В. Д. КОЛЕСНИК
Г. Ш. ПОЛТЫРЕВ

КУРС ТЕОРИИ ИНФОРМАЦИИ

В. Д. КОЛЕСНИК
Г. Ш. ПОЛТЫРЕВ

КУРС ТЕОРИИ ИНФОРМАЦИИ

*Допущено Министерством
высшего и среднего специального образования СССР
в качестве учебного пособия для студентов
высших технических учебных заведений*

МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1982

32.81
К60
УДК 62-50

Курс теории информации. Колесник В. Д., Полтырев Г. Ш. — М.: Наука. Главная редакция физико-математической литературы. 1982. — 416 с.

Теория информации представляет собой ветвь статистической теории связи, круг проблем которой можно охарактеризовать как исследования кодирования для обработки и передачи сообщений. Книга состоит из следующих пяти разделов: кодирование дискретных источников, кодирование в дискретных каналах, кодирование в непрерывных каналах, кодирование непрерывных источников и кодирование в системах с многими пользователями. Основные параграфы книги задуманы как пособие для студентов, впервые знакомящихся с теорией информации. Дополнительные параграфы, отмеченные звездочкой, предназначены для углубленного изучения «традиционной» теории информации и могут быть полезны аспирантам. Особое место в книге занимает глава, посвященная кодированию в системах с многими пользователями, содержащая наиболее поздние результаты теории информации. Для чтения этой части нужна определенная теоретико-информационная эрудиция. Каждая глава снабжена рядом задач и упражнений.
Табл. 8, илл. 56, библ. 70 назв.

Виктор Дмитриевич Колесник, Григорий Шоулович Полтырев

Курс теории информации
Редактор *Г. Л. Кацман*
Техн. редактор *И. Ш. Аксельрод*
Корректоры *О. А. Бутусова; Л.С. Сомова*

ИБ № 11892

Сдано в набор 27.01.82. Подписано к печати 11.11.82. Т-20315.
Формат 60×90¹/₁₆. Бумага типографская № 1. Гарнитура литературная.
Печать высокая. Усл. печ. л. 26. Уч.-изд. л. 28,99.
Тираж 14 000 экз. Заказ 119. Цена 1 р. 20 к.

Издательство «Наука»
Главная редакция физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

Ленинградская типография № 6 ордена Трудового Красного Знамени
Ленинградского объединения «Техническая книга» им. Евгении Соколовой
Союзполиграфпрома при Государственном комитете СССР
по делам издательств, полиграфии и книжной торговли.
193144, г. Ленинград, ул. Моисеенко, 10.

К $\frac{1502000000-157}{053(02)-82}$ 119-82

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Теория информации представляет собой ветвь статистической теории связи (ее часто с нею отождествляют), основы которой были заложены классическими трудами Н. Винера, А. Н. Колмогорова, В. А. Котельникова и К. Шеннона. Круг проблем, составляющих основное содержание теории информации (проблем «шенноновской теории информации»), можно охарактеризовать как исследование методов кодирования для экономного представления сообщений различных источников и для надежной передачи сообщений по каналам связи с шумом.

В основе теории информации лежит статистическое описание (статистические модели) источников сообщений и каналов связи, а также основанное на этом описании измерение количества информации между сообщениями по Шеннону, т. е. такое, при котором количество информации определяется только вероятностными свойствами сообщений и ни от каких других их свойств не зависит. В отличие от других разделов теории связи, например, теории обнаружения, теории оценивания, теории модуляции, алгебраической теории кодирования и т. д., предметом теории информации, как правило, являются теоремы, устанавливающие предельные возможности различных методов обработки и передачи сообщений. Эти предельные возможности зависят только от статистических свойств источников и каналов.

В качестве примеров можно привести три типичные задачи теории информации.

1. Предположим, что задан источник сообщений. Требуется найти наименьшее количество символов (например, двоичных), которое необходимо для указания последовательности сообщений, порожденных источником. При этом может быть задан критерий качества восстановления сообщений источника и требоваться указание последовательности сообщений с ошибкой относительно данного критерия качества, не превосходящей заданную величину.

2. Предположим, что задан канал связи. Требуется найти наибольшую возможную скорость передачи по этому каналу, при которой вероятность ошибочного приема сообщений может быть сделана произвольно малой.

