

Лекция 3. Энтропия как мера неопределённости (ч.2)

Энтропия непрерывных источников информации (дифференциальная энтропия)

Ранее была рассмотрена мера неопределенности выбора для дискретного источника информации. На практике, часто встречаются с источниками информации, множество возможных состояний которых составляет континуум. Такие источники называют непрерывными источниками информации.

Во многих случаях они преобразуются в дискретные посредством использования устройств дискретизации и квантования. Вместе с тем существует немало и таких систем, в которых информация передается и преобразуется непосредственно в форме непрерывных сигналов. Примерами могут служить системы телефонной связи и телевидения (аналоговые системы)).

Оценка неопределенности выбора для непрерывного источника информации имеет определенную специфику. Во-первых, значения, реализуемые источником, математически отображаются непрерывной случайной величиной. Во-вторых, вероятности значений этой случайной величины не могут использоваться для оценки неопределенности, поскольку в данном случае вероятность любого конкретного значения равна нулю.

Естественно, однако, связывать неопределенность выбора значения непрерывной случайной величины с плотностью распределения вероятностей этих значений. Учитывая, что для совокупности значений, относящихся к любому сколь угодно малому интервалу непрерывной случайной величины, вероятность конечна, попытаемся найти формулу для энтропии непрерывного источника информации, используя операции квантования и последующего предельного перехода при уменьшении кванта до нуля.

С этой целью разобьем диапазон изменения непрерывной случайной величины U , характеризующейся плотностью распределения вероятностей $p(u)$, на конечное число n малых интервалов шириной Δu (рис. 3.4). При реализации любого значения u , принадлежащего интервалу $(u_i, u_i + \Delta u)$, будем считать, что реализовалось значение u_i дискретной случайной величины U . Поскольку Δu мало, вероятность $p(u_i \leq u \leq u_i + \Delta u)$ реализации значения u из интервала $u_i, u_i + \Delta u$ (как площадь под кривой плотности распределения)

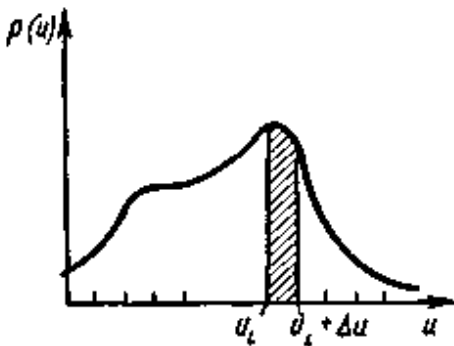


Рис. 3 4

$$p(u_i \leq u \leq u_i + \Delta u) = \int_{u_i}^{u_i + \Delta u} p(u) du \approx p(u_i) \Delta u.$$

Тогда энтропия дискретной случайной величины \tilde{U} может быть записана в виде:

$$H(\tilde{U}) = - \sum_{i=1}^n p(u_i) \Delta u \log [p(u_i) \Delta u],$$

или

$$H(\tilde{U}) = - \sum_{i=1}^n p(u_i) \Delta u \log p(u_i) - \sum_{i=1}^n p(u_i) \Delta u \log \Delta u.$$

Так как

$$\sum_{i=1}^n p(u_i) \Delta u = 1,$$

то

$$H(\tilde{U}) = - \sum_{i=1}^n p(u_i) \Delta u \log p(u_i) - \log \Delta u.$$

По мере уменьшения Δu $p(u_i \leq u \leq u_i + \Delta u)$ все больше приближается к вероятности $p(u_i)$, а свойства дискретной величины \tilde{U} — к свойствам непрерывной случайной величины U .

Переходя к пределу при $\Delta u \rightarrow 0$, получаем следующее выражение для энтропии $H(U)$ непрерывного источника:

$$\lim_{\Delta u \rightarrow 0} H(\tilde{U}) = \lim_{\Delta u \rightarrow 0} \left\{ - \sum_{i=1}^n p(u_i) \Delta u \log p(u_i) \right\} - \lim_{\Delta u \rightarrow 0} \log \Delta u$$

или

$$H(U) = - \int_{-\infty}^{\infty} p(u) \log p(u) du - \lim_{\Delta u \rightarrow 0} \log \Delta u. \quad (3.31)$$

Эта величина при $\Delta u \rightarrow 0$ стремится к бесконечности, что полностью соответствует интуитивному представлению о том, что неопределенность выбора из бесконечно большого числа возможных состояний (значений) бесконечно велика.

