Энтропия, свойства энтропии.

Энтропия характеризует недостающую информацию. В изучение любых систем человек вводит человеческий фактор, что приводит к неточностям. Для измерения, определения этих неточностей и существует понятие неопределённости.

Для количественной оценки этой неопределённости было введено понятие, называемое энтропией, т.е. **энтропия – это количественная мера неопределённости.**

Количество информации, приходящееся на один элемент сообщения (знак, букву), называется **энтропией.**

Впервые понятие энтропии ввёл Клаузилус.

Свойства энтропии

Энтропия всегда неотрицательна, так как значения вероятностей выражаются дробными величинами, а их логарифмы – отрицательными величинами.

Энтропия равна нулю в том крайнем случае, когда одно событие равно единице, а все остальные – нулю. Это положение соответствует случаю, когда состояние источника полностью определено.

Энтропия имеет наибольшее значение при условии, когда все вероятности равны между собой.

Свойства энтропии

1) $H(\overline{p})$ непрерывна на интервале $0 \le p \le 1$, т.е. при малых изменениях p величина H изменяется мало.

2) $H(\overline{p})$ симметрична относительно p, т.е. не изменяется при любой перемене аргументов p_i.

3) Если событие U_N состоит из двух событий $U_N^{'}$ и $U_N^{''}$ с Вероятностями q_1 и q_2, $q_1 + q_2 = p_N$, то

$$H(p_1, p_2, \ldots, p_{N-1}, q_1, q_2) = H(p_1, p_2, \ldots, p_N) + p_N * H(q_1 / p_N, q_2 / p_N)$$

4) Энтропия всегда неотрицательна.

5) Энтропия равна нулю, когда вероятность одного из событий равна единице.

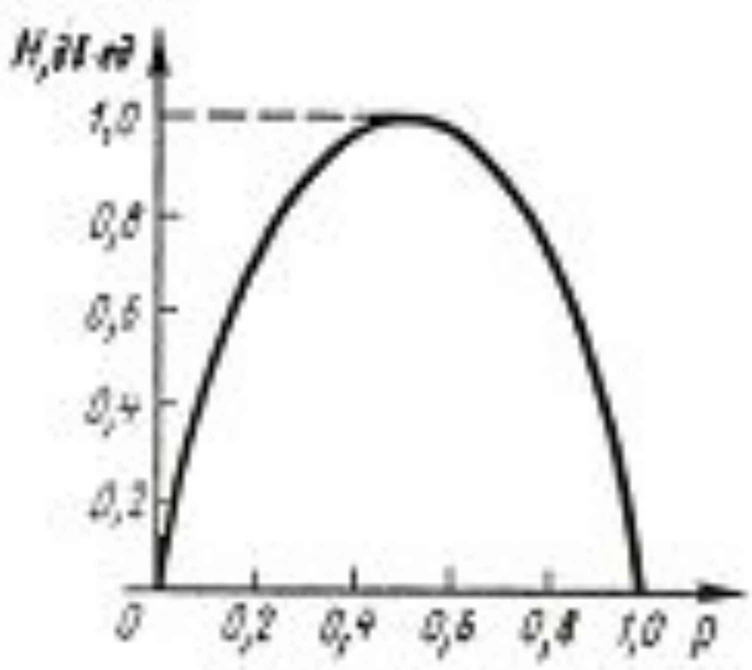
6) Энтропия имеет наибольшее значение, когда все вероятности равны между собой.

$$p_1 = p_2 = \ldots = p_N,$$

при этом

$$H = -\log \frac{1}{N} = \log N.$$

Изменение энтропии в зависимости от вероятности


1,0
0,8
0,6
0,4
0,2
0
0,2
0,4
0,6
0,8
1,0
p

Энтропия ансамбля

Ансамблем называется полная совокупность состояний с вероятностями их появлений, составляющими в сумме единицу:

$$X = \begin{pmatrix} X_1 & X_2 & \cdots & X_j & \cdots & X_k \\ P_1 & P_2 & \cdots & P_j & \cdots & P_k \end{pmatrix},$$

причем $\sum_{i=1}^{k} P_i = 1.$

Энтропия ансамбля вычисляется по формуле Шеннона.

Объединением называется совокупность двух и более взаимозависимых ансамблей дискретных случайных переменных.

Условная энтропия

Понятие условной энтропии в теории информации используется при определении взаимозависимости между:

а) символами кодируемого алфавита (между состояниями последовательно выбираемыми одним источником);

б) для определения потерь при передаче информации по каналам связи;

в) при вычислении энтропии объединения.