3. Предположим, что заданы источник и канал, а также задан критерий качества. Требуется определить наименьшую возмож-

ную относительно данного критерия качества величину ошибки, которую можно достичь, передавая сообщения данного источника по данному каналу связи.

Всякий раз, решая подобные задачи, пытаются не только найти предельные значения количества двоичных символов, скорости передачи или величины ошибки, но и найти некоторый способ обработки сообщений (некоторый способ кодирования и декодирования), который позволяет достичь указываемых пределов. Однако очень часто не удается указать наилучший способ кодирования и декодирования. Поэтому теория информации, как правило, не дает непосредственных практических рекомендаций инженерам, проектирующим аппаратуру обработки и передачи сообщений. Тем не менее, она является важным инструментом анализа различных технических систем: телеметрических систем, систем передачи речи или телевизионных изображений, систем передачи данных, банков данных, различных систем управления и т. д.

На основе теории информации можно ответить на вопросы о предельных возможностях перечисленных систем, определить, в какой мере проектируемая система уступает теоретически возможной. Следует отметить также, что в некоторых случаях логика вывода, используемая в теории информации, подсказывает путь, на котором может быть найдено конструктивное решение для данной реальной системы.

Первоначально теория информации возникла из инженерных задач радиосвязи и телеграфии. Датой ее рождения считают 1948 год, год появления двух основополагающих статей американского инженера и математика Клода Шеннона «Математическая теория связи» и «Связь при наличии шума» (см.: Шеннон К., Сборник работ по теории информации и кибернетике. — М.: ИЛ, 1963). Начиная с этого времени, теория информации бурно развивалась, главным образом благодаря работам математиков и математически образованных инженеров. Нельзя не отметить огромный вклад, который внесли в теорию информации Дж. Вольфовиц, Р. Галлагер, Р. Л. Добрушин, А. Н. Колмогоров, М. С. Пинскер, В. И. Сифоров, А. Файнстейн, Р. Фано, А. А. Харкевич, А. Я. Хинчин и многие другие. В результате развития теории информации основная часть теоретических работ стала носить математически сложный характер, и образовался определенный разрыв между инженерами-практиками и адресованной в первую очередь им прикладной математической теорией. К сожалению, до настоящего времени этот разрыв не имеет тенденции уменьшаться; желание его преодолеть было одним из основных стимулов при написании этой книги, которая по замыслу авторов должна помочь студенту технического вуза познакомиться с теорией информации или ее отдельными разделами.

Сегодня можно указать две основные книги, которые могут служить учебниками по теории информации. Это книга Р. Фано «Передача информации. Статистическая теория связи» (Мир, 1965) и книга Р. Галлагера «Теория информации и надежная связь» (Сов. радио, 1974). Обе эти книги написаны известными американскими учеными, внесшими существенный вклад в теорию информации. Однако обе они в значительной степени носят монографический характер и предназначены достаточно подготовленным читателям. Следует также отметить одну из первых книг на русском языке — книгу Ф. П. Тарасенко «Введение в курс теории информации» (изд. Томского ун-та, 1963) и книгу Р. Л. Стратоновича «Теория информации» (Сов. радио, 1975), посвященную нетрадиционному изложению шенноновской теории информации с позиций статистической термодинамики.

Настоящая книга состоит из следующих пяти основных частей: кодирование дискретных источников, кодирование в дискретных каналах, кодирование в непрерывных каналах, кодирование непрерывных источников и кодирование в системах с многими пользователями.

В первой главе рассматривается задача точного или сколь угодно точного кодирования дискретных источников. В ней дается ответ на вопрос, каково наименьшее количество двоичных символов на сообщение, по которому можно точно или с какой угодно малой вероятностью ошибки восстановить последовательность сообщений на выходе дискретного источника.

Во второй главе задачи кодирования не рассматриваются. Она посвящена исследованию свойств количества информации для различных вероятностных объектов. Кроме того, в этой главе приводятся математические сведения, необходимые для чтения этой и последующих глав книги.

В третьей главе рассматривается задача кодирования в дискретных каналах связи и дается ответ на вопрос, каково наибольшее количество информационных двоичных символов, которое может быть передано по каналу связи в единицу времени при условии, что вероятность ошибки при определении переданных сообщений может быть сделана сколь угодно малой величиной. Кроме того, в этой главе строятся верхняя и нижняя границы вероятности ошибки декодирования для дискретных каналов без памяти.