Первый член в правой части соотношения (3.31) имеет конечное значение, которое зависит только от закона распределения непрерывной случайной величины U и не зависит от шага квантования Δu . Он имеет точно такую же структуру, как энтропия дискретного источника.

Второй член того же соотношения, наоборот, зависит лишь от шага квантования случайной величины U . Именно в нем кроется причина того, что величина $H(U)$ обращается в бесконечность.

К использованию и трактовке соотношения (3.31) для получения конечной характеристики информационных свойств непрерывного источника известны два подхода.

- Один подход состоит в том, что в качестве меры неопределенности непрерывного источника принимают первый член соотношения (3.31):

$$h(U) = - \int_{-\infty}^{\infty} p(u) \log p(u) du. \quad (3.32)$$

Поскольку для определения этой величины используется только функция плотности вероятности, т. е. дифференциальный закон распределения, она получила название относительной дифференциальной энтропии или просто дифференциальной энтропии непрерывного источника информации (непрерывного распределения случайной величины U).

Ее можно трактовать как среднюю неопределенность выбора случайной величины U с произвольным законом распределения по сравнению со средней неопределенностью выбора случайной величины U' , изменяющейся в диапазоне, равном единице, и имеющей равномерное распределение.

Действительно, запишем соотношение (3.31) для случайной величины U' , равномерно распределенной в интервале δ :

$$H(U') = - \int_{-\infty}^{\infty} \frac{1}{\delta} \log \frac{1}{\delta} du' - \lim_{\Delta u' \rightarrow 0} \log \Delta u'.$$

При $\delta=1$

$$H(U') = - \lim_{\Delta u' \rightarrow 0} \log \Delta u',$$

откуда при $\Delta u = \Delta u'$

$$H(U) - H(U') = - \int_{-\infty}^{\infty} p(u) \log p(u) du = h(U).$$

Аналогично, используя операции квантования и предельного перехода, найдем выражение для условной энтропии непрерывного источника информации:

$$H_v(U) = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(u, v) \log \left[\frac{p(u, v)}{p(v)} \right] dudv - \lim_{\Delta u \rightarrow 0} \log \Delta u. \quad (3.33)$$

Отметим, что второй член в первой части выражения (3.33) идентичен соответствующему члену в соотношении (3.31). Обозначим первый член правой части выражения (3.33) через $h_v(U)$:

$$h_v(U) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(uv) \log \left[\frac{p(uv)}{p(v)} \right] dudv. \quad (3.34)$$

Эта величина конечна и называется относительной дифференциальной условной энтропией или

просто *дифференциальной условной энтропией* непрерывного источника. Она характеризует неопределенность выбора непрерывной случайной величины U при условии, что известны результаты реализации значений другой статистически связанной с ней непрерывной случайной величины V , и по сравнению со средней неопределенностью выбора случайной величины U' , изменяющейся в диапазоне, равном единице, и имеющей равномерное распределение вероятностей.

- При втором подходе к использованию соотношения (3.31) для количественного определения информационных свойств непрерывного источника информации предлагается принять во внимание практическую невозможность обеспечения бесконечно высокой точности различения определенных значений непрерывной величины U . Поэтому все бесконечное число значений U в пределах заданной точности измерений следует рассматривать как одно значение.

Из средней неопределенности выбора источником u некоторого значения в этом случае необходимо вычесть среднюю неопределенность того же источника, полученную при условии, что мы знаем результаты определения U с некоторой определенной точностью ε . Тогда информационные свойства непрерывного источника будут оцениваться разностью безусловной и условной энтропий, определяемых соотношениями (3.31) и (3.33) соответственно.

Такая разность, как будет показано позже, является мерой снятой неопределенности, называемой количеством информации.

Таким образом, при втором подходе безусловная и условная энтропии непрерывного источника рассматриваются лишь как некоторые вспомогательные величины, с помощью которых можно определить количество информации. Соотношение между понятиями энтропии и количества информации для непрерывного источника информации подобно соотношению между потенциалом, определенным с привлечением понятия бесконечности, и напряжением, определенным как разность потенциалов.

Поскольку вторые члены в правых частях соотношений (3.31) и (3.33) одинаковы, разность безусловной и условной энтропий непрерывного источника информации равна разности дифференциальных безусловной и условной энтропий того же источника, причем относительность их уже несущественна, так как разность не зависит от стандарта, с которым они сравнивались.

