ГЛОССАРИЙ

§ 3 Дискретные ансамбли

Нумерация соответствует нумерации в лекциях.

Говорят, что *на конечном множестве* $X = \{x_1, \ldots, x_m\}$ *задано распределение вероятностей* $p(x)$, если каждому сообщению $x_t \in X$ сопоставлено число $p(x_t)$, причем

$$\sum_{t=1}^{m} p(x_t) = 1,\ p(x_t) \geq 0,\ t = 1, 2, \ldots, m. \tag{3.1}$$

Конечное множество X вместе с заданным на нем распределением вероятностей $p(x)$ называется *дискретным вероятностным ансамблем* или коротко — дискретным ансамблем сообщений и обозначается $\{X, p(x)\}$.

Пусть A — подмножество множества X. *Вероятностью множества* A называется число

$$Pr(A) =: \sum_{x \in A} p(x).$$

Пусть $X = \{x_1, \ldots, x_m\}$ и $Y = \{y_1, \ldots, y_n\}$ — два конечных множества. Множество, элементы которого представляют собой все возможные упорядоченные пары

$$(x_i, y_j),\ x_i \in X,\ y_j \in Y,\ i = 1, \ldots, m;\ j = 1, \ldots, n,$$

называется *прямым произведением множеств* X и Y и обозначается через XY. Аналогичным образом определяются произведения более чем двух множеств. Произведение $X_1 X_2 \ldots X_n$ представляет собой множество всех последовательностей

$$X_1 \ldots X_n = \{(x_1, \ldots, x_n) \mid x_i \in X_i\}.$$

При задании ансамбля $\{XY, p(x, y)\}$, фактически задаются еще два ансамбля $\{X, p_1(x)\}$ и $\{Y, p_2(y)\}$:

$$X = \{x_1, \ldots, x_m\},\ p_1(x_i) =: \sum_{y_j \in Y} p(x_i, y_j), i = 1, 2, \ldots, m, \tag{3.2}$$

$$Y = \{y_1, \ldots, y_n\},\ p_2(y_j) =: \sum_{x_i \in X} p(x_i, y_j),\ j = 1, 2, \ldots, n \tag{3.2$'$}$$

Если распределение вероятностей на произведение двух множеств X и Y удовлетворяет условию

$$p(x_i, y_j) = p_1(x_i)p_2(y_j) \text{ для всех } x_i \in X, y_j \in Y, \tag{3.3}$$

то ансамбли X и Y называются *статистически независимыми*. В противном случае говорят, что эти ансамбли *статистически зависимы*.

Пусть задан ансамбль $\{XY, p(x, y)\}$. Предположим, что x_i — такой элемент множества X, для которого $p_1(x_i) \neq 0$. Число

$$p(y_j \mid x_i) =: \frac{p(x_i, y_j)}{p_1(x_i)} = \frac{p(x_i, y_j)}{\sum_{y_j \in Y} p(x_i, y_j)} \tag{3.4}$$

называется *условной вероятностью сообщения y_j при условии, что сообщение x_i известно*. Иногда это число называют *условной вероятностью сообщения y_j относительно сообщения x_i*. Множество $\{Y, p(y \mid x_i)\}$, где значения $p(y \mid x_i)$ определяются по формуле (3.4), называется *условным ансамблем относительно сообщения x_i*.

Пусть задан ансамбль $\{XY, p(x, y)\}$. Предположим, что y_j — такой элемент множества Y, для которого $p_2(y_j) \neq 0$. Число

$$p(x_i \mid y_j) =: \frac{p(x_i, y_j)}{p_2(y_j)} = \frac{p(x_i, y_j)}{\sum_{x_i \in X} p(x_i, y_j)} \tag{3.4$'$}$$

называется *условной вероятностью сообщения x_i при условии, что сообщение y_j известно*. Иногда это число называют *условной вероятностью сообщения x_i относительно сообщения y_j*. Множество $\{X, p(x \mid y_j)\}$, где значения $p(x \mid y_j)$ определяются по формуле (3.4′), называется *условным ансамблем относительно сообщения y_j*.

