}$ при $\begin{matrix} -\infty < x < 0 \\ 0 < x < \infty \end{matrix}$	$H^{*}=\log m!-$ $-m\log e\left(\sum\limits_{k=2}^{m}\frac{1}{k}-C\right)+$ $+\log e$ $C=0{,}5772..$ - число Эйлера
Эрлан-га	$W(x)=\begin{cases} 0 & \\ \frac{\beta^{\alpha}x^{\alpha-1}e^{-\beta x}}{(\alpha-1)!} & \end{cases}$ при $\begin{matrix} -\infty < x < 0 \\ 0 < x < \infty \end{matrix}$ $(\alpha=1,2,3,\ldots)$	$H^{*}=$ $=\log[(\alpha-1)!]-$ $-\log\beta+[\alpha-(\alpha-$ $-1)\psi(\alpha)\}\log e,$ $\psi(\alpha)=[\ln\tilde{A}(\alpha)]'$ - пси-функция Эйлера
Пир-сона	$W(x)=\begin{cases} 0 & \\ \frac{\alpha^{\lambda}x^{\lambda-1}e^{-\beta x}}{\tilde{A}(\lambda)} & \end{cases}$ при $\begin{matrix} -\infty < x < 0 \\ 0 < x < \infty \end{matrix}$	$H^{*}=\log\tilde{A}(\lambda)-$ $-\log\alpha+[\lambda-(\lambda-$ $-1)\psi(\lambda)]\log e,$ $\psi(\lambda)$ - пси-функция

	$\lambda=\frac{n}{2}, \quad (n=1,2,3,...)$	Эйлера
Гамма распре деле ние	$W(x)=\begin{cases} 0 & \text{при } -\infty<x<0 \\ \dfrac{x^{\alpha}e^{-\frac{x}{\beta}}}{\beta^{\alpha+1}\tilde{A}(\alpha+1)} & \text{при } 0<x<\infty \end{cases}$ $\alpha>-1,\ \beta>0$	$H^{*}=\log\tilde{A}(\alpha+1)-$ $-\alpha\log(e)\psi(\alpha+1)+$ $+(\alpha+1)\log e+\log\beta$ $\psi(\alpha+1)$- пси-функция Эйлера
Вей-була	$W(x)=\begin{cases} 0 & \text{при } -\infty<x<0 \\ \alpha\beta x^{\alpha-1}e^{-\beta x^{\alpha}} & \text{при } 0<x<\infty \end{cases}$	$H^{*}=\log e[1+$ $+\frac{\alpha-1}{\alpha}(C+\ln\beta)-$ $-\log\alpha\beta$, C - число Эйлера
Нор-маль-ный	$W(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-a)^{2}}{2\sigma^{2}}}$	$H^{*}=\log\sigma\left[\sigma\sqrt{2\pi e}\right]$
Одно-сторон ний нор-маль-ный	$W(X)=\begin{cases} 0 & \text{при } -\infty<x<0 \\ \sqrt{\frac{2}{\pi\sigma^{2}}}e^{-\frac{x^{2}}{2\sigma^{2}}} & \text{при } 0<x<\infty \end{cases}$	$H^{*}=\log\left[\sigma\sqrt{\frac{\pi e}{2}}\right]$
Усечен ный нор-маль-ный	$W(x)=\begin{cases} 0 & -\infty<x<x_{1} \\ \dfrac{Ae^{-\frac{(x-a)^{2}}{2\sigma^{2}}}}{\sqrt{2\pi}\sigma} & \text{при } x_{1}<x<x_{2}, \\ 0 & x_{2}<x<\infty \end{cases}$ $A=\dfrac{1}{\frac{1}{\sqrt{2\pi}}\left(\int\limits_{0}^{\frac{x_{2}-a}{\sigma}}e^{-\frac{u}{2}}du-\int\limits_{0}^{\frac{x_{1}-a}{\sigma}}e^{-\frac{u}{2}}du\right)}$	$H^{*}=\log\frac{\sqrt{2\pi}\sigma}{A}+$ $+\frac{1}{2}\left[1-A\frac{(x_{2}-a)}{\sqrt{2\pi}\sigma}\times\right.$ $\times e^{-\frac{(x_{2}-a)^{2}}{2\sigma^{2}}}-A\frac{x_{1}-a}{\sigma}\times$ $\left.\times e^{-\frac{(x_{1}-a)^{2}}{2\sigma^{2}}}\right]\log e$