Свойства дифференциальной энтропии

1. Дифференциальная энтропия в отличие от энтропии дискретного источника является относительной мерой неопределенности. Ее значение зависит от масштаба случайной величины U , а, следовательно, - и от выбора единицы ее измерения.

Изменим масштаб случайной величины U , например, в k раз, оставив неизменным масштаб равномерно распределенной в единичном интервале случайной величины U' , принятой за эталон. Если $u_k = ku$, то $p(u_k) = p(u)/k$.

Тогда

$$\begin{aligned} h(u) &= - \int_{-\infty}^{\infty} p(u_k) \log p(u_k) du_k = \\ &= - \int_{-\infty}^{\infty} \frac{p(u)}{k} \log \left[\frac{p(u)}{k} \right] k du = - \int_{-\infty}^{\infty} p(u) \times \\ &\times \log p(u) du + \log k \int_{-\infty}^{\infty} p(u) du = h(u) + \log k. \end{aligned} \quad (3.35)$$

Если одновременно изменить масштаб величины U' , то относительная неопределенность также изменится, так как значение эталона будет уже иным.

Из относительности дифференциальной энтропии следует, что энтропия может принимать положительные, отрицательные и нулевые значения.

2. Дифференциальная энтропия не зависит от конкретных значений случайной величины U и, в

частности, от изменения всех ее значений на постоянное. Действительно, масштаб U при этом не меняется и справедливо равенство

$$\begin{aligned} h(U + \Theta) &= - \int_{-\infty}^{\infty} p(u + \Theta) \log p(u + \Theta) d(u + \Theta) = \\ &= - \int_{-\infty}^{\infty} p(u) \log p(u) du = h(U). \end{aligned} \quad (3.36)$$

3. Какие же непрерывные распределения обладают максимальной дифференциальной энтропией?

А). Если единственным ограничением для случайной величины U является область ее возможных значений $[\alpha, \beta]$, то максимальной дифференциальной энтропией обладает равномерное распределение вероятностей в этой области.

При доказательстве решается задача определения плотности распределения $p(u)$, обеспечивающей максимальное значение функционала

$$h(U) = - \int_{\alpha}^{\beta} p(u) \log p(u) du$$

при ограничении

$$\int_{\alpha}^{\beta} p(u) du = 1. \quad (3.37)$$

Используя, например, метод неопределенных множителей Лагранжа, получим

$$p(u) = 1/(\beta - \alpha), \quad \alpha \leq u \leq \beta. \quad (3.38)$$

Нетрудно убедиться в том, что найденная функция $p(u)$ обеспечивает максимум функционала $h(U)$, причем

$$h(U) = \log(\beta - \alpha). \quad (3.39)$$

Б). Если ограничения на область значений непрерывной случайной величины U отсутствуют, но известно, что дисперсия ее ограничена, то максимальной дифференциальной энтропией обладает нормальное распределение величины U .

При доказательстве решается задача определения функции $p(u)$, обеспечивающей максимальное значение функционала

$$h(U) = - \int_{-\infty}^{\infty} p(u) \log p(u) du \quad (3.40)$$

при ограничениях

$$\int_{-\infty}^{\infty} p(u) du = 1 \quad \text{и} \quad \int_{-\infty}^{\infty} u^2 p(u) du = \sigma^2,$$

где σ — среднеквадратическое отклонение от математического ожидания $\bar{U} = 0$ (σ — заданное ограничение).

Искомую плотность распределения $p(u)$ находят, например, методом неопределенных множителей Лагранжа.

Она оказывается гауссовской:

$$p(u) = \frac{1}{\sigma\sqrt{2\pi}} e^{-u^2/2\sigma^2}. \quad (3.41)$$

Вычислив функционал (3.40) для этого распределения, получим значение максимальной дифференциальной энтропии $h_{\max}(U)$. В двоичных единицах неопределенности

$$h_{\max}(U) = \log_2 \sigma \sqrt{2\pi e}. \quad (3.42)$$

Поскольку в информационных системах сигнал, описываемый случайной величиной U , часто представляет собой электрическое напряжение (или ток), дисперсия U пропорциональна средней мощности сигнала. Тогда в соответствии с (3.41) можно утверждать, что при заданной мощности

наибольшей средней неопределенностью выбора будет обладать источник, генерирующий сигналы, амплитуды которых распределены по нормальному закону.