Пусть A — произвольное подмножество элементов из X такое, что

$$Pr_1(A) =: \sum_{x_i \in A} p_1(x_i) \neq 0,$$

где распределение $p_1(x_i)$ определено формулой (3.2). Число

$$p(y_j \mid A) =: \frac{1}{Pr_1(A)} \sum_{x_i \in A} p(x_i, y_j) \tag{3.5}$$

называется *условной вероятностью сообщения y_j относительно множества A*.

Пусть B — произвольное подмножество элементов из Y такое, что

$$Pr_2(B) =: \sum_{y_j \in B} p_2(y_j) \neq 0,$$

где распределение $p_2(y_j)$ определено формулой $(3.2')$. Число

$$p(x_i \mid B) =: \frac{1}{Pr_2(B)} \sum_{y_j \in B} p(x_i, y_j) \qquad (3.5')$$

называется *условной вероятностью сообщения x_i относительно множества B*.

Пусть задан ансамбль $\{XY, p(x, y)\}$ и A — подмножество элементов из X такое, что $Pr_1(A) \neq 0$. Множество $\{Y, p(y \mid A)\}$, где значения $p(y \mid A)$ определяются по формуле (3.5), называется *условным ансамблем относительно подмножество A из X*.

Пусть задан ансамбль $\{XY, p(x, y)\}$ и B — подмножество элементов из Y такое, что $Pr_2(B) \neq 0$. Множество $\{X, p(x \mid B)\}$, где значения $p(x \mid B)$ определяются по формуле $(3.5')$, называется *условным ансамблем относительно подмножество B из Y*.

ВЫВОД. Ансамбль $\{XY, p(x, y)\}$, заданный на произведение двух множеств X и Y, порождает три пары ансамблей:

1) $\{X, p_1(x)\}$ и $\{Y, p_2(y)\}$, где значения $p_1(x)$ и $p_2(y)$ определяются формулами (3.2) и $(3.2')$;

2) $\{X, p(x \mid y)\}$ и $\{Y, p(y \mid x)\}$, где значения $p(x \mid y)$ и $p(y \mid x)$ определяются формулами (3.4) и $(3.4')$;

3) $\{X, p(x \mid B)\}$, $B \subseteq Y$, и $\{Y, p(y \mid A)\}$, $A \subseteq X$, где значения $p(x \mid B)$ и $p(y \mid A)$ определяются формулами (3.5) и $(3.5')$.

§ 4 Энтропия

Пусть $\{X = \{x_1, \ldots, x_m\}, p(x)\}$ — ансамбль сообщений. *Собственной информацией* $I(x_i)$ сообщения x_i называется число

$$I(x_i) =: -\log p(x_i), \; i = 1, 2, \ldots, m. \qquad (4.1)$$

Число $I(x_i)$ также называют *количество собственной информации* в сообщении $x_i \in X$. В этой формуле подразумевается, что логарифмы вычисляются по основанию 2, что соответствует измерению информации в *битах*. Если бы логарифм вычислялся по натуральному основанию, единицей измерения информацией стал бы *нат*. Если основание десятичное — информация измеряется в *хартли*.

Свойства собственной информации

1. Неотрицательность. Собственная информация неотрицательна, т. е., $I(x) \geq 0$, $x \in X$. Она равна нулю только в том случае, когда сообщение имеет вероятность 1. Такое сообщение можно рассматривать как неслучайное и известное заранее до проведения опыта.