Общая условная энтропия сообщения В относительно сообщения А характеризует количество информации, содержащееся в любом символе алфавита, и определяется усреднением по всем символам, т.е. по всем состояниям с учетом вероятности появления каждого из состояний, и равна сумме вероятностей появления символов алфавита на неопределенность, которая остается после того, как адресат принял сигнал.

$$H(B/A) = -\sum_i p(a_i) H(b_j/a_i) = \\ = -\sum_i \sum_j p(a_i) p(b_j/a_i) \log p(b_j/a_i) \qquad (1)$$

где $$\mathbf{H(b_j/a_i) = -\sum_j p(b_j/a_i) \log p(b_j/a_i)} \qquad (2)$$

формула условной энтропии.

Выражение (1) является общим выражением для определения количества информации на один символ сообщения для случая неравномерных и взаимозависимых символов.

Так как p(ai)p(bj/ai) – вероятность совместного появления двух событий p(ai,bj), то формулу (1) можно записать следующим образом:

$$\mathbf{H(B/A) = -\sum_i \sum_j p(a_i, b_j) \log p(b_j / a_i)} \qquad (3)$$

В общем случае, если передаем m сигналов A и ожидаем получить m сигналов B, влияние помех в канале связи полностью описывается канальной матрицей размерностью **m×m**, содержащей элементы p(bj/ai). Сумма вероятностей распределения для каждой строки матрицы должна равняться единице.

Потери информации, которые приходятся на долю сигнала a_i, описываются при помощи частной условной энтропии вида (2).

Если исследуем канал со стороны приемника сообщений, то с получением сигнала bj предполагаем, что был послан какой – то из сигналов a1, a2, …, ai,…, am. При этом канальная матрица имеет размерность m×m и содержит элементы p(ai/bj). Единице должны равняться суммы условных вероятностей по столбцам канальной матрицы.

Частная условная энтропия при этом будет:

$$H(a_i/b_j) = -\sum_{i=1}^{m} p(a_i/b_j)\log p(a_i/b_j) \qquad (4)$$

Общая условная энтропия:

$$H(A/B) = -\sum_{j} p(b_j)\sum_{i} p(a_i/b_j)\log p(a_i/b_j) \qquad (5)$$

Если заданы безусловные вероятности источника и канальная матрица, то может быть вычислена энтропия **приемника**

$$H(B) = -\sum_{j} p(b_j)\log p(b_j) \qquad (6)$$

Энтропия источника сообщений может быть вычислена по формуле

$$\mathbf{H(A) = -\sum_{i} p(a_i) \log p(a_i)} \qquad (7)$$

Энтропия объединения используется для вычисления энтропии совместного появления статистически зависимых сообщений. H(A,B) – неопределенность того, что будет послано A, а принято B, и представляет собой сумму вида:

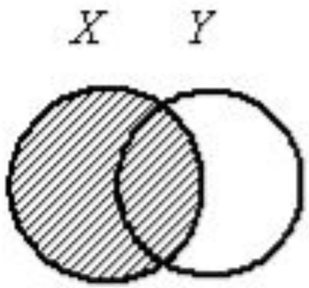
$$\mathbf{H(A,B) = -\sum_{i}\sum_{j} p(a_i, b_j) \log_2 p(a_i, b_j) \quad (бит/2 \;\; символа)} \qquad (8)$$

Элементы канальной матрицы имеют вид p(ai, bj), что позволяет вычислять энтропию как источника, так и приемника, непосредственно по канальной матрице

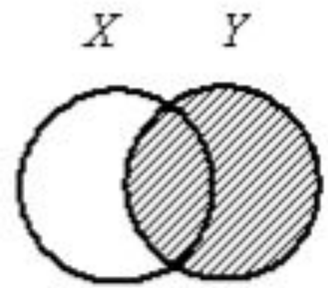
$$\mathbf{H(A) = -\sum_{i}\sum_{j} p(a_i, b_j) \log \sum_{j} p(a_i, b_j)} \qquad (9)$$

$$\mathbf{H(B) = -\sum_{i}\sum_{j} p(b_j, a_i) \log \sum_{i} p(b_j, a_i)} \qquad (10)$$