4. Соотношения для дифференциальной энтропии объединения статистически зависимых непрерывных источников аналогичны соответствующим формулам для дискретных источников:

$$h(UV) = h(U) + h_v(V) = h(V) + h_u(U), \quad (3.43)$$

где

$$h(UV) = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(uv) \log p(uv) du dv.$$

Справедливость соотношения (3.43) легко проверить подстановкой выражения (3.32), заданного для $h(V)$, и выражения (3.34) — для $h_v(U)$.

Так как

$$h_v(U) \leq h(U) \text{ и } h_u(V) \leq h(V),$$

то

$$h(UV) \leq h(U) + h(V),$$

причем равенство имеет место только в случае отсутствия статистической связи между U и V .

Пример 3.6. Определить, насколько мы выиграем в мощности, используя для организации мешающего воздействия, характеризующегося энтропией, источник шума с гауссовской плотностью распределения по сравнению с источником, имеющим в интервале $[\alpha, \beta]$ равномерную плотность распределения.

В соответствии с (3.42) дифференциальная энтропия гауссовского распределения

$$h_{\max}(u) = \log \sigma_r \sqrt{2\pi e},$$

где σ_r^2 — дисперсия, характеризующая мощность, выделяемую на резисторе с сопротивлением в 1 Ом.

Для равномерного распределения энтропия определена соотношением (3.39):

$$h_{\max}(U) = \log(\beta - \alpha).$$

Вычислим дисперсию σ_p^2 равномерного на интервале $[\alpha, \beta]$ распределения:

$$\sigma_p^2 = \int_{\alpha}^{\beta} \left(u - \frac{\alpha + \beta}{2} \right)^2 \frac{1}{\beta - \alpha} du = \frac{(\beta - \alpha)^2}{12}. \quad (3.44)$$

Из условия обеспечения равенства энтропий следует

$$\log \sigma_r \sqrt{2\pi e} = \log(\beta - \alpha), \quad (3.45)$$

$$\sigma_r \sqrt{2\pi e} = \beta - \alpha \quad (3.46)$$

Возведя (3.46) в квадрат и подставив в (3.44), получим

$$\sigma_p^2 = \frac{\pi e}{6} \sigma_r^2 = 1,42 \sigma_r^2.$$

Следовательно, искомый выигрыш составляет 42 %.

Количество информации как мера снятия энтропии (неопределённости)

Передача информации инициируется либо самим источником информации, либо осуществляется по запросу. Она диктуется желанием устранить неопределенность относительно последовательности состояний, реализуемых некоторым источником информации. Обычно запрос обусловлен отсутствием возможности наблюдать состояния источника непосредственно. Поэтому абонент обращается к информационной системе, которая извлекает интересующую его информацию из источника посредством некоторого первичного преобразователя и направляет ее по каналу связи абоненту.

Информация проявляется всегда в форме сигналов. Сигналы z , поступающие с выхода первичного преобразователя источника информации на вход канала связи, принято называть

сообщениями в отличие от сигнала u , формирующихся на входе линии связи. В зависимости от формы создаваемых сообщений различают источники дискретных и непрерывных сообщений.

Отдельные первичные сигналы с выхода источника дискретных сообщений называют *элементами сообщения*. Каждому элементу сообщения соответствует определенное состояние источника информации. В случае параллельной реализации источником информации множества состояний, как это имеет место, например, в документах с печатным текстом, первичный преобразователь, в частности, обеспечивает их последовательное отображение элементами сообщения. Таким преобразователем может быть, как автоматическое читающее устройство, так и человек.

Основное понятие теории информации — *количество информации* — рассматривается в данном параграфе применительно к передаче отдельных статистически несвязанных элементов сообщения. Дискретный источник сообщений при этом полностью характеризуется ансамблем

$$Z = \begin{pmatrix} z_1 & \dots & z_i & \dots & z_N \\ p(z_1) & \dots & p(z_2) & \dots & p(z_N) \end{pmatrix},$$

а непрерывный — одномерной плотностью распределения случайной величины — z — $p(z)$. Особенности определения количества информации при передаче сообщений изложены в § 4.2.

Передача информации от дискретного источника. Выясним, насколько будет изменяться неопределенность относительно состояния источника сообщения при получении адресатом элемента сообщения с выхода канала связи. Алфавиты передаваемых и принимаемых элементов сообщения считаем идентичными.