Четвертая глава посвящена обобщению результатов третьей главы на случай различных непрерывных каналов.

В пятой главе рассматривается задача кодирования источников при заданном критерии качества (теория эпсилон-энтропии). В ней дается ответ на вопрос, каково наименьшее количество двоичных символов на сообщение, по которым можно восстановить с заданной ошибкой относительно выбранного критерия качества

последовательность сообщений на выходе некоторого источника.

Шестая глава посвящена задачам кодирования в системах с многими пользователями (многими источниками, каналами связи и получателями сообщений). Здесь также даются ответы на вопросы о минимальном числе двоичных символов на сообщение источника или о максимальном числе информационных двоичных символов, передаваемых по каналу в единицу времени, в ситуации, когда имеется несколько источников и они зависимы или когда имеется несколько каналов и передача по одному каналу мешает передаче по другим.

В книге имеются основные и дополнительные параграфы. Основные задуманы как пособие для читателя, который впервые знакомится с теорией информации и который не рискнул бы считать себя хорошо владеющим теорией вероятностей. Тем не менее, эта часть книги позволяет читателю проникнуть в проблематику теории информации и познакомиться с ее основными результатами, пройдя через все трудности доказательств. Следующие параграфы являются основными: вся гл. I, §§ 2.1—2.3, вся гл. III (кроме пп. 3.5.2, 3.12.4, 3.12.5 и § 3.13), §§ 4.1, 4.2, 5.1—5.5. Эти разделы могут составить материал для односеместрового курса лекций по теории информации. Курсы лекций примерно с таким содержанием читаются в течение ряда лет в Ленинградском институте авиационного приборостроения. При рассмотрении основных разделов сделана попытка упростить изложение за счет сужения круга рассматриваемых вопросов, использования в связи с этим более простой математической техники, более подробного обсуждения постановок задач и примеров. Кроме того, здесь даются все необходимые математические сведения, что делает эту часть книги в определенной мере самостоятельной.

Дополнительных параграфов несколько (в книге они помечены звездочкой над номером параграфа). Все, что не вошло в перечисленные выше основные параграфы (в пределах первых пяти глав), можно рассматривать как дополнительный материал, предназначенный для углубленного изучения «традиционной» теории информации. Хотя многие математические сведения здесь также приводятся, предполагается, что читатель знаком с элементами функционального анализа, с элементами теории случайных процессов и с элементами нелинейного программирования. Кроме того, для чтения этих разделов требуется несколько бо́льшая математическая тренированность.

Особое место в книге занимает шестая глава, посвященная задачам кодирования в системах с многими пользователями. В ней представлены результаты, полученные в теории информации в течение последнего десятилетия и отражающие развитие современных систем обработки и передачи информации. Содержание

шестой главы адресовано в первую очередь читателям, хорошо знакомым с традиционными вопросами теории информации, например, по книге Р. Галлагера, или хорошо овладевшим содержанием первых пяти глав настоящей книги. Для успешного чтения этой главы от читателя требуется наличие определенной теоретико-информационной эрудиции.

Каждая глава снабжена рядом задач, упражнений и дополнений. Среди задач и упражнений имеются очень простые, предназначенные только для проверки того, что читатель правильно понял формулировки определений и теорем. Имеются задачи, которые требуют овладения техникой доказательств. В ряде случаев приводятся дополнительные сведения, затрагивающие как методы, так и интересные и важные результаты теории информации, которые по тем или другим соображениям не включены в основной текст, но которые могут быть полезны при более глубоком изучении теории или при использовании теории на практике.

В заключение отметим, что хотя в книге содержатся основные математические сведения, которые используются при изложении рассматриваемых теоретико-информационных задач, их явно недостаточно для глубокого понимания математических основ теории информации. Читателю, желающему более детально познакомиться с этими основами, мы рекомендуем обратиться к следующим книгам. С теорией вероятностей лучше знакомиться по книгам Б. В. Гнеденко «Курс теории вероятностей» (Наука, 1965) и В. Феллера «Введение в теорию вероятностей и ее приложения», т. I (Мир, 1967). С элементами теории случайных процессов можно познакомиться по книге В. Б. Давенпорта и В. Л. Рута «Введение в теорию случайных сигналов и шумов» (ИЛ, 1960). Основы матричной алгебры и теория операторов в конечномерных пространствах лучше всего изложены в книгах Ф. Р. Гантмахера «Теория матриц» (Наука, 1967) или Р. Беллмана «Введение в теорию матриц» (Наука, 1969). По книге Б. З. Вулиха «Введение в функциональный анализ» (Наука, 1967) можно познакомиться с элементами функционального анализа.