2. Монотонность. Если $x_1, x_2 \in X$, $p(x_1) \geq p(x_2)$, то $I(x_1) \leq I(x_2)$. *То сообщение имеет большую собственную информацию, которое имеет меньшую вероятность.*

3. Аддитивность. Рассмотрим ансамбль $\{XY, p(x, y)\}$. Если сообщения x_i и y_j статистически независимы, то

$$I(x_i, y_j) = -\log\ p(x_i, y_j) = -\log\ p_1(x_i) - \log\ p_2(y_j) = I(x_i) + I(y_j).$$

Энтропия

Энтропией дискретного ансамбля $\{X, p(x)\}$ называется число

$$H(X) =: \sum_{x \in X} I(x) p(x) = -\sum_{x \in X} p(x) \log\ p(x). \tag{4.2}$$

Энтропия представляет собой среднее количество собственной информации в сообщениях ансамбля $\{X, p(x)\}$. *Избыточностью* дискретного ансамбля $\{X, p(x)\}$ называется число

$$R =: \frac{\log\ |X| - H(X)}{\log\ |X|}.$$

Если у дискретного ансамбля $\{X = \{x_1, \ldots, x_m\}, p(x_i)$ все сообщения равновероятны: $p(x_1) = \ldots = p(x_m) = 1/m$, то

$$H(X) = \log\ m. \tag{4.3}$$

Свойства энтропии

Свойство 4.1. $H(X) \geq 0$, т. е. энтропия всякого дискретного ансамбля неотрицательна. Равенство нулю возможно в том и только в том случае, когда существует некоторое сообщение $x_i \in X$, для которого $p(x_i) = 1$, а вероятности остальных сообщений равны нулю.

Свойство 4.2. $H(X) \leq \log\ \mid X \mid$. Равенство имеет место в том и только в том случае, когда элементы ансамбля равновероятны.

Свойство 4.3. Если у ансамблей X и Y распределения вероятностей представляют собой одинаковые наборы чисел, то $H(X) = H(Y)$.

Свойство 4.4. Если ансамбли X и Y статистически независимы, то $H(XY) = H(X) + H(Y)$.

Свойство 4.5. Энтропия — выпуклая $\cap$ функция распределения вероятностей на элементах ансамбля X.

Свойство 4.6. Пусть задан ансамбль $\{X, p(x)\}$ и $A \subseteq X$. Введем ансамбль $\{X, p'(x)\}$, задав распределение вероятностей $p'(x)$ следующим образом: $p'(x) =: Pr(A)/ \mid A \mid$, если $x \in A$, $p'(x) = p(x)$, если $x \notin A$. Тогда $H(X') \geq H(X)$. Иными словами, «выравнивание» вероятностей элементов ансамбля не приводит к уменьшению энтропии.

Свойство 4.7. Пусть задан ансамбль $\{X, p(x)\}$ и на множестве его элементов определена функция $g(x)$. Для множества $Y = \{y = g(x) \mid x \in \in X\}$ положим $p(g(x)) = p(x)$ и введем ансамбль $\{Y, p(g(x))\}$. Тогда $H(Y) \leq H(X)$. Равенство имеет место тогда и только тогда, когда функция $g(x)$ обратима.

§ 5 Условная энтропия

Пусть задано произведение $\{XY, p(x, y)\}$ двух ансамблей и $y_j \in \in Y$ — такой элемент, для которого $p_2(y_j) \neq 0$. *Условной собственной информацией сообщения x_i при фиксированном y_j* называют число

$$I(x_i \mid y_j) =: -\log p(x_i \mid y_j) \overset{(3.4')}{=} -\log \frac{p(x_i, y_j)}{\sum_{x \in X} p(x, y_j)}. \tag{5.1}$$

Условной энтропией ансамбля X при фиксированном сообщении $y_j \in Y$ называется число

$$H(X \mid y_j) =: \sum_{x \in X} p(x \mid y_j) I(x_i \mid y_j) = -\sum_{x \in X} p(x \mid y_j) \log p(x \mid y_j). \tag{5.2}$$

Пусть задано произведение $\{XY, p(x, y)\}$ двух ансамблей и $x_i \in \in X$ — такой элемент, для которого $p_1(x_i) \neq 0$. *Условной собственной информацией сообщения y_j при фиксированном x_i* называют величину

$$I(y_j \mid x_i) =: -\log p(y_j \mid x_i) \overset{(3.4)}{=} -\log \frac{p(x_i, y_j)}{\sum_{y \in Y} p(x_i, y)}. \tag{5.1'}$$