Вследствие воздействия помех полученный элемент сообщения в общем случае отличается от переданного. Подчеркнем это различие. Обозначив принимаемые элементы сообщения другими буквами: $w_1, \dots, w_j, \dots, w_N$.

Априорная неопределенность (неопределенность до получения элемента сообщения) относительно состояния источника не является полной. Предполагается, что адресату известен алфавит элементов сообщения, а из прошлого опыта он знает вероятности их появления. Считая, что состояния источника реализуются независимо, *априорная* частная неопределенность появления элемента сообщения z_i

$$H(z_i) = -\log p(z_i),$$

где $p(z_i)$ — априорная вероятность появления элемента сообщения z_i .

Предполагаются также известными некоторые сведения относительно помехи в канале связи. Обычно считают, что между элементами сообщения и помехой статистические связи отсутствуют, искажения отдельных элементов сообщения являются событиями независимыми и адресату известна совокупность условных вероятностей $p(z_i/w_j)$ ($1 \leq i \leq N, 1 \leq j \leq N$) того, что вместо элемента сообщения z_i , будет принят элемент сообщения w_j . При получении конкретного элемента сообщения w_j , адресату становится известным значение условной вероятности $p(z_i/w_j)$, называемой *апостериорной* (послеопытной) вероятностью реализации источником элемента сообщения z_i . Это позволяет найти апостериорную частную неопределенность, остающуюся у адресата относительно выдачи источников элемента сообщения z_i после получения конкретного элемента сообщения w_j :

$$H(z_i/w_j) = -\log p(z_i/w_j). \quad (3.46)$$

Поскольку получение информации мы связываем с уменьшением неопределенности, естественно определить частное количество информации $I(z_i)$, получаемое при приеме элемента сообщения w_j относительно некоторого реализованного источником элемента сообщения z_i как разность частных неопределенностей, имевшихся у адресата до и после получения элемента сообщения (априорной и апостериорной):

$$I(z_i) = H(z_i) - H(z_i/w_j) = \log \frac{p(z_i/w_j)}{p(z_i)}. \quad (3.47)$$

Анализ формулы (3.47) позволяет сделать следующие заключения:

1) частное количество информации растет с уменьшением априорной и увеличением апостериорной вероятностей реализации элемента сообщения источником, что находится в полном соответствии с нашими интуитивными представлениями;

2) частное количество информации об элементе сообщения z_i может быть не только положительным, но и отрицательным, а также нулем, что зависит от соотношения априорной $p(z_i)$ и апостериорной $p(z_i/w_j)$ вероятностей. Если вероятность того, что источником был реализован элемент сообщения z_i увеличилась после приема элемента сообщения w_j , т. е. $p(z_i/w_j) > p(z_i)$, то полученное частное количество информации положительно. Если эта вероятность не изменилась, т. е. $p(z_i/w_j) = p(z_i)$, то имевшая место неопределенность тоже не изменилась и частное количество информации равно нулю.

Наконец, случай $p(z_i/w_j) < p(z_i)$ соответствует увеличению неопределенности относительно реализации z_i после получения элемента сообщения w_j , и, следовательно, частное количество информации отрицательно;

3) в случае отсутствия помехи апостериорная вероятность $p(z_i/w_j) = 1$. При этом частное количество информации численно совпадает с частной априорной неопределенностью реализации данного элемента сообщения z_i :

$$I(z_i) = H(z_i) = -\log p(z_i).$$

Это максимальное частное количество информации, которое можно получить об элементе сообщения z_i ;

4) частное количество информации относительно реализации источником элемента сообщения z_i , содержащееся в принятом элементе сообщения w_j , равно частному количеству информации относительно w_j , содержащемуся в элементе сообщения z_i :

$$I(z_i, w_j) = \log \frac{p(z_i/w_j)}{p(z_i)} = \log \frac{p(z_i, w_j)}{p(z_i)p(w_j)} = I(w_j, z_i). \quad (3.48)$$

Хотя имеют место случаи, где важно оценить частное количество информации $I(z_i, w_j)$, для задач анализа и оптимизации функционирования информационных систем более рациональны усредненные характеристики, отражающие статистические свойства источника информации и канала связи.

Найдем среднее количество информации, содержащееся в любом принятом элементе сообщения относительно переданного (реализованного) источником. До получения конкретного элемента сообщения средняя неопределенность, имеющаяся у адресата, относительно реализации источником любого элемента сообщения равна энтропии источника. Ее называют *априорной энтропией* источника.