Лога-рифми чески-нор-маль-ный	$W(x)=\begin{cases} 0 & -\infty<x<0 \\ \dfrac{e^{-\frac{(\ln x-a)^2}{2\sigma^2}}}{x\sigma\sqrt{2\pi}} \text{ при} & 0<x<\infty \end{cases}$	$H^*=\log\left[\sigma e^a\sqrt{2\pi e}\right]$
Рэлея	$W(x)=\begin{cases} 0 & -\infty<x<0 \\ \dfrac{x}{\sigma}e^{-\frac{x^2}{2\sigma^2}} \text{ при} & 0<x<\infty \end{cases}$	$H^*=\left(\dfrac{C}{2}+1\right)\log e,$ $C=0.5772..$-число Эйлера
Макс-велла	$W(x)=\begin{cases} 0 & -\infty<x<0 \\ \dfrac{4x^2e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{\pi}\left(2\sigma^2\right)^{3/2}} \text{ при} & 0<x<\infty \end{cases}$	$H^*=\dfrac{1}{2}\log\dfrac{2\pi\sigma^2}{e}+$ $+C\log e,$ C - число Эйлера
Пир-сона $\left(\chi^2\right)$	$W(x)=\begin{cases} 0 & -\infty<x<0 \\ \dfrac{x^{\frac{n}{2}-1}e^{-\frac{x}{2\sigma^2}}}{\left(2\sigma^2\right)^{\frac{n}{2}}\tilde{A}\left(\frac{n}{2}\right)} \text{ при} & 0<x<\infty \end{cases}$	$H^*=\log\tilde{A}\left(\dfrac{n}{2}\right)-$ $-\log\dfrac{1}{2\sigma^2}+\left[\dfrac{n}{2}-\left(\dfrac{n}{2}-\right.\right.$ $\left.\left.-1\right)\psi\left(\dfrac{n}{2}\right)\right]\log e,$ $\psi\left(\dfrac{n}{2}\right)$-пси-функция Эйлера
Распре деле-ние Нака-гами (m-рас-преде-ление)	$W(x)=\begin{cases} 0 & -\infty<x<0 \\ \dfrac{2m^mx^{2m-1}e^{-\frac{mx^2}{\sigma^2}}}{\tilde{A}(m)\sigma^{2m}}, & 0<x<\infty \end{cases},$ $m\geq\dfrac{1}{2}$	$H^*=$ $=\dfrac{\log\tilde{A}(m)\sigma e^m}{2\sqrt{m}}-$ $-\dfrac{2m-1}{2}\log'\tilde{A}(m),$ $\tilde{A}(m)$-гамма-функция

Распределение модуля многомерного вектора	$W\left(x\right)=\begin{cases} 0 & -\infty < x < 0 \\ \dfrac{2x^{n-1}e^{-\frac{x^2}{2\sigma^2}}}{\left(2\sigma^2\right)^{\frac{n}{2}}\tilde{A}\left(\frac{n}{2}\right)}, & 0 < x < \infty \end{cases},$ $n=1,2,3,...$	$H^* =$ $=\log\dfrac{\sigma e^{\frac{n}{2}}\tilde{A}\left(\frac{n}{2}\right)}{\sqrt{2}}-$ $-\dfrac{n-1}{2}\log'\tilde{A}\left(\dfrac{n}{2}\right)$