Условной энтропией ансамбля Y при фиксированном сообщении $x_i \in X$

называется число

$$H(Y \mid x_i) =: \sum_{y \in Y} p(y \mid x_i) I(y \mid x_i) = -\sum_{y \in Y} p(y \mid x_i) \log p(y \mid x_i). \quad (5.2')$$

Условной энтропией ансамбля X при фиксированном ансамбле Y называется число

$$H(X \mid Y) =: -\sum_{x \in X} \sum_{y \in Y} p(x \mid y) \log p(x \mid y). \quad (5.3)$$

Условной энтропией ансамбля Y при фиксированном ансамбле X называется число

$$H(Y \mid X) =: -\sum_{y \in Y} \sum_{x \in X} p(y \mid x) \log p(y \mid x). \quad (5.3')$$

Свойства условной энтропии

Свойство 5.1. Условная энтропия неотрицательна: $H(X \mid Y) \geq 0$.

Свойство 5.2. Условная энтропия ансамбля не превосходит безусловной энтропии того же ансамбля: $H(X \mid Y) \leq H(X)$, причем равенство имеет место в том и только в том случае, когда ансамбли X и Y статистически независимы.

Свойство 5.3. $H(XY) = H(X) + H(Y \mid X) = H(Y) + H(X \mid Y)$. Это свойство называют *свойством аддитивности энтропии*. В случае статистически независимых ансамблей выполняются равенства:

$$H(X \mid Y) = H(X),\ H(Y \mid X) = H(Y).$$

Поэтому из свойства 5.3 получаем свойство 4.4 безусловной энтропии:

Свойство 4.4. Если дискретные ансамбли $\{X, p(x)\}$ и $\{Y, p(y)\}$ статистически независимы, то $H(XY) = H(X) + H(Y)$.

Произведения трех ансамблей

Пусть $\{XYZ, p(x, y, z)\}$ — три совместно заданных ансамбля. По аналогии с формулами (3.2) для произведения трех ансамблей соотношения

$$p_1(x) =: \sum_{y \in Y} \sum_{z \in Z} p(x, y, z),\ p_2(y) =: \sum_{x \in X} \sum_{z \in Z} p(x, y, z), \quad (5.4)$$

$$p_3(z) =: \sum_{x \in X} \sum_{y \in Y} p(x, y, z), \tag{5.4}$$

задают распределения вероятностей $p_1(x)$, $p_2(y)$ и $p_3(z)$ на множествах X, Y и Z соответственно. Будем говорить, что вместе с ансамблем $\{XYZ, p(x, y, z)\}$ *совместно заданы три ансамбля*

$$\{X, p_1(x)\},\ \{Y, p_2(y)\},\ \{Z, p_3(z)\}.$$

Это означает, что распределения вероятностей на множествах X, Y и Z определяются по формулам (5.4).

Если распределение вероятностей на произведение множеств X, Y и Z удовлетворяет условию

$$p(x, y, z) = p_1(x) \cdot p_2(y) \cdot p_3(z) \text{ для всех } x \in X,\ y \in Y,\ z \in Z, \tag{5.5}$$

то ансамбли X, Y и Z называются *статистически независимыми*. В противном случае говорят, что эти ансамбли *статистически зависимы.*

При совместном задании трех ансамблей оказываются *совместно заданными совокупности по два ансамбля*. Так ансамбль $\{XY, p(x, y)\}$ определяется с помощью следующего соотношения:

$$p(x, y) =: \sum_{z \in Z} p(x, y, z). \tag{5.6}$$

Здесь под знаком суммы только переменная z пробегает все элементы множества Z, а пара (x, y) — фиксированная.