Средняя неопределенность относительно любого состояния источника, остающаяся у адресата после получения конкретного элемента сообщения w_j , характеризуется частной условной энтропией $H_{w_j}(z)$:

$$H_{w_j}(Z) = - \sum_{i=1}^N p(z_i/w_j) \log p(z_i/w_j). \quad (3.49)$$

Это случайная величина, зависящая от того, какой конкретно элемент сообщения принят.

Средняя неопределенность по всему ансамблю принимаемых элементов сообщений равна условной энтропии источника $H_w(Z)$:

$$H_W(Z) = \sum_{j=1}^N p(w_j) H_{w_j}(Z),$$

или

$$H_W(Z) = - \sum_{j=1}^N \sum_{i=1}^N p(w_j) p(z_i/w_j) \log p(z_i/w_j). \quad (3.50)$$

Эту условную энтропию называют *апостериорной энтропией* источника информации.

Таким образом, при наличии помех среднее количество информации, содержащееся в каждом принятом элементе сообщения, относительно любого переданного равно разности априорной и апостериорной энтропии источника:

$$I(Z) = H(Z) - H_W(Z). \quad (3.51)$$

Представив априорную и апостериорную энтропии соответственно выражениями (3.6) и (3.50) и проведя несложные преобразования, получим формулу для количества информации непосредственно через вероятности:

$$\begin{aligned} I(ZW) &= - \sum_{i=1}^N p(z_i) \log p(z_i) + \sum_{i=1}^N \sum_{j=1}^N p(z_i w_j) \log p(z_i/w_j) = \\ &= - \sum_{i=1}^N \sum_{j=1}^N p(z_i w_j) \log p(z_i) - \sum_{i=1}^N \sum_{j=1}^N p(z_i w_j) \log p(z_i/w_j) = \\ &= \sum_{i=1}^N \sum_{j=1}^N p(z_i w_j) \log \frac{p(z_i w_j)}{p(z_i) p(w_j)}. \end{aligned} \quad (3.52)$$

Если частный характер количества информации специально не оговаривается, мы всегда имеем дело с количеством информации, приходящимся в среднем на один элемент сообщения. Поэтому указание об усреднении опускается.

Передача информации от непрерывного источника. Количество информации, получаемой от непрерывного источника по каналу с помехами, определяется так же, как в случае, рассмотренном выше, но с использованием понятия дифференциальной энтропии.

Для источника, имеющего непрерывное множество состояний, среднее количество информации, содержащееся в каждом принятом значении случайной величины W относительно переданного значения случайной величины Z , можно получить как разность априорной и апостериорной дифференциальных энтропий:

$$I(ZW) = h(Z) - h_W(Z). \quad (3.53)$$

Соотношение (3.53) несложно выразить в виде, подобном (3.52):

$$I(ZW) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(zw) \log \frac{p(zw)}{p(z)p(w)} dz dw. \quad (3.54)$$

Относительность дифференциальных энтропий в этом случае не принимается во внимание, поскольку количество информации не зависит от выбранного стандарта сравнения.

Основные свойства количества информации. 1. Несмотря на то, что частное количество информации может быть величиной отрицательной, количество информации неотрицательно.

Действительно, согласно выражению

$$H_W(Z) \leq H(Z). \quad (3.55)$$

Тогда

$$I(ZW) = H(Z) - H_W(Z) \geq 0.$$

2. При отсутствии статистической связи между случайными величинами Z и W

$$H_W(Z) = H(Z), \quad (3.56)$$

следовательно, в этом случае

$$I(ZW) = 0$$

(принятые элементы сообщения не несут никакой информации относительно переданных).

3. Количество информации в W относительно Z равно количеству информации в Z относительно W .

Для доказательства этого утверждения воспользуемся выражением

$$H(ZW) = H(Z) + H_z(W).$$

Аналогично можно записать

$$H(WZ) = H(W) + H_w(Z).$$

Так как

$$H(ZW) = H(WZ),$$

то

$$H(Z) + H_z(W) = H(W) + H_w(Z),$$

откуда

$$H(Z) - H_w(Z) = H(W) - H_z(W), \quad I(ZW) = I(WZ) \quad (3.57)$$

4. При взаимно однозначном соответствии между множествами передаваемых и принимаемых элементов сообщений, что имеет место в отсутствии помехи, апостериорная энтропия равна нулю и количество информации численно совпадает с энтропией источника:

$$I(ZW) = H(Z). \quad (3.58)$$

Это максимальное количество информации о состоянии дискретного источника. Для непрерывного источника оно равно бесконечности.