5.4.2 Энтропия некоторых дискретных законов распределения

Закон распределения	Аналитическое выражение распределения $P\left(x\right)$	Энтропия $H=H\left(X\right)$
Равномерный	$P\left(x_i\right)=\begin{cases} 0 & -\infty < x_i < 1 \\ \dfrac{1}{m} \text{ при } & 1 \le x_i \le m \\ 0 & m < x_i < \infty \end{cases}$	$H=\log m$
Геометрический	$P\left(x_i=k\right)=\begin{cases} 0 & -\infty < x_i \le 0 \\ p\left(1-p\right)^{k-1} & 0 < x_i < \infty \end{cases}$	$H=\dfrac{p\log p}{p}+$ $+\dfrac{\left(1-p\right)\log\left(1-p\right)}{p}$
Биномиальный	$P\left(x_i=k\right)=\begin{cases} 0 & -\infty < x_i < 0 \\ C_m^k p^k g^{m-k}, & 0 \le x_i \le m \\ 0 & m < x_i < \infty \end{cases}$	$H=$ $=-m\left(p\log p-g\log g\right)-$ $-\sum_{k=1}^{m-1} C_m^k p^k g^{m-k}\log C_m^k$

Гипер гео- метри чес- кий	$P\left(x_i=k\right)=\begin{cases} 0 & -\infty<x_i<0 \\ \dfrac{C_m^k C_{N-m}^{r-k}}{C_N^r}, & 0\le x_i\le m \\ 0 & m<x_i<\infty \end{cases}$	$H=\log C_N^r-$ $-\frac{1}{C_N^r}\sum_{k=0}^{m} C_m^k C_{N-m}^{r-k}\log C_m^k-$ $-\frac{1}{C_N^r}\sum_{k=0}^{m} C_m^k C_{N-m}^{r-k}\log C_{N-m}^{r-k}$
Пуас- сона	$P\left(x_i=k\right)=\begin{cases} 0 & -\infty<x_i<0 \\ \dfrac{\lambda^k}{k!}e^{-\lambda}, & 0\le x_i<\infty \end{cases}$	$H=\lambda\log\frac{e}{\lambda}+$ $+\sum_{k=1}^{\infty}\frac{\lambda^k e^{-\lambda}}{k!}\log k!$

5.4.3 Таблица значений функции $-P\log P$

P	$-P\log P$	P	$-P\log P$	P	$-P\log P$	P	$-P\log P$
0	0	0,26	0,5053	0,51	0,4954	0,76	0,3009
0,01	0,0664	0,27	0,5100	0,52	0,4906	0,77	0,2930
0,02	0,1128	0,28	0,5142	0,53	0,4854	0,78	0,2796
0,03	0,1518	0,29	0,5179	0,54	0,4800	0,79	0,2687
0,04	0,1858	0,30	0,5211	0,55	0,4744	0,80	0,2575
0,05	0,2161	0,31	0,5238	0,56	0,4685	0,81	0,2462
0,06	0,2435	0,32	0,5260	0,57	0,4623	0,82	0,2348
0,07	0,2686	0,33	0,5278	0,58	0,4558	0,83	0,2231
0,08	0,2915	0,34	0,5292	0,59	0,4491	0,84	0,2112
0,09	0,3126	0,35	0,5301	0,60	0,4422	0,85	0,1992
0,10	0,3322	0,36	0,5306	0,61	0,4350	0,86	0,1871
0,11	0,3503	0,37	0,5307	0,62	0,4276	0,87	0,1748
0,12	0,3671	0,38	0,5305	0,63	0,4199	0,88	0,1623
0,13	0,3826	0,39	0,5298	0,64	0,4121	0,89	0,1496
0,14	0,3971	0,40	0,5288	0,65	0,4040	0,90	0,1368
0,15	0,4105	0,41	0,5274	0,66	0,3957	0,91	0,1238
0,16	0,4230	0,42	0,5256	0,67	0,3871	0,92	0,1107
0,17	0,4346	0,43	0,5236	0,68	0,3784	0,93	0,0974
0,18	0,4453	0,44	0,5210	0,69	0,3694	0,94	0,0839
0,19	0,4552	0,45	0,5184	0,70	0,3602	0,95	0,0703
0,20	0,4644	0,46	0,5153	0,71	0,3508	0,96	0,0565
0,21	0,4728	0,47	0,5120	0,72	0,3412	0,97	0,0426
0,22	0,4806	0,48	0,5083	0,73	0,3314	0,98	0,0286
0,23	0,4877	0,49	0,5043	0,74	0,3215	0,99	0,0144
0,24	0,4941	0,50	0,5000	0,75	0,3113	1,00	0
0,25	0,5000						