Безусловная энтропия

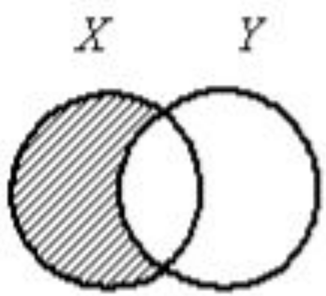


$H(X)=H(X/Y)+H(X\cdot Y)$

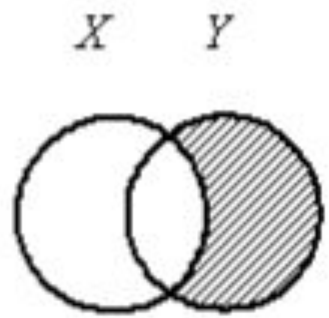


$H(Y)=H(Y/X)+H(X\cdot Y)$

Условная энтропия

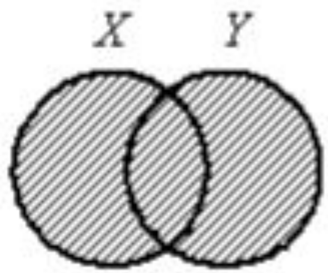


$H(X/Y)=H(X)-H(X\cdot Y)$



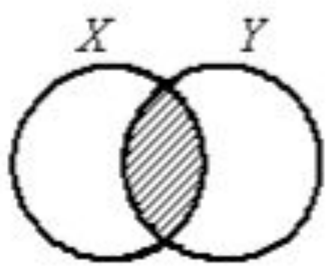
$H(Y/X)=H(Y)-H(X\cdot Y)$

Совместная энтропия

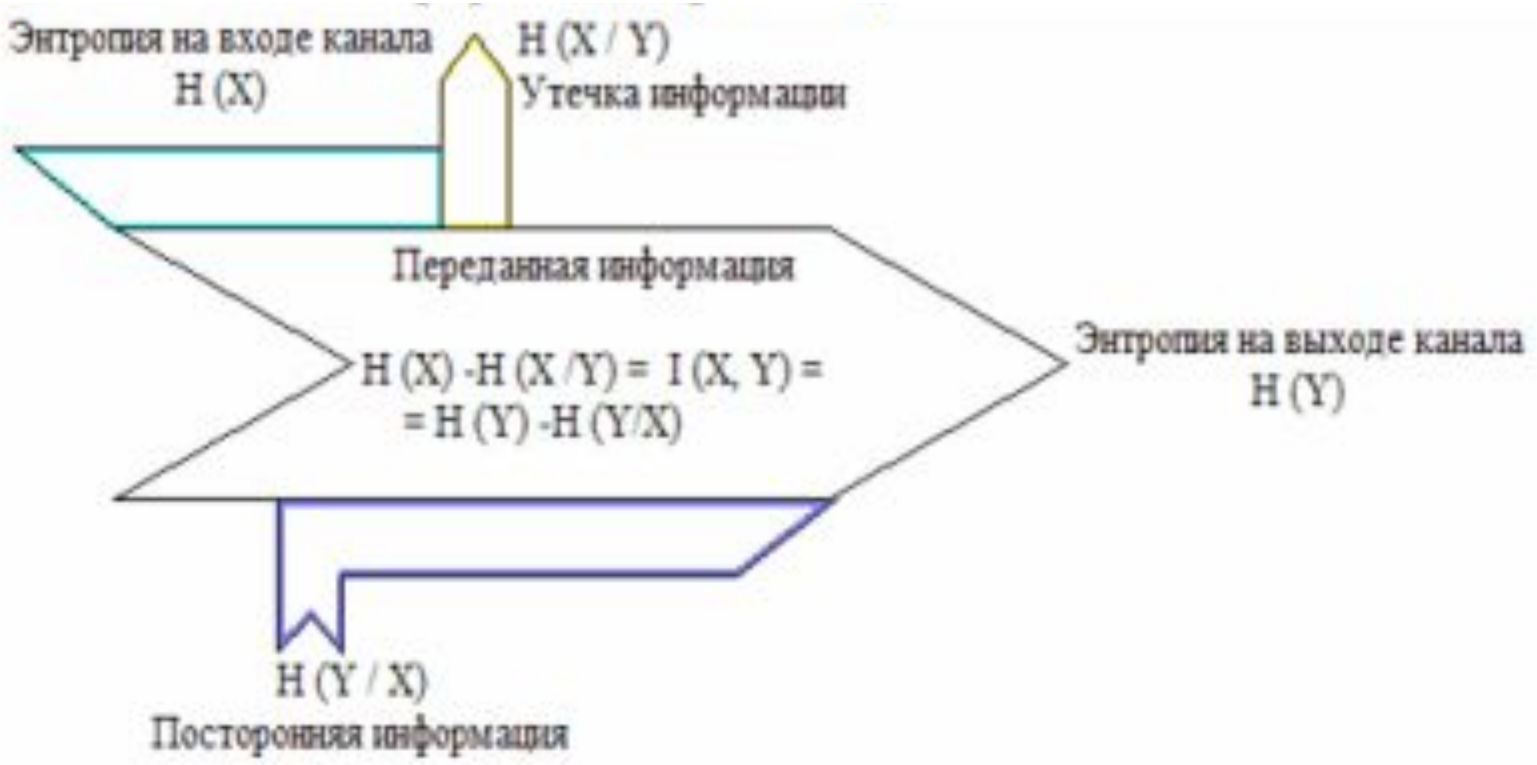


$$H(X,Y)=H(X)+H(Y/X)=H(Y)+H(X/Y)=H(X)+H(Y)-H(X\cdot Y)$$

Взаимная энтропия



$$H(X\cdot Y)=H(X)-H(X/Y)=H(Y)-H(Y/X)=H(X,Y)-H(X/Y)-H(Y/X)$$


Энтропия на входе канала
H (X)
H (X / Y)
Утечка информации
Переданная информация
H (X) -H (X /Y) = I (X, Y) =
= H (Y) -H (Y/X)
Энтропия на выходе канала
H (Y)
H (Y / X)
Посторонняя информация

Для вычисления среднего количества информации, содержащегося в принятом ансамбле сообщений В относительно переданного ансамбля сообщений А в условиях действия помех, пользуются следующими выражениями:

$$I(B,A)=\sum_i\sum_j p(a_i)p(b_j/a_i)\log\frac{p(b_j/a_i)}{p(b_j)} \qquad (11)$$

$$I(A,B)=\sum_i\sum_j p(b_j)p(a_i/b_j)\log\frac{p(a_i/b_j)}{p(a_i)} \qquad (12)$$

$$I(B,A)=I(A,B)=\sum_i\sum_j p(a_i,b_j)\log\frac{p(b_j/a_i)}{p(b_j)}=\sum_i\sum_j p(b_j,a_i)\log\frac{p(a_i/b_j)}{p(a_i)}=$$
$$=\sum_i\sum_j p(a_i,b_j)\log\frac{p(a_i,b_j)}{p(a_i)p(b_j)} \qquad (13)$$

I(A,B) показывает какое (в среднем) количество информации содержит сообщение *A* о сообщении *B* или наоборот сообщение *B* о сообщении *A*.

- $I(A,B)= H(A)-H(A/B)$
- $I(A,B)= I(B,A)= H(B)-H(B/A)$
- $I(A,B)= H(A)+ H(B)-H(B,A)$

Если *A* и *B* независимы, то $I(A,B)= 0$.

Если *A* и *B* полностью зависимы (содержат одну и ту же информацию), то $I(A,B)= H(A)= H(B)$.

Основные свойства количества информации