Аналогично, ансамбль $\{XZ, p(x, z)\}$ определяется соотношениями:

$$p(x, z) =: \sum_{y \in Y} p(x, y, z), \tag{5.6$'$}$$

а ансамбль $\{YZ, p(y, z)\}$ — соотношениями:

$$p(y, z) =: \sum_{x \in X} p(x, y, z). \tag{5.6$''$}$$

Условная вероятность $p(x \mid y, z)$ *сообщения* x *относительно пары сообщений* (y, z) определяется так:

$$p(x \mid y, z) =: \frac{p(x, y, z)}{\sum_{x \in X} p(x, y, z)}. \tag{5.7}$$

Условная собственная информация сообщения x *относительно па-*

ры сообщений (y, z) определяется так:

$$I(x \mid y, z) =: -\log p(x \mid y, z) =: -\log \frac{p(x, y, z)}{\sum_{x \in X} p(x, y, z)}. \tag{5.8}$$

Условной энтропией ансамбля X *относительно пары ансамблей* YZ называется число

$$H(X \mid YZ) =: -\sum_{XYZ} p(x, y, z) \log p(x \mid y, z) =: \sum_{XYZ} p(x, y, z) I(x \mid y, z). \tag{5.9}$$

По аналогии с этими определениями легко получить условные энтропии $H(Y \mid XZ)$, $H(Z \mid XY)$,

Свойство 5.5. $H(X \mid YZ) \leq H(X \mid Y)$ причем равенство имеет место в том и только в том случае, когда ансамбли X и Y условно независимы при всех $z \in Z$.

§ 6 Взаимная информация

Количеством информации в сообщении $x \in X$ *о сообщении* $y \in Y$ называется число

$$I(x; y) =: I(y) - I(y \mid x) \overset{(4.1);(5.1')}{=} -\log p(y) - (-\log p(y \mid x)) = \log \frac{p(y \mid x)}{p(y)}. \tag{6.1}$$

Количеством информации в сообщении $y \in Y$ *о сообщении* $x \in X$ называется число

$$I(y; x) =: I(x) - I(x \mid y) \overset{(4.1);(5.1)}{=} -\log p(x) - (-\log p(x \mid y)) = \log \frac{p(x \mid y)}{p(x)}. \tag{6.1'}$$

Количество информации $I(x; y)$ может принимать различные по знаку и величине конечные и бесконечные значения, но может быть не определено для некоторых пар сообщений. Неопределенность можно устранить, либо произвольным образом доопределив количество информации, либо исключив из рассмотрения сообщения $x \in X$, $y \in Y$, вероятности которых равны нулю.

Так как для любых $x \in X$ и $y \in Y$ таких, что $p(x) \neq 0$ и $p(y) \neq 0$, имеют места равенства

$$p(x, y) \overset{(3.4')}{=} p(x \mid y) p(y) \overset{(3.4)}{=} p(y \mid x) p(x), \text{ то}$$

$$\frac{p(x \mid y)}{p(x)} = \frac{p(y \mid x)}{p(y)}. \tag{6.2}$$

Поэтому

$$I(y; x) \overset{(6.1')}{=} \log \frac{p(x \mid y)}{p(x)} \overset{(6.2)}{=} \log \frac{p(y \mid x)}{p(y)} \overset{(6.1)}{=} I(x; y). \tag{6.3}$$

Таким образом, доказана следующая

Теорема 6.1. *Количество информации в сообщении x о сообщении y равно количеству информации в сообщении y о сообщении x:*

$$I(y; x) = I(x; y). \tag{6.3}$$

Теорема 6.1 показывает, что *количество информации есть симметрическая функция пары сообщений.* Поэтому величину $I(x; y)$ называют *количеством взаимной информации между сообщениями* x и y или просто *взаимной информацией между этими сообщениями.*

Формуле (6.1) можно придать симметричную форму:

$$I(x; y) \overset{(6.1)}{=}: \log \frac{p(y \mid x)}{p(y)} \overset{(6.2)}{=} \log \frac{p(x, y)}{p(x)p(y)}. \tag{6.4}$$

Средним количеством взаимной информации или просто *средней взаимной информацией между ансамблями* $\{X, p(x)\}$ и $\{Y, p(y)\}$ называется число

$$\mathbf{I}(X; Y) =: \sum_{XY} p(x, y) \log \frac{p(x \mid y)}{p(x)}. \tag{6.5}$$

Здесь $p(x)$ и $p(y)$ определены соотношениями (3.2) и (3.2′).