Пример 3.7. Выстрел из орудия не поражает цель с вероятностью p . Через какое число выстрелов следует поинтересоваться у разведчика-корректировщика, уничтожена ли цель, чтобы в результате ответа получить максимальное количество информации?

Ансамбль интересующих нас событий включает: z_1 — цель поражена; z_2 — цель не поражена. Вероятность того, что цель не поражена после k выстрелов, равна p^k . Вероятность противоположного события $(1 - p^k)$. Поскольку после ответа корректировщика неопределенность устраняется полностью, количество информации равно энтропии, а она максимальна при равновероятности событий. Следовательно,

$$p^k = 1 - p^k,$$

откуда

$$k = -1/\log_2 p.$$

Пример 3.8. Определить среднее количество информации, получаемое при передаче элемента сообщения по каналу, описанному матрицей совместных вероятностей передачи и приема элементов сообщения

$$p(wz) = \begin{vmatrix} 0,4 & 0,1 & 0 \\ 0 & 0,2 & 0,1 \\ 0 & 0 & 0,2 \end{vmatrix}$$

Безусловные вероятности посылаемых z и принимаемых w элементов сообщения определены при рассмотрении примера 3 4. Там же получены значения для априорной $H(Z)$ и апостериорной $H_w(Z)$ энтропий.

В соответствии с (3.51)

$$I(ZW) = H(Z) - H_w(Z) = 1,485 - 0,55 = 0,935 \text{ дв. ед}$$

Ранее было показано, что неопределенность реализации непрерывным источником информации состояния в конкретный момент времени (отсчета) равна бесконечности. Тем более равна бесконечности неопределенность реализации непрерывным источником конкретного сигнала длительности T .

Однако такой результат получен в предположении возможности фиксировать любые сколь

угодно малые различия между реализациями. На практике такая возможность отсутствует. Это объясняется тем, что воспринимающие информацию датчики, включая человека, обладают ограниченной чувствительностью и конечной разрешающей способностью, а также тем, что процесс восприятия сопровождается помехами.

Если учесть, что нас интересует приближенное восприятие реализации, то количество информации, приходящееся на отсчет или на единицу времени, можно вычислить.

Ограничимся рассмотрением простейшего случая, когда отдельные состояния источника информации представляют собой независимые реализации случайной величины U . Ансамбль реализаций случайной величины U описывается плотностью распределения вероятностей $p(u)$. О значениях случайной величины U можно судить по значениям другой случайной величины Z , если мера их различия не превышает заданной верности воспроизведения. В этом случае говорят, что Z воспроизводит U .

Для количественной оценки степени сходства сигналов целесообразно ввести какую-либо функцию $\rho(z, u)$, имеющую природу «расстояния». Тогда удобным критерием верности $V(Z, U)$ является среднее значение функции $\rho(z, u)$, взятое по всему множеству значений z и u :

$$V(ZU) = \iint \rho(zu) p(zu) dz du, \quad (3.59)$$

где $\rho(z, u)$ — плотность совместного распределения вероятностей величин Z и U .

Наиболее широко используется среднеквадратический критерий, при котором $\rho(z, u)$ представляет собой квадрат обычного евклидова расстояния между точками в соответствующем пространстве. Требование к верности в данном случае задается с использованием критерия $V(ZU)$:

$$\tilde{V}(ZU) = \iint p(u) p_u(z) (z - u)^2 du dz \leq \varepsilon^2, \quad (3.60)$$

где $p_u(z)$ — условная плотность распределения — функция правдоподобия того, что конкретный сигнал u будет воспроизведен как сигнал z ; ε — заданное значение верности.

Так как плотность $p(u)$ определена, то для выполнения условия (3.60) варьировать можно только условной плотностью распределения $p_u(z)$.

Если случайная величина Z воспроизводит случайную величину U с некоторой верностью ε , то количество информации, содержащееся в воспроизводящей величине Z относительно U , конечно и в соответствии с (3.54) может быть записано в форме

$$I(ZU) = \iint p(u) p_u(z) \log \frac{p_u(z)}{p(z)}. \quad (3.61)$$

где

$$p(z) = \int p(u) p_u(z) du$$

- плотность воспроизводящей величины Z . Нужно обеспечить заданную верность воспроизведения при минимальном количестве получаемой информации. Поэтому среди множества функций $p_u(z)$, удовлетворяющих условию (3.60), целесообразно выбрать такую, которая обеспечивает наименьшее $I(ZU)$ [10, 35].