Литература

1. Хартли Р. Передача информации. Теория информации и ее применение. Сборник переводов под редакцией А.А.Харкевича, Физматгиз, М., 1959.
2. Шеннон К. Работы по теории информации и кибернетики, ИЛ, М., 1963, 820с.
3. Стратонович Р.Л. Теория информации, Сов. Радио, М., 1975, 424с.
4. Гречко Л.Г., Сугаков В.И., Томасевич О.Ф., Федорченко А.М. Сборник задач по теоретической физике, Высшая школа, М., 1984, 320с.
5. Колмогоров А.Н. Теория информации и теория алгоритмов, Наука, М., 1987, 305с.
6. Лебедев Д.С., Гармаш В.А. О возможности передачи телеграфных сообщений, Электросвязь,1958, №1, 68-69**с.**
7. Харкевич А.А. Очерки общей теории связи, Гос.изд. теорет. лит., М.,1955, 268с.
8. Фельдбаум А.А., Дудыкин А.Д., Мановцев А.П.,Миролюбов Н.Н. Теоретические основы связи и управления, Физматгиз, М., 1963, 932с.
9. Гольдман С. Теория информации, ИЛ, М., 1957.
10. Шеннон К. Статистическая теория передачи электрических сигналов. Теория передачи электрических сигналов при наличии помех. Сборник статей под редакцией А.Н. Железнова, ИЛ, М., 1953, 288с.
11. Фано Р. Передача информации. Статистическая теория связи, Мир, М., 1965, 438с.
12. Липкин И.А. Основы статистической радиотехники, теории информации и кодирования, Сов. Радио, М., 1978, 238с.
13. Левин Б.Р. Теоретические основы статистической радиотехники, т.1, Сов. Радио, М., 1966, 728с.
14. Клюев Н.И. Информационные основы передачи сообщений, Сов. Радио, М., 1966, 360с.
15. Горяинов В.Т., Журавлев А.Г., Тихонов В.И. Примеры и задачи по статистической радиотехнике, Сов. Радио, М., 1970, 567с.

В 2009 году Университет стал победителем многоэтапного конкурса, в результате которого определены 12 ведущих университетов России, которым присвоена категория «Национальный исследовательский университет». Министерством образования и науки Российской Федерации была утверждена Программа развития государственного образовательного учреждения высшего профессионального образования «Санкт-Петербургский государственный университет информационных технологий, механики и оптики» на 2009–2018 годы.

КАФЕДРА ОПТИКО-ЭЛЕКТРОННЫХ ПРИБОРОВ И СИСТЕМ

История кафедры началась в 1936 году с организации в Ленинградском институте точной механики и оптики (ЛИТМО) кафедры военных оптических приборов. Первым заведующим кафедрой был К.Е. Солодилов, до этого возглавлявший Центральное конструкторское бюро (ЦКБ) Всесоюзного объединения оптико-механической промышленности (ВООМП). Преподавателями кафедры стали сотрудники этого ЦКБ - М.А. Резунов, М.Я. Кругер, С.Т. Цуккерман, В.А. Егоров, Б.М. Кулежнов.

В годы Великой Отечественной войны кафедра была эвакуирована в Черепаново, где ее объединили с кафедрой оптико-механических приборов под руководством профессора А.И. Захарьевского. После возвращения в Ленинград кафедрой в 1945-46 годах по совместительству заведовал начальник конструкторского бюро (КБ) Государственного оптического института им. С.И. Вавилова (ГОИ) М.А. Резунов.

В начале 1947 года кафедру возглавил профессор С.Т. Цуккерман, который руководил ею до 1972 года. В 1958 году кафедра была реорганизована в кафедру специальных оптических приборов, а в 1967 году в кафедру оптико-электронных приборов (ОЭП).