Средней взаимной информацией между ансамблем X и сообщением $y \in Y$ называется число

$$\mathbf{I}(X; y) =: \sum_{X} p(x \mid y) \log \frac{p(x \mid y)}{p(x)}. \tag{6.6}$$

Средней взаимной информацией между ансамблем Y и сообщением $x \in$ $\in X$ называется число

$$\mathbf{I}(x; Y) =: \mathbf{M}_x[I(x; y)] = \sum_{Y} p(y \mid x) \log \frac{p(y \mid x)}{p(x)}. \tag{6.6′}$$

Теорема 6.2. *Средняя взаимная информация между сообщением,* *вероятность которого отлична от нуля, и ансамблем, а также средняя взаимная информация между двумя ансамблями неотрицательна.*

Теорема 6.3. *Для любого сюръективного отображения $\varphi : X \to Z$ ансамбля X на ансамбль Z справедливо равенство:*

$$\mathbf{I}(X;Y) \geq \mathbf{I}(Z;Y), \tag{6.10}$$

причем равенство имеет место всегда, когда отображение биективно, т. е. когда каждому элементу $z \in Z$ соответствует единственный элемент $x \in \in X$.

§ 7 Непрерывные ансамбли

Непрерывный ансамбль $\{X, p(x)\}$ задается бесконечным множеством сообщений и плотностью распределения вероятностей $p(x)$.

Энтропией ансамбля $\{X, p(x)\}$ называется величина

$$H(X) =: -\int_{-\infty}^{\infty} p(x) \log p(x) dx. \tag{7.1}$$

Для произведения $\{XY, p(x,y)\}$ энтропия вводится аналогично:

$$H(XY) =: -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y) dx\, dy. \tag{7.2}$$

Частные распределения вероятностей, т. е. вероятности порожденных ансамблей $\{X, p_1(x)\}$ и $\{Y, p_2(y)\}$, находятся по формулам:

$$p_1(x) =: \int_{-\infty}^{\infty} p(x,y) dy;\ p_2(y) =: \int_{-\infty}^{\infty} p(x,y) dx. \tag{7.3}$$

Условные распределения вероятностей находятся по формулам:

$$p(x \mid y) =: \frac{p(x,y)}{p_2(y)};\ p(y \mid x) =: \frac{p(x,y)}{p_1(x)}. \tag{7.4}$$

Условная энтропия ансамбля $\{Y, p(y)\}$ относительно ансамбля $\{X, p(x)\}$ вычисляется по формуле:

$$H(Y \mid X) =: -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y) p(y \mid x) dx\, dy. \tag{7.5}$$

Взаимная информация сообщений $x \in X$ и $y \in Y$ определяется следующим образом:

$$I(x;y) =: \log \frac{p(x \mid y)}{p_1(x)p_2(y)}, \tag{7.6}$$

при этом предполагается, что $p_1(x) \neq 0 \neq p_2(y)$.

Интегрируя (7.6), получим выражение для средней взаимной информации:

$$\mathbf{I}(X;Y) =: -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y) \log \frac{p(x,y)}{p_1(x)p_2(y)} dx\, dy. \tag{7.7}$$

Введенные величины связаны между собой следующими соотношениями:

$$\mathbf{I}(X;Y) = H(X) - H(X \mid Y) = H(Y) - H(Y \mid X), \tag{7.8}$$

$$\mathbf{I}(X;Y) + H(X \mid Y) + H(Y \mid X) = H(XY). \tag{7.9}$$

Если для конечных ансамблей введенные характеристики имели определенный информационный смысл, то в ряде случаев для непрерывных ансамблей такого толкования провести не удается. Более того, при определенном подборе параметров, задающих плотности распределения вероятностей, энтропийные характеристики могут принимать и отрицательные значения. Они не инвариантны по отношению к преобразованию ансамблей и не могут быть истолкованы как математические ожидания от соответствующих собственных информаций.