Авторы выражают огромную признательность всем, кто знакомился с многочисленными вариантами рукописи этой книги и делился своими замечаниями. Особенно большое влияние на работу авторов оказали замечания и советы Ю. М. Штарькова и рецензентов — Р. Л. Добрушина и Э. М. Габидулина.

Глава 1

КОДИРОВАНИЕ ДИСКРЕТНЫХ ИСТОЧНИКОВ

В этой главе будут даны основные определения: дискретного вероятностного ансамбля, дискретного источника, дискретной случайной величины на ансамбле и кода для дискретного источника. Дискретные источники представляют собой наиболее простой объект теории информации. Начинать именно с этого типа источников удобно не только потому, что здесь требуется наименьшее количество определений и вспомогательных результатов, но и потому, что на этом простом объекте можно показать методологию теории информации и продемонстрировать ее основные технические приемы.

Задача, которая рассматривается в этой главе, весьма часто встречается на практике и иногда называется задачей сжатия данных. Предположим, что некоторый источник порождает последовательность дискретных сообщений и требуется представить эту последовательность с помощью некоторых символов, скажем, с помощью нулей и единиц. Не вызывает сомнения то, что это можно сделать для любой последовательности сообщений. Вопрос может заключаться в том, как это сделать наиболее экономным образом, т. е. как затратить на это наименьшее количество двоичных символов. Ответ на этот вопрос лежит в изучении различных статистических моделей источников и определении для этих моделей величины, называемой скоростью создания информации. Будет показано, что скорость создания информации равна энтропии источника на сообщение, величине, которая определяется с помощью вводимого в этой главе понятия количества информации в сообщении.

§ 1.1. Дискретные ансамбли и источники

Основным объектом изучения в этой главе будут дискретные источники сообщений. Здесь мы дадим определения, необходимые для описания математических моделей источников.

Описание источников удобно начать с определения дискретных вероятностных ансамблей. Пусть $X = \{x_1, \ldots, x_M\}$ — множество, состоящее из M элементов. Прописные латинские буквы X, Y и т. д. будут обозначать сами множества, а соответствующие строчные буквы x, y и т. д. будут обозначать элементы множеств.

Иногда элементы множеств мы будем снабжать подстрочными индексами, как это сделано выше. Такой индекс представляет собой номер элемента в множестве. Хотя для большинства случаев природа элементов несущественна, мы будем называть элементы множеств X сообщениями, подчеркивая тем самым область приложения теории.

Говорят, что на конечном множестве X задано распределение вероятностей $p(x)$, если каждому элементу $x_i \in X$ сопоставлено число $p(x_i)$, причем

$$p(x_i) \geqslant 0, \quad i = 1, 2, \ldots, M, \tag{1.1.1}$$

$$\sum_{i=1}^{M} p(x_i) = 1.$$

Пусть A есть подмножество множества X, $A \subseteq X$. Число *)

$$\Pr(A) \triangleq \sum_{x_i \in A} p(x_i)$$

представляет собой вероятность того, что при случайном выборе сообщения из множества X в соответствии с распределением $p(x)$, будет выбрано сообщение, принадлежащее множеству A. Число $\Pr(A)$ называют также вероятностью множества A.

Пример 1.1.1. Пусть X — множество сообщений о результатах бросания правильной игральной кости. Тогда $X = \{x_1, \ldots, x_6\}$, $p(x_i) = 1/6$, $i = 1, \ldots, 6$, причем x_i есть сообщение о том, что выпало i очков. Если $A = \{x_2, x_4, x_6\}$, то $\Pr(A) = 3 \cdot 1/6 = 1/2$ есть вероятность того, что при бросании кости выпало четное число очков.

Сообщения $x_i \in X$ иногда называют элементарными событиями. Как показывает предыдущая формула, могут одновременно рассматриваться как элементарные события, так и более сложные события, являющиеся объединением некоторого числа элементарных. Мы используем различные обозначения для вероятностей таких событий: $p(x_i)$ — для элементарных событий и $\Pr(A)$ — для множества A, образованного элементарными событиями $x_i \in A$. Это различие не принципиально, но делает некоторые формулы наглядными.