В 1958 г. при кафедре была организована отраслевая лаборатория "Специальные оптические приборы" с достаточно сильной группой конструкторов-разработчиков. С.Т. Цуккерман и старший научный сотрудник А.С. Гридин руководили разработкой приборов управления по лучу (ПУЛ), предназначенных для управления движением различных подвижных объектов по прямой линии или по программе.

В начале 60-х годов старший научный сотрудник Г.Г. Ишанин занимался разработкой фотометрической аппаратуры, предназначенной для паспортизации оптико-электронных приборов и систем различного назначения.

Значительное влияние на содержание подготовки специалистов и научных исследований оказало привлечение к работе на кафедре выдающегося специалиста в области оптико-электронного приборостроения, члена-корреспондента Российской академии наук (РАН), Героя Социалистического Труда, лауреата Ленинской премии профессора М.М. Мирошникова, который, работая на кафедре ОЭП с 1969 года по 1976 год в должности профессора по совместительству, поставил и читал курс "Теория оптико-электронных приборов".

С 1972 года по 1992 год кафедрой ОЭП заведовал заслуженный деятель науки и техники РСФСР, профессор Л.Ф. Порфирьев, известный специалист в области автоматических ОЭПиС в комплексах навигации и управления авиационной и космической техникой. Соответственно тематика выполнения научно-исследовательских работ на кафедре приобрела новые направления, существенно увеличилось число тем, носящих поисковый фундаментальный характер. Были разработаны новый учебный план и программы учебных дисциплин.

Л.Ф. Порфирьев является автором 19 учебников, учебных пособий и монографий, среди которых можно выделить такие как "Теория оптико-электронных приборов и систем" (Л.: Машиностроение, 1980), "Основы теории преобразования сигналов в оптико-электронных системах" (Л.: Машиностроение, 1989). Результаты его работ можно оценить как значительный вклад в разработку общей теории оптико-электронных систем.

Л.Ф. Порфирьев как руководитель проводил достаточно жесткую кадровую политику, при которой на кафедре оставались работать только те сотрудники, которые отличались преданностью делу. При этом он оказывал всемерную поддержку сотрудникам кафедры по разработке ими различных направлений теории и практики оптико-электронного приборостроения. По результатам научно-исследовательских работ в этот период защитили диссертации на соискание ученой степени доктора технических наук Г.Н. Грязин (1983 г.), Е.Г. Лебедько (1985 г.), Э.Д. Панков (1986 г.), Г.Г. Ишанин (1988 г.), защищено много диссертаций на соискание ученой степени кандидата технических наук.

В этот период под руководством Э.Д. Панкова начали проводиться исследования по разработке новых оптико-электронных систем измерения взаимного положения разнесенных в пространстве объектов.

Г.Н. Грязин, перешедший на кафедру с радиотехнического факультета в конце 60-х годов, продолжил свои работы в области прикладного телевидения, в частности, по разработке систем наблюдения за быстродвижущимися объектами и быстропротекающими процессами.

С 1975 года заведующим отраслевой лабораторией стал старший научный сотрудник А.Н. Тимофеев, который продолжил исследования по разработке методов и средств контроля пространственного положения объектов с помощью ОЭП с оптической равносигнальной зоной для машиностроения, энергетики, строительства, судостроения и железнодорожного транспорта.

С 1975 года, после увольнения в запас, из Ленинградской военной инженерной краснознаменной академии (ЛВИКА) им. А.Ф. Можайского на кафедру пришел работать в должности профессора С.П. Авдеев, известный специалист в области ОЭПиС космических аппаратов. Он поставил курсы и читал лекции по учебным дисциплинам "Оптико-электронные приборы", "Оптико-электронные приборы систем управления", "Оптико-электронные приборы для научных исследований".

Существенное влияние на содержание подготовки специалистов и научных исследований оказало привлечение к работе на кафедре лауреата Ленинской и Государственной премий профессора Б.А. Ермакова, известного специалиста в области физической оптики и оптико-электронного приборостроения. Б.А. Ермаков работал на кафедре ОЭП с 1979 года по 1992 год в должности профессора по совместительству и поставил курс "Оптико-электронные приборы с лазерами".