Минимальное количество информации в одной случайной величине Z относительно другой U , при котором удовлетворяется заданное требование к верности воспроизведения величины U , называется ε -энтропией величины U и обозначается $H_\varepsilon(U)$:

$$H_\varepsilon(U) = \min_{\{p_u(z)\}} I(ZU) \quad (3.62)$$

При

$$V(ZU) \leq \varepsilon^2. \quad (3.63)$$

Используя безусловную $h(U)$ и условную $h_z(U)$ дифференциальные энтропии величины U , выражение (3.63) можно представить в виде

$$H_\varepsilon(U) = h(U) - \max_{\{p_z(u)\}} h_z(U), \quad (3.64)$$

где $p_z(u)$ — условная плотность вероятности того, что в тех случаях, когда был принят сигнал z , передавался сигнал u .

Пример 3.9. Найти $H_\varepsilon(U)$ источника информации, ансамбль состояний которого описывается нормально распределенной случайной величиной U с дисперсией σ^2 при верности воспроизведения $V'(ZU) \leq \varepsilon^2$.

Будем считать, что заданная верность воспроизведения обусловлена действием аддитивной статистически не связанной с сигналом помехой Ξ , причем $M[\Xi] = 0$ и $M[\Xi^2] = \varepsilon^2$. Передаваемый сигнал u рассматриваем как сумму воспроизводящего сигнала z и помехи: $u = z + \xi$.

Так как в данном случае $h_z(U)$ в выражении (3.64) полностью определяется помехой

$$[h_z(U) = h_z(Z + \Xi) = h(\Xi)],$$

то

$$H_\varepsilon(U) = h(U) - \max_{\rho(\xi)} h(\Xi),$$

где $h(\Xi)$ — дифференциальная энтропия помехи; $\rho(\xi)$ — плотность распределения помехи Ξ .

Ранее [см. (3.41)] установлено, что при ограничении на дисперсию случайной величины максимальной дифференциальной энтропией обладает нормальное распределение. Поэтому в соответствии с (3.42) получаем

$$\max_{\rho(\xi)} h(\Xi) = \log \varepsilon \sqrt{2\pi e},$$

$$h(U) = \log \sigma \sqrt{2\pi e},$$

откуда

$$H_\varepsilon(U) = \log \sigma \sqrt{2\pi e} - \log \varepsilon \sqrt{2\pi e} = \frac{1}{2} \log \frac{\sigma^2}{\varepsilon^2}. \quad (3.65)$$

Так как σ^2 определяет среднюю мощность P_u сигнала, а ε^2 — среднюю мощность P_ξ помехи Ξ , то выражение (3.65) характеризует зависимость энтронии от величины P_u / P_ξ , называемой *отношением сигнал/помеха*.

При заданном отношении сигнал/помеха значение $H_\varepsilon(U)$ для нормально распределенной случайной величины является максимально возможным.

Для произвольно распределенной случайной величины U при том же критерии верности и малых ε [когда $H_\varepsilon(U)$ велико] справедливо приближенное равенство

$$H_\varepsilon(U) \approx h(U) - \log \varepsilon \sqrt{2\pi e}. \quad (3.66)$$

Контрольные вопросы

1. В чем сущность требования аддитивности к мере неопределенности выбора?
3. В каких единицах измеряется неопределенность выбора?
4. Какие требования предъявляются к мере неопределенности выбора из дискретного ансамбля?
5. Охарактеризуйте сущность понятия энтропии.
6. Изложите основные свойства энтропии дискретного ансамбля.
7. Запишите выражение для условной энтропии и поясните ее смысл.
8. Как определить энтропию нескольких взаимозависимых ансамблей?
9. Какова особенность определения энтропии непрерывного источника информации?
10. Дайте определение дифференциальной энтропии и сформулируйте ее основные свойства.
11. Какие распределения обладают максимальной дифференциальной энтропией:
 - а) при ограничении на диапазон изменения случайной величины?
 - б) при ограничении на дисперсию случайной величины?
12. Как связаны между собой понятия количества информации и энтропии?
13. Чем различаются понятия частного и среднего количества информации?
14. Сформулируйте основные свойства количества информации.