Определение 1.1.1. Конечное множество X вместе с заданным на нем распределением вероятностей $p(x)$ называется *дискретным вероятностным ансамблем* или коротко — *дискретным ансамблем* (сообщений) и обозначается символом $\{X, p(x)\}$. В тех случаях, когда из контекста видно, о каком распределении вероят-

*) Здесь и ниже знак $\triangleq$ используется для обозначения того, что правая и левая части равны по определению. Иногда для упрощения обозначений мы будем писать $\Pr(A) = \sum_{A} p(x_i)$.

ностей идет речь или когда точное описание распределения несущественно, мы будем обозначать ансамбль через 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 *произведением множеств* X и Y и обозначается через XY. Согласно этому определению XY и YX суть различные множества. Если множества X и Y совпадают, $X = Y$, то произведение XY обозначается как X^2. Аналогичным образом определяются произведения более чем двух множеств. Произведение $X_1X_2 \ldots X_n$ представляет собой множество всех последовательностей $(x^{(1)}, x^{(2)}, \ldots, x^{(n)})$ длины n таких, что первый элемент $x^{(1)}$ принадлежит множеству X_1, второй $x^{(2)}$ — множеству X_2 и т. д., n-й элемент принадлежит множеству X_n *). Если все множества совпадают между собой и с множеством X, то такое произведение обозначается как X^n. Таким образом, X^n — это множество всех последовательностей длины n, образованных из элементов множества X.

Пусть XY есть произведение двух конечных множеств X и Y и на множестве XY задано совместное распределение вероятностей $p(x, y)$, которое каждой паре (x_i, y_j), $x_i \in X$, $y_j \in Y$, сопоставляет вероятность $p(x_i, y_j)$. Очевидно, что соотношения

$$p_1(x_i) \triangleq \sum_{y_j \in Y} p(x_i, y_j), \quad i = 1, 2, \ldots, M, \tag{1.1.2}$$

$$p_2(y_j) \triangleq \sum_{x_i \in X} p(x_i, y_j), \quad j = 1, 2, \ldots, N, \tag{1.1.3}$$

задают распределения вероятностей $p_1(x)$ и $p_2(y)$ на множествах X и Y соответственно. Таким образом, при задании ансамбля $\{XY, p(x, y)\}$ фактически задаются еще два ансамбля $\{X, p_1(x)\}$ и $\{Y, p_2(y)\}$. Иногда, имея в виду ансамбль $\{XY, p(x, y)\}$, мы будем говорить, что *совместно заданы* два ансамбля X и Y. Это будет означать, что распределения вероятностей на множествах X и Y определяются по формулам (1.1.2) и (1.1.3), исходя из распределения вероятностей $p(x, y)$ на множестве XY.

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

$$p(x_i, y_j) = p_1(x_i)\, p_2(y_j) \text{ для всех } x_i \in X,\ y_j \in Y, \tag{1.1.4}$$

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

---

*) Здесь и в аналогичных обозначениях дальше надстрочный индекс обозначает номер элемента в последовательности, другими словами, $x^{(i)}$ — элемент, расположенный на i-м месте последовательности

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

$$p(y_j \mid x_i) \triangleq \frac{p(x_i, y_j)}{p_1(x_i)} \tag{1.1.5}$$

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*). Легко увидеть, что множество условных вероятностей относительно фиксированного сообщения x_i, которое получается, когда индекс j пробегает все возможные значения, удовлетворяет определению распределения вероятностей (1.1.1). Такое распределение называется *условным распределением на множестве Y относительно фиксированного сообщения x_i*; заметим, что (1.1.3) определяет так называемое *безусловное распределение на Y*. Понятно, что аналогичные распределения могут быть определены также на множестве X. Таким образом, задание ансамбля $\{XY, p(x, y)\}$ определяет также условные ансамбли $\{X, p(x \mid y)\}$, $p_2(y) \neq 0$, и $\{Y, p(y \mid x)\}$, $p_1(x) \neq 0$.