В 70-80 годах под руководством доцента Е.Г. Лебедько проводились исследования законов отражения лазерного излучения от нестационарных поверхностей и протяженных объектов, исследования в области теории идентификации объектов по их излучению в сложной фоновой ситуации. Создан комплекс для лазерной локации крупногабаритных морских объектов сложной конфигурации и водной поверхности. В этих работах принимали участие доценты О.П. Тимофеев и С.Б. Лукин.

В 70-90 годах под руководством Л.Ф. Порфирьева был разработан ряд астродатчиков, систем астроориентации и космической навигации (В.И. Калинчук, А.Л. Андреев, С.Н. Ярышев).

С 1992 г. заведующим кафедрой является заслуженный деятель науки Российской Федерации, профессор Э.Д. Панков. В 1992 году кафедра была переименована в кафедру оптико-электронных приборов и систем (ОЭПиС).

Под руководством Э.Д. Панкова в 70-90-х годах были проведены разработки ряда оптико-электронных приборов и систем специального и гражданского применения, нашедших практическое внедрение и способствующих научно-техническому прогрессу и укреплению обороноспособности нашей страны.

В рамках указанных работ доцентом И.А. Коняхиным проводились исследования, результаты которых можно классифицировать как разработку теории построения автоколлимационных систем с компонентами нарушенной типовой конфигурации.

В то же время доцентом В.В. Коротаевым разработан ряд поляризационных приборов и измерительных установок. Теоретическим результатом работ явилась разработка методологии анализа поляризационных свойств оптических систем с изменяющейся ориентацией элементов. По результатам указанных работ В.В. Коротаев (в 1997 г.) и И.А. Коняхин (в 1998г.) защитили диссертации на соискание ученой степени доктора технических наук.

Применение многоэлементных приемников в системах пеленгации дало толчок развитию телевизионных систем технического зрения, измерительных телевизионных систем и систем обработки изображений. Результаты этих исследований были использованы доцентом А.Л. Андреевым при постановке учебных курсов "Оптико-электронные системы с ЭВМ", "Специализированные аппаратные и программные средства ОЭП", "Автоматизированные телевизионные вычислительные комплексы", а также доцентом С.Н. Ярышевым при постановке им в 1993 году учебной дисциплины "Видеотехника".

На основе обобщения методик расчета оптико-электронных систем различного назначения и принципа действия в 1981 году были развернуты работы по созданию элементов систем автоматизированного проектирования ОЭП. За период с 1981 по 1987 год под руководством И.А. Коняхина были разработаны оригинальные пакеты прикладных программ расчета параметров систем измерения пространственного положения объектов.

Развитие компьютерной техники и программного обеспечения общего назначения позволило создать проблемно-ориентированное программное обеспечение поддержки проектирования ОЭП на системотехническом уровне.

С 2007 г. заведующим кафедрой является почетный работник высшего профессионального образования Российской Федерации, профессор В.В. Коротаев.

По результатам научных работ сотрудниками кафедры ОЭПиС выпущено в свет 15 монографий, 11 учебников и учебных пособий. На кафедре подготовлено 14 докторов наук, а также более 110 кандидатов наук.

На разработки кафедры получены авторские свидетельства СССР и патенты Российской Федерации на более чем 200 изобретений. Наибольший вклад в изобретательскую деятельность внес Э.Д. Панков – автор 123 изобретений, из которых 33 внедрены в промышленности.

Евгений Георгиевич Лебедько

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ПЕРЕДАЧИ ИНФОРМАЦИИ
(часть 5)

Учебное пособие

В авторской редакции
Дизайн Е.Г. Лебедько
Верстка Е.Г. Лебедько
Редакционно-издательский отдел Санкт-Петербургского государственного университета информационных технологий, механики и оптики
Зав. РИО Н.Ф. Гусарова
Лицензия ИД № 00408 от 05.11.99
Подписано к печати
Заказ № ______
Тираж 100 экз
Отпечатано на ризографе

Редакционно-издательский отдел
Санкт-Петербургского государственного
университета информационных технологий,
механики и оптики
197101, Санкт-Петербург, Кронверкский пр., 49