1. $I(X,Y) = I(Y,X)$, т.е. количество информации, содержащееся в случайном объекте Y о случайном объекте X, равно количеству информации, содержащемуся в случайном объекте X о случайном объекте Y.

2. $I(X,Y) \geq 0$, причём знак равенства имеет место, когда объекты X и Y независимы. Положительность $I(X,Y)$ следует из свойства энтропии : если события xi и yj статистически зависимы, то всегда $H(Y/X) < H(Y)$ и $H(X/Y) < H(X)$.

3. $I(X,Y) = H(X)$, т.е. энтропия может быть истолкована как информация, содержащаяся в объектах относительно самих себя.

Из этого также непосредственно вытекает, что энтропия есть максимальное количество информации, которое можно получить об объекте. Это возможно при взаимно однозначном соответствии между множествами передаваемых и принимаемых сообщений, что имеет место в отсутствии помехи, апостериорная энтропия равна нулю и количество информации численно совпадает с энтропией источника.

Если *A* и *B*- это сообщения на входе и на выходе канала связи с помехами, то для получения возможно большей информации ее получателем необходимо, чтобы взаимная информация была наибольшей. Тогда условная энтропия H(A/B) – это потери информации в канале связи (ненадежность канала). Условная энтропия H (B/A) – это информация о помехах (энтропия источника помех H(п)), поступающая в канал извне или создаваемая внутренними помехами в канале.

Потери в реальном канале связи

$$\Delta I = kH(B/A)$$

Среднее количество полученной информации

$$I = k[H(B) - H(B/A)] = k\,[\,H(B) - \Delta I]$$

В бесшумном канале H (A/B) = 0, т. е. информация передается от источника к приемнику и обратно без потерь. Если канал полностью зашумлен, то H (A/B) = H (A) и передачи информации не происходит. H (B/A) определяет среднюю неопределенность принятого символа при известных посланных символах, поэтому ее называют «посторонней» шумовой информацией.