Опишем теперь общее семейство условных ансамблей, которые образуются при совместном задании двух ансамблей. Пусть A — произвольное подмножество элементов из X такое, что $\mathbf{Pr}_1(A) \triangleq \triangleq \sum_{x \in A} p_1(x) \neq 0$, где распределение $p_1(x)$ определено выше. Число

$$p(y_j \mid A) \triangleq \frac{1}{\mathbf{Pr}_1(A)} \sum_{x_i \in A} p(x_i, y_j) \tag{1.1.6}$$

называется *условной вероятностью сообщения y_j при условии, что сообщение x_i принадлежит множеству A (или условной вероятностью относительно множества A)*. Легко видеть, что множество условных вероятностей относительно фиксированного множества A, которое получается, когда индекс j пробегает все возможные значения, снова удовлетворяет определению распределения вероятностей. Такое распределение называется *условным распределением на множестве Y относительно фиксированного множества A*. Если выбрать $A = X$, то $p(y_j \mid A) = p_2(y_j)$. Таким образом, условное распределение относительно множества X — это просто безусловное распределение на Y. Если A состоит из одного элемента, скажем x_i, то $p(y_j \mid A)$ равно вероятности, определенной в (1.1.5).

Аналогичные условные распределения могут быть определены также для множества X. Тем самым определены условные ансамбли $\{X, p(x \mid B)\}$, $B \subseteq Y$, $\mathbf{Pr}_2(B) \neq 0$ и $\{Y, p(y \mid A)\}$, $A \subseteq \subseteq X$, $\mathbf{Pr}_1(A) \neq 0$.

Рассмотрим произведение n множеств $X_1 \ldots X_n$ и распределение вероятностей $p(x^{(1)}, \ldots, x^{(n)})$, $x^{(i)} \in X_i$, $i = 1, \ldots, n$, задан-

ное на этом произведении. Другими словами, рассмотрим вероятностный ансамбль $\{X_1 \dots X_n,\ p(x^{(1)}, \dots, x^{(n)})\}$. Пусть

$$\begin{aligned} p_1(x^{(1)}) &\triangleq \sum_{X_2} \sum_{X_3} \dots \sum_{X_n} p(x^{(1)}, \dots, x^{(n)}), \\ p_2(x^{(2)}) &\triangleq \sum_{X_1} \sum_{X_3} \dots \sum_{X_n} p(x^{(1)}, \dots, x^{(n)}), \\ &\dots\dots\dots\dots\dots\dots \\ p_n(x^{(n)}) &\triangleq \sum_{X_1} \sum_{X_2} \dots \sum_{X_{n-1}} p(x^{(1)}, \dots, x^{(n)}). \end{aligned} \tag{1.1.7}$$

Соотношения (1.1.7) задают безусловные распределения вероятностей на множествах X_1, X_2, ..., X_n соответственно. Если для любых $x^{(1)} \in X_1$, ..., $x^{(n)} \in X_n$ имеет место равенство

$$p(x^{(1)}, \dots, x^{(n)}) = p_1(x^{(1)}) \dots p_n(x^{(n)}), \tag{1.1.8}$$

то ансамбли X_1, ..., X_n называют *статистически независимыми*.

При совместном задании n ансамблей X_1, ..., X_n оказываются совместно заданными всевозможные совокупности по $m < n$ ансамблей. Так, ансамбль $\{X_{i_1} \dots X_{i_m},\ p(x^{(i_1)}, \dots, x^{(i_m)})\}$ определяется с помощью следующего соотношения, которое дает безусловное распределение вероятностей на произведении множеств $X_{i_1} \dots X_{i_m}$:

$$p(x^{(i_1)}, \dots, x^{(i_m)}) \triangleq \sum_{X_{j_1}} \dots \sum_{X_{j_{n-m}}} p(x^{(1)}, \dots, x^{(n)}), \tag{1.1.9}$$

где суммирование производится по всем множествам X_{j_1}, ..., $X_{j_{n-m}}$ таким, что произведение соответствующим образом упорядоченных множеств X_{i_1}, ..., X_{i_m}, X_{j_1}, ..., $X_{j_{n-m}}$ есть множество $X_1 \dots X_n$ ($\{i_1, \dots, i_m\} \cap \{j_1, \dots, j_{n-m}\} = \phi$, $\{i_1, \dots, i_m\} \cup \{j_1, \dots, j_{n-m}\} = \{1, 2, \dots, n\}$). Другими словами, множество X_j участвует в суммировании в (1.1.9), если j не содержится среди чисел i_1, ..., i_m.

Вводя по аналогии с (1.1.5) и (1.1.6) условные распределения на множестве X_{i_1}, ..., X_{i_m}, можно получить различные условные ансамбли.

Пример 1.1.2. Пусть $\{X,\ p_1(x)\}$ — ансамбль сообщений из примера 1.1.1, а $\{Y,\ p_2(y)\}$ — ансамбль сообщений о результате бросания монеты, $Y = \{y_1, y_2\}$ и $p_2(y_1) = p_2(y_2) = 1/2$. Элементы множества XY (произведения множеств X и Y) представляют собой пары сообщений (x_i, y_j), причем первый элемент пары есть сообщение x_i о результате бросания кости, а второй элемент y_j есть сообщение о результате бросания монеты. В случае независимого бросания кости и монеты все пары имеют одинаковые вероятности: $p(x_i, y_j) = 1/12$ для всех i, j.

Можно представить себе более сложный эксперимент с костью и монетой. Предположим, что вначале бросается кость. Если выпало четное число очков, бросается неправильная монета, у которой вероятность выпадения стороны,

соответствующей y_1, равна $^1/_4$ (а стороны, соответствующей y_2, — $^3/_4$). Если же выпало нечетное число очков, то бросается другая неправильная монета, у которой вероятность выпадения стороны, соответствующей y_1, равна $^3/_4$ (а стороны, соответствующей y_2, — $^1/_4$). При таком эксперименте на множестве Y имеются два условных распределения: $p(y_1|x) = {}^1/_4$, $p(y_2|x) = {}^3/_4$ при четных $x \in X$ и $p(y_1|x) = {}^3/_4$, $p(y_2|x) = {}^1/_4$ при нечетных $x \in X$. Распределение вероятностей на множестве XY указано в следующей таблице:

$p(x_i, y_j)$	x_1	x_2	x_3	x_4	x_5	x_6
y_1	1/8	1/24	1/8	1/24	1/8	1/24
y_2	1/24	1/8	1/24	1/8	1/24	1/8

Нетрудно видеть, что безусловные распределения в этом случае такие же, как и в случае одной симметричной (правильной) монеты, но ансамбли X и Y статистически независимыми не являются.

Пример 1.1.3. Рассмотрим n последовательных бросаний некоторой монеты. Обозначим через Y_i множество сообщений о результате i-го бросания, $i = 1, 2, \ldots, n$, $Y_i = \{y_1, y_2\}$. Множество $Y^n \triangleq \prod_{i=1}^{n} Y_i$ содержит 2^n последовательностей длины n, являющихся сообщениями о результатах n-кратного бросания монеты. Если положить $y_1 = 0$ и $y_2 = 1$, то множество Y^n будет состоять из всех двоичных последовательностей длины n. Рассмотрим теперь два случая, соответствующие независимым и зависимым бросаниям.

а) Предположим, что бросается правильная монета. Тогда все возможные исходы имеют одинаковые вероятности, равные 2^{-n}. При каждом бросании сообщения y_1 и y_2 могут появиться с одинаковыми вероятностями. Поэтому для любой последовательности $(y^{(1)}, \ldots, y^{(n)})$

$$2^{-n} = p(y^{(1)}, \ldots, y^{(n)}) = \prod_{i=1}^{n} p(y^{(i)}).$$

Следовательно, подбрасывание правильной монеты соответствует последовательности независимых испытаний.

б) Теперь предположим, что имеются две неправильные монеты, описанные в примере 1.1.2. Сначала наудачу выбирается одна из двух монет. Если результат бросания есть y_1, то для следующего бросания выбирается монета с $p(y_1) = {}^3/_4$ и другая монета в противном случае. Затем в зависимости от результата второго подбрасывания аналогичным образом выбирается очередная монета и этот процесс повторяется. Соответствующее такому эксперименту распределение вероятностей задается следующим образом:

$$p(y^{(1)}, \ldots, y^{(n)}) \triangleq p(y^{(1)})\, p(y^{(2)} | y^{(1)}) \ldots p(y^{(n)} | y^{(n-1)}),$$

где

$$p(y^{(i)} = y_1 | y^{(i-1)} = y_1) = 3/4, \quad p(y^{(i)} = y_2 | y^{(i-1)} = y_1) = 1/4,$$
$$p(y^{(i)} = y_1 | y^{(i-1)} = y_2) = 1/4, \quad p(y^{(i)} = y_2 | y^{(i-1)} = y_2) = 3/4.$$